ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ
ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΤΜΗΜΑ ΕΠΙΣΤΗΜΗΣ ΥΠΟΛΟΓΙΣΤΩΝ

Υλοποίηση Συστήματος Σταθεροποίησης της Οπτικής Προσήλωσης Στερεοσκοπικής Κεφαλής Ενεργού Όρασης

Βαρβάρα Ε.Μ. Μαστή

Μεταπτυχιακή Εργασία

Ηράκλειο, 1996

ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ
ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΤΜΗΜΑ ΕΠΙΣΤΗΜΗΣ ΥΠΟΛΟΓΙΣΤΩΝ

Υλοποίηση Συστήματος Σταθεροποίησης της Οπτικής Προσήλωσης Στερεοσκοπικής Κεφαλής Ενεργού Όρασης

Εργασία που υποβλήθηκε από την
Βαρβάρα Ε.Μ. Μαστή
ως μερική εκπλήρωση των απαιτήσεων
για την απόκτηση
ΜΕΤΑΠΤΥΧΙΑΚΟΥ ΔΙΠΛΩΜΑΤΟΣ ΕΙΔΙΚΕΥΣΗΣ

Συγγραφέας:

Βαρβάρα Ε.Μ. Μαστή
Τμήμα Επιστήμης Υπολογιστών

Εισηγητική Επιτροπή:

Στέλιος Ορφανουδάκης, Επόπτης

Πάνος Τραχανιάς, Μέλος

Απόστολος Τραγανίτης, Μέλος

Δεκτή:

Πάνος Κωνσταντόπουλος
Πρόεδρος Επιτροπής Μεταπτυχιακών Σπουδών

Ηράκλειο, 1996

Ευχαριστίες

Θα ήθελα να ευχαριστήσω όλους όσους με βοήθησαν για να ολοκληρώσω αυτή την εργασία.

Ευχαριστώ θερμά τον επόπτη μου, κ. Στέλιο Ορφανουδάκη, για τη συνεργασία, την υποστήριξη και τις συμβουλές του κατά τη διάρκεια των μεταπτυχιακών μου σπουδών.

Ευχαριστώ, επίσης, τους καθηγητές Πάνο Τραχανιά, Νίκο Αλβέρτο και Γιώργο Τζίριτα για την προθυμία τους να με βοηθήσουν όσες φορές απευθύνθηκα σ' αυτούς.

Επίσης, οι συνεργάτες μου στο εργαστήριο Μηχανικής Όρασης και Ρομποτικής, Μανόλης Λουράκης και Αντώνης Αργυρός με βοήθησαν σε πολλά σημεία της εργασίας μου.

Για την ολοκλήρωση του κειμένου, ήταν πολύτιμα τόσο τα σχόλια του επόπτη μου κ. Στέλιου Ορφανουδάκη και του καθηγητή Πάνου Τραχανιά ως μέλους της εξεταστικής επιτροπής, όσο και των φίλων μου Γιάννη Σουρλαντζή και Δέσποινας Βαμβακά.

Όσον αφορά την ολοκλήρωση της προσομοίωσης, ευχαριστώ τον Γιάννη Σουρλαντζή που με ιδιαίτερη ευκολία έβλεπε ότι κάτι "δεν δουλεύει". Εξάλλου, χωρίς την βοήθεια του, πιθανόν να έγραφα ακόμα τις διαφάνειες για την παρουσίαση και όχι τις ευχαριστίες.

Σημαντική ήταν η βοήθεια και η ψυχολογική υποστήριξη που είχα από πολλούς φίλους και συναδέλφους μου μεταπτυχιακούς φοιτητές. Ιδιαίτερα θα ήθελα να ευχαριστήσω τη Μιράντα Κυπριώτη, τη Δέσποινα Βαμβακά, το Θοδωρή Χατσιούλη, το Γιάννη Σουρλαντζή τη Χριστίνα Γκριτζάπη, τη Νατάσα Βασιλά και τη Φωτεινή Σετάκη.

Θέλω ακόμη να πω ένα μεγάλο ευχαριστώ στους γονείς μου, Μιράντα και Μανώλη, και την αδελφή μου, Χριστίνα, για το κουράγιο και την ηθική στήριξη που μου πρόσφεραν όλα τα χρόνια των σπουδών μου.

Θα ήθελα επίσης να ευχαριστήσω τον Λουκά Μιχελή για την πίστη του σε μένα και για την ενθάρρυνση του στο να ξεκινήσω και να ολοκληρώσω τις μεταπτυχιακές μου σπουδές.

Τέλος, πρέπει να ευχαριστήσω το Τμήμα Επιστήμης Υπολογιστών του Πανεπιστημίου Κρήτης και το Ινστιτούτο Πληροφορικής του Ιδρύματος Τεχνολογίας και Έρευνας για την υλικοτεχνική υποστήριξη και την οικονομική ενίσχυση που μου παρείχαν κατά τη διάρκεια των μεταπτυχιακών μου σπουδών.

Υλοποίηση Συστήματος Σταθεροποίησης της Οπτικής Προσήλωσης Στερεοσκοπικής Κεφαλής Ενεργούς Όρασης.

Βαρβάρα Ε.Μ. Μαστή
Μεταπτυχιακή Εργασία

Τμήμα Επιστήμης Υπολογιστών
Πανεπιστήμιο Κρήτης

Περίληψη

Ο έλεγχος των παραμέτρων ενός οπτικού συστήματος είναι ένα σημαντικό θέμα στην μηχανική όραση. Ειδικότερα, σύμφωνα με την θεωρία της ενεργού όρασης, ένα σύστημα όρασης πρέπει να ελέγχει συνεχώς τις παραμέτρους της διαδικασίας λήψης του οπτικού σήματος, έτσι ώστε οι απαιτούμενοι υπολογισμοί για την εκδήλωση μιας συμπεριφοράς να είναι ευκολότεροι. Ένα σημαντικό ερευνητικό θέμα στα πλαίσια της θεωρίας αυτής είναι ο έλεγχος προσήλωσης (gaze control), δηλαδή ο έλεγχος των μηχανικών και οπτικών παραμέτρων του συστήματος. Ο έλεγχος προσήλωσης έχει δύο μορφές : την σταθεροποίηση (stabilization) και την μετατόπιση (change). Η σταθεροποίηση, η οποία μας απασχολεί στην παρούσα εργασία, συνίσταται στον έλεγχο των διαθέσιμων βαθμών ελευθερίας για τη λήψη καθαρών εικόνων κάποιου σημείου του κόσμου το οποίο μπορεί είτε να είναι σταθερό, είτε να κινείται σε σχέση με το οπτικό σύστημα. Ειδικότερα, το πρόβλημα της σταθεροποίησης, είναι ένα πρόβλημα ελέγχου αρνητικής ανάδρασης λάθους και επομένως οποιαδήποτε λύση του απαιτεί την σχεδίαση ενός εκτιμητή λάθους, ενός ελεγχόμενου μηχανικού συστήματος και ενός νόμου ελέγχου που να οδηγεί το λάθος στο μηδέν.

Στην παρούσα εργασία υλοποίηθηκε ένα σύστημα σταθεροποίησης της οπτικής προσήλωσης (ΣΣΠ) για μια στερεοσκοπική κεφαλή ενεργού όρασης που διαθέτει τέσσερις βαθμούς ελευθερίας. Το σύστημα στηρίχθηκε στην αρχιτεκτονική που έχει προταθεί από τον D.J. Coombs [22]. Αποτελείται από δυο υποσυστήματα τα οποία συνεργάζονται: το υποσύστημα σύγκλισης και το υποσύστημα καταδίωξης. Το υποσύστημα σύγκλισης υπολογίζει το λάθος στην κατεύθυνση του βάθους και το υποσύστημα καταδίωξης

το λάθος θέσης πάνω στην οροπτερική επιφάνεια (γεωμετρικός τόπος των σημείων που προβάλλονται σε σύστοιχα σημεία στα επίπεδα των καμερών). Και τα δύο υποσυστήματα με αυτές τις εκτιμήσεις προσπαθούν να κατευθύνουν τις κάμερες έτσι ώστε οι οπτικοί άξονες τους να τέμνονται πάνω στην επιφάνεια του στόχου.

Για καθένα από τα υποσυστήματα υλοποιήσαμε διάφορες ήδη προτεινόμενες στην βιβλιογραφία τεχνικές με σκοπό την αξιολόγηση τους ως προς το πως συνεργάζονται για την επίτευξη της σταθεροποίησης της οπτικής προσήλωσης. Για τον έλεγχο του συστήματος εξετάσαμε διάφορους τρόπους κατανομής των λαθών στους βαθμούς ελευθερίας του κεφαλιού με τους οποίους μπορεί να οδηγηθεί το λάθος παρακολούθησης στο μηδέν.

Για την αξιολόγηση της συμπεριφοράς του συστήματος αναπτύξαμε μια πλατφόρμα για την προσομοίωση της κινηματικής ρομποτικών συστημάτων. Χρησιμοποιώντας την πλατφόρμα αυτή έγινε προσομοίωση της στερεοκοπικής κεφαλής όπου το ρόλο της κάμερας παίζει ένας raytracer (Rayshade). Η πλατφόρμα αυτή μπορεί να υποστηρίξει συνθετότερα συστήματα ρομποτικής και να επεκταθεί εύκολα έτσι ώστε να λάβει υπ' όψην τη δυναμική τους.

Επόπτης: Στέλιος Ορφανουδάκης
Καθηγητής,
Πανεπιστήμιο Κρήτης.

Implemantation of a Gaze Stabilization System for a Stereo Active Vision Head.

Barbara E.M. Masti
Master's Thesis

Computer Science Department
University of Crete, Greece

Abstract

The control of imaging parameters of a visual system is an important research area of computer vision. According to the theory of active vision, a visual system must constantly control its imaging parameters to aid in the performance of visual tasks. An important research topic in this context is gaze control, the control of mechanical parts and imaging parameters of the system. Gaze control can be subdivided into two mechanisms: gaze stabilization and gaze change. Gaze stabilization, which is the topic of this thesis, relates to the control of the degrees of freedom so that a world point can be stabilized on the image plane (which is either stable or moving with respect to the visual system). In particular, the gaze stabilization problem, is a problem of negative error feedback control. Thus, it requires the design of an error estimator, a controlled mechanical system and a control law which minimizes the error.

In this thesis, a gaze stabilization system has been implemented for a stereo active vision head with four degrees of freedom. This system is based on the architecture proposed by D.J. Coombs [22]. It consists of two cooperating subsystems, the vergence subsystem and the pursuit subsystem. The vergence subsystem estimates the error in the depth direction and, similarly, the pursuit subsystem estimates the positional error on the horopter (locus of points that stimulate exactly corresponding points). Both subsystems use these errors to direct the cameras, so that their optic axes instersect on the target's surface.

For each subsystem we have implemented several techniques, already prosposed in the literature, in order to evaluate them in the context of gaze stabilization. For system control, we tested different ways by which errors can be distributed among the head's degrees of freedom in order to minimize the error of tracking.

In order to evaluate the system, we developed a platform for the simulation of robot system kinematics. Using this platform, we simulated a stereoscopic head where a raytracer (Rayshade) plays the role of a camera. This platform can be used to represent more complex robot systems and can be easily extended to take into account their dynamics.

Advisor: Stelios Orphanoudakis
Professor,
University of Crete.

Περιεχόμενα

1 Εισαγωγή **1**

1.1 Σταθεροποίηση προσήλωσης 1

1.2 Οριοθέτηση της εργασίας 2

1.3 Οργάνωση του κειμένου 3

2 Έλεγχος προσήλωσης **5**

2.1 Θεωρίες όρασης και έλεγχος προσήλωσης 5

2.2 Μορφές και πλεονεκτήματα ελέγχου προσήλωσης 6

2.2.1 Πλεονεκτήματα 6

2.2.2 Σταθεροποίηση προσήλωσης 7

2.2.3 Μετατόπιση προσήλωσης 8

2.3 Βιολογικά συστήματα ελέγχου προσήλωσης 9

2.3.1 Σχέση νευροφυσιολογίας και μηχανικής όρασης 9

2.3.2 Το ανθρώπινο οφθαλμοκινητικό σύστημα 9

2.3.2.1 Αντανακλαστικό ομαλής καταδίωξης και οπτοκινητικό αντανακλαστικό 10

2.3.2.2 Σακκαδικές κινήσεις 11

2.3.2.3 Αντανακλαστικό σύγκλισης, αντανακλαστικό προσαρμογής και αντανακλαστικό κοντινής τριάδας . . 11

2.3.2.4 Θαλαμιαίο οφθαλμικό αντανακλαστικό 11

2.4 Ενεργά συστήματα μηχανικής όρασης και έλεγχος προσήλωσης 12

2.4.1 Εξέλιξη αυτόνομων συστημάτων 12

2.4.2 Συστήματα ελέγχου προσήλωσης 13

2.4.2.1 Συστήματα σταθεροποίησης της προσήλωσης 13

2.4.2.2 Συστήματα μετατόπισης της προσήλωσης και μοντέλα εστίασης της οπτικής προσοχής 16
2.4.2.3 Συστήματα ελέγχου προσήλωσης με εφαρμογές στην όραση 17

3 Σύστημα σταθεροποίησης της προσήλωσης 19

3.1 Γεωμετρία διοφθαλμικής προσήλωσης 19
3.2 Το πρόβλημα 21
3.3 Αρχιτεκτονική του συστήματος 24
3.4 Υπολογισμός ανομοιότητας 27
3.5 Υπολογισμός θέσης 31
3.6 Έλεγχος του συστήματος 32
3.6.1 Μοντέλο καμέρας 32
3.6.2 Σφάλματα στις διορθώσεις 32
3.6.2.1 Κίνηση pan της κάμερας 33
3.6.2.2 Κίνηση tilt του κεφαλιού 33
3.6.2.3 Κίνηση pan του κεφαλιού 35
3.6.3 Σενάρια κατανομής των σφαλμάτων στους κινητήρες 38
3.7 Συνολική θεώρηση του συστήματος 43

4 Cepstral φίλτρο 45

4.1 Βιολογικό υπόβαθρο 45
4.2 Μαθηματικό υπόβαθρο 46
4.3 Υπολογισμός ανομοιότητας με cepstrum 47
4.3.1 Το cepstrum σήματος που περιέχει ηχώ 47
4.3.2 Ο αλγόριθμος υπολογισμού αραιού πεδίου ανομοιότητας των Yeshurun και Schwartz 49
4.3.3 Ο αλγόριθμος υπολογισμού ανομοιότητας του στόχου σύγκλισης των Olson και Coombs 51
4.3.4 Σύγκριση του cepstral φίλτρου με τη συσχέτιση φάσης 52
4.4 Προβλήματα και περιορισμοί του cepstral φίλτρου 53
4.4.1 Το πεδίο ανομοιότητας 53
4.4.2 Θόρυβος, παραμόρφωση του σήματος, ευθύγραμμα τμήματα ακμών και μεταβλητός φωτισμός 54
4.5 Αναζήτηση κορυφής με τεχνική πολλαπλών επιπέδων 55

5 Φίλτρα μηδενικής ανομοιότητας (ΦΜΑ) 61

5.1 ΦΜΑ σε ένα επίπεδο ανάλυσης 62
5.1.1 Απλό ΦΜΑ συσχέτισης κάθετων ακμών 63
5.1.2 ΦΜΑ συσχέτισης γειτονιών 64
5.1.2.1 Εικόνα διαφορών 64
5.1.2.2 Εικόνα ποιότητας των χαρακτηριστικών 65
5.1.3 ΦΜΑ συσχέτισης μέτρου-κατεύθυνσης ακμών 66

5.2 ΦΜΑ ζωνοδιαβατής πυραμίδας 67
6 Προσομοίωση και αποτελέσματα **71**
6.1 Προσομοίωση ενεργού στερεοσκοπικής κεφαλής 71
6.2 Αποτελέσματα σύγκλισης 71
6.3 Αποτελέσματα παρακολούθησης 72
7 Επίλογος **85**
7.1 Συμπεράσματα 85
7.2 Επεκτάσεις 87
A Σφάλμα διόρθωσης κατά την pan κίνηση του κεφαλιού **89**
A Ελληνοαγγλικό Ευρετήριο Όρων **93**
B Αγγλοελληνικό Ευρετήριο Όρων **97**
Βιβλιογραφία **100**

Κατάλογος Σχημάτων

3.1 Γεωμετρία διοφθαλμικής προσήλωσης 20
3.2 Παρακολούθηση σε μια διάσταση 22
3.3 Η παρακολούθηση ως σύστημα ανάδρασης λάθους 22
3.4 Ένα απλουστευμένο γεωμετρικό μοντέλο των κινήσεων των ματιών στον άνθρωπο. 24
3.5 Η γεωμετρία μιάς τυπικής κεφαλής ενεργού όρασης 25
3.6 Αντιστοιχία συμμετρικής ανομοιότητας και γωνίας σύγκλισης 26
3.7 Αρχιτεκτονική συστήματος σταθεροποίησης της προσήλωσης 28
3.8 Pan κίνηση της κάμερας με σκοπό να προσανατολιστεί πρός ένα σημείο P. 33
3.9 Tilt κίνηση του κεφαλιού με σκοπό να προσανατολιστεί πρός ένα σημείο P. 34
3.10 Σφάλμα διόρθωσης με την tilt κίνηση του κεφαλιού, ως συνάρτηση του λόγου d/h και της διόρθωσης. Το πεδίο θέας είναι $\pi/2$. 35
3.11 Σφάλμα διόρθωσης που οφείλεται στην pan κίνηση του κεφαλιού, ως συνάρτηση της tilt γωνίας και της διόρθωσης. Το πεδίο θέας είναι $\pi/6$. . . 36
3.12 Σφάλμα διόρθωσης που οφείλεται στην pan κίνηση του κεφαλιού, ως συνάρτηση της tilt γωνίας και της διόρθωσης. Το πεδίο θέας είναι $\pi/2$. . . 37
3.13 Σφάλμα διόρθωσης που οφείλεται στην pan κίνηση του κεφαλιού, ως συνάρτηση της γωνίας tilt. Η διόρθωση ορίζεται από ένα σημείο P στην άκρη της εικόνας. 39
3.14 Η $x-$συντεταγμένη του σφάλματος διόρθωσης με την pan κίνηση του κεφαλιού, ως συνάρτηση της γωνίας tilt. 40
3.15 Η $y-$συντεταγμένη του σφάλματος διόρθωσης με την pan κίνηση του κεφαλιού, ως συνάρτηση της γωνίας tilt. 41

3.16 Η τροχιά της προβολής του σημείου P (απ' το οποίο ορίζεται η διόρθωση) κατά τη διόρθωση με την pan κίνηση του κεφαλιού και στη συνέχεια με ζεύγη διορθώσεων tilt-pan (pan για την $x-$συντεταγμένη και tilt για την $y-$συντεταγμένη του σφάλματος). Η αρχική γωνία tilt είναι $\pi/2$. 42
3.17 Διάγραμμα λειτουργίας του συστήματος σταθεροποίησης της οπτικής προσήλωσης 44

4.1 Εφαρμογή του φίλτρου προσαρμογής CEP. 48
4.2 Βήματα στον υπολογισμό αραιού πεδίου τοπικής ανομοιότητας με cepstrum 50
4.3 Παραγωγή του $i+1$ επιπέδου από το i επίπεδο στην αναπαράσταση της Gaussian πυραμίδας 56
4.4 Παράδειγμα Gaussian πυραμίδας 56
4.5 Παράδειγμα περικομμένης Gaussian πυραμίδας 56
4.6 Συνδυασμός του cepstrum με το φίλτρο προσαρμογής CEP 58

5.1 Πυρήνες Sobel κάθετων ακμών ($2\times 2,\ 3\times 3,\ 4\times 4,\ 5\times 5$) 63
5.2 (α) ολική μάσκα (total mask), (β) κεντρική μάσκα (center mask). 66
5.3 Παραγωγή της Λαπλασιανής πυραμίδας. 68

6.1 Πείραμα σύγκλισης 73
6.2 Η ανομοιότητα του κυλινδρικού αντικειμένου κατά το πείραμα σύγκλισης 74
6.3 Σύγκλιση χωρίς άλλα αντικείμενα στη σκηνή με εντοπισμό της μέγιστης κορυφής στο cepstrum 75
6.4 Σύγκλιση χωρίς άλλα αντικείμενα στη σκηνή με αναζήτηση της κορυφής που αντιστοιχεί στο στόχο με την τεχνική πολλαπλών επιπέδων 75
6.5 Πείραμα παρακολούθησης 76
6.6 Λάθος παρακολούθησης χρησιμοποιώντας το απλό ΦΜΑ συσχέτισης κάθετων ακμών. 77
6.7 Λάθος παρακολούθησης χρησιμοποιώντας το ΦΜΑ συσχέτισης γειτονιάς με μέτρο ποιότητας των χαρακτηριστικών τις οριζόντιες διαφορές των εντάσεων 78
6.8 Λάθος παρακολούθησης χρησιμοποιώντας το ΦΜΑ συσχέτισης γειτονιάς με μέτρο ποιότητας των χαρακτηριστικών την οριζόντια τυπική απόκλιση 79
6.9 Λάθος παρακολούθησης χρησιμοποιώντας το ΦΜΑ συσχέτισης γειτονιάς με μέτρο ποιότητας των χαρακτηριστικών που προκύπτει από την συνέλιξη της στερεοσκοπικής εισόδου με την ολική μάσκα 81
6.10 Λάθος παρακολούθησης χρησιμοποιώντας το ΦΜΑ συσχέτισης γειτονιάς με μέτρο ποιότητας των χαρακτηριστικών που προκύπτει από τον πολλαπλασιασμό των γειτονιών με την κεντρική μάσκα 82
6.11 Λάθος παρακολούθησης χρησιμοποιώντας το ΦΜΑ συσχέτισης μέτρου-κατεύθυνσης των ακμών 83

Α.1 Γεωμετρική άποψη της περιστροφής στο μετασχηματισμό που εκφράζει την pan κίνηση του κεφαλιού. 90

Κεφάλαιο 1

Εισαγωγή

1.1 Σταθεροποίηση προσήλωσης

Ο έλεγχος των παραμέτρων ενός οπτικού συστήματος είναι ένα σημαντικό θέμα στην μηχανική όραση. Ειδικότερα, σύμφωνα με την θεωρία της ενεργού όρασης, ένα σύστημα όρασης πρέπει να ελέγχει συνεχώς τις παραμέτρους της διαδικασίας λήψης του οπτικού σήματος, έτσι ώστε οι απαιτούμενοι υπολογισμοί για την εκδήλωση μιας συμπεριφοράς να είναι ευκολότεροι. Ένα σημαντικό ερευνητικό θέμα στα πλαίσια της θεωρίας αυτής είναι ο έλεγχος προσήλωσης *(gaze control)*, δηλαδή ο έλεγχος των μηχανικών και οπτικών παραμέτρων του συστήματος.

Ο έλεγχος προσήλωσης έχει δύο μορφές : την σταθεροποίηση *(stabilization)* και την μετατόπιση *(change)*. Η σταθεροποίηση, η οποία μας απασχολεί στην παρούσα εργασία, συνίσταται στον έλεγχο των διαθέσιμων βαθμών ελευθερίας για τη λήψη καθαρών εικόνων κάποιου σημείου του κόσμου το οποίο μπορεί είτε να είναι σταθερό, είτε να κινείται σε σχέση με το οπτικό σύστημα. Η μετατόπιση της προσήλωσης είναι πιο σύνθετο πρόβλημα και σχετίζεται περισσότερο με τη μείωση της υπολογιστικής πολυπλοκότητας των οπτικών διεργασιών. Αποσκοπεί στην γρήγορη και ξαφνική μεταφορά της σταθεροποιημένης προσήλωσης σε νέα σημεία καθήλωσης *(fixation points)*, γεγονός το οποίο υποκινείται από οπτικές διεργασίες χαμηλού ή υψηλού επιπέδου. Η δυνατότητα ελέγχου προσήλωσης παρέχει διάφορα πλεονεκτήματα σ' ένα οπτικό σύστημα, όπως είναι η σταθεροποίηση της εικόνας των αντικειμένων που απαιτούν περισσότερη ανάλυση, το μη περιορισμένο πεδίο θέας, και η μικρότερη υπολογιστική πολυπλοκότητα για την διεκπεραίωση κάποιων οπτικών διεργασιών κ.ά.

Ειδικότερα, το πρόβλημα της σταθεροποίησης, είναι ένα πρόβλημα ελέγχου

αρνητικής ανάδρασης λάθους και επομένως οποιαδήποτε λύση του απαιτεί την σχεδίαση ενός εκτιμητή λάθους, ενός ελεγχόμενου μηχανικού συστήματος και ενός νόμου ελέγχου που να οδηγεί το λάθος στο μηδέν. Η είσοδος ενός συστήματος σταθεροποίησης μπορεί να προέρχεται από διαφορετικούς αισθητήρες οι οποίοι μπορούν να έχουν διαφορετικές ταχύτητες δειγματοληψίας και διαφορετικά σημεία συμφόρησης. Η έξοδος μπορεί να κατευθύνεται σε πολλούς αποδέκτες οι οποίοι συνεργάζονται για τη ρύθμιση των εσωτερικών και εξωτερικών παραμέτρων των καμερών.

1.2 Οριοθέτηση της εργασίας

Στην παρούσα εργασία υλοποίηθηκε ένα σύστημα σταθεροποίησης της προσήλωσης (ΣΣΠ) το οποίο παρακολουθεί το κέντρο βάρους του κινούμενου στόχου. Γενικότερα ένα ΣΣΠ μπορεί να έχει διάφορα ανεξάρτητα φωλιασμένα επίπεδα ελέγχου. Η είσοδος του συστήματος είναι οι εικόνες που λαμβάνονται από ένα στερεοσκοπικό σύστημα καμερών και η έξοδος του τροφοδοτεί τους κινητήρες μιας κεφαλής ενεργού όρασης που διαθέτει τέσσερις βαθμούς ελευθερίας.

Η αρχή λειτουργίας του συστήματος στηρίζεται στην πρόταση του Coombs [22] σύμφωνα με την οποία ένα υποσύστημα σύγκλισης και ένα υποσύστημα καταδίωξης συμβιώνουν κάτω από ένα κοινό σχήμα για την επίτευξη της σταθεροποίησης. Η σύγκλιση προϋποθέτει ότι ο στόχος βρίσκεται στο κέντρο των εικόνων και οδηγεί το λάθος στη γωνία σύγκλισης στο μηδέν, ενώ η καταδίωξη προϋποθέτει ότι ο στόχος έχει (σχεδόν) μηδενική ανομοιότητα και οδηγεί το λάθος στη γωνία στροφής στο μηδέν.

Για τον υπολογισμό του λάθους σύγκλισης χρησιμοποιούμε το cepstral φίλτρο, μια μέθοδο συσχέτισης η οποία έχει ζωνοδιαβατό χαρακτήρα, όπως προτείνουν οι Olson και Coombs στο [53]. Η έξοδος του φίλτρου είναι ένα σήμα του οποίου η μέγιστη κορυφή υποθέτουμε οτι αντιστοιχεί στην ανομοιότητα του στόχου. Το φίλτρο γενικά δίνει καλά αποτελέσματα αλλά έχει διάφορα προβλήματα. Ένα τέτοιο πρόβλημα, καθοριστικό για τη συμπεριφορά της σύγκλισης, είναι το πρόβλημα ύπαρξης πολλαπλών κορυφών στην έξοδο που οφείλονται σε αντικείμενα σε διαφορετικό βάθος απ' αυτό που βρίσκεται ο στόχος. Η λάθος επιλογή κορυφής κάνει το σύστημα να αποτυχαίνει. Ένας τρόπος βελτίωσης της σύγκλισης, ως προς το πρόβλημα αυτό, είναι η χρήση τεχνικής πολλαπλών επιπέδων που προτείνουν οι Taylor και Olson [66]. Ο βελτιωμένος αλγόριθμος σύγκλισης, αφ' ενός είναι υπολογιστικά ακριβότερος και αφ' ετέρου κάνει την υπόθεση ότι η θέση του στόχου καθορίζεται από την κεντρική περιοχή της μιας εικόνας. Η αντίστοιχη κάμερα οδηγείται από το σύστημα παρακολούθησης και η άλλη από τη διαδικασία σύγκλισης.

Για τον υπολογισμό του λάθους θέσης χρησιμοποιούμε διάφορα φίλτρα μηδενικής ανομοιότητας που προτείνονται από τους Coombs, Horswil και Kaenel [21, 22]. Τα φίλτρα αυτά είναι ευαίσθητα στην παρουσία υφής σε μια περιοχή κοντά στο σημείο καθήλωσης του στερεοσκοπικού συστήματος. Ένα σύνολο από τέτοια φίλτρα διαφορετικής

πολυπλοκότητας και υπολογιστικής δομής υλοποιήθηκε. Ωστόσο όλα στηρίζονται σε απλή, χαμηλού επιπέδου, επεξεργασία του οπτικού σήματος.

Ο προσανατολισμός της στερεοσκοπικής κεφαλής δοσμένων των εκτιμήσεων ανομοιότητας και θέσης του στόχου στο επίπεδο των καμερών γίνεται με διάφορα σενάρια κατανομής των λαθών σύγκλισης και θέσης στους βαθμούς ελευθερίας του κεφαλιού. Τα σενάρια αυτά εισάγουν διαφορετικά σφάλματα τόσο λόγω της εξάρτησης των βαθμών ελευθερίας μεταξύ τους όσο και λόγω της θέσης του συστήματος η οποία δεν είναι γνωστή με αποτέλεσμα οι διορθώσεις να γίνονται προσεγγιστικά.

Για την αξιολόγηση της συμπεριφοράς του συστήματος αναπτύξαμε μια πλατφόρμα για την προσομοίωση της κινηματικής ρομποτικών συστημάτων. Χρησιμοποιώντας την πλατφόρμα αυτή έγινε προσομοίωση της στερεοκοπικής κεφαλής όπου το ρόλο της κάμερας παίζει ένας raytracer (Rayshade) [34]. Η πλατφόρμα αυτή μπορεί να υποστηρίξει συνθετότερα συστήματα ρομποτικής και να επεκταθεί εύκολα έτσι ώστε να λάβει υπ' όψην τη δυναμική τους.

1.3 Οργάνωση του κειμένου

Στο κεφάλαιο 2 περιγράφεται το πρόβλημα του ελέγχου προσήλωσης στην μηχανική όραση, καθώς επίσης, και βιολογικά μοντέλα συστημάτων ελέγχου προσήλωσης τα οποία αποτελούν πηγή έμπνευσης για την έρευνα στη μηχανική όραση. Στο τέλος του κεφαλαίου γίνεται βιβλιογραφική επισκόπηση της περιοχής. Στο κεφάλαιο 3 αναλύουμε το πρόβλημα της σταθεροποίησης και εξετάζουμε τις επιλογές σχεδίασης που κάναμε για την υλοποίηση ενός ΣΣΠ. Στο κεφάλαιο 4 περιγράφεται το cepstral φίλτρο στο οποίο στηρίζεται ο εκτιμητής λάθους στην σύγκλιση. Στο κεφάλαιο 5 αναφέρονται διάφορα φίλτρα μηδενικής ανομοιότητας στα οποία στηρίζεται ο υπολογισμός λάθους θέσης. Στο κεφάλαιο 6 παρουσιάζουμε την προσομοίωση που υλοποιήθηκε και πειραματικά αποτελέσματα της εφαρμογής του ΣΣΠ και αποτελέσματα των φίλτρων που υλοποιήσαμε σε πραγματικές εικόνες. Τέλος στο κεφάλαιο 7 συνοψίζουμε τις επιλογές μας και τις δυνατότητες του ΣΣΠ που υλοποίηθηκε και προτείνουμε διάφορες κατευθύνσεις για την βελτίωση του.

Κεφάλαιο 2

Έλεγχος προσήλωσης

2.1 Θεωρίες όρασης και έλεγχος προσήλωσης

Η όραση σ'ένα πραγματικό περιβάλλον πρέπει να έχει τρόπους να ελέγχει τις παραμέτρους λήψης εικόνων. Αν φανταστούμε διάφορες απλές δουλειές που κάνουμε καθημερινά, δε θα μας φαίνονταν και τόσο απλές αν δεν είχαμε τρόπους να κινήσουμε τα μάτια, έτσι ώστε να τις υποστηρίξουμε. Μ' άλλα λόγια η στενή σχέση που έχουν η αντίληψη και η ενέργεια διευκολύνεται από τον έλεγχο των παραμέτρων λήψης των εικόνων, και αυτό είναι αναγκαίο για αυτόνομη αλληλεπίδραση με το περιβάλλον.

Η υπολογιστική θεωρία για την όραση του Marr [45] δεν είχε την εκφραστικότητα για να αναδείξει αυτή τη στενή σχέση αντίληψης και ενέργειας. Υπενθυμίζουμε εδώ, ότι η θεωρία του Marr είναι μια κλασική θεωρία στο χώρο της μηχανικής όρασης και αναφέρεται στην ανακατασκευή μιας πλήρους ποσοτικής περιγραφής του κόσμου ικανής να υποστηρίξει κάθε είδους λειτουργία που απαιτεί οπτικές πληροφορίες. Τελικός στόχος της μηχανικής όρασης είναι η απεικόνιση του πραγματικού κόσμου σε μια ακριβή τρισδιάστατη γεωμετρική αναπαράσταση της σκηνής και μια ποσοτική περιγραφή των χαρακτηριστικών των αντικειμένων που περιέχονται σ'αυτή. Η όραση είναι υπεύθυνη για την κάλυψη της πρώτης συνιστώσας του κύκλου : αντίληψη - σκέψη - δράση.

Η αδυναμία της θεωρίας του Marr οδήγησε στην ανάγκη μιας νέας προσέγγισης για την όραση η οποία ονομάζεται ενεργός όραση *(active vision ή animate vision ή utilitarian vision)* [5, 8, 9, 10]. Σύμφωνα με τη θεωρία αυτή, ένα σύστημα όρασης πρέπει να ελέγχει συνεχώς τις παραμέτρους της διαδικασίας απόκτησης των εικόνων, έτσι ώστε να είναι σε θέση να καταγράφει εικόνες που διευκολύνουν τους διάφορους υπολογισμούς που πρέπει να φέρει σε πέρας. Η διάκριση των υποσυστημάτων ενός συστήματος ενεργού

όρασης γίνεται ανάλογα με τις συμπεριφορές που το σύστημα καλείται να επιδείξει και όχι με βάση τη λειτουργική τους σημασία. Οι συμπεριφορές αυτές υλοποιούνται από ένα σύνολο διαδικασιών *(processes)* οι οποίες συνεργάζονται για την επίτευξη των στόχων του συστήματος. Στόχος της ενεργού όρασης είναι να μετασχηματίσει τα οπτικά ερεθίσματα σε ακολουθίες από δράσεις που αλληλεπιδρούν με τον κόσμο (συμπεριφορές *(behaviors)*). Για μια εκτενή σύγκριση των θεωριών για την όραση βλ. [1].

Η νέα υπολογιστική θεωρία για την όραση διαθέτει μια καλή σειρά από ενδιαφέροντα ερευνητικά θέματα. Ανάμεσα σ'αυτά είναι και ο έλεγχος προσήλωσης *(gaze control)* Στα πλαίσια της θεωρίας αυτής ο έλεγχος των μηχανικών και οπτικών παραμέτρων του συστήματος (ή έλεγχος προσήλωσης (*gaze control)*) έχει στόχο να παρέχει υπολογιστικό πλεονέκτημα.

Το πρόβλημα του ελέγχου προσήλωσης είναι στενά συνδεδεμένο με το πρόβλημα επιλογής νέων σημείων καθήλωσης, γνωστό ως 'where to look next'. Αυτό είναι πολύ σύνθετο και μπορεί να στηρίζεται σε πολλά επίπεδα επεξεργασίας. Λόγω της στενής σχέσης που έχουν τα προβλήματα ελέγχου προσήλωσης και 'where to look next', τα ερεθίσματα για αλλαγές της προσήλωσης σχετίζονται με τη μελέτη του ελέγχου προσήλωσης.

2.2 Μορφές και πλεονεκτήματα ελέγχου προσήλωσης

Ο έλεγχος προσήλωσης ως πρόβλημα έχει δύο μορφές :την σταθεροποίηση *(gaze stabilization)* και την μετατόπιση *(gaze change)*. Η σταθεροποίηση συνίσταται στον έλεγχο των διαθέσιμων βαθμών ελευθερίας για τη λήψη καθαρών εικόνων κάποιου σημείου του κόσμου (σε επιφάνεια αντικειμένου) το οποίο μπορεί να είναι σταθερό ή να κινείται σε σχέση με το οπτικό σύστημα. Στη δεύτερη περίπτωση μιλάμε για παρακολούθηση. Η μετατόπιση προσήλωσης είναι πιο σύνθετο πρόβλημα και είναι πιο στενά συνδεδεμένο με τη μείωση της υπολογιστικής πολυπλοκότητας των οπτικών διεργασιών. Μπορεί να χρησιμοποιηθεί για γρήγορη και ξαφνική μεταφορά της σταθεροποιημένης προσήλωσης σε νέα σημεία καθήλωσης, έτσι ώστε να υποστηρίξει τη λύση οπτικών διεργασιών χαμηλού ή υψηλού επιπέδου, δηλ. οι μετατοπίσεις της προσήλωσης υποκινούνται από τη διεργασία.

2.2.1 Πλεονεκτήματα

Στην ανθρώπινη όραση η ανάγκη για μηχανισμούς οπτικής προσήλωσης είναι προφανής λόγω της ανομοιόμορφης κατανομής των φωτοϋποδοχέων στον αμφιβληστροειδή και το μικρό μέγεθος της ωχρής κηλίδας *(fovea)*. Στα συστήματα ρομποτικής δεν υπάρχει αυτή η ανάγκη, παρ'όλο που υπάρχει εξέλιξη στην έρευνα εναλλακτικών αισθητήρων όρασης με ανομοιόμορφη κατανομή. Ωστόσο ο έλεγχος της προσήλωσης παρέχει αρκετά πλεονεκτήματα ούτως ή άλλως. Αυτά υπογραμμίστηκαν στο [65] και θα τα αναφέρουμε

κι εδώ:

- **σταθεροποίηση της εικόνας:** Η σταθεροποίηση της προσήλωσης σε κινούμενα αντικείμενα εμποδίζει την εικόνα του αντικειμένου να μεταφέρεται (translational image blur) κι έτσι προκύπτουν απλούστερα σχήματα ελέγχου. Εξάλλου το αντικείμενο ενδιαφέροντος είναι στο πεδίο θέας για μεγαλύτερο χρονικό διάστημα, πράγμα που επιτρέπει την χρήση υψηλότερης ανάλυσης έτσι ώστε να είναι δυνατή μια λεπτομερέστερη εξέταση των ιδιοτήτων του κινούμενου αντικειμένου..

- **μη περιορισμένο πεδίο θέας:** Νέα τμήματα της σκηνής απεικονίζονται και τα προβλήματα επικάλυψης *(σκίασης - occlusion)* αντιμετωπίζονται με κινήσεις των καμερών. Μια ειδική περίπτωση περιορισμένου πεδίου θέας προκύπτει από τη χρήση αισθητήρων που έχουν μια κεντρική περιοχή υψηλής ανάλυσης και μια χαμηλότερης ανάλυσης περιφέρεια. Δηλαδή ο έλεγχος προσήλωσης αναδεικνύει τα πλεονεκτήματα των foveal αισθητήρων.

- **μείωση της υπολογιστικής πολυπλοκότητας:** Σ'ένα ενεργό σύστημα η δυνατότητα μετατόπισης του σημείου λειτουργίας του συστήματος, έτσι ώστε οι παράμετροι λήψης των εικόνων να ταιριάζουν στο αντικείμενο προσήλωσης, απλοποιεί πολλούς υπολογισμούς. Για παράδειγμα, όταν ένα αντικείμενο βρίσκεται κοντά στον οπτικό άξονα της κάμερας η εικόνα μπορεί να προσεγγιστεί από ορθογώνια προβολή.

- **διαχωρισμός αντικειμένου-υποστρώματος:** Αν ένα αντικείμενο παρακολουθείται, η οπτική ροή γι' αυτό το αντικείμενο μειώνεται σε πολύ μικρές τιμές σε σχέση με την υπόλοιπη σκηνή. Οι διαφορές που εμφανίζονται στα πεδία ροής μπορούν να χρησιμοποιηθούν για την τμηματοποίηση του αντικειμένου που παρακολουθείται από το υπόστρωμα.

- **υπολογισμός διάταξης (range):** Όταν μια κινούμενη κάμερα σταθεροποιεί την προσήλωση σε στατικό σημείο, η οπτική ροή γι' αυτό το σημείο θα έχει μέτρο μηδέν. Οι κατευθύνσεις στο υπόλοιπο πεδίο οπτικής ροής μπορούν να χρησιμοποιηθούν για την διάταξη των υπόλοιπων σημείων της σκηνής σε σχέση με το σημείο προσήλωσης.

- **έλεγχος του βάθους πεδίου:** Το βάθος πεδίου μπορεί να αλλάξει με ρύθμιση του ανοίγματος (aperture) του φακού, του zoom και της ταχύτητας του διαφράγματος (shutter speed). Έτσι μπορούμε να έχουμε μεγάλο ή μικρό βάθος πεδίου ανάλογα με τις υποθέσεις που γίνονται για την υλοποίηση κάποιας οπτικής συμπεριφοράς.

2.2.2 Σταθεροποίηση προσήλωσης

Το πρόβλημα της σταθεροποίησης προσήλωσης είναι ένα πρόβλημα ελέγχου. Ως πρόβλημα ελέγχου αποτελείται από έναν εκτιμητή λάθους, ένα ελεγχόμενο μηχανικό

σύστημα και ένα νόμο ελέγχου ο οποίος προσπαθεί να οδηγήσει το λάθος στο μηδέν. Η έρευνα του προβλήματος αυτού κτίζεται πάνω στη δουλειά που έχει γίνει στα πλαίσια της παραδοσιακής θεωρίας ελέγχου, και σχετίζεται με προχωρημένα προβλήματα της θεωρίας αυτής.

Ωστόσο, το πρόβλημα είναι αρκετά ενδιαφέρον και από την άποψη της όρασης, δηλαδή την οπτική επεξεργασία που χρειάζεται για την εκτίμηση λάθους. Η εκτίμηση πρέπει να στηρίζεται σε τεχνικές μηχανικής όρασης για ακολουθίες εικόνων.Αυτό σημαίνει εξέλιξη στους αλγόριθμους δυναμικής υπολογιστικής όρασης. Η τμηματοποίηση είναι ένα σημαντικό πρόβλημα σ'αυτή την κατεύθυνση, αφού ο στόχος στον οποίο είναι επιθυμητή η σταθεροποίηση πρέπει να εξαχθεί και να διαχωριστεί από το υπόστρωμα. Επίσης σημαντικό πρόβλημα για μελέτη είναι τα μοντέλα θορύβου, η ευαισθησία, η σταθερότητα και η απόδοση για δειγματοληπτημένα χωρο-χρονικά σήματα.

