Πανεπιστήμιο Κρήτης
Σχολή Θετικών Επιστημών
Τμήμα Επιστήμης Υπολογιστών

Προβολικές-Μεταθετικές Αναλλοίωτες ως Οπτικά Ορόσημα του Χώρου για Τοπολογική Πλοήγηση

Βασίλειος Σ. Τσώνης

Μεταπτυχιακή Εργασία

Ηράκλειο, Σεπτέμβριος 1998

© Copyright 1998 by Vassilios S. Tsonis
All Rights Reserved

Πανεπιστήμιο Κρήτης
Σχολή Θετικών Επιστημών
Τμήμα Επιστήμης Υπολογιστών

Προβολικές-Μεταθετικές Αναλλοίωτες ως Οπτικά Ορόσημα του Χώρου για Τοπολογική Πλοήγηση

Εργασία που υποβλήθηκε από τον
Βασίλειο Σ. Τσώνη
ως μερική εκπλήρωση των απαιτήσεων για την απόκτηση
ΜΕΤΑΠΤΥΧΙΑΚΟΥ ΔΙΠΛΩΜΑΤΟΣ ΕΙΔΙΚΕΥΣΗΣ

Συγγραφέας:

Βασίλειος Σ. Τσώνης
Τμήμα Επιστήμης Υπολογιστών
Πανεπιστήμιο Κρήτης

Εισηγητική Επιτροπή:

Πάνος Τραχανιάς,
Αναπληρωτής Καθηγητής, Επόπτης

Στέλιος Ορφανουδάκης,
Καθηγητής, Μέλος

Γιώργος Τζιρίτας,
Αναπληρωτής Καθηγητής, Μέλος

Δεκτή:

Πάνος Κωνσταντόπουλος
Πρόεδρος Επιτροπής Μεταπτυχιακών Σπουδών

Σεπτέμβριος 1998

στην οικογένειά μου

Προβολικές-Μεταθετικές Αναλλοίωτες ως Οπτικά Ορόσημα του Χώρου για Τοπολογική Πλοήγηση

Βασίλειος Σ. Τσώνης

Μεταπτυχιακή Εργασία

Τμήμα Επιστήμης Υπολογιστών
Πανεπιστήμιο Κρήτης

Περίληψη

Η πλοήγηση ενός αυτόνομου συστήματος αποτελεί ένα από τα πιο προκλητικά, ως προς τη δυσκολία, προβλήματα της ρομποτικής. Μία από τις προσεγγίσεις στο πρόβλημα αυτό βασίζεται στη χρήση ορόσημων, δηλαδή χαρακτηριστικών σημείων του χώρου που μπορούν σχετικά εύκολα να αναγνωριστούν από το αυτόνομο σύστημα και να χρησιμοποιηθούν για την αυτο-οριοθέτησή του στο χώρο. Ωστόσο, η πολυπλοκότητα του προβλήματος του ορισμού και της αναγνώρισης των ορόσημων, οδηγεί πολλές φορές στην παραδοχή ύπαρξης προηγούμενης γνώσης, τόσο για το χώρο κίνησης, όσο και για το είδος των αντικειμένων που θα χρησιμοποιηθούν σαν ορόσημα. Έτσι, η εφαρμογή μιας τέτοιας μεθόδου παρουσιάζει προβλήματα σε δυναμικά μεταβαλλόμενα περιβάλλοντα.

Στην εργασία αυτή προτείνουμε μία μέθοδο για πλοήγηση με χρήση οπτικών ορόσημων, όπου περιορίζουμε την ανάγκη ύπαρξης προηγούμενης γνώσης. Η πλοήγηση βασισμένη στα ορόσημα συνήθως στηρίζεται στην ανίχνευση και μετέπειτα αναγνώριση ενός συνόλου αντικειμένων του περιβάλλοντος, τα οποία κρίνονται επαρκή για την περιγραφή της δομής του χώρου εργασίας. Η διαδικασία αυτή είναι εκ φύσεως δύσκολη στην πράξη , λόγω των ποικίλων οριοθετήσεων που μπορεί να έχει ένα αντικείμενο, σε πειράματα πλοήγησης.

Για να παρακάμψουμε το πρόβλημα αυτό, προτείνουμε μια προσέγγιση που υιοθετεί την έννοια των προβολικών αναλλοίωτων, υπολογιζόμενες σε πεντάδες σημείων, τις οποίες θεωρούμε ορόσημα του χώρου εργασίας. Τέτοιες ποσότητες παραμένουν αναλλοίωτες κάτω από διαφορετικές θέσεις της κάμερας, καθιστώντας έτσι δυνατή την αυτόματη αναγνώριση τους από ένα κινούμενο παρατηρητή.

Για την ανίχνευση δυνητικών αντίστοιχων πεντάδων σε ακολουθίες εικόνων, εισάγουμε μια απλή τεχνική που βασίζεται στον πίνακα συνδιασποράς της κάθε πεντάδας. Με αυτό τον τρόπο, προσπερνάμε ικανοποιητικά το πρόβλημα του υπολογισμού της αντιστοίχισης σημείου με σημείο. Μιας και η άνωθι τεχνική υποδηλώνει αντιστοίχιση

μεταξύ πεντάδων και όχι μεταξύ ανεξάρτητων σημείων, υιοθετούμε την έννοια της μεταθετικά και προβολικά αναλλοίωτης για την αναγνώριση πεντάδων.

Πριν την διαδικασία της πλοήγησης, απαιτείται το στάδιο της εκμάθησης. Σ' αυτό το στάδιο, το αυτόνομο σύστημα, συλλέγει ορόσημα από το χώρο εργασίας. Τα ορόσημα αποθηκεύονται στη βάση δεδομένων του συστήματος, μαζί με πληροφορίες για τη σχετική θέση τους στο χώρο. Το αποτέλεσμα είναι η κατασκευή ενός τοπολογικού χάρτη του χώρου, που βασίζεται στα ορόσημα που εξείχθησαν, ο οποίος αποτελεί το σημείο αναφοράς του μηχανικού συστήματος σε μεταγενέστερα στάδια.

Κατά τη διάρκεια της πλοήγησης, το αυτόνομο σύστημα συγκρίνει τα ορόσημα που συναντά με αυτά που είναι αποθηκευμένα στη μνήμη του. Η αναγνώριση ενός ορόσημου του επιτρέπει να εντοπίσει τη θέση του στο χώρο και να καθορίσει την επόμενη κίνησή του.

Η προσέγγιση μας αποτιμήθηκε σε στατικές ακολουθίες εικόνων με εσωτερικούς διαδρόμους κτηρίων. Τα αποτελέσματα που παρατηρήθηκαν, είναι αρκετά ενθαρυντικά και επιβεβαιώνουν την καλή απόδοση της μεθόδου.

Επόπτης: Πάνος Τραχανιάς
Αναπληρωτής Καθηγητής Επιστήμης Υπολογιστών
Πανεπιστήμιο Κρήτης

Autonomous Landmark Based Navigation using Projective and Permutation Invariants

Vassilios S. Tsonis

Master of Science Thesis

Computer Science Department
University of Crete

Abstract

Autonomous robot navigation is one of the most challenging problems in robotics. A promising approach to this problem is based on the use of landmarks, i.e. distinctive visual events which can be easily recognized by the robot and used for its self-localization in space. However, the complexity of landmark identification and recognition usually leads to assumptions concerning the existence of previous knowledge regarding the robot workspace as well as the objects used as landmarks. Therefore, such an approach cannot be easily applied in dynamically changing environments.

In this thesis we propose a method for indoor navigation using visual landmarks that is not based on *a priori* knowledge of the environment. Landmark-based navigation usually relies on the identification and subsequent recognition of a number of environment objects, that are deemed adequate in describing the workspace structure. This process is inherently difficult in practice, due to the many different poses of an object that may be encountered in navigational trials.

In our method we overcome this problem by employing projective invariants computed on quintuples of points, which are considered as workspace landmarks. Such quantities remain invariant under different camera positions and provide for effective description of the workspace structure.

In order to identify potential corresponding quintuples in image frames, we introduce a simple test based on the covariance matrix estimate of each quintuple. With this test, we effectively by-pass the calculation of point correspondences. Since the above test indicates correspondence between quintuples, and not between their individual points, we subsequently employ a permutation insensitive projective invariant for quintuple recognition.

Prerequisite to the navigation procedure is a learning phase. In this phase the robot extracts landmarks from the workspace. These landmarks are stored in the robot's database, along with information regarding their relative position in space. As a result, a topological map of the workspace is constructed, based on the extracted landmarks. This

map is consulted later, during a navigation session, in order to decide the next motion(s). During navigation, the robot compares the landmarks that it encounters with the ones stored in memory. Successful recognition of a landmark allows the robot to approximately localize itself and select its next movement.

Our approach has been evaluated using image sequences from indoor corridors. The results obtained are quite promising and serve as an indication of the method's effectiveness.

Supervisor: Panos Trahanias
Associate Professor of Computer Science
University of Crete

Ευχαριστίες

Στο σημείο αυτό, θα ήθελα καταρχάς να ευχαριστήσω τον επόπτη καθηγητή μου κ. Πάνο Τραχανιά, για την πολύτιμη και ουσιαστική καθοδήγηση που μου παρείχε στη διάρκεια των μεταπτυχιακών μου σπουδών και την ουσιαστική του συμμετοχή στην περάτωση της εργασίας. Μαζί του κατάφερα να αποκτήσω πολλές και ποικίλες γνώσεις τόσο σε επιστημονικό όσο και σε ερευνητικό επίπεδο.

Επίσης, ευχαριστώ τα μέλη της εισηγητικής επιτροπής κ. Στέλιο Ορφανουδάκη και κ. Γιώργο Τζιρίτα για τις εποικοδομητικές τους παρατηρήσεις και διορθώσεις στην εργασία αυτή.

Θα ήταν μεγάλη παράλειψη να μην ευχαριστήσω τον ερευνητή του Ινστιτούτου Πληροφορικής του Ι.Τ.Ε κ. Κωνσταντίνο Χανδρινό, που αποτέλεσε τον τεχνικό επιμελητή της εργασίας μου, για την άψογη συνεργασία που είχαμε, τη σωστή του καθοδήγηση, το χιούμορ του, τις εύστοχες υποδείξεις του και τις γνώσεις, που απέκτησα κοντά του.

Ευχαριστώ το Τμήμα Επιστήμης Υπολογιστών του Π.Κ. για το σύγχρονο περιβάλλον εργασίας που μου παρείχε και τις γνώσεις που μου έδωσε την ευκαιρία να αποκτήσω, καθώς και το Ινστιτούτο Πληροφορικής του Ίδρυματος Τεχνολογίας και Έρευνας, και ειδικότερα το εργαστήριο Μηχανικής Όρασης και Ρομποτικής, για την υλικοτεχνική υποδομή που μου παρείχε κατά τη διάρκεια των σπουδών μου.

Θα ήθελα επίσης, να ευχαριστήσω τους ‘‘συνοδοιπόρους’’ φίλους μου για την πολύτιμη βοήθεια, τη συμπαράστασή τους και την ποικιλλία εμπειριών, που γευτήκαμε, σ’ αυτό το εξάχρονο ταξίδι. Δεν ξέρω αν τελικά φτάσαμε στην πολυπόθητη Ιθάκη. Το σίγουρο είναι πάντως ότι το ταξίδι θα μείνει χαραγμένο στον χάρτη της ψυχής μου, αναμένοντας καρτερικά τη συνέχειά του.

Το μεγαλύτερο όμως **ευχαριστώ** απ’ όλους το αξίζει η οικογένειά μου. Αν έφτασα σε αυτό το σημείο το χρωστάω σε αυτή και μόνο. Ως μια ελάχιστη ένδειξη της ευγνωμοσύνης μου για τους κόπους, τις αγωνίες και τις θυσίες της για τη μόρφωση και πνευματική μου καλλιέργεια, της αφιερώνω την εργασία αυτή.

Βασίλειος Σ. Τσώνης

Περιεχόμενα

Περίληψη **i**

Abstract **iii**

Ευχαριστίες **v**

Περιεχόμενα **vii**

Κατάλογος Σχημάτων **xi**

Κατάλογος Πινάκων **xv**

1 Εισαγωγή **1**

1.1 Επισκόπηση ερευνητικού έργου 2

1.1.1 Τριδιάστατη δομή και προγραμματισμός κίνησης 3

1.1.1.1 Εξαγωγή δομής από το πεδίο οπτικής ροής 3

1.1.1.2 Δομή γνωστή εκ των προτέρων 4

1.1.1.3 Προγραμματισμός κίνησης 5

1.1.2 Πλοήγηση με ορόσημα 7

1.2 Η δική μας προσέγγιση 11

1.2.1 Συνεισφορά 14

1.3 Οργάνωση της γραπτής εργασίας 14

2 Στοιχεία Προβολικής γεωμετρίας **17**

2.1 Ο προβολικός μετασχηματισμός 18

2.2 Μαθηματική σημειολογία τού προβολικού κόσμου 19

2.2.1 Βασικοί ορισμοί 19

2.2.2 Προβολικά δομικά στοιχεία 20

2.2.2.1 Η προβολική ευθεία 20

2.2.2.2 Το προβολικό επίπεδο 20

2.2.2.3 Ο προβολικός χώρος 22

2.3 Είδη προβολικών μετασχηματισμών 22

2.3.1 Προοπτικός μετασχηματισμός 22

2.3.2 Προβολικότητα 23
2.4 Ιεραρχία γεωμετριών και μετασχηματισμών 24
2.5 Αναλλοίωτες ποσότητες κάτω από προβολικούς μετασχηματισμούς 26
2.5.1 Μονοδιάστατοι cross-ratio 26
2.5.2 Δύο ευθείες - Δύο σημεία αναλλοίωτη 27
2.5.3 Δύο ομοεπίπεδα κωνοειδή 28
2.5.4 Διδιάστατοι cross ratios 28
2.5.5 Αριθμητική ευστάθεια των γεωμετρικών αναλλοιώτων 30
2.5.6 Συνόψιση των χαρακτηριστικών της προβολικής γεωμετρίας 30
2.6 Μεταθετικές και Προβολικές αναλλοίωτες 30
2.6.1 Πρώτης τάξης $\mathbf{P}^2$-αναλλοίωτες του μονοδιάστατου cross-ratio . . . 33
2.6.2 Δεύτερης τάξης αναλλοίωτες στη μία διάσταση 36
2.6.3 $\mathbf{P}^2$-αναλλοίωτες των διδιάστατων cross ratios 39
2.7 Εφαρμογές της προβολικών αναλλοίωτων στην Μηχανική Όραση 43

3 Εκμάθηση και Αναγνώριση Ορόσημων **45**
3.1 Στάδιο εκμάθησης οπτικών ορόσημων 47
3.1.1 Ορισμός και αναπαράσταση ενός οπτικού ορόσημου 47
3.1.2 Ο ανιχνευτής χαρακτηριστικών 48
3.1.2.1 Βασικές αρχές του ανιχνευτή χαρακτηριστικών 49
3.1.3 Εύρεση περιοχών αναζήτησης ορόσημων 52
3.1.3.1 Ανατομία του χάρτη προεξοχής1 53
3.1.4 Περιορισμοί των υπο-ορόσημων 56
3.1.4.1 Έλεγχος τοπολογικής συνέπειας 56
3.1.4.2 Έλεγχος συγγραμικότητας 57
3.1.4.3 Έλεγχος αν πέντε σημεία είναι ομοεπίπεδα 58
3.1.5 Ο αλγόριθμος εξαγωγής ορόσημων 61
3.1.5.1 Εύρεση χαρακτηριστικών σε κάθε περιοχή αναζήτησης . . 62
3.1.5.2 Αναζήτηση υπο-ορόσημων σε δύο διαδοχικά πλαίσια . . . 62
3.1.5.3 Κριτήριο αντιστοίχισης μεταξύ δύο πεντάδων 63
3.1.5.4 Επιλογή των πιο εξαιρετικών πεντάδων 67
3.1.5.5 Επιβεβαίωση ‘‘χρονικής συνέπειας’’ 68
3.1.6 Κατασκευή του τοπολογικού χάρτη 70
3.2 Αναγνώριση ορόσημων 72

4 Πειραματικά αποτελέσματα **75**
4.1 Συνθετικά αποτελέσματα 75
4.2 Αποτελέσματα σε πραγματικά περιβάλλοντα 77
4.2.1 Παράδειγμα I 77
4.2.1.1 Περίοδος εκμάθησης 78

4.2.1.2 Περίοδος αναγνώρισης 83
4.2.2 Παράδειγμα ΙΙ 87
4.2.2.1 Περίοδος εκμάθησης 87
4.2.2.2 Περίοδος αναγνώρισης 90
4.2.3 Παράδειγμα ΙΙΙ 92
4.2.3.1 Περίοδος εκμάθησης 92
4.2.3.2 Περίοδος αναγνώρισης 96

5 Επίλογος **99**
5.1 Μελλοντικές κατευθύνσεις 99

Α Ευρετήριο όρων **101**

Βιβλιογραφία **104**

Κατάλογος Σχημάτων

2.1 Γεωμετρική ερμηνεία του cross-ratio 27
2.2 Γεωμετρικές ερμηνείες των διδιάστατων cross-ratios. 30
2.3 Η αναλλοίωτη $-\mathrm{I}_2$ 36
2.4 Η αναλλοίωτη $|\mathrm{I}_2|$ 36
2.5 Η αναλλοίωτη J_1 37
2.6 Η αναλλοίωτη J_2 37
2.7 Η αναλλοίωτη J_{14} 38
2.8 Ένα παράδειγμα της μονοδιάστατης $\mathbf{P}^2$-αναλλοίωτης αναπαράστασης . . 38
2.9 Ορισμοί του cross-ratio με κοινό σημείο το A_1 39
2.10 Η εξίσωση $\mathcal{K}$, με βάση τη συνάρτηση J_{14} 42

3.1 Διάγραμμα ροής της μεθόδου 46
3.2 Διάγραμμα ροής του σταδίου εκμάθησης 47
3.3 Ανατομία ενός Οπτικού Ορόσημου 48
3.4 Σχηματική απεικόνιση πεντάδων ‘‘κακής’’ τοπολογικής συνέπειας 56
3.5 Αλγοριθμική περιγραφή της διαδικασίας ελέγχου της τοπολογικής συνέπειας μιας πεντάδας 57
3.6 Αλγοριθμική περιγραφή της διαδικασίας που ελέγχει αν μια πεντάδα περιέχει τρία συνευθειακά σημεία. 58
3.7 Αλγοριθμική περιγραφή της διαδικασίας ελέγχου αν πέντε σημεία είναι ομοεπίπεδα 61
3.8 Αλγοριθμική περιγραφή της διαδικασίας εύρεσης χαρακτηριστικών στην περιοχή αναζήτησης R 62
3.9 Η διαδικασία εύρεσης της αντιστοίχου μιας πεντάδας 66
3.10 Αλγοριθμος επιλογής εξαιρετικών πεντάδων 68
3.11 Ο χρόνος ζωής ενός υπο-ορόσημου 69
3.12 Αλγοριθμική περιγραφή της διαδικασίας επιβεβαίωσης της ‘‘χρονικής συνέπειας’’ μιας πεντάδας Q. 70
3.13 Παράδειγμα μιας πλειάδας περιγραφής 71
3.14 Διάγραμμα ροής του σταδίου αναγνώρισης ορόσημων 73

4.1 Αποτελέσματα της νόρμας του πίνακα συνδιασποράς 76
4.2 Η ακολουθία εισόδου στο πρώτο παράδειγμα 78
4.3 Ανιχνευθέντα χαρακτηριστικά στην αρχική ακολουθία 79
4.4 Περιοχή αναζήτησης ορόσημων 79
4.5 Οι δώδεκα πιο εξαιρετικές πεντάδες 80
4.6 Η πιο εξαιρετική πεντάδα 81
4.7 Διατήρηση της πεντάδας στην αρχική ακολουθία. 82
4.8 Το υπο-ορόσημο που αποτελεί το ορόσημο της ακολουθίας 83
4.9 Ζεύγος εικόνων εισόδου στη διαδικασία αναγνώρισης 83
4.10 Ανιχνευθέντα χαρακτηριστικά στο ζεύγος εισόδου 84
4.11 Περιοχή αναζήτησης ορόσημων 84
4.12 Οι εννέα πιο εξαιρετικές πεντάδες 85
4.13 Οι δύο επιλεχθείσες, πιο εξαιρετικές πεντάδες 86
4.14 Ακολουθία εικόνων στο δεύτερο παράδειγμα 87
4.15 Οι εννέα πιο εξαιρετικές πεντάδες 88
4.16 Διατήρηση της πεντάδας στην αρχική ακολουθία. 89
4.17 Το υπο-ορόσημο που αποτελεί το ορόσημο της ακολουθίας 90
4.18 Ακολουθία εισόδου στο δεύτερο παράδειγμα 90
4.19 Οι εννέα πιο εξαιρετικές πεντάδες 91
4.20 Η επιλεχθείσα, πιο εξαιρετική πεντάδα 92
4.21 Ακολουθία εικόνων στο τρίτο παράδειγμα 93
4.22 Οι δύο πιο εξαιρετικές πεντάδες που επιλέγουμε 93
4.23 Διατήρηση της πρώτης πεντάδας στην αρχική ακολουθία. 94
4.24 Διατήρηση της δεύτερης πεντάδας στην αρχική ακολουθία. 94
4.25 Τα δύο υπο-ορόσημα, που συνθέτουν το ορόσημο 95
4.26 Ακολουθία εικόνων στο τρίτο παράδειγμα 96
4.27 Οι δύο πιο εξαιρετικές πεντάδες που επιλέγουμε 97

Κατάλογος Πινάκων

2.1 Γεωμετρίες - Μετασχηματισμοί - Αναλλοίωτες ποσότητες 31

4.1 Οι τιμές της σωρευτικής απόστασης για τις δώδεκα πιο εξαιρετικές πεντάδες 81
4.2 Οι τιμές της αναλλοίωτης και της νόρμας για κάθε αντίστοιχη πεντάδα. 82
4.3 Η πλειάδα περιγραφής του υπο-ορόσημου στη βιβλιοθήκη 83
4.4 Οι τιμές της σωρευτικής απόστασης για τις εννέα πιο εξαιρετικές πεντάδες 85
4.5 Η πλειάδα περιγραφής της πρώτης πεντάδας 86
4.6 Η πλειάδα περιγραφής της δεύτερης πεντάδας 86
4.7 Οι τιμές της σωρευτικής απόστασης για τις εννέα πιο εξαιρετικές πεντάδες 88
4.8 Τιμές αναλλοίωτης, νόρμας και νόρμας διαφοράς για τις αντίστοιχες πεντάδες 89
4.9 Η πλειάδα περιγραφής του υπο-ορόσημου στη βιβλιοθήκη 90
4.10 Οι σωρευτικές αποστάσεις των εννέα πιο εξαιρετικών πεντάδων 91
4.11 Η πλειάδα περιγραφής της πρώτης πεντάδας 92
4.12 Αναλλοίωτες, νόρμες και νόρμες διαφοράς για την πρώτη πεντάδα 95
4.13 Αναλλοίωτες, νόρμες και νόρμες διαφοράς για την δεύτερη πεντάδα . . . 95
4.14 Η πλειάδα περιγραφής του πρώτου υπο-ορόσημου στη βιβλιοθήκη 96
4.15 Η πλειάδα περιγραφής του δεύτερου υπο-ορόσημου στη βιβλιοθήκη . . . 96
4.16 Η πλειάδα περιγραφής για την πρώτη πεντάδα 97
4.17 Η πλειάδα περιγραφής για τη δεύτερη πεντάδα 97

Κεφάλαιο 1

Εισαγωγή

Ένας από τους πιο φιλόδοξους στόχους της ρομποτικής, είναι η απόδοση ικανοτήτων *αντίληψης του χώρου (space perception)* σε *μηχανικά συστήματα (robots)*, προκειμένου να επιτελούν σύνθετες εργασίες, με αποτελεσματικότητα και ευελιξία σε οποιαδήποτε πιθανή κατάσταση. Η πλήρης κατανόηση των αρχών της αντίληψης του χώρου, αναμένεται να βοηθήσει σημαντικά στην επίλυση αρκετών προβλημάτων στους τομείς, της Μηχανικής Όρασης, της Τεχνητής Νοημοσύνης και της Μηχανικής Μάθησης. Ωστόσο, αν και η αντίληψη του χώρου αποτελεί μια στοιχειώδη νοητική διεργασία των έλλογων βιολογικών οργανισμών, εντούτοις οι αρχές στις οποίες στηρίζεται δεν έχουν πλήρως διασαφηνιστεί, λόγω της ποικιλίας και της πολυπλοκότητας των μορφών της, σε σχέση με το είδος και τις επιτελούμενες ενέργειες των διάφορων οργανισμών. Η παρούσα εργασία, αποτελεί μια μικρή συνεισφορά στην μακρόχρονη προσπάθεια κατανόησης των άνωθι αρχών μέσα από την μελέτη μιας από τις σπουδαιότερες λειτουργίες ενός αυτόνομου συστήματος, της αυτόνομης πλοήγησης (autonomous navigation). Η αυτόνομη πλοήγηση αναφέρεται στην ικανότητα ενός βιολογικού ή τεχνητού οργανισμού, να μετακινείται αυτόνομα και με επιτυχία στο περιβάλλον του, αξιολογόντας την πληροφορία, που αυτό του παρέχει. Προς την επίτευξη του γενικού αυτού σκοπού, έχουν αναπτυχθεί αρκετές ιδέες, που καλύπτουν ένα ευρύ φάσμα επιστημών[1] και εμπερικλύουν ένα πλούσιο σύνολο αισθητήρων,[2] προκειμένου να αντλήσουν και να επεξεργαστούν την πληροφορία του χώρου.

Σημαντικό χαρακτηριστικό της αυτόνομης πλοήγησης είναι ότι, δεν πρέπει να εμποδίζεται από αλλαγές στο περιβάλλον και να αποτυγχάνει σε πρωτόγνωρες καταστάσεις. Η ιδιότητα αυτή, απορρέει από τη μελέτη του ανθρώπινου τρόπου συμπεριφοράς [1], που αποτελεί, τις πιο πολλές φορές, τη βασικότερη πηγή αναφοράς για την άντληση ιδεών. Το χαρακτηριστικό αυτό, οδηγεί στο συμπέρασμα ότι, η αυτόνομη πλοήγηση είναι γενικότερο πρόβλημα από τον *προγραμματισμό κίνησης (motion planning)*, ο οποίος

[1] Αναπαράσταση γνώσης, Τεχνητή Νοημοσύνη, Μηχανική Μάθηση, Νευροφυσιολογία, Φυσική, κ.ά

[2] οπτικοί αισθητήρες, αισθητήρες υπερήχων (sonar), αφής, κ.ά.

σχετίζεται με τον υπολογισμό μιας ακολουθίας κινήσεων, που θα οδηγήσει στη μετακίνηση ενός μηχανικού συστήματος στο χώρο, με βάση μια αναπαράσταση του τελευταίου που παρέχεται από άλλες διαδικασίες. Αντίθετα, η αυτόνομη πλοήγηση, αντιμετωπίζει συνολικά το πρόβλημα της απόκτησης μιας αναπαράστασης του χώρου μέσω κάποιων αισθητήρων και στη μετέπειτα χρήση της για τον προγραμματισμό της κίνησης.

Στην εργασία αυτή, επικεντρωνόμαστε στην *οπτική πλοήγηση (visual navigation)*, στην ικανότητα δηλαδή ενός συστήματος να κινείται αυτόνομα με βάση τις πληροφορίες που του παρέχει μια ακολουθία εικόνων. Στην απόφαση αυτή, μας οδήγησε ο μεγάλος πλούτος πληροφοριών που μπορούν να εξαχθούν μέσω της όρασης και η μεγάλη διάδοσή της στη φύση. Το ενδιαφέρον μας εστιάζεται στην ανάπτυξη τεχνικών πλοήγησης κατάλληλων για χρήση σε δυναμικά ή/και άγνωστα περιβάλλοντα, προσπαθώντας να διατηρήσουμε την ευελιξία και την γενικότητα της αυτόνομης πλοήγησης.

Πολλές και σημαντικές εφαρμογές, μπορούν να γίνουν πραγματικότητα, μέσω της κατασκευής αυτόνομων μηχανικών συστημάτων. Ένα αυτόνομο μηχανικό σύστημα, που κινείται με ασφάλεια και επιτυχία στο χώρο εργασίας του, μπορεί να αποτελέσει ένα πολύτιμο βοήθημα σε άτομα με ειδικές ανάγκες[3], ένα ακούραστο φύλακα που φροντίζει για την ασφάλεια και παρακολούθηση των χώρων, ένα πρόθυμο συνεργάτη σε εργασίες ρουτίνας[4], ένα παράτολμο εξερευνητή επικίνδυνων, για τον άνθρωπο, περιβάλλοντων, κ.ά. Εκτός όμως από την πρακτική πλευρά του θέματος, σημαντικό είναι και το θεωρητικό ενδιαφέρον του. Πιο συγκεκριμένα, το πρόβλημα της αυτόνομης πλοήγησης, ορίζεται σε σχέση με το περιβάλλον δράσης, με αποτέλεσμα να αποτελεί ένα ειδικό πρόβλημα που πρέπει να λυθεί κάτω από γενικές συνθήκες. Οι μηχανισμοί πλοήγησης μπορούν να οργανωθούν σε μια ιεραρχία με βάση το βαθμό πολυπλοκότητας τους και να μελετηθούν μέσω μιας συνθετικής προσέγγισης, κατά την οποία ένα σύστημα ξεκινά με πολύ απλές ικανότητες και προοδευτικά αποκτά νέες με μεγαλύτερη πολυπλοκότητα, οι οποίες στηρίζονται στις ήδη υπάρχουσες. Μερικές από τις συνιστώσες της ικανότητας πλοήγησης ενός συστήματος είναι η ικανότητα εκτίμησης των παραμέτρων κίνησής του (egomotion estimation) καθώς και της σχετικής κίνησής του ως προς το σύστημα ενός κινούμενου αντικειμένου (independent motion estimation) [2, 3], η ικανότητα αποφυγής εμποδίων (obstacle avoidance) και η ικανότητα μετακίνησης του από σημείο σε σημείο (homing). Ένα τέτοιο σύστημα θα πρέπει να είναι σε θέση να εξάγει και να ‘‘απομνημονεύει’’ κάποια χαρακτηριστικά του περιβάλλοντος και αργότερα να τα χρησιμοποιεί για την μετακίνηση και την οριοθέτησή του μέσα σε αυτό.

1.1 Επισκόπηση ερευνητικού έργου

Μέχρι σήμερα, η πλειοψηφία των μελετών σχετικά με την οπτική πλοήγηση είναι επηρεασμένες από την ανακατασκευαστική θεωρία του Marr [4]. Δεδομένου ότι η

[3] Μια ‘‘έξυπνη’’ αναπηρική πολυθρόνα, που διευκολύνει τη μετακίνηση ανθρώπων με κινητικές αδυναμίες
[4] π.χ. καθαρισμός κτιρίων

θεωρία αυτή διαχωρίζει τις διαδικασίες της αντίληψης από αυτές της σκέψης, οι μελέτες αυτές ασχολούνται με την επίλυση δύο ξεχωριστών προβλημάτων: το πρώτο είναι ο υπολογισμός μιας τριδιάστατης, γεωμετρικής περιγραφής του χώρου μέσω μιας ακολουθίας εικόνων που καταγράφει ένας παρατηρητής που κινείται ως προς το περιβάλλον του και το δεύτερο η χρήση αυτής της περιγραφής για τον υπολογισμό μιας ακολουθίας κινήσεων που θα επιτύχει την μετακίνηση ενός μηχανικού συστήματος στο χώρο. Το πρώτο πρόβλημα είναι γνωστό σαν *δομή από κίνηση (structure from motion)* και το δεύτερο σαν πρόβλημα *προγραμματισμού της κίνησης (motion planning)*.

Μία εναλλακτική προσέγγιση είναι βασισμένη στην έννοια των *ορόσημων (landmarks)*. Με τον όρο ορόσημο εννοούμε κάποιο ιδιαίτερο σύνολο χαρακτηριστικών του κόσμου, που ξεχωρίζει από το υπόλοιπο περιβάλλον και είναι σχετικά εύκολο να αναγνωριστεί από το αυτόνομο σύστημα. Στην προσέγγιση αυτή, η γεωμετρική περιγραφή του χώρου δεν είναι απαραίτητη. Οι σχετικές θέσεις των ορόσημων χρησιμοποιούνται, κατά την εκμάθηση του χώρου, για τη δημιουργία ενός τοπολογικού χάρτη που τον περιγράφει. Έτσι, ένα ορόσημο μπορεί να χρησιμοποιηθεί για να αντληθούν πληροφορίες για τη σχετική θέση του αυτόνομου συστήματος μέσα στο χώρο καθώς και για τις πιθανές πλοηγητικές ενέργειες που μπορεί να επιτελέσει.

Στο τμήμα αυτό της εργασίας γίνεται μία επισκόπηση των μέχρι τώρα ερευνητικών προσπαθειών στον τομέα της οπτικής πλοήγησης και των επιμέρους χαρακτηριστικών της, που είναι και το θέμα της συγκεκριμένης εργασίας.

1.1.1 Τριδιάστατη δομή και προγραμματισμός κίνησης

Χαρακτηριστικό αυτής της προσέγγισης είναι ότι στοχεύει σε μία επακριβή αναπαράσταση του χώρου και χρησιμοποιεί αποκλειστικά μετρικές μεθόδους για τον προγραμματισμό κίνησης. Η λεπτομερής περιγραφή της δομής του κόσμου, με μεθόδους όπως αυτή του πεδίου οπτικής ροής, αποδεικνύεται δύσκολο πρόβλημα. Έτσι, πολλές προσεγγίσεις έχουν θεωρήσει τη δομή του κόσμου γνωστή εκ των προτέρων, σε μικρό ή μεγάλο βαθμό.

1.1.1.1 Εξαγωγή δομής από το πεδίο οπτικής ροής

Μία κλασική προσέγγιση στο πρόβλημα της εξαγωγής τριδιάστατης δομής είναι η ανάλυση της κίνησης μέσω ακολουθιών εικόνων με χρήση του πεδίου οπτικής ροής. Η προσέγγιση αυτή περιλαμβάνει δύο στάδια:

1. Αρχικά υπολογίζεται η μετατόπιση των προβολών ενός συνόλου χαρακτηριστικών, σε μια ακολουθία εικόνων, η οποία οφείλεται στην κίνηση του παρατηρητή ως προς αυτά. Η μετατόπιση υπολογίζεται μέσω της *αντιστοίχισης (correspondence)* των χαρακτηριστικών της τρέχουσας εικόνας, σε διαδοχικές εικόνες. Στην περίπτωση μικρής κίνησης, η ακολουθία των εικόνων θεωρείται μία συνάρτηση

δύο χωρικών μεταβλητών[5] και μίας χρονικής[6]. Οι χωρικές και χρονικές μερικές παράγωγοι της συνάρτησης αυτής σε συνδυασμό με κάποιες επιπλέον υποθέσεις σχετικά με την κίνηση, συνθέτουν το διάνυσμα της ταχύτητας κάθε σημείου στο επίπεδο της εικόνας. Το διανυσματικό πεδίο που προκύπτει είναι γνωστό σαν πεδίο οπτικής ροής (optical flow field) [5].

2. Στη συνέχεια, υπολογίζεται η τριδιάστατη ταχύτητα και η δομή της σκηνής που περιέχεται στην ακολουθία των εικόνων, μέσω των εξισώσεων που συσχετίζουν τη διδιάστατη ταχύτητα των σημείων της εικόνας με την τριδιάστατη κίνηση των σημείων του κόσμου που απεικονίζονται σ' αυτά.

Το πρόβλημα της δομής από κίνηση έχει μελετηθεί διεξοδικά σε θεωρητικό επίπεδο [6, 7, 8, 9, 10, 11, 12]. Είναι γνωστά τα αποτελέσματα που αφορούν το είδος των πληροφοριών σχετικά με τη δομή και την κίνηση των σημείων μιας σκηνής που μπορούν να εξαχθούν από μια ακολουθία εικόνων [13], τις συνθήκες που εξασφαλίζουν την μοναδικότητα των πληροφοριών αυτών και τη συμπεριφορά των μεθόδων αυτών στην περίπτωση ύπαρξης θορύβου [12, 14]. Επίσης, έχουν σχεδιαστεί επαναληπτικοί [15, 16, 17, 18] και αναδρομικοί [19, 20, 21, 22, 23, 24] αλγόριθμοι για τον υπολογισμό της δομής και της κίνησης μιας σκηνής. Παρά τις προσπάθειες αρκετών ετών όμως, μέχρι σήμερα δεν έχει καταστεί δυνατή η κατασκευή ενός συστήματος ικανού να αντιμετωπίζει με επιτυχία το πρόβλημα της δομής από την κίνηση [25].

Τα αίτια της μη ικανοποιητικής απόδοσης βρίσκονται στην ανακατασκευαστική προσέγγιση που έχει υιοθετηθεί. Πιο συγκεκριμένα, στη δυσκολία επακριβούς εκτίμησης της μετατόπισης των χαρακτηριστικών, λόγω των, ελλειπών να εξασφαλίσουν μοναδική λύση, γεωμετρικών περιορισμών της κάμερας και των αντικειμένων που απεικονίζονται στο επίπεδο της. Προκειμένου να επιλυθεί το πρόβλημα μονοσήμαντα, εισάγονται επιπλέον περιορισμοί[7], που όμως δεν είναι πάντα ρεαλιστικοί, βασισμένοι σε ειδικές υποθέσεις σχετικά με τα αντικείμενα που απεικονίζονται στις εικόνες. Έτσι, αν και λύνουμε το πρόβλημα μονοσήμαντα, τα αποτελέσματα δεν είναι ικανοποιητικά. Επίσης, η διαδικασία υπολογισμού των παραμέτρων κίνησης στο χώρο, με βάση τη μετατόπιση ή την ταχύτητα των σημείων στο επίπεδο της εικόνας, είναι πολύ ευαίσθητη στο θόρυβο και κατά συνέπεια αριθμητικά ασταθής. Τέλος, σημαντικό μειονέκτημα αποτελεί και η υπολογιστική πολυπλοκότητα που παρουσιάζει το πρόβλημα της αντιστοίχισης, αφού θα πρέπει να εφαρμοστούν κριτήρια ομοιότητας σε μεγάλα τμήματα της εικόνας.

1.1.1.2 Δομή γνωστή εκ των προτέρων

Λόγω των μειονεκτημάτων της εξαγωγής δομής από το πεδίο οπτικής ροής, αρκετοί ερευνητές έχουν προσεγγίσει το πρόβλημα διαφορετικά, θεωρώντας ότι κάποιο μοντέλο

[5] οι δύο διαστάσεις της εικόνας
[6] η χρονική ακολουθία των εικόνων
[7] π.χ. ομαλότητα του πεδίου οπτικής ροής

του κόσμου είναι γνωστό εκ των προτέρων (a priori knowledge). Η βασική στρατηγική συνίσταται στην αντίληψη του περιβάλλοντος από το σύστημα μέσω αισθητήρων και τη σύγκριση της συλλεχθείσης πληροφορίας με την προϋπάρχουσα γνώση για τον κόσμο. Η επιθυμητή *ταύτιση (matching)*, οδηγεί στην προσεγγιστική εκτίμηση της θέσης του συστήματος και κατ' επέκταση σε μια περιγραφή του άμεσου περιβάλλοντος. Το πρόβλημα της περιγραφής του περιβάλλοντος σε αυτή την προσέγγιση ισοδυναμεί με το πρόβλημα του εντοπισμού της θέσης του συστήματος (self-localization), αφού δοθείσης της θέσης και της, προηγούμενα καταγραφόμενης, περιγραφής του κόσμου, μπορούμε να εξάγουμε συμπεράσματα (reasoning) για την αναμενόμενη όψη του κόσμου από τη συγκεκριμένη θέση. Η τεχνική μπορεί να λειτουργήσει και αντίστροφα, χρησιμοποιώντας τις πληροφορίες που συλλέγουν οι αισθητήρες για να περιγράψουμε πληρέστερα τον κόσμο. Το μοντέλο του κόσμου μπορεί να αναπαρασταθεί με χάρτες [26, 27, 28, 29, 30] ή με γεωμετρικές δομές δεδομένων. Για παράδειγμα, ο Zelinsky [31], χρησιμοποιώντας αισθητήρες αφής, περιγράφει το περιβάλλον μέσω ενός τετραδικού δένδρου (quadtree), ενώ οι Iyengar et al. [32] χρησιμοποιούν γράφους που αναπαριστούν τον ελεύθερο χώρο του πεδίου δράσης. Στις εργασίες αυτές έχουν αναπτυχθεί μέθοδοι για την εκτίμηση της θέσης του αυτόνομου συστήματος, με βάση το γνωστό μοντέλο του κόσμου [33, 34]. Κατόπιν, το πρόβλημα της πλοήγησης αντιμετωπίζεται με μετρικές μεθόδους.

Μία ανάλογη τεχνική χρησιμοποιήθηκε από τους Kak et al. [35, 36] στη δημιουργία του αυτόνομου συστήματος πλοήγησης FINALE, όπου είναι γνωστή μία τριδιάστατη περιγραφή του χώρου, εκ της οποίας προκύπτει η όψη του κόσμου, όπως θα φαινόταν από μία υποθετική θέση του αυτόνομου συστήματος. Το αυτόνομο σύστημα προσπαθεί να εκτιμήσει την τρέχουσα θέση του στο χώρο, λαμβάνοντας υπόψιν του την προηγούμενη καθώς και τις εντολές κίνησης που έλαβαν χώρα ενδιάμεσα. Όπως γίνεται αντιληπτό, η εκτίμηση δεν μπορεί να είναι ακριβής λόγω των ατελειών και των σφαλμάτων στα συστήματα κίνησης. Με βάση αυτή την εκτίμηση και το γνωστό μοντέλο του κόσμου, προκύπτει μία αναμενόμενη όψη του περιβάλλοντος, η οποία συγκρίνεται με την εικόνα που λαμβάνουν οι κάμερες του συστήματος. Η επιτυχής ταύτιση των εικόνων οδηγεί σε μια, μειωμένης αβεβαιότητας, εκτίμηση της θέσης του συστήματος. Το ίδιο μονοπάτι ακολουθεί και η εργασία των Tsubouchi et al. [37], η οποία βασίζεται κυρίως στην πληροφορία που περιέχει το χρώμα των εικόνων. Όπως υποστηρίζουν, το χρώμα της εικόνας δεν επηρεάζεται από τη φωτεινότητα ή τους πιθανούς τύπους φωτισμού και σκίασης, που μπορεί να συναντηθούν σε μια σκηνή.

1.1.1.3 Προγραμματισμός κίνησης

Ο προγραμματισμός κίνησης ενός μηχανικού συστήματος, δεδομένης μίας γεωμετρικής περιγραφής του περιβάλλοντος, συνίσταται στον υπολογισμό ενός ελεύθερου μονοπατιού διαμέσου των εμπόδιων, ανάμεσα σε μια δεδομένη αρχική θέση και μια δεδομένη τελική [38, 39]. Το πρόβλημα του υπολογισμού αυτού του μονοπατιού, είναι γνωστό σαν

''piano movers problem'' ή ''generalised movers problem'' [40]. Οι μέθοδοι που έχουν προταθεί, αποσκοπούν στη μοντελοποίηση του ελεύθερου χώρου στο χώρο δράσης.

Στις προσεγγίσεις, που υιοθετούν την τεχνική του *σκελετού (skeleton)*, το σύνολο των εφικτών[8] θέσεων του συστήματος, απεικονίζεται σε ένα δίκτυο. Κατά συνέπεια, ο υπολογισμός μιας ακολουθίας κινήσεων ανάγεται σε ένα πρόβλημα αναζήτησης σε γράφους. Στην κατηγορία αυτή ανήκουν μέθοδοι που χρησιμοποιούν δομές όπως οι *γράφοι ορατότητας (visibility graphs)* και τα *διαγράμματα Voronoi (Voronoi diagrams)* [41, 42, 43, 44]. Στην ίδια κατηγορία, ανήκει και η μέθοδος του *διαχωρισμού θέσεων (cell decomposition)* [45, 46], στην οποία, ο ελεύθερος χώρος απεικονίζεται με ένα σύνολο απλών γεωμετρικών σχημάτων (π.χ. παραλληλογράμμων), για τα μέλη του οποίου υπολογίζονται σχέσεις γειτνίασης. Ο υπολογισμός μιας ακολουθίας κινήσεων ανάγεται στην αναζήτηση μιας ακολουθίας, ανά δύο γειτονικών, σχημάτων, που θα ενώνει τα σχήματα που αντιστοιχούν στην αρχική και την τελική θέση του συστήματος. Οι μέθοδοι της κατηγορίας βασίζονται σε τεχνικές δανεισμένες από την υπολογιστική γεωμετρία [47]. Ένα άλλο κοινό χαρακτηριστικό τους, είναι ότι απαιτούν κάποιας μορφής αναζήτηση στη δομή που παριστάνει τον ελεύθερο χώρο. Κλασικές τεχνικές αναζήτησης σε γράφους, όπως π.χ. depth-first, breadth-first, best-first, A^*, χρησιμοποιούνται για το σκοπό αυτό. Συχνά επίσης εφαρμόζονται τεχνικές ελαχιστοποίησης, όπως ο αλγόριθμος του Dijkstra, μέθοδοι simulated annealing, κ.ά.

Στη μέθοδο του *δυναμικού πεδίου (potential field)*, κάθε αντικείμενο του περιβάλλοντος ασκεί απωστικές δυνάμεις στα σημεία του χώρου που το περιβάλλουν, το μέτρο των οποίων εξαρτάται από την απόσταση των σημείων από αυτό. Ο στόχος, η επιθυμητή τελική θέση του συστήματος δηλαδή, ασκεί αντίστοιχες ελκτικές δυνάμεις [48, 49]. Έτσι, σε κάθε σημείο του χώρου εργασίας αντιστοιχεί ένα διάνυσμα το οποίο υπολογίζεται ως η συνισταμένη των ελκτικών και απωστικών δυνάμεων στο σημείο αυτό, ορίζοντας ένα διανυσματικό πεδίο. Το μέτρο της δύναμης είναι ελάχιστο στο σημείο που αντιστοιχεί στο στόχο και αυξάνεται με την απομάκρυνση από αυτόν. Υπολογίζοντας την *κλίση (gradient)* σε κάποιο σημείο του χώρου, ένα σύστημα μπορεί να επιλέξει τη διεύθυνση κατά την οποία πρέπει να κινηθεί ώστε να φτάσει στο στόχο του όσο γίνεται συντομότερα. Το βασικότερο πρόβλημα με την προσέγγιση αυτή, είναι ότι η διανυσματική συνάρτηση που υπολογίζεται παρουσιάζει αρκετά τοπικά ελάχιστα, τα οποία είναι πιθανό να *παγιδεύσουν* το σύστημα.

Τέλος, στην μέθοδο του *μαθηματικού προγραμματισμού (mathematical programming)*, το πρόβλημα υπολογισμού μιας ακολουθίας κινήσεων, αντιμετωπίζεται σαν πρόβλημα βελτιστοποίησης μιας συνάρτησης, κάτω από περιορισμούς που σχετίζονται με τα αντικείμενα του περιβάλλοντος. Μιας και οι περιορισμοί αυτοί καθώς και η προς βελτιστοποίηση συνάρτηση είναι μη γραμμικοί, συνήθως χρησιμοποιούνται αριθμητικές μέθοδοι για την εύρεση της, προσεγγιστικά, βέλτιστης λύσης [50].

Βασικό πλεονέκτημα των παραπάνω προσεγγίσεων είναι το γεγονός ότι τα μονο-

[8] εκείνων δηλαδή που δεν παρεμποδίζονται από αντικείμενα του περιβάλλοντος

πάτια που υπολογίζονται είναι βέλτιστα σύμφωνα με κάποια κριτήρια όπως είναι το συνολικό τους μήκος, η απόσταση του μηχανικού συστήματος από τα εμπόδια κ.ά. Η προϋπόθεση όμως ύπαρξης μιας ακριβούς αναπαράστασης του περιβάλλοντος είναι εξωπραγματική, αφού συνήθως, η όποια γνώση για το περιβάλλον είναι ελλειπής και χαρακτηρίζεται από αβεβαιότητα λόγω ανακριβών μετρήσεων των αισθητήρων, απρόβλεπτων αλλαγών του περιβάλλοντος και ατελή έλεγχο του συστήματος. Επιπλέον, ο διαχωρισμός του υπολογισμού μιας ακολουθίας κινήσεων από την εκτέλεση τους, καθώς και το υψηλό υπολογιστικό κόστος, καθιστά τις μεθόδους αυτές ανεπαρκείς σε δυναμικά περιβάλλοντα. Ακόμη, ο καθαρά μετρικός προγραμματισμός κίνησης είναι ευάλωτος όταν υπάρχει ελλιπής μηχανικός έλεγχος του συστήματος. Πολλά μικρά λάθη στις κινήσεις του αυτόνομου συστήματος συσσωρευμένα, μπορεί να προκαλέσουν μία εντελώς λανθασμένη εκτίμηση της θέσης του [34, 51]. Ένα άλλο πρόβλημα με αρκετές από τις αναφερθείσες προσεγγίσεις είναι ότι δεν είναι επεκτάσιμες σε γενικούς και πολύπλοκους χώρους εργασίας, αφού οι απαιτούμενοι υπολογισμοί και οι αποθηκευτικές απαιτήσεις για την αναπαράσταση του κόσμου αυξάνουν με γεωμετρική πρόοδο. Τα παραπάνω, περιορίζουν τη χρησιμότητα των μεθόδων αυτών σε πολύ ειδικές περιπτώσεις, όπου το περιβάλλον είναι στατικό και ελεγχόμενο ενώ η δομή του είναι πλήρως γνωστή εκ των προτέρων, όπως, για παράδειγμα, ο έλεγχος βιομηχανικών ρομπότ σε μια γραμμή παραγωγής.

Η αδυναμία κατασκευής μιας κατάλληλης αναπαράστασης του κόσμου ικανής να υποστηρίξει αλγόριθμους προγραμματισμού κίνησης, όπως αυτοί που παρουσιάστηκαν πιο πάνω, καθώς και η αδυναμία εφαρμογής των τελευταίων σε γενικής φύσεως περιβάλλοντα, αποτέλεσαν το κίνητρο για την εμφάνιση πιο ευέλικτων μεθόδων. Οι μέθοδοι αυτές δεν απαιτούν μία ακριβή και λεπτομερή αναπαράσταση του κόσμου, αλλά στηρίζονται σε διαδικασίες *επιλεκτικής* εκμάθησης των χαρακτηριστικών του κόσμου. Υιοθετώντας τις αρχές της *ενεργού και τελεολογικής όρασης (active and purposive vision)* [52, 53, 54, 55], χρησιμοποιούν μηχανισμούς εστίασης της προσοχής δίνοντας προτεραιότητα στις πιο ‘‘χαρακτηριστικές’’ περιοχές του περιβάλλοντος, στις οποίες επιτελείται η αναζήτηση των ορόσημων. Ο εντοπισμός των ορόσημων από το αυτόνομο σύστημα, μπορεί να συμβάλλει στη μείωση της αβεβαιότητας εκτίμησης της θέσης του συστήματος [56] και στην *ποιοτική* πλοήγησή[9] του. Τα ζητήματα ωστόσο που προκύπτουν από αυτή την προσέγγιση είναι πολλά και αφορούν τον τρόπο επιλογής των κατάλληλων ορόσημων, το είδος τους, τα απαραίτητα στοιχεία για την περιγραφή τους και τους τρόπους αναγνώρισης από το αυτόνομο σύστημα.

1.1.2 Πλοήγηση με ορόσημα

Το πρόβλημα της επιλογής και αναγνώρισης ορόσημων στο χώρο, με στόχο κυρίως την πλοήγηση ρομπότ, απασχολεί τους ερευνητές από τα μέσα της δεκαετίας του ’80. Λόγω

[9]Για παράδειγμα, στο ορόσημο Α στρίψε δεξιά

όμως της μεγάλης του δυσκολίας στη γενική περίπτωση, το πρόβλημα αυτό συνήθως παρακάμπτεται ή ακόμη δεν αντιμετωπίζεται καθόλου, θεωρώντας ως δεδομένες κάποιες γεωμετρικές αναπαραστάσεις του χώρου που επιτρέπουν τον προγραμματισμό κίνησης του μηχανικού συστήματος [38, 41].

Σε ελάχιστες περιπτώσεις, η αντίληψη του χώρου μέσω οπτικών ορόσημων έχει προσεγγισθεί κάνοντας χρήση *πολύ απλών ορόσημων*. Ο λόγος γι' αυτό είναι η προφανής δυσκολία του προβλήματος όσο αναφορά την επιλογή ρεαλιστικών ορόσημων [57, 58]. Στη συνέχεια περιγράφονται συνοπτικά οι βασικότερες προσεγγίσεις που έχουν αναφερθεί στη διεθνή βιβλιογραφία.

- *Προσχεδιασμένα ορόσημα (predesigned landmarks)*: Κατά την προσέγγιση αυτή, το πρόβλημα της επιλογής και στη συνέχεια της αναγνώρισης των ορόσημων στην ουσία παρακάμπτεται, μιας και ο χώρος εμπλουτίζεται με προκαθορισμένα πρότυπα (patterns) που διαφέρουν σαφώς μεταξύ τους καθώς και με το υπόλοιπο περιβάλλον.

  Στην κατηγορία αυτή ανήκει η εργασία των Taylor et al. [59], οι οποίοι χρησιμοποιούν σαν ορόσημα bar codes αποτελούμενα από πέντε κάθετες λωρίδες και χρωματισμένα με ένα από τα τρία επιτρεπτά χρώματα. Η αναγνώριση ενός ορόσημου γίνεται εύκολα με την εξαγωγή των κάθετων ακμών της εικόνας και τα χρώματα χρησιμοποιούνται για τον διαχωρισμό των διαφορετικών ορόσημων. Στην εργασία αυτή προτείνεται ένας αλγόριθμος ''εξερεύνησης'' του χώρου, όπου με βάση τα γνωστά ορόσημα και την πληροφορία που συλλέγεται μέσω ένος αισθητήρα υπερήχων, δημιουργείται ένας γράφος που αναπαριστά τα εμπόδια του περιβάλλοντος. Ο γράφος αυτός μπορεί να χρησιμοποιηθεί στη συνέχεια για την πλοήγηση του αυτόνομου συστήματος. Επίσης, οι Lazanas et al. [56, 60] χρησιμοποιούν σαν ορόσημα απλά ασπρόμαυρα σχέδια, τοποθετημένα στην οροφή του χώρου κίνησης του αυτόνομου συστήματος. Αξίζει να σημειωθεί ότι τα ορόσημα είναι τοποθετημένα σε θέσεις γνωστές εκ των προτέρων, στο επίσης γνωστό μοντέλο του κόσμου. Γνωστές είναι επίσης και οι θέσεις των αντικειμένων που ενδέχεται να παρεμποδίσουν το αυτόνομο σύστημα στην κίνησή του. Με βάση όλη αυτή τη γνώση, οι ερευνητές παρουσιάζουν ένα αλγόριθμο πλοήγησης που είναι ανεκτικός σε λάθη, αφού ήδη τα έχουν προβλέψει, καθώς και ατέλειες του μηχανικού ελέγχου του συστήματος. Την ίδια οδό ακολουθούν στην εργασία τους και οι Magge et al. [61], που ασχολούνται με το πρόβλημα του οριοθέτησης του αυτόνομου συστήματος στο χώρο. Για να το αντιμετωπίσουν, χρησιμοποιούν σαν ορόσημο μία σφαίρα, που έχει σχεδιασμένους πάνω της δύο (κάθετους) κύκλους. Το μέγεθος της προβολής της σφαίρας στην εικόνα παρέχει την απόσταση από τη σφαίρα, ενώ η μετατόπιση των κύκλων πάνω της παρέχει τον προσανατολισμό του αυτόνομου συστήματος ως προς τη σφαίρα. Έτσι, με γνωστή τη θέση της σφαίρας, ο καθορισμός της θέσης του συστήματος απαιτεί μόνο τη λύση απλών

γραμμικών εξισώσεων.

Προφανώς, οι προσεγγίσεις αυτές κάθε άλλο παρά επιτυχείς είναι σε άγνωστα ή δυναμικά μεταβαλλόμενα περιβάλλοντα.

- *Επιλεγμένα ορόσημα (Selected landmarks)*: Σε ορισμένες περιπτώσεις, όταν είναι γνωστοί οι χώροι εργασίας του μηχανικού συστήματος, είναι δυνατό να επιλεχθούν δια χειρός (manually) μέσα από το περιβάλλον διάφορα αντικείμενα από το σχεδιαστή του αυτόνομου συστήματος [29, 62, 63, 64, 65]. Το κριτήριο επιλογής των κατάλληλων ορόσημων είναι η *άμεση διαχωρισιμότητα (distinctiveness)* των από το υπόλοιπο περιβάλλον, ούτως ώστε να παρέχουν μια ικανοποιητική περιγραφή του χώρου. Στην κατηγορία αυτή, αναφέρουμε την προσπάθεια των Yeh et al. [64, 65], η οποία μελετά την πιθανοκρατική επιλογή των ορόσημων. Συγκεκριμένα, με δεδομένο ένα καθορισμένο σύνολο οπτικών χαρακτηριστικών, επιλέγουν ένα ορόσημο που να έχει όσο το δυνατόν περισσότερες πιθανότητες να αναγνωριστεί και ταυτόχρονα να υπάρχουν λίγες πιθανότητες σύγχυσης του με κάποιο άλλο χαρακτηριστικό του χώρου. Για να υπολογιστούν αυτές οι πιθανότητες, γίνονται στατιστικές μετρήσεις σε πραγματικές εικόνες. Έτσι το κάθε ορόσημο είναι μία υπόθεση σε ένα χώρο υποθέσεων. Το πιο *ικανοποιητικών προδιαγραφών* ορόσημο, είναι αυτό που ελαχιστοποιεί το κριτήριο του Bayes στο δεδομένο χώρο υποθέσεων. Επίσης οι Trahanias et al. [66, 67] παρουσιάζουν μια μέθοδο *πλοήγησης προς κάποιο στόχο (targeted navigation)*, όπου καταβάλλεται προσπάθεια ανακάλυψης μιας διαδρομής, που να αποφεύγει τα πιθανά εμπόδια, και να καταλήγει σε ένα αντικειμενικό στόχο. Αν και σε πρώτη ματιά, η μέθοδος τους φαίνεται να ανήκει στην κατηγορία των μεθόδων του *προγραμματισμού κίνησης*, την αναφέρουμε εδώ, διότι ο αντικειμενικός στόχος του μηχανικού συστήματος δεν είναι παρά μια περιοχή με σημαντική πληροφορία, η οποία επιλέγεται από τον χρήστη. Θεωρώντας τον στόχο σαν ένα ορόσημο, η προσέγγισή τους είναι εν μέρει ισοδύναμη με το πρόβλημα αναγνώρισης ενός προ-επιλεγμένου ορόσημου.

  Οι προσέγγισεις της κατηγορίας αυτής, αφενός μεν περιορίζουν τη χρησιμότητα τους εφόσον δύναται να εφαρμοσθούν μόνο σε γνωστούς χώρους, αφετέρου βασίζονται σε ανθρωπομορφικά κριτήρια για τον καθορισμό των ορόσημων, τα οποία δεν είναι κατ' ανάγκη τα πιο κατάλληλα για αυτοματοποιημένες προσεγγίσεις.

- *Ευθείες γραμμές (Straight lines)*: Οι ευθείες γραμμές μπορούν σε ορισμένες περιπτώσεις να περιγράψουν το περιβάλλον και συνεπώς να χρησιμοποιηθούν ως ορόσημα του χώρου. Η προσέγγιση αυτή έχει υιοθετηθεί στις εργασίες που περιγράφονται στις αναφορές [28, 36, 64, 68].

  Παρόλο που κάποια ενθαρυντικά αποτελέσματα έχουν αναφερθεί με αυτή την προσέγγιση, κυρίως για περιπτώσεις γραφείων και διαδρόμων, δεν μπορεί να θεωρηθεί κατάλληλη για γενικότερους χώρους λόγω της εγγενούς ασάφειας στην

περιγραφή γενικών χώρων, που προκύπτει από τις ευθείες γραμμές.

- *Πανοραμική άποψη (Panoramic view)*: Μια μέθοδος προσανατολισμένη σε πλοήγηση μεγάλων αποστάσεων βασίζεται στην αναπαράσταση πανοραμικής άποψης του χώρου, η οποία είναι στην ουσία μια μνήμη οπτικής πληροφορίας όλης της διαδρομής που διανύθηκε κατά την εκμάθηση [69, 70, 71]. Το ρόλο των ορόσημων εδώ παίζουν ολόκληρα τμήματα από την παραπάνω αναπαράσταση, ανάλογα με το πόσο είναι άμεσα διαχωρίσιμα και άρα πόσο ικανοποιητικά μπορούν να αναπαραστήσουν το τμήμα του περιβάλλοντος στο οποίο αναφέρονται [71]. Ωστόσο, όπως γίνεται αντιληπτό, οι απαιτήσεις σε αποθηκευτικούς πόρους αυξάνουν με γεωμετρική πρόοδο, οδηγώντας στην μη αποδοτική διαχείριση της πληροφορίας.

- *Οπτικά πρότυπα (Visual patterns)*: Σε μια πρόσφατη εργασία τους, οι Trahanias et al. [72], ανέπτυξαν μια γεωμετρική μέθοδο για αναγνώριση οπτικών ορόσημων με ταυτόχρονη εφαρμογή της στην διαδικασία πλοήγησης μιας ρομποτικής πλατφόρμας. Στην προσέγγιση αυτή, ο χώρος κίνησης του μηχανικού συστήματος τμηματοποιείται ποιοτικά σε πέντε βασικά επίπεδα[10] και στη συνέχεια μέσω ενός μηχανισμού ‘‘εστίασης της προσοχής’’ γίνεται επιλεκτική εκμάθηση των ορόσημων στις προέχουσες περιοχές της εικόνας. Ως ορόσημα, θεωρούνται χαρακτηριστικά αντικείμενα με σημαντική πληροφορία στις περιοχές αυτές. Στην ίδια εργασία, οι συγγραφείς αναπτύσσουν και ένα μετασχηματισμό θέσης-παρατηρητή, ο οποίος λύνει το πρόβλημα των διαφορετικών θεάσεων ενός ορόσημου. Αν και αρκετά καινοτόμες ιδέες υλοποιηθήκανε στην εργασία αυτή, ωστόσο, η προτεινόμενη λύση στηρίζεται υπερβολικά στην υπόθεση ότι ο χώρος κίνησης του μηχανικού συστήματος αποτελείται από ορθογώνια παραλληλεπίπεδα, περιορίζοντας την αυτονομία της ικανότητας πλοήγησης του. Στη δική μας προσέγγιση, όπως θα εξηγήσουμε, χρησιμοποιούμε τον μηχανισμό ‘‘εστίασης της προσοχής’’, που αναπτύχθηκε στην εργασία αυτή, διότι μας ενδιαφέρουν οι σημαντικές περιοχές μιας εικόνας και όχι αυτές που είναι φτωχές σε πληροφορία.

- *Προβολικές αναλλοίωτες (Projective invariants)*: Στη διεθνή βιβλιογραφία, δεν υπάρχουν αρκετές εργασίες που να ασχολούνται με το θέμα της αυτόνομης πλοήγησης βάσει προβολικών αναλλοίωτων, μιας και η πλειονότητα αυτών εστιάζουν την προσοχή τους, όπως θα δούμε και στη συνέχεια, σε προβλήματα δημιουργίας *συναρτήσεων ευρετηριασμού*, αντιστοίχισης χαρακτηριστικών ιδωμένων από διαφορετικές οπτικές γωνίες, αναγνώρισης αντικειμένων κ.ά. Ωστόσο, μια προσέγγιση στο θέμα επιχείρησαν οι Roh et al. [73]. Οι συγγραφείς υποθέτουν ότι το περιβάλλον τους αποτελείται από τελείως επίπεδα πατώματα και ότι οι τοίχοι με το πάτωμα δημιουργούν δύο τελείως παράλληλες ευθείες. Χρησιμοποιούν τα σημεία τομής του πατώματος και των κάθετων ευθειών των πλαισίων των πορτών

[10] πάτωμα, αριστερός και δεξύς τοίχος, ταβάνι και εσωτερικό επίπεδο

σαν χαρακτηριστικά σημεία για τον υπολογισμό των cross-ratios[11]. Αυτό δεν είναι πάντα αποδεκτό, μιας και στην περίπτωση που ένας διάδρομος δεν περιέχει πόρτες ή κάποιο άνοιγμα ο αλγόριθμος τους είναι καταδικασμένος να αποτύχει. Επίσης στην περίπτωση που υπάρχει κάποιο εμπόδιο (π.χ. μία καρέκλα), στη μια πλευρά του τοίχου ο αλγόριθμος τους θα αποτύχει ξανά μιας και δεν θα μπορούν να δημιουργηθούν οι παράλληλες ευθείες γραμμές. Ακόμη, το γεγονός ότι κατασκευάζουν μια βάση δεδομένων με προ-υπολογισμένες τιμές των cross-ratios, καθιστά τη μέθοδο αρκετά περιοριστική ως προς την δυνατότητα εφαρμογής της σε δυναμικά περιβάλλοντα. Τέλος, η μέθοδος δε συμβαδίζει πλήρως και με τις αρχές της ενεργούς και τελεολογικής όρασης, μιας και δεν εστιάζει την προσοχή της σε σημαντικές περιοχές στην εικόνα αλλά σε μάλλον ασήμαντες. Το θετικό της εργασίας τους είναι η χρήση του cross-ratio 4 συνευθειακών σημείων, καταργώντας έτσι την ανάγκη για βαθμονόμηση της κάμερας και απλοποιώντας την διαδικασία αντιστοίχισης των σημείων μιας νέας σκηνής με τα προαποθηκευμένα στη βάση.

1.2 Η δική μας προσέγγιση

Γίνεται εμφανές ότι το πρόβλημα της αυτόνομης πλοήγησης έχει αντιμετωπισθεί αποσπασματικά με τις μέχρι τώρα προσεγγίσεις και οι ερευνητές που έχουν ασχοληθεί με αυτό έχουν κάνει πολλές παραδοχές για να στηρίξουν την προσπάθεια τους. Ο λόγος είναι η εγγενής δυσκολία που υπάρχει στην αξιόπιστη αναγνώριση χαρακτηριστικών του περιβάλλοντος σε πραγματικές συνθήκες πλοήγησης μέσα στο χώρο. Σημαντικό πρόσκομμα αποτελεί επίσης το επιχείρημα να αποδώσουμε σε ένα μηχανικό σύστημα πραγματικές ικανότητες αντίληψης του χώρου και λήψης αποφάσεων για την πλοήγηση μέσα σ' αυτόν.

Με τη μεθοδολογία, που προτείνουμε στην παρούσα εργασία, επιχειρούμε να δώσουμε λύσεις στα παραπάνω προβλήματα, αναπτύσσοντας μια μέθοδο πλοήγησης βασισμένη στα *οπτικά ορόσημα* και στην αναπαράσταση των μέσω γεωμετρικά αναλλοίωτων ποσοτήτων. Η προσέγγιση αυτή εντάσσεται στη θεωρία της ενεργούς και τελεολογικής όρασης [52, 53], σύμφωνα με την οποία οι διαδικασίες όρασης ενός μηχανικού συστήματος είναι προσανατολισμένες στην επίτευξη συγκεκριμένων στόχων.

Στα πλαίσια της θεωρίας αυτής, θα αποφύγουμε την προσπάθεια επακριβούς αναπαράστασης της γεωμετρικής δομής του περιβάλλοντος, παρακάμπτοντας τα προβλήματα που αυτή συνεπάγεται. Έτσι, το αυτόνομο σύστημα δε διαθέτει εξ' αρχής πληροφορίες για τη δομή ή τη διάταξη του κόσμου, ή για το είδος των αντικειμένων που υπάρχουν σ' αυτόν. Στην προτεινόμενη μέθοδο, οι αναγκαίες αυτές πληροφορίες εξάγονται από το αυτόνομο σύστημα χρησιμοποιώντας οπτικούς αισθητήρες.

Τα οπτικά ορόσημα δεν είναι προεπιλεγμένα ή προκαθορισμένα. Προσπαθώντας να

[11] η βασική προβολική αναλλοίωτη

διατηρήσουμε την ευελιξία του συστήματος και την ικανότητα κίνησης του μέσα σε άγνωστα περιβάλλοντα, χρησιμοποιούμε ένα μηχανισμό *εστίασης της προσοχής (focus of attention)* σε *προέχοντα (salient)* σημεία της εικόνας. Γι' αυτό το λόγο υιοθετούμε την έννοια του *χάρτη προεξοχής (saliency map)* [55], όπου με χρήση διάφορων κριτηρίων εντοπίζουμε στην εικόνα τα πιο ''χαρακτηριστικά'' σημεία. Τα σημεία αυτά αποτελούν τους ''σπόρους'' ανάπτυξης περιοχών στις οποίες θα αναζητηθούν τα ορόσημα.

Τα *οπτικά ορόσημα*, στην προσέγγιση μας, είναι σύνολα από *υπο-ορόσημα*. Το κάθε υπο-ορόσημο αποτελείται από πέντε ομοεπίπεδα σημεία, που δεν είναι ανά τρία συνευθειακά. Τα σημεία που συνθέτουν τα υπο-ορόσημα εξάγονται κατευθείαν από την εικόνα χρησιμοποιώντας ένα εύρωστο ανιχνευτή χαρακτηριστικών. Σε κάθε περιοχή αναζήτησης που έχει προκύψει από τα προέχοντα σημεία, εντοπίζουμε τα χαρακτηριστικά που ενυπάρχουν. Στη συνέχεια για κάθε δυνατή πεντάδα χαρακτηριστικών προσπαθούμε να εντοπίσουμε την αντίστοιχη της στο διαδοχικό πλαίσιο. Η εύρεση της αντιστοίχου πεντάδας γίνεται χρησιμοποιώντας γεωμετρικά αναλλοίωτες ποσότητες και τη νόρμα του πίνακα συνδιασποράς της πεντάδας, που εγγυάται την συνοχή της τοπολογικής δομής της.

Πολλοί ερευνητές έχουν χρησιμοποιήσει εκτεταμένα τις προβολικές αναλλοίωτες για να υπολογίσουν αντιστοιχίσεις χαρακτηριστικών, ιδωμένα κάτω από διαφορετικές οπτικές γωνίες, ταύτισης προτύπων και αναγνώρισης αντικειμένων. Η χρήση τους κρίνεται αναγκαία, μιας και στην αντίθετη περίπτωση, θα πρέπει για κάθε δυνατή οπτική γωνία θέασης ενός αντικειμένου να αποθηκεύεται πληροφορία σχετικά με το αντικείμενο, μιας και οι ποσότητες, που αποθηκεύονται και αναζητούνται, δεν είναι αμετάβλητες κάτω από προβολικούς μετασχηματισμούς. Το μειονέκτημα του γεγονότος αυτού, είναι η αυξημένη πιθανότητα να καταλήξουμε στην κατάσταση, όπου ένα αντικείμενο ταυτίζεται με ένα πλήθος από τα, ιδωμένα από διαφορετικές οπτικές γωνίες, αντικείμενα που έχουν αποθηκευθεί. Οπότε, δεν υπάρχει μονοσήμαντος τρόπος αναγνώρισης, του αντικειμένου που αναζητούμε, με βάση την αποθηκευμένη πληροφορία. Επίσης, η πληροφορία που αποθηκεύεται είναι υπέρογκη, επιβαρύνοντας υπολογιστικά και δυσχαιρένοντας το σκοπό της αναγνώρισης.

Ένας άλλος σημαντικός λόγος για την χρήση των προβολικών αναλλοίωτων είναι ότι μπορούμε να λύσουμε το πρόβλημα της αυτόνομης πλοήγησης, χωρίς τη χρήση καμερών που έχουν βαθμονομηθεί. Επίσης, μας δίνεται η δυνατότητα να αντιστοιχήσουμε χαρακτηριστικά ιδωμένα κάτω από αρκετά διαφορετικές οπτικές γωνίες. Το γεγονός αυτό είναι αρκετά σημαντικό, μιας και παρέχει ένα αποδοτικό τρόπο να αναγνωρίζουμε χαρακτηριστικά που έχουν υποστεί πολυμορφικούς[12] (generic) μετασχηματισμούς[13].

Ωστόσο, στην δική μας προσέγγιση δεν ακολουθούμε αυστηρά αυτή την οδό, μιας και συνήθως προϋποθέτει την επίλυση του *προβλήματος της αντιστοίχισης (correspondence problem)*. Ίσως η μεγαλύτερη δυσκολία έγκειται στο γεγονός ότι οι πιο πολλές προ-

[12]Ο όρος χρησιμοποιείται για να υποδηλώσει την ποικιλία του είδους των μετασχηματισμών
[13]όπως για παράδειγμα είναι οι προβολικοί μετασχηματισμοί

βολικά αναλλοίωτες ποσότητες είναι ευαίσθητες ως προς την μετάθεση των χαρακτηριστικών. Αυτό σημαίνει πως μια διαφορετική διάταξη των χαρακτηριστικών συνήθως παράγει μια διαφορετική τιμή για την προβολική αναλλοίωτη. Σ' αυτή την περίπτωση, θα πρέπει να αποθηκευτούν οι τιμές των προβολικών αναλλοίωτων για όλες τις μεταθέσεις των υπό-μελέτη χαρακτηριστικών. Ακόμη και σε αυτή την περίπτωση, η ανάλυση του χώρου των τιμών των προβολικών αναλλοίωτων καθίσταται αρκετά λεπτομερής για να ληφθεί μια μονοσήμαντη απόφαση για την ''εύρεση αντιστοίχου''. Υπερνικούμε αυτό το πρόβλημα προσπαθώντας να ταιριάξουμε πεντάδες μη διατεταγμένων σημείων χρησιμοποιώντας *προβολικά και μεταθετικά αναλλοίωτες ποσότητες*.

Μόλις έχουμε ανακαλύψει κάποιες υποψήφιες πεντάδες, των οποίων οι αντίστοιχες υπάρχουν στο διαδοχικό πλαίσιο δεν περατώνεται η διαδικασία εκμάθησης αλλά ελέγχουμε το πλήθος των επόμενων, διαδοχικών στο χρόνο, πλαισίων, στα οποία διατηρούνται. Αν το πλήθος είναι ικανοποιητικό, θεωρούμε τις πεντάδες ικανές να χαρακτηριστούν υπο-ορόσημα. Ο λόγος που επιτελείται το πρόσθετο τούτο βήμα είναι για να ελεγχθεί ο *χρόνος ζωής*[14] μιας πεντάδας. Αυτό είναι λογικό, διότι η ανίχνευση και αναγνώριση ορόσημων θα πρέπει να επεκτείνεται και στη διάσταση του χρόνου, αφού η είσοδος στο μηχανικό σύστημα είναι μία ακολουθία εικόνων. Με την παραπάνω τεχνική εισάγουμε την έννοια της ''χρονικής συνέπειας'' στη θέση και τα χαρακτηριστικά ενός ορόσημου[15]. Έτσι, αφενός μεν επιτυγχάνεται μεγαλύτερη αξιοπιστία στην αναγνώριση των ορόσημων, αφού οι διαδοχικές εικόνες ''συναινούν'' ή όχι στις προηγούμενες αποφάσεις, αφετέρου το υπολογιστικό κόστος παραμένει χαμηλό με το να περιορίζονται κατά πολύ οι απαιτούμενοι υπολογισμοί.

Γίνεται λοιπόν εμφανές ότι τα οπτικά ορόσημα, που χρησιμοποιούμε, δεν είναι προεπιλεγμένα ή προκαθορισμένα. Με τον τρόπο αυτό επιτρέπουμε στο αυτόνομο σύστημα την επιλογή των ορόσημων, δίνοντας του την ευχέρεια να προσαρμόζεται σε αλλαγές στο περιβάλλον του.

Η πληροφορία που συλλέγεται σε σχέση με τα οπτικά ορόσημα, αποθηκεύεται σε μια *Βιβλιοθήκη Ορόσημων (Landmark Library)*, μαζί με πληροφορίες για τη σχετική τους θέση στο χώρο. Με τον τρόπο αυτό πραγματοποιείται η δημιουργία ενός τοπολογικού χάρτη του χώρου και η αντιστοίχιση των ορόσημων πάνω σ' αυτόν.

Πέρα από την κάθ' αυτή αναπαράσταση των ορόσημων, επιπλέον πληροφορία μπορεί να συλλέγεται[16] στη φάση της εκμάθησης, που σχετίζεται με τις πιθανές ενέργειες του μηχανικού συστήματος, όταν αναγνωρίσει το συγκεκριμένο ορόσημο. Επίσης ένα ορόσημο μπορεί να αποτελεί ένδειξη για άλλα γειτονικά ορόσημα. Με τον τρόπο αυτό, οι εντολές πλοήγησης μπορούν να δίνονται με ευκολότερο και σαφέστερο τρόπο.

Έχοντας εκπαιδευτεί το αυτόνομο σύστημα με ένα σύνολο ορόσημων, καλείται να αποδείξει ότι πράγματι η πληροφορία που σταχυολόγησε αποτελεί σημαντικό παρά-

[14]ένας όρος δανεισμένος από τον τομέα των κατανεμημένων συστημάτων

[15]Ένα ορόσημο θα βρίσκεται σε γειτονική θέση και θα έχει παρόμοια χαρακτηριστικά σε διαδοχικές εικόνες μέσα στο χρόνο.

[16]π.χ. με τη βοήθεια ενός δασκάλου, που επιβλέπει τη διαδικασία της εκμάθησης

γοντα στην επιτυχία της αυτόνομης πλοήγησής του. Θα πρέπει να επιτρέψουμε στο μηχανικό σύστημα να κινείται στο περιβάλλον του και να αναγνωρίζει ορόσημα. Κατά τη φάση της πλοήγησης, ακολουθείται πρώτα μια διεργασία, πανομοιότυπη με αυτή που ακολουθείται κατά τη φάση της εκμάθησης, ώστε να εντοπισθούν πιθανά ορόσημα. Έχοντας εντοπίσει κάποια πιθανά υπο-ορόσημα, αρχίζουμε να εξετάζουμε υποθέσεις, συγκρίνοντας τα χαρακτηριστικά του κάθε υπο-ορόσημου, που ανιχνεύεται στην φάση αναγνώρισης με τα αποθηκευμένα υπο-ορόσημα στη βιβλιοθήκη ορόσημων. Η επιτυχής αναγνώριση ενός υπο-ορόσημου οδηγεί τον μηχανισμό αναγνώρισης υπο-ορόσημων σε επιπλέον υπο-ορόσημα που αναμένεται να υπάρξουν στο ίδιο πλαίσιο. Έτσι περιορίζεται αισθητά ο χώρος αναζήτησης των νέων υπο-ορόσημων, αλλά και ελαχιστοποιείται αισθητά το υπολογιστικό κόστος.

Αν κατά τη διάρκεια της πλοήγησης συναντηθεί ένα υπο-ορόσημο, που δεν υπάρχει στη βιβλιοθήκη ορόσημων, το αυτόνομο σύστημα μπορεί να επιστρέψει στο στάδιο της εκμάθησης και να αποθηκεύσει το καινούριο ορόσημο στη μνήμη του. Με τον τρόπο αυτό, το σύστημα έχει την ευχέρεια να προσαρμόζεται σε καινούρια δεδομένα και αλλαγές του περιβάλλοντος του. Η αναγνώριση ενός ορόσημου, σηματοδοτεί και τις αντίστοιχες ενέργειες που έχουν αποδωθεί στον τρέχοντα σκοπό πλοήγησης.