Το πρόβλημα της σταθεροποίησης εμπλέκει **εισόδους** από διαφορετικούς αλγορίθμους όρασης και διαφορετικούς αισθητήρες οι οποίοι με τη σειρά τους μπορούν να έχουν διαφορετικές ταχύτητες δειγματοληψίας και διαφορετικά υπολογιστικά bottlenecks (αδύνατα σημεία - σημεία συμφόρησης). Η **έξοδος** πρέπει να οδηγεί πολλούς αποδέκτες *(effectors)* που συνεργάζονται για να ρυθμίσουν εσωτερικές και εξωτερικές παραμέτρους των καμερών. Η διαδικασία της σταθεροποίησης μπορεί να εμπλέκει μόνο έλεγχο χαμηλού επιπέδου (παρακολούθηση centroid) ή διάφορα ανεξάρτητα φωλιασμένα στρώματα - επίπεδα ελέγχου (παρακολούθηση ενός ατόμου μέσα σε πλήθος) για να συνεργαστούν με διάφορους τύπους και επίπεδα οπτικών διεργασιών.

Η απόδοση ενός συστήματος σταθεροποίησης προσήλωσης επηρεάζεται από το θόρυβο, τις καθυστερήσεις και την ανοχή *(dropout)* των αλγορίθμων. Αν το σύστημα εμπλουτιστεί με ικανότητες πρόβλεψης μπορεί να γίνει πιο εύρωστο. Αυτές οι ικανότητες αποδίδονται με τη δημιουργία και εκμετάλλευση μοντέλων του ελεγχόμενου συστήματος, του σήματος, και της διαδικασίας ελέγχου ανάδρασης. Ο έλεγχος πρέπει να προσαρμόζεται στην γνώση που αποκτούμε για το στόχο και το ελεγχόμενο σύστημα κατά την πάροδο του χρόνου.

2.2.3 Μετατόπιση προσήλωσης

Οι μηχανισμοί μετατόπισης της προσήλωσης διευκολύνουν την δυναμική ολοκλήρωση διαφορετικών μορφών οπτικής πληροφορίας.

Στην ενεργό όραση, ο ανεξάρτητος χειρισμός των παραμέτρων λήψης εικόνων δημιουργεί την ανάγκη για σχήματα βέλτιστης ενοποίησης των διαφόρων πηγών ολικής πληροφορίας. Οι πηγές αυτές πρέπει να συνδυάζονται με συμβατό τρόπο. Συνήθως η αξιοπιστία της πληροφορίας που παίρνουμε με τη χρήση των πηγών αυτών διαφέρει ανάλογα με τη διάταξη του ενεργού συστήματος. Η μετατόπιση προσήλωσης μπορεί να υποστηρίξει τον έλεγχο, ενοποίηση και ανάλυση της σύγκρουσης αυτών των οπτικών

πηγών.

Η εξέταση διαφορετικών σημείων στη σκηνή απαιτεί διάφορους συνδυασμούς των παραμέτρων λήψης εικόνων. Το γεγονός αυτό υποδηλώνει ότι είναι αναγκαία η μελέτη και ανάπτυξη στρατηγικών για τη λήψη και επεξεργασία εικόνων. Οι στρατηγικές αυτές πρέπει να αναφέρονται στο περιεχόμενο των εικόνων και την υπολογιστική πολυπλοκότητα της οπτικής διαδικασίας. Οι στρατηγικές ελέγχου των παραμέτρων του συστήματος λήψης εικόνων σχεδιάζονται σύμφωνα με κάποιο μοντέλο ανάλυσης. Οι αισθητήρες ανομοιόμορφης ανάλυσης απαιτούν διαφορετική δομή στρατηγικών ελέγχου. Από τα παραπάνω φαίνεται ότι η μετατόπιση προσήλωσης μπορεί να απαιτεί και τον έλεγχο του βαθμού ανάλυσης της εικόνας όπως στην περίπτωση του ελέγχου εστίασης *(focus control)*.

Η υπολογιστική πολυπλοκότητα και η παραχώρηση των πόρων επίσης εμπλέκονται στο πρόβλημα της μετατόπισης προσήλωσης γιατί η επιλογή παραμέτρων και η ανακατανομή *(reallocation)* υπολογιστικών πόρων απαιτεί κάποιους συμβιβασμούς. Η συνεργασία της επεξεργασίας και του ελέγχου μπορεί να οδηγήσει σε πιο αποδοτικές και αποτελεσματικές μεθόδους ελέγχου προσήλωσης.

2.3 Βιολογικά συστήματα ελέγχου προσήλωσης

2.3.1 Σχέση νευροφυσιολογίας και μηχανικής όρασης

Πηγή έμπνευσης για πολλούς ερευνητές στην ενεργό μηχανική όραση αποτέλεσε το ανθρώπινο οφθαλμοκινητικό σύστημα και γενικότερα τα συστήματα όρασης που υπάρχουν στη φύση. Αυτό είναι πολύ λογικό αφού το οφθαλμοκινητικό σύστημα στα πρωτεύοντα θηλαστικά είναι ευέλικτο και λειτουργικό. Η σχέση νευροφυσιολογίας και μηχανικής όρασης είναι όμως αμφίδρομη. Η ανάπτυξη ρομποτικών συστημάτων που υλοποιούν υπολογιστικά μοντέλα ελέγχου προσήλωσης παρέχει πειραματική δυνατότητα που συμπληρώνει τις αντίστοιχες που χρησιμοποιούνται στην ψυχοφυσική και στη νευροεπιστήμη. Για παράδειγμα οι βιολόγοι δεν μπορούν να χαρακτηρίσουν την ποιότητα της απόδοσης της κίνησης του ματιού με παραδοσιακά μέτρα ελέγχου. Αυτό οφείλεται εν μέρει στο γεγονός ότι οι άνθρωποι είναι πολύ εξειδικευμένοι στο να εμφανίζουν προβλεπτικότητα σ' ένα ερέθισμα (π.χ. εκφράζουν την ποιότητα της ομαλής καταδίωξης με ένα αριθμό catch-up σακκαδικών κινήσεων που εκδηλώνει το υποκείμενο σε ορισμένη χρονική περίοδο.)

2.3.2 Το ανθρώπινο οφθαλμοκινητικό σύστημα

Η σημαντικότερη διαφορά του ανθρώπινου ματιού από τις υπάρχουσες ηλεκτρονικές κάμερες είναι το ότι διαθέτει πολύ καλύτερη ανάλυση σε μια πολύ μικρή περιοχή κοντά στον οπτικό άξονα. Αυτή η περιοχή είναι γνωστή ως ωχρή κηλίδα *(fovea)* και καταλαμβάνει λιγότερο από 0.01% του οπτικού πεδίου. Στην ωχρή κηλίδα η

ανάλυση είναι καλύτερη κατά μια τάξη μεγέθους σε σχέση με την περιφέρεια. Αυτή η διάταξη χαρακτηρίζεται από την αναπαράσταση ενός μεγάλου πεδίου θέας και μεγάλη ακρίβεια τοπικά. Με μια τέτοια ανομοιόμορφη κατανομή των φωτοϋποδοχέων στον αμφιβληστροειδή δεν είναι παράξενο που το ανθρώπινο μάτι έχει ειδικές συμπεριφορές για πολύ γρήγορη κίνηση έτσι ώστε να απεικονίζονται στην ωχρή κηλίδα διαφορετικά αντικείμενα στο χώρο.

Το ανθρώπινο οφθαλμοκινητικό σύστημα είναι πολύ σύνθετο και δεν έχει κατανοηθεί μέχρι τώρα πλήρως. Περιέχει πολλά λειτουργικά κομμάτια τα οποία αλληλεπιδρούν μεταξύ τους κι έτσι είναι δύσκολο να δώσει κανείς αυστηρούς και γρήγορους κανόνες για το τι μπορούν και τι δεν μπορούν να διεκπεραιώσουν. Ωστόσο, μπορούμε να υπογραμμίσουμε κάποια σημαντικά σημεία της έρευνας στη νευροφυσιολογία που στοχεύει στην κατανόηση του συστήματος αυτού. Το οπτικό σύστημα εκδηλώνει κάποιες συμπεριφορές οι οποίες υποκινούνται από οπτικά καθώς επίσης και από μη οπτικά ερεθίσματα. Θα αναφερόμαστε σ'αυτές τις συμπεριφορές με τον όρο **αντανακλαστικά**. Από την έρευνα στη νευροφυσιολογία μας αφορούν τα σημεία που αναφέρονται στα αντανακλαστικά.

2.3.2.1 Αντανακλαστικό ομαλής καταδίωξης και οπτοκινητικό αντανακλαστικό

Ο μηχανισμός ομαλής καταδίωξης *(smooth pursuit)* εκδηλώνεται στην παρακολούθηση των διαδρομών που κάνουν μικροί στόχοι και συνήθως καθοδηγείται από το ολίσθημα στην ωχρή κηλίδα του αμφιβληστροειδή *(foveal retinal slip)*. Ο μηχανισμός αυτός δεν μπορεί να αρχικοποιηθεί εκούσια. Το οπτο-κινητικό αντανακλαστικό *(opto-kinetic reflex (OKR))* κάνει τα μάτια να ακολουθούν μια κίνηση όλου του οπτικού πεδίου, και καθοδηγείται από το ολίσθημα στον αμφιβληστροειδή (retinal slip), ή οπτική ροή. Στα πρωτεύοντα θηλαστικά *(primates)* το OKR έχει δύο φάσεις: μία γρήγορη (κατ' ευθείαν) και μία αργή (παρεπόμενη). Η άμεση φάση είναι η κυρίαρχη στο ανθρώπινο οπτικό σύστημα. Κάποιοι μάλιστα πιστεύουν ότι το σύστημα ομαλής παρακολούθησης είναι το άμεσο κομμάτι του OKR [19].

Οι ομαλές κινήσεις κανονικά δεν μπορούν να επαχθούν από ένα μη ομαλά κινούμενο ερέθισμα, αλλά εξακολουθούν μετά την εξαφάνιση του. Το γεγονός αυτό είναι ένδειξη για το ότι υπάρχει κάποιου είδους πρόβλεψη [25]. Η ομαλή παρακολούθηση στους πιθήκους φαίνεται να καθοδηγείται όχι μόνο από το λάθος ταχύτητας *(velocity error)* αλλά επίσης από τη θέση και τα λάθη επιτάχυνσης *(acceleration errors)*.

Ο έλεγχος στο σύστημα ομαλής παρακολούθησης συχνά προσεγγίζεται σαν σύστημα κλειστού βρόγχου συνεχούς χρόνου με αρνητική ανάδραση. Ευστάθεια στο σύστημα επιτυγχάνεται με μιά εσωτερική θετική ανάδραση του σήματος κίνησης του ματιού. Έτσι η κίνηση του στόχου στο περιβάλλον ανακατασκευάζεται από το λάθος ταχύτητας του ερεθίσματος στον αμφιβληστροειδή κι ένα καθυστερημένο αντίγραφο τού σήματος του κινητήρα (αποδέκτη).

2.3.2.2 Σακκαδικές κινήσεις

Σακκαδικές κινήσεις *(saccades)* είναι οι γρήγορες κινήσεις οι οποίες μετατοπίζουν την προσήλωση πολύ γρήγορα (300 - 400 μοίρες το sec). Ελέγχονται εκούσια και μπορούν να συμβούν στο σκοτάδι. Το σακκαδικό σύστημα έχει μεγαλύτερη καθυστέρηση από το σύστημα ομαλής παρακολούθησης. Το γεγονός αυτό αντανακλά τις αρχές ελέγχου υψηλότερου επιπέδου που το διέπουν. Το σύστημα αυτό συχνά μοντελοποιείται ως σύστημα δειγματοληπτημένων δεδομένων. Η ευστάθεια του οφείλεται σ' ένα μηχανισμό latency and trigger ο οποίος απαγορεύει την επαναπυροδότηση προτού το σύστημα ηρεμήσει.

2.3.2.3 Αντανακλαστικό σύγκλισης, αντανακλαστικό προσαρμογής και αντανακλαστικό κοντινής τριάδας

Σύγκλιση ονομάζουμε εδώ τη διαδικασία ρύθμισης της γωνίας που ορίζουν οι δύο οπτικοί άξονες των ματιών έτσι ώστε να κατευθύνονται στο ίδιο σημείο του χώρου. Το κύριο ερέθισμα για τις κινήσεις σύγκλισης είναι το στερεοσκοπικό disparity. Για το ανθρώπινο οπτικό σύστημα, η προσαρμογή *(accommodation)* και η σύγκλιση είναι ασθενείς απόλυτες πηγές βάθους.

Το ανθρώπινο οπτικό σύστημα χρησιμοποιεί αυτούς τους δύο μηχανισμούς μαζί. Μια όψη αυτού του συνδυασμού είναι η συγκλιτική προσαρμογή *(convergence accommodation)*. Καθώς τα μάτια συγκλίνουν, προσαρμόζονται σαν να ήταν να εστιάσουν σε αντικείμενα πιο κοντά. Η ανάποδη όψη αυτού του συνδυασμού είναι η προσαρμοστική σύγκλιση *(accommodative convergence)* : όταν το ένα μάτι προσαρμόζεται σ'ενα στόχο, οι οπτικοί άξονες συγκλίνουν για να πέσουν στο στόχο. Με λίγα λόγια, η σύγκλιση ελλείψει disparity μπορεί να οδηγήσει την προσαρμογή και η προσαρμογή ελλείψει disparity μπορεί να οδηγήσει τη σύγκλιση και την προσαρμογή στ' άλλο μάτι. Υπάρχει και τρίτη όψη αυτού του ταιριάσματος, η οποία ονομάζεται μεταβλητότητα (variability). Οι οπτικές παράμετροι αλλάζουν σύμφωνα με το οπτικό ερέθισμα. Οι Miles, Judge και Optican [46] μελέτησαν ανθρώπους που φόρεσαν διάφορες οπτικές συσκευές (περισκόπια και πρίσματα) πριν και μετά , και επιβεβαίωσαν την ύπαρξη στοιχείων προσαρμογής που ρυθμίζουν την απόκλιση των συστημάτων σύγκλισης και/ή προσαρμογής.

Το αντανακλαστικό κοντινής τριάδας (near triad reflex) αποτελείται από τα συστήματα σύγκλισης, προσαρμογής και το σακκαδικό. Κατά τη λειτουργία αυτού του αντανακλαστικού, όταν συμβαίνει μια σακκαδική κίνηση, η εστίαση και η σύγκλιση οδηγούνται στη σωστή κατεύθυνση και γρηγορότερα απ' ότι κανονικά.

2.3.2.4 Θαλαμιαίο οφθαλμικό αντανακλαστικό

Το θαλαμιαίο οφθαλμικό αντανακλαστικό *(vestibulo-ocular reflex (VOR))* σταθεροποιεί την προσήλωση εξουδετερώνοντας τις κινήσεις του κεφαλιού με τις κινήσεις του

ματιού και είναι το γρηγορότερο οπτικό αντανακλαστικό ($\approx 16ms$ καθυστέρηση). Μοντελοποιείται ως έλεγχος ανοικτού βρόγχου, με την έννοια ότι η έξοδος του θαλαμιαίου αισθητήρα μετατρέπεται σε είσοδο των μυών του ματιού και δίδεται μέσω ενός μονοπατιού με περίπου τρεις συνάψεις. Μπορεί να είναι έλεγχος με υψηλό κέρδος (≈ 1): μπορεί συχνά να ακυρώνει με ακρίβεια τις επιδράσεις της κίνησης του κεφαλιού. Γεννιέται λοιπόν το ερώτημα πως μοντελοποιείται το σύστημα εσωτερικά έτσι ώστε να είναι εφικτός ο έλεγχος του με ανοικτό βρόγχο.

Έρευνα για το VOR ανέδειξε τη γεωμετρική όψη αυτού του μοντέλου: την μετατροπή των σημάτων των αισθητήρων στα συστήματα συντεταγμένων των ημικυκλικών καναλιών, σε σήματα effectors για τους μυες των ματιών που βρίσκονται σε ποικίλες θέσεις.

Η είσοδος του VOR προέρχεται από τα γραμμικά και γωνιακά επιταχυνσιόμετρα των ωτολιθικών *(otolith)* οργάνων και των ημικυκλικών καναλιών.

Μια σημαντική άποψη του VOR είναι η προσαρμοστικότητα του. Το αντανακλαστικό προσαρμόζεται σε αλλαγές του οπτικού συστήματος στο χρόνο [60]. Το VOR αλληλεπιδρά με άλλα αντανακλαστικά και το ερέθισμα το οποίο τα προκάλεσε.

2.4 Ενεργά συστήματα μηχανικής όρασης και έλεγχος προσήλωσης

2.4.1 Εξέλιξη αυτόνομων συστημάτων

Στα αυτόνομα συστήματα κυριαρχούν δύο σημαντικά χαρακτηριστικά : Η κινητικότητα (mobility) και η αυτονομία (autonomy). Η κινητικότητα παρέχεται από τη μηχανική του συστήματος και η αυτονομία από το λογισμικό και το υπολογιστικό hardware. Τα τελευταία χρόνια υπάρχει πρόοδος στη κατασκευή μικρών και δυνατών κινητήρων και οι υπολογιστές γίνονται όλο και πιο γρήγοροι. Σαν αποτέλεσμα αυτής της εξέλιξης υπάρχουν σήμερα προηγμένα συστήματα ρομποτικής σε εργαστήρια έρευνας ([69, 28, 55, 62, 47]. Η αυτονομία υστερεί ως προς την εξέλιξη σε σχέση με τη κινητικότητα και αυτό σχετίζεται με την όχι επαρκή κατανόηση για την αντίληψη του κόσμου και την αναπαράσταση της γνώσης.

Η ενεργός όραση με τη πρόθεση να διευκολύνει την αλληλεπίδραση της ενέργειας και της αντίληψης δίνει ώθηση στην εξέλιξη της αυτονομίας στα συστήματα ρομποτικής. Στα πλαίσια της θεωρίας αυτής η έρευνα εφαρμόζεται κατά κόρον σε συστήματα τα οποία διαθέτουν ανθρωπομορφικά χαρακτηριστικά (π.χ. head-eye συστήματα). Η τάση αυτή έχει τις ρίζες της στην εξής διαπίστωση : αν ο στόχος της υπολογιστικής όρασης είναι η κατανόηση των οπτικών συστημάτων, τα συστήματα ρομποτικής πρέπει να έχουν τη δυνατότητα να μιμούνται όψεις της όρασης των θηλαστικών και να λειτουργούν κάτω από παρόμοιες συνθήκες (το σύστημα πρέπει να βλέπει το κόσμο και όχι εικόνες που έχουν προκαταγραφεί [55]), έτσι ώστε να επιτυγχάνουν έστω και χαλαρά το συνδυασμό

απόδοσης και ευελιξίας που διαθέτει το ανθρώπινο οπτικό σύστημα.

Στα πλαίσια αυτού του προβληματισμού για τον κατάλληλο εξοπλισμό υποστήριξης της υπολογιστικής όρασης υπάρχει εξέλιξη όσον αφορά και τους αισθητήρες. Οι επίπεδες κάμερες ομοιόμορφης ανάλυσης αμφισβητούνται για το ρόλο που μπορούν να παίξουν στο σχεδιασμό δυνατών αλγορίθμων. Η αμφισβήτηση αυτή έχει οδηγήσει στην έρευνα εναλλακτικών αισθητήρων και στην ανάπτυξη κατάλληλων αλγόριθμων που να εκμεταλλεύονται την ανομοιόμορφη ανάλυσή τους [67, 1].

2.4.2 Συστήματα ελέγχου προσήλωσης

Σύμφωνα με τα όσα είπαμε στις προηγούμενες παραγράφους για τη σημασία του ελέγχου προσήλωσης στα ενεργά συστήματα όρασης, είναι αναμενόμενη η εμφάνιση ποικίλων τέτοιων συστημάτων που έχουν να επιδείξουν διαφορετικές ικανότητες ή αδυναμίες. Ποικιλία υπάρχει τόσο στην φιλοσοφία των αλγορίθμων, με τους οποίους λειτουργούν, όσο και στις πλατφόρμες ανάπτυξης τους. Ο τρόπος με τον οποίο προσεγγίζουν το πρόβλημα καθώς επίσης και η έμφαση σε διαφορετικά σημεία έρευνας διακρίνει τα συστήματα αυτά μεταξύ τους.

Το πρόβλημα της σταθεροποίησης της προσήλωσης έχει ελκύσει το ενδιαφέρον αρκετών ερευνητών τα τελευταία χρόνια. Αντίθετα δεν έχει γίνει αρκετή έρευνα για το πρόβλημα μετατόπισης της προσήλωσης πράγμα που οφείλεται στην μεγαλύτερη πολυπλοκότητα του προβλήματος.

2.4.2.1 Συστήματα σταθεροποίησης της προσήλωσης

Στερεοσκοπικά συστήματα

Οι David Coombs και Christopher Brown [20] ήταν απ' του πρώτους που έδειξαν απόδοση πραγματικού χρόνου σε σύστημα ενεργού όρασης (συχνότητα απόκρισης 7.5-8 Hz), το οποίο είναι ένα στερεοσκοπικό σύστημα παρακολούθησης. Οι κάμερες παρακολουθούν ένα κινούμενο αντικείμενο με την υπόθεση ότι το σύστημα είναι αρχικά καθηλωμένο στο αντικείμενο προς παρακολούθηση. Ο εντοπισμός του στόχου δεν γίνεται με την αναγνώριση των ιδιοτήτων του (π.χ. σχήμα, χρώμα), αλλά με στερεοσκοπική πληροφορία : η σύγκλιση παράγει ένα ορόπτερο *(horopter)* το οποίο ορίζεται ως επιφάνεια μηδενικής στερεοσκοπικής ανομοιότητας (disparity), κι έτσι ο στόχος εντοπίζεται με ένα απλό φίλτρο μηδενικής ανομοιότητας *(zero disparity filter).* Αυτό το σύστημα έχει κάποιες αδυναμίες: δεν μπορεί να παρακολουθήσει πράγματα που βρίσκονται μακριά ή που έρπονται στο πάτωμα, και δεν έχει καλά αποτελέσματα σε σκηνές που περιέχουν αντικείμενα με ποικίλες ανομοιότητες (δηλ. σε διάφορα βάθη). Το υποσύστημα σύγκλισης στο οποίο οφείλεται η δεύτερη αδυναμία περιγράφεται αναλυτικότερα στο [53]. Για την αντιμετώπιση αυτού του προβλήματος οι Taylor, Olson kai Martin [66] εισάγουν την χρήση scale-space θεωρώντας ότι υπάρχει μία κύρια κάμερα που κάνει την παρακολούθηση και μία δευτερεύουσα με την οποία επιτυγχάνεται η σύγκλιση.

Ο έλεγχος των ανεξάρτητων pan, tilt κινητήρων γίνεται με PID (Proportional Integral Derivative) [1] ελεγκτές και φαίνεται να είναι αποτελεσματικός αλλά ωστόσο εισάγει κάποιες καθυστερήσεις. Στη διδακτορική διατριβή του David Coombs [22] αναλύεται περισσότερο το θέμα των καθυστερήσεων και η χρήση πρόβλεψης. Σε μια δημοσίευση των Coombs,Christopher και Soong [14] εκτός από τους παραπάνω οπτοκινητικούς ελέγχους περιγράφεται κι ένας μη οπτικός έλεγχος για την σταθεροποίηση της προσήλωσης ο οποίος αντισταθμίζει την μεταφορική και περιστροφική κίνηση του κεφαλιού με pan και tilt κινήσεις του κεφαλιού, χρησιμοποιώντας σαφή κινηματική πληροφορία.

Η ιδέα χρησιμοποίησης της σύγκλισης όπως την περιγράψαμε παραπάνω έχει χρησιμοποιηθεί ευρέως στη συνέχεια και από άλλους ερευνητές.

Στα [61, 33] λόγω των πόρων hardware που διαθέτουν αναπτύσσουν έναν οικονομικότερο αλγόριθμο σύγκλισης ο οποίος εισάγει την έννοια των "εικονικών ορόπτερων" ("virtual horopters"). Το εικονικό ορόπτερο είναι το ορόπτερο που δημιουργείται μετακινώντας οριζόντια την δεξιά κάμερα κατά ένα συγκεκριμένο αριθμό pixels. Μικρές μετατοπίσεις είναι σχεδόν ισοδύναμες με εικονικές περιστροφές.

Ο Kam [32] ανέπτυξε ένα στερεοσκοπικό σύστημα παρακολούθησης πραγματικού χρόνου το οποίο δεν είναι σύστημα ενεργού όρασης αλλά ακολουθεί την κλασική θεωρία της τεχνητής νοημοσύνης για τον κύκλο αντίληψη-σκέψη-ενέργεια. Υποθέτει ότι η ιδία κίνηση είναι γνωστή, γεγονός το οποίο δεν ισχύει στη γενική περίπτωση. Ο τρόπος που προσεγγίζεται σ' αυτή την εργασία η εξαγωγή του στόχου από το υπόστρωμα δεν είναι γενικός, με αποτέλεσμα το σύστημα να μην μπορεί να χρησιμοποιήσει ταυτόχρονα την κίνηση σύγκλισης και τις pan-tilt κινήσεις, γιατί χρειάζεται διαφορετικό χειρισμό ως προς την τμηματοποίηση.

Στο KTH [56, 57] επισημαίνουν ότι το θάμπωμα (blur) μπορεί να χρησιμοποιηθεί ως πληροφορία γιά βάθος και επιχειρηματολογούν με βάση το ανθρώπινο σύστημα σύγκλισης. Αναπτύσσουν μια τέτοια ολοκλήρωση των δύο αυτών πηγών πληροφορίας και αναλύουν την συνεργασία και τον ανταγωνισμό τους.

Στα [70, 51] περιγράφεται ένα σύστημα παρόμοιο με αυτό στο [20] (η σημαντικότερη ομοιότητα ίσως είναι το ότι επιχειρηματολογούν για εύρωστη συμπεριφορά με την χρήση απλών αλγορίθμων που ενοποιούνται με κάποιο εκφραστικό τρόπο) αλλά χρησιμοποιείται και ανίχνευση κίνησης απ' την οποία καθώς επίσης και από το πεδίο ανομοιότητας υπολογίζεται μια μάσκα (φίλτρο) του στόχου που αναδρά παίζοντας το ρόλο πρόβλεψης και το σύστημα γίνεται πιο ανθεκτικό.

[1] Η λειτουργία του PID ελεγκτή παριστάνεται από την εξίσωση :

$$m_n = K_g \left(K_p e_n + K_i \sum_{k=0}^{n} c^{n-k} e_k + K_d \frac{e_n - e_{n-1}}{\Delta t} \right) \tag{2.1}$$

όπου m_n είναι η έξοδος του ελεγκτή, K_g είναι το ολικό κέρδος (global gain), K_p είναι το αναλογικό κέρδος (proportional gain), K_i είναι το κέρδος ολοκλήρωσης (integral gain), K_d είναι το διαφορικό κέρδος (differential gain), c είναι η σταθερά καθυστέρησης της ολοκλήρωσης (integral decay constant), e_n είναι το λάθος στην n-οστή επανάληψη και Δt είναι το διάστημα δειγματοληψίας.

Χρησιμοποιώντας τις βασικές αρχές που διέπουν την προσέγγιση [20], οι Huber και Kortenkamp [30] κατασκεύασαν ένα σύστημα καταδίωξης για ένα κινούμενο ρομπότ. Ο αλγόριθμος οπτικής παρακολούθησης συνδυάζεται με έναν αλγόριθμο αποφυγής εμποδίων. Όσον αφορά την οπτική παρακολούθηση εισάγεται η έννοια εγγύς πεδίο *(proximity space)*, το οποίο είναι μια περιορισμένη τρισδιάστατη περιοχή του χώρου για την οποία γίνεται όλη η οπτική επεξεργασία και έχει το κέντρο της περίπου στο σημείο καθήλωσης. Η παρακολούθηση ορίζεται ως η μετακινηση του εγγέως πεδίου σύμφωνα με την κίνηση του καταδιωκόμενου. Το διάνυσμα της μετακίνησης προκύπτει από το άθροισμα των εξής διανυσμάτων: ένα δισδιάστατο διάνυσμα το οποίο επιχειρεί να κινήσει το εγγύς πεδίο σύμφωνα με την κατεύθυνση της κίνησης, ένα δισδιάστατο διάνυσμα το οποίο επιχειρεί να κινήσει το εγγύς πεδίο προς ή από την στερεοσκοπική κεφαλή, ένα δισδιάστατο διάνυσμα από το εγγύς πεδίο προς το κέντρο βάρους των texture-hits που αυτό περικλείει (texture attractor behavior), και ένα μονοδιάστατο διάνυσμα το οποίο πολώνει το εγγύς πεδίο προς τα πάνω.

Οι Vieville et al [71] ανέπτυξαν μια παρόμοια μέθοδο με αυτήν των [20]. Η ακριβής τοποθέτηση των pan, tilt, vergence εξασφαλίζεται από μία κινηματική αλυσίδα μερικών συνδέσμων (joints). Αξίζει να σημειωθεί ότι το κεφάλι ρομποτικής στο INRIA αποτελείται από δύο κάμερες απ' τις οποίες, η μία είναι σταθερή με υψηλή ανάλυση και ελεγξιμότητα (zoom, focus kai iris) και είναι το κύριο μάτι (παίζει το ρόλο foveal αισθητήρα), ενώ η άλλη κάμερα είναι ικανή μόνο για σύγκλιση και έχει σταθερό φακό (παίζει το ρόλο του περιφερειακού αισθητήρα). Το κεφάλι στο INRIA είναι ικανό να παρακολουθεί αντικείμενα σε συχνότητα περίπου 1Hz, η οποία περιορίζει σημαντικά την μέγιστη επιτρεπτή ταχύτητα του αντικειμένου στην οποία μπορεί να αποκριθεί το σύστημα.

Μονοφθαλμικά συστήματα

Οι Αλοίμονος και Tsakiris [6] παρουσιάζουν μια μαθηματική θεωρία για την οπτική παρακολούθηση ενός τρισδιάστατου στόχου γνωστού τοπικού σχήματος . Υποθέτουν ότι το αντικείμενο κινείται συμπαγώς, αρχικά είναι κεντραρισμένο και είναι συνέχεια ορατό κατά τη διάρκεια της παρακολούθησης. Από την κινηματική του συστήματος και τον φορμαλισμό της οπτικής ροής παράγεται ένα γενικό κριτήριο, το οποίο καλείται περιορισμός παρακολούθησης *(tracking constraint)*. Απ' αυτόν τον περιορισμό υπολογίζεται η βέλτιστη (ως προς την ελαχιστοποίηση της οπτικής ροής στο κέντρο της εικόνας) ανακατευθυνση της κάμερας. Δηλ. το πρόβλημα παρακολούθησης διατυπώνεται ως προβλημα βελτιστοποίησης.

Οπτική ροή χρησιμοποιεί και ο Tolg [68]. Κατασκευάζει ένα δυναμικό μοντέλο σχήματος του αντικειμένου που παρακολουθείται. Υποθέτει ότι η κίνηση είναι fronto-parallel, και η παρακολούθηση γίνεται μόνο στην οριζόντια κατεύθυνση. Ο Tolg δεν παρουσιάζει την απόδοση του συστήματος παρουσία άλλων κινούμενων αντικειμένων. Θετικός ωστόσο είναι ο τμηματικός και ιεραρχικός (modular and hierarchical) σχεδιασμός του συστήματος ο οποίος κάνει το σύστημα κατάλληλο να χρησιμοποιηθεί ως βάση για ένα περισσότερο εξελιγμένο σύστημα. Νέα στοιχεία είναι ο υπολογισμός ιδίας κίνησης

από την εικόνα και catch-up μηχανισμοί εμπλέκονται για να δώσουν γενικότητα και εύρωστη συμπεριφορά.

Οι Reddi και Loizou [58] προσδιορίζουν τις αλλαγές της σχετικής περιστροφικής κίνησης της κάμερας ως προς ένα ρομπότ, οι οποίες επιτρέπουν την κάμερα να παρακολουθεί ένα αντικείμενο στόχο χωρίς ν' αλλάζει η σχετική κίνηση του ρομπότ ως προς το περιβάλλον. Η παρακολούθηση δεν είναι πάντα δυνατή όταν ο στόχος βρίσκεται κοντά στο επίπεδο που ορίζεται από τους δύο άξονες περιστροφής.

Οι Bartholomeus et al [11] περιγράφουν ένα σύστημα όρασης το οποίο είναι ικανό να ανιχνεύει ένα κινούμενο στόχο και είναι εύρωστο όσον αφορά τις διακυμάνσεις στις παραμέτρους κίνησεις του στόχου και του υποστρώματος και διακυμάνσεις στην υφή της εικόνας. Η μέθοδος βασίζεται σε μια αναπαράσταση δεδομένων πολλαπλών αναλύσεων και εξάγει ψήφους για ένα περιορισμένο σύνολο ταχυτήτων στην εικόνα. Η τμηματοποίηση γίνεται με βάση αυτές τις ψήφους. Το σύστημα όρασης χρησιμοιείται σε συνδυασμό με έναν αλγόριθμο self-learning (αυτοεκπαίδευσης) ο οποίος αντιστοιχεί τα δεδομένα της εικόνας σε δεδομένα συνδέσμων για την οπτική πλοήγηση ενός robot manipulator.

Θέματα ελέγχου και προσανατολισμού

Αρκετοί έχουν δώσει έμφαση σε θέματα ελέγχου της σταθεροποίησης. Ο Christopher Brown [13] μελετά τον έλεγχο πολλών συμπεριφορών σε μια προσομοίωση ελέγχου προσήλωσης χρησιμοποιώντας σακκαδικές κινήσεις, σύγκλιση, καταδίωξη, θαλαμιαίο-οφθαλμικό αντανακλαστικό και έλεγχο κεφαλιού θεωρώντας γνωστό το λάθος ανάδρασης. Εισάγει κινηματική πρόβλεψη για το σύστημα και δυναμική για τα αντικείμενα του κόσμου για να αντιμετωπίσει το πρόβλημα των καθυστερήσεων, τον κορεσμό και τις αλληλεπιδράσεις των συμπεριφορών. Οι Clark και Ferrier [18] υλοποίησαν ένα μοντέλο το οποίο βασίζεται στο οφθαλμοκινητικό σύστημα των θηλαστικών. Απαντούν στο ερώτημα "τι κίνηση πρέπει να κάνει η κάμερα μια δεδομένη χρονική στιγμή" με τη χρήση μιας συνάρτησης σπουδαιότητας που ορίζεται στο χώρο διατάξεων της κάμερας.

Το πρόβλημα του προσανατολισμού συνίσταται στον προσδιορισμό μιας χωρικής σχέσης ανάμεσα σ' ένα σύστημα καμερών και ένα στατικό η κινούμενο αντικείμενο που παρακολουθείται χρονικά, και στην διατήρηση της. Οι στρατηγικές προσανατολισμού μπορούν να στηρίζονται σε ποιοτικά μέτρα [29].

2.4.2.2 Συστήματα μετατόπισης της προσήλωσης και μοντέλα εστίασης της οπτικής προσοχής

Ο Wixson [73] περιγράφει μια μέθοδο επιλογής για την εύρεση ενός αντικειμένου σε μια σκηνή. Η μέθοδος επιτρέπει πολλαπλές όψεις της σκηνής για να απλοποιήσει και να δώσει έμφαση στην οπτική αναζήτηση. Επιτρέπει επικαλύψεις και ελαχιστοποιεί τον αριθμό των κινήσεων προσήλωσης Οι παραχώρηση των πόρων γίνεται δυναμικά.

Οι Bradshaw et al [12] περιγράφουν ένα σύστημα περιπολίας *(surveillance)* ενεργού όρασης πραγματικού χρόνου το οποίο ανιχνεύει κινούμενα αντικείμενα σ' ένα καθημερινό περιβάλλον, κατευθύνει την προσήλωση μιας πλατφόρμας head/eye στα αντικείμενα και έπειτα τα ακολουθεί ομαλά. Η ανίχνευση του στόχου και η παρακολούθηση βασίζεται απλά στην κίνηση της εικόνας, και μπορεί να συνεχίζει για εκτεταμένες περιόδους. Δύο ανεξάρτητες παράλληλες διεργασίες παράγουν: α) κίνηση χαμηλής ανάλυσης για όλη την εικόνα για να κατευθύνουν τις σακκαδικές μετατοπίσεις σε εστίες προσοχής σ' ένα ευρύ πεδίο θέας, β) κίνηση υψηλής ανάλυσης σε μία μικρή περιοχή της εικόνας για την επίτευξη της ομαλής παρακολούθησης. Ένας ελεγκτής προσήλωσης επιλέγει τα αποτελέσματα αυτών των δύο οπτικών διεργασιών και ελέγχει την κίνηση του κεφαλιού. Ο ελεγκτής έχει υλοποιηθεί ως μηχανή πεπερασμένων καταστάσεων.

Ο Westelius στην διδακτορική διατριβή του [72] προτείνει μια θεωρία εστίασης της προσοχής (*focus of attention*). Η θεωρία εφαρμόζεται σε ένα preattentive σύστημα ελέγχου προσήλωσης τριων στρατηγικών ελέγχου που συνεργάζονται : α) εντοπισμός αντικειμένων (object finder) κατά τον οποίο χρησιμοποιούνται κυκλικές συμμετρίες ως ενδείξεις για πιθανά αντικείμενα, β) παρακολούθηση ακμών (edge tracker κατά την οποία ακολουθούνται δομές στη σκηνή ,και γ) σύγκλιση. Ο συγχρονισμός γίνεται με δυναμικά πεδία στο χώρο παραμέτρων του ρομπότ.

Ο Maki [41] παρουσιάζει ένα υπολογιστικό μοντέλο οπτικής προσοχής που αποτελείται απο δύο στάδια: ένα αρχικό παράλληλο (early parallel) και ένα επόμενο σειριακό (later serial). Στο δεύτερο στάδιο όπου συγκεντρώνεται η πληροφορία χρησιμοποιείται ένα κριτήριο που στηρίζεται στο βάθος έτσι ώστε η προσοχή να διατηρείται στο κοντινότερο κινούμενο αντικείμενο.

2.4.2.3 Συστήματα ελέγχου προσήλωσης με εφαρμογές στην όραση

Οι Fermuller και Αλοίμονος [26] παρουσιάζουν κάποια υπολογιστικά πλεονεκτήματα της σταθεροποίησης της προσήλωσης. Ο τρόπος με τον οποίο μεταβάλλεται το πεδίο οπτικής ροής δίνει λύσεις στα εξής προβλήματα : εκτίμηση τρισδιάστατης κίνησης, recovery της ιδίας κίνησης, και εκτίμηση του time-to-contact.

Οι Brunstrum, Eklundh και Uhlin [15] ασχολούνται με το πρόβλημα ενεργού επιλογής σημείων καθήλωσης στο πλαίσιο της αναγνώρισης τεχνητών αντικειμένων που χαρακτηρίζονται από το σχήμα τους. Η προσήλωση οδηγείται από μια στρατηγική ομαδοποίησης συνδέσμων (junctions) που διαχωρίζονται από ασυνέχειες βάθους. Επίσης ανέπτυξαν μία μεθοδολογία για γρήγορη ενεργή ανίχνευση και κατηγοριοποίηση των συνδέσμων με την επιλογή σημείων καθήλωσης.

Οι Ahuja και Abbott [4] υλοποιούν ένα ενεργό στερεοσκοπικό σύστημα το οποίο έχει πολλά από τα σημεία που υπογραμμίζονται στο [65] ως τμήματα ενός ενεργού συστήματος όρασης - σύγκλιση εστίαση, έλεγχος διαφράγματος και calibration - και

χρησιμοποιεί την ανομοιότητα σαν εργαλείο για την ανακατασκευή τρισδιάστατης επιφάνειας. Ο μηχανισμός αυτός απαιτεί μεταβολή της διάταξης των καμερών έτσι ώστε να είναι ορατά όλα τα τμήματα της σκηνής. Έτσι, συνδέεται η στερεοσκοπική επεξεργασία με τον έλεγχο της εστίασης και της σύγκλισης. Γίνεται αυτόνομη επιλογή νέων σημείων καθήλωσης με βάση τη συνολική επεξεργασία της σκηνής. Η προσέγγιση είναι τελεολογική *(goal-dependent).* και χρησιμοποιεί ένα σύνολο από κανόνες για προσήλωση σ' ένα σημείο με σκοπό την ομαλή επέκταση της περιοχής. Γίνεται πρόβλεψη των επικαλύψεων και ελαχιστοποίηση πολυπλοκότητας. Ένα συμπέρασμα που προκύπτει από τη λειτουργία του συστήματος αυτού είναι ότι η ανομοιότητα που υπολογίζεται χρησιμοποιώντας μόνο την σύγκλιση ή μόνο την εστίαση είναι μη κατάλληλη πληροφορία για μεγάλες σκηνές. Είναι προτιμότερο οι μηχανισμοί εστίασης και σύγκλισης να συνδυαστούν. Ο τρόπος με τον οποίο συνεργάζονται η σύγκλιση και η εστίαση καθορίζεται ενεργά με βάση την περιγραφή της σκηνής. Το αποτέλεσμα είναι ένα σύστημα χαμηλού επιπέδου όρασης το οποίο παράγει μια ασαφή περιγραφή της σκηνής με εκτιμήσεις βάθους (χάρτες επιφάνειας) οι οποίες μπορούν να βοηθήσουν την πλοήγηση ή άλλες διεργασίες ενεργού όρασης.

Οι Krotkov και Bajcsy [36] συνδυάζουν τον έλεγχο της εστίασης, stereo και σύγκλισης για να εξάγουν σχετική πληροφορία απο μια σκηνή. Η δουλειά τους είναι χαλαρά βασισμένη σε βιολογικά παραδείγματα, αλλά γίνεται σειριακή αντιστοίχηση παρά ακολουθούνται τα περισσότερο παράλληλα μοντέλα που προτείνει το βιολογικό σύστημα όρασης. Ο στόχος του συστήματος αυτού είναι η δυνατότητα αυτόνομης διάταξης (ranging).

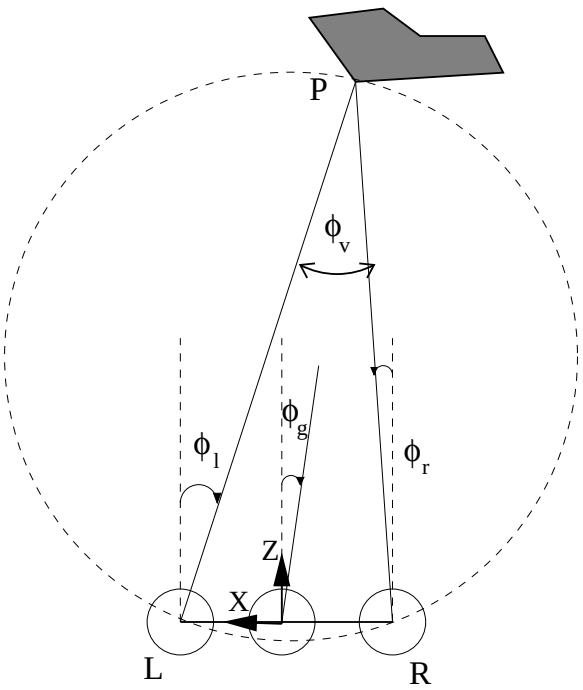
Κεφάλαιο 3

Σύστημα σταθεροποίησης της προσήλωσης

Σ' αυτό το κεφάλαιο θα περιγράψουμε το σύστημα σταθεροποίησης της οπτικής προσήλωσης (ΣΣΠ) που έχει υλοποιηθεί στα πλαίσια της παρούσας εργασίας. Αφού περιγράψουμε την γεωμετρία της διοφθαλμικής προσήλωσης θα δούμε πως ορίζεται το πρόβλημα της σταθεροποίησης και στη συνέχεια θα μιλήσουμε για την αρχή λειτουργίας του παρόντος συστήματος και τις επιλογές που έγιναν κατά τη σχεδίαση του. Όπως αναφέραμε και στο προηγούμενο κεφάλαιο, η σταθεροποίηση της προσήλωσης είναι ένα πρόβλημα ελέγχου αρνητικής ανάδρασης λάθους. Το λάθος που είναι κωδικοποιημένο στο οπτικό σήμα υπολογίζεται με οπτική επεξεργασία χαμηλού επιπέδου (δηλ. δεν φτάνει στο επίπεδο της αναγνώρισης των αντικειμένων που βρίσκονται στην σκηνή). Το ΣΣΠ έχει μελετηθεί σε μια προσομοίωση μιας μηχανικής πλατφόρμας που διαθέτει τέσσερις βαθμούς ελευθερίας. Αποτελέσματα της εφαρμογής του συστήματος στη συγκεκριμένη προσομοίωση θα παρουσιαστούν στο κεφάλαιο 6.