Η εφικτότητα και η ευρωστία της προτεινόμενης μεθόδους έχει αποδειχθεί διαμέσου των πειραμάτων σε ακολουθίες εικόνων διαδρόμων εσωτερικών χώρων. Εντούτοις, η μέθοδος μπορεί να αποδώσει ικανοποιητικά σε εικόνες γενικότερης μορφολογίας, μιας και δεν υπόκειται σε περιορισμούς ως προς τον χώρο δράσης. Επίσης, οι μαθηματικές ποσότητες που χρησιμοποιούνται στην εργασία, έχουν αποτιμηθεί και σε ακολουθίες συνθετικών εικόνων, προκειμένου να μελετηθεί η αριθμητική τους ευστάθεια καθώς και η συνέπεια τους όσο αναφορά τα αποτελέσματα με τα οποία συσχετίζονται.

1.2.1 Συνεισφορά

Σίγουρα η επίλυση ενός τόσο πολύπλοκου προβλήματος, όπως η αυτόνομη πλοήγηση, δεν μπορεί να επιλυθεί, στη γενικότητα του, στα πλαίσια μιας μεταπτυχιακής εργασίας. Ωστόσο θεωρούμε ότι η προσέγγιση που υλοποιήθηκε στην τρέχουσα εργασία είναι ένα μικρό βήμα και ταυτόχρονα συνεισφορά στην ανηφορική οδό της, σε βάθος, κατανόησης της αντίληψης του χώρου. Τα σημαντικότερα αποτελέσματα από την εργασία αυτή απετέλεσαν επιστημονικές δημοσίευσεις που έγιναν αποδεκτές σε δύο διεθνούς κύρους συνέδρια [74, 75].

1.3 Οργάνωση της γραπτής εργασίας

Μετά τη συνοπτική περιγραφή του περιεχομένου της εργασίας και την ένταξη του προβλήματος στην ευρύτερη ερευνητική περιοχή, στο επόμενο κεφάλαιο εισάγουμε τον αναγνώστη στον θαυμαστό κόσμο της Προβολικής γεωμετρίας και των προβολικών

μετασχηματισμών. Παραθέτουμε τις βασικές αρχές της και αναλύουμε τις ποσότητες που παραμένουν αναλλοίωτες κάτω από τους προβολικούς μετασχηματισμούς. Επίσης, εισάγουμε την έννοια των προβολικά και μεταθετικά αναλλοίωτων ποσοτήτων, που αποτελούν τον ακρογωνιαίο λίθο της μεθόδου μας. Οι προβολικά και μεταθετικά αναλλοίωτες ποσότητες διατηρούνται τόσο κατά την επίδραση των προβολικών μετασχηματισμών όσο και κάτω από τις τυχόν μεταθέσεις, που μπορεί να υποστεί το σύνολο των χαρακτηριστικών στο οποίο υπολογίζονται. Στη συνέχεια, στο τρίτο κεφάλαιο, γίνεται αναλυτική περιγραφή της προτεινόμενης προσέγγισης στο πρόβλημα της αυτόνομης πλοήγησης. Παρουσιάζονται οι τεχνικές με τις οποίες το αυτόνομο σύστημα:

- Εστιάζει την προσοχή του (focus of attention) σε *προέχουσες* περιοχές του χώρου για αναζήτηση ορόσημων.
- Εκπαιδεύεται βάσει των ορόσημων και αποθηκεύει κατάλληλα χαρακτηριστικά και περιγραφές τους, καθώς και άλλα στοιχεία χρήσιμα στη διεργασία της αναγνώρισης (π.χ. υπόδειξη για άλλα γειτονικά ορόσημα).
- Αναγνωρίζει τα ορόσημα και ταυτόχρονα οριοθετείται στο χώρο.

Βαδίζοντας ακούραστα, στο τέταρτο κεφάλαιο παρουσιάζουμε τα αποτελέσματα που προέκυψαν από την εφαρμογή της μεθόδου σε ακολουθίες εικόνων και στο τελευταίο κεφάλαιο, παραθέτουμε τα συμπεράσματα, που εξείχθησαν από την εργασία ερευνόντας παράλληλα τις πιθανές επεκτάσεις της. Κλείνουμε την εργασία με την παράθεση της βιβλιογραφίας που χρησιμοποιήθηκε, η οποία αποτελεί πηγή γνώσης για την περαιτέρω επεξήγηση και βαθύτερη κατανόηση ορισμένων εννοιών της.

Κεφάλαιο 2

Στοιχεία Προβολικής γεωμετρίας

Σε όλους μας είναι γνώστη η Ευκλίδεια γεωμετρία και το γεγονός ότι περιγράφει τον τριδιάστατο κόσμο μας τόσο καλά. Στην Ευκλίδεια γεωμετρία, οι πλευρές των αντικειμένων έχουν μήκη, οι γωνίες καθορίζονται από τεμνόμενες ευθείες, και δύο ευθείες είναι παράλληλες εάν κείνται στο ίδιο επίπεδο και δεν έχουν σημείο τομής. Επιπροσθέτως, οι ιδιότητες αυτές παραμένουν αναλλοίωτες κατά την εφαρμογή Ευκλίδειων μετασχηματισμών. Μιας και η Ευκλίδεια γεωμετρία περιγράφει τον κόσμο μας τόσο καλά, είναι σε πρώτη σκέψη δελεαστικό να κάνουμε τη σκέψη ότι είναι το μόνο είδος γεωμετρίας[1]. Η γεωμετρία των αντικείμενων όμως μεταβάλλεται αισθητά κατά την διαδικασία απεικόνισης μιας τριδιάστατης σκηνής στο επίπεδο της εικόνας. Ο προοπτικός μετασχηματισμός των γεωμετρικών σχημάτων δεν μπορεί να αντιμετωπιστεί με τους συνηθισμένους μηχανισμούς της Ευκλίδειας γεωμετρίας, αφού κατά την προοπτική προβολή, οι παράλληλες ευθείες τέμνονται[2] στο *σημείο διαφυγής (vanishing point).*

Καθίσταται λοιπόν αναγκαία η εύρεση ενός πολυμορφικού μετασχηματισμού, που να υπερνικά τις ατέλειες του προοπτικού και να οδηγεί σε μια αποτελεσματική αντιμετώπιση του θέματος. Ο μετασχηματισμός αυτός πηγάζει από την μελέτη της προβολικής γεωμετρίας.

Η προβολική γεωμετρία αποτελεί ένα μαθηματικό εργαλείο που, τελευταία, χρησιμοποιείται όλο και πιο συχνά για την επίλυση προβλημάτων στη Μηχανική Όραση [76, 77, 78]. Ένας βασικός λόγος είναι, ότι αν βασιστεί κανείς στην προβολική γεωμετρία, μπορεί να περιγραφεί επακριβώς η διαδικασία σχηματισμού εικόνων μετά από προοπτική προβολή. Όμως, ο πιο σημαντικός λόγος είναι ότι με βάση την προβολική γεωμετρία μπορούμε να επιλύσουμε σημαντικά προβλήματα στη Μηχανική Όραση χωρίς τη χρήση *βαθμονομημένων καμερών*[3] *(calibrated cameras).* Αυτή η προσέγγιση ακολουθείται αρκετά από ερευνητές στο χώρο της Μηχανικής Όρασης, μιας και παρέχει το

[1] Άλλωστε, η λέξη γεωμετρία σημαίνει *καταμέτρηση της γης.*

[2] Το γεγονός αυτό είναι μια από τις κύριες διαφορές της προβολικής γεωμετρίας από την Ευκλίδεια

[3] χωρίς γνώση δηλαδή των παραμέτρων της κάμερας με την οποία έχουν ληφθεί οι εικόνες

πλαίσιο αντιμετώπισης του θέματος χωρίς την εισαγωγή μη ρεαλιστικών περιορισμών. Επίσης, αποφεύγουμε ορισμένα πρακτικά προβλήματα που καθιστούν τη διαδικασία της βαθμονόμησης δύσκολη, όπως η ατέλεια των πληροφοριών που παρέχει ο κατασκευαστής και τα σφάλματα στην εκτίμηση των παραμέτρων λόγω των αριθμητικών μεθόδων που χρησιμοποιούνται [79]. Επιπλέον, σε πολλές εφαρμογές, οι παράμετροι της κάμερας πρέπει να μεταβάλλονται δυναμικά. Στις περιπτώσεις αυτές, ο επαναληπτικός προσδιορισμός των παραμέτρων της κάμερας αποτελεί ένα περιοριστικό παράγοντα για την απόδοση του συστήματος και το προσανατολίζει μόνο σε εφαρμογές που δεν ακολουθούν τη σύγχρονη οδό της ενεργού όρασης. Ένα άλλο σημαντικό στοιχείο είναι ότι, με βάση τη θεωρία της προβολικής γεωμετρίας μπορούν να διατυπωθούν ποσότητες που παραμένουν αναλλοίωτες κάτω από προβολικούς μετασχηματισμούς. Οι ποσότητες αυτές ονομάζονται προβολικές αναλλοίωτες (projective invariants)

Η μελέτη της προβολικής γεωμετρίας, ουσιαστικά, ξεκίνησε από τους καλλιτέχνες στην περίοδο της Ιταλικής Αναγέννησης, στην προσπάθεια τους να δημιουργήσουν μια ρεαλιστική αίσθηση του τριδιάστατου βάθους στις αρχιτεκτονικές τους δημιουργίες. Επινόησαν ενδιαφέρουσες γεωμετρικές κατασκευές, όπως για παράδειγμα ο υπολογισμός της προβολής ενός παραλληλογράμμου όταν η προβολή δύο πλευρών του είναι γνωστή και ο διαχωρισμός ενός προβολικού τετραγώνου σε τέσσερα ίσα υπό-τετράγωνα. Αν και πολλές θεμελιώδεις έννοιες της προβολικής γεωμετρίας μελετήθηκαν το $19^{\underline{ο}}$ αιώνα, εντούτοις οι σύγχρονες έρευνες προχώρησαν αρκετά πιο πέρα από τις αρχικές αυτές μελέτες, ιδιαίτερα όσο αναφορά την ικανότητα μας να διαχειριζόμαστε ακολουθίες εικόνων και άγνωστες παραμέτρους της κάμερας, καθώς και σε πρακτικά υπολογιστικά θέματα όπως η συνέπεια, η ευρωστία και η αριθμητική ακρίβεια [76].

2.1 Ο προβολικός μετασχηματισμός

Ο προβολικός μετασχηματισμός έχει τρία σημεία κλειδιά που τον χαρακτηρίζουν:

1. Κάθε προοπτική προβολή μπορεί να αναπαρασταθεί μέσω ενός προβολικού μετασχηματισμού.

2. Οιαδήποτε σύνθετη αλυσίδα προοπτικών προβολών δεν είναι απαραίτητα ένας προοπτικός μετασχηματισμός αλλά είναι πάντα ένας προβολικός μετασχηματισμός.

3. Κάθε προβολικός μετασχηματισμός μπορεί να διασπαστεί σε δύο προοπτικούς μετασχηματισμούς.

Έχοντας ορίσει λεκτικά τα χαρακτηριστικά του προβολικού μετασχηματισμού, το επόμενο βήμα είναι η ανάπτυξη ενός μαθηματικού πλαισίου εργασίας που να χαρακτηρίζει τις ιδιότητες του. Η βασική αξία μιας μαθηματικής θεμελίωσης είναι η συμπάγεια και η εύκολη διαχείριση των φαινομένων που στηρίζονται σε προβολικούς μετασχηματισμούς.

2.2 Μαθηματική σημειολογία τού προβολικού κόσμου

Ένα σημαντικό στοιχείο της προβολικής γεωμετρίας είναι η εισαγωγή των ομογενών συντεταγμένων που αναπαριστούν ένα προβολικό μετασχηματισμό σαν ένα πολλαπλασιασμό. Η συμπαγής αυτή μορφή επιτρέπει την παρουσίαση πολλών σημαντικών ιδιοτήτων των προβολικών μετασχηματισμών και της προβολικής γεωμετρίας, χρησιμοποιώντας απλούς υπολογισμούς από την άλγεβρα πινάκων. Με χρήση Ευκλίδειων συντεταγμένων, πολλές από αυτές τις παραγωγές καθίστανται δύσκολες, αν όχι αδύνατες.

2.2.1 Βασικοί ορισμοί

Αναπαριστούμε ένα σημείο του πραγματικού n-διάστατου *προβολικού χώρου* $\mathcal{P}^n$ με ένα διάνυσμα $(n+1)$ πραγματικών συντεταγμένων:

$$\mathbf{x} \;=\; \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n+1} \end{bmatrix} \tag{2.1}$$

Οι συντεταγμένες x_i ονομάζονται *προβολικές συντεταγμένες* του σημείου και το διάνυσμα $\mathbf{x}$ *διάνυσμα συντεταγμένων*. Η προϋπόθεση που πρέπει να τηρείται είναι ότι τουλάχιστο μία συντεταγμένη θα πρέπει να είναι μη-μηδενική. Επίσης τα διανύσματα:

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n+1} \end{bmatrix} \quad \text{και} \quad \begin{bmatrix} \lambda x_1 \\ \lambda x_2 \\ \vdots \\ \lambda x_{n+1} \end{bmatrix} \tag{2.2}$$

αναπαριστούν το ίδιο σημείο στον $\mathcal{P}^n$ για κάθε $\lambda \neq 0$, μιας και η κλιμάκωση των ομογενών συντεταγμένων δεν προκαλεί καμιά διαφορά.

Ένας $(n+1) \times (n+1)$ πίνακας $\mathbf{A}$ με $|A| \neq 0$, ορίζει ένα *γραμμικό μετασχηματισμό* από το $\mathcal{P}^n$ στον εαυτό του. Το σύνολο των γραμμικών μετασχηματισμών από το $\mathcal{P}^n$ στο $\mathcal{P}^n$ συνθέτει την *προβολική ομάδα*.

Ως *προβολική βάση* για τον προβολικό χώρο $\mathcal{P}^n$ ορίζεται ένα σύνολο $n+2$ διανυσμάτων του $\mathcal{P}^n$, που δεν περιέχει $n+1$ γραμμικά εξαρτημένα διανύσματα. Για παράδειγμα, το σύνολο

$$\mathcal{C}_n \;=\; \left\{ \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} \quad \vdots \quad \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} \quad \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \right\} \tag{2.3}$$

των $n+2$ διανυσμάτων συνθέτει μία προβολική βάση, που είναι γνωστή και ως *κανονικοποιημένη βάση*. Οποιαδήποτε προβολική βάση μπορεί να απεικονιστεί στην

πρότυπη αυτή μορφή μέσω ενός κατάλληλου μετασχηματισμού. Κάθε διάνυσμα $\mathbf{x} \in \mathcal{P}^n$ μπορεί να εκφραστεί ως γραμμικός συνδυασμός οποιονδήποτε $n+1$ διανυσμάτων της κανονικοποιημένης βάσης:

$$\mathbf{x} \;=\; \sum_{i=1}^{n+1} x_i \mathbf{e}_i$$

2.2.2 Προβολικά δομικά στοιχεία

Η προβολική γεωμετρία ενυπάρχει σε κάθε διάσταση, όπως ακριβώς και η Ευκλίδεια. Τα βασικά δομικά στοιχεία που εξετάζουμε είναι η προβολική ευθεία, το προβολικό επίπεδο και ο προβολικός χώρος.

2.2.2.1 Η προβολική ευθεία

Η προβολική ευθεία, την οποία συμβολίζουμε με $\mathcal{P}^1$, είναι ανάλογη με τον μονοδιάστατο Ευκλίδειο κόσμο. Με βάση την κανονικοποιημένη βάση $\mathcal{C}_1$ κάθε σημείο $\mathbf{x}$ στην ευθεία μπορεί να εκφραστεί με δύο συντεταγμένες x_1, x_2 τέτοιες ώστε

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \;=\; x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \text{με} \;\; |x_1| + |x_2| \neq 0 \tag{2.4}$$

2.2.2.2 Το προβολικό επίπεδο

Το προβολικό επίπεδο είναι μια μαθηματική έννοια που σκοπό έχει να μοντελοποιήσει τις γεωμετρικές ιδιότητες μιας ακολουθίας ενός ή περισσοτέρων προοπτικών προβολών. Στο μοντέλο του προβολικού επιπέδου, οι μετασχηματισμοί αναπαρίστανται ως απεικονίσεις του επιπέδου στον εαυτό του.

Το προβολικό επίπεδο, $\mathcal{P}^2$, αποτελεί μια γενίκευση του Ευκλίδειου επιπέδου. Η γενίκευση αυτή προχωρά σε δύο βήματα. Κατά πρώτον, η έννοια της απόστασης αποσύρεται, δημιουργώντας μια δομή που ονομάζεται *αφφινικό* επίπεδο. Η κύρια ιδιότητα που χαρακτηρίζει το αφφινικό επίπεδο είναι η διατήρηση της παραλληλίας κάτω από αφφινικούς μετασχηματισμούς. Το αφφινικό επίπεδο συχνά χρησιμοποιείται σε εφαρμογές της μηχανικής όρασης σαν μια καλή προσέγγιση της προοπτικής απεικόνισης εικόνας. Κατά δεύτερο, το μοντέλο αποσύρει την έννοια των παράλληλων ευθειών. Κάθε ζεύγος ευθειών τέμνεται σε κάποιο μοναδικό σημείο κάτω από προοπτική θέαση. Για την αντιμετώπιση της περίπτωσης που δύο ευθείες είναι πραγματικά παράλληλες, τέμνονται δηλαδή στο άπειρο, εισάγεται η έννοια του *ιδεατού* σημείου. Η διεύθυνση κάθε παράλληλης ευθείας ορίζει ένα διαφορετικό ιδεατό σημείο. Το σύνολο όλων των ιδεατών σημείων είναι μια ευθεία που προστίθεται στο αφφινικό επίπεδο.

Ένα αφφινικό επίπεδο με προσαρτημένη μια ιδεατή ευθεία ονομάζεται *προβολικό επίπεδο*. Στο μοντέλο του προβολικού επίπεδου, τα ιδεατά σημεία δεν διαχωρίζονται

από τα κανονικά, απλοποιώντας αισθητά την προσπάθεια ανάλυσης των προοπτικών επιδράσεων. Δύο σημαντικά αξίωματα του προβολικού επιπέδου είναι τα εξής:

Αξίωμα 1 *Δύο διαφορετικά σημεία καθορίζουν μια και μόνη ευθεία.*

Αξίωμα 2 *Δύο διαφορετικές ευθείες καθορίζουν ένα μοναδικό σημείο.*

Το αξίωμα 2 δεν θα ίσχυε στο αφφινικό επίπεδο αφού οι παράλληλες ευθείες δεν συναντώνται. Η πιο σημαντική ιδιότητα των πιο πάνω αξιωμάτων είναι ότι είναι πανομοιότυπα εκτός από τις λέξεις *σημείο* και *ευθεία*. Έτσι στο προβολικό επίπεδο, τα σημεία και οι ευθείες είναι *δυϊκά ανάλογες* μορφές. Οποιοδήποτε θεώρημα (ιδιότητα) που εφαρμόζεται στις ευθείες, μπορεί να εφαρμοστεί και στα σημεία, και αντίστροφα. Σε μια πρόταση, που περιλαμβάνει σημεία και ευθείες, οι δύο λέξεις μπορούν να εναλλαχθούν χωρίς να αλλοιωθεί η αλήθεια που την διέπει. Η ιδιότητα αυτή του δυϊσμού είναι πολύ σημαντική για τις εφαρμογές της προβολικής γεωμετρίας. Μόλις ένα αποτέλεσμα ισχύει για σημεία, ισχύει δωρεάν και για ευθείες. Χρησιμοποιώντας την κανονικοποιημένη βάση $\mathcal{C}_2$ κάθε σημείο $\mathbf{x}$ στο προβολικό επίπεδο μπορεί να εκφραστεί με τρεις συντεταγμένες x_1, x_2, x_3 τέτοιες ώστε

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad , \quad \text{με} \quad |x_1| + |x_2| + |x_3| \neq 0 \tag{2.5}$$

Μια ευθεία στον $\mathcal{P}^2$ ορίζεται από μια τριάδα αριθμών u_1, u_2, u_3 μέσω της εξίσωσης:

$$\sum_{i=1}^{3} u_i x_i = 0 \text{ με } |u_1| + |u_2| + |u_3| \neq 0 \tag{2.6}$$

Οι γραμμικοί μετασχηματισμοί στον $\mathcal{P}^2$ ορίζονται από αντιστρέψιμους 3×3 πίνακες.

Για να εκφράσουμε ένα σημείο του προβολικού επιπέδου σε Ευκλίδειες συντεταγμένες, απλά διαιρούμε με την τρίτη συντεταγμένη:

$$\begin{aligned} \mathrm{X}_1 &= \frac{x_1}{x_3} \\ \mathrm{X}_2 &= \frac{x_2}{x_3} \end{aligned}$$

Γίνεται άμεσα αντιληπτό, ότι το προβολικό επίπεδο περιλαμβάνει περισσότερα σημεία από ότι το Ευκλίδειο, ουσιαστικά τα σημεία των οποίων η τρίτη συντεταγμένη είναι μηδέν. Το σημείο $(x_1, x_2, 0)$ είναι ένα νόμιμο τριδιάστατο σημείο που ορίζει μια τελείως κάθετη οπτική ακτίνα, της οποίας το ίχνος δεν αντιστοιχεί σε κανένα εικονοστοιχείο. Τέτοιες ακτίνες ή ομογενή διανύσματα δεν μπορούν να χαρακτηριστούν ως πεπερασμένα σημεία στο διδιάστατο επίπεδο. Ωστόσο, μπορούν να ειδωθούν σαν επιπρόσθετα *ιδεατά σημεία* ή όρια καθώς το (x_1, x_2) συγκλίνει στο άπειρο σε μια δεδομένη κατεύθυνση:

$$\lim_{x_3 \to 0} (\frac{x_1}{x_3}, \frac{x_2}{x_3}, 1) \sim \lim_{x_3 \to 0} (x_1, x_2, x_3) = (x_1, x_2, 0)$$

Υπάρχει ένα ξεχωριστό ιδεατό σημείο που σχετίζεται με κάθε διάσταση του επιπέδου. Για παράδειγμα, τα σημεία (1, 0, 0) και (0, 1, 0) σχετίζονται με την οριζόντια και κάθετη κατεύθυνση αντίστοιχα. Όλα τα ιδεατά σημεία κείνται σε μια ευθεία, που ονομάζεται *ιδεατή* και αναπαρίσταται ως (0, 0, 1). Μπορούμε να προσθέσουμε τέτοια ιδεατά σημεία σε κάθε τριδιάστατο επίπεδο. Σε διδιάστατες εικόνες του επιπέδου, τα επιπρόσθετα σημεία στο άπειρο αποτελούν τον *ορίζοντα* του επιπέδου.

2.2.2.3 Ο προβολικός χώρος

Ο προβολικός χώρος, $\mathcal{P}^3$, σχετίζεται με τον τριδιάστατο Ευκλίδειο χώρο. Ένα σημείο $\mathbf{x}$ στον προβολικό χώρο $\mathcal{P}^3$ αναπαρίσταται από 4 συντεταγμένες x_1, x_2, x_3, x_4. Ένα *επίπεδο* στον $\mathcal{P}^3$ ορίζεται από μια τετράδα αριθμών u_1, u_2, u_3, u_4 μέσω της εξίσωσης:

$$\sum_{i=1}^{4} u_i x_i \;=\; 0 \text{ με } |u_1| + |u_2| + |u_3| + |u_4| \neq 0 \tag{2.7}$$

Σκεπτόμενοι όμοια, μπορούμε να εκφράσουμε κάθε σημείο του $\mathcal{P}^3$ με τρεις καρτεσιανές συντεταγμένες $X_i \; i = 1, \ldots, 3$

$$X_1 \;=\; \frac{x_1}{x_4}$$

$$X_2 \;=\; \frac{x_2}{x_4}$$

$$X_3 \;=\; \frac{x_3}{x_4}$$

Τα σημεία που εξαιρούνται από αυτήν την αναπαράσταση είναι τα σημεία του επιπέδου με εξίσωση $x_4 = 0$. Αυτό το επίπεδο συμβολίζεται με π_∞. Κάθε επίπεδο στον $\mathcal{P}^3$ τέμνει το π_∞ στην ευθεία του απείρου που του αντιστοιχεί. Όμοια με τον $\mathcal{P}^2$, μπορεί κανείς να φανταστεί τα σημεία που ανήκουν στο π_∞ ως κατευθύνσεις στον $\mathcal{P}^3$. Οι γραμμικοί μετασχηματισμοί στον $\mathcal{P}^3$ ορίζονται από 4×4 αντιστρέψιμους πίνακες.

2.3 Είδη προβολικών μετασχηματισμών

Με βάση τους παραπάνω ορισμούς, παραθέτουμε ορισμένα είδη μετασχηματισμών που έχουν ιδιαίτερο ενδιαφέρον σε εφαρμογές στον τομέα της Μηχανικής Όρασης.

2.3.1 Προοπτικός μετασχηματισμός

Η προβολή του $\mathcal{P}^3$ στον $\mathcal{P}^2$ προσομοιώνει το σχηματισμό εικόνων από την τριδιάστατη σκηνή στο επίπεδο της κάμερας. Μπορεί να εκφραστεί μέσω της σχέσης:

$$\mathcal{P}^3 \rightarrow \mathcal{P}^2 : \mathbf{x} \;=\; \mathcal{T}\,\mathbf{X} \tag{2.8}$$

όπου ο $\mathcal{T}$ συμβολίζει ένα 3×4 πίνακα μετασχηματισμού. Ο πίνακας $\mathcal{T}$ έχει 12 παραμέτρους, αλλά καθώς η συνολική κλιμάκωση δεν επηρεάζει το αποτέλεσμα οι ουσιαστικές παράμετροι είναι 11. Στο συγγραφικό του έργο, ο Roberts [80], απόδειξε, με βάση την παραπάνω διατύπωση, ότι για τον προσδιορισμό του πίνακα $\mathcal{T}$, είναι απαραίτητη η γνώση τουλάχιστο έξι αντιστοιχίσεων μεταξύ πραγματικά τριδιάστατων σημείων και των αντίστοιχων σημείων της εικόνας που σχηματίζεται στο επίπεδο της κάμερας.

2.3.2 Προβολικότητα

Όπως είναι γνωστό, ο προβολικός μετασχηματισμός μεταξύ δύο προβολικών επιπέδων μπορεί να αναπαρασταθεί με ένα γενικό γραμμικό μετασχηματισμό του *χώρου ακτινών (ray space)*, $(x_1, x_2, x_3)^{\mathrm{T}} = T(X_1, X_2, X_3)^T$ [78]. Ο γενικός αυτός μετασχηματισμός ονομάζεται προβολικότητα. Αποδεικνύεται ότι όλες οι ιδιότητες ενός γενικού προβολικού μετασχηματισμού μπορούν να απεικονιστούν μέσω του μετασχηματισμού αυτού του πίνακα. Αντίστροφα, οποιοσδήποτε μη-παθολογικός (nonsingular) γραμμικός μετασχηματισμός των ομογενών συντεταγμένων ορίζει ένα προβολικό μετασχηματισμό του προβολικού επιπέδου.

Η μορφή αυτή του προβολικού μετασχηματισμού είναι απόρεια της εισαγωγής ομογενών συντεταγμένων, μιας και αρκετά σημαντικά αποτελέσματα μπορούν να γίνουν αντιληπτά άμεσα αποτιμώντας απλές αλγεβρικές εκφράσεις πινάκων και διανυσμάτων.

Αφού το προβολικό επίπεδο έχει τρεις ομογενείς συντεταγμένες ο μετασχηματισμός αναπαρίσταται μέσω ενός 3×3 πίνακα $\mathcal{H}$ με 8 βασικές παραμέτρους. Λόγω του ότι η κλιμάκωση των ομογενών συντεταγμένων δεν προκαλεί καμιά διαφορά, ένα στοιχείο του $\mathcal{H}$ μπορεί να τεθεί στη μονάδα. Ο γενικός προβολικός μετασχηματισμός από ένα προβολικό επίπεδο, Π, σε ένα άλλο, π, αναπαρίσταται ως[4]:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} \tag{2.9}$$

ή

$$\mathbf{x} = \mathcal{H}\,\mathbf{X} \tag{2.10}$$

Αν ο μετασχηματισμός εκφραζόταν σε Καρτεσιανές συντεταγμένες στον Ευκλίδειο ή στον αφφινικό χώρο, η μη γραμμική φύση του είναι άμεσα εμφανής:

$$\begin{aligned} x &= \frac{x_1}{x_3} = \frac{h_{11}X + h_{12}Y + h_{13}}{h_{31}X + t_{32}Y + t_{33}} \\ y &= \frac{x_2}{x_3} = \frac{h_{21}X + t_{22}Y + h_{23}}{h_{31}X + t_{32}Y + t_{33}} \end{aligned} \tag{2.11}$$

[4]Τα μικρά γράμματα αναφέρονται στα αντικείμενα που έχουν υποστεί τον προβολικό μετασχηματισμό ενώ τα κεφαλαία αναφέρονται στα αντικείμενα πριν τον προβολικό μετασχηματισμό

Ο πίνακας προβολικού μετασχηματισμού, $\mathcal{H}$, απαιτεί 8 ανεξάρτητες παραμέτρους για να ορίσει μια μονοσήμαντη απεικόνιση. Το κάθε σημείο στο επίπεδο παρέχει δύο εξισώσεις καρτεσιανών συντεταγμένων, οπότε είναι απαραίτητο να βρούμε τέσσερις αντιστοιχίσεις σημείων μεταξύ δύο προβολικά μετασχηματισμένων επιπέδων για τον μονοσήμαντο ορισμό του πίνακα μετασχηματισμού. Έτσι αν τα τέσσερα αντιστοιχούντα σημεία αναπαρίσταντο ως, $(\lambda_i x_i, \lambda_i y_i, \lambda_i)^T = T\,(X_i, Y_i, 1)^T$, το παραγόμενο σύστημα των γραμμικών εξισώσεων θα ήταν [78],

$$\begin{bmatrix} X_1 & Y_1 & 1 & 0 & 0 & 0 & -x_1X_1 - x_1X_1 \\ 0 & 0 & 0 & X_1 & Y_1 & 1 & -y_1X_1 - y_1X_1 \\ X_2 & Y_2 & 1 & 0 & 0 & 0 & -x_2X_2 - x_2X_2 \\ 0 & 0 & 0 & X_2 & Y_2 & 1 & -y_2X_2 - y_2X_2 \\ X_3 & Y_3 & 1 & 0 & 0 & 0 & -x_3X_3 - x_3X_3 \\ 0 & 0 & 0 & X_3 & Y_3 & 1 & -y_3X_3 - y_3X_3 \\ X_4 & Y_4 & 1 & 0 & 0 & 0 & -x_4X_4 - x_4X_4 \\ 0 & 0 & 0 & X_4 & Y_4 & 1 & -y_4X_4 - y_4X_4 \end{bmatrix} \begin{bmatrix} h_{11} \\ h_{12} \\ h_{13} \\ h_{21} \\ h_{22} \\ h_{23} \\ h_{31} \\ h_{32} \end{bmatrix} = \begin{bmatrix} x_1 \\ y_1 \\ x_2 \\ y_2 \\ x_3 \\ y_3 \\ x_4 \\ y_4 \end{bmatrix} \tag{2.12}$$

Στην περίπτωση που έχουν βρεθεί περισσότερα από 4 αντιστοιχούντα ζεύγη σημείων, μπορεί κανείς να εφαρμόσει τεχνικές, όπως η μέθοδος ελαχίστων τετραγώνων, για την καλύτερη εκτίμηση της λύσης, καθώς προκύπτουν περισσότεροι περιορισμοί για το σύστημα από αυτούς που απαιτούνται για την επίλυσή του [78].

Η ύπαρξη του γραμμικού αυτού συστήματος βεβαιώνει την μοναδικότητα του $\mathcal{H}$, δοθέντος της αντιστοιχίας των 4 σημείων, με μόνο περιορισμό τρία εξ αυτών να μην είναι συνευθειακά. Το συμπέρασμα αυτό οδηγεί αμέσως σε ένα κανονικοποιημένο σύστημα συντεταγμένων, που βασίζεται σε τέσσερα σημεία, όπου οι ιδιότητες των γεωμετρικών εικόνων μπορούν να αναπαρασταθούν αναλλοίωτα. Μια προφανής επιλογή είναι αυτή της κανονικοποιημένης βάσης $\mathcal{C}$. Κάθε τετράπλευρο, μπορεί να μετασχηματιστεί προβολικά στην $\mathcal{C}$ και κατόπιν οι γεωμετρικές σχέσεις των άλλων σημείων και ευθειών μπορούν να αναπαρασταθούν αναλλοίωτα σε αυτή.

2.4 Ιεραρχία γεωμετριών και μετασχηματισμών

Ο μετασχηματισμός ενός σημείου $\mathbf{p} \in \pi$ στο $\mathbf{p}^{'} \in \Pi$ γράφεται ως

$$\mathbf{p}' \simeq \mathcal{H}\mathbf{p} \tag{2.13}$$

Η μελέτη του πίνακα $\mathcal{H}$ οδηγεί στην διελεύκανση της σχέσης ανάμεσα στις γεωμετρίες και των μετασχηματισμών που ισχύουν σε αυτές. Αναπαριστούμε τον πίνακα $\mathcal{H}$, σύμφωνα με τον Weiss [81], ως εξής:

$$\mathcal{H}_{projective} = \begin{bmatrix} a_1 & a_2 & t_x \\ a_3 & a_4 & t_y \\ p_1 & p_2 & 1 \end{bmatrix} \tag{2.14}$$

Τα στοιχεία a_i μαζί με τα στοιχεία t_x, t_y συνθέτουν τους αφφινικούς μετασχηματισμούς. Τα στοιχεία p_1, p_2 αντιπροσωπεύουν τις παραμέτρους τριδιάστατης κλίσης tilt και slant.

Το αφφινικό επίπεδο είναι απλά το προβολικό επίπεδο μείον την ιδεατή ευθεία. Οπότε, οι αφφινικοί μετασχηματισμοί[5] πρέπει να διατηρούν την ιδεατή ευθεία και τα ιδεατά σημεία, δηλαδή, οποιοδήποτε σημείο $[\mathrm{X}, \mathrm{Y}, 0]^{\mathrm{T}}$ πρέπει να μετασχηματιστεί στο $[\alpha\mathrm{X}, \alpha\mathrm{Y}, 0]^{\mathrm{T}}$ για κάποια αυθαίρετη σταθερά κλιμάκωσης α, δηλαδή:

$$\alpha \begin{bmatrix} X \\ Y \\ 0 \end{bmatrix} = \mathcal{H} \begin{bmatrix} X \\ Y \\ 0 \end{bmatrix} \tag{2.15}$$

από το οποίο συνεπάγεται ότι $p_1 = p_2 = 0$. Ο πίνακας για τον αφφινικό μετασχηματισμό είναι:

$$\mathcal{H}_{affine} = \begin{bmatrix} a_1 & a_2 & t_x \\ a_3 & a_4 & t_y \\ 0 & 0 & 1 \end{bmatrix} \tag{2.16}$$

Όπως γίνεται εμφανές ο μετασχηματισμός είναι γραμμικός:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} \begin{bmatrix} X \\ \mathrm{Y} \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix} \tag{2.17}$$

Αντίθετα με τους αφφινικούς μετασχηματισμούς, οι μετασχηματισμοί ομοιότητας[6] διατηρούν τις γωνίες και τους λόγους των μηκών. Ο πίνακας μετασχηματισμού είναι:

$$\mathcal{H}_{similarity} = \begin{bmatrix} \cos\theta & \sin\theta & t_x \\ -\sin\theta & \cos\theta & t_y \\ 0 & 0 & 1 \end{bmatrix} \tag{2.18}$$

όπου το θ είναι μια αυθαίρετη γωνία.

Τέλος, κάτω από Ευκλίδειους μετασχηματισμούς[7], η κλιμάκωση δεν έχει επιρροή, και έτσι το σημείο **p** πρέπει πρώτα να μετατραπεί σε Ευκλίδειες συντεταγμένες διαιρώντας με το τρίτο του στοιχείο. Έτσι ο μετασχηματισμός είναι:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix} \tag{2.19}$$

Όπως γίνεται εμφανές, οι μετασχηματισμοί ακολουθούν μια αντίστοιχη ιεραρχία με τις γεωμετρίες. Η γενική σχέση είναι:

Προβολικός $\supset$ Αφφινικός $\supset$ Ομοιότητας $\supset$ *Ευκλίδειος*

[5] περιστροφή, μεταφορά, κλιμάκωση, shearing
[6] μεταφορά, περιστροφή, κλιμάκωση
[7] μεταφορά, περιστροφή

2.5 Αναλλοίωτες ποσότητες κάτω από προβολικούς μετασχηματισμούς

Η προβολική γεωμετρία μοντελοποιεί πολύ καλά την διαδικασία απεικόνισης μιας τριδιάστατης σκηνής στο επίπεδο της εικόνας γιατί επιτρέπει πλουσιότερο σύνολο από μετασχηματισμούς εκτός των Ευκλίδειων, όπως οι προοπτικοί. Φυσικά το μειονέκτημα είναι ότι διατηρούνται λιγότερες μετρικές ποσότητες. Οι λιγοστές αυτές ποσότητες που διατηρούνται ονομάζονται *προβολικές αναλλοίωτες* και, στη συνέχεια, θα αναλύσουμε τα χαρακτηριστικά και το πλαίσιο των εφαρμογών που χρησιμοποιούνται.

Η ύπαρξη των προβολικών αναλλοίωτων δίνει την δυνατότητα ανάπτυξης αλγορίθμων στον τομέα της Μηχανικής Όρασης με αισθητή μείωση στο πλήθος και τη μορφή των περιορισμών.

Η βασική ιδιότητα τους είναι ότι παρέχουν τη δυνατότητα να διαπιστώσουμε κατά πόσο δύο υποσύνολα από γεωμετρικά χαρακτηριστικά, συνήθως σημεία, ευθείες και γεωμετρικές καμπύλες, μπορούν να σχετίζονται μεταξύ τους με ένα προβολικό μετασχηματισμό. Δηλαδή, κατά πόσο το δεύτερο υποσύνολο γεωμετρικών χαρακτηριστικών μπορεί να προκύψει εφαρμόζοντας ένα προβολικό μετασχηματισμό στο πρώτο. Η βασική εφαρμογή αυτής της ιδιότητας, όπως θα διαπιστώσουμε, είναι η αναγνώριση συνόλων γεωμετρικών χαρακτηριστικών από διαφορετικές όψεις που σχετίζονται μεταξύ τους με γενικούς προβολικούς μετασχηματισμούς.

Όπως είναι γνωστό, οι απλές αποστάσεις, οι γωνίες, οι λόγοι αποστάσεων και οι λόγοι γωνιών δεν διατηρούνται κάτω από γενικούς προβολικούς μετασχηματισμούς. Αποδεικνύεται, ότι όλες οι ποσότητες που παραμένουν αναλλοίωτες μετά την εφαρμογή γενικών προβολικών μετασχηματισμών ανάγονται σε μια ερμηνεία με βάση μια προβολική αναλλοίωτη, που καλείται *cross-ratio*.

2.5.1 Μονοδιάστατοι cross-ratio

Ίσως το πιο σημαντικό αποτέλεσμα είναι ότι το cross-ratio τεσσάρων συνευθειακών σημείων, που είναι ένας λόγος λόγων αποστάσεων, διατηρήται κάτω από προβολικούς μετασχηματισμούς. Υπάρχουν πολλά αποτελέσματα στην προβολική γεωμετρία, που επεξηγούνται μέσω του cross-ratio [78]. Είναι όντως αξιοσημείωτο ότι όλες οι αναλλοίωτες ποσότητες μιας γεωμετρικής συλλογής μπορούν να αναλυθούν μέσω κάποιων κατασκευών cross ratios. Για παράδειγμα στο [82], οι αναλλοίωτες δύο ομοεπίπεδων κωνοειδών εκφράζονται συναρτήσει των cross-ratios.