3.1 Γεωμετρία διοφθαλμικής προσήλωσης

Μια αναπαράσταση της γεωμετρίας διοφθαλμικής προσήλωσης φαίνεται στο σχήμα 3.1. Αναπαριστούμε τον οπτικό χώρο ως προς ένα εικονικό/υποθετικό κυκλωπικό μάτι το οποίο βρίσκεται στο μέσο της ευθείας που ενώνει τα δύο οπτικά κέντρα και ο οπτικός του άξονας τέμνει τον αριστερό και δεξί οπτικό άξονα. Οι X, Z άξονες καθώς και τα κέντρα των ματιών βρίσκονται στο ίδιο επίπεδο, το επίπεδο καθήλωσης *(fixation plane)*. Το σημείο P ονομάζεται σημείο καθήλωσης *(fixation point)*. Έστω ϕ_l και ϕ_r οι γωνίες που σχηματίζει η κύρια (ευθεία μπροστά) κατεύθυνση με τον αριστερό και τον δεξιό



Σχήμα 3.1: Γεωμετρία διοφθαλμικής προσήλωσης

άξονα αντίστοιχα. Γωνία σύγκλισης *(vergence angle)*, ϕ_v, ονομάζουμε την γωνία που σχηματίζουν οι οπτικοί άξονες και γωνία προσήλωσης *(gaze angle ή version angle)*, ϕ_g, την γωνία του οπτικού άξονα του κυκλωπικού ματιού με την κύρια κατεύθυνση. Ισχύει:

$$\phi_v = \phi_l - \phi_r \tag{3.1}$$

$$\tan\phi_g = \frac{1}{2}(\tan\phi_l + \tan\phi_r) \tag{3.2}$$

Επειδή, οι παραπάνω σχέσεις είναι διανυσματικές ο προσανατολισμός των γωνιών λαμβάνεται υπ' όψην. Για μικρές γωνίες προσήλωσης μπορούμε χρησιμοποιώντας την προσέγγιση $\phi \approx \tan\phi$ να ξαναγράψουμε την εξίσωση 3.2 ως εξής:

$$\phi_g = \frac{1}{2}(\phi_l + \phi_r) \tag{3.3}$$

Η στερεοσκοπική είσοδος είναι το πρότυπο (pattern) ζευγών σημείων ταυτόχρονου ερεθίσματος στην αριστερή και δεξιά κάμερα. Ορίζουμε ως στερεοσκοπική ανομοιότητα [1] *(stereo disparity)* ενός σημείου της σκηνής την διαφορά των θέσεων των προβολών του σημείου στην αριστερή και δεξιά εικόνα. Για καρτεσιανό σύστημα συντεταγμένων :

$$\vec{d} = (x_l - x_r, y_l - y_r) \tag{3.4}$$

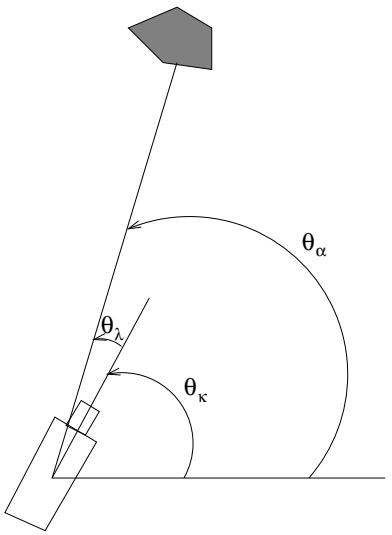
Το σημείο καθήλωσης έχει εξ' ορισμού ανομοιότητα μηδέν. Ο γεωμετρικός τόπος των σημείων με μηδενική ανομοιότητα ονομάζεται ορόπτερο *(horopter)*.

3.2 Το πρόβλημα

Σ' αυτή την ενότητα, θα περιγράψουμε το πρόβλημα της σταθεροποίησης, κατ' αρχήν, στη μία διάσταση για ένα μονοφθαλμικό σύστημα, και στη συνέχεια, στις τρεις διαστάσεις για ένα διοφθαλμικό σύστημα.

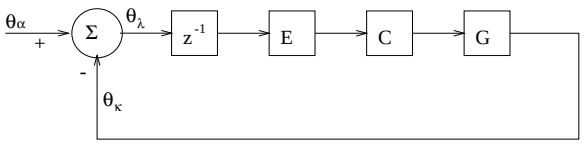
Έστω ότι έχουμε μια κάμερα κι ένα αντικείμενο να κινείται μπροστά απ' αυτήν όπως φαίνεται στο σχήμα 3.2. Ο σκοπός ενός συστήματος παρακολούθησης σε μιά διάσταση είναι να κατευθύνει την κάμερα έτσι ώστε να ακολουθεί την γωνία θ_α. Αν ονομάσουμε θ_κ την γωνία προσήλωσης, η διαφορά $\theta_\lambda = \theta_\alpha - \theta_\kappa$ είναι το λάθος στην απόδοση της παρακολούθησης. Είναι προφανές ότι το σύστημα είναι κλειστού βρόγχου με αρνητική ανάδραση, όπου το λάθος είναι κωδικοποιημένο στο οπτικό σήμα. Οποιαδήποτε λύση και να δοθεί στο πρόβλημα, θα έχει συνεπώς την μορφή του σχήματος 3.3. Το οπτικό σύστημα αποτελείται από την καθυστέρηση z^{-1} και τον εκτιμητή λάθους E. Ο ελεγκτής C , μετατρέπει το λάθος σ' ένα σύνολο από σήματα ελέγχου τα οποία τροφοδοτούν τον κινητήρα της κάμερας (ελεγχόμενο σύστημα) για να προκαλέσουν κατάλληλες αλλαγές στη γωνία θ_κ.

[1] Υπάρχει και ο όρος ανομοιότητα λόγω κίνησης (*motion disparity*). Από δω και στο εξής, θα αναφερόμαστε με τον όρο ανομοιότητα στη στερεοσκοπική ανομοιότητα για λόγους συντομίας.



Σχήμα 3.2: Παρακολούθηση σε μια διάσταση

Στόχος της παρακολούθησης είναι η κάμερα να κοιτάει πρός το αντικείμενο



Σχήμα 3.3: Η παρακολούθηση ως σύστημα ανάδρασης λάθους

Το οπτικό σύστημα αντιστοιχεί στην καθυστέρηση z^{-1} και τον εκτιμητή λάθους Ε. Ο ελεγκτής C , μετατρέπει το λάθος σ' ένα σύνολο από σήματα ελέγχου τα οποία τροφοδοτούν τον κινητήρα της κάμερας G (ελεγχόμενο σύστημα).

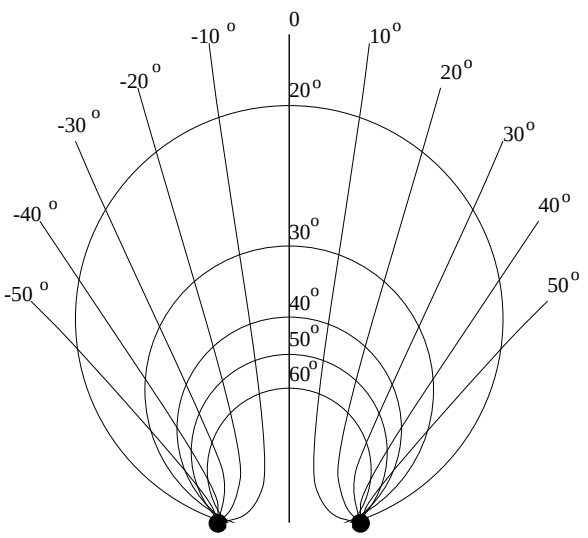
Το λάθος στην κατεύθυνση προσήλωσης είναι κωδικοποιημένο στο οπτικό σήμα, με τη μορφή θέσης της εικόνας του στόχου πάνω στην επιφάνεια του αισθητήρα. Επομένως, η κίνηση της κάμερας με σκοπό τη διόρθωση του λάθους θέσης θα προκαλέσει ολίσθηση και πιθανόν θάμπωμα της εικόνας του στόχου. Το σύστημα ελέγχου, δηλαδή, έχει να αντιμετωπίσει ταυτόχρονα το λάθος θέσης και το λάθος ταχύτητας. Μια απλή αντιμετώπιση στο δίλημμα αυτό είναι να θυσιάσει το ένα απ' τα δύο σε κάποιο βαθμό. Στο οφθαλμοκινητικό σύστημα των πρωτεύοντων θηλαστικών, το σύστημα παρακολούθησης ακολουθεί την ταχύτητα του στόχου με ομαλές κινήσεις των ματιών, και χρησιμοποιεί catch-up σακκαδικές κινήσεις για να διορθώσει το λάθος θέσης. Αντίθετα, στους πιθήκους και τα δύο είδη λαθών (θέσης/ταχύτητας) αντιμετωπίζονται με ομαλές κινήσεις.

Η απόδοση ενός συστήματος παρακολούθησης επηρεάζεται από τις καθυστερήσεις που οφείλονται στην οπτική επεξεργασία. Οι καθυστερήσεις αυτές έχουν σαν αποτέλεσμα το σύστημα να καθοδηγείται από το προηγούμενο λάθος. Το γεγονός αυτό κάνει το σύστημα να μην συμπεριφέρεται σωστά (overeact). Ένας απλός τρόπος αντιμετώπισης του προβλήματος αυτού είναι να μειωθεί η απόκριση του ελέγχου στα λάθη. Ωστόσο, ο έλεγχος θα ανταποκρίνεται στα λάθη καθυστερημένα, αν δεν υπάρχει κάποια πρόβλεψη της κίνησης του στόχου.

Το πρόβλημα της παρακολούθησης στις τρεις διαστάσεις μοντελοποιείται, επίσης, με έλεγχο αρνητικής ανάδρασης. Όμως, είναι πολλών εισόδων και πολλών εξόδων. Αυτό σημαίνει ότι μπορεί να υπάρχουν διάφορα ανεξάρτητα ή φωλιασμένα επίπεδα ελέγχου τα οποία συνεργάζονται έτσι ώστε να κατευθύνουν την προσήλωση προς το στόχο της παρακολούθησης και διάφοροι τρόποι κατανομής των λαθών στους βαθμούς ελευθερίας του συστήματος.

Στα πρωτεύοντα θηλαστικά, για παράδειγμα, η παρακολούθηση εκδηλώνεται με δύο ξεχωριστές κλάσεις κινήσεων: την σύγκλιση (vergence) και τη στροφή (version). Στους ανθρώπους οι δύο αυτές κινήσεις ολοκληρώνονται μ' ένα σύνθετο τρόπο και όχι πλήρως ανεξάρτητο. Ένα απλουστευμένο γεωμετρικό μοντέλο τους φαίνεται στο σχήμα 3.4([16] after [39]). Σ' αυτό το σχήμα οι γραμμές ίσης στροφής (δηλ. ο γεωμετρικός τόπος των σημείων συμμετρικής ανομοιότητας) σχηματίζουν μια σειρά από τετράγωνες υπερβολές οι οποίες περνάνε από το κέντρο θέας των δύο ματιών και το κοινό τους κέντρο είναι το μέσο της βασικής γραμμής. Οι γραμμές ίσης σύγκλισης (δηλ. ο γεωμετρικός τόπος των σημείων ίσης ανομοιότητας) αναπαρίστανται από τους κύκλους Vieth-Muller οι οποίοι περνάνε και από τα δύο μάτια. Στην πραγματικότητα, οι γραμμές ίσης σύγκλισης είναι κύκλοι διαφορετικής ακτίνας μόνο για ιδανικές κυκλικές φωτοευαίσθητες επιφάνειες. Στην περίπτωση επίπεδων επιφανειών, οι γραμμές αυτές είναι κωνικές τομές και ειδικότερα το ορόπτερο είναι κύκλος ενώ οι γραμμές ίσης σύγκλισης για μικρές ανομοιότητες είναι ελλείψεις. Η σύγκλιση και η στροφή σχετίζονται με δύο είδη ανομοιότητας. Η καθαρή στροφή σχετίζεται με μηδενική ανομοιότητα και η καθαρή σύγκλιση με συμμετρική ανομοιότητα. Αυτές οι ποιότητες είναι ειδικά ενδιαφέρουσες.

Η σταθεροποίηση της προσήλωσης με ένα στερεοσκοπικό σύστημα καμερών, κατά



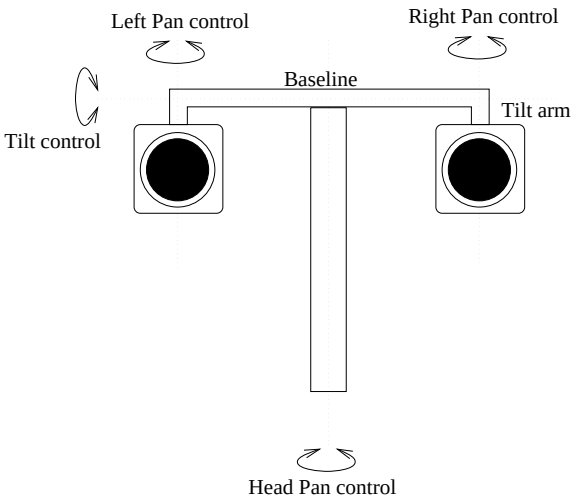
Σχήμα 3.4: Ένα απλουστευμένο γεωμετρικό μοντέλο των κινήσεων των ματιών στον άνθρωπο.

Οι γραμμές ίσης σύγκλισης, δηλ. ο γεωμετρικός τόπος των σημείων ίσης ανομοιότητας (οι κύκλοι που περνάνε από τα μάτια) και οι γραμμές ίσης μέσης στροφής, δηλ. ο γεωμετρικός τόπος των σημείων συμμετρικής ανομοιότητας (οι υπερβολές που περνάνε από τα μάτια).

αντιστοιχία με το ανθρώπινο σύστημα όρασης, απαιτεί μηχανισμούς για στροφή έτσι ώστε να επιτυγχάνεται η παρακολούθηση πάνω στο ορόπτερο και για σύγκλιση έτσι ώστε να επιτυγχάνεται η παρακολούθηση στην κατεύθυνση του βάθους. Στη συνέχεια θα δούμε πως μπορούμε να φτιάξουμε τέτοιους μηχανισμούς για μία τυπική κεφαλή ενεργού όρασης, η γεωμετρία της οποίας φαίνεται στο σχήμα 3.5. Η κεφαλή αυτή διαθέτει τους εξής βαθμούς ελευθερίας: ανεξάρτητη κίνηση κάθε κάμερας δεξιά-αριστερά (left pan και right pan), κίνηση του 'λαιμού' δεξιά-αριστερά (head pan) και κίνηση των δύο καμερών μαζί πάνω-κάτω (tilt).

3.3 Αρχιτεκτονική του συστήματος

Το ΣΣΠ που υλοποιήσαμε ακολουθεί την προσέγγιση του David J. Coombs [22]. Σύμφωνα μ' αυτή την προσέγγιση το ΣΣΠ αποτελείται από δύο υποσυστήματα: το υποσύστημα σύγκλισης, και το υποσύστημα καταδίωξης (pursuit). Το υποσύστημα σύγκλισης υπολογίζει το λάθος στη γωνία σύγκλισης και δίνει σήματα ελέγχου στις ανεξάρτητες pan αρθρώσεις των καμερών έτσι ώστε να το μηδενίσει, με σκοπό την παρακολούθηση στην κατεύθηνση του βάθους. Το υποσύστημα καταδίωξης υπολογίζει το λάθος θέσης της προβολής του αντικειμένου και ελέγχει τόσο τις ανεξάρτητες αρθρώσεις pan των καμερών, όσο και την κοινή tilt άρθρωση τους, με σκοπό την παρακολούθηση του



Σχήμα 3.5: Η γεωμετρία μιάς τυπικής κεφαλής ενεργού όρασης

Διαθέτει τέσσερις βαθμούς ελευθερίας : ανεξάρτητη κίνηση κάθε κάμερας δεξιά-αριστερά (left pan control και right pan control), κίνηση του 'λαιμού' δεξιά-αριστερά (head pan control) και κίνηση των δύο καμερών μαζί πάνω-κάτω (tilt control).

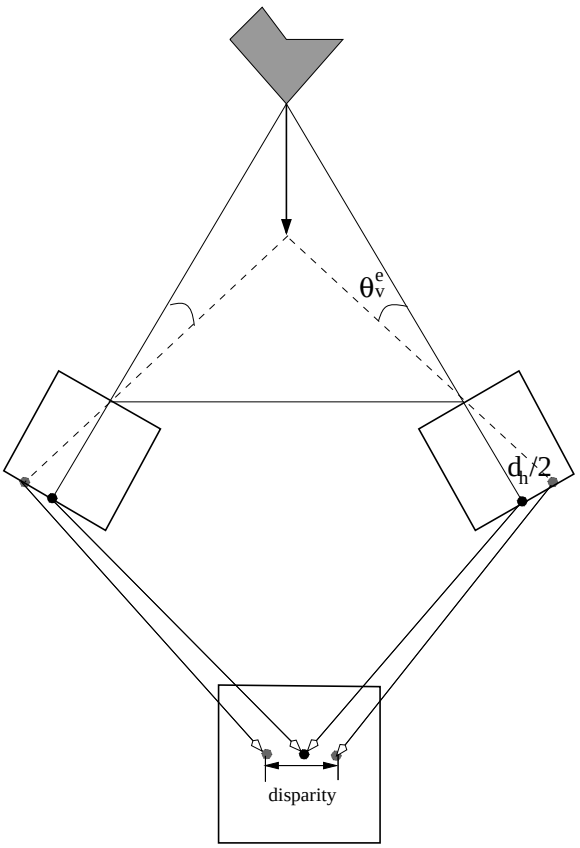
στόχου πάνω στην οροπτερική επιφάνεια. Τα δύο υποσυστήματα συνεργάζονται. Η σύγκλιση φροντίζει ο στόχος να έχει μηδενική ανομοιότητα. Ανιχνεύοντας τα σημεία μηδενικής ανομοιότητας το σύστημα καταδίωξης ξεχωρίζει το στόχο απ' το υπόστρωμα και υπολογίζει το λάθος θέσης. Η καταδίωξη φροντίζει ο στόχος να είναι στο κέντρο της εικόνας έτσι ώστε να είναι εφικτή η τοπική επεξεργασία για τον υπολογισμό της ανομοιότητας που γίνεται για την διατήρηση της σύγκλισης.

Η σύγκλιση στην γενικότερη της μορφή μπορεί να χρησιμοποιεί διάφορα είδη πληροφορίας. Στην εργασία αυτή, μας απασχολεί η εξαγωγή του λάθους στη γωνία σύγκλισης χρησιμοποιώντας μόνο οπτική πληροφορία. Για τον σκοπό αυτό, θα ήταν χρήσιμη οποιαδήποτε πληροφορία του βάθους, αν ήταν γνωστή η γωνία σύγκλισης αφού η γωνία σύγκλισης μπορεί να εκφραστεί συναρτήσει της απόστασης του στόχου από τον παρατηρητή. Ωστόσο, είναι δυνατός ο υπολογισμός του λάθους στην γωνία σύγκλισης γνωρίζοντας την οριζόντια ανομοιότητα και δεν απαιτεί την γνώση της γωνίας σύγκλισης. Όπως φαίνεται στο σχήμα 3.6 η αντιστοιχία οριζόντιας ανομοιότητας (d_h) και λάθους σύγκλισης (θ_v^e)είναι:

$$\theta_v^e = 2 \arctan \frac{d_h}{2f} \tag{3.5}$$

όπου f η εστιακή απόσταση της κάμερας. Όπως φαίνεται από την παραπάνω σχέση δεν χρειάζεται η γνώση της γωνίας σύγκλισης. Γενικά, η ανομοιότητα υπολογίζεται ευκολότερα και γρηγορότερα απ' ότι το βάθος. Για την διαδικασία της σύγκλισης είναι απαραίτητος μόνο ο υπολογισμός της ανομοιότητας του στόχου.

Η καταδίωξη διαχωρίζει το στόχο απ' το υπόστρωμα χρησιμοποιώντας την



Σχήμα 3.6: Αντιστοιχία συμμετρικής ανομοιότητας και γωνίας σύγκλισης

ανομοιότητα του. Ο στόχος ορίζεται από τα σημεία μηδενικής ανομοιότητας. Αν (x, y) είναι το κέντρο βάρους αυτών των σημείων τότε το λάθος στη γωνία προσήλωσης είναι :

$$\theta^e_{g_x} = \arctan\frac{x}{f} \tag{3.6}$$

στο επίπεδο προσήλωσης και

$$\theta^e_{g_y} = \arctan\frac{y}{f} \tag{3.7}$$

στην καθετη στο επίπεδο προσήλωσης κατεύθυνση.

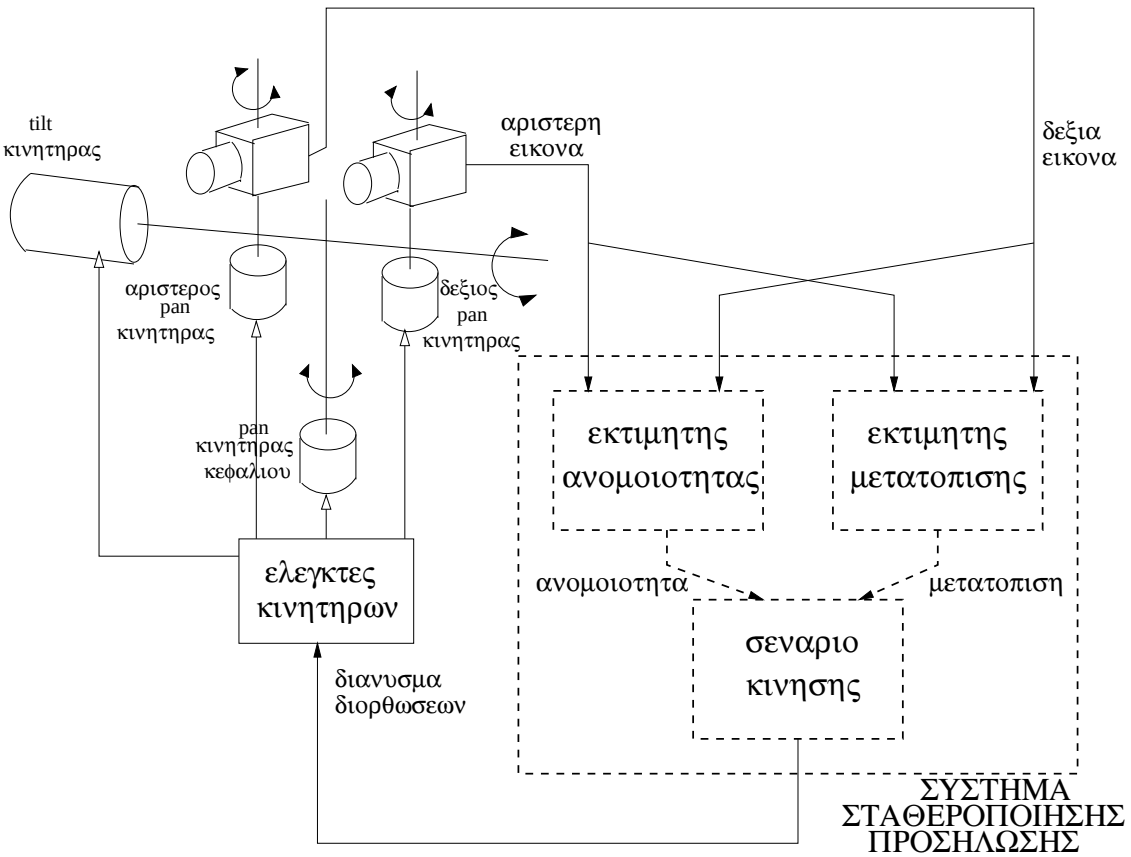
Περισσότερο σύνθετα διοφθαλμικά συστήματα σταθεροποίησης της προσήλωσης, απ' το ΣΣΠ που περιγράφουμε εδώ, έχουν διάφορες άλλες παραμέτρους να ρυθμίσουν, (όπως είναι το βάθος πεδίου) και αλληλεπιδράσεις μεταξύ τους. Ένας άλλος πιθανός βαθμός ελευθερίας είναι η στρέψη (*torsion*), δηλ. η περιστροφή γύρω από τον οπτικό άξονα. Ο έλεγχος της στρέψης μπορεί να χρησιμοποιηθεί για την ευθυγράμμιση των οριζόντιων γραμμών του πίνακα φωτοευαίσθητων στοιχείων με τις επιπολικές γραμμές, πράγμα που υποθέτουν οι περισσότεροι στερεοσκοπικοί αλγόριθμοι. Οι παράμετροι προσήλωσης $\theta^e_v, \theta^e_{g_x}, \theta^e_{g_y}$ του παρόντος συστήματος μπορούν να μετατραπούν σε διορθώσεις $\phi_l, \phi_r, \phi_p, \phi_t$ στους τέσσερις βαθμούς ελευθερίας που διαθέτει το οπτικό σύστημα με την εφαρμογή κάποιου σεναρίου κίνησης.

Το σύστημα που έχουμε υλοποιήσει φαίνεται στο σχήμα 3.7.Αποτελείται από έναν εκτιμητή ανομοιότητας και έναν εκτιμητή μετατόπισης. Η ανομοιότητα και η μετατόπιση μετατρέπονται σύμφωνα με κάποιο σενάριο κίνησης σε διάνυσμα λαθών στις γωνίες προσήλωσης.

Για την επιλογή της μεθόδου υπολογισμού ανομοιότητας θα μιλήσουμε στην ενότητα 3.4 και θα περιγράψουμε τις μεθόδους που επιλέξαμε εκτενώς στο κεφάλαιο 4. Αντίστοιχα για την μέθοδο υπολογισμού του λάθους θέσης θα μιλήσουμε στην παράγραφο 3.5 και στο κεφάλαιο 5. Επίσης, Θα εξετάσουμε διάφορα σενάρια κατανομής των παραμέτρων προσήλωσης στους βαθμούς ελευθερίας (ενότητα 3.6.3). Τα σενάρια αυτά μπορούν να ενεργοποιηθούν από διαδικασίες υψηλότερου επιπέδου.

3.4 Υπολογισμός ανομοιότητας

Σ' αυτήν την ενότητα θα παρουσιάσουμε διάφορες μεθόδους για τον υπολογισμό της ανομοιότητας που όπως αναφέραμε χρησιμοποιείται για την εκτίμηση του λάθους στην κατεύθυνση του βάθους. Θα καταλήξουμε στην προσέγγιση που ακολουθήσαμε στο ΣΣΠ που υλοποιήσαμε. Οι περισσότερες προσεγγίσεις στο πρόβλημα υπολογισμού ανομοιότητας εντάσσονται σε δύο υπολογιστικά μοντέλα : την αντιστοίχηση *(correspondence)* και την συσχέτιση *(correlation)* (για μια αναλυτική επισκόπηση των τεχνικών αυτών βλ. [24]).



Σχήμα 3.7: Αρχιτεκτονική συστήματος σταθεροποίησης της προσήλωσης

Το σύστημα σταθεροποίησης της προσήλωσης χρησιμοποιεί μόνο οπτική πληροφορία και υπολογίζει ένα διάνυσμα λαθών στις γωνίες προσήλωσης. Αποτελείται από έναν εκτιμητή ανομοιότητας και έναν εκτιμητή μετατόπισης. Η ανομοιότητα και η μετατόπιση μετατρέπονται σύμφωνα με κάποιο σενάριο κίνησης σε διάνυσμα λαθών στις γωνίες προσήλωσης. Η επιλογή του σεναρίου κίνησης γίνεται από διεργασίες υψηλότερου επιπέδου.

Οι τεχνικές ανίχνευσης ανομοιότητας που βασίζονται στη συσχέτιση κάνουν ουσιαστικά την υπόθεση ότι η αριστερή και δεξιά εικόνα είναι απλά μετατοπισμένη έκδοση η μια της άλλης - σύμφωνα μ' αυτήν την υπόθεση η επιφάνεια είναι μετωπική (frontoparallel), το μοντέλο της κάμερας ακολουθεί την ορθογραφική προβολή, και οι συνθήκες φωτισμού είναι πανομοιότυπες ως προς τις δύο κάμερες - και ότι ο υπολογισμός αυτής της μετατόπισης (ή ολίσθησης) είναι ισοδύναμος με τον προσδιορισμό της ανομοιότητας. Επίσης, γίνεται η υπόθεση ότι η σωστή μετατόπιση είναι η μετατόπιση που θέτει το ολοκλήρωμα της τετραγωνικής διαφοράς μεταξύ δύο εικόνων στο μηδέν. Το ολοκλήρωμα αυτό έχει ένα ελάχιστο όταν οι δύο εικόνες είναι μετατοπισμένες κατά το σωστό ποσοστό μεταξύ τους. Με αντίστοιχο τρόπο, η συσχέτιση (ολοκλήρωμα του γινομένου τους) έχει ένα μέγιστο. Το μοντέλο της συσχέτισης γενικεύεται για τον υπολογισμό της σχετικής μετατόπισης δύο αντίστοιχων περιοχών στις εικόνες. Η ανομοιότητα που προσδίδεται στο σημείο x μπορεί να ευρεθεί ως :

$$\max_{disp} \int W(z)L(x + disp + z)R(x + z)dz \tag{3.8}$$

όπου W είναι μια συνάρτηση παραθύρου που χρησιμοποιείται λόγω τοπικότητας της μέτρησης, L και R η αριστερή και δεξιά εικόνα αντίστοιχα. Η επιλογή της συνάρτησης W μπορεί να αυξήσει την απόδοση του αλγορίθμου σε περιπτώσεις θορύβου.

Γενικά η αριστερή και δεξιά όψη μιας σκηνής δεν είναι απλά μια μετατοπισμένη έκδοση η μια της άλλης. Πρακτικά, η συσχέτιση δεν μπορεί να διευθετήσει περιπτώσεις αλλοιώσεων που οφείλονται στην προοπτική προβολή ή δια-οφθαλμικών διαφορών φωτισμού. Εκτός από τα προβλήματα γεωμετρίας και φωτισμού η τεχνική συσχέτισης έχει ενδογενή προβλήματα. Για μεγάλες περιοχές σταθερής έντασης είναι δύσκολος ο προσδιορισμός ενός μοναδικού τοπικού ελάχιστου. Ακόμα και σε περιπτώσεις με μεγάλη διακύμανση έντασης η συσχέτιση αντιμετωπίζει προβλήματα. Ένα άλλο πρόβλημα με τις τεχνικές συσχέτισης πηγάζει από το γεγονός ότι το υπόστρωμα του οπτικού ερεθίσματος δεν έχει ένταση μηδέν, με αποτέλεσμα να προκύπτουν κορυφές στο χώρο της συσχέτισης οι οποίες δεν οφείλονται σε δομή επιφάνειας. Μια αντιμετώπιση για το πρόβλημα αυτό είναι η εφαρμογή ζωνοδιαβατών (bandpass) φίλτρων για την μείωση των περιοχών σταθερής ή αργά μεταβαλλόμενης έντασης. Υπάρχουν πολύ πιο εκλεπτυσμένες τεχνικές συσχέτισης όπως η κανονικοποιημένη ετεροσυσχέτιση. Ωστόσο όλες οι τεχνικές αντιμετωπίζουν δυσκολίες ως προς: α) την προοπτική προβολή, β) τις δια-οφθαλμικές διαφορές φωτισμού, γ) τις παραγώγους (gradients) ανομοιότητας, δ) πολύ υφή στην εικόνα, και ε) λίγη υφή στην εικόνα.

Οι τεχνικές ανίχνευσης ανομοιότητας που βασίζονται στην αντιστοίχιση γενικά αναλύονται σε τρεις φάσεις: α) ανεξάρτητη επεξεργασία των δύο εικόνων που αποσκοπεί στην εξαγωγή μονοφθαλμικών στοιχείων (παρουσία zero-crossing σε μια bandpassed έκδοση της εικόνας (ακμή), β) σύγκριση των μονοφθαλμικών χαρακτηριστικών της μιας εικόνας με τα αντίστοιχα στην άλλη, γ) απαλοιφή των λανθασμένων αντιστοιχήσεων (matches) που έχουν προκύψει απ' τη δεύτερη φάση. Επομένως, στις τεχνικές αυτές

βλέπουμε τα εξής θέματα σχεδίασης: α) επιλογή χαρακτηριστικών, β) επιλογή κριτηριών σύγκρισης, και γ) τρόπος αποφυγής/απαλοιφής εσφαλμένων αντιστοιχήσεων (false targets). Τα χαρακτηριστικά πρέπει: α) να αντιστοιχούν σε ιδιότητες επιφάνειας, β) να παράγουν αρκετά πυκνές αντιστοιχήσεις, γ) να κατανέμονται πάνω σε μια τυπική εικόνα έτσι ώστε οι εσφαλμένες αντιστοιχήσεις να μπορούν να αποφευχθούν/εξαλειφθούν σχετικά εύκολα. Πολλές και από τις τεχνικές αντιστοίχισης επειδή υποθέτουν ότι το ίδιο χαρακτηριστικό επιφάνειας φαίνεται το ίδιο στην αριστερή και δεξιά όψη θεωρούν frontoparallel επιφάνειες και ορθογραφική προβολή στη δειγματοληψία του συστήματος απεικόνισης. Εκλεπτυσμένα μονοφθαλμικά χαρακτηριστικά που εξάγονται από ζωνοδιαβατά κανάλια (φίλτρα) μπορούν να είναι λιγότερο ευαίσθητα όσον αφορά τις μηδενικές τιμές gradients ανομοιότητας ή επιδράσεις δια-οφθαλμικού φωτισμού. Αλλά τα προβλήματα αυτά εξακολουθούν να υπάρχουν και δεν μπορούν να αποφευχθούν τελείως.

Συγκριτικά οι τεχνικές αντιστοίχισης έχουν χαμηλότερη πυκνότητα απόκρισης από τις τεχνικές συσχέτισης αλλά παρέχουν ένα περισσότερο ανθεκτικό σύνολο αποκρίσεων. Οι τεχνικές συσχέτισης και αντιστοίχισης, από μια άποψη είναι συμπληρωματικές.

Μια τρίτη υπολογιστική δομή για τον υπολογισμό τοπικής ανομοιότητας είναι οι τεχνικές τοπικής διαφοράς φάσης [31, 27, 42]. Το γεγονός ότι η φάση μεταβάλλεται με τον ίδιο τρόπο που μεταβάλλεται και η θέση (equivariant) μπορεί να χρησιμοποιηθεί στην εκτίμηση τοπικής μετατόπισης μεταξύ δύο εικόνων. Το πλεονέκτημα στις τεχνικές αυτές είναι ότι η τοπική φάση είναι μια συνεχής μεταβλητή ευαίσθητη σε αλλαγές πολύ μικρότερες από το κβαντισμό στο χώρο. Η ανομοιότητα εκφράζεται ως διαφορά φάσης στην έξοδο τοπικών ζωνοδιαβατών φίλτρων που εφαρμόζονται στην αριστερή και δεξιά εικόνα. Μια φυσική επιλογή των ζωνοδιαβατών φίλτρων είναι τα Gabor φίλτρα [72, 74, 40] αλλά αρκεί οποιοδήποτε quadrature ζεύγος ζωνοδιαβατών, φίλτρων σταθερής φάσης.

Στο ΣΣΠ που αναπτύξαμε χρησιμοποιήσαμε μια μέθοδο συσχέτισης, το cepstral φίλτρο για τον υπολογισμό της ανομοιότητας του στόχου. Το φίλτρο αυτό εισήγαγαν στη στερεοσκοπική όραση οι Yeshurun και Schwartz [75], για τον υπολογισμό αραιού πεδίου ανομοιότητας, αλλά έχει χρησιμοποιηθεί κατα κώρον στην ανάλυση ηχητικών σημάτων [23]. Οι Olson και Coombs [53] προσάρμοσαν το φίλτρο αυτό για τον υπολογισμό της ανομοιότητας του στόχου σύγκλισης, χρησιμοποιώντας το σε κεντρικά τμήματα των εικόνων. Η τελική φάση του αλγορίθμου ουσιαστικά είναι η ανίχνευση της μέγιστης κορυφής στην έξοδο. Αυτό αποτελεί πρόβλημα σε σκηνές με πολλαπλές ανομοιότητες και συνδέεται με την επιλογή του μεγέθους του παράθυρου δειγματοληψίας (sample window size). Τα παράθυρα πρέπει να είναι αρκετά μεγάλα για να είναι δυνατός ο χειρισμός του προσδοκόμενου διαστήματος μέσα στο οποίο παίρνει τιμές η ανομοιότητα. Έτσι όμως σίγουρα θα εμφανίζονται πολλά μέγιστα στην έξοδο τουλάχιστον μερικές φορές κατά τη διάρκεια της λειτουργίας του συστήματος κάτω από πραγματικές συνθήκες. Ένας προφανής τρόπος για την βελτίωση της απόδοσης της μεθόδου είναι η χρήση

κάποιου φίλτρου στην είσοδο έτσι ώστε να τονιστεί το αντικείμενο. Αυτό προϋποθέτει περισσότερη πληροφορία για τη θέση του αντικειμένου, έτσι ώστε να πολλαπλασιαστεί με την είσοδο μια μάσκα που να το τονίζει (αυτό κάνουν στο εργαστήριο υπολογιστικής όρασης και ενεργού αντίληψης (CVAP) στο KTH [41, 43, 44]). Μια πιο απλή λύση, την οποία υιοθετήσαμε, είναι να θεωρηθεί ότι το αντικείμενο ορίζεται από τη θέση της μιας κάμερας. Ακολουθήσαμε την προσέγγιση των Taylor, Olson και Martin [66], η οποία στηρίζεται σε μια τεχνική εφαρμογής του cepstral φίλτρου σε πολλαπλά επιπέδα.

Στο επόμενο κεφάλαιο θα εξετάσουμε αναλυτικότερα τη συμπεριφορά του cepstral φίλτρου και θα περιγράψουμε τους αλγορίθμους σύγκλισης που στηρίζονται σ' αυτό.

3.5 Υπολογισμός θέσης

Στην προηγούμενη παράγραφο είδαμε τεχνικές που μας βοηθούν στον υπολογισμό του λάθους παρακολούθησης στην κατεύθυνση του βάθους. Σ' αυτή την ενότητα, κατ' αντιστοιχία, θα αναφερθούμε σε μεθόδους που μπορούν να χρησιμοποιηθούν στον υπολογισμό του λάθους θέσης του στόχου.

Ένας τρόπος εύρεσης της θέσης και ολίσθησης του στόχου πάνω στην εικόνα είναι η ανάλυση της οπτικής ροής (φαινόμενη κίνηση των σημείων του χώρου, όπως αυτή παρουσιάζεται στις διαδοχικές εικόνες που καταγράφει η κάμερα ενός κινούμενου παρατηρητή). Η οπτική ροή εξαρτάται από την απόσταση των αντικειμένων και την κίνηση του ρομπότ. Όταν το ρομπότ είναι ακίνητο και το μόνο πράγμα που κινείται είναι ο στόχος, η ανάλυση είναι εύκολη. Η κίνηση, ωστόσο, των καμερών κατά την διαδικασία της παρακολούθησης εισάγει δυσκολίες στην ανάλυση της οπτικής ροής αφού οι κάμερες βλέπουν όλη τη σκηνή να κινείται. Οι τεχνικές ανάλυσης οπτικής ροής κυμαίνονται από την ανίχνευση κινούμενων αντικειμένων [3, 2, 49, 48, 52] ως την ανακατασκευή της σκηνής. Γενικά το σήμα της οπτικής ροής είναι σύνθετο και είναι δύσκολος ο διαχωρισμός της ολίσθησης του στόχου από την υπόλοιπη σκηνή.

Ένας άλλος τρόπος εντοπισμού του στόχου είναι τα φίλτρα σύγκρισης (matched filters) [18, 35, 17, 7, 37, 59]. Στηρίζονται στην επιλογή ενός κατάλληλου συνόλου από σημαντικά χαρακτηριστικά και μιας συνάρτησης βάρους η οποία εκφράζει τη σχέση τους με τη διαδικασία. Το σύνολο των χαρακτηριστικών μπορεί να αρχικοποιηθεί με σύζευξη των χαρακτηριστικών που βρίσκονται σ' ένα παράθυρο της εικόνας που είναι γνωστό ότι περιέχει το στόχο. Ένα πρόβλημα με τα φίλτρα αυτά είναι ότι δεν δουλεύουν σε περιπτώσεις που το αντικείμενο περιστρέφεται (view-dependent). Μια αντιμετώπιση αυτού του προβλήματος είναι η συσχέτιση του περιγράμματος του στόχου με την καινούργια εικόνα του στόχου. Ωστόσο χρειάζεται και κάποιος άλλος τρόπος κλειδώματος (locking) πάνω στο στόχο αλλιώς η προσήλωση είναι επιρρεπής να παρασυρθεί πάνω στο υπόστρωμα.

Θεωρώντας ότι η προσήλωση διατηρείται πάνω στο στόχο, είναι δυνατή η διατύπωση περιορισμών της θέσης του στόχου στην εικόνα και της ανομοιότητας του ([26]). Σ'

αυτή την κατηγορία ανήκουν τα φίλτρα μηδενικής ανομοιότητας που χρησιμοποιούμε εδώ και ουσιαστικά αγνοούν τα σημεία που βρίσκονται σε διαφορετικό βάθος από το σημείο καθήλωσης. Είναι μη γραμμικά φίλτρα, τα οποία πλεονεκτούν σε απόδοση και δυνατότητα παραλληλοποίησης σε σχέση με τις παραπάνω μεθόδους, αλλά μειονεκτούν σε ανθεκτικότητα στο θόρυβο. Θα εξετάσουμε τα φίλτρα αυτά, αναλυτικότερα, στο κεφάλαιο 5.

3.6 Έλεγχος του συστήματος

Το ελεγχόμενο σύστημα διαθέτει, όπως είπαμε, τέσσερις βαθμούς ελευθερίας (δύο για την κίνηση δεξιά-αριστερά των καμερών, έναν για την κίνηση του κεφαλιού πάνω-κάτω και έναν για την κίνηση του κεφαλιού δεξιά-αριστερά). Θα δούμε στην συνέχεια τι σφάλματα εισάγονται από την γεωμετρία του συστήματος αν κάνουμε την υπόθεση ότι όλοι οι άξονες περιστροφής περνάνε από τα οπτικά κέντρα των καμερών, πράγμα που δεν ισχύει εν γένει. Για μια τέτοια μελέτη χρειάζεται να ορίσουμε ένα μοντέλο κάμερας. Στο τέλος θα καταλήξουμε σε δύο εναλλακτικά σενάρια κίνησης της οπτικής πλατφόρμας τα οποία έχουν διαφορετική ευαισθησία σ' αυτά τα σφάλματα.