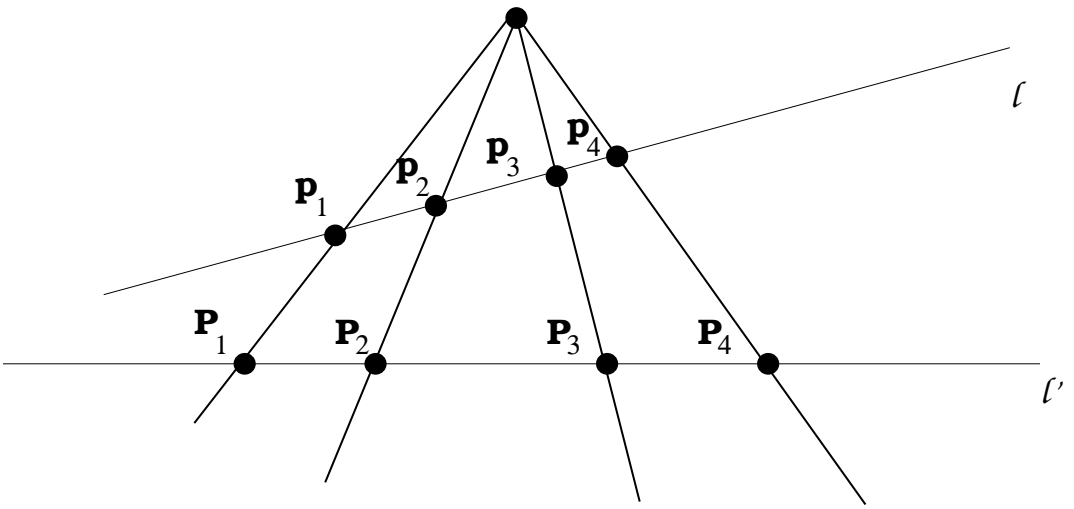
Στον προβολικό χώρο τα σημεία και οι ευθείες είναι δυϊκές μορφές. Χρησιμοποιούμε μόνο συλλογές σημείων, αλλά τα παρατηρούμενα αποτελέσματα είναι νόμιμα και για τις συλλογές των δυϊκών ευθειών. Η βασική προβολική αναλλοίωτη τεσσάρων συνευθειακών σημείων A_i, $i = 1, \ldots, 4$, όπως αναφέραμε, είναι το *cross ratio*.

Το cross-ratio εμφανίζεται σε διάφορες μορφές ανάλογα με τον προβολικό χώρο

στον οποίο αναφερόμαστε. Η βασική του μορφή δίνεται όταν έχουμε τέσσερα σημεία p_i, $i = 1, \ldots, 4$ που κείνται στην ευθεία $\mathcal{L}$ [83]. Θεωρώντας την ευθεία αυτή ως τον προβολικό χώρο $\mathcal{P}^1$, το κάθε σημείο μπορεί να εκφραστεί μέσω μιας καρτεσιανής συντεταγμένης x_i που να εκφράζει τη θέση του στην ευθεία. Το *cross ratio* της συλλογής των τεσσάρων σημείων ορίζεται ώς:

$$\lambda \;=\; [A_1, A_2, A_3, A_4] \;=\; \frac{(x_3 - x_1)(x_4 - x_2)}{(x_3 - x_2)(x_4 - x_1)} \tag{2.20}$$

Η γεωμετρική ερμηνεία του cross-ratio επεξηγείται μέσω της εικόνας 2.1, όπου απεικονίζονται τέσσερις συντρέχουσες ευθείες που κείνται σε ένα επίπεδο[8]. Τότε



Σχήμα 2.1: Γεωμετρική ερμηνεία του cross-ratio

οποιαδήποτε ευθεία $\mathcal{L}$, από μια δέσμη ευθειών, επιλέξουμε, που να τέμνει τις τέσσερις ευθείες στα σημεία $\mathbf{p}_i$, $i = 1, \ldots, 4$, η ποσότητα της εξίσωσης 2.20 παραμένει αναλλοίωτη και είναι ίση με το cross-ratio των τεσσάρων ευθειών.

Ομοίως, έστω στον $\mathcal{P}^3$ μια δέσμη επιπέδων π_i, $i = 1, \ldots, 4$ που έχουν μια κοινή ευθεία τομής. Το cross-ratio των επιπέδων είναι ίσο με το cross-ratio των ευθειών που προκύπτουν από την τομή ενός τυχαίου επιπέδου $\pi \in \mathcal{P}^3$ με τη δέσμη.

Μπορεί να αποδειχθεί [78, 84], ότι το cross-ratio είναι αναλλοίωτο κάτω από μονοδιάστατους προβολικούς μετασχηματισμούς της συλλογής των τεσσάρων σημείων. Το cross-ratio μπορεί να γενικευθεί για ομοεπίπεδες συλλογές στον N−διάστατο χώρο [85].

[8] Η κατασκευή είναι γνωστή και σαν *pencil*

2.5.2 Δύο ευθείες - Δύο σημεία αναλλοίωτη

Πρόσφατα, έχει ανακαλυφθεί ότι η συλλογή δύο ομοεπίπεδων ευθειών και δύο σημείων έχει την ακόλουθη αναλλοίωτη:

$$I = \frac{\mathbf{l}_1\mathbf{x}_1}{\mathbf{l}_2\mathbf{x}_1} \cdot \frac{\mathbf{l}_2\mathbf{x}_2}{\mathbf{l}_1\mathbf{x}_2} \tag{2.21}$$

όπου $\mathbf{l}_i\mathbf{x}_j = \alpha_i x_j + b_i y_j + c$ είναι η αλγεβρική απόσταση από την ευθεία $\mathbf{l}_i$ και το σημείο $\mathbf{x}_j$. Μπορεί εύκολα να αποδειχθεί ότι η αναλλοίωτη αυτή είναι μια διαφορετική αναπαράσταση του cross ratio [86].

2.5.3 Δύο ομοεπίπεδα κωνοειδή

Όπως ένας κύκλος στην Ευκλίδεια γεωμετρία ορίζεται σαν το γεωμετρικό τόπο των σημείων που ισαπέχουν από ένα δεδομένο σημείο, έτσι και το κωνοειδές στην προβολική γεωμετρία ορίζεται σαν το γεωμετρικό τόπο των σημείων με σταθερό cross ratio σε τέσσερα δεδομένα σημεία, τρία εκ των οποίων δεν είναι συνευθειακά. Παρατηρούμε ότι και στις δύο περιπτώσεις η μορφή της καμπύλης ορίζεται αναφορικά με μια αναλλοίωτη ποσότητα της συγκεκριμένης γεωμετρίας, της απόστασης στην περίπτωση της Ευκλίδειας, και του cross ratio στην περίπτωση της προβολικής. Η εξίσωση ενός κωνοειδούς έχει την εξής μορφή:

$$\mathbf{p}^T C \mathbf{p} = 0 \tag{2.22}$$

όπου το $\mathbf{p}$ είναι ένα 3×1 διάνυσμα και ο C είναι ένας συμμετρικός 3×3 πίνακας συντελεστών για την κωνική τομή. Ένα ζεύγος ομοεπίπεδων κωνοειδών έχει δύο προβολικές αναλλοίωτες. Οι δύο αναλλοίωτες ενός ζεύγους ομοεπίπεδων κωνοειδών με $|C_i| = 1$ έχουν τους εξής τύπους:

$$I_1 = Trace\left[C_1^{-1}C_2\right] \tag{2.23}$$

$$I_2 = Trace\left[C_2^{-1}C_1\right] \tag{2.24}$$

2.5.4 Διδιάστατοι cross ratios

Στις δυο διαστάσεις απαιτείται μια συλλογή πέντε σημείων $\mathrm{A}_1, \ldots, A_5$. Χρησιμοποιώντας ομογενείς συντεταγμένες τα σημεία περιγράφονται από τις τριάδες (x_i, y_i, z_i) και οιοσδήποτε μη-παθολογικός 3×3 πίνακας ορίζει ένα επίπεδο προβολικό μετασχηματισμό της συλλογής. Τα σημεία υπόκεινται μόνο στον περιορισμό ότι τρία από αυτά δεν πρέπει να είναι συνευθειακά. Υπάρχουν πολλοί ισοδύναμοι ορισμοί για το cross-ratio στις δύο διαστάσεις. Ο υπολογιστικά πιο ασφαλής είναι ο λόγος των εμβαδών τεσσάρων τριγώνων, που έχουν μία κοινή κορυφή. Η κορυφή αυτή ονομάζεται *κοινό*

σημείο της συλλογής. Αν το κοινό σημείο είναι το A_1 τότε το cross-ratio, μιας συλλογής σημείων $(\mathrm{A}_1, \mathrm{A}_2, \mathrm{A}_3, \mathrm{A}_4, A_5)$, ορίζεται ως:

$$\begin{aligned}\mu &= [\mathrm{A}_1, \mathrm{A}_2, \mathrm{A}_3, \mathrm{A}_4, A_5] \\ &= \frac{(\Delta A_1 A_2 A_3)(\Delta A_1 A_3 A_5)}{(\Delta A_1 A_3 A_4)(\Delta A_1 A_2 A_5)} \\ &= \frac{\begin{vmatrix} x_1 & x_2 & x_4 \\ y_1 & y_2 & y_4 \\ z_1 & z_2 & z_4 \end{vmatrix} \begin{vmatrix} x_1 & x_3 & x_5 \\ y_1 & y_3 & y_5 \\ z_1 & z_3 & z_5 \end{vmatrix}}{\begin{vmatrix} x_1 & x_3 & x_4 \\ y_1 & y_3 & y_4 \\ z_1 & z_3 & z_4 \end{vmatrix} \begin{vmatrix} x_1 & x_2 & x_5 \\ y_1 & y_2 & y_5 \\ z_1 & z_2 & z_5 \end{vmatrix}}\end{aligned} \tag{2.25}$$

όπου το $\Delta A_i A_j A_k$ είναι το προσανατολισμένο εμβαδόν της τριάδας που ορίζεται από τα σημεία A_i, A_j και A_k. Αν το A_i είναι ένα σημείο του καρτεσιανού επιπέδου, τότε μπορούμε να επιλέξουμε $z_i = 1$ ειδάλλως, αν είναι ιδεατό, επιλέγουμε $z_i = 0$. Στη δική μας περίπτωση, θεωρούμε ότι τα σημεία μας είναι καρτεσιανά, με αποτέλεσμα τα z_i να έχουν πάντα τιμή ένα.

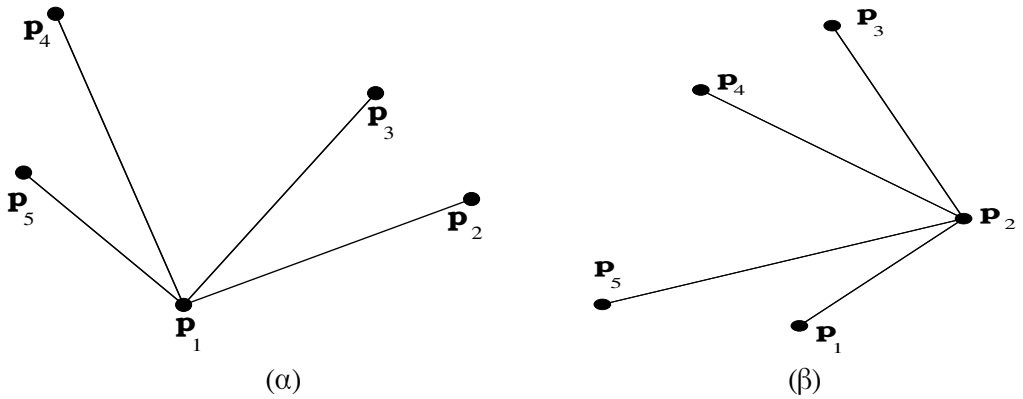
Θεωρώντας σαν κοινό σημείο όλα τα σημεία της συλλογής, μπορούν να υπολογιστούν πέντε διαφορετικοί διδιάστατοι cross-ratios για μια συλλογή από πέντε ομοεπίπεδα σημεία. Εκφράσεις των τεσσάρων υπολοίπων cross-ratios μπορούν να παραχθούν από την εξίσωση 2.25 εναλλάσοντας τους δείκτες 1 και k για $k = 2, \ldots, 5$. Έτσι το cross-ratio που σχετίζεται με το σημείο A_2 είναι:

$$\begin{aligned}\nu &= [\mathrm{A}_1, \mathrm{A}_2, \mathrm{A}_3, \mathrm{A}_4, A_5] \\ &= \frac{(\Delta A_2 A_1 A_4)(\Delta A_2 A_3 A_5)}{(\Delta A_2 A_3 A_4)(\Delta A_2 A_1 A_5)}\end{aligned} \tag{2.26}$$

Όπως θα δειχθεί πιο κάτω, οι δύο αυτοί ορισμοί σχετίζονται και η επιλογή αυτή μας οδηγεί σε απλούστερες εκφράσεις αναλλοίωτων. Μόνο δύο από τους cross-ratios είναι λειτουργικά ανεξάρτητοι[9] αφού τα πέντε σημεία συνθέτουν ένα 10-διάστατο χώρο, ενώ ένας επίπεδος προβολικός μετασχηματισμός έχει 8 βαθμούς ελευθερίας. Οποιοδήποτε ζεύγος cross-ratios, από τους πέντε, μπορεί να αποτελέσει το βασικό ανεξάρτητο ζεύγος, βάσει του οποίου, θα εκφραστούν οι υπόλοιποι τρεις.

Δύο κατασκευές [78] μπορούν να χρησιμοποιηθούν για να παρουσιαστεί η γεωμετρική ερμηνεία τη προβολικής αναλλοίωτης πέντε ομοεπίπεδων σημείων $\mathbf{p}_1 \ldots \mathbf{p}_5$. Στην εικόνα 2.2 απεικονίζονται οι δύο αυτές κατασκευές. Για την πρώτη κατασκευή ας θεωρήσουμε τις ευθείες από το $\mathbf{p}_1$ σε οποιοδήποτε άλλο σημείο, όπως φαίνεται στο σχήμα 2.2-α. Αυτό οδηγεί σε ένα σύστημα τεσσάρων ευθειών που διέρχονται από το ίδιο σημείο

[9] οι υπόλοιποι τρεις εκφράζονται συναρτήσει αυτών



Σχήμα 2.2: Γεωμετρικές ερμηνείες των διδιάστατων cross-ratios.

και το cross ratio αυτών των ευθειών είναι μια αναλλοίωτη ποσότητα. Η δεύτερη αναλλοίωτη ποσότητα παρατηρείται κατασκευάζοντας τις ευθείες με αφετηρία το σημείο $\mathbf{p}_2$, όπως φαίνεται στο σχήμα 2.2-β και υπολογίζοντας το cross ratio αυτών των ευθειών.

Μια τρίτη κατασκευή βασίζεται στο γεγονός ότι, στη γενικότητα, τέσσερα γνωστά σημεία έχουν οκτώ βαθμούς ελευθερίας οι οποίοι απαιτούνται για τον μονοσήμαντο μετασχηματισμό στο κανονικοποιημένο πλαίσιο. Το πέμπτο σημείο έχει δύο βαθμούς ελευθερίας, οι οποίοι αποτελούν τις τιμές των δύο αναλλοίωτων [78, 87].

2.5.5 Αριθμητική ευστάθεια των γεωμετρικών αναλλοιώτων

Στο άρθρο των Coelho et al. [86], επιτελείται μια πειραματική αξιολόγηση της αριθμητικής αξιοπιστίας των τεσσάρων προαναφερθέντων επίπεδων προβολικών αναλλοίωτων, παραθέτοντας τις τιμές των σφαλμάτων τους σε υπολογισμούς πληθυσμών σημείων ιδωμένων κάτω από αρκετά διαφορετικές οπτικές γωνίες.

Το συμπέρασμα στο οποίο καταλήγουν οι συγγραφείς, είναι ότι οι συγκεκριμένες γεωμετρικές αναλλοίωτες μεταβάλλουν αμυδρά τις τιμές τους για σημαντικές αλλαγές στην οπτική γωνία, δεν επηρεάζονται από τον θόρυβο κβαντοποίησης και ο βασικός παράγοντας καθορισμού της ακρίβειας και της σταθερότητας των είναι η σχετικά χαμηλή απόδοση των αλγορίθμων εντοπισμού χαρακτηριστικών. Το γεγονός αυτό φαίνεται και από το γεγονός ότι ο Sinclair et al. [87] μοντελοποιεί τη διαφορά ανάμεσα σε δύο διδιάστατες αναλλοίωτες σαν συνάρτηση της διασποράς της αβεβαιότητας των συντεταγμένων των χαρακτηριστικών της εικόνας.

2.5.6 Συνόψιση των χαρακτηριστικών της προβολικής γεωμετρίας

Οι σχέσεις ανάμεσα στις διαφορετικές γεωμετρίες, οι επιτρεπόμενοι μετασχηματισμοί στην καθεμιά καθώς και οι μετρικές ποσότητες που παραμένουν αναλλοίωτες κάτω από αυτούς τους μετασχηματισμούς συνοψίζονται στον πίνακα 2.1

	Γεωμετρίες			
	Ευκλίδεια	Ομοιότητας	Αφφινική	Προβολική
Μετασχηματισμοί				
Ευκλίδειος	X	X	X	X
ομοιόμορφη κλιμάκωση		X	X	X
ανομοιόμορφη κλιμάκωση			X	X
shear			X	X
προοπτική προβολή				X
σύνθεση προβολικών				X
Αναλοίωτες ποσότητες				
μήκος	X			
γωνία	X	X		
λόγος μηκών	X	X		
παραλληλία	X	X	X	
πρόσπτωση	X	X	X	X
cross ratio	X	X	X	X

Πίνακας 2.1: Γεωμετρίες - Μετασχηματισμοί - Αναλλοίωτες ποσότητες

2.6 Μεταθετικές και Προβολικές αναλλοίωτες

Στη συνέχεια, εισάγουμε τον αναγνώστη στον κόσμο των προβολικά και μεταθετικά αναλλοίωτων εκφράσεων, οι οποίες αποτελούν τον ακρογωνιαίο λίθο της εργασίας μας. Οι αναλλοίωτες ενός συνόλου χαρακτηριστικών έχουν αποτελέσει ένα συχνά χρησιμοποιούμενο εργαλείο στον τομέα της Μηχανικής Όρασης. Δοθέντος του συνόλου χαρακτηριστικών s και ενός τελεστή g που ανήκει σε μια ομάδα μετασχηματισμών, η αναλλοίωτη $\mathrm{I}[\cdot]$ ικανοποιεί την $\mathrm{I}[s] = \mathrm{I}[gs]$ για κάθε g στην ομάδα. Αν το g είναι ένας προβολικός μετασχηματισμός, παρατηρούνται οι προβολικές αναλλοίωτες. Βασικές εφαρμογές των προβολικών αναλλοίωτων στην μηχανική όραση είναι η αναγνώριση αντικειμένων και η ανακατασκευή της τριδιάστατης δομής από μη βαθμονομημένες κάμερες. Ένα απλό παράδειγμα μπορεί να αναδείξει την χρησιμότητα τους. Ας υποθέσουμε ότι αρκετά ομοεπίπεδα αντικείμενα είναι ορατά στο οπτικό μας πεδίο. Δύο εικόνες της σκηνής λαμβάνονται από δύο διαφορετικές θέσεις. Οι δύο εικόνες συνδέονται μέσω ενός προβολικού μετασχηματισμού g. Αν s_A το σύνολο χαρακτηριστικών της πρώτης εικόνας και s_B το σύνολο των χαρακτηριστικών της δεύτερης, τότε η προβολική αναλλοίωτη $\mathrm{I}[\cdot]$, που σχετίζεται με το s_A, ικανοποιεί την $\mathrm{I}[s_\mathrm{A}] = \mathrm{I}[g\, s_\mathrm{A}]$ για κάθε αλλαγή στην σχετική θέση των δύο καμερών. Έτσι το πρόβλημα της ταύτισης του συνόλου των χαρακτηριστικών s_A στην δεύτερη εικόνα ανάγεται στην διερεύνηση των σχέσεων ανάμεσα στις προβολικές αναλλοίωτες $\mathrm{I}[s_\mathrm{A}]$ και $\mathrm{I}[s_\mathrm{B}]$, ανεξάρτητα από τους μετασχηματισμούς μεταξύ των δύο εικόνων.

Οι περισσότερες προβολικές αναλλοίωτες είναι ευαίσθητες ως προς την μετάθεση.

Αυτό σημαίνει, ότι η τιμή τους εξαρτάται από την διάταξη με την οποία θεωρούνται τα χαρακτηριστικά στον υπολογισμό. Μια διαφορετική διάταξη στο σύνολο, επιφέρει συνήθως και μια διαφορετική τιμή στην προβολικά αναλλοίωτη ποσότητα. Στο παράδειγμα που αναφέραμε, η ευαισθησία της μετάθεσης της προβολικής αναλλοίωτης επιφέρει τεράστια υπολογιστική επιβάρυνση, αφού για την αναγνώριση του συνόλου των χαρακτηριστικών s_A στη δεύτερη εικόνα, τα χαρακτηριστικά θα πρέπει να θεωρηθούν με την ίδια διάταξη και στις δύο εικόνες κατά τον υπολογισμό των αναλλοίωτων ποσοτήτων. Τούτο όμως σημαίνει ότι θα πρέπει να γνωρίζουμε την μεταξύ των αντιστοίχιση πριν τον υπολογισμό των αναλλοίωτων. Ωστόσο η πληροφορία αυτή δεν είναι πάντα διαθέσιμη, με αποτέλεσμα να απαιτείται ο έλεγχος όλων των πιθανών συνδυασμών για τα δύο σύνολα, ο οποίος, επιπρόσθετα, επαυξάνει την πιθανότητα λήψης εσφαλμένων αποφάσεων. Η εναλλαγή των δεικτών των στοιχείων είναι ισοδύναμη με μια ομάδα μεταθέσεων που επιδρά στο σύνολο. Η ομάδα των μεταθέσεων έχει τις δικές της μεταθετικές αναλλοίωτες[10].

Έστω ο προβολικός μετασχηματισμός T και η μετάθεση π που επιδρά στο σύνολο s. Μια προβολική αναλλοίωτη $Q[\cdot]$ ικανοποιεί την συνθήκη $Q[s] = Q[Ts]$ για κάθε, λόγου χάρη, επίπεδο προβολικό μετασχηματισμό. Η μεταθετική αναλλοίωτη $P[\cdot]$ ικανοποιεί την συνθήκη $P[s] = P[\pi s]$ για όλες τις πιθανές μεταθέσεις των δεικτών που συσχετίζονται με τα χαρακτηριστικά στο σύνολο. Μια προβολική και μεταθετική αναλλοίωτη[11] $J[\cdot]$ ενός συνόλου s πρέπει να ικανοποιεί την συνθήκη:

$$J[s] = J[\pi s] = J[Ts] = J[\pi Ts] = J[T\pi s]$$

για κάθε T στην ομάδα των μετασχηματισμών και κάθε π στην ομάδα των μεταθέσεων. Η τιμή του $J[\cdot]$ παραμένει αναλλοίωτη όταν το σύνολο υφίσταται ένα προβολικό μετασχηματισμό. Είναι επίσης ανεξάρτητη από τον τρόπο με τον οποίο σχετίζονται οι δείκτες με τα χαρακτηριστικά στο πρωταρχικό ή στο μετασχηματισμένο σύνολο.

Η μεταθετική ευαισθησία των προβολικών αναλλοίωτων σημαίνει ότι θα πρέπει να αποθηκεύεται μια αναλλοίωτη για κάθε μία από τις πιθανές διατάξεις του συνόλου. Όταν δύο σύνολα προσπαθούν να αντιστοιχιθούν μέσω παρόμοιων τιμών μιας $\mathbf{P}^2$-αναλλοίωτης, η διάταξη των στοιχείων εντός του συνόλου δεν συνεισφέρει στην απόφαση. Τα σύνολα συγκρίνονται ανεξάρτητα από την διάταξή των στοιχείων τους και επαρκεί η αποθήκευση μιας μόνο $\mathbf{P}^2$-αναλλοίωτης για κάθε σύνολο. Η μελέτη των $\mathbf{P}^2$-αναλλοίωτων στηρίζεται στις ιδιότητες της βασικής προβολικής αναλλοίωτης, του cross-ratio.

Οι αναλλοίωτες μιας ομάδας μετασχηματισμών μπορούν να κατασκευαστούν μέσω της *Θεωρίας Αναπαράστασης*, στην οποία τα στοιχεία μιας ομάδας απεικονίζονται σε πίνακες. Οι λειτουργίες στην ομάδα ισοδυναμούν με πολλαπλασιασμούς πινάκων και οι αναλλοίωτες της ομάδας, υπολογίζονται, αναγνωρίζοντας τον υπόχωρο των

[10] εκφράσεις των οποίων η τιμή παραμένει σταθερή κατά την αναδιάταξη των στοιχείων ενός συνόλου

[11] Για λόγους συντομίας την συμβολίζουμε $\mathbf{P}^2$-αναλλοίωτη

αναλλοίωτων στον αντίστοιχο χώρο διανυσμάτων. Για μια εισαγωγή στην γενική θεωρία της αναπαράστασης ομάδων, ο αναγνώστης μπορεί να συμβουλευτεί το [88] και για μια συνοπτική περιγραφή των πεπερασμένων ομάδων μια καλή πηγή αποτελεί το [89]

2.6.1 Πρώτης τάξης $\mathbf{P}^2$-αναλλοίωτες του μονοδιάστατου cross-ratio

Η $\mathbf{P}^2$-αναλλοίωτη τεσσάρων συνευθειακών σημείων κατασκευάζεται σε δύο βήματα. Κατά πρώτον, η επιρροή των προβολικών μετασχηματισμών αφαιρείται μέσω της χρήσης των cross-ratios. Η επιρροή των μεταθέσεων εξαφανίζεται, υπολογίζοντας την αναλλοίωτη που σχετίζεται με το σύνολο των cross-ratio εκφράσεων. Το cross-ratio είναι ευαίσθητο ως προς τις μεταθέσεις των σημείων μιας και μια μετάθεση π των δεικτών των τεσσάρων σημείων αλλάζει την τιμή του. Παρατηρήστε ότι μια μετάθεση των δεικτών των σημείων σημαίνει μια διαφορετική διάταξη των ίδιων τεσσάρων συγγραμικών σημείων στον υπολογισμό του cross-ratio. Η αντίστοιχη αλλαγή στο cross-ratio δηλώνεται μέσω της απεικόνισης Π:

$$\Pi(\lambda) \;=\; \left[A_{\pi(1)}, A_{\pi(2)}, A_{\pi(3)}, A_{\pi(4)}\right] \tag{2.27}$$

όπου τα $\pi(i)$ είναι άνισα και λαμβάνουν τιμές μεταξύ 1 και 4. Το σύνολο όλων των μεταθέσεων των δεικτών των σημείων αποτελεί τη συμμετρική ομάδα $\mathcal{G}_4$. Έστω η μετάθεση στο $\mathcal{G}_4$ συμβολίζεται σαν π_k. Αν π_{k_1} και π_{k_2} είναι δύο μεταθέσεις στο $\mathcal{G}_4$, η σύνθεσή τους $\pi_{k_1}\pi_{k_2}$ είναι επίσης μια μετάθεση στο $\mathcal{G}_4$. Η επίδραση της σύνθεσης των δύο μεταθέσεων στον cross-ratio είναι η σύνθεση των απεικονίσεων $\Pi_{k_1}(\Pi_{k_2}(\lambda))$

Θεώρημα 1 *[90] Όλα τα στοιχεία μιας συμμετρικής ομάδας $\mathcal{G}_4$ είναι πεπερασμένα γινόμενα τριών βασικών μεταθέσεων:*

$$\begin{aligned}
\pi_1 : [1234] &\;\mapsto\; [2134] \\
\pi_2 : [1234] &\;\mapsto\; [1324] \\
\pi_3 : [1234] &\;\mapsto\; [1243]
\end{aligned}$$

Οι βασικές αυτές μεταθέσεις εναλλάσουν μόνο τα δύο σημεία k και k+1 της εξίσωσης 2.20. Έτσι, διαμέσου μιας ακολουθίας των βασικών αυτών μεταθέσεων, τα τέσσερα σημεία μπορούν να αναδιαταχθούν σε οποιαδήποτε διάταξη κατά τον υπολογισμό του cross-ratio.

Υπάρχουν $4! = 24$ στοιχεία στην ομάδα $\mathcal{G}_4$. Ωστόσο, παράγουν μόνο έξι ξεχωριστές τιμές cross-ratio. Αυτό μπορεί να αποδειχθεί παρατηρώντας ότι οι τρεις βασικές μεταθέσεις π_k παράγουν μόνο δύο ξεχωριστές απεικονίσεις Π_k για το cross-ratio λ:

$$\Pi_1(\lambda) \;=\; \tfrac{1}{\lambda} \quad \Pi_2(\lambda) \;=\; 1-\lambda \quad \Pi_3(\lambda) \;=\; \tfrac{1}{\lambda}$$

Από το Θεώρημα 1 συνεπάγεται ότι όλες οι πιθανές απεικονίσεις Π μπορούν να παρατηρηθούν εφαρμόζοντας τις πράξεις $1/q$ και $1-q$ στο λ. Για παράδειγμα, η μετάθεση $\pi = \pi_1\pi_2$ οδηγεί στην εξής απεικόνιση για τον cross-ratio:

$$\Pi(\lambda) \;=\; (\Pi_1\Pi_2)(\lambda) \;=\; \Pi_1\left(\Pi_2(\lambda)\right) \;=\; \Pi_1(1-\lambda) \;=\; \frac{1}{1-\lambda}$$

Εκτελώντας τη διαδικασία αυτή για όλες τις δυνατές συνθέσεις, παράγονται μόνο έξι διακριτικές απεικονίσεις. Αυτές παράγουν τις ακόλουθες τιμές για τα cross-ratios:

$$\lambda_1 = \lambda, \quad \lambda_2 = \frac{1}{\lambda}, \quad \lambda_3 = \frac{1}{1-\lambda}, \quad \lambda_4 = 1-\lambda \quad \lambda_5 = \frac{\lambda-1}{\lambda}, \quad \lambda_6 = \frac{\lambda}{1-\lambda} \tag{2.28}$$

Οι έξι εκφράσεις λ_i ορίζουν ένα 6-αδιάστατο διάνυσμα Λ. Οι συναρτήσεις στο διάνυσμα Λ συνδέουν ένα 6-αδιάστατο χώρο διανυσμάτων $\mathbf{V}_1$.

Οι δύο απεικονίσεις $\Pi_1(\lambda) = \Pi_3(\lambda)$ και $\Pi_2(\lambda)$ γίνονται οι 6×6 πίνακες $P_1 = P_3$ και P_2:

$$P_1 \;=\; P_3 \;=\; \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 \end{bmatrix} \qquad P_2 \;=\; \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \tag{2.29}$$

Για παράδειγμα, ο πολλαπλασιασμός του P_1 με το 6-αδιάστατο διάνυσμα των λ_i περιγράφει την επίδραση της εναλλαγής των δεικτών του πρώτου και του δεύτερου ζεύγους σημείων στον cross-ratio.

$$P_1\Lambda \;=\; \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \\ \lambda_5 \\ \lambda_6 \end{bmatrix} \;=\; \begin{bmatrix} \lambda_2 \\ \lambda_1 \\ \lambda_4 \\ \lambda_3 \\ \lambda_6 \\ \lambda_5 \end{bmatrix} \;=\; \begin{bmatrix} \Pi_1(\lambda_1) \\ \Pi_1(\lambda_2) \\ \Pi_1(\lambda_3) \\ \Pi_1(\lambda_4) \\ \Pi_1(\lambda_5) \\ \Pi_1(\lambda_6) \end{bmatrix} \;=\; \Pi_1(\Lambda) \tag{2.30}$$

Από τον ορισμό των λ_i στην εξίσωση 2.28 παρατηρούμε άμεσα ότι η εξίσωση 2.30 αντιστοιχεί στην αντικατάσταση του λ με το 1/λ στις εκφράσεις για τα λ_i.

Ορισμός 1 *Μια απεικόνιση* Φ *από το* $\mathcal{G}_4$ *στο σύνολο των μη-παθολογικών* $n \times n$ *πινάκων που ικανοποιεί την:*

$$\Phi\left(\pi_{k_1}\pi_{k_2}\right) \;=\; \Phi\left(\pi_{k_1}\right)\Phi\left(\pi_{k_2}\right) \tag{2.31}$$

για όλα τα $\pi_{k_1}, \pi_{k_2} \in \mathcal{G}_4$, *ονομάζεται n-διάστατη αναπαράσταση του* $\mathcal{G}_4$.

Στην δική μας περίπτωση έχουμε, $\mathbf{\Phi}(\pi_1) = \mathbf{\Phi}(\pi_3) = P_1$, $\mathbf{\Phi}(\pi_2) = P_2$ και τους 24 πίνακες $\mathbf{\Phi}(\pi)$ των μεταθέσεων $\pi \in \mathcal{G}_4$ που είναι γινόμενα των P_1 και P_2 (βλέπε Θεώρημα 1).

Ο υπόχωρος $\mathbf{U}$ του $\mathbf{V}_1$ είναι ένας *υπόχωρος αναλλοίωτων* αν για όλα τα στοιχεία $u \in \mathbf{U}$ και όλα τα $\pi \in \mathcal{G}_4$ έχουμε $\mathbf{\Phi}(\pi)\, u \in \mathbf{U}$. Ένας υπόχωρος αναλλοίωτων είναι *ανάγωγος* αν δεν περιέχει κατάλληλους υπόχωρους αναλλοίωτων. Μπορεί να αποδειχθεί, με χρήση κλασικών εργαλείων από τη θεωρία αναπαράστασης [91], ότι ο χώρος $\mathbf{V}_1$ μπορεί να αποσυντεθεί σε τέσσερις ανάγωγους υποχώρους με διαστάσεις: 1, 1, 2, και 2. Ο απλούστερος υπόχωρος αναλλοίωτων είναι ο μονοδιάστατος υπόχωρος που συνδέεται με το διάνυσμα $(1, 1, 1, 1, 1, 1)$. Σε αυτό τον υποχώρο έχουμε $\mathbf{\Phi}(\pi)\, u = u$ για κάθε $\pi \in \mathcal{G}_4$.

Οι ανάγωγοι υπόχωροι περιέχουν διανύσματα που είναι αναλλοίωτα κάτω από την απεικόνιση $\mathbf{\Phi}$ της μεταθετικής ομάδας $\mathcal{G}_4$. Η προβολή του Λ στον πρώτο μονοδιάστατο ανάγωγο υποχώρο παράγει μια αναγκαία $\mathbf{P}^2$-αναλλοίωτη. Η προβολή προκύπτει υπολογίζοντας το εσωτερικό γινόμενο μεταξύ του διανύσματος $(1, 1, 1, 1, 1, 1)$ και του Λ:

$$I_1[\lambda] \;=\; \sum_{i=1}^{6} \lambda_i \;=\; 3 \tag{2.32}$$

Η αναλλοίωτη αυτή είναι τετριμμένη και άνευ σημασίας στις εφαρμογές. Καταλήγουμε δηλαδή στο ότι οι γραμμικοί συνδυασμοί των εκφράσεων των cross-ratios παρέχουν μια τετριμμένη p^2-αναλλοίωτη κάτω από την απεικόνιση $\mathbf{\Phi}$ της μεταθετικής ομάδας $\mathcal{G}_4$ στο χώρο διανυσμάτων.

Ένα πιο ενδιαφέρον αποτέλεσμα μπορεί να παρατηρηθεί αν δεν υπάρξει απαίτηση για πλήρη αναλλοίωση άλλα μόνο αναλλοίωση μέχρι ένα πολλαπλασιασμό με το 1 ή -1. Το είδος αυτό της αναλλοίωσης συνδέεται με το διάνυσμα $(1, -1, 1, -1, 1, -1)$ και παράγει την συνάρτηση $\mathrm{I}_2[\lambda]$:

$$\mathrm{I}_2[\lambda] \;=\; \frac{2\lambda^3 \,-\, 3\lambda^2 \,-\, 3\lambda \,+\, 2}{\lambda(\lambda \,-\, 1)} \tag{2.33}$$

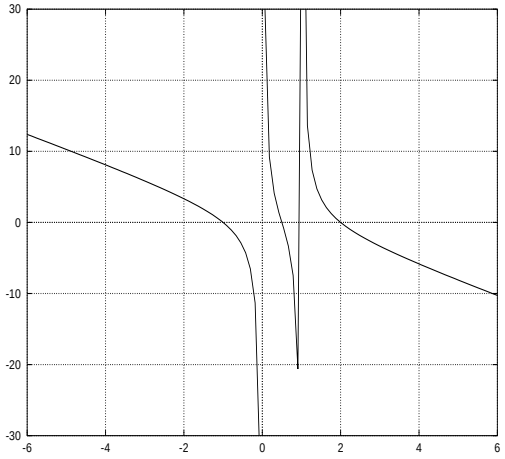
που μετασχηματίζεται σαν την εναλλακτική αναπαράσταση του $\mathcal{G}_4$

$$\mathrm{I}_2[\lambda] \;=\; -\mathrm{I}_2\left[\frac{1}{\lambda}\right] \;=\; \mathrm{I}_2\left[\frac{1}{1-\lambda}\right] \;=\; -\mathrm{I}_2\left[1-\lambda\right] \;=\; \mathrm{I}_2\left[\frac{\lambda-1}{\lambda}\right] \;=\; -\mathrm{I}_2\left[\frac{\lambda}{\lambda-1}\right] \tag{2.34}$$

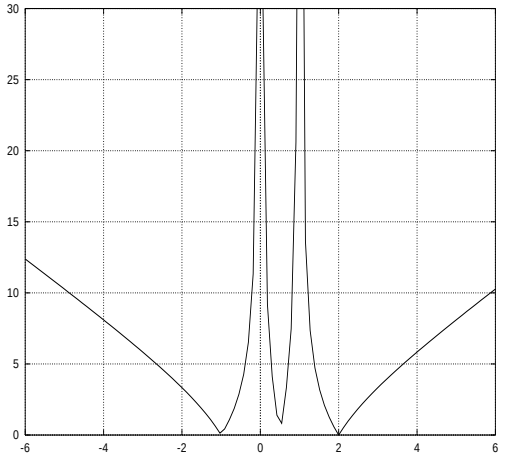
Στις εφαρμογές, η πληροφορία που περιέχεται στο πρόσημο του $\mathrm{I}_2[\lambda]$ είναι συνήθως άνευ σημασίας. Οπότε μπορούμε να χρησιμοποιήσουμε το $|\mathrm{I}_2[\lambda]|$. Οι γραφικές παραστάσεις των συναρτήσεων $-\mathrm{I}_2[\lambda]$ και $|\mathrm{I}_2[\lambda]|$ απεικονίζονται στις εικόνες 2.3 και 2.4 αντίστοιχα.

Η μελέτη άλλων ανάγωγων υπόχωρων οδηγεί στις συναρτήσεις:

$$\mathrm{I}_3[\lambda] \;=\; \frac{\lambda^2 \,-\, \lambda \,+\, 1}{\lambda(\lambda \,-\, 1)}, \qquad \mathrm{I}_4[\lambda] \;=\; \frac{\lambda^2 \,-\, \lambda \,+\, 1}{(\lambda \,-\, 1)} \tag{2.35}$$



Σχήμα 2.3: Η αναλλοίωτη $-\mathrm{I}_2$



Σχήμα 2.4: Η αναλλοίωτη $|\mathrm{I}_2|$

Οι υπόλοιπες αναλλοίωτες συνήθως εκφράζονται σαν γραμμικοί συνδυασμοί των $\mathrm{I}_1, \ldots, I_4$, με αποτέλεσμα να καθίστανται ακατάλληλοι για πρακτικές εφαρμογές.

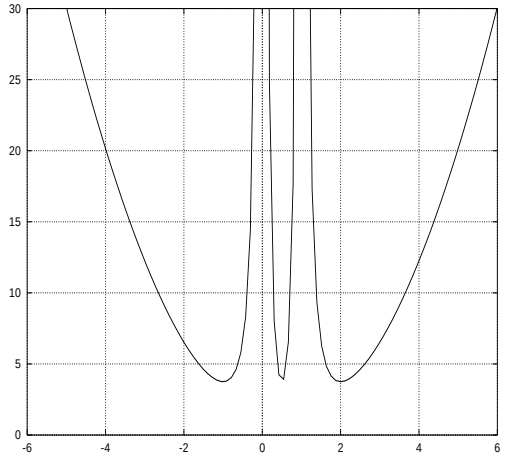
Στην επόμενη ενότητα δείχνουμε πως μέσω μιας διαφορετικής αναπαράστασης του $\mathcal{G}_4$, που βασίζεται σε ανά δύο γινόμενα $\lambda_i\lambda_j$, μπορούν να βρεθούν χρήσιμες $\mathbf{P}^2$-αναλλοίωτες για το σύνολο των τεσσάρων συνευθειακών σημείων.