3.6.1 Μοντέλο καμέρας

Στην παρακάτω ανάλυση θα θεωρήσουμε κάμερες απειροελαχίστου ανοίγματος (pinhole camera model), με τετράγωνους αισθητήρες και με το οριζόντιο και κάθετο πεδίο θέας ίσα μεταξύ τους. Οι παράμετροι του πεδίου απεικόνισης είναι το πεδίο θέας, ϕ και ο αριθμός, N_p των φωτοευαίσθητων στοιχείων του αισθητήρα (οριζόντια/κάθετα). Αν w είναι το φυσικό πλάτος του αισθητήρα και f η εστιακή απόσταση ισχύει :

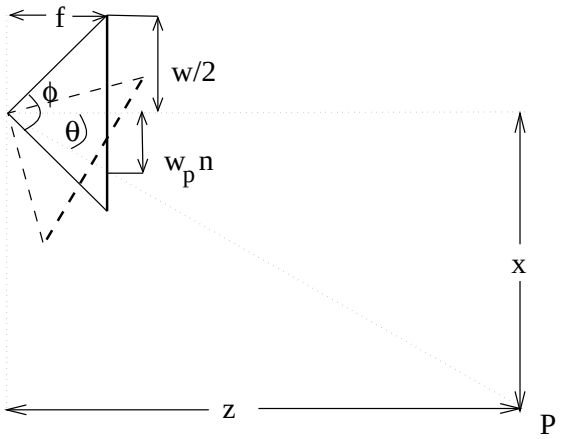
$$\tan\left(\frac{\phi}{2}\right) = \frac{w}{2\,f} \tag{3.9}$$

και λύνοντας ως προς ϕ :

$$\phi = 2\,\arctan\left(\frac{w}{2\,f}\right) \tag{3.10}$$

3.6.2 Σφάλματα στις διορθώσεις

Σ' αυτήν την ενότητα θα επιχείρησουμε να κάνουμε μια ανάλυση των σφαλμάτων τα οποία οφείλονται στη γεωμετρία του συστήματος αν κάνουμε την υπόθεση ότι όλοι οι άξονες περιστροφής περνάνε από τα οπτικά κέντρα των καμερών. Με τον όρο διόρθωση θα αναφερόμαστε στην κίνηση της πλατφόρμας με σκοπό οι οπτικοί άξονες των καμερών να συναντιούνται στην επιφάνεια του στόχου και χρησιμοποιώντας την παραπάνω υπόθεση. Με τον όρο σφάλμα διόρθωσης θα εννοούμε την ευκλείδια απόσταση της προβολής του σημείου προς σκόπευση από το κέντρο της εικόνας, αφού έχει ολοκληρωθεί η διόρθωση.



Σχήμα 3.8: Pan κίνηση της κάμερας με σκοπό να προσανατολιστεί πρός ένα σημείο P.

Πάνω όψη της κάμερας. Η απαιτούμενη αλλαγή θ στην pan γωνία της κάμερας είναι ανεξάρτητη της απόστασης του οπτικού της κέντρου από το σημείο P.

3.6.2.1 Κίνηση pan της κάμερας

Θεωρούμε ότι το οπτικό κέντρο των καμερών συμπίπτει με τις ανεξάρτητες αρθρώσεις pan. Το γεγονός αυτό σημαίνει ότι η κατάλληλη αλλαγή, θ, στην pan γωνία για να κεντραριστεί ένα σημείο P είναι ανεξάρτητη από την απόσταση του οπτικού κέντρου από το σημείο P (σχήμα 3.8). Αν w_p $(= \frac{w}{N_p})$ είναι το μέγεθος του pixel και n η x συντεταγμένη της προβολής του σημείου P στο σύστημα συντεταγμένων της εικόνας :

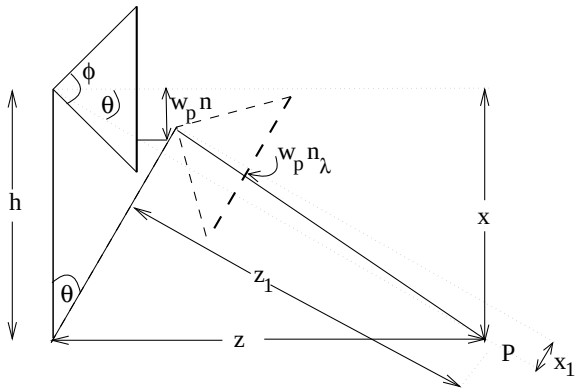
$$\tan(\theta) = \frac{w_p\, n}{f} \tag{3.11}$$

και χρησιμοποιώντας την εξίσωση 3.9 :

$$\theta = \arctan\left(\frac{2\, n}{N_p} \tan\left(\frac{\phi}{2}\right)\right) \tag{3.12}$$

3.6.2.2 Κίνηση tilt του κεφαλιού

Το οπτικό κέντρο των καμερών δεν συμπίπτει με τις tilt αρθρώσεις του κεφαλιού. Αυτό έχει σαν αποτέλεσμα η κατάλληλη αλλαγή στην γωνία tilt για να κεντραριστεί ένα σημείο P να εξαρτάται από την απόσταση του tilt άξονα από το οπτικό κέντρο. Η πληροφορία του βάθους όμως μπορεί να εξαχθεί με επιπλέον επεξεργασία του οπτικού σήματος, πράγμα που θα έκανε το σύστημα παρακολούθησης πιο σύνθετο. Θα δούμε τι σφάλμα εισάγει η χρήση της εξίσωσης 3.12 στον έλεγχο της γωνίας tilt. Σύμφωνα με το σχήμα 3.9 :



Σχήμα 3.9: Tilt κίνηση του κεφαλιού με σκοπό να προσανατολιστεί πρός ένα σημείο P.

Πλαϊνή όψη της κάμερας. Αν χρησιμοποιηθεί η εξίσωση 3.12, η προβολή του σημείου P δεν συμπίπτει με το κέντρο της εικόνας μετά την κίνηση

$$x_1 = h - h\cos(\theta) \tag{3.13}$$

$$z_1 = d - h\sin(\theta) \tag{3.14}$$

όπου $d = \sqrt{x^2 + z^2}$. Συνεπώς μετά την tilt κίνηση η προβολή του σημείου P θα είναι (pinhole camera equation):

$$w_p\, n_\lambda = f\,\frac{x_1}{z_1} \tag{3.15}$$

Συνδυάζοντας τις εξισώσεις 3.13, 3.14 και 3.15 :

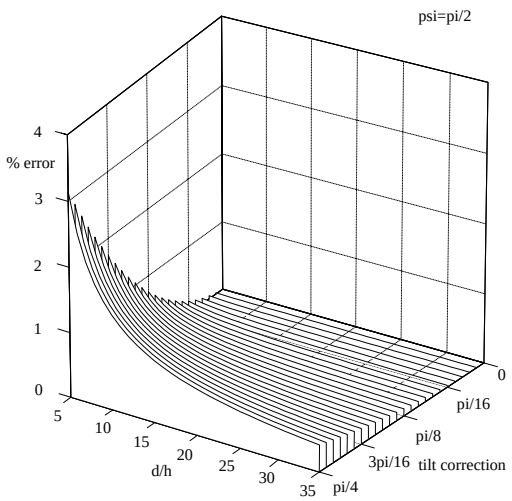
$$\begin{aligned} n_\lambda &= \frac{f\,h(1-\cos(\theta))}{w_p(d - h\sin(\theta))} \\ &= \frac{N_p(1-\cos(\theta))}{2\tan(\frac{\phi}{2})(\frac{d}{h} - \sin(\theta))} \end{aligned} \tag{3.16}$$

Από την εξίσωση 3.16, η οποία απεικονίζεται στο σχήμα 3.10 για $\pi/2$ γωνία θέας, μπορούμε να παρατηρήσουμε τα εξής για την συμπεριφορά του σφάλματος διόρθωσης :

1. Το σφάλμα είναι μικρότερο για μακρινά αντικείμενα και μεγαλύτερο για κοντινά αντικείμενα.

2. Το σφάλμα είναι μικρότερο για μεγάλα πεδία θέας και μεγαλύτερο για μικρά πεδία θέας.

3. Το σφάλμα είναι μεγαλύτερο για μεγάλες διορθώσεις στη γωνία tilt. Αυτό ως προς την παρακολούθηση σημαίνει ότι έχουμε μεγάλο σφάλμα για μεγάλη κίνηση στην κάθετη κατεύθυνση.



Σχήμα 3.10: Σφάλμα διόρθωσης με την tilt κίνηση του κεφαλιού, ως συνάρτηση του λόγου d/h και της διόρθωσης. Το πεδίο θέας είναι $\pi/2$.

Η διόρθωση ορίζεται από την σχέση 3.12.

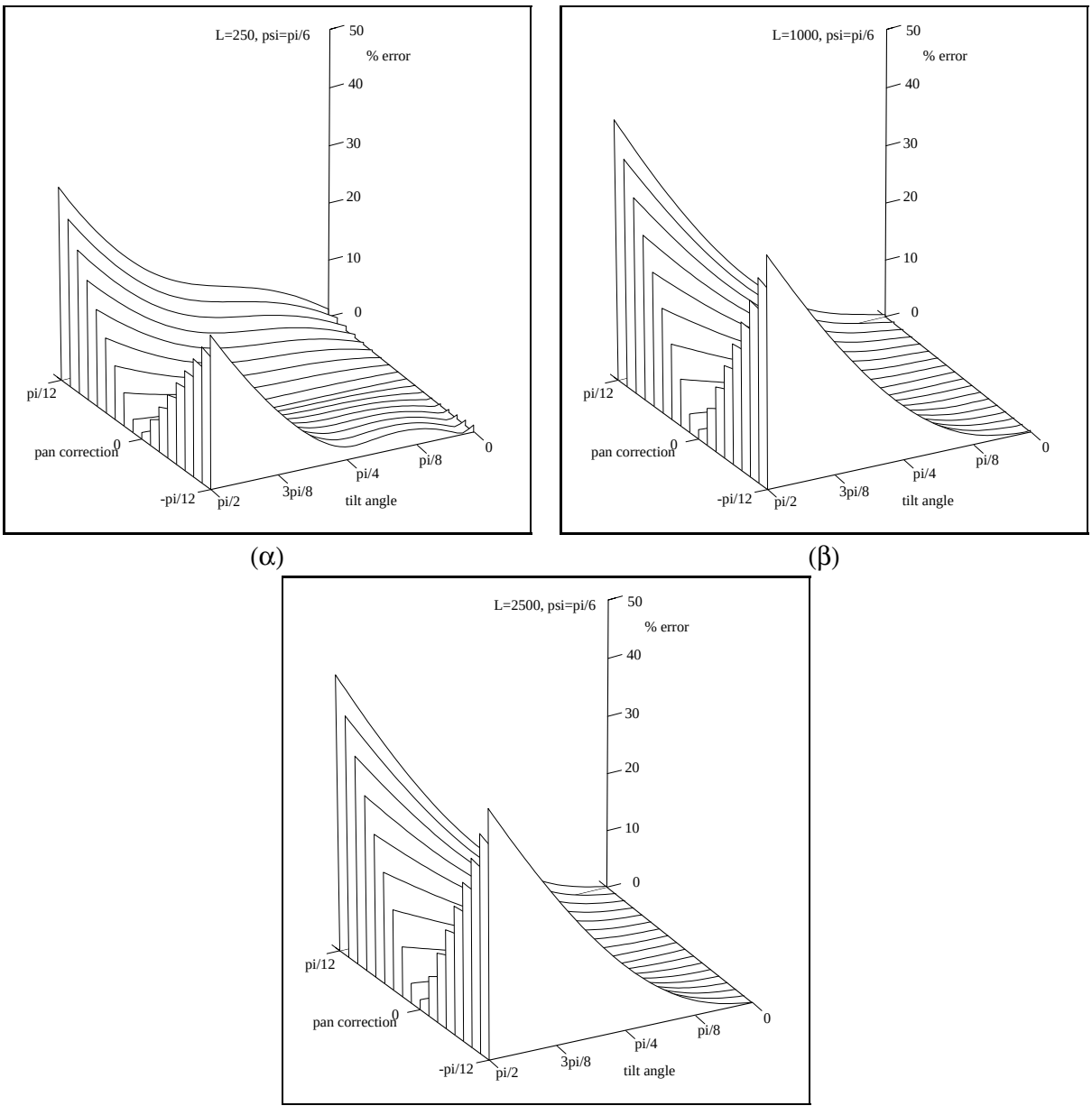
3.6.2.3 Κίνηση pan του κεφαλιού

Ο υπολογισμός του σφάλματος (Παράρτημα Α) που εισάγεται για οριζόντιες διορθώσεις, σύμφωνα με την σχέση 3.12, με χρήση της pan κίνησης του κεφαλιού στην γενική περίπτωση είναι πιο σύνθετος απ' ότι για την tilt κίνηση. Όταν η γωνία tilt είναι μηδέν ισχύει η σχέση 3.16, όπου h είναι στην προκειμένη περίπτωση το μισό μήκος της βασικής γραμμής (baseline). Μη μηδενική γωνία tilt σημαίνει όμως ότι η διόρθωση δεν γίνεται ως προς την οριζόντια κατεύθυνση των καμερών. Για να γίνει κατανοητό αυτό, ας σκεφτούμε την περίπτωση που η γωνία tilt είναι 90° (μέγιστη). Ο pan άξονας είναι παράλληλος με την κατεύθυνση z (βάθος).

Στα σχήματα 3.11, 3.12 βλέπουμε το σφάλμα που οφείλεται στην διόρθωση με το pan του κεφαλιού σε ποσοστό επί τοις εκατό του πλάτους του αισθητήρα ως προς την γωνία tilt και την διόρθωση σε rads. Οι γραφικές παραστάσεις στο 3.11 αντιστοιχούν σε πεδίο θέας $\frac{\pi}{6}$ ενώ στο 3.12 $\frac{\pi}{2}$. Η απόσταση του σημείου απ' το οπτικό κέντρο της κάμερας είναι για τα (α),(β),(γ): $250mm$, $1000mm$ και $2500mm$ αντίστοιχα. Ο tilt βραχίονας είναι $100mm$ καθώς επίσης και το μήκος της βασικής γραμμής. Από τις γραφικές παραστάσεις του σφάλματος μπορούμε να παρατηρήσουμε τα εξής :

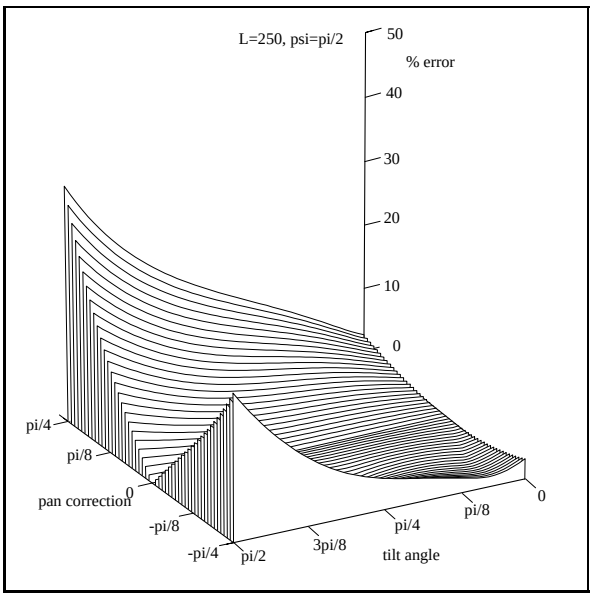
1. Το σφάλμα είναι μεγαλύτερο για μεγάλες γωνίες tilt.

2. Το σφάλμα αυξάνει με την απόσταση για tilt γωνία μεγαλύτερη από $\frac{\pi}{4}$ περίπου.

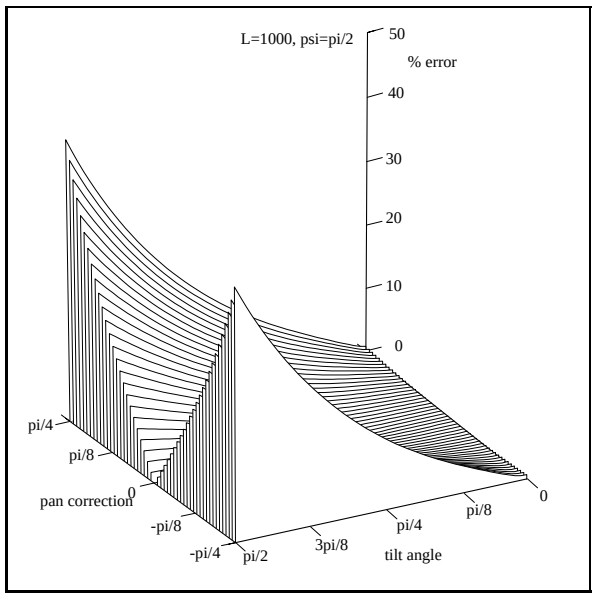


Σχήμα 3.11: Σφάλμα διόρθωσης που οφείλεται στην pan κίνηση του κεφαλιού, ως συνάρτηση της tilt γωνίας και της διόρθωσης. Το πεδίο θέας είναι $\pi/6$.

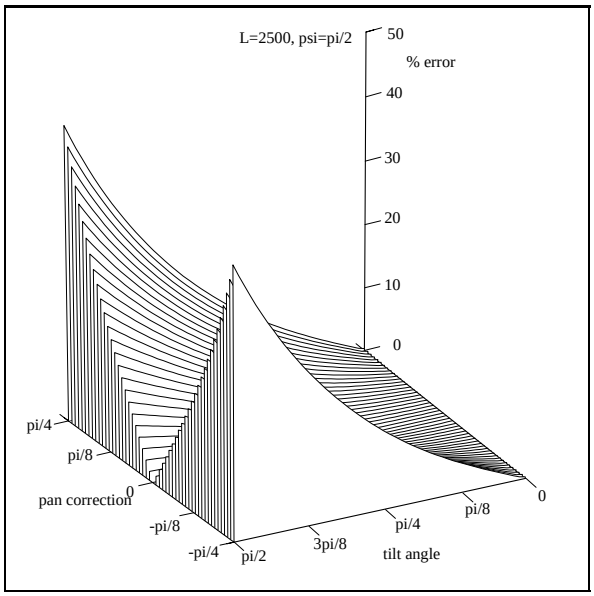
Η διόρθωση ορίζεται από την σχέση 3.12. Το σημείο P βρίσκεται σε απόσταση (α) $250mm$, (β) $1000mm$, (γ)$2500mm$, από το οπτικό κέντρο της κάμερας



(α)



(β)



(γ)

Σχήμα 3.12: Σφάλμα διόρθωσης που οφείλεται στην pan κίνηση του κεφαλιού, ως συνάρτηση της tilt γωνίας και της διόρθωσης. Το πεδίο θέας είναι $\pi/2$.

Η διόρθωση ορίζεται από την σχέση 3.12. Το σημείο P βρίσκεται σε απόσταση (α) $250mm$, (β) $1000mm$, (γ)$2500mm$, από το οπτικό κέντρο της κάμερας

3. Η συμπεριφορά του σφάλματος δεν είναι ίδια για διορθώσεις ίσες κατά απόλυτη τιμή αλλά αντίθετου προσήμου. Αυτο οφείλεται στο ότι ο αισθητήρας στη μία περίπτωση πλησιάζει το σημείο και στην άλλη απομακρύνεται απ' αυτό. Το φαινόμενο αυτό γίνεται πιο έντονο για μικρότερες αποστάσεις του σημείου από το οπτικό κέντρο.

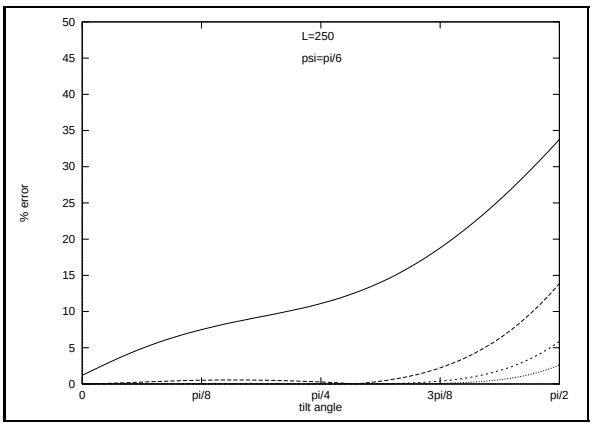
Στα σχήματα 3.13, 3.14, 3.15 βλέπουμε το σφάλμα, την $x-$συντεταγμένη του, και την $y-$συντεταγμένη του αντίστοιχα, που εισάγεται για μέγιστη διόρθωση (δηλ. το σημείο P αρχικά προβάλλεται στην άκρη της εικόνας) ως συνάρτηση της γωνίας tilt. Οι συνεχείς γραμμές παριστάνουν το σφάλμα που οφείλεται στην αρχική διόρθωση ενώ οι υπόλοιπες έχουν προκύψει από τρεις επόμενες διορθώσεις (pan για την $x-$συντεταγμένη και tilt για την $y-$συντεταγμένη του σφάλματος). Η ανάλυση του σφάλματος που παρουσιάζεται στο σχήμα 3.13 είναι τελείως αντίστοιχη με αυτήν που γίνεται από τον Carl-Johan Westelius στην διδακτορική διατριβή του [72]. Στο σχήμα 3.16 βλέπουμε την τροχιά της προβολής του σημείου P στην ακραία περίπτωση που το σημείο αρχικά προβάλλεται στην άκρη της εικόνας (μέγιστη διόρθωση) και η αρχική tilt γωνία είναι μέγιστη (90°). Η τροχιά έχει προκύψει από τέσσερις διορθώσεις pan πού εναλλάσσονται από tilt διορθώσεις προκειμένου να μηδενιστεί το σφάλμα στην κάθετη κατεύθυνση. Παρατηρούμε ότι :

1. Το σφάλμα συγκλίνει γρηγορότερα στο μηδέν για κοντινότερα σημεία.
2. Το σφάλμα συγκλίνει γρηγορότερα στο μηδέν για μεγαλύτερες διορθώσεις.
3. Το σφάλμα συγκλίνει γρηγόροτερα όσο μεγαλύτερο είναι το λάθος που εισάγει στην $y-$κατεύθυνση η αρχική pan κίνηση. Αυτό συμβαίνει γιατί η αρχική μέγιστη γωνία tilt (για την οποία εξετάζουμε τις τροχιές) μειώνεται πιο σημαντικά κατά τις ενδιάμεσες tilt κινήσεις οπότε οι επόμενες pan κινήσεις εισάγουν μικρότερο σφάλμα.

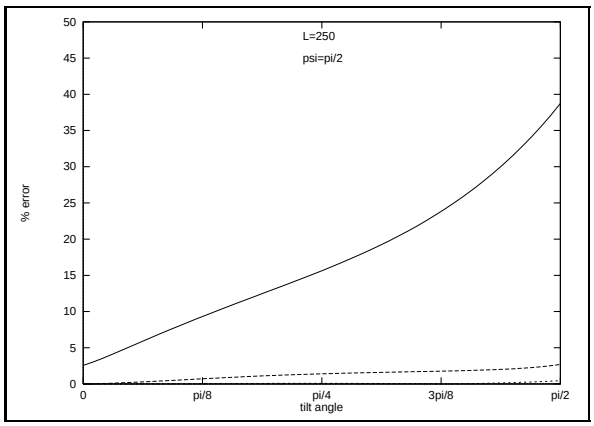
3.6.3 Σενάρια κατανομής των σφαλμάτων στους κινητήρες

Ο προσανατολισμός μιας μηχανικής πλατφόρμας είναι ένα σημαντικό θέμα έρευνας για την μηχανική όραση. Όσο πιο σύνθετες είναι οι ικανότητες με τις οποίες θέλουμε να εμπλουτίσουμε ένα ρομπότ, τόσο πιο σύνθετο είναι το πρόβλημα του προσανατολισμού. Για την παρακολούθηση μας ενδιαφέρει να ορίσουμε ένα απλό αντανακλαστικό τρόπο προσανατολισμού της μηχανικής πλατφόρμας.

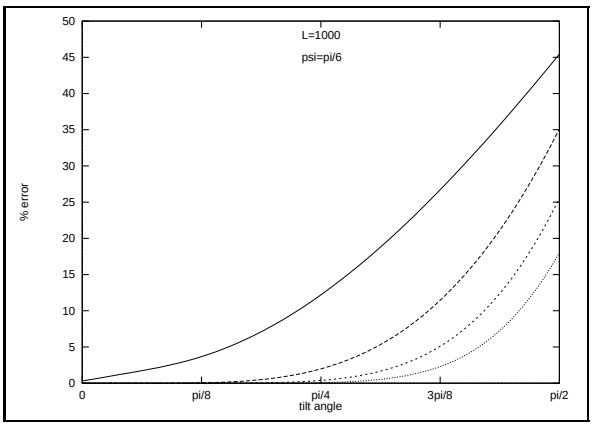
Όπως είπαμε και στην ενότητα 3.3 θέλουμε να ορίσουμε μια αντιστοιχία από τις παραμέτρους προσήλωσης στις διορθώσεις $\phi_l, \phi_r, \phi_p, \phi_t$. με τις οποίες θα τροφοδοτήσουμε τους ελεγκτές. Οι ελεγκτές, στη συνέχεια, θα μετατρέψουν τις διορθώσεις σε σήματα ταχυτήτων και/ή επιταγχύνσεων για να οδηγήσουν το ελεγχόμενο σύστημα. Θα ορίσουμε δύο τρόπους να γίνει η αντιστοίχιση αυτή από τις εξόδους των εκτιμητών λάθους σε εισόδους στο σύστημα ελέγχου.



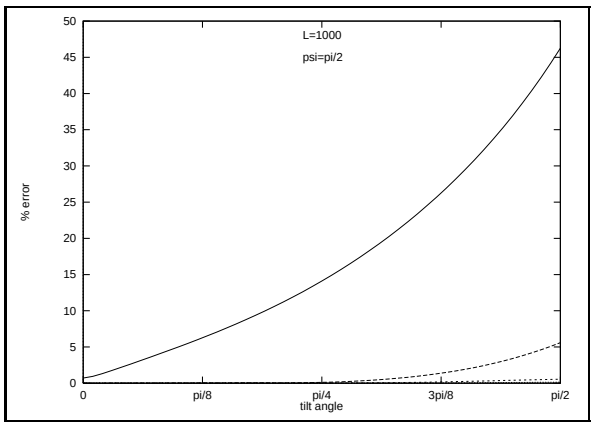
(α)



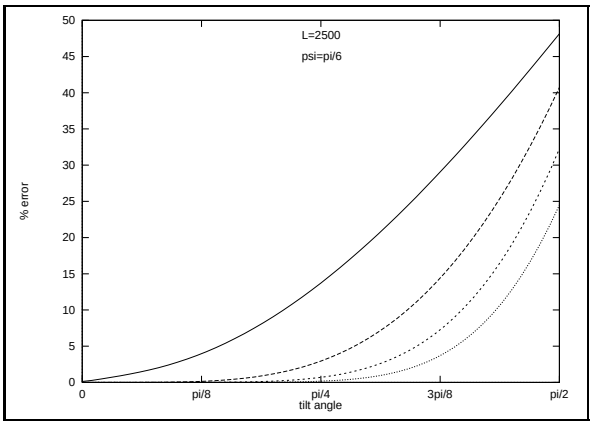
(δ)



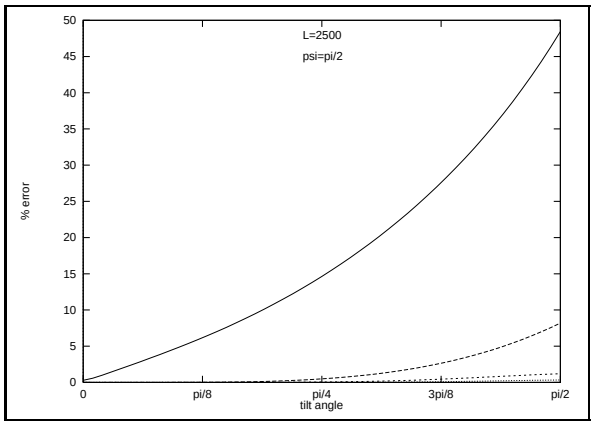
(β)



(ε)



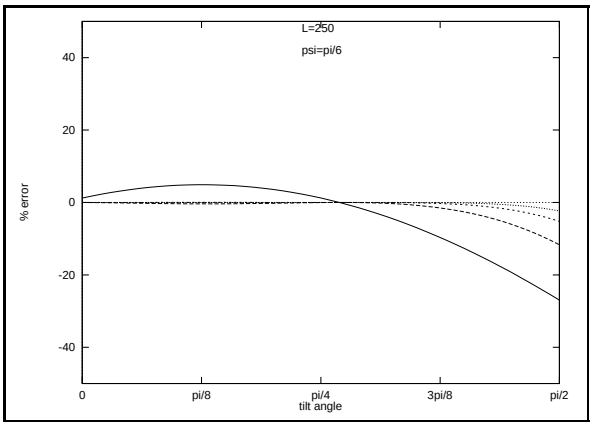
(γ)



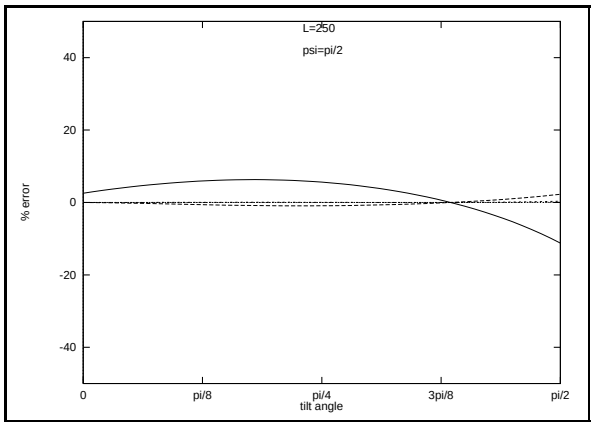
(στ)

Σχήμα 3.13: Σφάλμα διόρθωσης που οφείλεται στην pan κίνηση του κεφαλιού, ως συνάρτηση της γωνίας tilt. Η διόρθωση ορίζεται από ένα σημείο P στην άκρη της εικόνας.

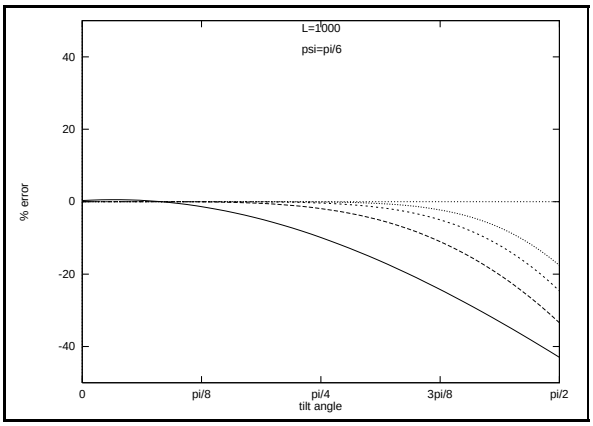
Στο σχήμα αυτό απεικονίζεται στην αριστερή στήλη η τομή του σχήματος 3.11 για διόρθωση $\frac{\pi}{12}$ και στην δεξιά στήλη η τομή του σχήματος 3.12 για διόρθωση $\frac{\pi}{4}$. (δηλ. στα α, β, γ (δ, ε, στ) το πεδίο θέας είναι $\frac{\pi}{6}$ ($\frac{\pi}{2}$), στα α,δ το σημείο P βρίσκεται σε απόσταση $250mm$ ενώ στα β,ε και γ,στ, $1000mm$ και $2500mm$ αντίστοιχα). Η συνεχής καμπύλη αντιστοιχεί στο σφάλμα που οφείλεται στην 1η διόρθωση με την pan κίνηση του κεφαλιού. Οι υπόλοιπες καμπύλες έχουν προκύψει από επόμενες διορθώσεις (pan για την $x-$συντεταγμένη και tilt για την $y-$ συντεταγμένη του σφάλματος).



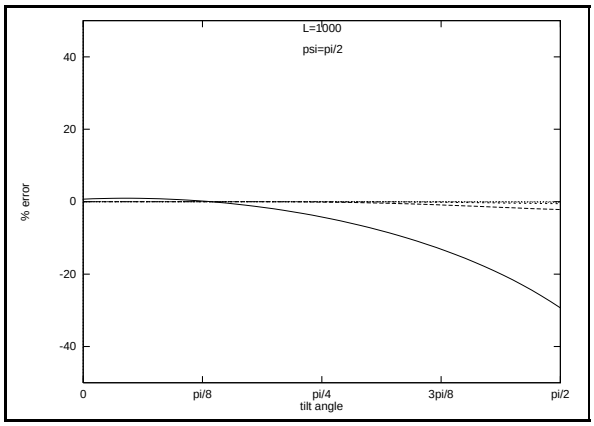
(α)



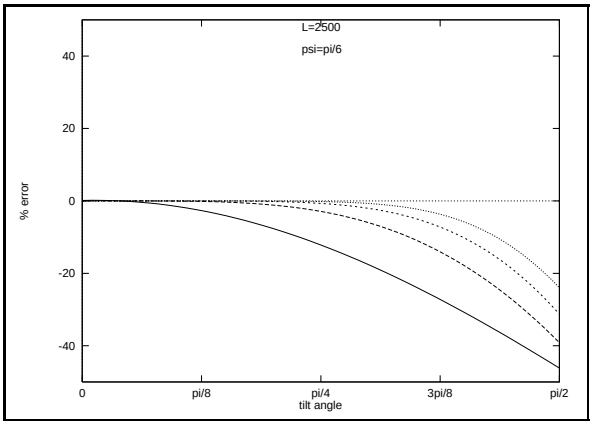
(δ)



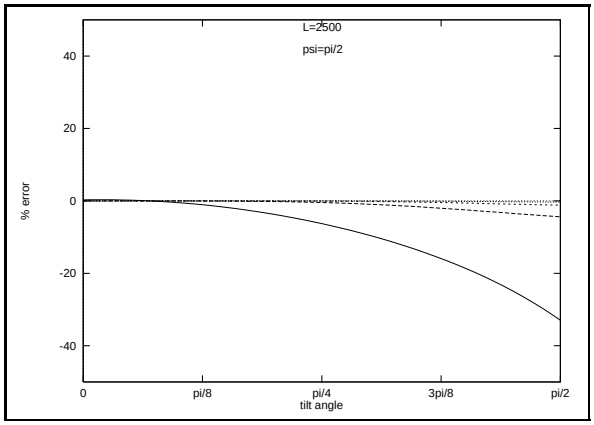
(β)



(ε)



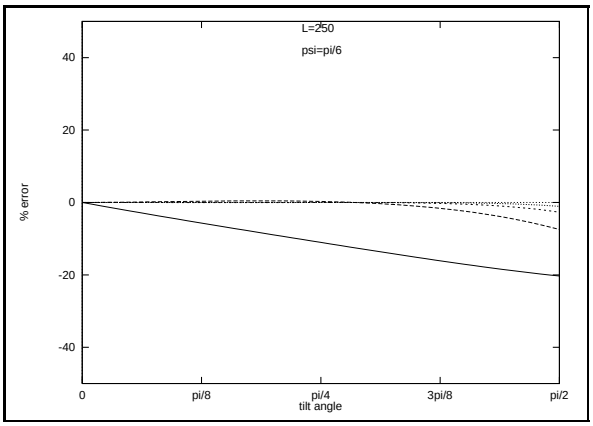
(γ)



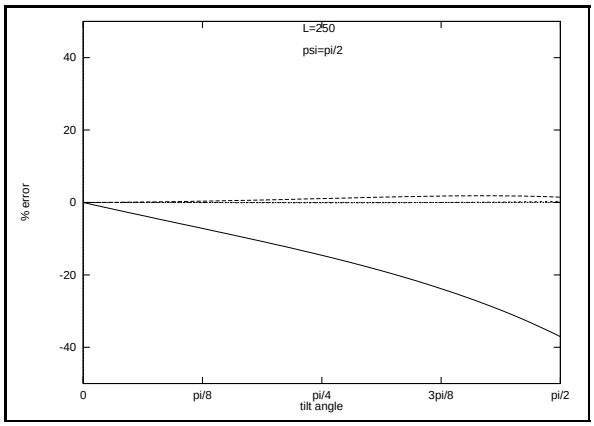
(στ)

Σχήμα 3.14: Η $x-$συντεταγμένη του σφάλματος διόρθωσης με την pan κίνηση του κεφαλιού, ως συνάρτηση της γωνίας tilt.

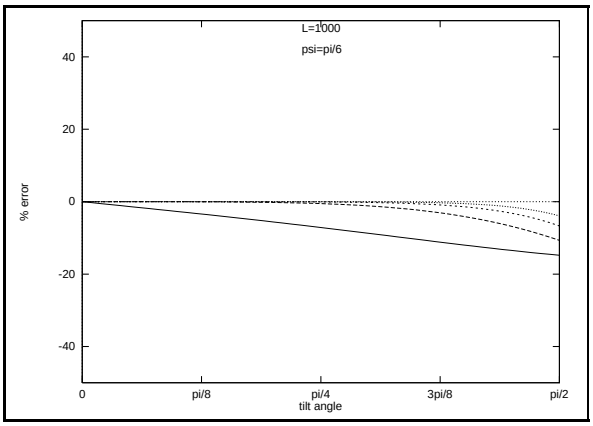
Στο σχήμα αυτό βλέπουμε την $x-$συντεταγμένη του σφάλματος που απεικονίζεται στο σχήμα 3.13



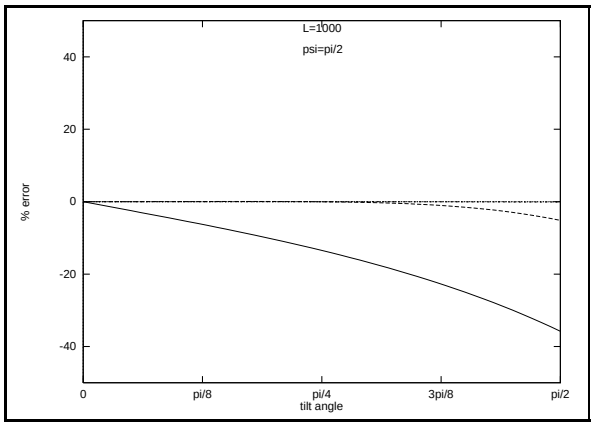
(α)



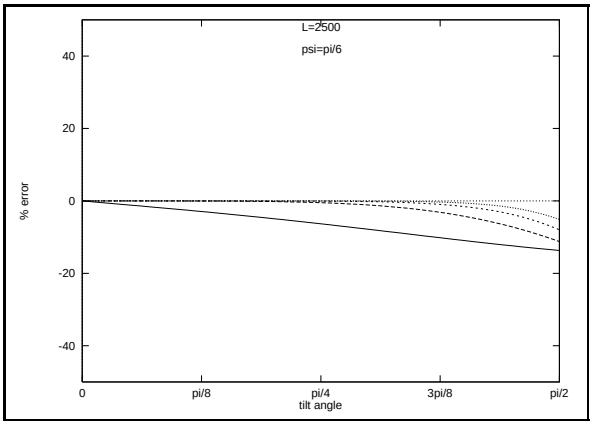
(δ)



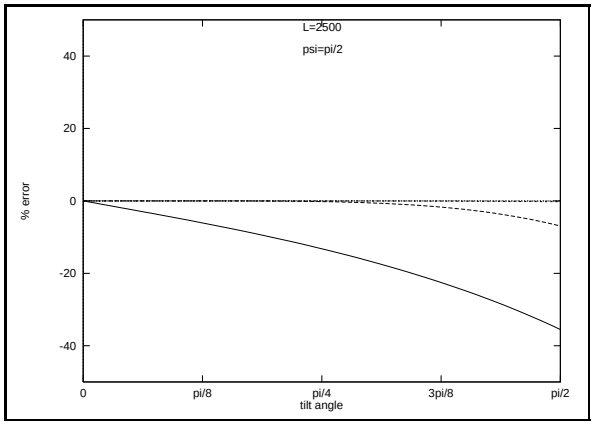
(β)



(ε)



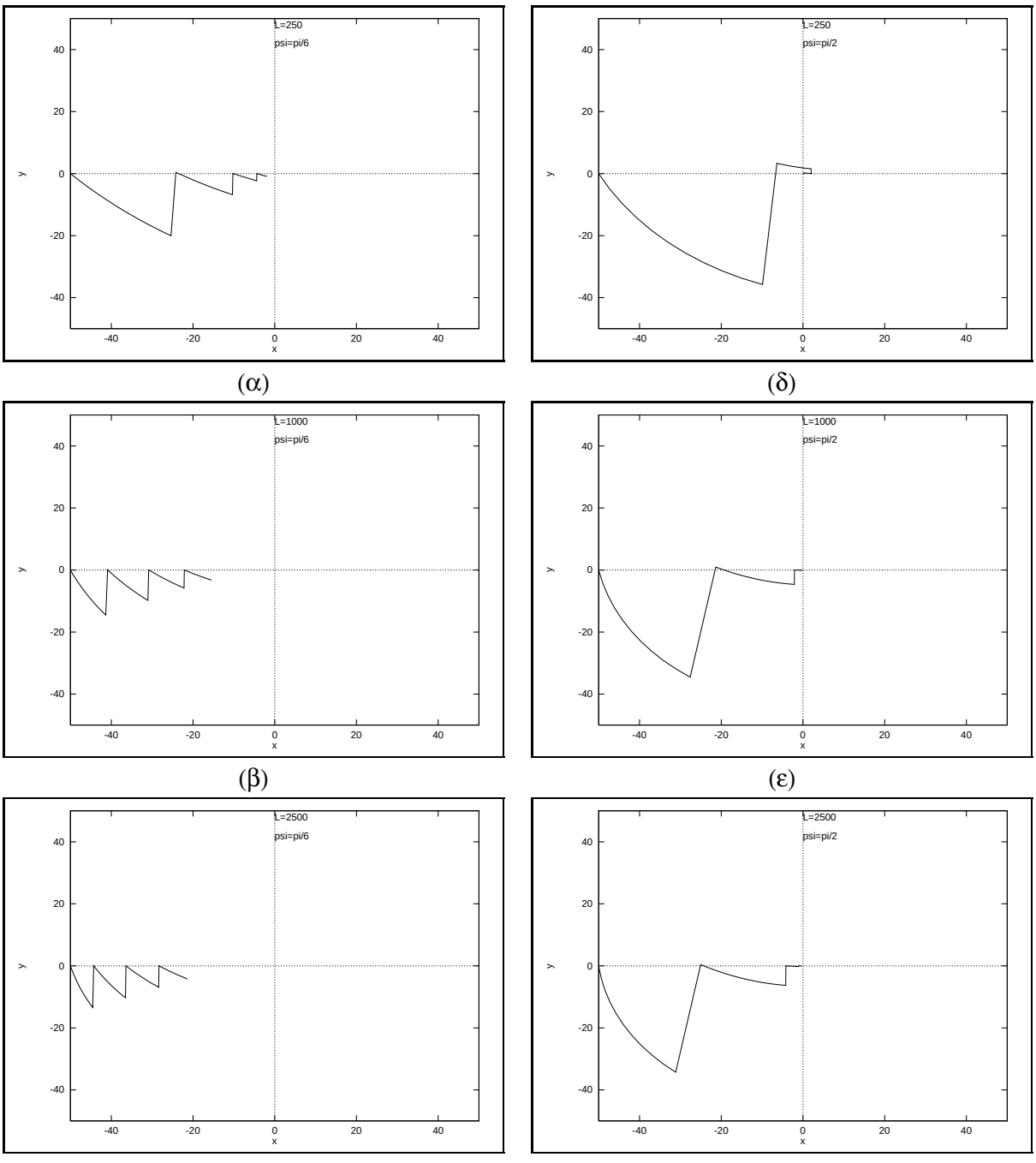
(γ)



(στ)

Σχήμα 3.15: Η $y-$συντεταγμένη του σφάλματος διόρθωσης με την pan κίνηση του κεφαλιού, ως συνάρτηση της γωνίας tilt.