2.6.2 Δεύτερης τάξης αναλλοίωτες στη μία διάσταση

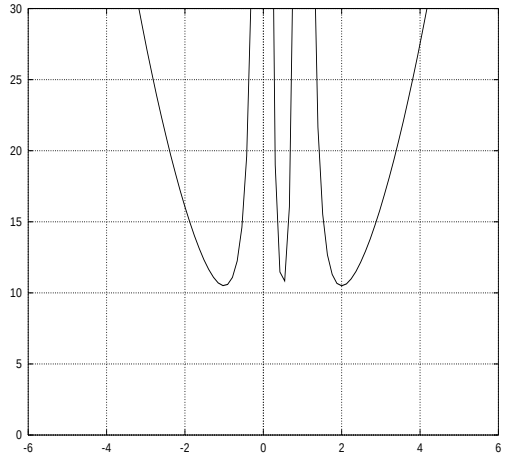
Ορίζουμε μια νέα απεικόνιση από τον $\mathcal{G}_4$ σε πίνακες που επιδρούν σε ένα χώρο διανυσμάτων $\mathbf{V}_2$. Από τα 21 γινόμενα $\lambda_i\lambda_j$, $i \geq j$ $i, j = 1 \ldots 6$, τρία είναι ίσα με ένα. Τα εναπομείναντα, συνθέτουν ένα χώρο διανυσμάτων, $\mathbf{V}_2$, διάστασης 18. Ακολουθώντας την προηγούμενη διαδικασία παρατηρούμε μία 18-διάστατη αναπαράσταση του $\mathcal{G}_4$. Η αναπαράσταση αυτή οδηγεί στις παρακάτω $\mathbf{P}^2$-αναλλοίωτες:

$$\begin{aligned} J_1[\lambda] &= \frac{\lambda^6 - 3\lambda^5 + 3\lambda^4 - \lambda^3 + 3\lambda^2 - 3\lambda + 1}{\lambda^2(\lambda - 1)^2} \\ J_2[\lambda] &= \frac{2\lambda^6 - 6\lambda^5 + 9\lambda^4 - 8\lambda^3 + 9\lambda^2 - 6\lambda + 2}{\lambda^2(\lambda - 1)^2} \\ J_3[\lambda] &= 3 \\ J_4[\lambda] &= -3 \end{aligned} \tag{2.36}$$

Οι $\mathbf{P}^2$-αναλλοίωτες θεωρούνται δευτέρας τάξης μιας και κατασκευάστηκαν χρησιμοποιώντας γινόμενα εκφράσεων cross-ratios. Όπως φαίνεται, μόνο οι δύο πρώτες είναι μη-τετριμμένες. Στις εικόνες 2.5 και 2.6, όπου απεικονίζονται οι συναρτήσεις $J_1[\lambda]$ και $J_2[\lambda]$, είναι εμφανείς οι ασυνέχειες στο 0 και στο 1.



Σχήμα 2.5: Η αναλλοίωτη J_1



Σχήμα 2.6: Η αναλλοίωτη J_2

Μπορεί να αποδειχθεί ότι όλες οι δευτεροβάθμιες $\mathbf{P}^2$-αναλλοίωτες τεσάρων συνευθειακών σημείων είναι γραμμικοί συνδυασμοί των τεσσάρων αναλλοίωτων της εξίσωσης 2.36. Μερικοί από αυτούς τους συνδυασμούς είναι αρκετά γνωστοί. Για παράδειγμα η $\mathbf{P}^2$-αναλλοίωτη $J_{11}[\lambda]$, με τύπο

$$J_{11}[\lambda] \;=\; \frac{\lambda^2 - \lambda + 1^2}{\lambda^2(\lambda - 1)^2} \;=\; J_1[\lambda] + J_3[\lambda] \tag{2.37}$$

έχει συχνά αναφερθεί στη μαθηματική βιβλιογραφία [92] και χρησιμοποιήθηκε στην Μηχανική Όραση από τον Maybank [93] για την διερεύνηση μη-επιπέδων κωνικών καμπύλων. Ο γραμμικός συνδυασμός

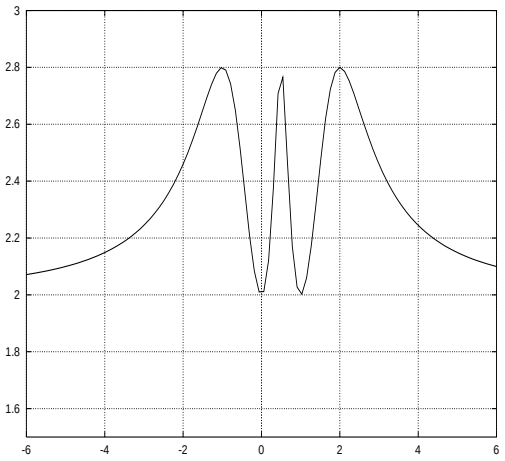
$$J_{12}[\lambda] \;=\; \sum_{i=1}^{6} \sum_{j=1,\, j\neq i}^{6} \lambda_i \lambda_j \;=\; J_1[\lambda] + J_4[\lambda] \tag{2.38}$$

είναι μια διαισθητική $\mathbf{P}^2$-αναλλοίωτη που μπορεί να παραχθεί σαν μια συμμετρική συνάρτηση των έξι λ_i [94]. Μια άλλη συμμετρική συνάρτηση είναι η:

$$J_{13}[\lambda] \;=\; \sum_{i=1}^{6} \lambda_i^2 \;=\; -J_1[\lambda] + J_2[\lambda] - \tfrac{1}{2} \cdot J_3[\lambda] \tag{2.39}$$

Μπορούμε να δημιουργήσουμε αναλλοίωτες χρησιμοποιώντας μη γραμμικούς συνδυασμούς των τεσσάρων βασικών αναλλοίωτων. Με τον τρόπο αυτό προκύπτουν φραγμένες $\mathbf{P}^2$-αναλλοίωτες των τεσσάρων σημείων. Μία περίπτωση είναι να χρησιμοποιήσουμε λόγους των $J_1[\lambda]$ και $J_2[\lambda]$, αφού καμία από αυτές τις συναρτήσεις δεν έχει πραγματικές ρίζες. Μια άλλη $\mathbf{P}^2$-αναλλοίωτη είναι ο λόγος των $J_{11}[\lambda]$ και $J_{12}[\lambda]$. Στην εικόνα 2.7 απεικονίζεται η φραγμένη $\mathbf{P}^2$-αναλλοίωτη:

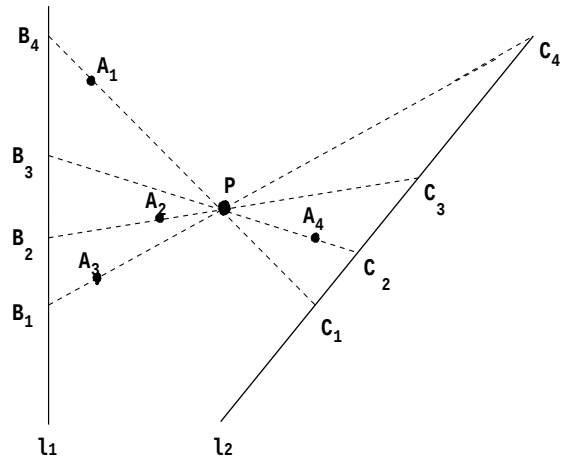
$$J_{14}[\lambda] \;=\; \frac{J_2[\lambda]}{J_1[\lambda]} \;=\; \frac{2\lambda^6 - 6\lambda^5 + 9\lambda^4 - 8\lambda^3 + 9\lambda^2 - 6\lambda + 2}{\lambda^6 - 3\lambda^5 + 3\lambda^4 - \lambda^3 + 3\lambda^2 - 3\lambda + 1} \tag{2.40}$$



Σχήμα 2.7: Η αναλλοίωτη J_{14}

Μια σύγκριση των εικόνων 2.4, 2.5, 2.6 και 2.7 δείχνει ότι όλες οι αναλλοίωτες έχουν την ίδια δομή που αποτελείται από έξι μονότονους κλάδους στα διαστήματα: $(-\infty,-1]\,,(-1,0)\,,\left[0,\frac{1}{2}\right]\,,\left(\frac{1}{2},1\right]\,,(1,2)\,,[2,\infty]$. Όλα τα ακρότατα είναι πάντα στα όρια αυτών των διαστημάτων.

Η χρησιμότητα των $\mathbf{P}^2$-αναλλοίωτων στη μία διάσταση μπορεί να εκφραστεί γραφικά μέσω ενός απλού παραδείγματος, που απεικονίζεται στην εικόνα 2.8. Έστω τέσσερα



Σχήμα 2.8: Ένα παράδειγμα της μονοδιάστατης $\mathbf{P}^2$-αναλλοίωτης αναπαράστασης

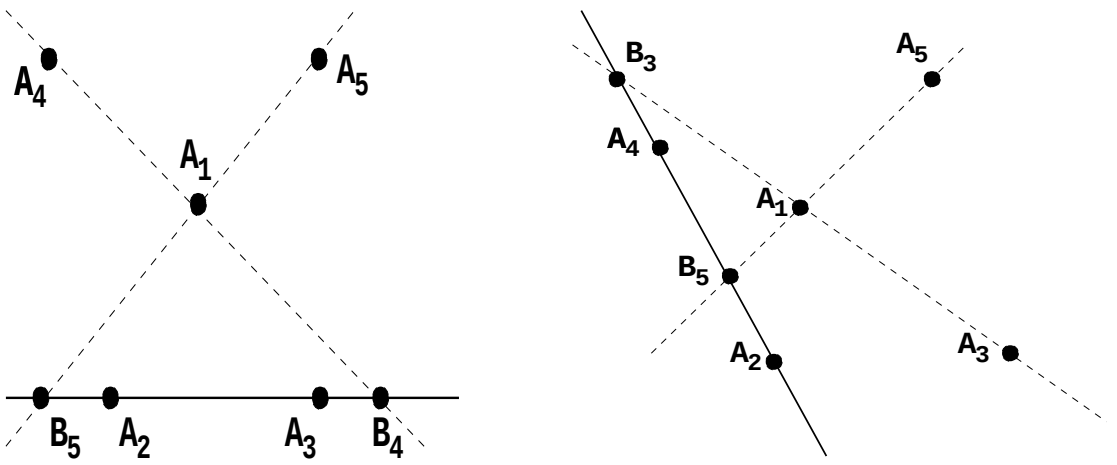
ομοεπίπεδα σημεία $\mathrm{A}_1(2.5, 14.5)$, $\mathrm{A}_2(5, 11)$, $\mathrm{A}_3(9, 10)$, $\mathrm{A}_4(3, 9.5)$, που προβάλλονται στον άξονα y και σε μια ευθεία που διέρχεται από τη αρχή των αξόνων με κλίση 45 μοίρες. Το κέντρο της προβολής είναι το σημείο $P(7.5, 10.5)$. Τα δύο σύνολα των συνευθειακών σημείων $B_1, \ldots, B_4$ και $C_1, \ldots, C_4$ βρίσκονται σε προβολική αντιστοίχιση. Ωστόσο, τα σημεία επιγράφονται διαφορετικά. Ας υποθέσουμε ότι μόνο οι δύο μονοδιάστατες εικόνες των συνόλων είναι διαθέσιμες. Όταν το cross-ratio υπολογιστεί για τα δύο σύνολα, σύμφωνα με την εξίσωση 2.20, οι τιμές που θα παράγει θα είναι -1.446 και

2.446 αντίστοιχα. Για να δείξουμε ότι τα σύνολα των σημείων είναι πράγματι σε προβολική αντιστοίχιση πρέπει να βρεθεί ο επαρκής μετασχηματισμός της εξίσωσης 2.28. Μια $\mathbf{P}^2$-αναλλοίωτη, ωστόσο, θα δώσει την ίδια τιμή για το cross-ratio, 7.464, αποδεικνύοντας την αντιστοίχιση με μόνο μία σύγκριση. Για το συγκεκριμένο παράδειγμα, χρησιμοποιήσαμε την αναλλοίωτη της εξίσωσης 2.37.

Οι $\mathbf{P}^2$-αναλλοίωτες για τα τέσσερα συνευθειακά σημεία κατασκευάστηκαν χρησιμοποιώντας την ειδική δομή των παραγόμενων εκφράσεων των έξι cross-ratios. Στην περίπτωση των δύο διαστάσεων, δεν υπάρχει κάποια αντίστοιχη δομή. Ωστόσο η κατασκευή τους επιτυγχάνεται, με συνδυασμούς των μονοδιάστατων $\mathbf{P}^2$-αναλλοίωτων, που περιγράψαμε.

2.6.3 $\mathbf{P}^2$-αναλλοίωτες των διδιάστατων cross ratios

Ο διδιάστατος cross-ratio είναι και αυτός ευαίσθητος ως προς την μετάθεση των σημείων. Η ευαισθησία αυτή μπορεί να αναλυθεί με την βοήθεια του σχήματος 2.9. Το κοινό σημείο των τεσσάρων τριγώνων, έστω το A_1, λαμβάνεται σαν το κέντρο



(α) (β)

Σχήμα 2.9: Ορισμοί του cross-ratio με κοινό σημείο το A_1

προβολής από το οποίο τα υπόλοιπα τέσσερα σημεία προβάλλονται στην ευθεία που ορίζεται από δύο από αυτά, έστω τα A_2 και A_3, όπως φαίνεται στην εικόνα 2.9-α. Ο μονοδιάστατος cross-ratio των τεσσάρων συνευθειακών σημείων $[B_5, A_2, A_3, B_4]$ συνδέεται με τον διδιάστατο cross-ratio της εξίσωσης 2.25. Όταν ένα διαφορετικό ζεύγος σημείων, έστω τα A_2 και A_4, χρησιμοποιούνται για να ορίσουν την ευθεία, το νέο σύνολο σημείων $[B_3, A_2, A_3, B_4]$, που απεικονίζεται στην εικόνα 2.9-β, είναι σε προβολική αντιστοίχιση με το σύνολο $[B_5, A_2, A_3, B_4]$, που φαίνεται στην εικόνα 2.9-α.

Ωστόσο, η διάταξη των αντίστοιχων σημείων αλλάζει και μπορεί να οδηγήσει σε μια διαφορετική τιμή για το cross-ratio, όπως είδαμε.

Ο Haralick [95] χρησιμοποίησε τον μονοδιάστατο ορισμό για να ερευνήσει τις σχέσεις ανάμεσα στους πέντε διδιάστατους cross-ratios που υπολογίζονται για τα πέντε σημεία. Για να λύσει το πρόβλημα της μεταθετικής ευαισθησίας λάμβανε υπόψιν του τις πιο πιθανές τοπολογίες του συνόλου των πέντε σημείων. Οι $\mathbf{P}^2$-αναλλοίωτες που θα αναλυθούν πιο κάτω είναι αμετάβλητες ως προς την μετάθεση των δεικτών των πέντε σημείων και κατά συνέπεια ανεξάρτητες από την τοπολογία του συνόλου. Αν μ ο διδιάστατος cross-ratio μιας δοθείσας συλλογής σημείων, τότε όπως και στην μονοδιάστατη περίπτωση, ορίζουμε σαν π_k την μετάθεση που εναλλάσει το k και το k+1 σημείο και σαν $\Pi_k(\mu)$ τον αντίστοιχο μετασχηματισμό του διδιάστατου cross-ratio της εξίσωσης 2.25.

Θεώρημα 2 *[90] Οι βασικές μεταθέσεις π_2, π_3 και π_4 οδηγούν σε δύο διαφορετικούς μετασχηματισμούς για τους διδιάστατους cross-ratios:*

$$\Pi_2(\mu) \;=\; \Pi_4(\mu) \;=\; \tfrac{1}{\mu} \quad \textit{και} \quad \Pi_3(\mu) \;=\; 1 - \mu \tag{2.41}$$

Μια αυθαίρετη μετάθεση των πέντε σημείων μπορεί να γραφτεί σαν το γινόμενο των δύο μεταθέσεων $\pi\pi^{(k)}$ όπου π είναι μια μετάθεση των τελευταίων τεσσάρων σημείων και η $\pi^{(k)}$ είναι η ειδική μετάθεση που εναλλάσει το πρώτο και το k-οστό σημείο. Έτσι $\pi^{(k)} \neq \pi_k$ αν $k \neq 1$. Οι αντίστοιχοι μετασχηματισμοί των cross-ratios είναι $\Pi(\mu)$ και $\Pi^{(k)}(\mu)$ και έτσι ο μετασχηματισμός του cross ratio μιας αυθαίρετης μετάθεσης μπορεί να γραφτεί ως: $\Pi\left(\Pi^{(k)}(\mu)\right)$. Οι μεταθέσεις των τεσσάρων τελευταίων σημείων μπορούν πάντα να εκφραστούν σαν γινόμενα των π_2, π_3 και π_4. Από το θεώρημα 2 παρατηρούμε ότι η επίδραση του Π στον μ είναι ένας από τους έξι βασικούς μετασχηματισμούς του μονοδιάστατου cross-ratio της εξίσωσης 2.28. Οι συναρτήσεις $J[\cdot]$ που έχουν διερευνηθεί, είναι αμετάβλητες κάτω από τους μετασχηματισμούς αυτούς και έτσι έχουμε:

$$J\left[\left(\Pi\Pi^{(k)}\right)(\mu)\right] \;=\; J\left[\Pi\left(\Pi^{(k)}(\mu)\right)\right] \;=\; J\left[\Pi^{(k)}(\mu)\right] \;=\; \mathcal{J}^{(k)}. \tag{2.42}$$

Η σχέση 2.42 δείχνει ότι όταν η τιμή του διδιάστατου cross-ratio αντικατασταθεί στην έκφραση μιας μονοδιάστατης αναλλοίωτης $J[\cdot]$, η παρατηρούμενη εξίσωση εξαρτάται μόνο από το σημείο που χρησιμοποιήται ως κέντρο προβολής. Οι μεταθέσεις των άλλων σημείων δεν έχουν καμία επιρροή στην τιμή.

Συνεχίζουμε τώρα στο κύριο αποτέλεσμα για την συλλογή των πέντε ομοεπίπεδων σημείων:

Θεώρημα 3 *[90] Μια μετάθεση πέντε σημείων στον υπολογισμό του διδιάστατου cross-ratio οδηγεί στην ίδια μετάθεση των τιμών $\mathcal{J}^{(1)}, \ldots \mathcal{J}^{(5)}$*

Σημειώσαμε προηγουμένως ότι υπάρχουν μόνο δύο ανεξάρτητοι cross-ratios και οι υπόλοιποι είναι συναρτήσεις αυτών. Με βάση το θεώρημα 3 αυτό συνεπάγεται, ότι

υπάρχουν μόνο δύο ανεξάρτητες συναρτήσεις ανάμεσα στις $\mathcal{J}^{(1)}, \ldots \mathcal{J}^{(5)}$. Η σχέση ανάμεσα στις συναρτήσεις αυτές μπορεί να εξαχθεί παρατηρώντας πως ο cross-ratio μετασχηματίζεται κάτω από τις τέσσερις ειδικές μεταθέσεις $\pi^{(k)},\ k = 2, \ldots, 4$

Θεώρημα 4 *[90] Αν μ και ν είναι οι cross-ratios των εξισώσεων 2.25 και 2.26 αντίστοιχα, τότε οι πέντε $\mathcal{J}$-τιμές είναι:*

$$\begin{aligned} \mathcal{J}^{(1)} &= J[\mu] & \mathcal{J}^{(2)} &= J[\nu] \\ \mathcal{J}^{(3)} &= J\left[\frac{\mu}{\nu}\right] & \mathcal{J}^{(4)} &= J\left[\frac{\nu-\mu}{\nu-1}\right] \\ \mathcal{J}^{(5)} &= J\left[\frac{\mu(\nu-\mu)}{\nu-\mu}\right] \end{aligned} \tag{2.43}$$

Μιας και οι συναρτήσεις $\mathcal{J}^{(k)}$ είναι αναλλοίωτες κάτω από μεταθέσεις των υπολοίπων τεσσάρων σημείων στη συλλογή, μπορούμε να επιλέξουμε περισσότερο συμμετρικές εκφράσεις σαν ορίσματα $\mathcal{J}^{(4)}$ και $\mathcal{J}^{(5)}$:

$$\begin{aligned} \mathcal{J}^{(4)} &= J\left[\frac{\nu-1}{\mu-1}\right] & \mathcal{J}^{(5)} &= J\left[\frac{\mu(\nu-1)}{\nu(\mu-1}\right] \end{aligned} \tag{2.44}$$

Το θεώρημα 3 διαβεβαιώνει ότι οι σχέσεις 2.43 και 2.44 διατηρούν την ισχύ τους όταν τα μ και ν υπολογίζονται χρησιμοποιώντας σαν κοινά σημεία οποιαδήποτε από τα πέντε σημεία. Για κάθε μονοδιάστατη αναλλοίωτη $J[\cdot]$ μπορούμε να κατασκευάσουμε ένα σύνολο από πέντε συναρτήσεις για την συλλογή των πέντε ομοεπίπεδων σημείων.

Οι πέντε $\mathcal{J}^{(k)}$ τιμές μπορούν να συνδιασθούν με πολλούς τρόπους για να οριστούν $\mathbf{P}^2$-αναλλοίωτες για πρακτικούς σκοπούς. Όταν το 5-αδιάστατο διάνυσμα χρησιμοποιείται για ευρετηριασμό [94] ή αντιστοίχιση χαρακτηριστικών [96], το θεώρημα 3, διαβεβαιώνει ότι η αντιστοίχιση των σημείων μεταξύ των συλλογών εξάγεται επίσης. Στην δική μας προσέγγιση χρησιμοποιούμε μια βαθμωτή $\mathbf{P}^2$-αναλλοίωτη [90] για την συλλογή των πέντε ομοεπίπεδων σημείων:

$$\boxed{\mathcal{K}[\mu,\nu] = J[\mu] + J[\nu] + J\left[\frac{\mu}{\nu}\right] + J\left[\frac{\nu-1}{\mu-1}\right] + J\left[\frac{\mu(\nu-1)}{\nu(\mu-1}\right]} \tag{2.45}$$

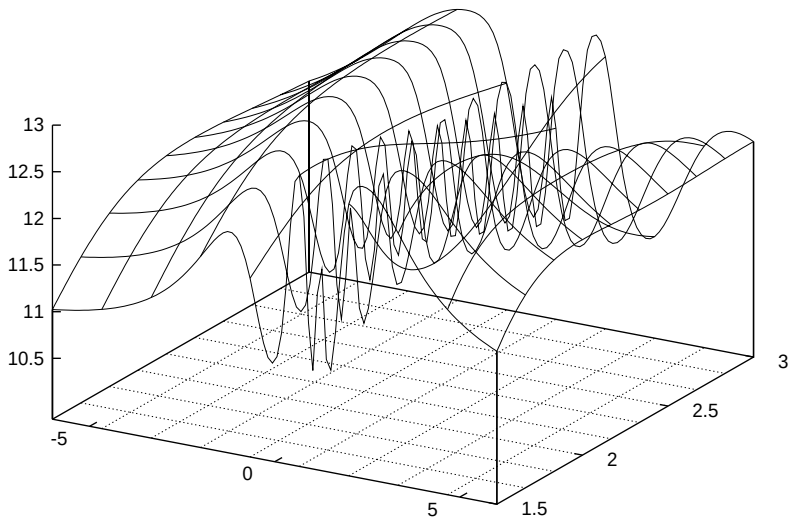
Η συνάρτηση 2.45 περιγράφει μια ολόκληρη οικογένεια από διαφορετικές $\mathbf{P}^2$-αναλλοίωτες, μιας και οποιεσδήποτε από τις συναρτήσεις $J[\cdot], I[\cdot]$, που κατασκευάστηκαν μπορούν να χρησιμοποιηθούν.

Οι αναπαραστάσεις με χρήση αναλλοίωτων είναι πολύ ευαίσθητες στην ύπαρξη θορύβου[12] [96]. Την ευαισθησία αυτή την εντοπίσαμε και εμείς σε πειράματα που

[12]σε λάθη συντεταγμένων που αφορούν την θέση των χαρακτηριστικών

επιτελέσαμε προκειμένου να επιλέξουμε τη συνάρτηση J, με τη βέλτιστη συμπεριφορά. Συγκεκριμένα, υπολογίσαμε την αναλλοίωτη $\mathcal{K}$ σε ποικίλες συλλογές σημείων και καταγράψαμε τη σταθερότητα και τη διατήρηση των τιμών της με κάθε μια από τις προαναφερθείσες συναρτήσεις. Τελικά, καταλήξαμε στο συμπέρασμα, ότι οι φραγμένες συναρτήσεις, και πιο συγκεκριμένα η J_{14} (βλέπε εξίσωση 2.40), συμπεριφέρονται πολύ πιο αποδοτικά αφού δεν παρουσιάζουν ασυνέχειες και λοιπές παθολογικές καταστάσεις. Στο ίδιο συμπέρασμα, καταλήγουν και οι συγγραφείς στο [90], όπου διερευνούν την αριθμητική ευστάθεια των αναλλοίωτων ως προς τον θόρυβο, εργαζόμενοι σε τυχαία παραγόμενες συλλογές σημείων. Επίσης, παρατηρήσαμε ότι η έκφραση 2.45 είναι αρκετά πιο εύρωστη στην παρουσία θορύβου, συγκρινόμενη με το διάνυσμα των πέντε $\mathcal{J}$ συνιστωσών. Αυτό γίνεται άμεσα εμφανές, μιας και αν κάποια από τις εξισώσεις $J[\cdot]$ επηρεαστεί από το θόρυβο, τότε όλο το διάνυσμα επηρεάζεται, ενώ στο άθροισμα η επιρροή του θορύβου διαμοιράζεται και εξομαλύνεται, με αποτέλεσμα να μην είναι έντονη στην εξαγωγή των αποτελεσμάτων.

Η γραφική παράσταση της εξίσωσης 2.45, με βάση τη συνάρτηση J_{14}, απεικονίζεται στο σχήμα 2.10 Όπως φαίνεται, από τη γραφική παράσταση, το πεδίο τιμών της



Σχήμα 2.10: Η εξίσωση $\mathcal{K}$, με βάση τη συνάρτηση J_{14}

αναλλοίωτης $\mathcal{K}$, που χρησιμοποιούμε, είναι το διάστημα $[10, 14]$. Αυτός ο περιορισμός ισχύει λόγω του ότι η συνάρτηση J_{14} είναι φραγμένη στο διάστημα $[2, 2.8]$. Το πρόβλημα που δημιουργείται με το περιορισμένο πεδίο τιμών, είναι η σύγκριση των τιμών των αναλλοίωτων, μέσω της οποίας εξάγουμε συμπεράσματα για το αν μια συγκεκριμένη πεντάδα περιέχει ομοεπίπεδα σημεία. Ωστόσο, καταφέρνουμε να μοντελοποιήσουμε το

άνω όριο της ανεκτής διαφοράς μεταξύ των τιμών δύο αναλλοίωτων, χρησιμοποιώντας ως βάση τη διασπορά της αναλλοίωτης.

2.7 Εφαρμογές της προβολικών αναλλοίωτων στην Μηχανική Όραση

Προκειμένου να αξιολογήσουμε την δύναμη της Προβολικής γεωμετρίας στον τομέα της Μηχανικής Όρασης παραθέτουμε αντιπροσωπευτικές εφαρμογές της.

- *Εύρεση ομοεπίπεδων περιοχών*: Οι Sinclair et al. [97], αναπτύσσουν μια μέθοδο τμηματοποίησης ομοεπίπεδων περιοχών από ένα στερεοσκοπικό ζεύγος εικόνων, θεωρώντας ως γνωστή την αντιστοίχιση των χαρακτηριστικών μεταξύ των εικόνων του ζεύγους. Η αρχική επιλογή ομάδων ομοεπίπεδων σημείων βασίζεται στο γεγονός της διατήρησης[13] των διδιάστατων cross-ratios των εξισώσεων 2.25 και 2.26. Οι γνωστές αντιστοιχίσεις των τεσσάρων από τα πέντε σημεία χρησιμοποιούνται στον ορισμό της προβολικότητας, η οποία χρησιμοποιείται για την πρόβλεψη στην αλλαγή της θέσης για τα υπόλοιπα σημεία, υποθέτοντας ότι κείνται στο ίδιο επίπεδο με τα τέσσερα αρχικά. Μέσω ενός κατωφλιού απόστασης μεταξύ των πραγματικών και των εκτιμώμενων θέσεων καθίσταται δυνατή η εύρεση εκτεταμένων επίπεδων περιοχών. Το μειονέκτημα, είναι η προϋπόθεση γνώσης της αντιστοίχισης σημείων μεταξύ των εικόνων.

- *Κατασκευή συναρτήσεων δεικτοδότησης*: Οι προβολικές αναλλοίωτες έχουν χρησιμοποιηθεί με επιτυχία σε προβλήματα κατασκευής συναρτήσεων δεικτοδότησης σε εφαρμογές όπου υπάρχει μία *εκ των προτέρων* γνώση για το περιβάλλον δράσης[14]. Πολλά χειροποίητα αντικείμενα περιέχουν ομοεπίπεδες ευθείες και κωνοειδή [98, 99], επιτρέποντας άμεση εφαρμογή των πιο πάνω αναλλοίωτων. Πολλές από τις εργασίες που ακολουθούν συνοψίζονται στο συγγραφικό έργο του Rothwell [100, 101]. Αναφέρουμε εποπτικά ορισμένες.

  Οι Rothwell et al. [102] παρουσιάζουν τεχνικές για τον υπολογισμό γεωμετρικών αναλλοίωτων για συμμετρικά ως προς περιστροφή αντικείμενα (rotationally symmetric objects) [103], για αντικείμενα που εμφανίζουν επαναλαμβανόμενη δομή και διμερή συμμετρία (bilateral symmetry) [104]. Οι Mundy et al. [105], επεκτείνουν τις πιο πάνω τεχνικές σε περιπτώσεις, όπου υπάρχει μερική απόκρυψη (partial occlusion) των αντικειμένων, επαυξάνοντας την ισχύ των γεωμετρικών αναλλοίωτων. 0ι προσέγγισεις αυτές, αν και παρουσιάζουν ικανοποιητική απόδοση όσο αναφορά την ταχύτητα και την ευρωστία στην αναγνώριση, εντούτοις εφαρμόζονται ικανοποιητικά μόνο σε συγκεκριμένες κλάσεις τριδιάστατων αντικειμένων.

[13] Ομάδες σημείων για τις οποίες η αναλλοίωτη διατηρείται, χαρακτηρίζονται ως ομοεπίπεδες
[14] Η, όπως είναι πιο γνωστό, στη model-based όραση

Το σύστημά που παρουσιάζεται στην εργασία [97] στηρίζεται σε εξαγωγή κωνοειδών και σύγκριση αναλλοίωτων βασιζόμενες σ' αυτές. Το μειονέκτημα αυτής της μεθόδου είναι αφενός η αποτυχία της μεθόδου στις περιπτώσεις που δεν υπάρχουν κωνοειδή αντικείμενα στην εικόνα και αφετέρου η αδυναμία της στις περιπτώσεις που υπάρχουν σημαντικές επικαλύψεις. Εξέλιξη της ιδέας αυτής αποτελεί το σύστημα LEWIS [106], που χρησιμοποιείται για την αναγνώριση αντικειμένων σε επίπεδα.

Στην εργασία του [107], ο Carlsson, προκειμένου να αναγνωρίσει αντικείμενα, ανεξάρτητα της οπτικής γωνίας λήψης των, υπολογίζει προβολικά αναλλοίωτες ελλείψεις, με βάση τις συντεταγμένες των ακμών του κάθε αντικειμένου. Το πρόβλημα με τη μέθοδο, είναι ότι δεν μπορεί να αντιμετωπίσει σημαντικές επικαλύψεις. Στο [108], η μέθοδος προσπαθεί να καλύψει και περιπτώσεις που δεν μπορούν να εξαχθούν ελλείψεις με βάση την πληροφορία ακμών του κάθε αντικειμένου.

Η πλειονότητα των άνωθι προσεγγίσεων εφαρμόζεται κυρίως σε βάσεις αντικειμένων (π.χ. μηχανικών εργαλείων), όπου η κύρια απαίτηση είναι η ταχεία εύρεση ενός αντικειμένου. Η λειτουργικότητά τους όμως σε γενικής φύσεως περιβάλλοντα, δεν έχει δυστυχώς ερευνηθεί. Επίσης η χρήση αναλλοίωτων βασιζόμενων σε κωνοειδή, εγείρει αρκετά ζητήματα ασυνεχειών και αριθμητικής ευστάθειας στην αξιόπιστη προσαρμογή των σε τμήματα καμπυλών [78].

- *Ανάπτυξη βέλτιστων φίλτρων για σκοπούς αναγνώρισης προτύπων [109, 110]*: Οι τεχνικές αυτές βασίζονται στη Θεωρία αναπαράστασης, στην οποία, όπως αναφέραμε, τα στοιχεία μιας ομάδας αναπαρίστανται μέσω πινάκων. Τα φίλτρα αυτά λειτουργούν ανεξάρτητα από την οπτική γωνία λήψης ενός αντικειμένου, διευκολύνοντας αρκετά την διαδικασία αναγνώρισης. Η κατασκευή των φίλτρων βασίζεται στο γεγονός ότι οι μετασχηματισμοί που υφίσταται ένα αντικείμενο, κατά την διαδικασία απεικόνισης του στο επίπεδο της κάμερας, μοντελοποιούνται σαν την επίδραση μιας ''πράξης'' στα στοιχεία μιας ομάδας.

- *Αντιστοίχιση χαρακτηριστικών*: Σε μια πρόσφατη μεταπτυχιακή εργασία, ο Halkidis [111] κατασκεύασε μια μέθοδο αντιστοίχισης ομοεπίπεδων χαρακτηριστικών ιδωμένα κάτω από δύο διαφορετικές θεάσεις, που έχουν υποστεί προβολικούς μετασχηματισμούς. Η μέθοδός του ανήκει στη γενικότερη κατηγορία τεχνικών μη βαθμονομημένης όρασης και στηρίζεται σε ένα μηχανισμό πρόβλεψης χαρακτηριστικών στη δεύτερη θέαση, καθιστώντας δυνατή την εύρεση υποσυνόλων χαρακτηριστικών στις δύο θεάσεις, για τα οποία υπάρχει ένα μέτρο εμπιστοσύνης για τη μεταξύ των αντιστοιχία.

Κεφάλαιο 3

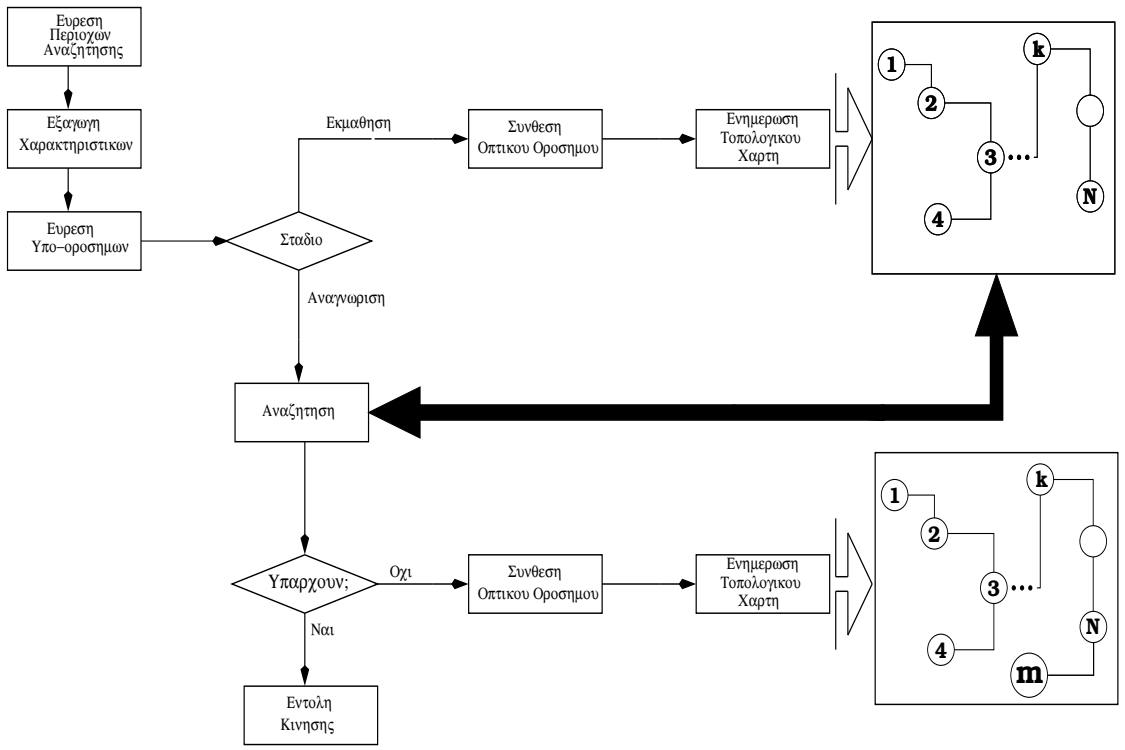
Εκμάθηση και Αναγνώριση Ορόσημων

Η προτεινόμενη μέθοδος, όπως αναφέραμε στην εισαγωγή της εργασίας, αντιμετωπίζει το πρόβλημα της αυτόνομης πλοήγησης ενός μηχανικού συστήματος μέσω της χρήσης οπτικών ορόσημων. Τα οπτικά ορόσημα έχουν μπει πρόσφατα στον τομέα της επίλυσης των προβλημάτων, που σχετίζονται με το γενικότερο πρόβλημα της ικανοποιητικής *αντίληψης του χώρου* και αποτελούν μια αρκετά υποσχόμενη τεχνική. Στο κεφάλαιο αυτό, επικεντρώνουμε την προσοχή μας στην ανατομία της μεθόδου, που αναπτύξαμε, παρέχοντας στον αναγνώστη μια αναλυτική εικόνα της. Επεξηγούμε τον μηχανισμό εξαγωγής των ορόσημων κατά τη διάρκεια μιας αρχικής περιόδου εκμάθησης και την ικανότητα αναγνώρισης των κατά την διάρκεια μιας μεταγενέστερης φάσης αναγνώρισης. Εξετάζουμε τη διαδικασία της τοπολογικής πλοήγησης, που χρησιμοποιεί τις προβολικά και μεταθετικά αναλλοίωτες ποσότητες, που περιγράψαμε στο κεφάλαιο 2, και αναλύουμε τους μηχανισμούς που συνεισφέρουν στην επιτυχή εκμάθηση και αναγνώριση των ορόσημων.

Είναι φανερό ότι οποιαδήποτε μέθοδος πλοήγησης που βασίζεται σε οπτικά ορόσημα, πρέπει να υπόκεινται σε κάποιους περιορισμούς όσο αναφορά την σύσταση του χώρου δράσης [57]. Στην αντίθετη περίπτωση, το σύνολο των ορόσημων που απαιτείται να *μάθει* το σύστημα και να αναγνωρίζει στη συνέχεια, είναι απαγορευτικά μεγάλο και πολύπλοκο. Για τον παραπάνω λόγο και λόγω της χρήσης των προβολικά και μεταθετικά επίπεδων αναλλοίωτων, θεωρούμε ότι για την ικανοποιητική απόδοση της μεθόδου, θα πρέπει να υπάρχουν επίπεδες επιφάνειες στον χώρο δράσης. Ο *περιορισμός* αυτός δεν περιορίζει τη γενικότητα της μεθόδου, αφού η αρχιτεκτονική της πλειοψηφίας των εσωτερικών χώρων απαρτίζεται από επίπεδες επιφάνειες. Ακόμη, όμως και αν θελήσουμε να εφαρμόσουμε τη μέθοδο σε κάποιο *ιδιόμορφο* χώρο, δεν θα υπάρχει σημαντική διαφορά, μιας και οποιοδήποτε μη-επίπεδο αντικείμενο μιας σκηνής, μπορεί να προσεγγιστεί από μια πλειάδα επίπεδων επιφανειών.

Όπως έχουμε αναφέρει, η προσέγγιση μας αποτελείται από δύο λειτουργικές μονάδες που αναλαμβάνουν την *εκμάθηση* και την *αναγνώριση* των ορόσημων. Η μονάδα εκμάθησης ορόσημων, είναι υπεύθυνη για την εκπαίδευση του μηχανικού συστήματος.

Πιο συγκεκριμένα, αναζητά οπτικά ορόσημα σε ακολουθίες εικόνων και τα αποθηκεύει στη *μνήμη* του μηχανικού συστήματος. Η μονάδα αναγνώρισης, αναλαμβάνει την πλοήγηση και την αυτο-οριοθέτηση του μηχανικού συστήματος στο χώρο, κάνοντας χρήση της πληροφορίας, που έχει αποθηκευτεί, στο στάδιο της εκμάθησης. Κατά το στάδιο της εκμάθησης, το μηχανικό σύστημα κατασκευάζει περιοχές αναζήτησης υποψήφιων ορόσημων, με βάση μηχανισμούς *εστίασης της προσοχής* και ανιχνεύει σε αυτές σύνολα χαρακτηριστικών, από τα οποία, επιλέγει τις πιο ικανοποιητικές πεντάδες, που θα αποτελέσουν τα συστατικά των ορόσημων. Τα ορόσημα που εντοπίζονται με αυτό τον τρόπο, αποθηκεύονται στη μνήμη, μαζί με πληροφορίες για τη σχετική τους θέση στο χώρο. Με τον τρόπο αυτό πραγματοποιείται η δημιουργία ενός *τοπολογικού χάρτη* του χώρου, ο οποίος αποτελεί σημαντική πηγή πληροφορίας στο στάδιο της αναγνώρισης, όπου το μηχανικό σύστημα, ακολουθώντας μια παρόμοια συλλογιστική, προσπαθεί να αναγνωρίσει *γνωστά* ορόσημα, τα οποία θα το βοηθήσουν στην αυτο-οριοθέτησή του στο χώρο και στην εκτέλεση πλοηγητικών εντολών, που έχουν συσχετιστεί με κάθε ορόσημο. Για παράδειγμα, η αναγνώριση ενός ορόσημου, μπορεί να αποτελεί ένδειξη για δεξιά στροφή του μηχανικού συστήματος. Η μέθοδος που αναπτύσσουμε, εμφανίζεται, μέσω ενός διαγράμματος ροής, στο σχήμα 3.1, όπου απεικονίζεται η διαδρομή που ακολουθούμε για την επιτυχή πλοήγηση του μηχανικού συστήματος. Τα δύο στάδια, όπως φαίνεται και από το σχήμα, δεν είναι ανεξάρτητα,



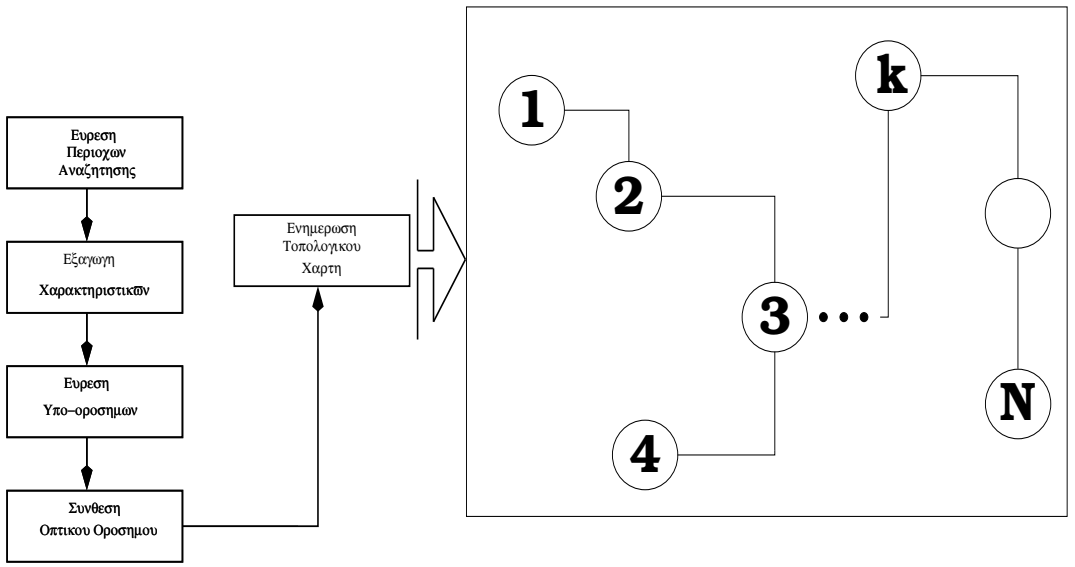
Σχήμα 3.1: Διάγραμμα ροής της μεθόδου

αλλά συνεργάζονται αποδοτικά, αφού αν κατά τη διάρκεια της πλοήγησης συναντηθεί

ένα άγνωστο, με τα μέχρι τώρα, ορόσημο, ενεργοποιείται η διεργασία της εκμάθησης, η οποία αναλαμβάνει να ενημερώσει το σύστημα με τη νέα πληροφορία. Με τον τρόπο αυτό, το σύστημα μπορεί να προσαρμόζεται σε καινούρια δεδομένα και αλλαγές του περιβάλλοντος. Στη συνέχεια μελετούμε το κάθε τμήμα της μεθόδου, αναλύοντας λεπτομερώς τα χαρακτηριστικά του.

3.1 Στάδιο εκμάθησης οπτικών ορόσημων

Στην ενότητα αυτή περιγράφουμε το στάδιο εκμάθησης του μηχανικού συστήματος, κατά τη διάρκεια του οποίου εξάγονται τα πρότυπα των οπτικών ορόσημων και αποθηκεύονται σε μια βάση μοντέλων, η οποία αποτελεί μελλοντικό σημείο αναφοράς. Αρχικά, ορίζουμε τη μορφή και τις ιδιότητες ενός οπτικού ορόσημου. Κατόπιν, αναλύουμε το μηχανισμό εστίασης της προσοχής, μέσω του οποίου κατασκευάζονται οι περιοχές αναζήτησης και περιγράφουμε τον ανιχνευτή χαρακτηριστικών, ο οποίος παρέχει τα συστατικά σημεία των ορόσημων. Στη συνέχεια, περιγράφουμε λεπτομερειακά την ανατομία της μεθόδου εξαγωγής των ορόσημων, μιας και, ουσιαστικά, θα χρησιμοποιηθεί αυτούσια αργότερα, κατά την φάση της πλοήγησης, για την εξαγωγή των παρατηρούμενων ορόσημων, που θα αναζητηθούν στη βάση μοντέλων. Ένα γενικό διάγραμμα της μεθόδου εκμάθησης απεικονίζεται στο σχήμα 3.2.

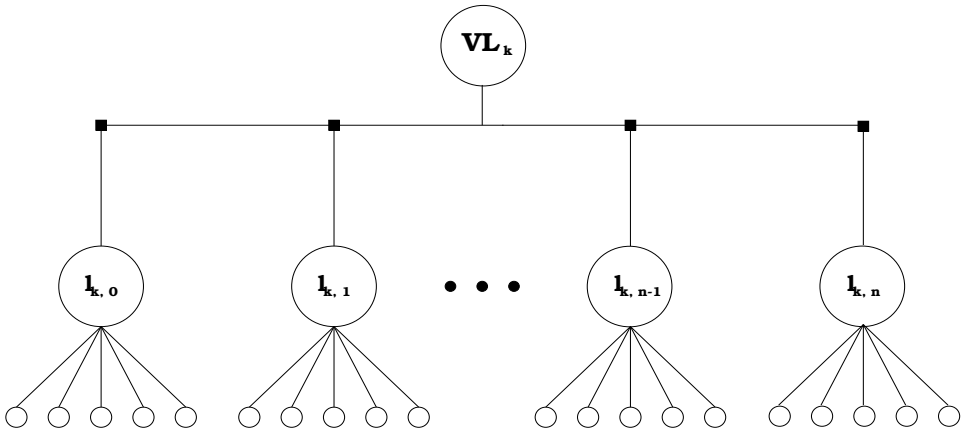


Σχήμα 3.2: Διάγραμμα ροής του σταδίου εκμάθησης

3.1.1 Ορισμός και αναπαράσταση ενός οπτικού ορόσημου

Στα πλαίσια αυτής της εργασίας ορίζουμε τα *οπτικά ορόσημα*, VL_i, να είναι σύνολα που περιέχουν *υπό-ορόσημα* $l_{i,j}$. Τα υπό-ορόσημα είναι πεντάδες ομοεπίπεδων σημείων

$P_{i,j,k}$, $k = 1 \ldots 5$, που δεν είναι ανά τρία συνευθειακά. Στην εικόνα 3.3 απεικονίζεται ένα ορόσημο με τη μορφή μιας ιεραρχίας επιπέδων λεπτομέρειας. Στο πρώτο επίπεδο, εμφανίζεται το ορόσημο με το μορφή ενός αφηρημένου κόμβου, στο αμέσως επόμενο επίπεδο λεπτομέρειας εμφανίζονται τα υπο-ορόσημα, και στο τρίτο επίπεδο τα σημεία που αποτελούν τα υπο-ορόσημα. Τα οπτικά ορόσημα, όπως τονίσαμε, δεν είναι προ-



Σχήμα 3.3: Ανατομία ενός Οπτικού Ορόσημου

επιλεγμένα ή προκαθορισμένα. Η αναπαράσταση αυτή υιοθετήθηκε γιατί είναι γενική, δεν περιορίζεται από τον αριθμό των πιθανών σημείων αναφοράς μιας συγκεκριμένης σκηνής και είναι ευέλικτη στις πιθανές αλλαγές του περιβάλλοντος. Επίσης, η χρήση των συντεταγμένων των πέντε σημείων χρησιμοποιείται αρκετά για την ποιοτική περιγραφή προτύπων στην εικόνα, και διευκολύνει την διαδικασία αναγνώρισης. Αν χρησιμοποιούσαμε κομμάτια της εικόνας σαν ορόσημα του χώρου, όπως στο [72], τότε θα έπρεπε να εμπλακούμε σε τεχνικές *ταύτισης φορμών (template matching)*, οι οποίες είναι αρκετά πιο χρονοβόρες και λιγότερο αξιόπιστες [112].

Έχοντας ορίσει την σύσταση των ορόσημων, το επόμενο θέμα είναι να ερευνήσουμε πως παράγονται τα σημεία που συνθέτουν τις πεντάδες. Τα σημεία αυτά δεν είναι τυχαία επιλεγμένα, μιας και τότε τα συμπεράσματα που θα παράγονταν δεν θα είχαν γενικό ενδιαφέρον, αλλά θα εξαρτώνταν από τις ιδιότητες της γεννήτριας τυχαίων χαρακτηριστικών. Άλλωστε, ο ίδιος ο άνθρωπος, προκειμένου να αναγνωρίσει ένα χώρο και να προσανατολιστεί, δεν στηρίζεται σε συμπεράσματα τυχαιότητας, αλλά μάλλον στην αναγνώριση ιδιαίτερων χαρακτηριστικών του χώρου δράσης του. Τα ιδιαίτερα χαρακτηριστικά πρέπει να είναι ποιοτικώς διαχωρίσιμα από τα υπόλοιπα σημεία της σκηνής. Η εξαγωγή των χαρακτηριστικών πραγματοποιείται μέσω ενός κλασικού ανιχνευτή χαρακτηριστικών, του οποίου η περιγραφή ακολουθεί.

3.1.2 Ο ανιχνευτής χαρακτηριστικών

Ο ανιχνευτής χαρακτηριστικών είναι ένα από τα βασικά δομικά στοιχεία της προσέγγισής μας, μιας και παράγει τα χαρακτηριστικά που αποτελούν τους *δομικούς λίθους*

των ορόσημων. Τα χαρακτηριστικά που εντοπίζει ο ανιχνευτής θα πρέπει να είναι:

1. *σημαντικά* ως προς την ποιότητα τους
2. *συνεπή* ως προς την εμφάνιση τους σε διαδοχικά πλαίσια
3. *εύρωστα* ως προς την παρουσία θορύβου και τυχόν διαταραχών της κίνησης της κάμερας.
4. *ταχέως* ανιχνεύσιμα, ώστε η εξαγωγή τους να μην επιβραδύνει την διαδικασία πλοήγησης.

Στο αρχικό στάδιο της υλοποίησης, χρησιμοποιήσαμε τον ανιχνευτή κορυφών των Smith et al. [113], ο οποίος, σύμφωνα με τους συγγραφείς, ικανοποιούσε τις άνωθι προδιαγραφές. Ωστόσο, στην εξέλιξη της εργασίας, και μετά από πειράματα σε ακολουθίες εικόνων με διαδρόμους κτηρίων, διαπιστώσαμε ότι, αν και ήταν αρκετά ταχύς στον εντοπισμό των χαρακτηριστικών, εντούτοις δεν τηρούσε την απαίτηση της *συνέπειας*, αφού ορισμένα χαρακτηριστικά που ήταν εμφανή σε ένα πλαίσιο εξαφανίζονταν στο αμέσως διαδοχικό. Για το λόγο αυτό, χρησιμοποιήσαμε, ως εναλλακτική λύση, τον ανιχνευτή χαρακτηριστικών των Tomasi et al.[1] [114], ο οποίος, όπως αποδείχθηκε στα πειράματα που επιτελέσαμε, τηρούσε, σε ικανοποιητικό βαθμό, όλες τις επιθυμητές απαιτήσεις. Σύμφωνα με τους συγγραφείς, η ποιότητα ενός χαρακτηριστικού καθορίζεται από την ικανότητά του να εντοπίζεται αξιόπιστα στις ακολουθίες εικόνων της εισόδου. Με τον τρόπο αυτό, η επιλογή των χαρακτηριστικών είναι βέλτιστη εκ κατασκευής. Στη συνέχεια, παραθέτουμε τα βασικά χαρακτηριστικά του ανιχνευτή καθώς και τις αρχές στις οποίες στηρίζεται.

3.1.2.1 Βασικές αρχές του ανιχνευτή χαρακτηριστικών

Καθώς η κάμερα κινείται, οι τιμές των εντάσεων της εικόνας παρουσιάζουν διακυμάνσεις κατά ένα πολύπλοκο τρόπο. Γενικά μιλώντας, μια ακολουθία εικόνων μπορεί να αναπαρασταθεί μέσω μιας συνάρτησης $\mathrm{I}(x, y, t)$ δύο χωρικών μεταβλητών x, y και μιας χρονικής t, οι οποίες είναι διακριτές και κατάλληλα φραγμένες. Ωστόσο, εικόνες που λαμβάνονται σε διαδοχικές χρονικές στιγμές συνήθως σχετίζονται ισχυρά μεταξύ τους, μιας και αναφέρονται στην ίδια σκηνή ιδωμένη από αμυδρά διαφορετικές οπτικές γωνίες. Αν η συχνότητα της χρονικής δειγματοληψίας είναι επαρκώς υψηλή[2], μπορούμε να υποθέσουμε ότι οι μικρές περιοχές της εικόνας μετακινούνται αλλά οι εντάσεις των παραμένουν σταθερές. Αυτό σημαίνει ότι η συνάρτηση $\mathrm{I}(x, y, t)$ δεν είναι αυθαίρετη, αλλά ικανοποιεί την ακόλουθη σχέση:

$$\mathrm{I}(x, y, t + \tau) \;=\; \mathrm{I}\left(x - \xi(x, y, \tau), y - \eta(x, y, t, \tau), t\right) \tag{3.1}$$

[1] Έπειτα από μια εύστοχη υπόδειξη του συναδέλφου Σπύρου Χαλκίδη

[2] που σημαίνει ότι η κίνηση της κάμερας είναι μικρή

Έτσι, μια επόμενη εικόνα που έχει ληφθεί την χρονική στιγμή $t + \tau$ μπορεί να παρατηρηθεί μετακινώντας κάθε σημείο της εικόνας, που έχει ληφθεί την χρονική στιγμή t, κατά ένα κατάλληλο ποσό. Το ποσό αυτό της κίνησης $\delta = (\xi, \eta)$ ονομάζεται *μετατόπιση*[3] του σημείου στη θέση $\mathbf{x} = (x, y)$ μεταξύ των δύο χρονικών στιγμών t και $t + \tau$.

Ακόμη και σε μια στατική σκηνή, κάτω από σταθερές συνθήκες φωτισμού, η εξίσωση 3.1 συχνά παραβιάζεται. Ένα άλλο πρόβλημα είναι, ότι το διάνυσμα μετατόπισης δ είναι συνάρτηση της θέσης $\mathbf{x}$, πάνω στην εικόνα, και οι μεταβολές του είναι συχνά παρατηρήσιμες ακόμη και εντός των μικρών παραθύρων που χρησιμοποιούνται στην διαδικασία *συνεχούς εντοπισμού (tracking)*. Οπότε δεν έχει ουσία να κάνουμε λόγο για ''την'' μετατόπιση ενός παραθύρου χαρακτηριστικών, μιας και υπάρχουν διαφορετικές μετατοπίσεις εντός αυτού. Δυστυχώς, δεν μπορούμε να συρικνώσουμε τις διαστάσεις των παραθύρων στις αντίστοιχες διαστάσεις των εικονοστοιχείων για να αποφύγουμε αυτή την δυσκολία. Στην πραγματικότητα, η τιμή ενός εικονοστοιχείου μπορεί να μεταβληθεί λόγω επίδρασης θορύβου με αποτέλεσμα να μπερδευτεί με τις τιμές των γειτονικών εικονοστοιχείων, καθιστώντας δύσκολη, αν όχι αδύνατη, την απάντηση στο αν μετακινήθηκε στο επόμενο πλαίσιο.

Μια καλύτερη προσέγγιση είναι να εμπλουτίσουμε την περιγραφή της κίνησης εντός του παραθύρου, δηλαδή, να ορίσουμε ένα σύνολο από πιθανές συναρτήσεις μετατοπίσεων $\delta(\mathbf{x})$, για δοθέντα t και τ, το οποίο δεν περιλαμβάνει μόνο σταθερές συναρτήσεις του $\mathbf{x}$. Η υπόθεση της ταχείας χρονικής δειγματοληψίας (fast-sampling hypothesis), μας επιτρέπει να προσεγγίσουμε την κίνηση με ένα *μεταφορικό μοντέλο*, το οποίο αποτελεί ένα καλό συμβιβασμό μεταξύ απλότητας και ευελιξίας και περιγράφεται από τη συνάρτηση:

$$\delta(\mathbf{x}) \quad = \quad \mathbf{x} + \mathbf{d} \tag{3.2}$$

όπου το $\mathbf{d}$ είναι το διάνυσμα μετατόπισης (displacement vector), που εκφράζει τη μεταφορά του κέντρου του παραθύρου χαρακτηριστικών. Οι συντεταγμένες των εικονοστοιχείων $\mathbf{x}$ της εικόνας εκφράζονται σαν συνάρτηση του κέντρου του παραθύρου. Τότε, ένα σημείο $\mathbf{x}$ της πρώτης εικόνας I μετακινείται στο σημείο $\mathbf{x} + \mathbf{d}$ στην δεύτερη εικόνα J. Έτσι το μεταφορικό μοντέλο κίνησης συνοψίζεται στην ακόλουθη εξίσωση, που συσχετίζει τις εντάσεις της εικόνας:

$$J(\mathbf{x} + \mathbf{d}) \quad = \quad I(\mathbf{x}) \tag{3.3}$$

Δοθέντος δύο εικόνων I και J και ενός παραθύρου στην εικόνα I, ο συνεχής εντοπισμός σημαίνει να υπολογίσουμε τις παραμέτρους του διανύσματος μετατόπισης $\mathbf{d}$. Η ποιότητα της εκτίμησης εξαρτάται από το μέγεθος του παραθύρου χαρακτηριστικών, την ποιότητα της υφής της εικόνας εντός αυτού, και το ποσό της κίνησης της κάμερας

[3]Το δ είναι γνωστό και σαν *πεδίο κίνησης*

μεταξύ των πλαισίων. Όταν το παράθυρο είναι μικρό, εξαιτίας των διακυμάνσεων της κίνησης εντός αυτού, οι μετρήσεις είναι μικρότερες και άρα λιγότερο αξιόπιστες. Ωστόσο, τα μικρά παράθυρα είναι γενικά προτιμώτερα για σκοπούς συνεχούς εντοπισμού, μιας και είναι πιο πιθανό να περιέχουν χαρακτηριστικά σε παρόμοια βάθη. Επειδή το μοντέλο κίνησης της εικόνας δεν θεωρείται ακριβές, και επιπλέον λόγω του θορύβου, οι εξίσωσεις 3.1 και 3.3 δεν ικανοποιούνται επακριβώς. Το πρόβλημα, σ' αυτή την περίπτωση, ανάγεται στον υπολογισμό της μετατόπισης $\mathbf{d}$ που ελαχιστοποιεί το κριτήριο της *ανομοιότητας (dissimilarity)*:

$$\epsilon \;=\; \iint_{\mathcal{W}} [J(\mathbf{x}+\mathbf{d}) - I(\mathbf{x})]^2 w(\mathbf{x})\; d\mathbf{x} \tag{3.4}$$

όπου το $\mathcal{W}$ είναι το δοθέν παράθυρο χαρακτηριστικών και η $w(\mathbf{x})$ είναι μια συνάρτηση βάρους. Στην απλούστερη περίπτωση, $w(\mathbf{x}) = 1$. Εναλλακτικά, το w μπορεί να είναι μια Γκαουσιανή συνάρτηση, αν θέλουμε να δώσουμε έμφαση στην κεντρική περιοχή του παραθύρου.

Για να ελαχιστοποιήσουμε το υπόλοιπο 3.4, το διαφοροποιούμε ως προς τις άγνωστες παραμέτρους του διάνυσματος της μετατόπισης $\mathbf{d}$, θέτοντας το αποτέλεσμα ίσο με το μηδέν. Η διαδικασία αυτή παράγει την κάτωθι εξίσωση:

$$\frac{1}{2}\frac{\partial \epsilon}{\partial \mathbf{d}} \;=\; \iint_{\mathcal{W}} [J(\mathbf{x}+\mathbf{d}) - I(\mathbf{d})]\, \mathbf{g} w(\mathbf{x})\; d\mathbf{x} \;=\; 0 \tag{3.5}$$

όπου

$$\mathbf{g} \;=\; \left(\frac{\partial J}{\partial x}, \frac{\partial J}{\partial y}\right)^T$$

είναι το χρονικό ανάδελτα των εντάσεων της εικόνας. Χάρη στην υπόθεση της ταχείας δειγματοληψίας ο όρος $J(\mathbf{x}+\mathbf{d})$, μπορεί να προσεγγισθεί από την ανάπτυξη του σε σειρά Taylor, αποκοπτόμενη στην γραμμική της μορφή:

$$J(\mathbf{x}+\mathbf{d}) \;=\; J(\mathbf{x}) + \mathbf{g}^T(\mathbf{d}) \tag{3.6}$$

η οποία, συνδιαζόμενη με την εξίσωση 3.5, μας οδηγεί στην παρακάτω εξίσωση:

$$\iint_{\mathcal{W}} \mathbf{g}\left(\mathbf{g}^T\mathbf{d}\right) w\, d\mathbf{x} \;=\; \iint_{\mathcal{W}} [I(\mathbf{x}) - J(\mathbf{x})]\, \mathbf{g} w\, d\mathbf{x} \tag{3.7}$$

Για να κάνουμε αυτό τον υπολογισμό πιο ακριβή, εκφράζουμε την εξίσωση 3.7, σε πιο συμπαγή μορφή απλοποιώντας, ως κοινό παράγοντα, τον άγνωστο $\mathbf{d}$. Αυτό οδηγεί [115, 116] στο ακόλουθο γραμμικό σύστημα:

$$G\mathbf{d} \;=\; \mathbf{e} \tag{3.8}$$

όπου

$$G \;=\; \iint_{\mathcal{W}} \begin{bmatrix} g_x^2 & g_x g_y \\ g_x g_y & g_y^2 \end{bmatrix} w\, d\mathbf{x}$$

είναι ο πίνακας που μπορεί να υπολογιστεί από μια εικόνα και το

$$\mathbf{d} \;=\; \begin{bmatrix} d_x \\ d_y \end{bmatrix}$$

είναι το διάνυσμα μετατόπισης, ενώ το

$$\mathbf{e} \;=\; \iint_{\mathcal{W}} \left[I\left(\mathbf{x}\right) - J\left(\mathbf{x}\right) \right] \begin{bmatrix} g_x \\ g_y \end{bmatrix} w \, d\mathbf{x}$$

είναι το διάνυσμα λάθους που εξαρτάται από την διαφορά ανάμεσα στις δύο εικόνες.

Η επίλυση του συστήματος 3.8 οδηγεί στην λύση $\hat{\mathbf{d}}$, η οποία έχει την εξής μορφή:

$$\hat{\mathbf{d}} \;=\; G^{-1}\,\mathbf{e} \tag{3.9}$$

και χρησιμοποιείται για την πρόβλεψη ενός νέου πλαισίου. Η διαδικασία επαναλαμβάνεται σύμφωνα με την αριθμητική μέθοδο Newton-Raphson, μέχρι την σύγκλιση της εκτίμησης των τιμών του διανύσματος μετατόπισης.

Μπορούμε να διατηρούμε σε συνεχή εντοπισμό αξιόπιστα ένα χαρακτηριστικό, σε διαδοχικά πλαίσια, αν μια αριθμητικά ευσταθής λύση μπορεί να υπολογιστεί για το σύστημα 3.8. Για να επιτευχθεί αυτό, απαιτείται η *καλή κατάσταση* του 2×2 συμμετρικού πίνακα G του συστήματος καθώς και η απαίτηση των τιμών των στοιχείων του να είναι πάνω από το επίπεδο θορύβου της εικόνας. Πρακτικά, μιας και η μεγαλύτερη ιδιοτιμή φράσσεται από τη μέγιστη επιτρεπόμενη τιμή του εικονοστοιχείου, οι δύο απαιτήσεις ισοδυναμούν με το να είναι η ελάχιστη ιδιοτιμή επαρκώς υψηλή. Αν οι δύο ιδιοτιμές του G είναι οι λ_1 και λ_2, δεχόμαστε ένα χαρακτηριστικό αν:

$$\min\left(\lambda_1, \lambda_2\right) \;>\; \lambda \tag{3.10}$$

όπου το λ είναι ένα προκαθορισμένο κατώφλι.

3.1.3 Εύρεση περιοχών αναζήτησης ορόσημων

Το επόμενο θέμα που μελετούμε, είναι η εύρεση των περιοχών της εικόνας στις οποίες θα αναζητηθούν τα υπο-ορόσημα, και κατ' επέκταση τα ορόσημα. Το σίγουρο είναι ότι όλα τα μέρη μιας σκηνής, δεν περιέχουν σημαντική[4] πληροφορία. Ο άνθρωπος, προκειμένου να προσανατολιστεί, εστιάζει την προσοχή του στις περιοχές της σκηνής που έχουν κάποια σημασία γι' αυτόν.

Ακολουθώντας τη συμπεριφορά αυτή και προσπαθώντας να διατηρήσουμε την ευελιξία του συστήματος και την ικανότητα κίνησης του μέσα σε άγνωστα περιβάλλοντα, χρησιμοποιούμε ένα μηχανισμό *εστίασης της προσοχής (focus of attention)* σε *σημαντικά* σημεία της εικόνας. Με τον τρόπο αυτό, ακολουθούμε τη σύγχρονη οδό της

[4] ότι και αν σημαίνει αυτό κατά περίπτωση

ενεργούς και τελεολογικής όρασης [52, 53], σύμφωνα με την οποία οι διαδικασίες όρασης ενός μηχανικού συστήματος είναι προσανατολισμένες στην επίτευξη συγκεκριμένων στόχων και η αντίληψη του χώρου, ισοδυναμεί με τον ποιοτικό προσδιορισμό της σημασιολογικής του συνιστώσας. Στα πλαίσια της θεωρίας αυτής, αποφεύγουμε την προσπάθεια επακριβούς αναπαράστασης της γεωμετρικής δομής του περιβάλλοντος, παρακάμπτοντας τα προβλήματα που αυτή συνεπάγεται. Έτσι, το αυτόνομο σύστημα δε διαθέτει εξ' αρχής πληροφορίες για τη δομή ή τη διάταξη του κόσμου, ή για το είδος των αντικειμένων που υπάρχουν σ' αυτόν. Στην προτεινόμενη μέθοδο, οι αναγκαίες αυτές πληροφορίες συλλέγονται από το αυτόνομο σύστημα χρησιμοποιώντας οπτικούς αισθητήρες.

Ο μηχανισμός εστίασης της προσοχής υλοποιείται μέσω του *χάρτη προεξοχής (saliency map)* [55], όπου με χρήση διάφορων κριτηρίων, αποδίδονται υψηλές τιμές στα *προέχοντα* σημεία της εικόνας και χαμηλές τιμές στα σημεία των ομαλών (smoothed) περιοχών. Τα προέχοντα σημεία αποτελούν αφετηρίες ανάπτυξης περιοχών αναζήτησης υπο-ορόσημων. Μιας και οι περιοχές αυτές θεωρούνται σημαντικές, τα υπο-ορόσημα που θα αναζητηθούν εκεί, θα αποτελούν διακριτικά πρότυπα της σκηνής, διευκολύνοντας την μετέπειτα εξαγωγή και αναγνώρισή τους. Στη συνέχεια, μελετούμε την ανατομία του χάρτη προεξοχής, προκειμένου να κατανοήσουμε τον τρόπο με τον οποίο αποδίδονται ''σημασιολογικά βάρη'' στα σημεία της εικόνας.

3.1.3.1 Ανατομία του χάρτη προεξοχής[5]

Ο χάρτης αυτός υπολογίζεται σε κάθε σημείο της εικόνας, ως ένας γραμμικός συνδιασμός άθροισμα μιας οικογένειας συναρτήσεων $\mathcal{F}_i$, που εκφράζουν με ποσοτικό τρόπο την *ανομοιογένεια* της εικόνας στην περιοχή του σημείου. Έτσι, η τιμή του χάρτη σε κάποιο σημείο $\mathbf{p}$ είναι

$$\mathcal{SM}(\mathbf{p}) \;=\; \sum_i w_i \mathcal{F}_i(\mathbf{p}) \tag{3.11}$$

όπου w_i το βάρος της συνάρτησης $\mathcal{F}_i$ στο τελικό άθροισμα. Οι συναρτήσεις, που χρησιμοποιούνται στην κατασκευή του χάρτη προεξοχής της εικόνας είναι οι εξής:

1. Μία συνάρτηση *συσχέτισης*, R, σε ένα παράθυρο N της εικόνας με τύπο:

$$R_N \;=\; \frac{\mu_{mn} - \mu_m\,\mu_n}{\sigma_n\,\sigma_m} \tag{3.12}$$

όπου μ_n συμβολίζει τη μέση τιμή στην περιοχή N, μ_{nn} η μέση τιμή του κατά εικονοστοιχείου γινομένου των τιμών των εικονοστοιχείων σε δύο θέσεις του παραθύρου, M, N και $\sigma_n \;=\; \mu_{nn} - \mu_n\,\mu_n$ είναι η τυπική απόκλιση του N. Είναι φανερό ότι σε ομαλές περιοχές η R λαμβάνει υψηλές τιμές, ενώ σε μη-ομαλές, λαμβάνει χαμηλές τιμές. Για το λόγο αυτό, η συνάρτηση χρησιμοποιείται ανεστραμμένη στον υπολογισμό του χάρτη προεξοχής.

[5]Το τμήμα αυτό αποτελεί μέρος της πτυχιακής εργασίας του Θοδωρή Γκαραβέλου

2. Μια δεύτερη συνάρτηση *συσχέτισης*, C, στο παράθυρο N, που δίνεται από τον τύπο:

$$C_{\mathrm{N}} = \sum_i \sum_j \rho(i,j), \quad (i,j) \in N \tag{3.13}$$

με

$$\rho(x,y) = \frac{1}{|S|} \sum_i \sum_j \left[I(x,y) - I(x+i, y+j) \right]^2, \quad (i,j) \in \mathrm{N} \tag{3.14}$$

όπου S ένα παράθυρο όμοιων διαστάσεων με κέντρο το (x,y) και $|S|$ ο αριθμός των εικονοστοιχείων σε αυτό. Η τιμή της C_N είναι υψηλή σε περιοχές που διαφέρουν μεταξύ τους.

3. Η *εντροπία* H σε ένα παράθυρο N, υπολογιζόμενη ως:

$$H_{\mathrm{N}} = -\sum_i P(h_i) \log_2 P(h_i) \tag{3.15}$$

όπου $P(h_i)$ συμβολίζει την πιθανότητα εμφάνισης της τιμής h_i εντός του N. Η εντροπία στην προσέγγισή μας αποτελεί ένα μετρικό εκτιμητή της *αταξίας* ενός συνόλου σημείων και λειτουργεί σαν μέτρο της ανομοιότητας, αφού λαμβάνει υψηλές τιμές σε περιοχές με διαφορετικές φωτεινότητες.

4. Η *τυπική απόκλιση* του ιστογράμματος φωτεινοτήτων στο παράθυρο N. Αν είναι μεγαλύτερη από κάποιο κατώφλι υποδεικνύει αυξημένη ανομοιογένεια στο παράθυρο.

5. Η *τυπική απόκλιση* των *κατά σημείων διαφορών* κατά μήκος των διαδοχικών στο χρόνο πλαισίων. Αυτό που υπολογίζεται είναι αρχικά η απόλυτη διαφορά μεταξύ των φωτεινοτήτων των εικόνων στο παράθυρο N:

$$f_i = | f_{1i} - f_{2i} | \tag{3.16}$$

Στη συνέχεια, στο παράθυρο N της εικόνας υπολογίζεται η τυπική απόκλιση των διαφορών αυτών:

$$D = \sum_i \frac{(f_i - \mu)^2}{|N|} \tag{3.17}$$

όπου $|N|$ το πλήθος των σημείων του παραθύρου και μ ο μέσος όρος των διαφορών φωτεινοτήτων f_i. Όσο πιο μεγάλη η τυπική απόκλιση των φωτεινοτήτων, τόσο μεγαλύτερη η διαφορά των διαδοχικών εικόνων και κατ' επέκταση η ανομοιογένεια του παραθύρου. Το μέτρο αυτό εγγυάται τη διακριτικότητα του παραθύρου με βάση την πληροφορία της τρέχουσας εικόνας.

Όπως γίνεται κατανοητό, ο χάρτης προεξοχής δεν είναι παρά ένας διδιάστατος *τελεστής ενδιαφέροντος*. Οι τελεστές αυτοί, συνήθως, βασίζονται σε μια προκαθορισμένη και αυθαίρετη αντίληψη του τι συνιστά ένα σημαντικό σημείο [114]. Για το λόγο αυτό, χρησιμοποιούμε την πληροφορία του ποιοτικά. Οι Trahanias et al. [72], χρησιμοποιούσαν την πληροφορία του χάρτη, αφού είχε προηγηθεί μια ποιοτική τμηματοποίηση του χώρου δράσης. Ένας σημαντικός λόγος γι' αυτό, είναι ότι ο χάρτης προεξοχής στηρίζεται σε καθαρά μετρικά κριτήρια των εντάσεων της εικόνας, με αποτέλεσμα οι υψηλές τιμές του να αποδίδονται είτε στο ταβάνι, λόγω των φώτων ή σε μέρη όπου υπάρχουν αντανακλάσεις (π.χ. πάτωμα). Η ποιοτική τμηματοποίηση, ουσιαστικά, εξαιρούσε τις περιπτώσεις αυτές. Ωστόσο, η προσέγγισή τους στηρίζονταν αρκετά στην παραδοχή της παραλληλεπίπεδης δομής του χώρου δράσης, προκειμένου να εξάγει τα βασικά επίπεδα της εικόνας. Στη δική μας προσέγγιση, η χρήση της ποιοτικής τμηματοποίησης, θα περιόριζε την εφαρμογή της μεθόδου σε περιβάλλοντα, που έχουν μία εκ των προτέρων γνωστή δομή. Το γεγονός αυτό, επιβαρύνει το έργο της αυτόνομης πλοήγησης με ένα επιπλέον μηχανισμό, ο οποίος θα πρέπει να καθορίζει το σχήμα της δομής της κάθε σκηνής. Προκειμένου, λοιπόν, να διατηρήσουμε την γενικότητα της μεθόδου, δεν κάνουμε κάποια παραδοχή για το χώρο εργασίας, αλλά ούτε και για τις πιθανές θέσεις, που μπορεί να υπάρχει ένα ορόσημο. Η ευαισθησία του χάρτη προεξοχής αντιμετωπίζεται, συνδιάζοντας την πληροφορία του με την πληροφορία του ανιχνευτή χαρακτηριστικών. Πιο συγκεκριμένα, για την ανίχνευση ''προέχοντων'' σημείων, εξετάζουμε τις τιμές του χάρτη προεξοχής μόνο στα σημεία, που υπάρχουν τα χαρακτηριστικά της εικόνας. Με τον τρόπο αυτό, προσδίδουμε ένα μέτρο βεβαιότητας στην επιλογή των σημαντικών σημείων. Τελικά, από το σύνολο των χαρακτηριστικών, επιλέγονται αυτά, στα οποία ο χάρτης προεξοχής αποδίδει τιμές, που ανήκουν σε ένα διάστημα τιμών. Έτσι αποφεύγουμε αφενός το (άσκοπο) κόστος της τμηματοποίησης και αφετέρου εστιάζουμε την προσοχή μας στα σημεία που πράγματι κρύβουν το ενδιαφέρον. Τα όρια του διαστήματος έχουν προκύψει έπειτα από πειραματικές μετρήσεις σε ακολουθίες εικόνων με διαδρόμους.

Έχοντας λάβει τα σημεία που ο χάρτης προεξοχής παράγει εκτελούμε τα εξής βήματα:

1. Ελέγχουμε για σχέσεις γειτνίασης μεταξύ των σημείων. Αυτό οδηγεί σε μείωση του υπολογιστικού κόστους, διότι η δημιουργία μιας περιοχής αναζήτησης με αφετηρία ένα σημείο P θα εμπερικλύει, με αυξημένη πιθανότητα, όμοια χαρακτηριστικά αντικείμενα με την περιοχή, που δημιουργείται με αφετηρία ένα γειτονικό σημείο του P.

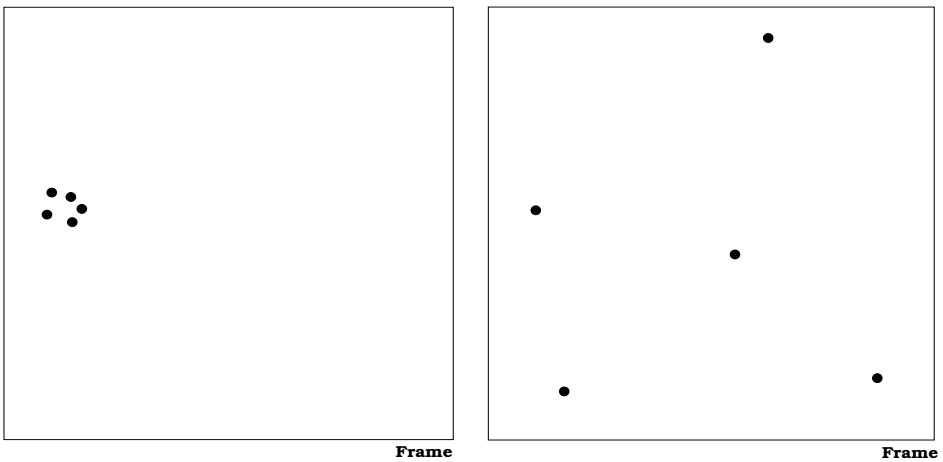
2. Κατόπιν, για κάθε ένα από τα εναπομείναντα σημεία δημιουργούμε περιοχές διαστάσεων $W_a \times W_b$, οι οποίες αποτελούν τους χώρους αναζήτησης των ορόσημων.