Στο σχήμα αυτό βλέπουμε την $y-$συντεταγμένη του σφάλματος που απεικονίζεται στο σχήμα 3.13



(α) (δ)

(β) (ε)

(γ) (στ)

Σχήμα 3.16: Η τροχιά της προβολής του σημείου P (απ' το οποίο ορίζεται η διόρθωση) κατά τη διόρθωση με την pan κίνηση του κεφαλιού και στη συνέχεια με ζεύγη διορθώσεων tilt-pan (pan για την $x-$συντεταγμένη και tilt για την $y-$συντεταγμένη του σφάλματος). Η αρχική γωνία tilt είναι $\pi/2$.

Το σημείο P που ορίζει την διόρθωση βρίσκεται σε απόσταση $250mm$ στα α,δ, $1000mm$ στα β,ε, και $2500mm$ στα γ,στ. Το πεδίο θέας είναι $\frac{\pi}{6}$ στα α,β,γ και $\frac{\pi}{2}$ στα δ,ε,στ.

Το πρώτο σενάριο κίνησης δεν χρησιμοποιεί την pan κίνηση του κεφαλιού (πράγμα που ενθαρρύνει η ύπαρξη μεγάλων σφαλμάτων κατά τη χρήση της κίνησης αυτής). Έτσι οι ανεξάρτητες pan αρθρώσεις αναλαμβάνουν να διορθώσουν και το λάθος στη γωνία προσήλωσης στην οριζόντια κατεύθυνση :

$$\phi_l = \theta^e_{g_x} - \frac{\theta^e_v}{2} \tag{3.17}$$

$$\phi_r = \theta^e_{g_x} + \frac{\theta^e_v}{2} \tag{3.18}$$

ενώ το λάθος στην κάθετη κατεύθυνση (δηλ. στην γωνία του επιπέδου προσήλωσης αναιρείται με χρήση της κοινής tilt άρθρωσης :

$$\phi_t = \theta^e_{g_y} \tag{3.19}$$

Σύμφωνα με το δεύτερο σενάριο κίνησης χρησιμοποιείται η pan άρθρωση του κεφαλιού :

$$\phi_p = \theta^e_{g_x} \tag{3.20}$$

και ενώ η διόρθωση με τον tilt βαθμό ελευθερίας δίνεται από την εξίσωση 3.19, οι pan αρθρώσεις των καμερών διορθώνουν μόνο το λάθος σύγκλισης :

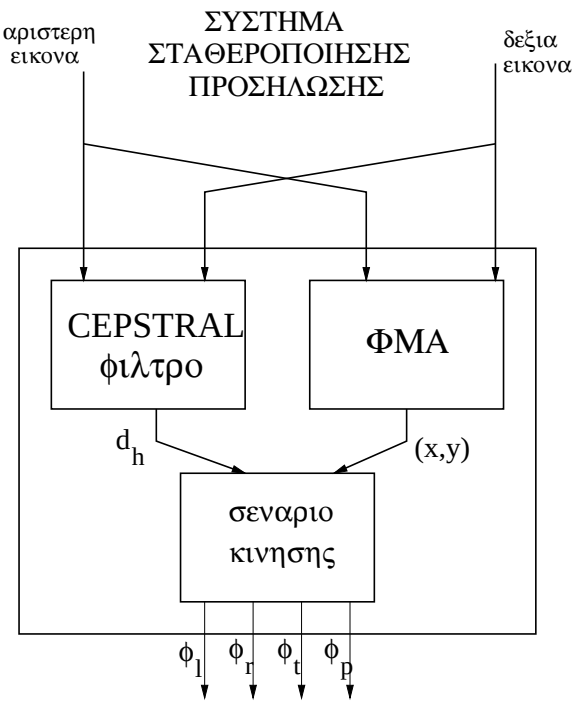
$$\phi_l = -\frac{\theta^e_v}{2} \tag{3.21}$$

$$\phi_r = \frac{\theta^e_v}{2} \tag{3.22}$$

Θα δούμε στο κεφάλαιο 6 πως συμπεριφέρεται το σύστημα παρακολούθησης ανάλογα με το σενάριο κίνησης με το οποίο κατευθύνει την προσήλωση.

3.7 Συνολική θεώρηση του συστήματος

Στην ενότητα αυτή ανακεφαλαιώνουμε τις επιλογές στην σχεδίαση του ΣΣΠ που υλοποιήσαμε. Ο εκτιμητής ανομοιότητας έχει υλοποιηθεί με τεχνικές συσχέτισης που στηρίζονται στο cepstral φίλτρο. Για τον εκτιμητή θέσης έχουν υλοποιηθεί διάφορα φίλτρα μηδενικής ανομοιότητας. Τα δύο αυτά κομμάτια του συστήματος θα περιγράψουμε στα κέφαλαια 4 και 5. Οι εκτιμήσεις αυτές χρησιμοποιούνται από τα δύο σενάρια κίνησης που περιγράψαμε στην ενότητα 3.6.3 για τον προσανατολισμό του στερεοσκοπικού κεφαλιού. Συνολικά το σύστημα σταθεροποίησης φαίνεται στο σχήμα 3.17.



Σχήμα 3.17: Διάγραμμα λειτουργίας του συστήματος σταθεροποίησης της οπτικής προσήλωσης

Το cepstral φίλτρο καθώς και το φίλτρο μηδενικής ανομοιότητας (ΦΜΑ) επεξεργάζονται τη στερεοσκοπική είσοδο και τροφοδοτούν το σενάριο κίνησης της οπτικής πλατφόρμας με τις εκτιμήσεις της ανομοιότητας d_h και της θέσης (x, y) του στόχου. Το σενάριο κίνησης μετατρέπει αυτές τις εκτιμήσεις σ' ένα διάνυσμα διορθώσεων $(\phi_l, \phi_r, \phi_t, \phi_p)$ το οποίο κατανέμεται στους τέσσερις βαθμούς ελευθερίας του κεφαλιού.

Κεφάλαιο 4

Cepstral φίλτρο

Σ' αυτό το κεφάλαιο θα περιγράψουμε τεχνικές υπολογισμού της στερεοσκοπικής ανομοιότητας που στηρίζονται στο φίλτρο cepstrum. Όπως είπαμε στο προηγούμενο κεφάλαιο ο υπολογισμός της ανομοιότητας χρησιμοποιείται για την εκτίμηση του λάθους παρακολούθησης στην κατεύθυνση του βάθους. Θα περιγράψουμε δύο μεθόδους εξαγωγής της ανομοιότητας του στόχου στις ενότητες 4.3.3 και 4.5 τις οποίες χρησιμοποιήσαμε στο σύστημα σταθεροποίησης της προσήλωσης.

4.1 Βιολογικό υπόβαθρο

Τα τελευταία χρόνια ένα σοβαρό βήμα στην νευροβιολογία είναι η ανακάλυψη της λειτουργικής εξειδίκευσης *(functional specialization)* του εγκεφάλου, δηλ. η εξειδίκευση διαφορετικών περιοχών του εγκεφάλου για συγκεκριμένες λειτουργίες του οπτικού συστήματος. Αυτό σημαίνει ότι υπάρχουν διαφορετικά κέντρα του εγκεφάλου για την ανίχνευση χρώματος, κίνησης [76] ή άλλων στοιχείων που συνθέτουν την αίσθηση της όρασης, πράγμα που έρχεται σε σύγκρουση με την προηγούμενη θεωρία σύμφωνα με την οποία η λειτουργία της όρασης σχετίζεται μόνο με μια περιοχή του φλοιού. Πολλοί ερευνητές έχουν ασχοληθεί με την αντιστοίχηση των οπτικών λειτουργιών στα διάφορα μέρη του εγκεφάλου και πως μεταφέρεται σ' αυτά η οπτική πληροφορία από τον αμφιβληστροειδή. Η έρευνα που έχει γίνει υποστηρίζει την ύπαρξη τριων (χοντρικά) καναλιών επεξεργασίας που εξυπηρετούν διαφορετικούς στόχους. Ένα απ' αυτά είναι το κανάλι *magno*, το οποίο δεν επεξεργάζεται την πληροφορία για το χρώμα (*colour-blind*) αλλά κάνει υπολογισμούς που αφορούν την αντίληψη του βάθους και της κίνησης

μεταφέροντας μόνο την πληροφορία που είναι σχετική με το φωτισμό (luminance-specific). Η πληροφορία αυτή καταλήγει στον ραβδωτό φλοιό *(striate cortex)* σε μορφή λωρίδων που είναι γνωστές ως *ocular dominance columns*. Οι λωρίδες αυτές εναλλάσσονται ως προς την προέλευση - δεξί/αριστερό μάτι- της πληροφορίας που φέρουν και διαχωρίζονται από ωχρές ζώνες (pale bands).

Δεν είναι ξεκάθαρο ακόμα το γιατί είναι χρήσιμη μια τετοια δομή. Μια απάντηση σ' αυτό το ερώτημα πηγάζει από την γεωμετρία διοφθαλμικής προσήλωσης. Πιθανόν η περιοχή του Panum (μια μικρή ζώνη γύρω από το ορόπτερο, σταθερού μεγέθους) να έχει σχέση με το σταθερό πλάτος των λωρίδων. Μ' άλλα λόγια η διάταξη της στερεοσκοπικής πληροφορίας σε μορφή εναλλάξ λωρίδων διευκολύνει την συγχώνευση (fusion) της για την αντίληψη του βάθους τοπικά.

Πέρα από την έρευνα στην νευροβιολογία, η ύπαρξη των ocular dominance columns αποτελεί έναυσμα για έρευνα και στην μηχανική όραση. Παρακάτω θα αναφερθούμε αναλυτικότερα στον αλγόριθμο των Yeshurun και Shwartz [75] για τον υπολογισμό ενός αραιού πεδίου ανομοιότητας με χρήση του cepstral φίλτρου και την προσαρμογή αυτού του αλγορίθμου στα πλαίσια της ενεργού όρασης για να παρέχει τη δυνατότητα σύγκλισης σε στερεοσκοπικά συστήματα [53, 66, 38].

4.2 Μαθηματικό υπόβαθρο

Το *cepstrum* είναι αναγραμματισμός της λέξης spectrum (φάσμα). Το *cepstrum* $\mathcal{C}\{\cdot\}$ ενός σήματος $s(\vec{x})$ ορίζεται ως το φάσμα ισχύος του λογαρίθμου του φάσματος ισχύος του (εξίσωση 4.3). Το φάσμα ισχύος *(power spectrum)* είναι ο μετασχηματισμός Fourier της συνάρτησης αυτοσυσχέτισης ϕ_{ss} του σήματος [1] οπότε προκύπτει η εξίσωση 4.4.

$$\mathcal{C}\{s\} = \|\mathcal{F}\{\log(\mathcal{F}\{\phi_{ss}\})\}\|^2 \tag{4.3}$$

$$= \|\mathcal{F}\left\{\log\left(\|\mathcal{F}\{s\}\|^2\right)\right\}\|^2 \tag{4.4}$$

Οπως φαίνεται από τον παραπάνω ορισμό είναι ένα μη γραμμικό φίλτρο. Αρχικά χρησιμοποιήθηκε σαν εργαλείο ανάλυσης μονοδιάστατων σεισμικών σημάτων για την ανίχνευση του χρόνου άφιξης της ηχούς *(echo)* - echo arrival times.

[1] Η αυτοσυσχέτιση μπορεί να εκφραστεί ως :

$$\phi_{ss}(\vec{x}) = s(\vec{x}) \star s^*(-\vec{x}) \tag{4.1}$$

όπου $\star$ ο τελεστής της συνέλιξης. Λαμβάνοντας υπ' όψην τη σχέση 4.1, παίρνουμε το μετασχηματισμό Fourier της αυτοσυσχέτισης:

$$\begin{aligned}\mathcal{F}\{\phi_{ss}(\vec{x})\} &= \mathcal{F}\{s(\vec{x})\}\,\mathcal{F}\{s^*(-\vec{x})\} \\ &= \mathcal{F}\{s(\vec{x})\}\,\mathcal{F}^*\{s(\vec{x})\} \\ &= \|\mathcal{F}\{s(\vec{x})\}\|^2\end{aligned} \tag{4.2}$$

Οι Olson και Coombs [53] υποδεικνύουν πως μπορεί να κατανοηθεί η συμπεριφορά του φίλτρου αυτού προσπαθώντας να προσδιορίσουν τον ρόλο του λογαρίθμου στον ορισμό του φίλτρου. Εδώ παρουσιάζουμε την ανάλυση αυτή πιστεύοντας ότι έχει σημασία στο να αποκτήσουμε μία όσο γίνεται πιο ολοκληρωμένη εικόνα για τους αλγορίθμους σύγκλισης που θα παρουσιάσουμε στις επόμενες ενότητες.

Το φάσμα ισχύος ενός σήματος έχει τιμές στο σύνολο $\Re$ των πραγματικών και είναι άρτιας συμμετρίας όπως και η συνάρτηση αυτοσυσχέτισης του. Επομένως ο ευθύς και ο αντίστροφος μετασχηματισμός Fourier είναι ισοδύναμοι. Αυτό σημαίνει ότι χωρίς το βήμα της λογαρίθμησης ο αλγόριθμος υπολογισμού του cepstrum θα έδινε απλά την ενέργεια της αυτοσυσχέτισης του σήματος. Ποια είναι όμως ακριβώς η επίδραση του λογαρίθμου. Αν ξαναγράψουμε το λογάριθμο του φάσματος ισχύος ως εξής :

$$\begin{aligned} \log \|\mathcal{F}\{s\}\|^2 &= \frac{\log \|\mathcal{F}\{s\}\|^2}{\|\mathcal{F}\{s\}\|^2} \mathcal{F}\{s\}\, \mathcal{F}^*\{s\} \\ &= \| \frac{\sqrt{\log \|\mathcal{F}\{s\}\|^2}}{\|\mathcal{F}\{s\}\|} \mathcal{F}\{s\} \|^2 \\ &= \|H\|^2 \end{aligned} \tag{4.5}$$

Προκύπτει :

$$\begin{aligned} \mathcal{C}\{s\} &= \|\mathcal{F}\left\{\|H\|^2\right\}\|^2 \\ &= \|\mathcal{F}\{H^* \cdot H\}\|^2 \\ &= \|h^*(-x,-y) \star h(x,y)\|^2 \\ &= \|h(x,y) \circ h(x,y)\|^2 \end{aligned} \tag{4.6}$$

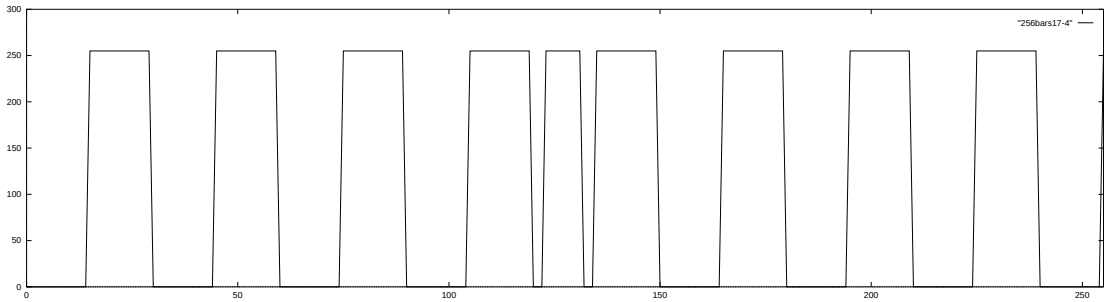
όπου $h = CEP \star s$, με h^* συμβολίζουμε το μιγαδικό συζυγή του h, με $\star$ τη συνέλιξη και με $\circ$ την συσχέτιση. Το δεξί μέλος στην εξίσωση 4.5 είναι το φάσμα ισχύος μιας φιλτραρισμένης μορφής της αρχικής συνάρτησης. Δηλ. το cepstrum είναι το φάσμα ισχύος της αυτοσυσχέτισης αφού προηγηθεί συνέλιξη του αρχικού σήματος μ' ένα μη γραμμικό φίλτρο προσαρμογής CEP. Αυτό το φίλτρο τείνει να κάνει το φάσμα ισχύος περισσότερο ομοιόμορφο, μειώνοντας την συμβολή των narrowband σημάτων και αφήνοντας τα broadband σήματα σχεδόν απείραχτα. Γι' αυτό, το φίλτρο κάνει την είσοδο λιγότερο διφορούμενη ως προς τη συσχέτιση (σχήμα 4.2).

Αν στον ορισμό του cepstrum στη θέση του λογαρίθμου είχαμε ρίζα τέταρτης τάξεως ή τόξο εφαπτομένης το αποτέλεσμα δεν θα είχε μεγάλη διαφορά [22].

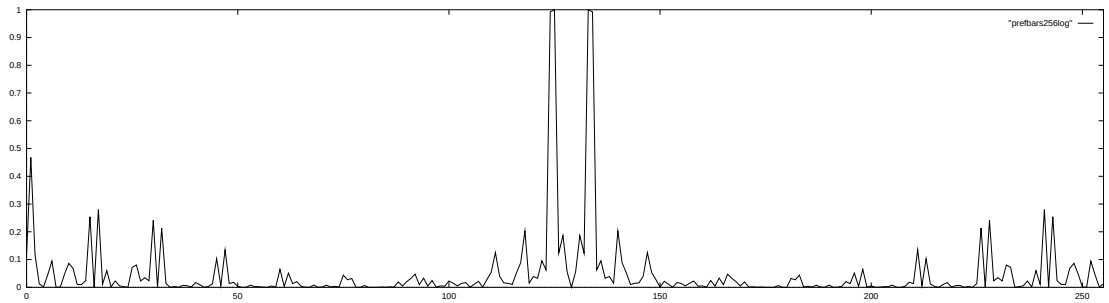
4.3 Υπολογισμός ανομοιότητας με cepstrum

4.3.1 Το cepstrum σήματος που περιέχει ηχώ

Δοσμένου ενός δισδιάστατου σήματος $s(x,y)$ ορίζουμε το $f(x,y) = s(x,y) + s(x-x_0, y-y_0)$ δηλαδή ένα σήμα που παράγεται από το άθροισμα του $s(x,y)$ με ένα αντίγραφο του



(α)



(β)

Σχήμα 4.1: Εφαρμογή του φίλτρου προσαρμογής CEP.

Το περιοδικό μέρος της εισόδου (α) το οποίο δίνει εσφαλμένες αντιστοιχήσεις με τις μεθόδους συσχέτισης έχει συμπιεστεί (β) με την εφαρμογή του φίλτρου προσαρμογής CEP.

μετατοπισμένο κατά (x_0, y_0).

Αποδεικνύεται [38] ότι :

$$\mathcal{C}\{f(x,y)\} = \mathcal{C}\{s(x,y)\} + \sum_{n=-\infty}^{\infty} \frac{\delta(x - n \cdot x_0, y - n \cdot y_0)}{n} \tag{4.7}$$

Το cepstrum του f, λοιπόν, είναι το άθροισμα του cepstrum του s και ενός περιοδικού όρου. Ο περιοδικός όρος στην παραπάνω εξίσωση θα έχει βασικές συχνότητες x_0 στην x κατεύθυνση και y_0 στην y κατεύθυνση, ενώ ο όρος που εξαρτάται από το σήμα s αποτελείται από χαμηλότερες συχνότητες στις περισσότερες καταστάσεις πραγματικών εικόνων. Είναι γενικά λοιπόν εφικτός ο προσδιορισμός της μετατόπισης (x_0, y_0) με ανίχνευση των κορυφών στη συνάρτηση του cepstrum (της περιοδικότητας του δεύτερου όρου). Στην επόμενη ενότητα θα δούμε πως το γεγονός αυτό χρησιμοποιήθηκε για την εκτίμηση αραιού πεδίου ανομοιότητας από το σήμα που λαμβάνεται από μια στερεοσκοπική διάταξη.

4.3.2 Ο αλγόριθμος υπολογισμού αραιού πεδίου ανομοιότητας των Yeshurun και Schwartz

θεωρώντας μια στερεοσκοπική διάταξη μπορούμε να προσεγγίσουμε το ζεύγος των εικόνων που λαμβάνεται σε μια δεδομένη χρονική στιγμή με το εξής απλό μοντέλο :

$$r(x,y) = l(x,y) \star \delta(x - d_h, y - d_v) \tag{4.8}$$

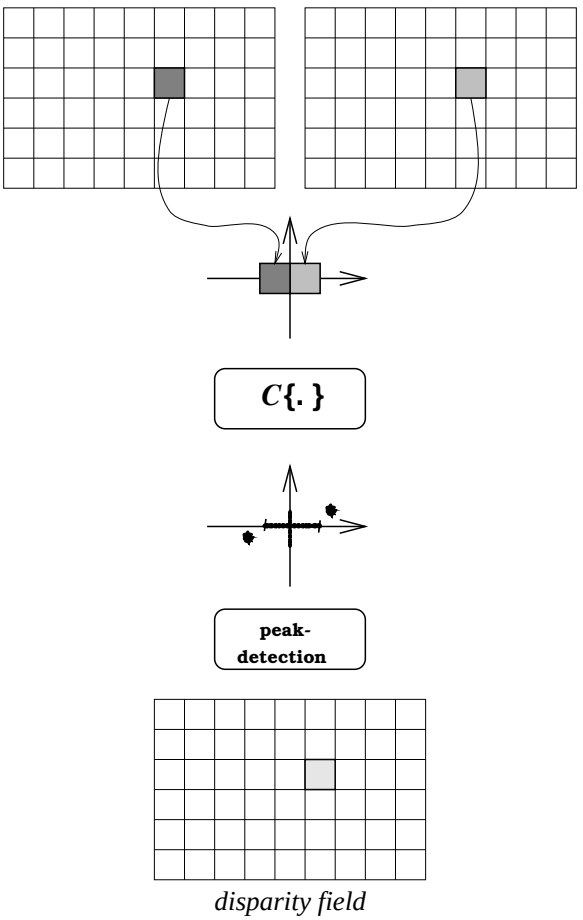
όπου $l(x,y)$ ένα ορθογώνιο κομμάτι διάστασης $(h \times w)$ της αριστερής εικόνας, $r(x,y)$ το αντίστοιχο κομμάτι (ίσης διάστασης) στην δεξιά εικόνα, και (d_h, d_v) το ζητούμενο διάνυσμα ανομοιότητας [2]. Εδώ θα συνδέσουμε το μοντέλο αυτό με την ιδέα ανίχνευσης της μετατόπισης που αναφέραμε παραπάνω. Αν συρράψουμε τα δύο υποσήματα $l(x,y)$ και $r(x,y)$ το ένα διπλά στο άλλο, δηλ. συνθέσουμε το σήμα :

$$f(x,y) = l(x,y) + r(x,y) \star \delta(x - w, y) \tag{4.9}$$

τότε μπορούμε να ανιχνεύσουμε το διάνυσμα (d_h, d_v) σύμφωνα με όσα ειπώθηκαν στην προηγούμενη ενότητα.

Αυτή η ιδέα διατυπώθηκε και υλοποιήθηκε σε μια παράλληλη αρχιτεκτονική από τους Yeshurun και Schwartz [75]. Μολονότι τα παραπάνω δεν αποτελούν ένα υπολογιστικό μοντέλο για την επεξεργασία της πληροφορίας στις ocular dominance columns, σαν ιδέα πηγάζει από την συγκεκριμένη έρευνα στη νευροβιολογία [75] (εναλλαγή των λωρίδων στον φλοιό <--> επεξεργασία σε concatenated αντίστοιχα παράθυρα των δύο εικόνων).

[2] για την εκφραστικότητα αυτού του μοντέλου αναφερθήκαμε στο προηγούμενο κεφάλαιο στην περιγραφή των μεθόδων συσχέτισης που στηρίζονται σ' αυτό.



Σχήμα 4.2: Βήματα στον υπολογισμό αραιού πεδίου τοπικής ανομοιότητας με cepstrum

Οι αρχικές εικόνες χωρίζονται σε παράθυρα κατάλληλου μεγέθους. Τα αντίστοιχα παράθυρα συρράπτονται το ένα δίπλα στο άλλο, και εφαρμόζεται ο cepstrum μετασχηματισμός $\mathcal{C}\{\cdot\}$. Η ανομοιότητα που αντιστοιχεί σ' αυτά τα παράθυρα υπολογίζεται με ανίχνευση κορυφών στο σήμα έξοδου του μετασχηματισμού και αποθηκεύεται στο πεδίο ανομοιότητας.

4.3.3 Ο αλγόριθμος υπολογισμού ανομοιότητας του στόχου σύγκλισης των Olson και Coombs

Οι Olson και Coombs [53], επιδιώκοντας να υπολογίσουν από ένα στερεοσκοπικό ζεύγος εικόνων την ανομοιότητα του στόχου σύγκλισης, εφάρμοσαν τον αλγόριθμο των Yeshurun και Schwartz σε κεντρικά υποδειγματοληπτημένα (32×32 από 256×256) παράθυρα των εικόνων στα οποία φροντίζει το σύστημα καταδίωξης να βρίσκεται ο στόχος.

Ο αλγόριθμος που υπολογίζει το cepstrum αποτελείται από έναν 2D FFT, ένα σημειακό μετασχηματισμό (τον λογάριθμο του φάσματος ισχύος) και ένα δεύτερο 2D FFT. Το κύριο σημείο καθυστέρησης αυτού του αλγορίθμου είναι επομένως ο FFT. Με στόχο την μείωση του υπολογιστικού κόστους των μονοδιάστατων 1D FFT, που αποτελούν το 2D FFT,προτείνονται στο [53] οι εξής βελτιώσεις:

- Αντιμετώπιση του wraparound στο πεδίο του cepstral μετασχηματισμού. Ένας διακριτός μετασχηματισμός Fourier (DFT) μεγέθους $h \times 2w$ χειρίζεται συχνότητες από $-\frac{h}{2}$ ως $\frac{h}{2}$ κάθετα και από $-w$ ως w οριζόντια. Στη συνεχή περίπτωση ο cepstral μετασχηματισμός παράγει κορυφές στα $(\pm(w + d_h), \pm d_v)$. Έτσι θετικές οριζόντιες ανομοιότητες θα δώσουν κορυφές που θα πέφτουν έξω από την κλίμακα του DFT. Ο DFT όμως είναι κυκλικός, οπότε οι θετικές ανομοιότητες θα περάσουν από την άλλη μεριά του οριζόντιου άξονα συχνοτήτων. Έτσι δεν θα είναι εφικτός ο διαχωρισμός των ανομοιοτήτων (d_h, d_v) και $(-d_h, -d_v)$. Η κλασική λύση σ' αυτό το πρόβλημα είναι η διεύρυνση της εισόδου με μηδενικά. Αυτό όμως έχει παραπάνω κόστος σε χρόνο και μνήμη. Μια πιο οικονομική λύση είναι η εξής : αφού η κάθετη ανομοιότητα στην προκειμένη περίπτωση είναι μικρή μπορούμε να υποθέσουμε ότι είναι στο διάστημα $\pm\frac{h}{4}$. Συνεπώς μπορούμε να ολισθήσουμε τη δεξιά εικόνα προς τα πάνω κατά $\frac{1}{4}$ του ύψους της (δηλ. να μετακινήσουμε την γραμμή m στην γραμμή $\left(m + \frac{h}{4}\right) mod\, h$. Μ' αυτόν τον τρόπο μια αρνητική κάθετη ανομοιότητα στην έξοδο δείχνει ότι η οριζόντια ανομοιότητα έχει περάσει από την άλλη μεριά του άξονα και επομένως έχει λάθος πρόσημο.

- Μείωση των μετασχηματισμών στηλών. Στον δεύτερο 2D FFT το σύνολο των 1D FFT στηλών χρειάζεται μόνο στην περιοχή που θα ψάξουμε για κορυφές ($\frac{w}{2}$ ως w). Δηλαδή δεν γίνονται $\frac{3}{4}$ των 1D FFT στηλών.

Θα μπορούσε στα παραπάνω να προστεθεί η χρήση του γρήγορου μετασχηματισμού Hartley (FHT) αντί για τον FFT. Η χρήση FHT θα είχε σαν αποτέλεσμα διπλάσια ταχύτητα στον υπολογισμό της εξόδου του φίλτρου όπως ισχυρίζονται οι Kai-Oliver Ludwig, Heiko Neumann και Bernd Neumann [38].

Η ακρίβεια της μεθόδου περιορίζεται από την υποδειγματοληψία. Για την αντιμετώπιση του προβλήματος αυτού, οι Olson και Coombs, μετά την ανίχνευση υψηλότερης κορυφής, εφαρμόζουν μονοδιάστατη παρεμβολή στη γραμμή στην οποία

έχει εντοπιστεί η κορυφή. Η μέθοδος παρεμβολής είναι :

$$x = \frac{peak\ x_{peak} + larger\ x_{larger}}{peak + larger} \tag{4.10}$$

όπου $peak$ είναι η τιμή της κορυφής, $larger$ είναι η μεγαλύτερη γειτονική της τιμή και x_{peak}, x_{larger} οι θέσεις στις οποίες βρίσκονται πάνω στη γραμμή όπου εφαρμόζεται η παρεμβολή Το x δηλαδή είναι το κέντρο βάρους της υψηλότερης κορυφής και του μεγαλύτερου γειτονικού της σημείου. Η υπόθεση στην οποία στηρίζεται η μέθοδος, είναι ότι η έξοδος του φίλτρου για ένα δείγμα, σαν συνάρτηση της ανομοιότητας, είναι η συνέλιξη ενός ορθογώνιου παλμού πλάτους ενός pixel και αγνώστου ύψους (παλμός δειγματοληψίας) με τον ορθογώνιο παλμό της ανομοιότητας. Δηλαδή, πρόκειται για έναν τριγωνικό παλμό πλάτους δύο pixels. Στην πραγματικότητα στη βάση του ο τριγωνικός παλμός είναι ελαφρώς πεπλατυσμένος (broadened), με αποτέλεσμα η μέθοδος παρεμβολής να δίνει ασυνέχειες στο σημείο τομής (crossing point) των καμπυλών [22]. Ένας τρόπος αποφυγής αυτών των ασυνεχειών είναι η συμμετοχή του μικρότερης γειτονικής τιμής $smaller$ κατά κάποιο ποσοστό στην παρεμβολή :

$$x = \frac{peak\ x_{peak} + larger\ x_{larger} + a\ smaller\ x_{smaller}}{peak + larger + a\ smaller} \tag{4.11}$$

όπου :

$$a = 1 - \frac{larger - smaller}{peak - smaller} \tag{4.12}$$

4.3.4 Σύγκριση του cepstral φίλτρου με τη συσχέτιση φάσης

Μια πιο άμεση προσέγγιση από το cepstral φίλτρο για τον υπολογισμό ανομοιότητας είναι εφικτή. Θεωρώντας πάλι ότι ένα τετραγωνικό τμήμα στη δεξιά εικόνα έχει προκύψει από το αντίστοιχο στην αριστερή με μετατόπιση κατά (d_h, d_v) (4.8), και εφαρμόζοντας μετασχηματισμό Fourier παίρνουμε:

$$e^{-j2\pi(ud_h - vd_v)} = \frac{R(u,v)}{L(u,v)} \tag{4.13}$$

ή

$$\delta\left(x - d_h, y - d_v\right) = \mathcal{F}^{-1}\left(\frac{R(u,v)L^*(u,v)}{\|L(u,v)\|^2}\right) \tag{4.14}$$

Αλλά το μέτρο των μετασχηματισμών των δύο εικόνων είναι το ίδιο, μόνο η φάση διαφέρει λόγω της μετατόπισης. Άρα :

$$\delta\left(x - d_h, y - d_v\right) = \mathcal{F}^{-1}\left(\frac{R(u,v)L^*(u,v)}{\|R(u,v)\|\|L^*(u,v)\|}\right) \tag{4.15}$$

Η παράσταση στην παρένθεση στην εξίσωση 4.15 είναι όμως ο μετασχηματισμός Fourier της κανονικοποιημένης ετεροσυσχέτισης *(cross-correlation)* έτσι ώστε να έχει απόλυτη τιμή < 1.

Αυτή η μέθοδος φαίνεται να είναι γρηγορότερη από το cepstral φίλτρο γιατί χρησιμοποιεί μετασχηματισμό Fourier μικρότερων παραθύρων και όπως είναι γνωστό το κόστος του FFT αυξάνει περισσότερο από γραμμικά όταν αυξάνει το παράθυρο. Ωστόσο το cepstral φίλτρο με την συγκεκριμένη αντιμετώπιση για το wraparound, που περιγράψαμε στην παράγραφο 4.2, είναι οικονομικότερο, ενώ η ίδια αντιμετώπιση δεν είναι κατάλληλη για την μέθοδο συσχέτισης φάσης [22].

4.4 Προβλήματα και περιορισμοί του cepstral φίλτρου

4.4.1 Το πεδίο ανομοιότητας

Οι Kai-Oliver Ludwig, Heiko Neumann και Bernd Neumann [38], μελετώντας το πεδίο ανομοιότητας που δημιουργείται από ένα επίπεδο τετραγωνικό πλέγμα και το πεδίο ανομοιότητας που δημιουργείται από μια σαγματική επιφάνεια (saddle surface), κατέληξαν στα εξής συμπεράσματα για την μορφή του πεδίου και κατά πως αυτό επηρεάζει τα αποτελέσματα του υπολογισμού ανομοιότητας με τον cepstral μετασχηματισμό :

- Η ανομοιότητα που υπολογίζεται από την έξοδο του cepstrum είναι η μέση τιμή όλων των διαφορετικών ανομοιοτήτων που βρίσκονται στο παράθυρο εφαρμογής του φίλτρου. Αυτό σημαίνει ότι, το φίλτρο είναι εφαρμόσιμο μόνο κοντά στο σημείο καθήλωσης όπου υπάρχει μικρή διακύμανση τιμών ανομοιότητας.

- Η μορφή του πεδίου ανομοιότητας, γενικά, δεν εξαρτάται από τη συγκεκριμένη δομή της επιφάνειας αλλά κυρίως από το slant και tilt (όρος πρώτης τάξης στην προσέγγιση με ανάπτυγμα Taylor) της επιφάνειας στο σημείο καθήλωσης.

- Η οριζόντια συνιστώσα του διανύσματος ανομοιότητας δεν εξαρτάται από την $y-$συντεταγμένη του τρισδιάστατου σημείου. Η κάθετη συνιστώσα εξαρτάται και από τις τρεις συντεταγμένες του σημείου. Στη γειτονιά του σημείου καθήλωσης κυριαρχεί η οριζόντια συνιστώσα.

- Η τοπική επίπεδη προσέγγιση μιας επιφάνειας, στην οποία έχει καθηλωθεί το σύστημα, δίνει μια απλή συνάρτηση μετατόπισης/ολίσθησης για την ανομοιότητα. Ο γεωμετρικός τόπος των σημείων με ανομοιότητα μηδέν είναι το σημείο καθήλωσης στην περίπτωση foveal ανάλυσης. Αυτό δεν ισχύει για non-foveal ανάλυση.

- Η διακύμανση του μέτρου ανομοιότητας μπορεί να θεωρηθεί αμελητέα μόνο για πολύ μικρή γωνία θέας, στον υπολογισμό με λωρίδες σταθερού πλάτους. Γι' αυτό το μέγεθος των παραθύρων που περιέχουν κομμάτια του σήματος από την αριστερή και δεξιά εικόνα δεν πρέπει να είναι πολύ μεγάλο, γιατί διαφορετικά οι τιμές της ανομοιότητας που υπολογίζονται με τη μέθοδο δεν είναι αξιόπιστες.

4.4.2 Θόρυβος, παραμόρφωση του σήματος, ευθύγραμμα τμήματα ακμών και μεταβλητός φωτισμός

Ο λόγος σήμα προς θόρυβο (SNR) δεν είναι επαρκές μέτρο για την απόδοση του φίλτρου στο θόρυβο. Το σχετικό εύρος (*bandwidth*) του σήματος και του θορύβου παίζει επίσης ρόλο [64, 63]. Παρά το γεγονός αυτό ένας πρακτικός κανόνας μπορεί να διατυπωθεί για πραγματικές εικόνες : η σωστή κορυφή της εξόδου μπορεί να ανιχνευθεί παρουσία λευκού προσθετικού θορύβου έως 100% του πλάτους του σήματος [38].

Γενικά για την προοπτική προβολή δεν ισχύει η υπόθεση της μεταφορικής ολίσθησης στο μετασχηματισμό μεταξύ των δύο υποσημάτων που εξάγονται από την αριστερή και δεξιά εικόνα. Οι Kai-Oliver Ludwig, Heiko Neumann και Bernd Neumann [38] προκειμένου να μοντελοποιήσουν τέτοιου είδους παραμορφώσεις του ενός από τα δύο υποσήματα εξέτασαν τις περιστροφές και τις αλλαγές στο μέγεθος και συνδυασμούς και των δύο. Ωστόσο, ο υπολογισμός της ανομοιότητας με το cepstrum παραμένει σωστός για περιστροφή $< 3^{\circ}$ ή μεγέθυνση/σμίκρυνση $< 4\%$ ενός από τα υποσήματα. Για την περίπτωση πραγματικών εικόνων αυτά τα όρια επεκτείνονται σε 6° και 6% αντίστοιχα [38]. Οι Yeshurun και Shwartz αναφέρουν ότι τα όρια αυτά είναι 10° και 15% αντίστοιχα αλλά αυτές οι τιμές φαίνονται να είναι οι διπλάσιες για την περίπτωση εικόνων με θόρυβο.

Το cepstrum, όπως κάθε εκτιμητής τοπικής ανομοιότητας, έχει να αντιμετωπίσει το aperture πρόβλημα. Εκτός απ' αυτό, ο αρχικός αλγόριθμος παρουσιάζει προβλήματα με γραμμικά τμήματα ακμών, που δίνουν πρόσθετες κορυφές στην έξοδο. Αν αυτές οι κορυφές είναι αρκετά ισχυρές για να διαταράξουν την εκτίμηση της ανομοιότητας εξαρτάται από την περιοχή εύρεσης που χρησιμοποιείται και πόσο το ιδανικό μοντέλο ταιριάζει στη συγκεκριμένη γραμμική δομή στην εικόνα. Για την αντιμετώπιση του προβλήματος αυτού μπορούν να χρησιμοποιηθούν διαφορετικές συναρτήσεις υποστήριξης (support functions) των παραθύρων (π.χ. Gaussian, Hamming) [38].

Οι Yeshurun και Schwartz ισχυρίζονται ότι ο αλγόριθμος δεν διαταράσσεται από διαφορές στην ένταση, ούτε από απλές προσθετικές αυξήσεις έντασης 50% στην μια εικόνα του ζεύγους. Ωστόσο η υλοποίηση του αλγορίθμου εισάγει προβλήματα. Διαφορετικά επίπεδα φωτισμού (ή αναστροφή του σήματος) σε κάποιο από τα υποσήματα μπορεί να διαταράξει μικρές ανομοιότητες όταν χρησιμοποιούνται gray level εικόνες, αφού μια πρόσθετη κορυφή μπορεί να αναδειχθεί σε μια θέση όπου είναι κωδικοποιημένη μηδενική ανομοιότητα. Μπορεί να μετατοπιστεί το σημείο αναφοράς για μηδενική ανομοιότητα από το $(D, 0)$ σ' ένα άλλο σημείο του cepstral επιπέδου για να αποφευχθούν οι πρόσθετες κορυφές [38].

Το cepstrum είναι συμβατό με κάθε λειτουργία προ-επεξεργασίας. Έτσι, αν εφαρμοστεί LoG (Laplacian of Gausssian) φιλτράρισμα στην είσοδο με μικρό $\sigma(0.35$ έως $0.71)$ βελτιώνει τα αποτελέσματα αφού το LoG αντιστοιχεί στον υπολογισμό της παραγώγου του εξομαλυμένου σήματος [38].

Όπως είπαμε στην ενότητα 4.2 το cepstrum μπορεί να θεωρηθεί ως η αυτοσυσχέτιση μιας φιλτραρισμένης έκδοσης της εικόνας. Η φιλτραρισμένη εικόνα δείχνει σαν να έχει υποστεί ζωνοδιαβατό φιλτράρισμα. Αν συγκρίνουμε τα αποτελέσματα του cepstrum με τα αντίστοιχα της αυτοσυσχέτισης εικόνων που έχουν υποστεί LoG φιλτράρισμα είναι καλύτερα [38].

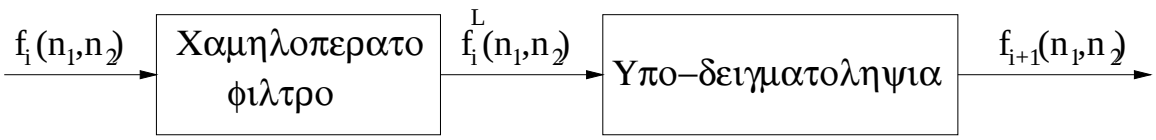
4.5 Αναζήτηση κορυφής με τεχνική πολλαπλών επιπέδων

Πολλές τεχνικές υπολογισμού ανομοιότητας χρησιμοποιούν κάποια μορφή ιεραρχικής ή coarse-to-fine στρατηγική ελέγχου για να βελτιώσουν τις εκτιμήσεις ανομοιότητας [50]. Οι coarser scales επιτρέπουν τα συστήματα να χειρίζονται μεγάλες ανομοιότητες αποτελεσματικά περιορίζοντας τον αριθμό των δυνατών συγκρίσεων ενώ οι finer scales παρέχουν υψηλότερη ανάλυση. Οι Taylor, Olson και Martin [66] προσάρμοσαν τον αλγόριθμο υπολογισμού ανομοιότητας του στόχου σύγκλισης που περιγράψαμε στην ενότητα 4.3.3 χρησιμοποιώντας coarse-to-fine στρατηγική, με σκοπό την αντιμετώπιση του προβλήματος των πολλαπλών ανομοιοτήτων. Ο εντοπισμός της ανομοιότητας με την τεχνική αυτή γίνεται ανθεκτικότερος με την εφαρμογή κριτηρίων που σχετίζονται με τη διαδικασία της σύγκλισης. Επιπλέον, επιτρέποντας υψηλότερη ανάλυση της εισόδου έχει καλύτερη ακρίβεια στον εντοπισμό της ανομοιότητας.