3.1.4 Περιορισμοί των υπο-ορόσημων

Στην αρχή του κεφαλαίου δώσαμε ένα γενικό ορισμό των οπτικών ορόσημων, που ακολουθούμε στη δική μας προσέγγιση. Παρατηρούμε ότι ο δομικός λίθος των ορόσημων είναι τα πέντε σημεία που συνθέτουν το κάθε υπο-ορόσημο. Τα σημεία αυτά υπόκεινται σε κάποιους περιορισμούς, οι οποίοι καθιστούν τον υπολογισμό των $\mathbf{P}^2$-αναλλοίωτων αριθμητικά ευσταθή και ελέγχουν το κριτήριο της διακριτικής ικανότητας ενός υπο-ορόσημου. Στην υπο-ενότητα αυτή, περιγράφουμε τους περιορισμούς αυτούς καθώς και τους μηχανισμούς ελέγχους, που αναπτύσσουμε, για να διαπιστώσουμε αν μια πεντάδα σημείων τους ικανοποιεί.

3.1.4.1 Έλεγχος τοπολογικής συνέπειας

Θεωρούμε ότι η τοπολογία ενός ορόσημου είναι ικανοποιητική, αν οι αποστάσεις των σημείων, ανά δύο, είναι μεταξύ μίας ελάχιστης και μιας μέγιστης απόστασης. Δηλαδή, αν κάποια πεντάδα περιέχει σημεία, των οποίων η ελάχιστη/μέγιστη απόσταση δεν υπακούει στις προδιαγραφές, τότε θεωρούμε ότι είναι ακατάλληλη να αποτελέσει ένα υπο-ορόσημο. Η απαίτηση αυτή απορρέει από το γεγονός ότι ένα χαρακτηριστικό αντικείμενο του χώρου δεν μπορεί να είναι ούτε μηδαμινών ούτε τεραστίων διαστάσεων. Τα σημεία της πεντάδας, ουσιαστικά, χαρακτηρίζουν το αντικείμενο, οπότε θα πρέπει



(α) (β)

Σχήμα 3.4: Σχηματική απεικόνιση πεντάδων ‘‘κακής’’ τοπολογικής συνέπειας

να έχουν μια τοπολογική συνέπεια. Αυτό σημαίνει, πως δεν πρέπει να είναι ούτε πολύ κοντινά, όπως στην περίπτωση της εικόνας 3.4-α, μιας και τότε χάνεται η διακριτικότητα του αντικειμένου, αλλά ούτε και πολύ μακρινά, όπως στην εικόνα 3.4-β, μιας και τότε χάνεται η έννοια του αντικειμένου. Να τονίσουμε βέβαια ότι το ‘‘κοντινό’’ και το ‘‘μακρινό’’ εκφράζονται με βάση τις διαστάσεις της εικόνας.

Συνοπτικά, η αλγοριθμική μορφή της διαδικασίας ελέγχου της τοπολογικής συνοχής των πέντε σημείων συνοψίζεται στο σχήμα 3.5:

```
function Topology (Quintuple Q, Distance Minimum, Distance Maximum)
begin
    foreach pair P_i, P_j from Q do
    begin
        dist := || P_i − P_j ||;  /* Ευκλίδεια απόσταση των δύο σημείων */
        if dist ∉ [Minimum, Maximum] then
        begin
            return bad − topology
        end
    end
    return good − topology
end
```

Σχήμα 3.5: Αλγοριθμική περιγραφή της διαδικασίας ελέγχου της τοπολογικής συνέπειας μιας πεντάδας

3.1.4.2 Έλεγχος συγγραμικότητας

Όπως αναφέραμε, τα σημεία στην πεντάδα μπορούν να υπάρχουν σε γενικές θέσεις, χωρίς τρία από αυτά να κείνται στην ίδια ευθεία. Αν τρία σημεία είναι ημι-συνευθειακά (quasi-collinear), το εμβαδόν του αντίστοιχου τριγώνου είναι κοντά στο μηδέν και ο υπολογισμός της αναλλοίωτης καθίσταται αριθμητικά ασταθής. Αν $\mathbf{x}_1$, $\mathbf{x}_2$ και $\mathbf{x}_3$ είναι τρία σημεία, η τιμή του εμβαδού του αντίστοιχου τριγώνου δεν αποτελεί αξιόπιστη ένδειξη ότι ο υπολογισμός τείνει να χαρακτηριστεί παθολογικός [117]. Αντίθετα, ακολουθώντας τον Kanatani [118], ορίζουμε το πίνακα ροπής $\mathcal{M}_{123}$ ως εξής:

$$\mathcal{M}_{123} = \sum_{i=1}^{3} \mathbf{x}_i \mathbf{x}_i^t \tag{3.18}$$

Ο πίνακας αυτός είναι γνωστός ως πίνακας Gram στη διεθνή βιβλιογραφία. Η ελάχιστη ιδιοτιμή του $\mathcal{M}_{123}$ είναι επίσης η ελάχιστη τετραγωνική παθολογική τιμή του πίνακα $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$ [119]. Η εγγύτητα της ελάχιστης ιδιοτιμής στο μηδέν καθορίζει το πόσο συνευθειακά είναι τα τρία σημεία[6]. Στην δική μας προσέγγιση θεωρούμε ότι μια πεντάδα είναι ακατάλληλη, όταν η ελάχιστη ιδιοτιμή όλων των δυνατών τριάδων είναι μικρότερη ενός κατωφλιού συγγραμικότητας. Ενώ η συγγραμικότητα είναι μια προβολικά αναλλοίωτη ιδιότητα, η ημι-συγγραμμικότητα, που περιλαμβάνει μετρικές σχέσεις δεν είναι. Συνοπτικά, η διαδικασία ελέγχου της συγγραμικότητας των πέντε σημείων απεικονίζεται στο σχήμα 3.6

[6]δηλαδή, την γραμμική εξάρτηση των διανυσμάτων τους

```
function IsCollinear (Quintuple Q, Collinearity − threshold T)
begin
        min := 10^6;
        foreach tripple (P_i, P_j, P_k) from Q do
        begin
                ev := eigen−value (P_i, P_j, P_k) ;
                if (ev < min) then
                begin
                        min := ev;
                end
        end
        if (min < T) then
        begin
                return they − are − collinear
        end else
        begin
                return they − are − not − collinear
        end
end
```

Σχήμα 3.6: Αλγοριθμική περιγραφή της διαδικασίας που ελέγχει αν μια πεντάδα περιέχει τρία συνευθειακά σημεία.

3.1.4.3 Έλεγχος αν πέντε σημεία είναι ομοεπίπεδα

Οι προβολικές αναλλοίωτες έχουν χρησιμοποιηθεί για τον χαρακτηρισμό άγνωστης γεωμετρικής δομής [106]. Στην περίπτωση αυτή, μπορούμε να εξετάσουμε αν πέντε σημεία της εικόνας είναι ομοεπίπεδα ή όχι. Ο έλεγχος αυτός μπορεί να επιτευγχθεί μέσω του υπολογισμού του διδιάστατου cross-ratio για δύο πεντάδες σημείων, ιδωμένες από δύο διαφορετικές οπτικές γωνίες, χωρίς εκ των προτέρων γνώση και βαθμονόμηση της κάμερας. Αν οι τιμές των αναλλοίωτων, που έχουν υπολογιστεί για τα σημεία, διαφέρουν *σημαντικά*, τότε τα σημεία δεν είναι ομοεπίπεδα. Αντιθέτως, αν η διαφορά των αναλλοίωτων είναι μικρή, τότε έχουμε μια ισχυρή ένδειξη, αλλά δυστυχώς όχι και μια απόδειξη, ότι τα σημεία είναι ομοεπίπεδα. Στο άρθρο τους [87] οι Sinclair et al. μοντελοποιούν την *σημαντική* διαφορά, χρησιμοποιώντας τη διασπορά των δύο προβολικά αναλλοίωτων ποσοτήτων. Πιο συγκεκριμένα, υπολογίζουν τη διασπορά των πέντε σημείων στην εικόνα. Η διασπορά αυτή χρησιμοποιείται για την εκτίμηση της διασποράς των δύο προβολικά αναλλοίωτων ποσοτήτων. Μελετούμε το μαθηματικό υπόβαθρο της μεθόδου, μιας και η δική μας τεχνική αποτελεί γενίκευσή της.

Η αβεβαιότητα στην εκτίμηση της θέσης των χαρακτηριστικών σε κάθε εικόνα, θεωρείται ότι ακολουθεί την κανονική κατανομή με διασπορές σ_x και σ_y στις αντίστοιχες συντεταγμένες. Οι διασπορές αυτές θα διαφέρουν ανάλογα με τον ανιχνευτή χαρα-

κτηριστικών που χρησιμοποιείται. Η γραμμική διασπορά των αναλλοίωτων μπορεί να υπολογιστεί ως ακολούθως:

$$Var(\mu) = \sum_{i=1}^{5} \frac{\partial \mu}{\partial \mathbf{x}_i} \begin{pmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{pmatrix} \frac{\partial \mu}{\partial \mathbf{x}_i}^{\mathrm{T}} \tag{3.19}$$

ή μιας και ο θόρυβος είναι ασυσχέτιστος του λάθους [86],

$$Var(\mu) = \sigma^2 \sum_{i=1}^{5} \left| \frac{\partial \mu}{\partial \mathbf{x}_i} \right|^2 \tag{3.20}$$

όπου ο συμβολισμός $|\mathbf{x}|$ αναφέρεται στο μέτρο του διανύσματος $\mathbf{x}$.

Οι μερικές παράγωγοι του μ ως προς τις τεταγμένες των σημείων της πεντάδας έχουν ως εξής:

$$\begin{aligned}
\frac{\partial \mu}{\partial x_1} &= \mu \left(\frac{|(\mathbf{e}_1, \mathbf{x}_2, \mathbf{x}_4)|}{|(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_4)|} + \frac{|(\mathbf{e}_1, \mathbf{x}_3, \mathbf{x}_5)|}{|(\mathbf{x}_1, \mathbf{x}_3, \mathbf{x}_5)|} - \frac{|(\mathbf{e}_1, \mathbf{x}_3, \mathbf{x}_4)|}{|(\mathbf{x}_1, \mathbf{x}_3, \mathbf{x}_4)|} - \frac{|(\mathbf{e}_1, \mathbf{x}_2, \mathbf{x}_5)|}{|(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_5)|} \right) \\
\frac{\partial \mu}{\partial x_2} &= \mu \left(\frac{|(\mathbf{x}_1, \mathbf{e}_1, \mathbf{x}_4)|}{|(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_4)|} - \frac{|(\mathbf{x}_1, \mathbf{e}_1, \mathbf{x}_5)|}{|(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_5)|} \right) \\
\frac{\partial \mu}{\partial x_3} &= \mu \left(\frac{|(\mathbf{x}_1, \mathbf{e}_1, \mathbf{x}_5)|}{|(\mathbf{x}_1, \mathbf{x}_3, \mathbf{x}_5)|} - \frac{|(\mathbf{x}_1, \mathbf{e}_1, \mathbf{x}_4)|}{|(\mathbf{x}_1, \mathbf{x}_3, \mathbf{x}_4)|} \right) \\
\frac{\partial \mu}{\partial x_4} &= \mu \left(\frac{|(\mathbf{x}_1, \mathbf{x}_2, \mathbf{e}_1)|}{|(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_4)|} - \frac{|(\mathbf{x}_1, \mathbf{x}_3, \mathbf{e}_1)|}{|(\mathbf{x}_1, \mathbf{x}_3, \mathbf{x}_4)|} \right) \\
\frac{\partial \mu}{\partial x_5} &= \mu \left(\frac{|(\mathbf{x}_1, \mathbf{x}_3, \mathbf{e}_1)|}{|(\mathbf{x}_1, \mathbf{x}_3, \mathbf{x}_5)|} - \frac{|(\mathbf{x}_1, \mathbf{x}_2, \mathbf{e}_1)|}{|(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_5)|} \right)
\end{aligned} \tag{3.21}$$

Με παρόμοια συλλογιστική, οι μερικές παράγωγοι του μ ως προς τις τετμημένες των σημείων της πεντάδας έχουν ως εξής:

$$\begin{aligned}
\frac{\partial \mu}{\partial y_1} &= \mu \left(\frac{|(\mathbf{e}_2, \mathbf{x}_2, \mathbf{x}_4)|}{|(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_4)|} + \frac{|(\mathbf{e}_2, \mathbf{x}_3, \mathbf{x}_5)|}{|(\mathbf{x}_1, \mathbf{x}_3, \mathbf{x}_5)|} - \frac{|(\mathbf{e}_2, \mathbf{x}_3, \mathbf{x}_4)|}{|(\mathbf{x}_1, \mathbf{x}_3, \mathbf{x}_4)|} - \frac{|(\mathbf{e}_2, \mathbf{x}_2, \mathbf{x}_5)|}{|(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_5)|} \right) \\
\frac{\partial \mu}{\partial y_2} &= \mu \left(\frac{|(\mathbf{x}_1, \mathbf{e}_2, \mathbf{x}_4)|}{|(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_4)|} - \frac{|(\mathbf{x}_1, \mathbf{e}_2, \mathbf{x}_5)|}{|(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_5)|} \right) \\
\frac{\partial \mu}{\partial y_3} &= \mu \left(\frac{|(\mathbf{x}_1, \mathbf{e}_2, \mathbf{x}_5)|}{|(\mathbf{x}_1, \mathbf{x}_3, \mathbf{x}_5)|} - \frac{|(\mathbf{x}_1, \mathbf{e}_2, \mathbf{x}_4)|}{|(\mathbf{x}_1, \mathbf{x}_3, \mathbf{x}_4)|} \right) \\
\frac{\partial \mu}{\partial y_4} &= \mu \left(\frac{|(\mathbf{x}_1, \mathbf{x}_2, \mathbf{e}_2)|}{|(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_4)|} - \frac{|(\mathbf{x}_1, \mathbf{x}_3, \mathbf{e}_2)|}{|(\mathbf{x}_1, \mathbf{x}_3, \mathbf{x}_4)|} \right) \\
\frac{\partial \mu}{\partial y_5} &= \mu \left(\frac{|(\mathbf{x}_1, \mathbf{x}_3, \mathbf{e}_2)|}{|(\mathbf{x}_1, \mathbf{x}_3, \mathbf{x}_5)|} - \frac{|(\mathbf{x}_1, \mathbf{x}_2, \mathbf{e}_2)|}{|(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_5)|} \right)
\end{aligned} \tag{3.22}$$

όπου

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad \text{και} \quad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

Οι μερικές παράγωγοι του ν, ως προς τις συντεταγμένες των σημείων της πεντάδας, μπορούν να κατασκευαστούν παρόμοια. Οι εκφράσεις αυτές επιτρέπουν τον υπολογισμό της διασποράς των δύο αναλλοίωτων κατευθείαν από τις συντεταγμένες των πέντε σημείων $\mathbf{x}_i$ στην πρώτη εικόνα. Για να διαπιστώσουμε αν πέντε σημεία, ορατά σε δύο διαδοχικές εικόνες, κείνται στο αυτό επίπεδο, θα πρέπει η διαφορά των αναλλοίωτων να ικανοποιεί τις εξής συνθήκες:

$$\begin{aligned} |\mu' - \mu| &< 2\sqrt{Var(\mu)} \\ |\nu' - \nu| &< 2\sqrt{Var(\nu)} \end{aligned} \tag{3.23}$$

όπου το μ δηλώνει την αναλλοίωτη υπολογισμένη στην πρώτη εικόνα και το μ' την αναλλοίωτη υπολογισμένη στη δεύτερη εικόνα. Το όριο του διπλασίου της εκτιμώμενης τυπικής απόκλισης σε μια αναλλοίωτη επιλέχθηκε ως η μέγιστη επιτρεπόμενη διαφορά στις αναλλοίωτες δύο συνόλων με πέντε στοιχεία.

Ωστόσο, η παραπάνω ανάλυση έχει βάση μόνο όταν είναι εκ των προτέρων γνωστή η αντιστοίχιση μεταξύ των σημείων των δύο συνόλων[7]. Η δική μας προσέγγιση αποτελεί γενίκευση της ιδέας, αφού στη θέση των μεταθετικά ευαίσθητων cross-ratios χρησιμοποιούμε την προβολικά και μεταθετικά αναλλοίωτη της εξίσωσης 2.45. Αναλογικά με την εξίσωση 3.23, δύο πεντάδες σημείων θα κείνται στο αυτό επίπεδο αν:

$$\boxed{|\mathcal{K}[\mu,\nu] - \mathcal{K}[\mu',\nu']| < 2\sqrt{Var(\mathcal{K}[\mu,\nu])}} \tag{3.24}$$

όπου η μορφή της $Var(\mathcal{K}[\mu,\nu])$, αναλογικά με την εξίσωση 3.20 είναι η εξής:

$$Var(\mathcal{K}[\mu,\nu]) = \sigma^2 \sum_{i=1}^{5} \left| \frac{\partial \mathcal{K}[\mu,\nu]}{\partial \mathbf{x}_i} \right|^2 \tag{3.25}$$

Όπως γίνεται εμφανές, θα πρέπει να υπολογίσουμε τις μερικές παραγώγους της διδιάστατης παράστασης $\mathcal{K}$ ως προς τα σημεία της πεντάδας στο πρώτο πλαίσιο. Σύμφωνα με τον κανόνα της αλυσίδας έχουμε:

$$\frac{\partial \mathcal{K}[\mu,\nu]}{\partial \mathbf{x}_i} = \frac{\partial \mathcal{K}[\mu,\nu]}{\partial \mu} \cdot \frac{\partial \mu}{\partial \mathbf{x}_i} + \frac{\partial \mathcal{K}[\mu,\nu]}{\partial \nu} \cdot \frac{\partial \nu}{\partial \mathbf{x}_i} \tag{3.26}$$

όπου

$$\frac{\partial \mathcal{K}[\mu,\nu]}{\partial \mu} = \frac{\partial J[\mu]}{\partial \mu} + \frac{\partial J\left[\frac{\mu}{\nu}\right]}{\partial \mu} + \frac{\partial J\left[\frac{\nu-1}{\mu-1}\right]}{\partial \mu} + \frac{\partial J\left[\frac{\mu(\nu-1)}{\nu(\mu-1)}\right]}{\partial \mu} \tag{3.27}$$

[7]Κάτι που απαιτεί προϋπόθεση στο άρθρο

και

$$\frac{\partial \mathcal{K}[\mu,\nu]}{\partial \nu} = \frac{\partial J[\nu]}{\partial \nu} + \frac{\partial J\left[\frac{\mu}{\nu}\right]}{\partial \nu} + \frac{\partial J\left[\frac{\nu-1}{\mu-1}\right]}{\partial \nu} + \frac{\partial J\left[\frac{\mu(\nu-1)}{\nu(\mu-1}\right]}{\partial \nu} \quad (3.28)$$

Υποθέτοντας ότι:

$$G[\lambda] = \frac{\partial J[\lambda]}{\partial \lambda} \quad (3.29)$$

με λίγες αλγεβρικές πράξεις καταλήγουμε ότι:

$$\frac{\partial \mathcal{K}[\mu,\nu]}{\partial \mu} = G[\mu] + \frac{1}{\nu} G\left[\frac{\mu}{\nu}\right] - \frac{\nu-1}{(\mu-1)^2}\left(G\left[\frac{\nu-1}{\mu-1}\right] - \frac{1}{\nu} G\left[\frac{\mu(\nu-1)}{\nu(\mu-1)}\right]\right) \quad (3.30)$$

και

$$\frac{\partial \mathcal{K}[\mu,\nu]}{\partial \nu} = G[\nu] - \frac{\mu}{\nu^2} G\left[\frac{\mu}{\nu}\right] + \frac{1}{\mu-1}\left(G\left[\frac{\nu-1}{\mu-1}\right] + \frac{\mu}{\nu^2} G\left[\frac{\mu(\nu-1)}{\nu(\mu-1)}\right]\right) \quad (3.31)$$

Συνδυάζοντας τις εξισώσεις 3.26, 3.30 και 3.31 η εξίσωση 3.25 είναι άμεση και το κριτήριο ελέγχου ομοεπίπεδων πεντάδων μπορεί να υπολογιστεί. Η συνάρτηση J που χρησιμοποιούμε είναι, όπως έχουμε τονίσει, η J_{14}, μιας και αυτή παρουσιάζει την καλύτερη συμπεριφορά από όλες τις άλλες. Άρα συνοπτικά για να ελέγξουμε αν πέντε σημεία, ορατά σε δύο διαδοχικά πλαίσια, κείνται στο αυτό επίπεδο, απλά εκτελούμε τα βήματα του σχήματος 3.7

```
function IsCoplanar (Quintuple Q, Quintuple Q')
begin
        K1 := K(Q);
        K2 := K(Q');
        if ( |K1 − K2| < 2 · sqrt(Var(K1)) ) then
        begin
                return they − are − coplanar
        end else
        begin
                return they − are − not − coplanar
        end
end
```

Σχήμα 3.7: Αλγοριθμική περιγραφή της διαδικασίας ελέγχου αν πέντε σημεία είναι ομοεπίπεδα

3.1.5 Ο αλγόριθμος εξαγωγής ορόσημων

Μέχρι τώρα περιγράψαμε βασικές έννοιες, που είναι χρήσιμες στην κατανόηση του αλγορίθμου εξαγωγής ορόσημων. Το επόμενο βήμα είναι να αναλύσουμε τη ροή λειτουργίας του αλγόριθμου εξαγωγής. Αρχικά εξετάζουμε τον τρόπο εξαγωγής των χαρακτηριστικών σε κάθε περιοχή αναζήτησης.

3.1.5.1 Εύρεση χαρακτηριστικών σε κάθε περιοχή αναζήτησης

Όπως έχουμε ήδη αναφέρει, οι περιοχές αναζήτησης είναι ορθογώνιες. Η επιλογή αυτή δεν έγινε τυχαία, αλλά αποσκοπούσε στον εύκολο έλεγχο των *εσωτερικών σημείων* μιας περιοχής. Ουσιαστικά, για να ελέγξουμε αν τα χαρακτηριστικά είναι εσωτερικά σε μια περιοχή διαστάσεων $W_a \times W_b$, απλά ελέγχουμε αν οι συντεταγμένες τους είναι εντός των ορίων που σχηματίζονται με βάση το κέντρο και τις διαστάσεις της περιοχής. Η διαδικασία αυτή συνοψίζεται, μέσω αλγοριθμικής περιγραφής, στο σχήμα 3.8

```
function SelectFeatures (Feature−List fl, Region R)
begin
        center := CENTER(R)
        A_x := center_x − W_a;
        A_y := center_y − W_b;
        B_x := center_x + W_a;
        B_y := center_y + W_b;
        foreach point P from fl do
        begin
                if (P_x ∈ [A_x, B_x] ∧ P_y ∈ [A_y, B_y]) then
                begin
                        flist := Insert(P);
                end
        end
        return flist
end
```

Σχήμα 3.8: Αλγοριθμική περιγραφή της διαδικασίας εύρεσης χαρακτηριστικών στην περιοχή αναζήτησης R

3.1.5.2 Αναζήτηση υπο-ορόσημων σε δύο διαδοχικά πλαίσια

Έχοντας εξάγει τη λίστα των χαρακτηριστικών σε μια περιοχή αναζήτησης, εκτελούμε τα εξής βήματα. Αν n ο αριθμός των χαρακτηριστικών αυτών, τότε κατασκευάζουμε όλους τους $\binom{n}{5}$ συνδυασμούς πεντάδων. Η προσέγγιση μας είναι εξαντλητική, γιατί δεν έχουμε καμία εκ των προτέρων γνώση για κάποιο από τα χαρακτηριστικά της περιοχής που να μας οδηγεί στο να μην το λάβουμε υπόψιν μας. Οι Sinclair et al. [87] υποστηρίζουν ότι το $\binom{n}{5}$ αποτελεί ένα υπερβολικό μεγάλο πλήθος περιπτώσεων. Ο ισχυρισμός τους είναι σωστός, εφόσον το n λαμβάνει υψηλές τιμές, σύμφωνα με το παράδειγμα που παραθέτουν. Ωστόσο, η αναζήτηση ποιοτικών χαρακτηριστικών σε ''προέχουσες'' περιοχές της εικόνας μειώνει αισθητά το πλήθος των χαρακτηριστικών που συμμετέχουν στους υπολογισμούς. Με άλλα λόγια, πετυχαίνουμε να περιορίσουμε το n σε μικρές τιμές. Στα πειράματα που επιτελέσαμε ο μέγιστος αριθμός για το n ήταν

το 20, κάτι που καθιστά το υπολογιστικό κόστος ανεκτό. Αν και τα χαρακτηριστικά δεν υπόκεινται σε κάποιους περιορισμούς, εντούτοις οι πεντάδες των σημείων μας υπόκειντε στους περιορισμούς που αναφέραμε, περιορίζοντας αισθητά το υπολογιστικό κόστος.

Για κάθε μία πεντάδα l_{ij} με σημεία $P_{i,j,k}$, που ικανοποιεί το κριτήριο της μη-συγγραμικότητας και της τοπολογικής συνοχής, υπολογίζουμε την $\mathbf{P}^2$-αναλλοίωτη $\mathcal{K}$. Προτού αποφανθούμε για το αν η πεντάδα l_{ij} μπορεί να αποτελέσει ένα υπο-ορόσημο, θα πρέπει να αποδείξουμε ότι τα πέντε σημεία της $P_{i,j,k}$ είναι ομοεπίπεδα.

Μιας και η ρομποτική πλατφόρμα υποτίθεται ότι κινείται αργά σε σχέση με την καταγραφή των εικόνων, μπορούμε με ασφάλεια, να υποθέσουμε ότι τα διαδοχικά πλαίσια $\mathcal{F}, \mathcal{F}'$ δεν διαφέρουν σε μεγάλο ποσοστό στην δομή τους. Η υπόθεση αυτή επιτρέπει την ανίχνευση νέων πεντάδων ${l'}_{ij}$ στο διαδοχικό πλαίσιο $\mathcal{F}'$, οι οποίες είναι οι βασικές αντίστοιχες υποψήφιες για κάθε πεντάδα l_{ij} στο $\mathcal{F}$, μέσω επανάληψης των άνωθι βημάτων.

Όπως έχει ήδη αναφερθεί το *κριτήριο διατήρησης της προβολικής αναλλοίωτης*, δηλαδή το γεγονός ότι ο υπολογισμός της αναλλοίωτης θα πρέπει να παράγει την ίδια τιμή για διαδοχικές πεντάδες σε διαφορετικά πλαίσια, είναι μια ισχυρή ένδειξη, αλλά δεν αποτελεί απόδειξη, ότι τα πέντε σημεία είναι ομοεπίπεδα και οι πεντάδες l_{ij} και l'_{ij} είναι αντίστοιχες. Αυτό σημαίνει, πως δύο πεντάδες σημείων, που έχουν παρόμοιες τιμές στις αναλλοίωτες τους, μπορεί να διαφέρουν στην τοπολογία τους και να μην είναι αντίστοιχες. Εμείς όμως επιθυμούμε, να βρούμε την αντίστοιχη μιας πεντάδας στο διαδοχικό πλαίσιο. Στην προσπάθειά μας αυτή, δεν θέλουμε να εκτιμήσουμε την αντιστοίχιση του κάθε σημείου ανεξάρτητα, αλλά την αντιστοίχιση της πεντάδας των σημείων στο σύνολό της. Για να αντιμετωπίσουμε τις αριθμητικές αστάθειες της γεωμετρικής αναλλοίωτης, θα πρέπει να περιορίσουμε τον χώρο αναζήτησης στο $\mathcal{F}'$ σε μερικές[8] πεντάδες για τις οποίες υπάρχει η ένδειξη της αντιστοίχισης. Για την επίτευξη του σκοπού αυτού, θα πρέπει να χρησιμοποιήσουμε μία μετρική που να μας παρέχει μια ένδειξη[9] ότι δύο πεντάδες σημείων είναι αντίστοιχες. Περιγράφουμε, την συλλογιστική πορεία, που ακολουθήσαμε για την εξαγωγή της απαιτούμενης μετρικής.

3.1.5.3 Κριτήριο αντιστοίχισης μεταξύ δύο πεντάδων

Η επιλογή του κριτηρίου αντιστοίχισης που χρειαζόμαστε, δεν θα πρέπει να λαμβάνει υπόψιν την αντιστοίχιση σημείου-προς-σημείου. Προτού καταλήξουμε στο κριτήριο, μελετήσαμε και δοκιμάσαμε αρκετές μεθόδους, οι οποίες απορρίφθηκαν είτε λόγω των μη ικανοποιητικών αποτελεσμάτων τους είτε λόγω των συνθηκών που έπρεπε να ικανοποιούνται, για την αποδοτική τους λειτουργία. Περιγράφουμε ορισμένες από τις μεθόδους, που δοκιμάσαμε, για να αιτιολογήσουμε την επιλογή μας.

[8] Στην ιδανική περίπτωση σε μία

[9] με μικρό βαθμό αβεβαιότητας

Αρχικά, δοκιμάσαμε το κριτήριο της απόστασης *Hausdorff* [120], το οποίο μέσω μιας μετρικής τύπου *min-max* υπολογίζει το βαθμό ομοιότητας δύο συνόλων σημείων. Το μειονέκτημα του είναι ότι λειτουργεί ικανοποιητικά μόνο για σύνολα σημείων που έχουν υποστεί μέχρι και αφφινικό μετασχηματισμό. Στην δική μας περίπτωση, όμως, οι μετασχηματισμοί είναι γενικότερης φύσεως, με αποτέλεσμα η μέθοδος να αδυνατεί να τους αντιμετωπίσει ικανοποιητικά. Επίσης, μιας και η μέθοδος βασίζεται αυστηρά σε κριτήρια απόστασης, είναι λογικό να είναι ευαίσθητη στο θόρυβο θέσης των στοιχείων των συνόλων. Το θεωρητικό αποτέλεσμα, επιβεβαιώθηκε και στην πράξη, μέσω των πειραμάτων που επιτελέσαμε σε συλλογές σημείων σε διαδοχικά πλαίσια ακολουθιών με διαδρόμους. Στην ίδια κατηγορία με την απόσταση Hausdorff, ανήκουν και οι περιπτώσεις των ροπών Zernike (Zernike moments) [121], των αναλλοίωτων ροπών [122] και των περιγραφών Fourier (Fourier descriptors) [123].

Στην πορεία αναζήτησης ικανοποιητικής μεθόδου, δοκιμάσαμε διάφορα στατιστικά κριτήρια στα δύο σύνολα σημείων, σε μια προσπάθεια σύγκρισης των διδιάστατων κατανομών των. Στα πειράματα αυτά εφαρμόσαμε τα Kolmogoroff-Smirnoff και χ^2 τεστ, όπως διατυπώνονται στο [124]. Ωστόσο η εφαρμογή των, δεν απέδωσε τα αναμενόμενα αποτελέσματα. Η βασική αιτία ήταν ότι το πλήθος των σημείων του κάθε συνόλου ήταν μικρό (=5), με αποτέλεσμα το τεστ να μην αποδίδει ικανοποιητικά, λόγω της αδυναμίας του στατιστικού εκτιμητή να παράγει αξιόπιστα αποτέλεσμα για την κατανομή του κάθε συνόλου.

Κατόπιν, δοκιμάσαμε να συγκρίνουμε τις διασπορές των συντεταγμένων των στοιχείων των δύο συνόλων, αλλά διαπιστώσαμε πως δεν ήταν σημαντική η πληροφορία που μας έδιναν, μιας και η τοπολογία των σημείων καθόριζε την τιμή τους. Τελικά, καταλήξαμε στη χρήση της τεχνικής του *πίνακα συνδιασποράς (covariance matrix)*. Στην επιλογή μας αυτή, υιοθετούμε την προσέγγιση των [36, 125, 126], οι οποίοι θεωρούν ότι η αβεβαιότητα θέσης μπορεί να περιγραφεί μέσω μιας Γκαουσιαννής κατανομής με μέση τιμή $\overline{\mathbf{p}}$ και πίνακα συνδιασποράς Σ_p. Οι συγγραφείς θεωρούν ότι ο πίνακας συνδιασποράς Σ'_p της αβεβαιότητας μετά τον μετασχηματισμό κίνησης συνδέεται με τον πίνακα συνδιασποράς Σ_p της αβεβαιότητας πριν τον μετασχηματισμό κίνησης, μέσω της παρακάτω σχέσης:

$$\Sigma'_p = \left(\frac{\partial \mathbf{h}}{\partial \mathbf{p}}\right) \Sigma_p \left(\frac{\partial \mathbf{h}}{\partial \mathbf{p}}\right)^T \tag{3.32}$$

Οι παράγωγοι $\frac{\partial \mathbf{h}}{\partial \mathbf{p}}$ είναι οι Ιακωβιανές του μετασχηματισμού κίνησης $\mathbf{h}$. Με τον τρόπο αυτό μοντελοποιείται η αβεβαιότητα της θέσης του μηχανικού συστήματος κατά την κίνηση του στο χώρο. Στη δική μας περίπτωση, η κίνηση της ρομποτικής πλατφόρμας είναι μικρή, με αποτέλεσμα η συνεισφορά των Ιακωβιανών να είναι αμελητέα στην εκτίμηση της αβεβαιότητας. Για το λόγο αυτό, παρακάμπτουμε την συνεισφορά του μετασχηματισμού στη μοντελοποίηση της αβεβαιότητας, προσπαθώντας να έχουμε απλά μια προσεγγιστική εκτίμησή της. Έτσι η δική μας έκδοση της μοντελοποίησης της

αβεβαιότητας είναι η ελαχιστοποίηση του ακόλουθου υπολοίπου:

$$\epsilon \;=\; \left|\Sigma'_p - \Sigma_p\right| \tag{3.33}$$

Ο *πίνακας συνδιασποράς*, λειτουργεί σαν μετρική αντιστοίχισης μεταξύ της πεντάδας l_{ij} και των υποψήφιων πεντάδων ${l'}_{ij}$, διακριτοποιόντας και συγκρίνοντας την χωρική κατανομή των σημείων των. Υιοθετώντας τον μηχανισμό αυτό, προσπερνάμε ικανοποιητικά το πρόβλημα της σημείο-προς-σημείο αντιστοίχισης μεταξύ των σημείων στην πεντάδα l_{ij} και των υποψήφιων πεντάδων ${l'}_{ij}$, λύνοντας το πιο εύκολο πρόβλημα της αντιστοίχισης της χωρικής κατανομής μεταξύ πεντάδων σημείων. Ουσιαστικά, η τεχνική του πίνακα συνδιασποράς, αποτελεί ένα ποσοτικό κριτήριο ομοιότητας μεταξύ των χωρικών κατανομών των σημείων των δύο πεντάδων. Η εισαγωγή του δικαιολογείται από το γεγονός ότι η κατανομή αυτή μεταβάλλεται αμυδρά όταν η ίδια σκηνή απεικονίζεται από κοντινές οπτικές γωνίες.

Για να ανιχνεύσουμε υποψήφιες πεντάδες, δημιουργούμε μια περιοχή η οποία περιέχει την πεντάδα l_{ij} στο πρώτο πλαίσιο. Στην περιοχή αυτή, που προκύπτει υπολογίζοντας το *Ελάχιστο Περίβλημα (Minimum Bounding Box)* της πεντάδας l_{ij}, εστιάζουμε την προσοχή μας στο διαδοχικό πλαίσιο $\mathcal{F}'$, επεκτείνοντάς την κατάλληλα για να αντιμετωπίσουμε την κίνηση της πλατφόρμας στο διάστημα λήψης των διαδοχικών εικόνων. Για να επιβεβαιώσουμε ότι η ταύτιση των πεντάδων σε διαδοχικά πλαίσια είναι σωστή και να καταλήξουμε σε μία, ή μερικές, αντίστοιχες πεντάδες στο $\mathcal{F}'$ ενεργούμε ως εξής. Για κάθε δυνατή πεντάδα χαρακτηριστικών στο πλαίσιο $\mathcal{F}$ κατασκευάζουμε όλα τα πιθανά ζεύγη μεταξύ κάθε πεντάδας l_{ij} που αναγνωρίζεται στο $\mathcal{F}$ και των υποψήφιων αντίστοιχων πεντάδων της, ${l'}_{ij}$, στο $\mathcal{F}'$. Για κάθε τέτοιο ζεύγος $(l_{ij}, {l'}_{ij})$ εξετάζουμε τις αποστάσεις των πινάκων συνδιασποράς των. Κατασκευάζουμε τον πίνακα συνδιασποράς για κάθε πεντάδα σύμφωνα με τον μη-προκατειλλημένο εκτιμητή [127]:

$$\hat{\Sigma}_u \;=\; \frac{1}{n-1}\sum_{i=1}^{n}(\mathbf{x}_k \,-\, \mu)\,(\mathbf{x}_k \,-\, \mu)^T \tag{3.34}$$

όπου το διάνυσμα $\mathbf{x}_k \;=\; (x_k, y_k)\,,\; k \;=\; 1,\ldots,5$ υποτίθεται ότι είναι η εκτιμώμενη κατανομή που ακολουθείται από τα σημεία $P_{i,j,k}$ σε κάθε πεντάδα και μ το διάνυσμα των *μέσων τιμών*.

Για να συγκρίνουμε τους δύο πίνακες συνδιασποράς, χρειαζόμαστε ένα μέτρο που να μας παρέχει αξιόπιστα τις διαφορές των δύο πινάκων. Αρχικά, η σκέψη, ήταν να συγκρίνουμε τα αντίστοιχα στοιχεία των δύο πινάκων. Αυτό όμως θα απαιτούσε να επιτελούμε ταυτόχρονα τρεις[10] συγκρίσεις, κάτι που δεν μας διευκολύνει στην εξαγωγή συμπερασμάτων. Ο αντικειμενικός σκοπός είναι να αναχθούμε στην σύγκριση μόνο μιας ποσότητας και όχι όλου του πίνακα. Η ποσότητα αυτή είναι η *νόρμα απείρου*

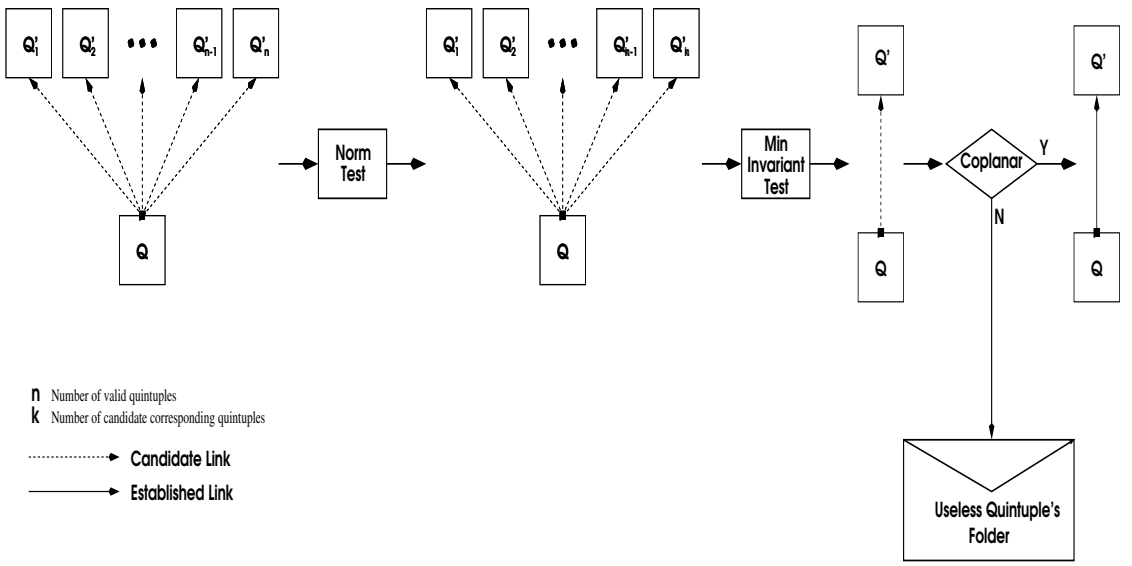
[10]Μιας και ο πίνακας είναι συμμετρικός

του πίνακα διαφοράς των δύο πινάκων συνδιασποράς, εκφραζόμενη κατά γραμμές[11], η οποία μας παρέχει και το μέγιστο δυνατό σφάλμα στις εκτιμήσεις μας [128]. Ο τύπος της είναι ο εξής:

$$\| C \|_{\infty} \quad = \quad \max \sum_{q=1}^{2} \sum_{r=1}^{2} \left| |a_{q,r}| \; - \; \left| a'_{q,r} \right| \right| \tag{3.35}$$

όπου τα $a_{q,r}$ και $a'_{q,r}$ είναι τα αντίστοιχα στοιχεία των δύο πινάκων συνδιασποράς, που έχουν υπολογιστεί για κάθε ζεύγος πεντάδων (l_{ij}, l'_{ij}) υπό θεώρηση. Όσο πιο μικρή είναι η τιμή της νόρμας, τόσο πιο υψηλή η ποιότητα της ταύτισης.