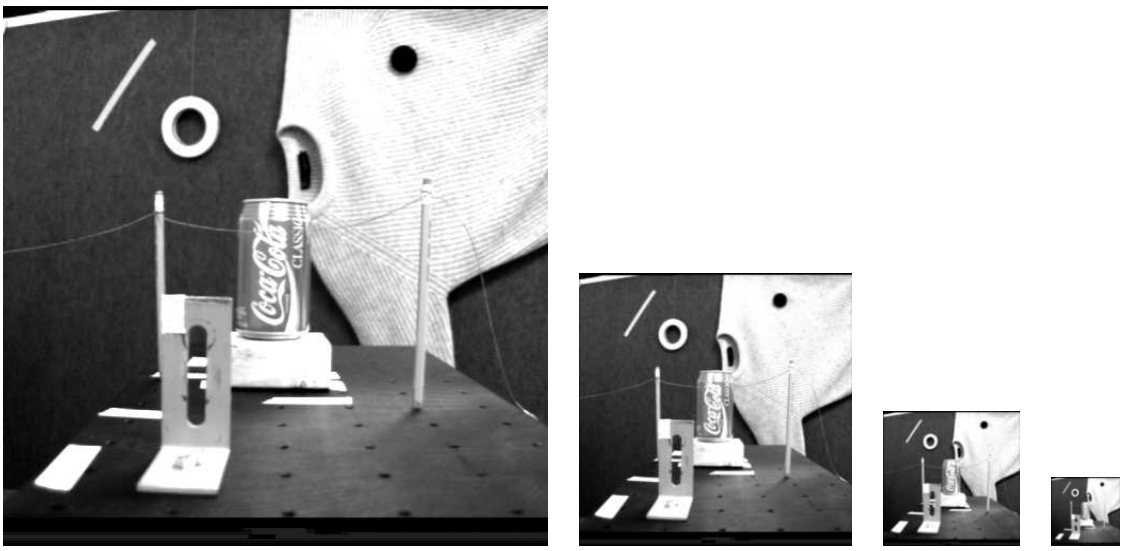
Η εφαρμογή της τεχνικής αυτής προυποθέτει ότι ο στόχος σύγκλισης προσδιορίζεται από τον προσανατολισμό της μιας κάμερας (π.χ αριστερή) η οποία ονομάζεται κύρια *(dominant)*. Η κύρια κάμερα καθοδηγείται από την διαδικασία της παρακολούθησης. Η άλλη κάμερα της στερεοσκοπικής διάταξης καλείται μη κύρια *(non-dominant)* και καθοδηγείται από τη σύγκλιση. Μια άλλη διαφορά από τον αλγόριθμο των Olson και Coombs είναι η χρήση Hamming συναρτήσεων υποστήριξης των παραθύρων έτσι ώστε να μην χάνεται πληροφορία κατα τη συρραφή των μονοφθαλμικών υποσημάτων.

Η *Gaussian* πυραμίδα (σχήμα 4.4) είναι μια σειρά από εκδόσεις μιας εικόνας με προοδευτικά μικρότερη ανάλυση και πυκνότητα δειγματοληψίας. Από μια εικόνα ομοιόμορφης ανάλυσης η Gaussian πυραμίδα μπορεί να κατασκευαστεί από διαδοχικό χαμηλοπερατό φιλτράρισμα και υποδειγματοληψία της εικόνας (σχήμα 4.3). Ο αλγόριθμος των Taylor-Olson-Martin χρησιμοποιεί την δομή της περικομμένης *(truncated)* Gaussian πυραμίδας (σχήμα 4.5), η οποία είναι μια στήλη από παράθυρα ίσου μεγέθους που έχουν ληφθεί από το κέντρο των αντίστοιχων επιπέδων της Gaussian πυραμίδας. Τα χαρακτηριστικά που εμφανίζονται στο κέντρο των coarser επιπέδων εμφανίζονται στα finer επίπεδα με μεγαλύτερη πυκνότητα δειγματοληψίας. Ωστόσο, τα coarser επίπεδα της δειγματοληπτούν ένα μεγαλύτερο πεδίο θέας, δηλ., τα χαρακτηριστικά που εμφανίζονται στα όρια ενός δοσμένου επιπέδου δεν θα εμφανίζονται σε finer επίπεδα. Η δομή αυτή μπορεί να θεωρηθεί ως προσέγγιση ενός foveated αισθητήρα.

Το cepstrum όπως κάθε εκτιμητής ανομοιότητας που βασίζεται στη συσχέτιση μπορεί να παράγει πολλαπλές κορυφές. Όσον αφορά τη διαδικασία της σύγκλισης είναι



Σχήμα 4.3: Παραγωγή του $i+1$ επιπέδου από το i επίπεδο στην αναπαράσταση της Gaussian πυραμίδας



Σχήμα 4.4: Παράδειγμα Gaussian πυραμίδας

Σχήμα 4.5: Παράδειγμα περικομμένης Gaussian πυραμίδας

επιθυμητή η αναγνώριση μόνο της κορυφής που οφείλεται στην ανομοιότητα του στόχου. Οι υπόλοιπες προέρχονται από άλλα αντικείμενα ή από λάθος αντιστοιχήσεις. Έτσι η επιλογή της υψηλότερης κορυφής όπως αναφέραμε παραπάνω δεν αντιστοιχεί σίγουρα στην ανομοιότητα του στόχου. Ο αλγόριθμος που προτείνεται από τους Taylor-Olson-Martin στηρίζεται στα εξής τρία κριτήρια:

1. Οι υψηλότερες κορυφές είναι περισσότερο πιθανές από τις μικρότερες.
2. Οι κορυφές με μεγάλες κάθετες ανομοιότητες δεν λαμβάνονται υπ' όψην αφού η γεωμετρία του συστήματος αναγκάζει τις επιπολικές γραμμές για τα χαρακτηριστικά του στόχου να είναι σχεδόν οριζόντιες. Έτσι οι μεγάλες κάθετες ανομοιότητες οφείλονται σε λάθος αντιστοιχήσεις.
3. Οι κορυφές πρέπει να προέρχονται από χαρακτηριστικά της εικόνας τα οποία βρίσκονται στην κεντρική περιοχή (όπου βρίσκεται ο στόχος) της κύριας κάμερας. Αυτό γίνεται με δύο τρόπους: α) coarse-to-fine αναζήτηση στην περικομμένη Gaussian πυραμίδα, β) προβολή από το cepstrum στην εικόνα.

Θα περιγράψουμε στην συνέχεια τους τρόπους α) και β) για την υλοποίηση του τρίτου κριτηρίου, και στη συνέχεια θα περιγράψουμε την λειτουργία του αλγορίθμου και ως προς τα τρία κριτήρια.

coarse-to-fine search

Έστω μια κορυφή στο επίπεδο k της πυραμίδας. Αν η θέση, (dx, dy), της κορυφής αντιστοιχεί σε μικρή ανομοιότητα και προήλθε από χαρακτηριστικά στο κέντρο της κύριας εικόνας, τότε τα ίδια χαρακτηριστικά θα υπάρχουν στο πεδίο θέας και των δύο καμερών στο επόμενο επίπεδο $k-1$ της πυραμίδας και η αντίστοιχη θέση της κορυφής θα είναι στην γειτονιά του $(2dx, 2dy)$. Αν, αντίθετα, η κορυφή βρίσκεται σε τέτοια θέση έτσι ώστε να αντιστοιχεί σε μεγάλη ανομοιότητα, τότε στο επίπεδο $k-1$ της μη κύριας κάμερας δεν περιέχεται ο στόχος. Αυτό σημαίνει ότι, για μεγάλες ανομοιότητες δεν μπορούμε να αποφανθούμε για το αν τα χαρακτηριστικά της εικόνας από τα οποία προήλθε η κορυφή βρίσκονται στο κέντρο της κύριας κάμερας στο επόμενο επίπεδο.

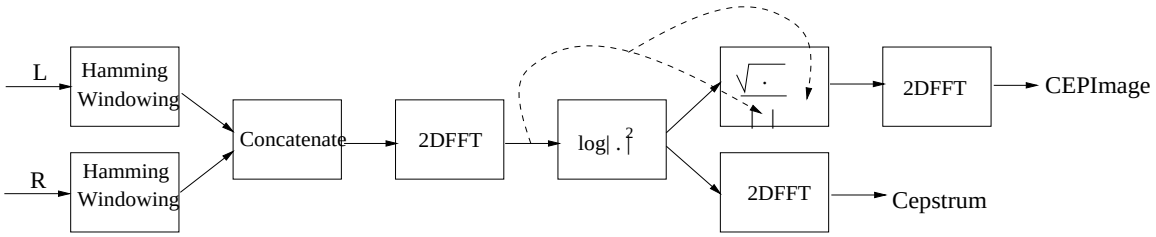
Προβολή από το cepstrum στην εικόνα

Το cepstrum σύμφωνα με όσα είπαμε στην παράγραφο 4.2 είναι φάσμα ισχύος της αυτοσυσχέτισης της συνελιγμένης εισόδου με το φίλτρο προσαρμογής, CEP, Η συνάρτηση μεταφοράς του CEP, που αναφέραμε στην παραγραφο 4.2 είναι :

$$\frac{\sqrt{\log |I|^2}}{|I|} \tag{4.16}$$

όπου I ο μετασχηματισμός Fourier του σήματος εισόδου.

Αν I′ είναι η έξοδος του CEP, w το πλάτος της εικόνας πριν την συρραφή και (dx, dy) η υποψήφια ανομοιότητα, εξετάζοντας το γινόμενο $I'(i,j) \times I'(i+w+dx, j+dy)$ ως προς ένα κατώφλι μπορούμε να αποφανθούμε για το αν το pixel (i,j) συνέβαλε στην



Σχήμα 4.6: Συνδυασμός του cepstrum με το φίλτρο προσαρμογής CEP

κορυφή που αντιστοιχεί στην ανομοιότητα (dx, dy) Το φίλτρο CEP μπορεί επίσης να χρησιμοποιηθεί για να καθορίσει αν μια συγκεκριμένη περιοχή στην εικόνα συνέβαλε σημαντικά σε μια κορυφή αυτοσυσχέτισης, για μια δοσμένη ανομοιότητα. Αυτό γίνεται με υπολογισμό της αυτοσυσχέτισης, CEV (Cepstrum Equivalent Value), της εξόδου του φίλτρου CEP, στην περιοχή αυτή:

$$CEV(dx, dy) = \sum_{i=mini}^{maxi} \sum_{j=minj}^{maxj} \frac{I'(i,j)I'(i+w+dx, j+dy)}{N} \tag{4.17}$$

όπου τα $mini, maxi, minj, maxj$ ορίζουν την περιοχή πάνω στην εικόνα.

Αν ο υπολογισμός του CEV γινόταν σε όλη την κύρια εικόνα για κάθε πιθανή ανομοιότητα θα προσέγγιζε την υπολογιστική πολυπλοκότητα του cepstral φίλτρου, μέσω ενός χωρικού υπολογισμού. Έτσι, είναι προτιμότερος ο υπολογισμός της CEV μόνο για μια κεντρική περιοχή της εξόδου του CEP. Επιπλέον η εφαρμογή του CEP μπορεί να συνδυαστεί με την εφαρμογή του cepstrum έτσι ώστε η έξοδος του πρώτου δισδιάστατου FFT να χρησιμοποιηθεί και από τα δύο φίλτρα (C και CEP) (σχήμα 4.6).

Αλγόριθμος

Ο αλγόριθμος σύγκλισης κάνει τις εξής υποθέσεις:

- Αν ο στόχος σύγκλισης απεικονίζεται σε οποιοδήποτε επίπεδο της μη κύριας κάμερας, τουλάχιστον μια κορυφή της συσχέτισης αντιστοιχεί στην ανομοιότητα μεταξύ των χαρακτηριστικών του στόχου στις δύο εικόνες.

- Όταν ο στόχος απεικονίζεται στο επίπεδο 0 της μη κύριας κάμερας, η υψηλότερη κορυφή σ' αυτό το επίπεδο αντιστοιχεί στο στόχο.

Η λειτουργία του αλγορίθμου εξαρτάται από το αν η κύρια κάμερα έχει κατευθυνθεί σε καινούργιο σημείο καθήλωσης ή όχι. Ανάλογα με το κριτήριο αυτό διακρίνονται δύο επίπεδα λειτουργίας : α) όταν η κύρια κάμερα κινηθεί σύμφωνα με τη διαδικασία της παρακολούθησης προσπαθεί να ανακτήσει τη σύγκλιση ψάχνοντας στην πυραμίδα την κορυφή που αντιστοιχεί στην ανομοιότητα του στόχου. (acquisition mode) β) όταν η

κύρια κάμερα παραμένει ακίνητη προσπαθεί να διατηρήσει την σύγκλιση λαμβάνοντας υπ' όψη μόνο το επίπεδο 0 (finer) της διοφθαλμικής πυραμίδας. (fineadjustment mode)

Κατά την διατήρηση της σύγκλισης εφαρμόζεται το cepstral φίλτρο όπως έχει προταθεί από τους Olson και Coombs στο επίπεδο υψηλότερης ανάλυσης. Κατά την ανάκτηση της σύγκλισης, το cepstral φίλτρο εφαρμόζεται διαδοχικά από το επίπεδο χαμηλότερης ανάλυσης (3) ως το επίπεδο υψηλότερης ανάλυσης (0). Ανιχνεύονται όλες οι κορυφές στην έξοδο του φίλτρου και εξετάζονται με φθίνουσα σειρά είτε με coarse-to-fine αναζήτηση, είτε με προβολή στην εικόνα αν δεν είναι εφαρμόσιμη η coarse-to-fine αναζήτηση. Αν μια κορυφή επιβεβαιωθεί με την coarse-to-fine στρατηγική και από το επίπεδο 0 τότε το σύστημα σύγκλισης κινεί την μη κύρια κάμερα σύμφωνα με την ανομοιότητα που έχει υπολογιστεί μ' αυτόν τον τρόπο. Αν όμως, η coarse-to-fine στρατηγική δεν ήταν δυνατό να εφαρμοστεί σε κάποιο ενδιάμεσο επίπεδο, υπολογίζεται η CEV τιμή και εξετάζεται ως προς ένα ελάχιστο κατώφλι. Αν είναι μεγαλύτερη απ' αυτό το κατώφλι τότε το σύστημα σύγκλισης κατευθύνει την μη κύρια κάμερα σύμφωνα με την ανομοιότητα που εξετάστηκε και θέτει μια τοπική μεταβλητή $Current$ ίση με το επίπεδο στο οποίο σταμάτησε η αναζήτηση. Αυτό γίνεται για οικονομία στους υπολογισμούς. Υπάρχει περίπτωση, μια συγκεκριμένη κορυφή να μην επιβεβαιωθεί ούτε με την coarse-to-fine στρατηγική, ούτε με την στρατηγική της προβολής της στο επίπεδο της εικόνας. Σ' αυτή την περίπτωση εξετάζονται οι υπόλοιπες κορυφές που έχουν ανιχνευθεί στο επίπεδο χαμηλότερης ανάλυσης. Αν καμία απ' αυτές δεν εγκριθεί σύμφωνα με τον παραπάνω συλλογισμό, τότε το σύστημα σύγκλισης δεν εκτελεί καμία κίνηση και θέτει το $Current$ ίσο με 3, δηλώνοντας αποτυχία.

Τα παραπάνω περιγράφονται με τον εξής ψευδοκώδικα:

Ανάκτηση σύγκλισης με την τεχνική πολλαπλών επιπέδων

Αν ισχύει $Current <> 0$ κάνε τα εξής :
$TruncatedGaussianPyramid(L, Current)$
$TruncatedGaussianPyramid(R, Current)$
$Cepstrum_{Current} \nwarrow$
$C\text{-}CEP(L_{Current}, R_{Current})$
$CEPImage_{Current} \swarrow$
$Peaks_{Current} = FindSortPeaksIn(Cepstrum_{Current})$
/* Αναζήτηση */
$CurPeak = 0$
Όσο ισχύει $CurPeak < MaxNumberOfPeaks$ κάνε τα εξής
Ονόμασε (dx, dy) την θέση του $Peaks_{Current}(CurPeak)$ στο $Cepstrum$
$state = FAILURE$
$Level = Current$
Όσο ισχύει ότι $Level > 0$ κάνε τα εξής

Αν δεν έχουν υπολογιστει τα $Cepstrum_{Level-1}$ και $CEPImage_{Level-1}$

$Cepstrum_{Level-1}$ $\searrow$

$C\text{-}CEP(L_{Level-1}, R_{Level-1})$

$CEPImage_{Level-1}$ $\swarrow$

αν $dx < HorizontalThreshold$

αν υπάρχει κορυφή στο $Cepstrum_{Level-1}$

στη γειτονιά του σημείου $(2dx, 2dy)$

$disparity \approx (2dx, 2dy)$

$dx \approx 2dx$

$dy \approx 2dy$

αλλιώς

$CEV = ComputeCEV((dx, dy), CEPImage_{Current})$

αν $CEV > CEVThreshold$

$disparity = (dx, dy)$

state = SUCCESS

break

$Level = Level - 1$

Αν $Level == 0$ ή $state == SUCCESS$

break

αλλιώς $CurPeak = CurPeak + 1$

Αν $CurPeak <> MaxNumberOfPeaks$

pan την μη κύρια κάμερα σύμφωνα με την τιμή του $disparity$

Current = Level

Κεφάλαιο 5

Φίλτρα μηδενικής ανομοιότητας (ΦΜΑ)

Σ' αυτό το κεφάλαιο θα περιγράψουμε διάφορα φίλτρα μηδενικής ανομοιότητας τα οποία συμμετέχουν στο σύστημα σταθεροποίησης της προσήλωσης για την εκτίμηση του λάθους παρακολούθησης πάνω στην οροπτερική επιφάνεια.

Ένα φίλτρο ανομοιότητας *(disparity filter)* έχει σαν στόχο την εύρεση των σημείων της σκηνής που έχουν μια δοσμένη στερεοσκοπική ανομοιότητα ή έχουν ανομοιότητα μέσα σε μια δοσμένη γκάμα. Η διαφορά του από την κανονική στερεοσκοπική επεξεργασία (π.χ. κατασκευή πεδίου ανομοιοτήτων), είναι ότι δεν γίνεται εξαντλητική αναζήτηση στο χώρο των δυνατών τιμών ανομοιότητας για κάθε pixel, αλλά κάθε pixel εξετάζεται ως προς την συμβατότητα του με μια δοσμένη ανομοιότητα. Ωστόσο, και στις δύο περιπτώσεις χρειάζεται κάποιο είδος σύγκρισης *(matching)* για την αντιμετώπιση του προβλήματος αντιστοίχισης *(correspondence problem)*.

Αν θεωρήσουμε τις εικόνες ως αντιστοιχήσεις από σημεία εικόνας σε ένα χώρο από χαρακτηριστικά (τιμές gray-scale, πρόσημο ακμής (+,-,none), προσανατολισμός ακμής, γωνίες, gray scale patches) τότε μπορούμε να ορίσουμε το φίλτρο, DF_d, ως :

$$\begin{aligned} DF_d(L,R) &= f \\ \text{όπου } f(x,y) &= \begin{cases} 1, & \text{αν } L(x,y) \approx R(x+d,y) \\ 0, & \text{αλλιώς} \end{cases} \end{aligned} \tag{5.1}$$

όπου $\approx$ τελεστής ομοιότητας ως προς τα χαρακτηριστικά της αντιστοίχισης.

Το φίλτρο για μια περιοχή από ανομοιότητες ορίζεται σαν το λογικό OR των εξόδων ενός συνόλου φίλτρων DF_d :

$$DF_{[a,b]}(L,R) = \bigvee_{d \in [a,b]} DF_d(L,R) \tag{5.2}$$

Το πλεονέκτημα των φίλτρων ανομοιότητας σε σχέση με άλλου είδους στερεοσκοπική επεξεργασία είναι το ότι είναι εξαιρετικά αποδοτικά και έχουν υψηλό βαθμό παραλληλοποίησης. Μειονεκτεί ωστόσο σε ανθεκτικότητα στο θόρυβο. Οι πηγές θορύβου είναι : η συμπτωματική ευθυγράμμιση (accidental alignement) σημείων στις δύο εικόνες τα οποία αντιστοιχούν σε διαφορετικά σημεία της σκηνής (π.χ. κάθετες άσπρες μπάρες σε μαύρο υπόστρωμα), και η ύπαρξη σημειών χωρίς υφή που βρίσκονται στο ίδιο βάθος. Το τελευταίο αποτελεί πρόβλημα για οποιοδήποτε στερεοσκοπικό σύστημα, και αντιμετωπίζεται, σε περισσότερο πολύπλοκα συστήματα, με παρεμβολή στο πεδίο βάθους. Το φίλτρο με είσοδο μόνο τις δύο εικόνες περιορίζεται στην εφαρμογή τοπικής, χαμηλού επιπέδου σύγκρισης. Ο αριθμός των λάθος καταφάσεων, που προκύπτουν από τη σύγκριση (συμπτωματική ευθυγράμμιση) στερεοσκοπικού ζεύγους πραγματικών εικόνων, ελαττώνεται όσο μεγαλύτερη είναι η περιοχή τοπικής επεξεργασίας και όσο περισσότερο εκλεπτυσμένο είναι το κριτήριο σύγκρισης.

Η παρουσία υφής σε μια περιοχή στον τρισδιάστατο χώρο διεγείρει το φίλτρο. Το μέγεθος και το σχήμα της περιοχής εξαρτάται από τις παραμέτρους των καμερών. Για επίπεδη προβολή, κάμερες να κοιτάνε μπροστά και μικρό ως μεσαίο πεδίο θέας, είναι μια σχετικά επίπεδη επιφάνεια μπροστά από τις κάμερες σε βάθος το οποίο καθορίζεται από την ανομοιότητα. Για την ακρίβεια, λόγω της πεπερασμένης ανάλυσης των κάμερων, δεν πρόκειται για επιφάνεια αλλά για φλοιό.

Ένα φίλτρο ανομοιότητας είναι χρήσιμο στην τμηματοποίηση *(segmentation)* και στην ανίχνευση εγγύτητας, δηλ. στο αν υπάρχουν αντικείμενα κοντά στον παρατηρητή. Τμηματοποίηση μπορεί να επιτευχθεί με μια συγκλίνουσα στερεοσκοπική διάταξη και το αποτέλεσμα είναι μια πρώτη προσέγγιση της σιλουέτας του αντικειμένου. Ανίχνευση εγγύτητας μπορεί να γίνει με thresholding της μέσης τιμής (DC component) της εξόδου του φίλτρου (η μέση τιμή εκφράζει την ποσότητα υφής στην περιοχή που είναι ρυθμισμένο το φίλτρο).

Ειδικότερα, ένα φίλτρο μηδενικής ανομοιότητας *(zero disparity filter)* - στην συνέχεια θα αναφερόμαστε στον όρο αυτό με την συντομογραφία ΦΜΑ - έχει σάν είσοδο δύο εικόνες και η έξοδος του είναι ένας πίνακας από σημεία τα οποία εκφράζουν την γνώμη του φίλτρου για το αν δύο σύστοιχα σημεία των δύο εικόνων είναι πράγματι σημεία μηδενικής ανομοιότητας.

5.1 ΦΜΑ σε ένα επίπεδο ανάλυσης

Σ' αυτή την ενότητα θα περιγράψουμε διάφορα ΦΜΑ που έχουν προταθεί στη βιβλιογραφία [21, 20, 22, 14] και εφαρμόζονται σε μια συγκεκριμένη ανάλυση της στερεοσκοπικής εισόδου.

-1	1
-1	1

-1	0	1
-2	0	2
-1	0	1

0	-1	1	0
-1	-2	2	1
-1	-2	2	1
0	-1	1	0

0	-1	0	1	0
-1	-2	0	2	1
-2	-4	0	4	2
-1	-2	0	2	1
0	-1	0	1	0

Σχήμα 5.1: Πυρήνες Sobel κάθετων ακμών ($2 \times 2,\ 3 \times 3,\ 4 \times 4,\ 5 \times 5$)

5.1.1 Απλό ΦΜΑ συσχέτισης κάθετων ακμών

Ένα χαρακτηριστικό το οποίο φέρει πληροφορία για την οριζόντια ανομοιότητα είναι οι κάθετες ακμές. Αντίθετα οι οριζόντιες ακμές δεν παρέχουν χρήσιμη πληροφορία γιατί μεγάλα τμήματα τους μπορούν να συμπέσουν ακόμα και αν η ανομοιότητα τους είναι μεγάλη. Το ΦΜΑ συσχέτισης κάθετων ακμών [21, 20, 22, 14], λοιπόν, χρησιμοποίει ως χαρακτηριστικά για τη σύγκριση των δύο εικόνων που λαμβάνουμε από τη συγκλίνουσα στερεοσκοπική διάταξη τις κάθετες ακμές. Οι κάθετες ακμές μπορούν να ανιχνευθούν με συνέλιξη του οπτικού σήματος με ένα φίλτρο ανίχνευσης κάθετων ακμών (π.χ Sobel (σχήμα 5.1). Αν συμβολίσουμε με E_L και E_R το αποτέλεσμα της συνέλιξης με την αριστερή (L) και δεξιά εικόνα (R) αντίστοιχα η έξοδος του ΦΜΑ ορίζεται σύμφωνα με το κριτήριο :

$$Z(i,j) = \begin{cases} \frac{a+b}{2}, & \text{αν } \ |a-b| < D \times 2^B \ \text{ και } \ \left|\frac{a+b}{2}\right| > S \times 2^B \\ 0, & \text{αλλιώς} \end{cases} \tag{5.3}$$

όπου $a = E_L(i,j)$, $b = E_R(i,j)$, $D =$ κατώφλι % διαφοράς (η μέγιστη διαφορά έντασης των ακμών), $S =$ κατώφλι % ισχύος (ελάχιστη % ισχύ ακμών που λαμβάνονται υπ' όψη), B $=$ αριθμός bits κωδικοποίησης των εικόνων ακμών. Το φίλτρο, δηλαδή, έχει μη μηδενική έξοδο για ακμές που εμφανίζονται στην ίδια θέση στις δύο εικόνες E_L και E_R με σχεδόν το ίδιο μέτρο και έχουν κάποια ελάχιστη ένταση έτσι ώστε να μην πρόκειται για θόρυβο.

Η μέθοδος αυτή παρουσιάζει τα εξής προβλήματα :

- Μεγάλη πιθανότητα αλληλοεπικάλυψης *(aliasing)* των ακμών.
- Οι ακμές δεν είναι πάντα το κατάλληλο χαρακτηριστικό για σύγκριση σε μια συγκεκριμένη ανάλυση. Εξάλλου επειδή ο ανιχνευτής ακμών είναι γραμμικός

είναι πιθανότερο να δώσει ψηλότερη έξοδο σε περιοχές μεγάλης αντίθεσης της εικόνας.

- Η επιλογή κατωφλιού είναι μία μη ακριβής μέθοδος να αφαιρεθεί ο θόρυβος και μπορεί να επηρεάσει τα τελικά αποτελέσματα θετικά ή αρνητικά, ανάλογα με τις αρχικές εικόνες.

5.1.2 ΦΜΑ συσχέτισης γειτονιών

Τα ΦΜΑ συσχετίσης γειτονιών [21, 22] είναι επέκταση του απλού ΦΜΑ που περιγράψαμε στην προηγούμενη υποενότητα. Μ' αυτά εξαλείφεται το πρόβλημα επιλογής του κατωφλιού και η σύγκριση των εικόνων του στερεοσκοπικού ζεύγους στηρίζεται σε πιο σύνθετα κριτήρια.

Δύο απαραίτητες έννοιες, για τον ορισμό των φίλτρων αυτών είναι η εικόνα διαφορών *(difference image)*, D, και η εικόνα ποιότητας των χαρακτηριστικών *(feature quality image)*, Q. Η εικόνα διαφορών εκφράζει πόσο μοιάζουν αντίστοιχα τμήματα των στερεοσκοπικών εικόνων. Η εικόνα ποιότητας των χαρακτηριστικών αναπαριστά πόσο αξιόπιστη είναι η διαφορά αυτή ως έκφραση ομοιότητας. Οι υπολογισμοί για την κατασκευή των D,Q γίνονται σε τμήματα, $P_{i,j}$, μεγέθους $m \times m$ των εικόνων ($P_{i,j}(k,l) = I(i+k, j+l)$, $k,l : 1,..,m$. Οι υπολογισμοί αυτοί χρησιμοποιούν μη γραμμικές πράξεις για την εξαγωγή της D και γραμμικές και μη, πράξεις για την εξαγωγή της Q. Τα αποτελέσματα αυτών των υπολογισμών συνδυάζονται μη γραμμικά για να δώσουν την τελική έξοδο. Η έξοδος δεν υπολογίζεται αν το τμήμα δεν βρίσκεται ολόκληρο μέσα στην εικόνα. Γι' αυτό το λόγο, οι D και Q έχουν μικρότερο μέγεθος κατά $m-1$ σε κάθε διάσταση από τις εικόνες εισόδου.

Όταν κατασκευαστούν οι εικόνες D, Q, κανονικοποιούνται στην κλίμακα $[0, 2^{\mathrm{B}}-1]$ και δίνουν την εικόνα μηδενικής ανομοιότητας, Z :

$$Z(i,j) = (2^{\mathrm{B}} - 1 - D(i,j))Q(i,j) \tag{5.4}$$

Όσο μεγαλύτερη είναι η διαφορά D τόσο περισσότερο διαφέρουν οι εικόνες. Η ποσότητα $2^{\mathrm{B}} - 1 - D(i,j)$ παίρνει μεγάλες τιμές και σε περιοχές του υποστρώματος. Με την ποσότητα Q οι μεγάλες αυτές τιμές δεν περνάνε στην έξοδο του φίλτρου. Το απλό ΦΜΑ είναι ειδική περίπτωση των ΦΜΑ συσχέτισης γειτονιών αν το μέγεθος της γειτονιάς θεωρηθεί 1×1 και η ποιότητα των χαρακτηριστικών οριστεί κατάλληλα.

5.1.2.1 Εικόνα διαφορών

Μια απλή μέθοδος για την κατασκευή της εικόνας διαφορών είναι η γνωστή μέθοδος του αθροίσματος απόλυτων διαφορών *(sum of absolute differences (SAD))*.

$$D(i,j) = \sum_{k=1}^{m}\sum_{l=1}^{m} |L(k+i, l+j) - R(k+i, l+j)| \tag{5.5}$$

Ένα πρόβλημα της μεθόδου είναι ότι τα σημεία των εικόνων που η φωτεινότητα τους διαφέρει ελαφρά (λόγω φωτισμού) δεν θα αναδειχτούν στην εικόνα διαφορών. Για να ξεπεράσουμε το πρόβλημα αυτό μπορούμε να υπολογίσουμε την μέση τιμή κάθε τμήματος και στη συνέχεια να την αφαιρέσουμε από την τιμή κάθε pixel που ανήκει στο τμήμα.

5.1.2.2 Εικόνα ποιότητας των χαρακτηριστικών

Η εικόνα ποιότητας των χαρακτηριστικών καθορίζει αν τα τμήματα που έχουμε πάρει για τον υπολογισμό της εικόνας διαφορών είναι πιθανό να περιέχουν χαρακτηριστικά τα οποία κάνουν τη σύγκριση αξιόπιστη η όχι. Η τιμή της ποιότητας του χαρακτηριστικού είναι ανάλογη της αξιοπιστίας (ένα patch με μικρή διακύμανση φωτεινότητας δεν μπορεί να δώσει αξιόπιστη ομοιότητα). Γενικά :

$$Q(L,R) = \min\{Q(L), Q(R)\} \tag{5.6}$$

Οι συναρτήσεις Q εκφράζουν οριζόντια κλίση. Παίρνοντας το ελάχιστο εξασφαλίζουμε ότι έχουμε κάθετη ακμή σε σύστοιχα παράθυρα και στις δύο εικόνες που πιθανόν να προέρχονται από το ίδιο αντικείμενο.

Θα δούμε τρία μέτρα ποιότητας των χαρακτηριστικών:

- Οριζόντιες διαφορές

$$Q_p(P_{i,j}) = \sum_{k=1}^{m}\sum_{l=1}^{m} P_{i,j} \star \boxed{\text{-1 1}} \tag{5.7}$$

  όπου με $\star$ συμβολίζουμε την συνέλιξη.

- Οριζόντια τυπική απόκλιση (horizontal standard deviation)
  Το κριτήριο αυτό έχει προκύψει από την παραδοχή ότι η οριζόντια τυπική απόκλιση υποδεικνύει την παρουσία κάθετων στοιχείων.

$$Q_{sd}(P_{i,j}) = \sum_{k=1}^{m} \sqrt{\frac{\sum_{l=1}^{m} P_{i,j}(k,l)^2 - m\overline{x_i}^2}{m-1}} \tag{5.8}$$

$$\text{όπου } \overline{x_i} = \frac{\sum_{l=1}^{m} P_{ij}(i,l)}{m}$$

- Παρουσία ακμών διάφορων scales (edgeness)
  Το κριτήριο αυτό εκφράζει την παρουσία ακμής ανεξαρτήτου scale σε μια δοσμένη θέση. Για την ανίχνευση ακμών μπορούμε να χρησιμοποιήσουμε πυρήνες Sobel για ευκολία.

$$Q_e(P_{i,j}) = \sqrt{\sum_{n=2}^{5} |Sobel_{n\times n} \star P_{i,j}|} \tag{5.9}$$

Προκειμένου να εξοικονομήσουμε υπολογιστική ισχύ που δαπανούν οι πολλαπλές συνελίξεις στο παραπάνω κριτήριο αναιρούμε την μη γραμμικότητα του :

$$Q_e(P_{i,j}) = \sqrt{\left(\sum_{n=2}^{5} Sobel_{n\times n}\right) \star P_{i,j}} \tag{5.10}$$

Ένα δίλλημα που έχουμε να αντιμετωπίσουμε στην εφαρμογή του κριτηρίου αυτού είναι αν η συνέλιξη με όλους τους πυρήνες θα προχωρήσει μέχρι το τελευταίο σημείο του $P_{i,j}$. Σ' αυτή την περίπτωση θα έχουμε αλληλοκάλυψη μεταξύ διαφορετικών παραθύρων της εικόνας. Ανάλογα με τον τρόπο ευθυγράμμισης των πυρήνων Sobel ορίζονται δύο καινούριοι πυρήνες [22]:

-- Η ολική μάσκα *(total mask)* είναι το αποτέλεσμα του αθροίσματος τεσσάρων πυρήνων Sobel κάθετων ακμών (2×2 ως 5×5) (σχήμα 5.1) ευθυγραμμισμένων στην πάνω αριστερή γωνία (σχήμα 5.2 (α)).

-- Η κεντρική μάσκα *(center mask)* είναι το αποτέλεσμα του αθροίσματος τεσσάρων πυρήνων Sobel κάθετων ακμών (2×2 ως 5×5) ευθυγραμμισμένων στο κέντρο της μεγαλύτερης (σχήμα 5.2 (β)). Η κεντρική μάσκα δεν συνελίσσεται

-2	-1	2	1	0
-5	-3	4	3	1
-4	-6	3	5	2
-1	-3	1	2	1
0	-1	0	1	0

(α)

0	-2	1	1	0
-2	-6	3	4	1
-3	-9	3	7	2
-1	-4	1	3	1
0	-1	0	1	0

(β)

Σχήμα 5.2: (α) ολική μάσκα (total mask), (β) κεντρική μάσκα (center mask).

με το παράθυρο αλλά πολλαπλασιάζεται με τις αντίστοιχες τιμές τοποθετώντας την στο κέντρο του.

5.1.3 ΦΜΑ συσχέτισης μέτρου-κατεύθυνσης ακμών

Ένα πρόβλημα στο απλό ΦΜΑ είναι ότι ακμές διαφορετικής κατεύθυνσης μπορεί να θεωρηθούν όμοιες στο σημείο τομής τους. Η ανάπτυξη μιας μεθόδου που λαμβάνει υπόψη της την κατεύθυνση των ακμών περιγράφεται εδώ για να αποφευχθούν τέτοιου

είδους λάθος καταφάσεις του φίλτρου. Η μέθοδος προτείνεται από τους David Coombs, Ian Horswill και Peter von Kaenel [21].

Το φίλτρο συσχέτισης μέτρου-κατεύθυνσης ακμών εφαρμόζει συνέλιξη πυρήνων Sobel κάθετων και οριζόντιων ακμών στις δύο εικόνες και στη συνέχεια σχηματίζει την εικόνα μέτρου και την εικόνα κατεύθυνσης για κάθε εικόνα ως εξής :

$$M(i,j) = \sqrt{E_0^2(i,j) + E_{90}^2(i,j)} \tag{5.11}$$

$$O(i,j) = \arctan\left(\frac{E_0(i,j)}{E_{90}(i,j)}\right) \tag{5.12}$$

όπου E_0 και E_{90} είναι οι οριζόντιες και κάθετες ακμές αντίστοιχα.

Κατά αναλογία με τα ΦΜΑ συσχέτισης γειτονιών, για την αποφυγή χρήσης κατωφλιών, οι M_L, O_L, M_R, O_R συνδυάζονται για να δώσουν μια εικόνα διαφορών και μια εικόνα χαρακτηριστικών ποιότητας:

$$D(i,j) = \sqrt{(O_L(i,j) - O_R(i,j) + 1) \times (M_L(i,j) - M_R(i,j) + 1)} \tag{5.13}$$

$$Q(i,j) = \begin{cases} \min\left(M_L(i,j), M_R(i,j)\right) & (|O_L(i,j)| \wedge |O_R(i,j)| \in [45°, 135°]) \\ 0 & \text{αλλιώς} \end{cases} \tag{5.14}$$

ή

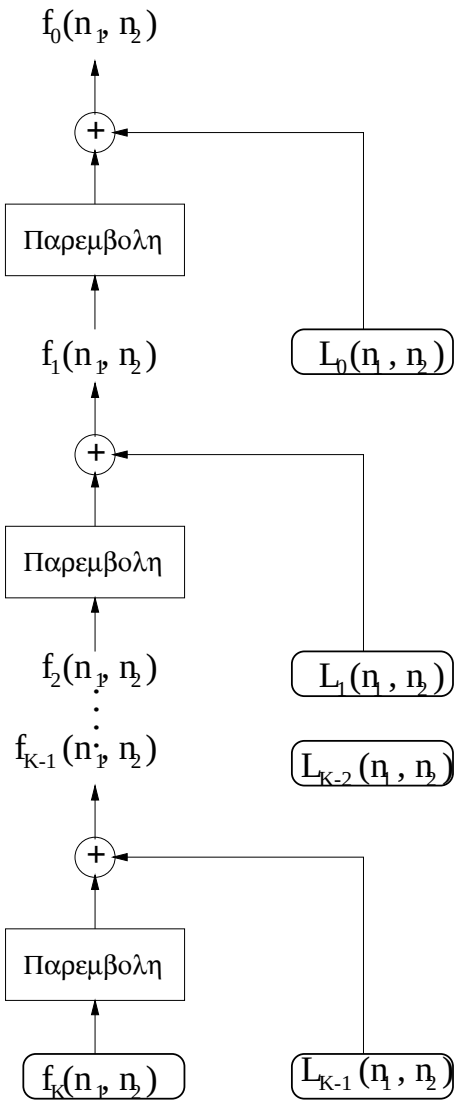
$$Q(i,j) = \min\left(\sin|O_L| \times M_L, \sin|O_R| \times M_R\right) \tag{5.15}$$

Σύμφωνα με την σχέση 5.14 οι μικρές γωνίες κατεύθυνσης (σχεδόν οριζόντιες ακμές) δεν λαμβάνονται υπόψη ενώ σύμφωνα με τη σχέση 5.15 λαμβάνονται υπόψη με τη μορφή μικρότερης τιμής στην εικόνα ποιότητας των χαρακτηριστικών.

5.2 ΦΜΑ ζωνοδιαβατής πυραμίδας

Σ' αυτή την ενότητα θα περιγράψουμε ένα ΦΜΑ που έχει προταθεί από τον Lockwood και Olson [54]. Η υπολογιστική δομή του στηρίζεται στην Λαπλασιανή πυραμίδα (Laplacian pyramid). Η ειδοποιός διαφορά του φίλτρου αυτού από τα προηγούμενα είναι ότι δεν λειτουργεί σε μία μόνο χωρική ανάλυση και η έξοδος του δεν είναι κάθετες ακμές αλλά η εικόνα του αντικειμένου στο οποίο συγκλίνει η στερεοσκοπική διάταξη.

Η Λαπλασιανή πυραμίδα αποτελείται από μια σειρά υποδειγματοληπτημένων ζωνοδιαβατών εικόνων με εκθετικά φθίνουσα ανάλυση (σχήμα 5.3). Το πρώτο επίπεδο της πυραμίδας L_0 κατασκευάζεται με blurring, υποδειγματοληψία της αρχικής εικόνας f_0, διαστολή της προσέγγισης στο αρχικό μέγεθος της με (Gaussian) παρεμβολή και αφαίρεση από την αρχική. Η L_0 δηλ. είναι το σφάλμα της προσέγγισης $n/2 \times n/2$ (η πληροφορία στην αρχική εικόνα που δεν υπάρχει στην προσέγγιση). Κατά την κατασκευή της πυραμίδας, η διαδικασία επαναλαμβάνεται ξεκινώντας από την προσέγγιση $n/2 \times n/2$. Το τελικό επίπεδο είναι μια προσέγγιση 1×1 της αρχικής εικόνας (μέση τιμή της εικόνας).



Σχήμα 5.3: Παραγωγή της Λαπλασιανής πυραμίδας.

Η αρχική εικόνα $f_0(n_1, n_2)$ μπορεί να ανακατασκευαστεί από το $L_i(n_1, n_2)$ για $0 \leq i \leq K-1$ και την $f_K(n_1, n_2)$

Η διαδικασία της επιλεκτικής απόσβεσης μερών της πυραμίδας, μοντελοποιείται με πολλαπλασιασμό pixel με pixel της πυραμίδας με μια πυραμίδα εμπιστοσύνης (confidence pyramid). Η τελευταία είναι ισομορφική ως προς την πυραμίδα της εικόνας, αλλά τα pixels της παίρνουν πραγματικές τιμές από 0 εως 1. Οι τιμές αυτές αναπαριστούν το μέτρο εμπιστοσύνης ως προς το αν τα αντίστοιχα σημεία της σκηνής βρίσκονται σε μια περιοχή βαθών κοντά στο σημείο καθήλωσης (πάνω στον στόχο καθήλωσης πιο συγκεκριμένα). Ο πολλαπλασιασμός των δύο πυραμίδων δίνει σαν αποτέλεσμα μια πυραμίδα στην οποία σημεία μη μηδενικής ανομοιότητας σε συγκεκριμένη κλίμακα έχει συμπιεστεί. Η ανίχνευση περιοχών μικρής αξιοπιστίας είναι ευκολότερη σε υψηλή χωρική ανάλυση. Γι' αυτό η ανακατασκευή της σταθμισμένης (weighted) πυραμίδας παράγει μια εικόνα στην οποία οι περιοχές μεγάλης ανομοιότητας έχουν υποστεί χαμηλοπερατό φιλτράρισμα.

Το σημαντικότερο σημείο και στο σχεδιασμό του ΦΜΑ ζωνοδιαβατής πυραμίδας είναι η επιλογή του μέτρου σύγκρισης. Το πιο απλό μέτρο είναι (κανονικοποιημένη διαφορά αντίστοιχων σημείων):

$$C_k(i,j) = 1 - \frac{|L_k(i,j) - R_k(i,j)|}{|L_k(i,j)| + |R_k(i,j)|} \tag{5.16}$$

όπου με k συμβολίζουμε το επίπεδο της πυραμίδας.

Ένα περισσότερο ανθεκτικό μέτρο σύγκρισης είναι (ανίχνευση κορυφών στην κανονικοποιημένη συσχέτιση 3x3 γειτονιών οριζόντιας πόλωσης (offset) d :

$$C_{k,d}(i,j) = \frac{\sum_{l=-1}^{1}\sum_{m=-1}^{1} L_k(i+l,j+m)R_k(i+l,j+m+d)}{\sqrt{\sum_{l=-1}^{1}\sum_{m=-1}^{1}\left(L_k(i+l,j+m)\right)^2}\sqrt{\sum_{l=-1}^{1}\sum_{m=-1}^{1}\left(R_k(i+l,j+m+d)\right)^2}} \tag{5.17}$$

Με βάση αυτό το μέτρο η έξοδος του ΦΜΑ μπορεί να οριστεί ίση με 1 αν και μόνο αν, $C_{k,0}(i,j) > sqrt2/2$ και $C_{k,0}(i,j) > C_{k,-1}(i,j)$ και $C_{k,0}(i,j) > C_{k,1}(i,j)$. Το πρώτο κατώφλι εξασφαλίζει ότι η γωνία μεταξύ των τοπικών γειτονιών (ως διανύσματα) είναι μικρότερη από 45°. Τα άλλα κατώφλια εξασφαλίζουν ότι πρόκειται για τοπικό μέγιστο.

Κεφάλαιο 6

Προσομοίωση και αποτελέσματα

6.1 Προσομοίωση ενεργού στερεοσκοπικής κεφαλής

Αναπτύχθηκε ένας προσομοιωτής ρομποτικής όρασης όπου το ρόλο της κάμερας παίζει ένας raytracer (Rayshade [34]). Ο στόχος της προσομοιώσης ήταν η αξιολόγηση της ακρίβειας των αλγορίθμων αφού σε κάθε χρονική στιγμή είναι γνωστή η θέση του αντικειμένου-στόχου. Ο προσομοιωτής κατασκευάζει μια γεωμετρική περιγραφή της στερεοσκοπικής κεφαλής με 4 βαθμούς ελευθερίας. Οι σκηνές που κατασκευάστηκαν για τα πειραμάτα περιέχουν απλά γεωμετρικά αντικείμενα και δεν περιέχουν θόρυβο, δηλαδή απέχουν αρκετά από πραγματικές συνθήκες. Στόχος όμως ήταν να δούμε πως εξαρτάται η συμπεριφορά του συστήματος από την ακρίβεια των αλγορίθμων και τα γεωμετρικά σφάλματα και όχι η συνολική απόδοση που εξαρτάται σημαντικά από τη δυναμική του συστήματος και τον έλεγχο.

Η προσομοίωση, όπως υλοποιήθηκε, μπορεί να επεκταθεί για να υποστηρίξει και άλλα συστήματα ρομποτικής και κατ' επέκταση και περισσότερο σύνθετες συμπεριφορές πλοήγησης (πχ περισσότεροι βαθμοι ελευθερίας).

6.2 Αποτελέσματα σύγκλισης

Θα δούμε σ' αυτή την ενότητα αποτελέσματα από τις δύο τεχνικές σύγκλισης που περιγράψαμε στο κεφάλαιο 4. Και, οι δύο τεχνικές όπως είδαμε στηρίζονται στον cepstral μετασχηματισμό. Η πρώτη τεχνική έχει μικρότερη ακρίβεια λόγω της χαμηλής ανάλυσης στην οποία εφαρμόζεται το φίλτρο (32x32). Αντίθετα η δεύτερη τεχνική αφού

ανακτήσει τη σύγκλιση έχει πολύ καλύτερη ακρίβεια αφού ο cepstral μετασχηματισμός εφαρμόζεται στην αρχική ανάλυση εισόδου. Το μεγαλύτερο πλεονέκτημα της δεύτερης τεχνικής είναι ότι το κριτήριο ανίχνευσης της κορυφής είναι περισσότερο ποιοτικό. Στη πρώτη τεχνική γίνεται η υπόθεση ότι το μεγαλύτερο τοπικό μέγιστο οφείλεται στο αντικείμενο, η οποία είναι πολύ περιοριστική. Το μειονέκτημα της δεύτερης τεχνικής είναι η μεγαλύτερη υπολογιστική πολυπλοκότητα λόγω της ανάλυσης πολλών επιπέδων.

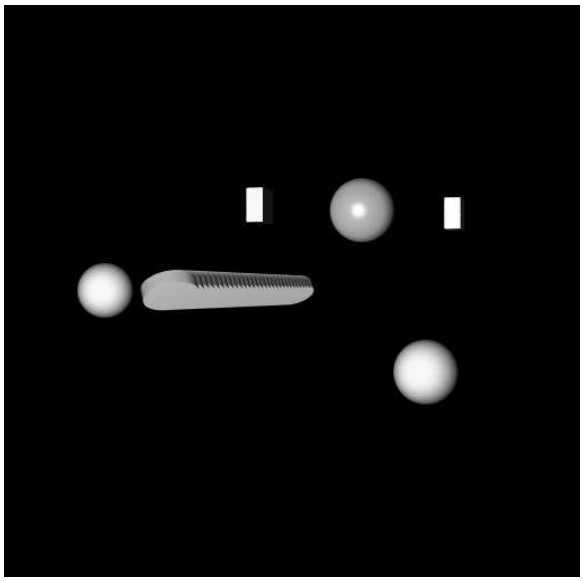
Στο σχήμα 6.1 βλέπουμε μια απλή είσοδο με την οποία συγκρίθηκαν οι δυο αλγόριθμοι. Για να απομονωθεί η σύγκλιση από τη παρακολούθηση η κίνηση του κυλινδρικού αντικειμένου γίνεται κατά μήκος του οπτικού άξονα της κύριας (αριστερής) κάμερας. Οι κάμερες έχουν πεδίο θέας 30^o και αρχικά είναι παράλληλες. Στο σχήμα 6.2 βλέπουμε την ανομοιότητα του κινούμενου αντικειμένου κατά την εξέλιξη του πειράματος. Παρατηρούμε ότι με τη πρώτη τεχνική δεν επιτυγχάνεται ποτέ σύγκλιση λόγω της παρουσίας των άλλων αντικειμένων και παραμένει στην αστάθεια. Αντίθετα με την τεχνική πολλών επιπέδων χρειάζεται κάποια αρχικά στιγμιότυπα για να ανακτήσει τη σύγκλιση την οποία διατηρεί μέχρι το τέλος του πειράματος. Η ανομοιότητα του κινούμενου αντικειμένου, όταν έχει επιτευχθεί η σύγκλιση, είναι της τάξης του ενός pixel της εικόνας εισόδου και οφείλεται στη χωρική κβαντοποίηση. Η ανομοιότητα παραμένει σ' αυτά τα επίπεδα μέχρι το τέλος του πειράματος αφού το επίπεδο ανάλυσης παραμένει σ' αυτό της εικόνας εισόδου. Η παρουσία των άλλων αντικειμένων δεν επηρεάζει την ακρίβεια στη σύγκλιση αλλά το χρόνο για να επιτευχθεί αυτή.

Στα σχήματα 6.3 και 6.4 συγκρίνεται η ακρίβεια σύγκλισης των δύο τεχνικών. Για να γίνει αυτό το προηγούμενο πείραμα έγινε απουσία ξένων αντικειμένων. Η ανομοιότητα του κινούμενου αντικειμένου για τις δύο τεχνικές κυμαίνεται στο επίπεδο του ενός pixel στην ανάλυση που λειτουργούν. Η πρώτη τεχνική εφαρμόζεται σε ανάλυση 8 φορές μικρότερη από αυτήν της εισόδου μ' αποτέλεσμα η ακρίβεια εντοπισμού της ανομοιότητας να είναι της τάξης των 8 pixels.

6.3 Αποτελέσματα παρακολούθησης

Το σύστημα σταθεροποίησης της οπτικής προσήλωσης που περιγράψαμε στο κεφάλαιο 3 στηρίζεται στην τμηματοποίηση του στόχου από τα γύρω αντικείμενα με τα ΦΜΑ, δηλαδή την ανίχνευση περιοχών στις εικόνες με μηδενική ανομοιότητα. Τα φίλτρα αυτά καθορίζουν την συμπεριφορά του συστήματος σταθεροποίησης. Αφ' ενός η εξάρτηση τους από την σωστή σύγκλιση πάνω στο στόχο και αφετέρου η χρησιμότητα της στερεοσκοπικής ανομοιότητας στην τμηματοποίηση μιας σκηνής, είναι δύο σημαντικοί παράγοντες για τη σωστή παρακολούθηση.

Στο σχήμα 6.5 βλέπουμε τη σκηνή που κατασκευάστηκε για τα πειράματα. Όπως βλέπουμε ένα κυλινδρικό αντικείμενο εκτελεί μια παραβολική κίνηση πλησιάζοντας το κεφάλι και την αριστερή κάμερα. Το αρχικό βάθος του κυλινδρικού αντικειμένου είναι 100 και το τελικό 70 ενώ το μέγιστο ύψος είναι 30. Η βασική γραμμή έχει μήκος 30 ο



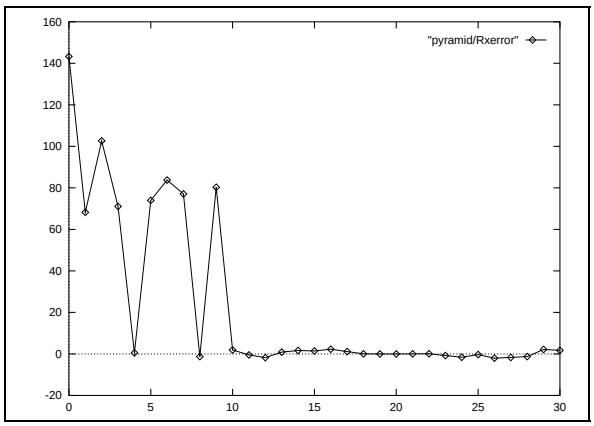
Σχήμα 6.1: Πείραμα σύγκλισης

Το κυλινδρικό αντικείμενο κινείται κατά μήκος του άξονα θέας της αριστερής κάμερας προς την κάμερα. Στην παραπάνω εικόνα φαίνονται οι διαδοχικές θέσεις του αντικειμένου από το σημείο τομής των αξόνων του συστήματος συντεταγμένων του κεφαλιού

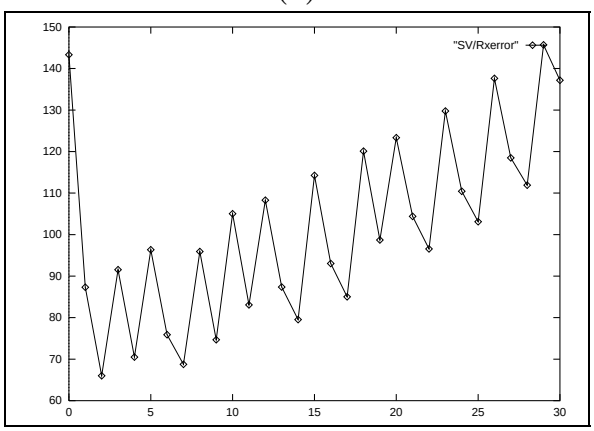
tilt βραχίονας 10 και οι κάμερες έχουν πεδίο θεας 30°. Τα ακίνητα αντικείμενα είναι σε βάθος 60 (κύβος) και 70 (σφαίρα).

Σε όλα τα αποτελέσματα παρακάτω (σχήμα 6.6 έως 6.11) απεικονίζεται η τροχιά του σημείου καθήλωσης όπως αυτό φαίνεται από την αριστερή (σχήματα (α), (β)) και δεξιά (σχήματα (γ),(δ)) κάμερα αν αυτή παρέμενε ακίνητη. Για τα αποτελέσματα της αριστερής στήλης έχει εφαρμοστεί το πρώτο σενάριο κίνησης, δηλ. χρησιμοποιούνται μόνο οι ανεξάρτητοι pan κινητήρες των καμερών και ο tilt κινητήρας του κεφαλιού, ενώ στη δεξιά στήλη βλεπουμε τα αντίστοιχα αποτελέσματα για το σενάριο κίνησης που χρησιμοποιεί και τους 4 βαθμούς ελευθερίας του κεφαλιού (οι ανεξάρτητοι pan κινητήρες των καμερών χρησιμοποιούνται για τη σύγκλιση ενώ ο pan κινητήρας του κεφαλιού για την παρακολούθηση, αναλυτικά τα σενάρια κατανομής σφαλμάτων στους βαθμούς ελευθερίας αναφέρθηκαν στη παράγραφο 3.6.3).

Στο σχήμα 6.6 βλέπουμε την συμπεριφορά του συστήματος σταθεροποίησης κατά την εφαρμογή του απλού ΦΜΑ συσχέτισης κάθετων ακμών. Με το δεύτερο σενάριο κίνησης επιτυγχάνεται καλύτερη παρακολούθηση. Αυτό οφείλεται στο γεγονός ότι το σύστημα είναι συμμετρικό οπότε η αντιστοιχία της ανομοιότητας στο λάθος της γωνίας σύγκλισης είναι σωστή. Αντίθετα στο πρώτο σενάριο όσο το αντικείμενο πλησιάζει την αριστερή κάμερα η υπόθεση συμμετρικής σύγκλισης παραβιάζεται οπότε η ανομοιοτητα που υπολογίζεται δεν αντιστοιχεί ακριβώς στο λάθος της γωνίας σύγκλισης. Έτσι στο πρώτο σενάριο στα τελευταία στιγμιότυπα χάνεται η σύγκλιση με αποτέλεσμα να αποτύχει



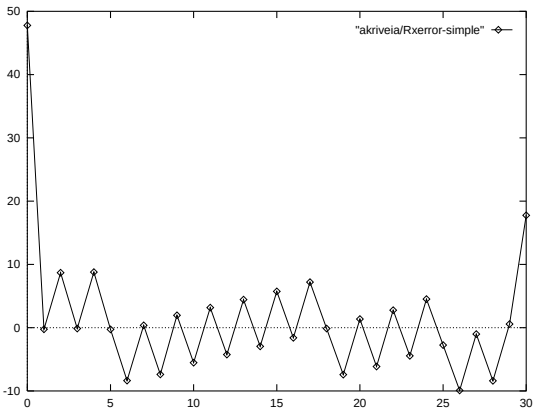
(α)



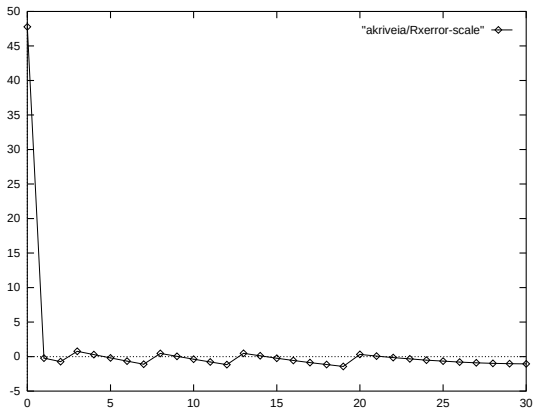
(γ)

Σχήμα 6.2: Η ανομοιότητα του κυλινδρικού αντικειμένου κατά το πείραμα σύγκλισης

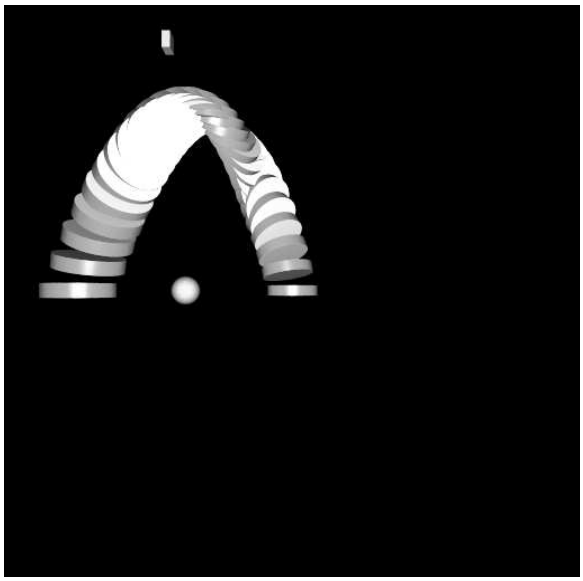
Στο σχήμα (α) έχει εφαρμοστεί ο αλγόριθμος αναζήτησης της κορυφής στην έξοδο του cepstrum με την τεχνική πολλαπλών επιπέδων, ενώ στο (β) επιλέγεται απλά η μέγιστη κορυφή. Στο (α) η ανομοιότητα μηδενίζεται σχεδόν μετά από 10 στιγμιότυπα και η τιμή της παραμένει κοντά στο 0 μέχρι το τέλος του πειράματος. Αντίθετα στο (β) το σύστημα δεν επιλέγει ποτέ την σωστή κορυφή στο cepstrum.



Σχήμα 6.3: Σύγκλιση χωρίς άλλα αντικείμενα στη σκηνή με εντοπισμό της μέγιστης κορυφής στο cepstrum



Σχήμα 6.4: Σύγκλιση χωρίς άλλα αντικείμενα στη σκηνή με αναζήτηση της κορυφής που αντιστοιχεί στο στόχο με την τεχνική πολλαπλών επιπέδων



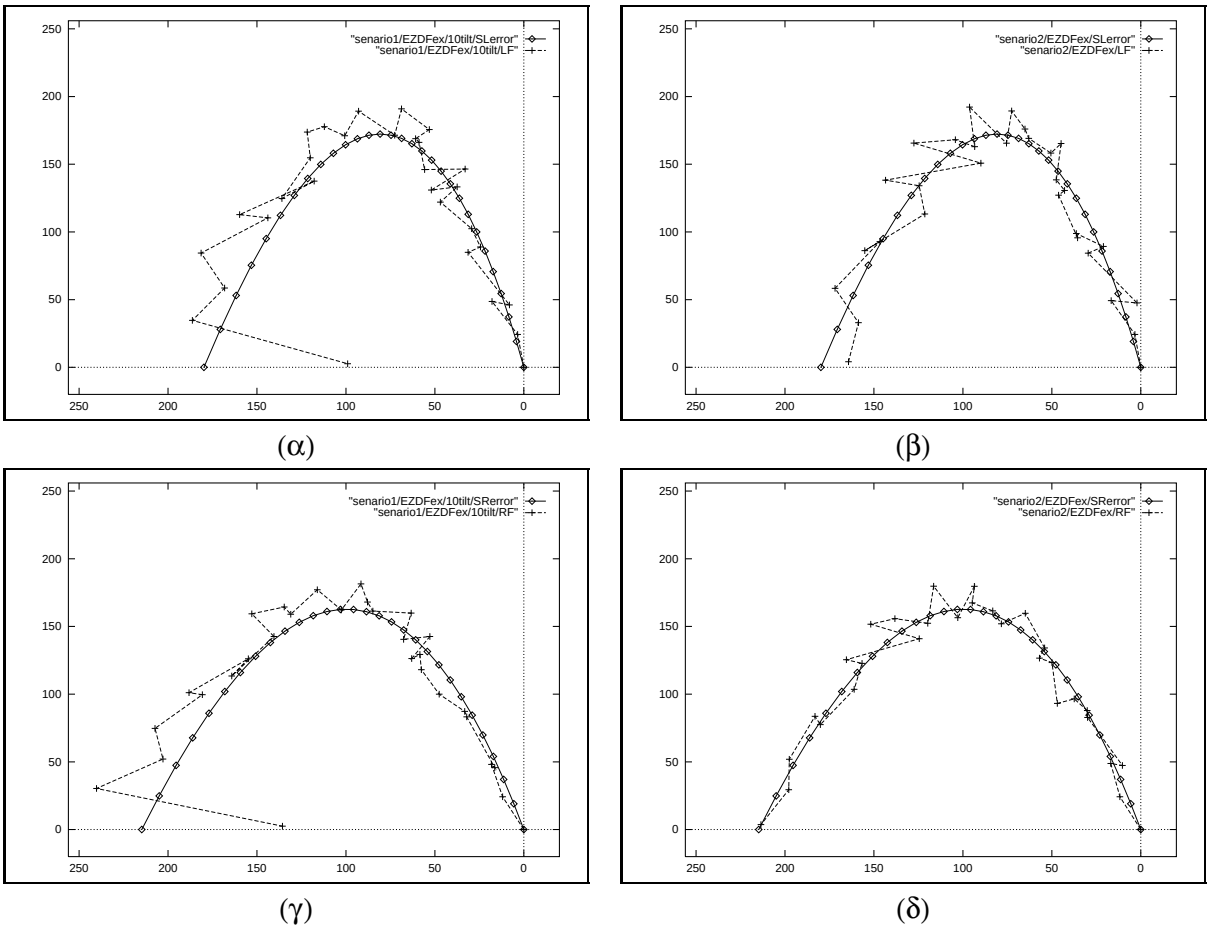
Σχήμα 6.5: Πείραμα παρακολούθησης

Το κυλινδρικό αντικείμενο εκτελεί μια παραβολική μη ομαλή κίνηση ενώ περιστρέφεται γύρω από έναν άξονα που περνάει από το κέντρο του. Στην παραπάνω εικόνα φαίνονται οι διαδοχικές θέσεις του αντικειμένου από το σημείο τομής των αξόνων του συστήματος συντεταγμένων του κεφαλιού

και η παρακολούθηση. Από τη συμπεριφορά αυτή φαίνεται ότι η σωστή απόκριση της παρακολούθησης εξαρτάται από τη σωστή απόκριση της σύγκλισης. Επειδή η σύγκλιση έχει χαμηλή ακρίβεια αυτό έχει σαν αποτέλεσμα να ταλαντώνεται η παρακολούθηση. Αυτό είναι εμφανές και στά επόμενα πειράματα.

Τα αποτελέσματα που παρουσιάζονται στο σχήμα 6.7 έχουν προκύψει εφαρμόζοντας το ΦΜΑ συσχέτισης γειτονιάς που χρησιμοποιεί ως μέτρο ποιότητας των χαρακτηριστικών τις οριζόντιες διαφορές των εντάσεων. Όπως είναι εμφανές η συμπεριφορά του συστήματος είναι χειρότερη από αυτήν στο προηγούμενο πείραμα. Αυτό οφείλεται στη κακή κατασκευή της εικόνας ποιότητας των χαρακτηριστικών. Η προσέγγιση της πρώτης παραγώγου με γειτονικές διαφορές δεν ειναι καλό μέτρο ποιότητας για την αντιστοίχιση των χαρακτηριστικών των δύο εικόνων. Όπως αναφέραμε και στο προηγούμενο πείραμα η εξάρτηση των μηχανισμών σύγκλισης και παρακολούθησης είναι άμεση. Έτσι στο δεύτερο σενάριο η κακή συμπεριφορά της μεθόδου οδηγεί σε αποτυχία της σύγκλισης το οποίο έχει σαν αποτέλεσμα το αντικείμενο να χάνεται τελείως από το οπτικό πεδίο.

Η χρήση της οριζόντιας τυπικής απόκλισης (σχήμα 6.8) σαν μέτρο ποιότητας χαρακτηριστικών έχει πολύ καλύτερα αποτελέσματα από ότι η χρήση οριζόντιων διαφορών. Η συμπεριφορά του συστήματος πλησιάζει περισσότερο αυτή στην οποία χρησιμοποιείται το απλό ΦΜΑ. Η συμπεριφορά του συστήματος δεν παρουσιάζει σημαντικές διαφορές όταν χρησιμοποιηθεί το ΦΜΑ συσχέτισης γειτονιάς με μέτρο

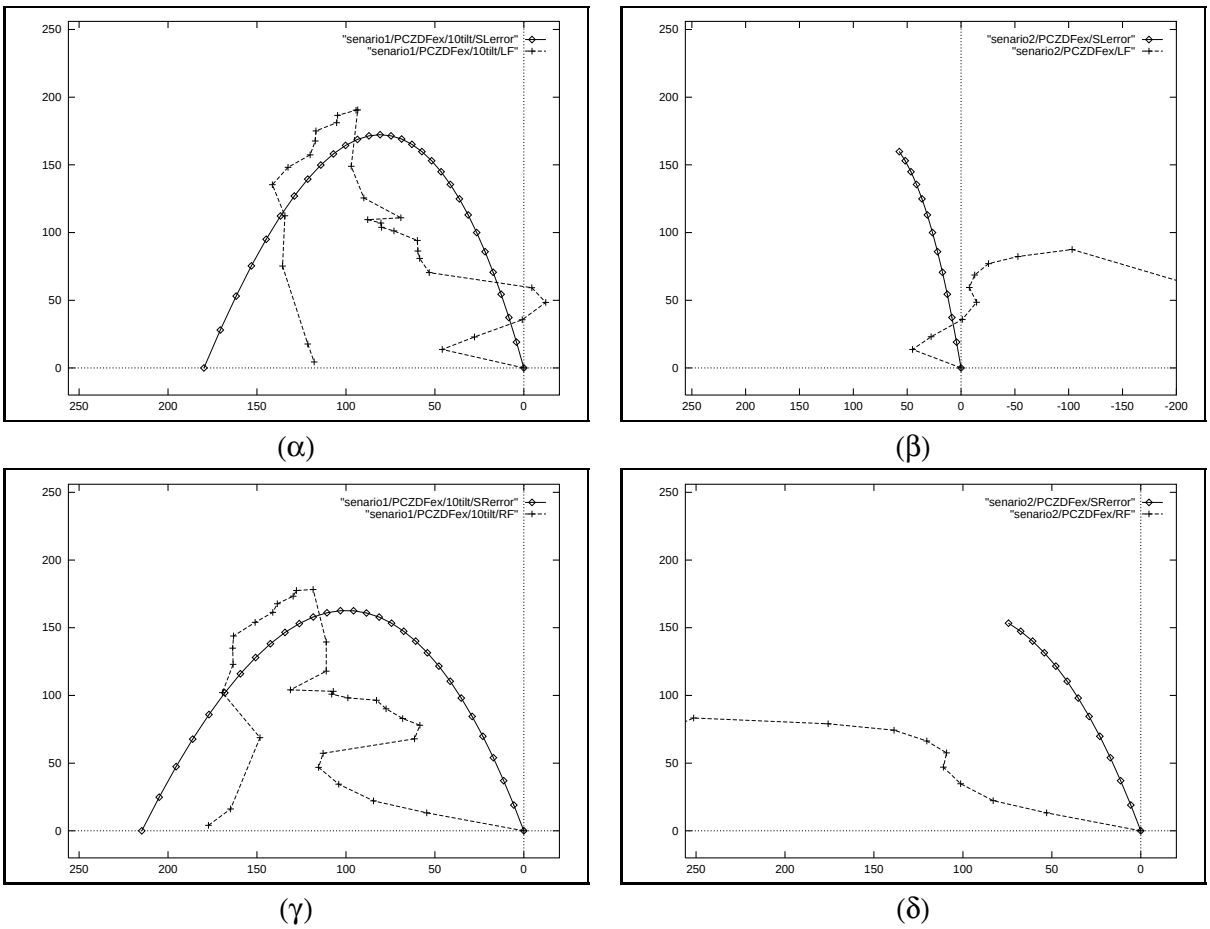


(α) (β)

(γ) (δ)

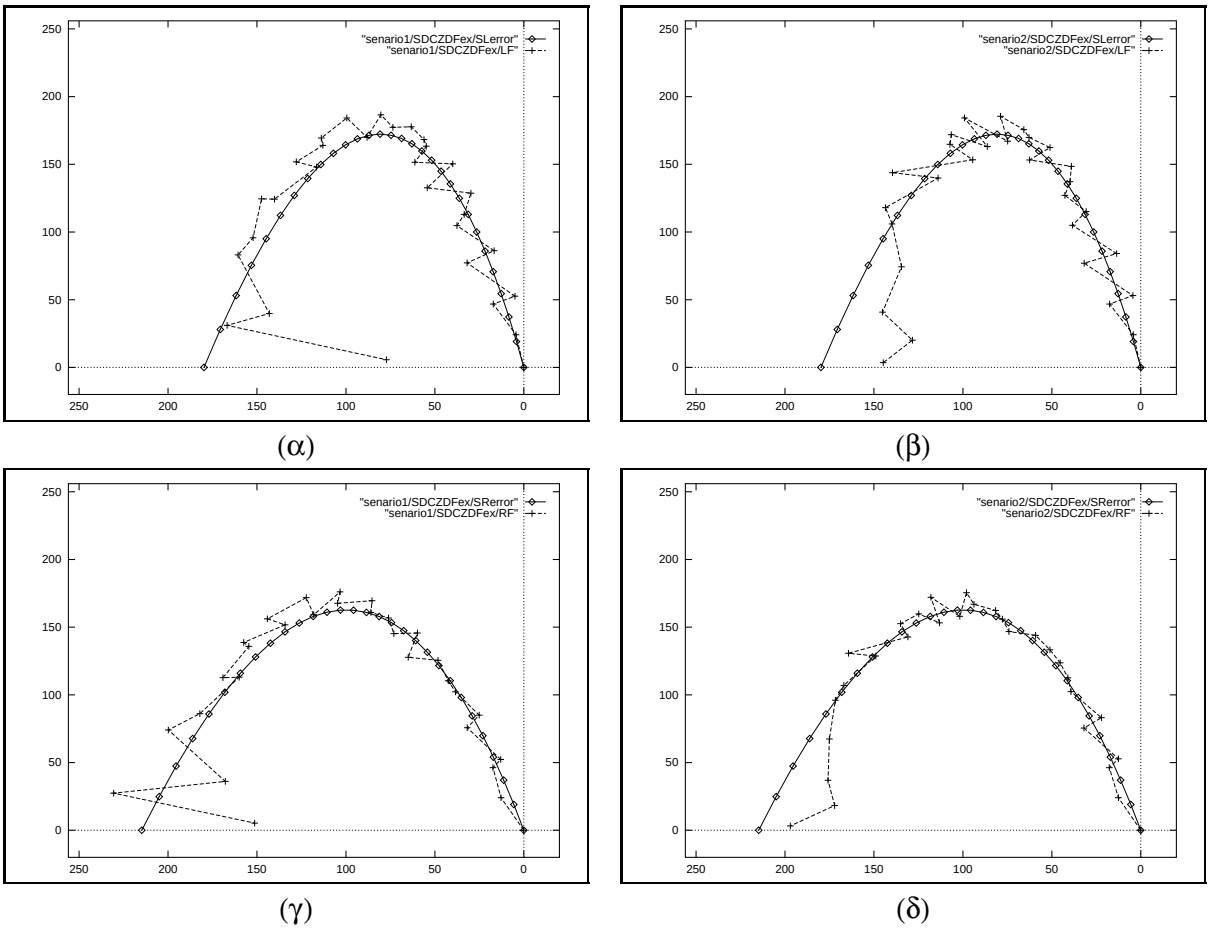
Σχήμα 6.6: Λάθος παρακολούθησης χρησιμοποιώντας το απλό ΦΜΑ συσχέτισης κάθετων ακμών.

Στα σχήματα (α),(β) ((γ),(δ)) παριστάνεται η τροχιά του σημείου καθήλωσης όπως αυτό φαίνεται από την αρχική θέση της αριστερής (δεξιάς) κάμερας. Τα σχήματα (α),(γ) ((β),(δ)) έχουν προκύψει με εφαρμογή του πρώτου (δεύτερου) σεναρίου κίνησης του κεφαλιού. Όπως φαίνεται από τα σχήματα με το δεύτερο σενάριο επιτυγχάνεται καλύτερη παρακολούθηση. Στο πρώτο σενάριο στα τελευταία στιγμιότυπα χάνεται η σύγκλιση με αποτέλεσμα να αποτύχει και η παρακολούθηση.



Σχήμα 6.7: Λάθος παρακολούθησης χρησιμοποιώντας το ΦΜΑ συσχέτισης γειτονιάς με μέτρο ποιότητας των χαρακτηριστικών τις οριζόντιες διαφορές των εντάσεων

Στα σχήματα (α),(β) ((γ),(δ)) παριστάνεται η τροχιά του σημείου καθήλωσης όπως αυτό φαίνεται από την αρχική θέση της αριστερής (δεξιάς) κάμερας. Τα σχήματα (α),(γ) ((β),(δ)) έχουν προκύψει με εφαρμογή του πρώτου (δεύτερου) σεναρίου κίνησης του κεφαλιού. Όπως είναι εμφανές η συμπεριφορά του συστήματος είναι χειρότερη από το προηγούμενο πείραμα. Στο δεύτερο σενάριο η κακή συμπεριφορά της μεθόδου οδηγεί σε αποτυχία της σύγκλισης το οποίο έχει σαν αποτέλεσμα το αντικείμενο να χάνεται τελείως από το οπτικό πεδίο.



Σχήμα 6.8: Λάθος παρακολούθησης χρησιμοποιώντας το ΦΜΑ συσχέτισης γειτονιάς με μέτρο ποιότητας των χαρακτηριστικών την οριζόντια τυπική απόκλιση

Στα σχήματα (α),(β) ((γ),(δ)) παριστάνεται η τροχιά του σημείου καθήλωσης όπως αυτό φαίνεται από την αρχική θέση της αριστερής (δεξιάς) κάμερας. Τα σχήματα (α),(γ) ((β),(δ)) έχουν προκύψει με εφαρμογή του πρώτου (δεύτερου) σεναρίου κίνησης του κεφαλιού. Όπως φαίνεται, η χρήση της οριζόντιας τυπικής απόκλισης σαν μέτρο ποιότητας χαρακτηριστικών έχει πολύ καλύτερα αποτελέσματα από ότι η χρήση τοπικών διαφορών. Η συμπεριφορά του συστήματος πλησιάζει περισσότερο αυτή στην οποία χρησιμοποιείται το απλό ΦΜΑ.

ποιότητας των χαρακτηριστικών την ύπαρξη ακμών σε διάφορες αναλύσεις (σχήματα 6.9, 6.10). Μια παράμετρος που επηρέαζει την απόδοση όλων των φίλτρων συσχέτισης γειτονιών είναι το μέγεθος του παραθύρου που χρησιμοποιείται για την εφαρμογή της συσχέτισης. Όσο μεγαλύτερο είναι το παράθυρο, τόσο περισσότερο εξομαλυμένη είναι η έξοδος του φίλτρου.

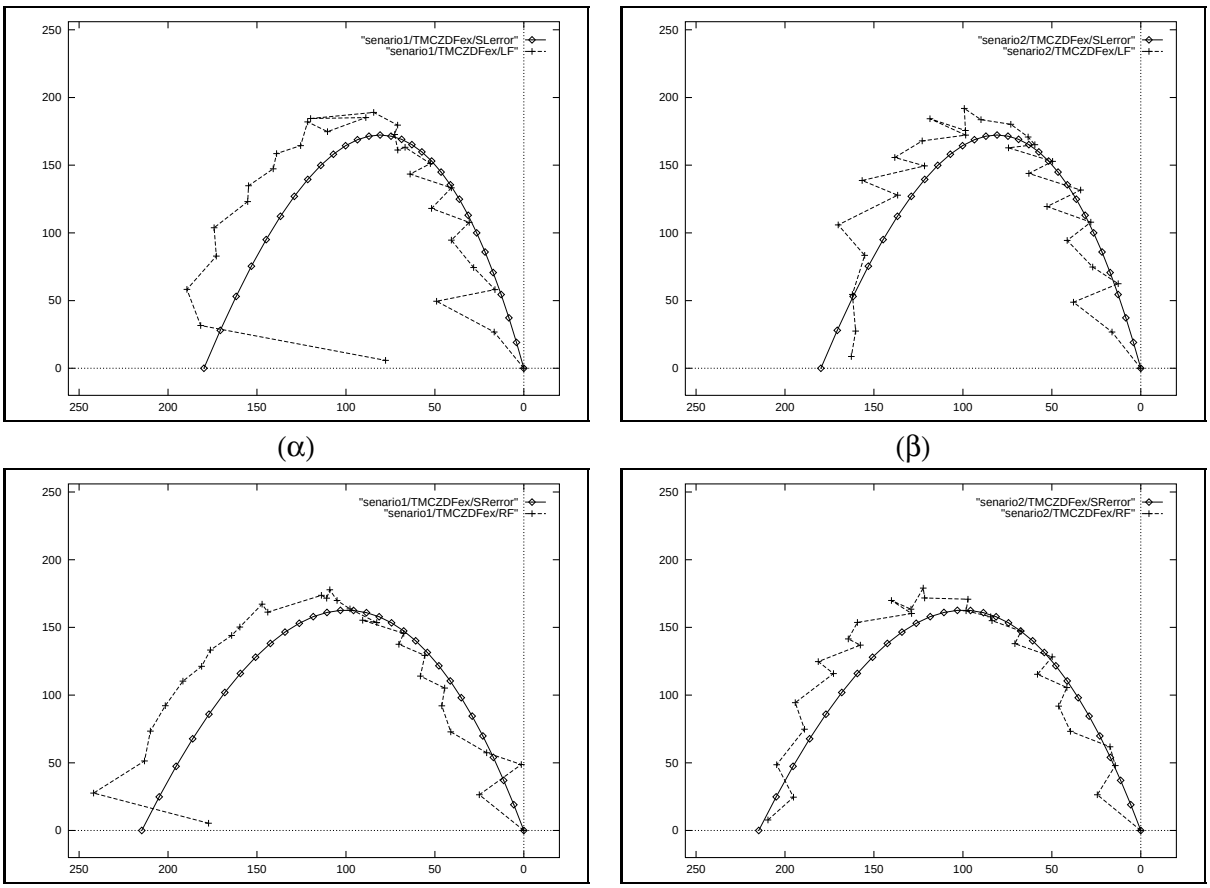
Με το ΦΜΑ συσχέτισης μέτρου-κατεύθυνσης ακμών (σχήμα 6.11) φαίνεται να έχουμε κάποια βελτίωση στην σταθεροποίηση. Η βελτίωση αυτή οφείλεται στο ότι το φίλτρο συγκεντρώνει πληροφορία και από ακμές που είναι σχεδόν κάθετες.

Από τα πειράματα μπορούν να βγουν γενικά συμπεράσματα και όχι ειδικά γιατί οι πηγές λάθους είναι πολλές και όχι ελεγχόμενες (γεωμετρικά σφάλματα, χαμηλή ακρίβεια των φίλτρων, έλλειψη έλεγχου).

Σ' όλα τα πειράματα με το πρώτο σενάριο η ασυμμετρία του συστήματος επηρεάζει αρνητικά τη σύγκλιση η οποία μερικές φορές αποτυγχάνει τελείως, και αυτό έχει σαν συνέπεια την αποτυχία και της παρακολούθησης. Η σύγκλιση επηρεάζεται σημαντικά από την ανάλυση στην οποία εφαρμόζεται το cepstal φίλτρο και αυτό έχει σαν αποτέλεσμα την ταλάντωση του σημείου καθήλωσης γύρω από την θέση στην οποία βρίσκεται το αντικείμενο. Η ταλάντωση αυτή θα μπορούσε να μειωθεί αν υπήρχε κάποιος έλεγχος που να αποκρίνεται λιγότερο στα λάθη που υπολογίζει το σύστημα.

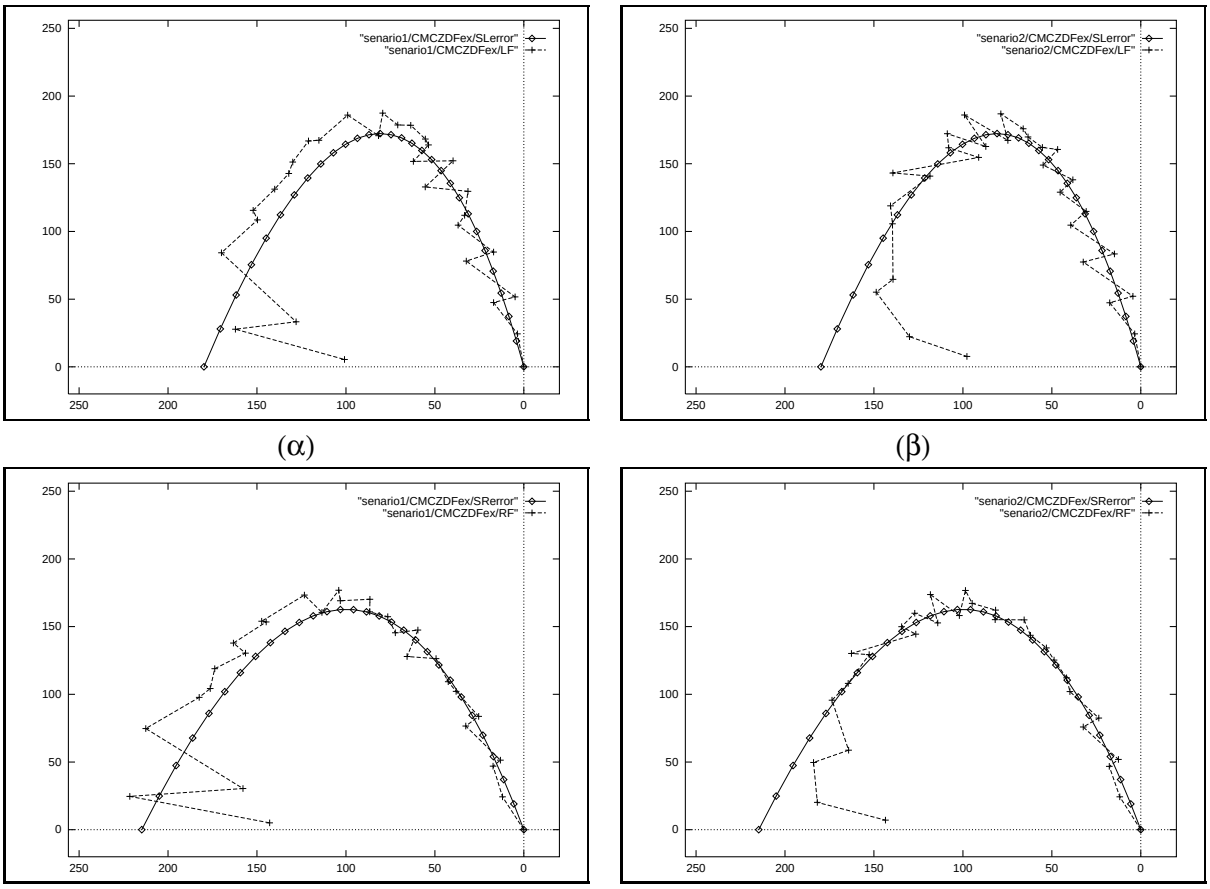
Ο διαχωρισμός του αντικειμένου από τα τριγύρω αντικείμενα εξαρτάται σε μεγάλο βαθμό από την γεωμετρία του συστήματος. Όσο μεγαλύτερη είναι η περιοχή στην οποία έχουμε μικρές ανομοιότητες τόσο μεγαλύτερη πιθανότητα υπάρχει το σύστημα να ξεφύγει από την επιφάνεια του αντικειμένου και να παρακολουθήσει άλλα αντικείμενα τα οποία βρίσκονται σ' αυτή την περιοχή. Επίσης, το σύστημα είναι ευαίσθητο στο πρόβλημα της ευθυγράμμισης των ακμών (aliasing), μ' αποτέλεσμα σε πειράματα με επαναλαμβανόμενα patterns το σύστημα να παρακολουθεί τυχαία τα patterns.

Η έλλειψη του ελέγχου κατά την εξέλιξη του φαινομένου έχει σαν συνέπεια το σύστημα να αποτυγχάνει πολύ εύκολα με την εμφάνιση κάποιου λάθους. Ειδικά, η χρήση ελέγχου σε συνδυασμό με κάποιου είδους πρόβλεψης της κίνησης του αντικειμένου και των καθυστερήσεων ενός πραγματικού συστήματος, είναι επιτακτική, μια και μπορεί να αντισταθμίσει τις αδυναμίες των φίλτρων.



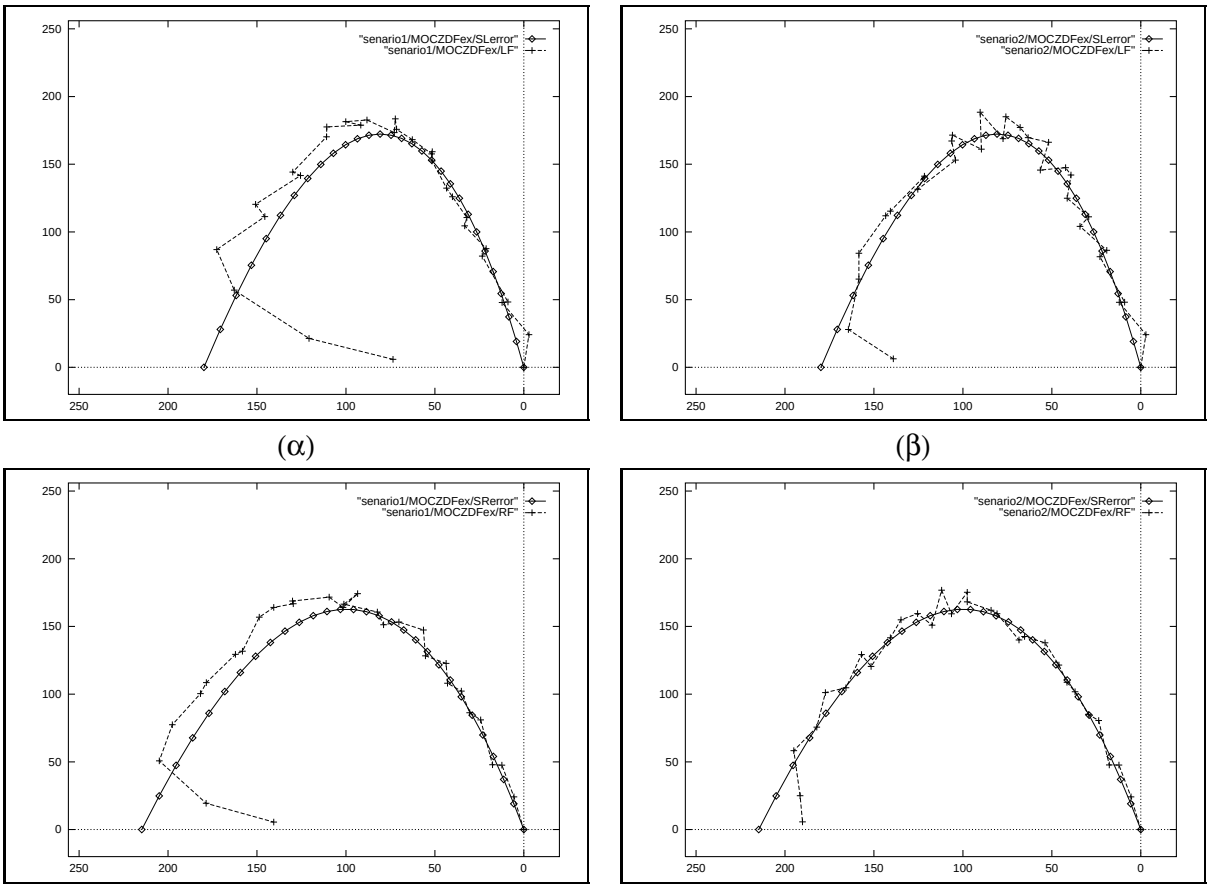
Σχήμα 6.9: Λάθος παρακολούθησης χρησιμοποιώντας το ΦΜΑ συσχέτισης γειτονιάς με μέτρο ποιότητας των χαρακτηριστικών που προκύπτει από την συνέλιξη της στερεοσκοπικής εισόδου με την ολική μάσκα

Στα σχήματα (α),(β) ((γ),(δ)) παριστάνεται η τροχιά του σημείου καθήλωσης όπως αυτό φαίνεται από την αρχική θέση της αριστερής (δεξιάς) κάμερας. Τα σχήματα (α),(γ) ((β),(δ)) έχουν προκύψει με εφαρμογή του πρώτου (δεύτερου) σεναρίου κίνησης του κεφαλιού. Με την εφαρμογή του φίλτρου αυτού ανιχνεύονται κάθετες ακμές σε διαφορετικές αναλύσεις και έτσι η αντιστοίχιση των χαρακτηριστικών της εικόνας βασίζεται σε περισσότερη πληροφορία. Η συμπεριφορά του λάθους είναι παρόμοια με αυτήν στο προηγούμενο σχήμα.



Σχήμα 6.10: Λάθος παρακολούθησης χρησιμοποιώντας το ΦΜΑ συσχέτισης γειτονιάς με μέτρο ποιότητας των χαρακτηριστικών που προκύπτει από τον πολλαπλασιασμό των γειτονιών με την κεντρική μάσκα

Στα σχήματα (α),(β) ((γ),(δ)) παριστάνεται η τροχιά του σημείου καθήλωσης όπως αυτό φαίνεται από την αρχική θέση της αριστερής (δεξιάς) κάμερας. Τα σχήματα (α),(γ) ((β),(δ)) έχουν προκύψει με εφαρμογή του πρώτου (δεύτερου) σεναρίου κίνησης του κεφαλιού. Και εδώ βλέπουμε ότι η συμπεριφορά του λάθους είναι παρόμοια με αυτήν στα δύο προηγούμενα σχήματα.



Σχήμα 6.11: Λάθος παρακολούθησης χρησιμοποιώντας το ΦΜΑ συσχέτισης μέτρου-κατεύθυνσης των ακμών

Στα σχήματα (α),(β) ((γ),(δ)) παριστάνεται η τροχιά του σημείου καθήλωσης όπως αυτό φαίνεται από την αρχική θέση της αριστερής (δεξιάς) κάμερας. Τα σχήματα (α),(γ) ((β),(δ)) έχουν προκύψει με εφαρμογή του πρώτου (δεύτερου) σεναρίου κίνησης του κεφαλιού. Το φίλτρο αυτό μοιάζει ως προς τη λειτουργία του με το απλό ΦΜΑ κάθετων ακμών αλλά συγκεντρώνει πληροφορία και από ακμές που είναι σχεδόν κάθετες. Η συμπεριφορά του λάθους είναι λίγο καλύτερη από ότι χρησιμοποιώντας τα προηγούμενα φίλτρα.

Κεφάλαιο 7

Επίλογος

7.1 Συμπεράσματα

Στα πλαίσια αυτής της εργασίας υλοποίηθηκε ένα σύστημα σταθεροποίησης της οπτικής προσήλωσης για μια στερεοσκοπική κεφαλή ενεργού όρασης. Το σύστημα στηρίχθηκε στην αρχιτεκτονική που έχει προταθεί από τον D.J. Coombs [22] και αποτελείται από δύο υποσυστήματα: το υποσύστημα σύγκλισης και το υποσύστημα καταδίωξης, τα οποία συνεργάζονται για να είναι δυνατή η παρακολούθηση κινούμενων αντικειμένων. Το υποσύστημα σύγκλισης στηρίζεται στον υπολογισμό της ανομοιότητας με τη χρήση του cepstral φίλτρου και το σύστημα καταδίωξης στην τμηματοποίηση του στόχου με φίλτρα μηδενικής ανομοιότητας.

Για τις απαιτήσεις της σύγκλισης, έχει υλοποιηθεί, όπως προτείνουν οι Olson και Coombs στο [53], το cepstral φίλτρο που είναι μια μέθοδος συσχέτισης. Το βασικότερο μειονέκτημα της μεθόδου είναι ο τρόπος που επιλέγεται το τοπικό μέγιστο στην έξοδο του φίλτρου το οποίο αντιστοιχεί στο στόχο. Η υπόθεση που γίνεται για την επιλογή είναι ότι η μεγαλύτερη κορυφή στο cepstrum είναι αυτή που οφείλεται στο στόχο. Αυτό όμως είναι περιοριστικό ως προς το τι πρέπει να περιέχει η σκηνή για να ισχύει αυτή η υπόθεση. Για το λόγο αυτό, υλοποιήθηκε ένας άλλος αλγόριθμος σύγκλισης που πρότειναν οι Olson, Taylor και Martin [66] ο οποίος εφαρμόζει αναζήτηση της κατάλληλης κορυφής σε πολλά επίπεδα ανάλυσης. Ο αλγόριθμος αυτός θέτει πιο χαλαρούς περιορισμούς ως προς τα αντικείμενα που περιέχονται στη σκηνή και εντοπίζει με μεγαλύτερη ακρίβεια την κορυφή. Ωστόσο, η τεχνική αυτή είναι ακριβότερη υπολογιστικά λόγω της εφαρμογής του cepstral μετασχηματισμού σε πολλά επίπεδα ανάλυσης.

Για το σύστημα καταδίωξης υλοποιήθηκαν φίλτρα μηδενικής ανομοιότητας που

προτάθηκαν από τους Coombs, Horswill και Kaenel [21, 22] και τα οποία στηρίζονται στην αντιστοίχιση κάθετων χαρακτηριστικών στις δύο εικόνες του στερεοσκοπικού ζεύγους.

Όσον αφορά το πρόβλημα προσανατολισμού της στερεοσκοπικής κεφαλής, δοσμένων των εκτιμήσεων ανομοιότητας και θέσης του στόχου στο επίπεδο των καμερών, εξετάσαμε διάφορα σενάρια κατανομής των λαθών σύγκλισης και θέσης στους βαθμούς ελευθερίας του κεφαλιού. Τα σενάρια αυτά εισάγουν διαφορετικά σφάλματα τόσο λόγω της εξάρτησης των βαθμών ελευθερίας μεταξύ τους όσο και λόγω της θέσης του συστήματος η οποία δεν είναι γνωστή με αποτέλεσμα οι διορθώσεις να γίνονται προσεγγιστικά.

Για την αξιολόγηση της λειτουργίας του συστήματος σταθεροποίησης, αναπτύξαμε μία προσομοίωση της ενεργού κεφαλής όπως περιγράψαμε στο προηγούμενο κεφάλαιο. Δεν είναι εφικτό να διατυπώσουμε γενικά συμπεράσματα για την απόδοση του συστήματος λόγω των πολλών και ανεξέλεγκτων πηγών λάθους. Ωστόσο, χρησιμοποιώντας εναλλακτικά τις διάφορες μεθόδους που διεκπεραιώνουν την σύγκλιση και την καταδίωξη, καθώς και τα διαφορετικά σενάρια κίνησης της εικονικής πλατφόρμας σημειώσαμε σημαντικά σημεία της συμπεριφοράς του συστήματος:

- Αν αποτύχει κάποιο υποσύστημα (σύγκλισης/καταδίωξης) το σύστημα σταθεροποίησης αποτυχαίνει και οδηγείται στην αστάθεια. Αυτό είναι αναμενόμενο αφού το σύστημα στηρίζεται στην συνεργασία των δύο υποσυστημάτων. Ο D.J Coombs στο [22] παρουσιάζει την συμπεριφορά των δύο υποσυστημάτων και την συμμετοχή καθενός απ' αυτά στην παρακολούθηση. Καταλήγει στο ότι κάθε υποσύστημα συμβάλλει στην απόδοση του συστήματος και στο ότι ο συνδυασμός των υποσυστημάτων επιτρέπει την χρήση απλών τεχνικών για την υλοποίηση τους. Ωστόσο, ισχυρίζεται ότι ο έλεγχος της σύγκλισης δεν είναι απαραίτητος αφού το φίλτρο ανομοιότητας μπορεί να σχεδιαστεί έτσι ώστε να αποκρίνεται σε οποιαδήποτε πεπερασμένη ανομοιότητα.

- Η διατήρηση της συμμετρικής σύγκλισης των καμερών είναι σημαντική έτσι ώστε να επιτυγχάνεται σύγκλιση χωρίς μεγάλα σφάλματα που να οδηγούν την παρακολούθηση σε αποτυχία.

- Η ανάλυση της εισόδου στην οποία εφαρμόζονται τα διάφορα φίλτρα σχετίζεται με την ταλάντωση του συστήματος γύρω από την ιδανική απόκριση παρακολούθησης.

- Ο διαχωρισμός του στόχου από τα τριγύρω αντικείμενα εξαρτάται σε μεγάλο βαθμό από τη γεωμετρία των καμερών. Όσο μεγαλύτερο είναι το πεδίο θέας τόσο πιο επιρρεπές είναι το σύστημα στην παρακολούθηση αντικειμένων που έχουν επίσης μικρή ανομοιότητα.

- Η παρουσία περιοδικών προτύπων (patterns) στη σκηνή κάνει τα φίλτρα μηδενικής ανομοιότητας να δίνουν εσφαλμένες θετικές απαντήσεις, με αποτέλεσμα το σύστημα να παρακολουθεί τυχαία τα πρότυπα αυτά.

- Η έλλειψη πρόβλεψης και δυναμικής στο σύστημα έχει σαν συνέπεια το σύστημα να αποτυγχάνει πολύ εύκολα με την εμφάνιση κάποιου λάθους. Ο D.J. Coombs στο [22] αναλύει εκτενώς την αλληλεπίδραση του ελέγχου και της αντίληψης λαμβάνοντας υπ' όψην την δυναμική του συστήματος και του ερεθίσματος. Υποστηρίζει ότι η χρήση πρόβλεψης μπορεί να βελτιώσει την απόδοση της παρακολούθησης στόχων, των οποίων η κίνηση μπορεί να προβλεφθεί, μετριάζοντας την επίδραση των καθυστερήσεων στο σύστημα.

7.2 Επεκτάσεις

Η λειτουργία του συστήματος στην προσομοίωση σίγουρα δεν είναι ενδεικτική για την εξαγωγή συμπερασμάτων όσον αφορά την ανθεκτικότητα του σε ένα δυναμικό περιβάλλον με διάφορες πηγές θορύβου. Η εφαρμογή του συστήματος σε μια μηχανική κεφαλή σ' αυτή τη φάση θα είχε ενδιαφέρον και θα έδινε ώθηση στη συνεργασία του και με άλλα συστήματα οπτικών συμπεριφορών στοχεύοντας τόσο στην καλύτερη απόδοση του όσο και στην ανάπτυξη συνθετότερων συμπεριφορών (π.χ. σε συνδυασμό με ένα σύστημα αποφυγής εμποδίων). Μια τέτοια εφαρμογή μπορεί να αναδείξει την χρησιμότητα του συστήματος σταθεροποίησης, πράγμα το οποίο αδυνατεί να κάνει η προσομοίωση.

Ωστόσο, κατά την ανάπτυξη διάφορων λειτουργιών για μια μηχανική πλατφόρμα είναι χρήσιμη η προσομοίωση για την καλύτερη κατανόηση της συμπεριφοράς των αλγορίθμων. Επιπλέον, επιτρέπει επανάληψη ακριβώς του ίδιου πειράματος και πραγματοποίηση πειραμάτων σε αργή κίνηση. Η προσομοίωση στην παρούσα εργασία καλύπτει μόνο την κινηματική της πλατφόρμας. Θα μπορούσε, ωστόσο, να επεκταθεί έτσι ώστε να υποστηρίξει την δυναμική του συστήματος και συστήματα ρομποτικής με περισσότερους οπτικούς ή μηχανικούς βαθμούς ελευθερίας. Επίσης θα ήταν πολύ χρήσιμη η οπτική αναπαράσταση της κατάστασης του συστήματος και η διαλογική ενεργοποίηση των οπτικών αλγορίθμων που συνθέτουν το σύστημα και των διάφορων σχημάτων ελέγχου.

Στο σύστημα που αναπτύχθηκε χρησιμοποιήθηκαν μόνο στερεοσκοπικές πηγές πληροφορίας. Θα μπορούσαν να χρησιμοποιηθούν και άλλες πηγές πληροφορίας είτε οπτικές (π.χ. κίνηση) είτε από άλλους αισθητήρες που μπορεί να βρίσκονται στην κινούμενη πλατφόρμα.

Από τα αποτελέσματα που παρουσιάστηκαν στο προηγούμενο κεφάλαιο είναι αισθητή η έλλειψη ελέγχου με κάποιο είδος πρόβλεψης της δυναμικής του κεφαλιού και της δυναμικής της σκηνής. Η πρόβλεψη θα μπορούσε να εξομαλύνει την ταλάντωση του λάθους που οφείλεται σε διάφορους παράγοντες καθώς και να αντιμετωπίσει κάποιες περιπτώσεις αποτυχίας.

Παράρτημα Α

Σφάλμα διόρθωσης κατά την pan κίνηση του κεφαλιού

Μπορούμε να εκφράσουμε την pan κίνηση με τον εξής μετασχηματισμό :

$$\mathrm{M}_{pan} = \mathrm{T}_{\vec{d}} \times R_{\vec{r},-\theta_p} \times T_{-\vec{d}} \tag{A.1}$$

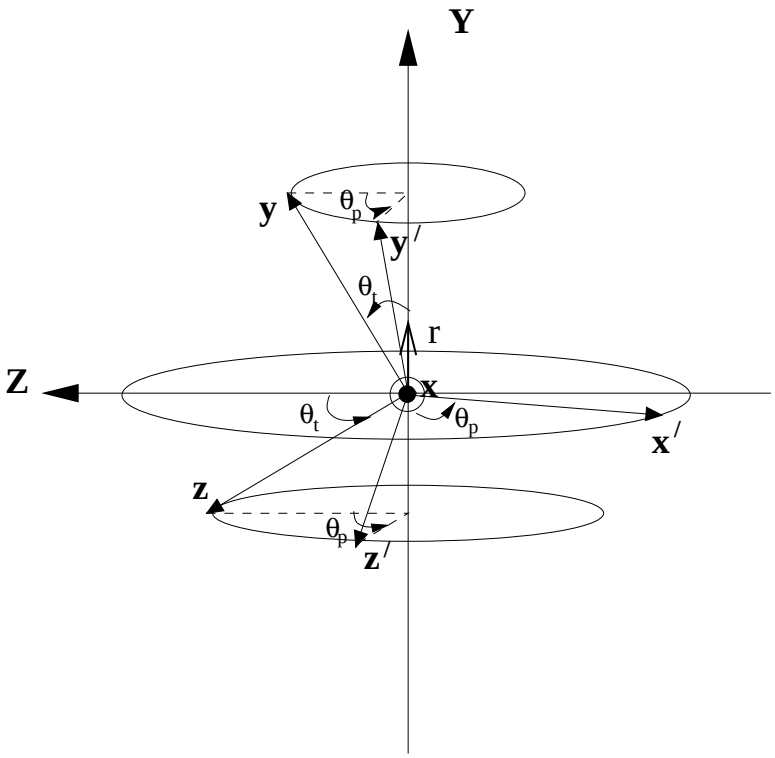
όπου $\vec{d} = (h, \frac{b}{2}, 0)$ για την αριστερή κάμερα, $\vec{d} = (h, -\frac{b}{2}, 0)$ για την δεξιά, και $\vec{r} = (0, \cos\theta_t, -\sin\theta_t)$ ως προς το σύστημα συντεταγμένων $Oxyz$ της κάμερας (σχήμα Α.1), θ_p η διόρθωση στην pan γωνία και θ_t η γωνία tilt. Το διάνυσμα $\vec{r}$ είναι ο pan άξονας του κεφαλιού, δηλ. ο OY άξονας του συστήματος συντεταγμένων OXYZ που βλέπουμε στο σχήμα.

Υπενθυμίζουμε ότι ο βασικός ομογενής πίνακας μεταφοράς κατά $\vec{d} = (d_x, d_y, d_z)$ είναι:

$$\mathrm{T}_{\vec{d}} = \begin{bmatrix} 1 & 0 & 0 & d_x \\ 0 & 1 & 0 & d_y \\ 0 & 0 & 1 & d_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{A.2}$$

ενώ ο βασικός ομογενής πίνακας περιστροφής κατά γωνία ϕ γύρω από τον άξονα $\vec{r} = (r_x, r_y, r_z)$ είναι:

$$R_{\vec{r}} = \begin{bmatrix} {r_x}^2(1-\cos\phi) + cos\phi & r_x r_y(1-\cos\phi) - r_z\sin\phi & r_x r_z(1-\cos\phi) + r_y\sin\phi & 0 \\ r_x r_y(1-\cos\phi) + r_z\sin\phi & {r_y}^2(1-\cos\phi) + cos\phi & r_y r_z(1-\cos\phi) - r_x\sin\phi & 0 \\ r_x r_z(1-\cos\phi) - r_y\sin\phi & r_y r_z(1-\cos\phi) + r_x\sin\phi & {r_z}^2(1-\cos\phi) + cos\phi & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{A.3}$$



Σχήμα Α.1: Γεωμετρική άποψη της περιστροφής στο μετασχηματισμό που εκφράζει την pan κίνηση του κεφαλιού.

Οποιοδήποτε σημείο P με συντεταγμένες $P_1 = (x_1, y_1, z_1, 1)^T$ (ομογενείς συντεταγμένες) ώς προς το σύστημα συντεταγμένων της κάμερας, μετά την pan κίνηση (δηλ. την εφαρμογή του μετασχηματισμού M_{pan}), θα ορίζεται από το διάνυσμα συντεταγμένων P_2 :

$$P_2 = \mathrm{M}_{pan} \times P_1 \tag{A.4}$$

Ο υπολογισμός του σφάλματος διόρθωσης μπορεί να γίνει παίρνοντας την προοπτική προβολή του σημείου P_2 στο επίπεδο της κάμερας και στη συνέχεια υπολογίζοντας την ευκλείδια απόσταση της προβολής από το κέντρο του επιπέδου.

Ευρετήριο Α

Ελληνοαγγλικό Ευρετήριο Όρων

α

αισθητήρας	sensor
ανομοιότητα	disparity
αντανακλαστικό κοντινης τριάδας	near triad reflex
αυτοεκπαίδευση	self-learning
αυτόνομο σύστημα	autonomous system
αυτοσυσχέτιση	autocorrelation

β

βασική γραμμή	baseline

δ

διοφθαλμικός	binocular

ε

εγγύς πεδίο	proximity space
εικόνα διαφορών	difference image
εικόνα κατεύθυνσης	orientation image
εικόνα μέτρου	magnitude image

εικόνα ποιότητας των χαρακτηριστικών	feature quality image
ελεγκτής	controller
έλεγχος ανάδρασης	feedback control
έλεγχος εστίασης	focus control
έλεγχος προσήλωσης	gaze control
ενεργός όραση	active vision
επικάλυψη	occlusion
επιπολικός	epipolar
ετεροσυσχέτιση	cross-correlation

θ

θαλαμιαίο οφθαλμικό αντανακλαστικό	vestibulo-ocular reflex

ι

ιδία κίνηση	egomotion

κ

καθήλωση	fixation
κάμερα απειροελαχίστου ανοίγματος	pinhole camera
καταδίωξη	pursuit
κινητήρας	motor
κινητικότητα	mobility

λ

λειτουργική εξειδίκευση	functional specialization

μ

μετατόπιση προσήλωσης	gaze change

O

ομαλή καταδίωξη	smooth pursuit
οπτική ροή	optical flow
οπτοκινητικό αντανακλαστικό	opto-kinetic reflex

ορόπτερο	horopter

π

παρακολούθηση	tracking
περιπολία	surveillance
περιστροφή γύρω από τον οπτικό άξονα	torsion
πλοήγηση	navigation
πρόβλημα αντιστοίχισης	correspondence problem
προοπτική προβολή	perspective projection
προσαρμογή	accommodation
προσαρμοστική σύγκλιση	accommodative convergence
πρωτεύοντα θηλαστικά	primates

ρ

ραβδωτός φλοιός	striate cortex

σ

σταθεροποίηση προσήλωσης	gaze stabilization
στροφή	version
σύγκλιση	vergence
συγκλιτικη προσαρμογή	convergence accommodation
συσχέτιση κατά ζεύγη	pairwise correlation

τ

τμηματοποίηση	segmentation
τυπική απόκλιση	standard deviation

υ

υπο-δειγματοληψία	subsampling
υφή	texture

φ

φάσμα ισχύος	power spectrum

ω

ωτολιθικός	otolith
ωχρή κηλίδα	fovea

Ευρετήριο Β

Αγγλοελληνικό Ευρετήριο Όρων

A

accommodation	προσαρμογή
accommodative convergence	προσαρμοστικη σύγκλιση
active vision	ενεργός όραση
autocorrelation	αυτοσυσχέτιση
autonomous system	αυτόνομο σύστημα

B

baseline	βασική γραμμή
binocular	διοφθαλμικός

C

controller	ελεγκτής
convergence accommodation	συγκλιτική προσαρμογή
correspondence problem	πρόβλημα αντιστοίχισης
cross-correlation	ετεροσυσχετιση

D

difference image	εικόνα διαφορών
disparity	ανομοιότητα

E

egomotion	ιδία κίνηση
epipolar	επιπολικός

F

feature quality image	εικόνα ποιότητας των χαρακτηριστικών
feedback control	έλεγχος ανάδρασης
fixation	καθήλωση
focus control	έλεγχος εστίασης
fovea	ωχρή κηλίδα
functional specialization	λειτουργική εξειδίκευση

G

gaze change	μετατόπιση προσήλωσης
gaze control	έλεγχος προσήλωσης
gaze stabilization	σταθεροποίηση προσήλωσης

H

horopter	ορόπτερο

M

magnitude image	εικόνα μέτρου
mobility	κινητικότητα
motor	κινητήρας

N

navigation	πλοήγηση
near triad reflex	αντανακλαστικό κοντινής τριάδας

O

occlusion	επικάλυψη
optical flow	οπτική ροή
opto-kinetic reflex	οπτοκινητικό αντανακλαστικό
orientation image	εικόνα κατεύθυνσης
otolith	ωτολιθικός

P

pairwise correlation	συσχέτιση κατά ζεύγη
perspective projection	προοπτική προβολή
pinhole camera	κάμερα απειροελαχιστου ανοίγματος
power spectrum	φάσμα ισχύος
primates	πρωτεύοντα θηλαστικα
proximity space	εγγύς πεδίο
pursuit	καταδίωξη

S

segmentation	τμηματοποίηση
self-learning	αυτοεκπαιδευση
sensor	αισθητήρας
smooth pursuit	ομαλή καταδίωξη
standard deviation	τυπική απόκλιση
striate cortex	ραβδωτός φλοιός
subsampling	υπο-δειγματοληψία
surveillance	περιπολία

T

texture	υφή
torsion	περιστροφή γύρω από τον οπτικό άξονα
tracking	παρακολούθηση

V

vergence	σύγκλιση
version	στροφή
vestibulo-ocular reflex	θαλαμιαίο οφθαλμικό αντανακλαστικό

Βιβλιογραφία

[1] Μανόλης Ι. Α. Λουράκης. *Μια εναλλακτική προσέγγιση για τη μελέτη της αντίληψης του χώρου : Χρήση διατακτικού αντί μετρικού βάθους.* Μεταπτυχιακό Δίπλωμα Ειδίκευσης, Πανεπιστήμιο Κρήτης, Σχολή Θετικών Επιστημών, Τμήμα Επιστήμης Υπολογιστών, Ιανουάριος 1995.

[2] Αντώνης Α. Αργυρός. *Οπτική Ανίχνευση Ανεξάρτητης Τρισδιάστατης Κίνησης από Κινούμενο Παρατηρητή.* Διδακτορική Διατριβή, Πανεπιστήμιο Κρήτης, Σχολή Θετικών Επιστημών, Τμήμα Επιστήμης Υπολογιστών, Οκτώβριος 1996.

[3] Νικόλαος Κ. Παραγυιός. *Δυναμική Ανίχνευση και Εντοπισμός Κινούμενων Αντικειμένων σε Ακολουθίες Εικόνων.* Μεταπτυχιακό Δίπλωμα Ειδίκευσης, Πανεπιστήμιο Κρήτης, Σχολή Θετικών Επιστημών, Τμήμα Επιστήμης Υπολογιστών, Ιούνιος 1996.

[4] Narendra Ahuja and A. Lynn Abbot. Active Stereo: Integrating Disparity, Vergence, Focus, Aperture, and Calibration for Surface Estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(10):1007--1029, 1993.

[5] John Aloimonos, Isaac Weiss, and Amit Bandyopadhyay. Active Vision. *International Journal of Computer Vision*, pages 333--356, 1988.

[6] John(Yannis) Aloimonos and Dimitris P. Tsakiris. On the Visual Mathematics of Tracking. *Image and Vision Computing*, 9(4):235--252, 1991.

[7] Magnus Andersson and Peter Nordlund. A Model-Based System for Localization and Tracking. In *Proc. Workshop on Computer Vision for Space Applications*, Antibes, France, September 1993.

[8] Ruzena Bajcsy. Active Perception. *Proceedings of the IEEE*, 76(8):996--1005, August 1988.

[9] Dana H. Ballard. Animate Vision. *Artificial Intelligence*, 48:57--86, 1991.

[10] Dana H. Ballard and Christopher M. Brown. Principles of Animate Vision. *Computer Vision Graphics and Image Processing : Image Understanding*, 56(1):3--21, 1992.

[11] M.G.P. Bartholomeus, B.J.A. Krose, and A.J.Noest. A Robust Multi-Resolution Vision System for Target Tracking with a Moving Camera. *Computer Science in The Netherlands, (H. Wijshof, ed.), CWI, Amsterdam*, pages 52--63, 1993. ftp://ftp.fwi.uva.nl/computer-systems/aut-sys/reports/BarKroNoe93.ps.gz.

[12] Kevin J. Bradshaw, Philip F. McLauchlan, Ian D. Reid, and David W. Murray. Saccade and Pursuit on an Active Head/Eye Platform. *Image and Vision Computing*, 12(3):155--163, April 1994.

[13] Christopher Brown. Gaze Controls with Interactions and Delays. *IEEE Transactions on Systems, Man, and Cybernetics*, 20(1):518--527, March 1990.

[14] Christopher Brown, David Coombs, and John Soong. Real-time Smooth Pursuit Tracking. In Blake A and Yuille A, editors, *Active vision*, chapter 8. MIT, 1992.

[15] Kjell Brunnstrφm, Jan-Olof Eklundh, and Tomas Uhlin. Active Fixation for Scene Exploration. *International Journal of Computer Vision*, 17(2):137--162, 1996.

[16] R.H.S. Carpenter. *Movements of the Eyes.* Pion Press: London, 1988.

[17] Qinsheng Chen, Michel Defrise, and F. Deconinck. Symmetric Phase-Only Matched Filtering of Fourier-Mellin Transforms for Image Registration and Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(12):1156--1168, December 1994.

[18] James J. Clark and Nicola J. Ferrier. Attentive Visual Servoing. In Blake A and Yuille A, editors, *Active vision*, chapter 9. MIT, 1992.

[19] H. Collewijn. Integration of Adaptive Changes of the Optokinetic Reflex, Pursuit and the Vestibulo-ocular Reflex. In A. Berthoz and G. Melvill Jones, editors, *Adaptive Mechanisms in Gaze Control.* New York: Elsevier, 1985.

[20] David Coombs and Christopher Brown. Real-Time Binocular Smooth Pursuit. *International Journal of Computer Vision*, 11(2):147--164, 1993.

[21] David Coombs, Ian Horswill, and Peter von Kaenel. Disparity Filtering : Proximity Detection and Segmentation. In *SPIE Vol. 1825 Intelligent Robots and Computer Vision XI*, pages 195--206, 1992.

[22] David John Coombs. *Real-time Gaze Holding in Binocular Robot Vision.* PhD thesis, Department of Compure Science, University of Rochester, June 1992.

[23] John R. Deller, John G. Proakis, and John H.L. Hansen. *Discrete-Time Processing of Speech Signals.* Macmillan Publishing Company, 1993.

[24] U. R. Dhond and J. K. Aggarwal. Structure from Stereo - A Review. *IEEE Transactions on Systems, Man, and Cybernetics*, 19(6), November/December 1989.

[25] R. Eckmiller. Neural Control of Foveal Pursuit Versus Saccadic Eye Movements in Primates - Single-unit Data and Models. *IEEE Transactions on Systems, Man, and Cybernetics*, 13:980--989, Sept./Oct. 1983.

[26] Cornelia Fermuller and Yannis Aloimonos. The Role of Fixation in Visual Motion Analysis. *International Journal of Computer Vision*, 11(2):165--186, 1993.

[27] David J. Fleet, Allan D. Jepson, and Michael R. M. Jenkin. Phase-Based Disparity Measurement. *Computer Vision Graphics and Image Processing : Image Understanding*, 53(2):198--210, March 1991.

[28] O. Fuentes, J. Karlsson, W. MEira, R. Rao, T. Riopk, J. Rosca, R. Surukkai, M. Van Wie, M. Zaki, T. Becker, R. Frank, B. Miller, and C. M. Brown. Mobile Robotics 1994. Technical Report TR 588, Computer Science Dept., U. Rochester, June 1995.

[29] Enrico Grosso and Dana H. Ballard. Head-Centered Orientation Strategies in Animate Vision. Technical Report TR 442, Computer Science Dept., U. Rochester, October 1992.

[30] Eric Huber and David Kortenkamp. Using Stereo Vision to Pursue Moving Agents with a Mobile Robot. In *Proc. 1995 IEEE Conference on Robotics and Automation*, 1995.

[31] Michael R. M. Jenkin, Allan D. Jepson, and John K. Tsotsos. Techniques for Disparity Measurement. *Computer Vision Graphics and Image Processing : Image Understanding*, 53(1):14--30, January 1991.

[32] Johnny Wai Yee Kam. A Real-Time 3D Motion Tracking System. Technical Report TR-93-16, Laboratory for Computational Intelligence, Department of Computer Science, The University of British Columbia, Vancouver, Canada, April 1993. http://www.cs.ubc.ca/tr/1993/TR-93-16.ps.gz.

[33] N. Kita, S. Rougeaux, Y. Kuniyoshi, and S. Sakane. Thorough ZDF-Based Localization for Binocular Tracking. In *Proc. of IAPR International Workshop MVA'94*, pages 190--195, 1994.

[34] Craig E. Kolb. Rayshade User's Guide and Reference Manual, 1992.

[35] D. Koller, K. Daniilidis, and H. H. Nagel. Model-Based Object Tracking in Monocular Image Sequences of Road Traffic Scenes. *International Journal of Computer Vision*, 10(3):257--281, 1993.

[36] Eric Krotkov and Ruzena Bajcsy. Active Vision for Reliable Ranging: Cooperating Focus, Stereo, and Vergence. *International Journal of Computer Vision*, 11(2):187--203, 1993.

[37] David G. Lowe. Robust Model-based Motion Tracking through the Integration of Search and Estimation. Technical Report TR-92-11, Laboratory for Computational Intelligence, Departm ent of Computer Science, The University of British Columbia, Vancouver, Canada, May 1992. Accepted for publication in the International Journal of Computer Vision. http://www.cs.ubc.ca/tr/1993/TR-92-11.ps.gz.

[38] Kai-Oliver Ludwig, Heiko Neumann, and Bernd Neumann. Local Stereoscopic Depth Estimation. *Image and Vision Computing*, 12(1):16--35, January/February 1994.

[39] R.K. Luneburg. *Mathematical Analysis of Binocular Vision*. Princeton University Press: Princeton, 1948.

[40] Bruce MacLennan. Gabor Representations of Spatiotemporal Visual Images. Technical Report CS-91-144, Computer Science Department, University of Tennessee, Knoxville, September 13 1991.

[41] Atsuto Maki. *Stereo Vision in Attentive Scene Analysis*. PhD thesis, Kungl Teknika Hogskolan, Stockholms Universitet, March 1996.

[42] Atsuto Maki, Lars Bretzner, and Jan-Olof Eklundh. Local Fourier phase and disparity estimates: An analytical study. Technical Report ISRN KTH/NA/P--95/16--SE, Computational Vision and Active Perception Laboratory(CVAP) Department of Numerical Analysis and Computing Science, Royal Institute of Tech nology, Stockholm, Sweden, September 1995. ftp://ftp.bion.kth.se/reports/cvap178. Shortened version in Proc. 6th International Conf. on Computer Analysis of Images and Patterns.

[43] Atsuto Maki and Tomas Uhlin. Disparity Selection in Binocular Pursuit. Technical Report ISRN KTH/NA/P--95/17--SE, Computational Vision and Active Perception Laborat ory(CVAP) Department of Numerical Analysis and Computing Science, Royal Institute of Technology, Stockholm, Sweden, September 1995. ftp://ftp.bion.kth.se/reports/cvap179.ps.Z. Also in IEICE(Institute of Electronics Information and Communication Engineers) Transactions on Information and Systems vol.E78-D, pp. 1591-1597 December 1995, and MVA'94 IAPR Workshop on Machine Vision Applications.

[44] Atsuto Maki, Tomas Uhlin, and Jan-Olof Eklundh. Disparity Selection for Vergence in Binocular Pursuit. In *4th IAPR International Workshop on Machine Vision Applications*, Kawasaki, Japan, 1994.

[45] David Marr. *Vision.* W. H. Freeman, San Francisco, 1982.

[46] F. Miles, S. Judge, and L. Optican. Optically Induced Changes in the Coupling Between Vergence and Accommodation. *J. Neuroscience*, 7&8:2576--2589, 1987.

[47] David W. Murray, Fenglei Du, Philip F. McLauchlan, Ian D. Reid, Paul M. Sharkey, and Michael Brady. Design of Stereo Heads. In Blake A and Yuille A, editors, *Active vision*, chapter 10. MIT, 1992.

[48] Don Murray and Anup Basu. Motion Tracking with an Active Camera. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(5):449--459, 1994.

[49] Randal C. Nelson. Qualitative Detection of Motion by a Moving Observer. *International Journal of Computer Vision*, 7(1):33--46, 1991.

[50] H. K. Nishihara. PRISM: A Practical Real-Time Imaging Stereo Matcher. Technical Report A.I. Memo No. 780, Massachusetts Institute of Technology, Artificial Intelligence Laboratory, May 1984.

[51] Peter Nordlund and Tomas Uhlin. Closing the Loop: Detection and Pursuit of a Moving Object by a Moving Observer. Technical Report ISRN KTH/NA/P--95/06--SE, Computational Vision and Active Perception Laboratory(CVAP) Department of Numerical Analysis and Computing Science, Royal Institute of Technology, Stockholm, Sweden, March 1995. ftp://ftp.bion.kth.se/reports/cvap173.ps.Z. Shortened version in Proc. 6th International Conf. on Computer Analysis of Images and Patterns. To appear in Image and Vision Computing.

[52] J. M. Odobez and P. Bouthemy. Detection of Multiple Moving Ojects Using Multiscale MRF with Camera Motion Compensation. In *1st IEEE Int. Conf. on Image Processing, ICIP*, Austin, 1994.

[53] Thomas J. Olson and David J. Coombs. Real-Time Vergence Control for Binocular Robots. *International Journal of Computer Vision*, 7(1):67--89, 1991.

[54] Thomas J. Olson and Robert J. Lockwood. Fixation-Based Filtering. In *SPIE Vol. 1825 Intelligent Robots and Computer Vision XI*, pages 685--695, 1992.

[55] Kourosh Pahlavan and Jan-Olof Eklundh. A Head-Eye System --- Analysis and Design. *Computer Vision Graphics and Image Processing : Image Understanding*, 56(1):41--56, July 1992.

[56] Kourosh Pahlavan, Tomas Uhlin, and Jan-Olof Eklundh. Fixation by Active Accommodation. In *SPIE Proceedings Vol. 1825, Intelligent Robots and Computer Vision XI: Algorithms*, pages 670--684, Boston, 1992.

[57] Kourosh Pahlavan, Tomas Uhlin, and Jan-Olof Eklundh. Dynamic fixation and active perception. *International Journal of Computer Vision*, 17(2):113--135, 1996.

[58] Swarup Reddi and George Loizou. Analysis of Camera Behavior During Tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(8):765--778, 1995.

[59] I. D. Reid and D. W. Murray. Active Tracking of Foveated Feature Clusters Using Affine Structure. http://www.robots.ox.ac.uk:5000/ lav/Papers/reid.murray.ijcv95.ps.gz. In press, IJCV, April 1996, 1995.

[60] D. A. Robinson. The Coordinates of Neurons in the Vestibulo-ocular Reflex. In A. Berthoz and G. Melvill Jones, editors, *Adaptive Mechanisms in Gaze Control*. New York: Elsevier, 1985.

[61] S. Rougeaux, N. Kita, Y. Kuniyoshi, S. Sakane, and F. Chavand. Binocular Tracking Based on Virtual Horopters. In *Proc. of IROS'94*, pages 2052--2057, 1994.

[62] P. M. Sharkey, D. W. Murray, S. Vandevelde, I. D. Reid, and P. F. McLauchlan. A Modular Head/Eye Platform for Real-Time Reactive Vision. *Mechatronics*, 3(4):517--535, 1993. http://www.robots.ox.ac.uk:5000/ lav/Papers/sharkey.etal.mech93.ps.gz.

[63] Philip W. Smith and N. Nandhakumar. An Accurate Stereo Correspondence Method for Textured Scenes Using Improved Power Cepstrum Techniques. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, New York, NY, June 1993.

[64] Phillip W. Smith and N. Nandhakumar. An Improved Power Cepstrum Based Stereo Correspondence Method for Textured Scenes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(3):338--348, March 1996.

[65] Michael J. Swain and Markus Stricker. Promising Directions in Active Vision. *International Journal of Computer Vision*, 11(2):109--126, 1993. University of Chicago, August 5-7, 1991.

[66] John R. Taylor, Thomas J. Olson, and Worthy N. Martin. Accurate Vergence Control in Complex Scenes. In *Proc. Computer Vision and Pattern Recognition*, pages 540--545, Seattle, June 1994.

[67] Massimo Tistarelli and Giulio Sandini. Dynamic Aspects in Active Vision. *Computer Vision Graphics and Image Processing : Image Understanding*, 56(1):108--129, July 1992.

[68] Sebastian Toelg. Gaze Control for an Active Camera System by Modeling Human Pursuit Eye Movements. In *SPIE Proceedings Vol. 1825, Intelligent Robots and Computer Vision XI: Algorithms*, pages 585--598, Boston, 1992.

[69] Tomas Uhlin and Kenneth Johhansson. Autonomous Mobile Systems : A Study of Current Research. Technical report, Computational Vision and Active Perception Laboratory(CVAP) De partment of Numerical Analysis and Computing Science, Royal Institute of Technology , Stockholm, Sweden, January 1996.

[70] Tomas Uhlin, Peter Nordlund, Atsuto Maki, and Jan-Olof Eklundh. Towards an Active Visual Observer. Technical Report ISRN KTH/NA/P--95/08--SE, Computational Vision and Active Perception Laborat ory(CVAP) Department of Numerical Analysis and Computing Science, Royal Institute of Technology, Stockholm, Sweden, March 1995. ftp://ftp.bion.kth.se/reports/cvap174.ps.Z. Shortened version in Proc. 5th International Conference on Computer Vision pp 679--686.

[71] Thierry Vievville, Emmanuelle Clergue, Reyes Enciso, and Herve Mathieu. Experimenting with 3D Vision on a Robotic Head. *Robotics and Autonomous Systems*, 14(1), Jan/Feb 1995.

[72] Carl-Johan Westelius. *Focus of Attention and Gaze Control for Robot Vision*. PhD thesis, Linkφping University, Computer Vision Laboratory, May 1995.

[73] Lambert E. Wixson. Gaze Selection for Visual Search. Technical Report TR 512 and Ph.D. Thesis, Computer Science Dept., U. Rochester, May 1994.

[74] Yalin Xiong and Steven A. Shafer. Variable Window Gabor Filters and Their Use in Focus and Correspondence. Technical Report CMU-RI-TR-94-06, The Robotics Institute, Carnegie Mellon University, March 1994. A short version of this technical report will appear in the Proceedings of Computer Vision and Pattern Recognition 1994.

[75] Yehezkel Yeshurun and Eric L. Schwartz. Cepstral Filtering on a Columnar Image Architecture : A Fast Algorithm for Binocular Stereo Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(7):759--767, 1989.

[76] Semir Zeki. *A Vision of the Brain*. Blackwell Scientific Publications, 1993.