Άρα ουσιαστικά οι πεντάδες l'_{ij}, των οποίων η διαφορά της νόρμας του πίνακα συνδιασποράς των και της νόρμας της πεντάδας l_{ij} είναι φραγμένη άνω από κάποιο κατώφλι, θεωρούνται υποψήφιες αντίστοιχες της, για τις οποίες υπάρχει η ένδειξη της αντιστοίχισης. Στη συνέχεια, από τις υποψήφιες αυτές πεντάδες, επιλέγουμε αυτή, της οποίας η αναλλοίωτη είναι πιο κοντά στην αναλλοίωτη της l_{ij}. Σαν ένα τελευταίο βήμα επιβεβαιώνουμε ότι η διαφορά των τιμών των $\mathbf{P}^2$-αναλλοίωτων της αρχικά επιλεχθείσας πεντάδας l_{ij} και και της πεντάδας, που η τεχνική του πίνακα συνδιασποράς χαρακτήρισε ως αντίστοιχη, υπακούουν στην εξίσωση 3.24. Το γεγονός αυτό αποτελεί μια ισχυρή ένδειξη, ότι τα πέντε υπό-μελέτη σημεία είναι ομοεπίπεδα και άρα τα σημεία $P_{i,j,k}$ που αρχικά επιλέχθηκαν ικανοποιούν τις απαιτήσεις ενός υπο-ορόσημου l_{ij}. Στην εικόνα 3.9 απεικονίζεται γραφικά η διαδικασία εύρεσης της αντίστοιχης πεντάδας l'_{ij} δοθέντος της αρχικής l_{ij}



Σχήμα 3.9: Η διαδικασία εύρεσης της αντιστοίχου μιας πεντάδας

[11]Η νόρμα απείρου κατά στήλες παράγει το ίδιο αποτέλεσμα, αφού ο πίνακας είναι συμμετρικός

3.1.5.4 Επιλογή των πιο εξαιρετικών πεντάδων

Από την μέχρι τώρα διαδικασία, αυτό που προκύπτει είναι ένα σύνολο ζευγών πεντάδων Q και Q'. Η αποθήκευση όλων αυτών των ζευγών πεντάδων οδηγεί αφενός μεν στην χωρική επιβάρυνση της βάσης και γενικότερα του αλγορίθμου, αφετέρου δε δυσχαιρένει εννοιολογικά και υπολογιστικά το έργο της αναγνώρισης, αφού οι τιμές των αναλλοίωτων ποσοτήτων έχουν περιορισμένο πεδίο τιμών, όπως φαίνεται από την γραφική παράσταση του σχήματος 2.10, με αποτέλεσμα τη συσσώρευση πολλών πεντάδων με τα ίδια χαρακτηριστικά. Για να αποφύγουμε αυτό το πρόβλημα, κατασκευάζουμε ένα μηχανισμό επιλογής, ο οποίος αναλαμβάνει να ανακαλύψει από τη λίστα των πεντάδων Q του πρώτου πλαισίου, τις πεντάδες εκείνες των οποίων ο συνδυασμός των τιμών της αναλλοίωτης και της νόρμας του πίνακα συνδιασποράς διαφέρει αισθητά από τις υπόλοιπες.

Ο μηχανισμός αυτός είναι δανεισμένος από την επιστήμη της Στατιστικής και προσπαθεί με μετρικά κριτήρια να επιλέξει ένα αριθμό ''επικρατέστερων'' διανυσμάτων από ένα πλήθος διαθέσιμων διανυσμάτων. Μιας και η τεχνική εργάζεται με διανύσματα, θα πρέπει με κάποιο τρόπο να αναχθούμε από την πεντάδα των σημείων και τα λοιπά χαρακτηριστικά της κάθε μιας, σε ένα διάνυσμα με κατάλληλα στοιχεία. Τα στοιχεία του διανύσματος δεν είναι τα πέντε στοιχεία, διότι τότε θα προσπαθούσαμε να επιλέξουμε τις πεντάδες εκείνες που έχουν την ''επικρατέστερη'' τοπολογία ή την ''επικρατέστερη'' συνοχή, κάτι που δυστυχώς δεν είναι μονοσήμαντα ερμηνεύσιμο. Έτσι, τροφοδοτούμε το διάνυσμα με τα κύρια χαρακτηριστικά της κάθε πεντάδας, την τιμή της αναλλοίωτης και τη νόρμα του πίνακα συνδιασποράς της. Αν το διάνυσμα το συμβολίσουμε ως $\mathbf{V}$, τότε θα είναι:

$$\mathbf{V} = \begin{bmatrix} I \\ \mathrm{N} \end{bmatrix} \tag{3.36}$$

όπου I η τιμή της $\mathbf{P}^2$-αναλλοίωτης της πεντάδας και N η νόρμα του πίνακα συνδιασποράς της. Επειδή οι τιμές των I και N έχουν διαφορετικά πεδία τιμών, προφανώς η ποσότητα που θα είναι πιο μεγάλη θα συνεισφέρει και πιο πολύ στην έκφανση της εκτίμησης. Για το λόγο αυτό, κανονικοποιούμε τις τιμές και των δύο στοιχείων του διανύσματος $\mathbf{V}$ στο διάστημα $[0, 1]$. Έτσι, και οι δύο ποσότητες συνεισφέρουν το ίδιο βάρος στη διαδικασία επιλογή της επικρατέστερης πεντάδας. Η κανονικοποίηση αυτή επιτελείται διαιρώντας όλες τις αναλλοίωτες I με τη μέγιστη τιμή της αναλλοίωτης των πεντάδων και όλες τις νόρμες N με τη μέγιστη τιμή των της νόρμας των πεντάδων. Στη συνέχεια, υπολογίζουμε την *σωρευτική απόσταση (aggregate distance)* κάθε διανύσματος, σύμφωνα με τον τύπο[12]:

$$O_q = \sum_{j=1}^{n} \| \mathbf{V}_q - \mathbf{V}_j \| \tag{3.37}$$

[12]το $\| \cdot \|$, είναι η ευκλίδεια απόσταση ανάμεσα σε δύο διανύσματα.

και ταξινομούμε τα διανύσματα σε αύξουσα σειρά σύμφωνα με τις τιμές της σωρευτικής απόστασής τους. Στο ταξινομημένο σύνολο, που προκύπτει, το τελευταίο στοιχείο αποτελεί το *πιο εξαιρετικό (most outlier)* διάνυσμα, που αντιστοιχεί στην πιο εξαιρετική πεντάδα. Στη συνέχεια, με αφετηρία το τελευταίο στοιχείο του συνόλου, ελέγχουμε ανά δύο διαδοχικά τα στοιχεία και αν η διαφορά των σωρευτικών αποστάσεών τους είναι σημαντική, τότε από κάθε ζεύγος επιλέγεται το στοιχείο που είναι πιο κοντά στο τέλος. Τελικώς, για τα στοιχεία που επιλέχθηκαν, με μια αντίστροφη διαδικασία απεικόνισης, ανακτούμε τις πληροφορίες που είχαν πριν την διανυσματική τους αναπαράσταση. Η αλγοριθμική όψη της διαδικασίας παρατίθεται στο σχήμα 3.10.

```
function OutlierEstimation (Quintuple−List ql)
begin
    foreach quintuple q_i from ql do
    begin
        V_i := EstimateVector(q_i);
        V_i := Normalize(V_i);
        vl  := Insert(V_i);
    end
    foreach vector V_i from vl do
    begin
        O_i := Aggregate(V_i);
        ol  := Insert(O_i);
    end
    sort (ol, increasing order);
    foreach pair (ol_i, ol_j) from end of ol do
    begin
        if (|ol_i − ol_j| ≥ T) then
        begin
            ValidQuintuples := RecoverInfo(ol_i);
        end
    end
    return ValidQuintuples
end
```

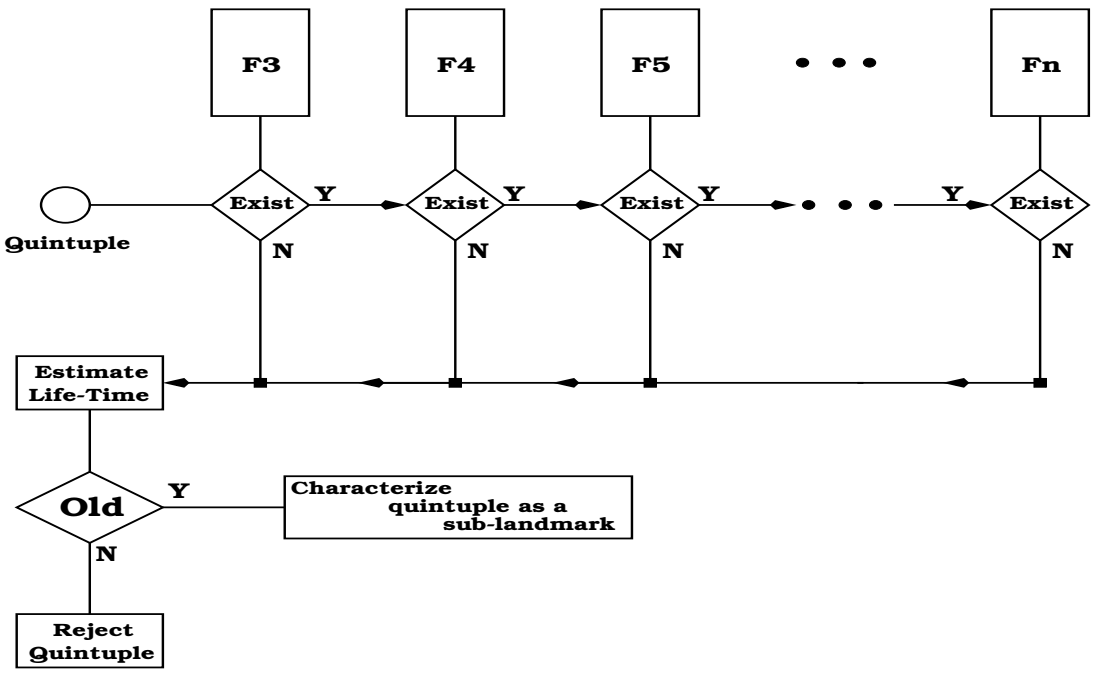
Σχήμα 3.10: Αλγοριθμος επιλογής εξαιρετικών πεντάδων

3.1.5.5 Επιβεβαίωση ‘‘χρονικής συνέπειας’’

Όπως αναφέραμε, στην εισαγωγή, η ανίχνευση και αναγνώριση ορόσημων θα πρέπει να επεκτείνεται και στη διάσταση του χρόνου, αφού η είσοδος στο μηχανικό σύστημα είναι μία ακολουθία εικόνων. Θα πρέπει λοιπόν να υπάρχει ένας μηχανισμός που να εγγυάται τη ‘‘χρονική συνέπεια’’ στη θέση και τα χαρακτηριστικά ενός ορόσημου, και

κατ' επέκταση ενός υπο-ορόσημου. Αυτό σημαίνει πως ένα υπο-ορόσημο, θα πρέπει να βρίσκεται σε γειτονική θέση και να έχει παρόμοια χαρακτηριστικά σε διαδοχικές εικόνες μέσα στο χρόνο. Με τον τρόπο αυτό, αφενός μεν επιτυγχάνεται μεγαλύτερη αξιοπιστία στην αναγνώριση των ορόσημων, αφού οι διαδοχικές εικόνες ''συναινούν'' ή όχι στις προηγούμενες αποφάσεις, αφετέρου το υπολογιστικό κόστος παραμένει χαμηλό με το να περιορίζονται κατά πολύ οι απαιτούμενοι υπολογισμοί.

Η μέχρι τώρα πορεία μας, έχει ανακαλύψει ένα σύνολο εξαιρετικών ζευγών πεντάδων, που ουσιαστικά είναι οι εξαιρετικές πεντάδες στο πρώτο πλαίσιο και οι αντίστοιχες αυτών στο δεύτερο. Η διατήρηση[13] όμως, σε δύο μόνο διαδοχικά πλαίσια δεν αποτελεί προϋπόθεση διατήρησης και στα επόμενα διαδοχικά πλαίσια. Για το λόγο αυτό, προτού αποφανθούμε αν οι πεντάδες που διαθέτουμε, αποτελούν υπο-ορόσημα, ελέγχουμε αν διατηρούνται και σε επόμενα διαδοχικά πλαίσια της ακολουθίας εισόδου. Επίσης, εισάγουμε την έννοια του *χρόνου ζωής* ή την *ηλικία* μιας πεντάδας, δανειζόμενοι τον όρο από τον τομέα των κατανεμημένων συστημάτων, όπου εκφράζει την ιστορία μιας διεργασίας. Ο χρόνος ζωής μιας πεντάδας πιστοποιεί τη διατήρηση της σε ένα προκαθορισμένο πλήθος από διαδοχικά πλαίσια και ορίζεται σαν το πλήθος των διαδοχικών πλαισίων που η πεντάδα αυτή διατηρείται. Η διαδικασία ελέγχου της



Σχήμα 3.11: Ο χρόνος ζωής ενός υπο-ορόσημου

διατήρησης, που απεικονίζεται στο σχήμα 3.11, αποτελεί επανάληψη της διαδικασίας εξαγωγής αντιστοίχου του σχήματος 3.9, όπου η πεντάδα Q είναι η υπο-μελέτη πεντάδα και το σύνολο $Q'_1, \ldots, Q'_n$, συνθέτουν οι πεντάδες που κατασκευάζονται στα διαδοχικά

[13] ουσιαστικά η ύπαρξη αντιστοίχου

πλαίσια της ακολουθίας. Για κάθε πλαίσιο που διατηρείται η πεντάδα, επαυξάνουμε το χρόνο ζωής της κατά ένα. Αν ο χρόνος ζωής που καταλήγουμε είναι μεγαλύτερος ενός αριθμού πλαισίων, θεωρούμε ότι η πεντάδα τηρεί τη συνθήκη της ''χρονικής συνέπειας'' και χαρακτηρίζεται υπο-ορόσημο. Η διαδικασία αυτή επιτελείται για κάθε μία από τις παραπάνω πεντάδες και τελικά καταλήγουμε σε ένα πλήθος από υπο-ορόσημα, του οποίου η σύνθεση παράγει ένα ορόσημο. Αλγοριθμικά, η διαδικασία επιβεβαίωσης της ''χρονικής συνέπειας'' μιας πεντάδας παρατίθεται στο σχήμα 3.12.

```
function LifeTime (Quintuple Q, Sequence I, LifeTimeThreshold T)
begin
        lt := 2;   /* Allready in first 2 frames */
        foreach frame I_i from I do
        begin
                if (Corresponding(Q, I_i) == EXIST) then
                begin
                        lt++;
                end
        end
        if (lt < T) then
        begin
                Characterize (Q, sub-landmark)
        end
end
```

Σχήμα 3.12: Αλγοριθμική περιγραφή της διαδικασίας επιβεβαίωσης της ''χρονικής συνέπειας'' μιας πεντάδας Q.

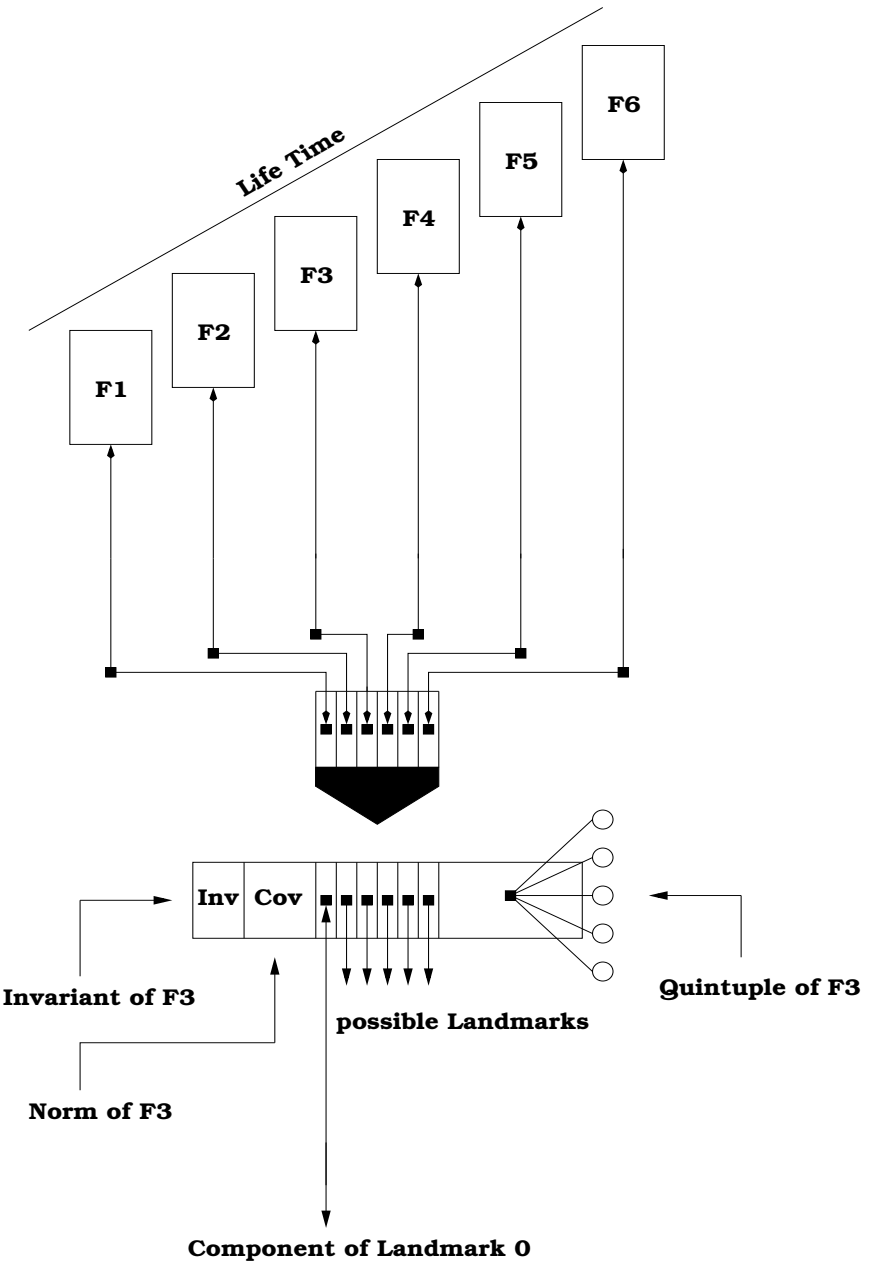
3.1.6 Κατασκευή του τοπολογικού χάρτη

Κατά την διάρκεια της περιόδου εκπαίδευσης, το μηχανικό σύστημα εξάγει και αποθηκεύει ορόσημα, καθώς και τα αντίστοιχα υπο-ορόσημά τους. Το κάθε υπο-ορόσημο αναπαρίσταται στην βιβλιοθήκη των υπο-ορόσημων χρησιμοποιώντας την αντίστοιχη *πλειάδα περιγραφής (descriptor tupple)*:

$$DT : \left\{ \begin{array}{c} I \\ N \\ L_n \\ \vdots \\ L_m \\ P_{i,j,k} \end{array} \right\}$$

όπου το I είναι η τιμή της αναλλοίωτης του υπο-ορόσημου, το N είναι η νόρμα του πίνακα συνδιασποράς του, το σύνολο $(L_n, \ldots, L_m)$ είναι το πλήθος των αμφίδρομων

συνδέσμων στα ορόσημα των οποίων αποτελεί συστατικό και τα $P_{i,j,k}$ είναι οι συντεταγμένες των χαρακτηριστικών που συνθέτουν την πεντάδα. Στην εικόνα 3.13 εμφανίζεται



Σχήμα 3.13: Παράδειγμα μιας πλειάδας περιγραφής

ένα στιγμιότυπο μιας πλειάδας περιγραφής. Ο χρόνος ζωής του υπο-ορόσημου, είναι έξι, μιας και διατηρείται σε έξι πλαίσια και τα στοιχεία της αντιστοιχούν στα στοιχεία του πλαισίου **F3**, μιας και αυτό είναι το μεσαίο στη συγκεκριμένη περίπτωση.

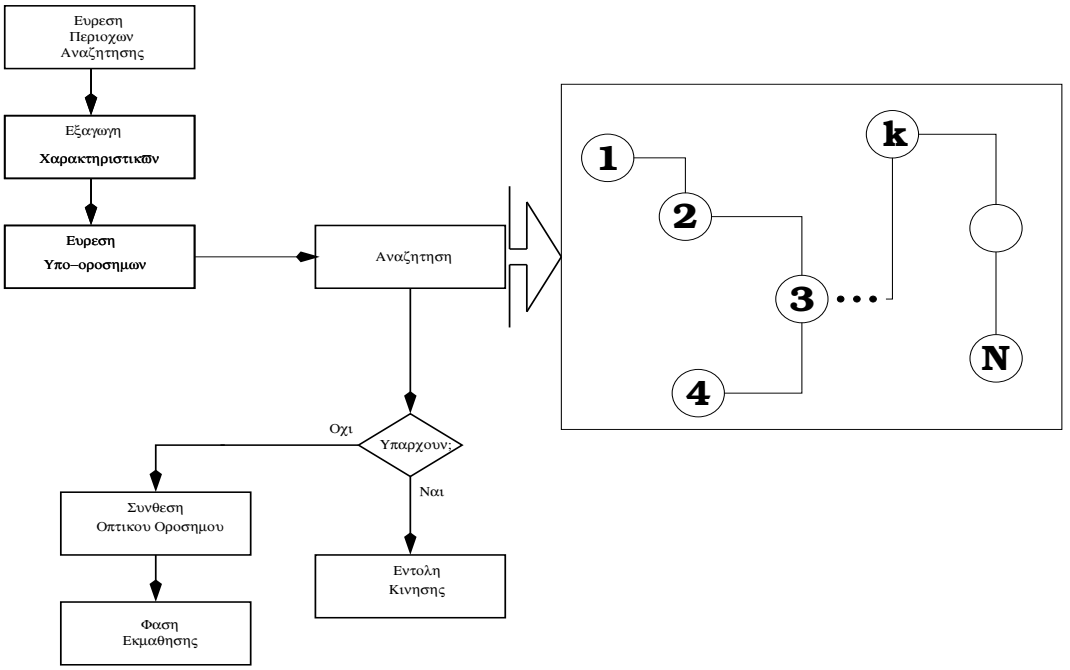
Οι αμφίδρομες συνδέσεις διευκολύνουν την άρση της όποιας σημασιολογικής σύγχισης, που παρατηρείται στην περίπτωση όπου ένα υπο-ορόσημο ανήκει σε ένα σύνολο από ορόσημα. Στην περίπτωση αυτή, με χρήση των συνδέσμων, μπορούμε να δια-

πιστώσουμε ευέλικτα το ορόσημο που κατέχει το συγκεκριμένο υπο-ορόσημο. Αυτό επιτυγχάνεται, υπολογίζοντας τα υπόλοιπα υπο-ορόσημα στο πλαίσιο που βρισκόμαστε και προσπαθώντας να τα αντιστοιχήσουμε στον τοπολογικό χάρτη. Τελικά, η αντιστοίχιση ενός ορόσημου συνίσταται στην επιτυχή αντιστοίχιση όλων των υπο-ορόσημων του στον τοπολογικό χάρτη. Η πληροφορία της πλειάδας περιγραφής, χρησιμοποιείται επίσης σημαντικά σε ζητήματα σχετικής οριοθέτησης των υπο-ορόσημων στη φάση πλοήγησης, προκειμένου να καθοριστούν περιοχές εστίασης της προσοχής καθώς και στη μείωση του χώρου αναζήτησης κατά τη διαδικασία επιβεβαίωσης των οπτικών ορόσημων. Επίσης αποθηκεύουμε αναφορές σε ενέργειες πλοήγησης ειδικά επιλεγμένες για κάθε ορόσημο, κατασκευάζοντας έτσι ένα τοπολογικό χάρτη, ο οποίος μπορεί να χρησιμοποιηθεί στην διαδικασία αναγνώρισης για λήψη αποφάσεων σχετικά με τη συγκεκριμένη αποστολή πλοήγησης (navigational task).

3.2 Αναγνώριση ορόσημων

Η επιτυχής λειτουργία ενός αυτόνομου μηχανικού συστήματος στο χώρο εργασίας του, σύμφωνα με την προτεινόμενη προσέγγιση, βασίζεται σε μεγάλο βαθμό στην αναγνώριση των οπτικών ορόσημων και εκμετάλλευση της πληροφορίας που παρέχεται από αυτά. Έχοντας λοιπόν εκπαιδευτεί το αυτόνομο σύστημα με ένα σύνολο ορόσημων, καλείται να αποδείξει ότι πράγματι η πληροφορία που σταχυολόγησε αποτελεί σημαντικό παράγοντα στην επιτυχία της αυτόνομης πλοήγησής του.

Στη φάση, λοιπόν, της πλοήγησης, όπως φαίνεται στο σχήμα 3.14, ακολουθείται πρώτα μια διεργασία, πανομοιότυπη με αυτή της φάσης εκμάθησης, μέχρις ότου εντοπιστεί ένα υπο-ορόσημο. Ουσιαστικά, κατασκευάζουμε τις περιοχές αναζήτησης στο πλαίσιο όπου θέλουμε να επιτελέσουμε την αναγνώριση. Ανιχνεύουμε τα χαρακτηριστικά στην πρώτη από αυτές, και στη συνέχεια αναζητούμε πεντάδες ομοεπίπεδων σημείων, που δεν είναι ανά τρία συνευθειακά και διέπονται από τοπολογική συνέπεια. Στις πεντάδες, που ικανοποιούν τα άνωθι κριτήρια, εφαρμόζουμε τον μηχανισμό επιλογής εξαιρετικών πεντάδων, προκειμένου να επιλέξουμε τις πιο *αντιπροσωπευτικές* πεντάδες. Με την παραπάνω διαδικασία, καταλήγουμε σε ένα σύνολο πεντάδων. Στη συνέχεια, προσπαθούμε να αντιστοιχήσουμε τα στοιχεία του συνόλου αυτού, στην βιβλιοθήκη ορόσημων, που κατασκευάστηκε στη φάση εκμάθησης. Έστω ότι συμβολίζουμε ως Q μια πεντάδα του πιο πάνω συνόλου. Η διαδικασία αναζήτησης της Q στη βιβλιοθήκη, συνίσταται στην κατασκευή της πλειάδας περιγραφής της, με βάση τις τιμές της αναλλοίωτης και της νόρμας της, και τη σειριακή σύγκρισή της με τις αντίστοιχες πλειάδες περιγραφής των υπό-ορόσημων της βιβλιοθήκης. Η σύγκριση πραγματοποιείται μέσω μιας μετρικής ομοιότητας διανυσματικών ποσοτήτων, η οποία συγκρίνει την πλειάδα περιγραφής T της Q με την πλειάδα T' κάθε υπο-ορόσημου της βιβλιοθήκης. Η μετρική αυτή ποσότητα είναι η *k-nearest-neighbour* [129], η οποία δίνεται από τον



Σχήμα 3.14: Διάγραμμα ροής του σταδίου αναγνώρισης ορόσημων

παρακάτω τύπο:

$$D(T, T') = \sqrt{(I - I')^2 + (N - N')^2} \tag{3.38}$$

Η πλειάδα T_k της βιβλιοθήκης, για την οποία η μετρική παράγει την μικρότερη τιμή[14], αποτελεί το υπο-ορόσημο στο οποίο αντιστοιχεί η Q. Με τον τρόπο αυτό, βεβαιωνόμαστε ότι η Q έχει παρόμοιες ιδιότητες με αυτές του υπο-ορόσημου της βιβλιοθήκης. Αυτό σημαίνει, πως η τιμή της αναλλοίωτής της και η τιμή του πίνακα συνδιασποράς της είναι *κοντά* στις αντίστοιχες τιμές του υπο-ορόσημου της βιβλιοθήκης, στο οποίο αντιστοιχεί. Ωστόσο, ο αντικειμενικός σκοπός μας είναι να ανακαλύψουμε ένα ορόσημο, που αποτελεί και το σημείο αναφοράς για το μηχανικό σύστημα. Η ανακάλυψη ενός ορόσημου, ανάγεται στην αναγνώριση όλων των υπο-ορόσημων του. Οπότε, από την πλειάδα περιγραφής T_k, αντλούμε την πληροφορία, που αναφέρεται στους *κατόχους* ορόσημα του συγκεκριμένου υπο-ορόσημου. Η πληροφορία αυτή, είναι πολύ σημαντική, γιατί οδηγεί τον μηχανισμό αναζήτησης υπο-ορόσημων σε πιθανά υπο-ορόσημα, που εμφανίζονται στο πλαίσιο αναγνώρισης. Ουσιαστικά, τα νέα υπο-ορόσημα θα πρέπει να είναι κάποια από τα υπο-ορόσημα των κατόχων. Οπότε, προσπαθούμε να ανακαλύψουμε τα εναπομείναντα υπο-ορόσημα του κάθε κατόχου στο πλαίσιο αναγνώρισης. Για να επιτευγχθεί η αντιστοίχιση αυτή, για κάθε ένα υπο-ορόσημο q από τα εναπομείναντα, επιτελείται η γνωστή διαδικασία, που έχουμε διατυπώσει, μόνο που τώρα οι περιοχές αναζήτησης καθορίζονται με βάση τις συντεταγμένες των πέντε σημείων

[14]η οποία είναι φραγμένη άνω από κάποιο κατώφλι

που απαρτίζουν το *q*. Με τον τρόπο αυτό *οδηγούμαστε* στην περιοχή αναζήτησης και δεν την κατασκευάζουμε εξ' αρχής, περιορίζοντας αισθητά το χώρο της, αλλά και ελαχιστοποιώντας σημαντικά το υπολογιστικό κόστος της αναζήτησης. Ουσιαστικά η νέα περιοχή αναζήτησης, κατασκευάζεται επαυξάνοντας το ελάχιστο περίβλημα της *q* κατά ένα ποσοστό σε όλες του τις διαστάσεις. Στην περιοχή αυτή, εφαρμόζεται η διαδικασία εύρεσης αντίστοιχου, και οι προκύπτουσες πεντάδες συγκρίνονται με το κάθε υπο-ορόσημο. Αν επιτευχθεί η αντιστοίχιση, τότε συνεχίζουμε στο επόμενο συστατικό υπο-ορόσημο του τρέχοντος κατόχου. Η διαδικασία αυτή, είναι ουσιαστικά μια model-driven προσέγγιση στην εύρεση υπο-ορόσημων μιας και η βιβλιοθήκη αποτελεί συνεχή επιτηρητή και συμμετέχει ενεργά στον καθορισμό της εύρεση υπο-ορόσημων. Αν ένας κάτοχος, απεικονιστεί στο πλαίσιο, τότε μεταβαίνουμε στην κατάσταση *επιτυχούς αναγνώρισης ενός ορόσημου* και ακόλουθες ενέργειες για το τρέχον σκοπό πλοήγησης μπορούν να πραγματοποιηθούν. Στην αντίθετη περίπτωση, μεταβαίνουμε στην κατάσταση της *ανίχνευσης* ενός νέου ορόσημου, οπότε η διαδικασία εκμάθησης εγείρεται και αναλαμβάνει την διαχείριση του και την ενημέρωση της βιβλιοθήκης σχετικά με αυτό. Με τον τρόπο αυτό, το σύστημα έχει την ευχέρεια να προσαρμόζεται σε καινούρια δεδομένα και αλλαγές του περιβάλλοντος του.

Κεφάλαιο 4

Πειραματικά αποτελέσματα

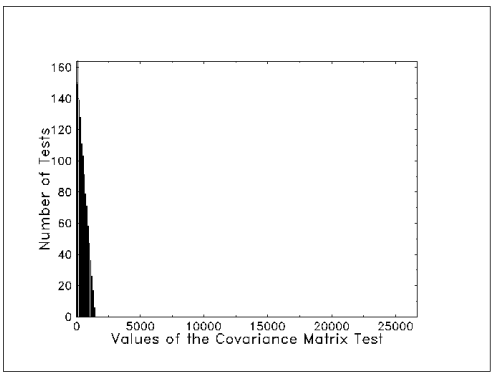
Για την ποσοτική αποτίμηση της μεθόδου ακολουθήσαμε μια δυϊκή προσέγγιση. Αρχικά, εκτελέσαμε ένα πλήθος πειραμάτων με *συνθετικούς κόσμους*, όπου η κάμερα καταγράφει σκηνές από διαφορές οπτικές γωνίες. Τα πειράματα αυτά χρησιμοποιήθηκαν στον έλεγχο της λειτουργικής ικανότητας της νόρμας του πίνακα συνδιασποράς, που αποτελεί την μετρική αντιστοίχισης στην προσέγγισή μας. Κατόπιν, επιτελέσαμε μια πλειάδα πειραμάτων σε στατικές ακολουθίες από εσωτερικούς διάδρομους κτιρίων, προκειμένου να αποτιμήσουμε την απόδοση και τη λειτουργικότητα της μεθόδου. Τα αποτελέσματα που παρείχθησαν συναινούν στην καλή απόδοση της μεθόδου και στην εφαρμοσιμότητά της σε εσωτερικούς διαδρόμους κτηρίων.

Στη συνέχεια περιγράφουμε τα πειράματα που επιτελέσαμε για την αξιολόγηση της μεθόδου. Ξεκινάμε με τα πειραματικά αποτελέσματα στους συνθετικούς κόσμους και κατόπιν μεταβαίνουμε στα αποτελέσματα που παρείχθησαν από την εφαρμογή της μεθόδου σε πραγματικά περιβάλλοντα.

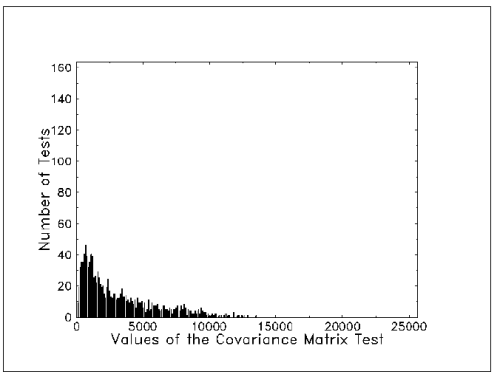
4.1 Συνθετικά αποτελέσματα

Μέσω των πειραμάτων στους συνθετικούς κόσμους προσπαθούμε να επιβεβαιώσουμε ότι οι ελάχιστες τιμές της νόρμας, που περιγράφεται στην εξίσωση 3.34, εμφανίζονται όταν μια πεντάδα συγκρίνεται με την πραγματικά αντίστοιχη της σε διαφορετικά πλαίσια. Στην προσπάθεια μας αυτή, επιβεβαιώνουμε ότι η τιμή της νόρμας είναι ευαίσθητη στην παρουσία θορύβου, γεγονός που είναι βασικό στη δική μας περίπτωση, μιας και η τιμή της αυξάνεται δραστικά, όταν οι υπο-σύγκριση πεντάδες δεν είναι πλήρως αντίστοιχες, αλλά η μία περιέχει σημεία που τα αντίστοιχα τους δεν ανήκουν στην άλλη πεντάδα. Η αντιστοίχιση των σημείων, δεν γίνεται διά χειρός, αλλά υπολογίζεται βάσει του είδους και των παραμέτρων κίνησης της κάμερας[1]. Τα πειράματα σε συνθετικά περιβάλλοντα, τα οποία περιγράφουμε, μας επιτρέπουν να ορίσουμε ένα άνω φράγμα των τιμών

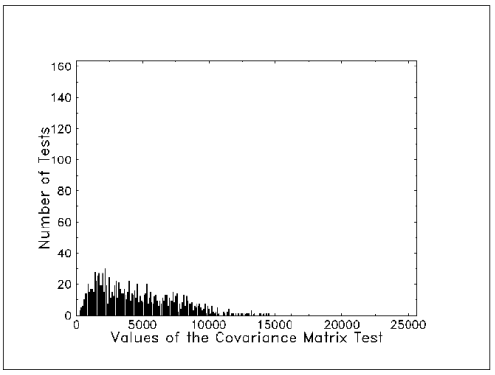
[1]Σε περιβάλλοντα κατασκευής συνθετικών κόσμων, υπάρχει αυτή η δυνατότητα



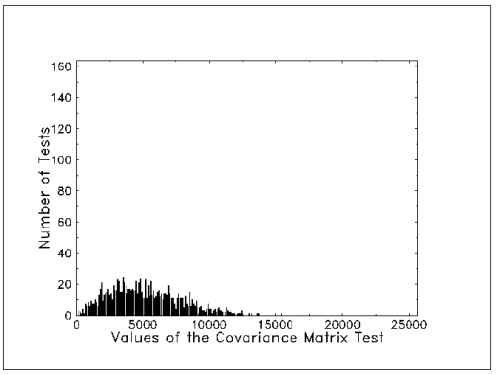
(α)



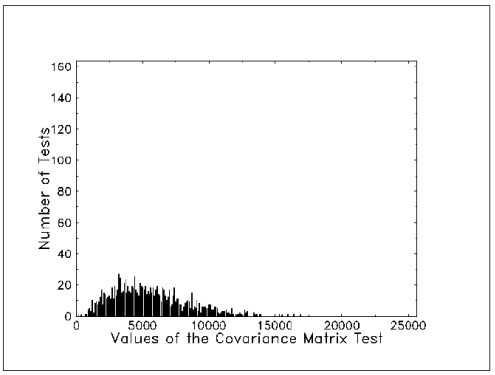
(β)



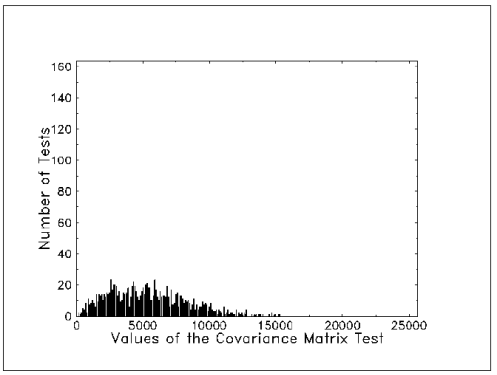
(γ)



(δ)



(ε)



(ζ)

Σχήμα 4.1: Αποτελέσματα της νόρμας του πίνακα συνδιασποράς

της νόρμας, μέχρι το οποίο η μετρική ποσότητα συμπεριφέρεται ικανοποιητικά όσο αναφορά τα αποτελέσματα που παράγει, τόσο στη διάρκεια της εξαγωγής όσο και της αναγνώρισης ορόσημων σε πραγματικά περιβάλλοντα.

Η κατασκευή των συνθετικών κόσμων επιτελέσθηκε μέσω της χρήσης του ray tracer POVRAY 3.0 σε μια πλατφόρμα Solaris 2.6. Ο συνθετικός ''κόσμος'', στον οποίο πειραματιστήκαμε, είναι μια ακολουθία 50 πλαισίων, που αποτελείται από πέντε επίπεδα[2]. Σε κάθε πλαίσιο, επιλέξαμε τυχαία 8 πεντάδες, δύο ανά επιφάνεια, εξαιρόντας το εσωτερικό επίπεδο, λόγω του ότι δεν είναι σταθερό. Για κάθε πιθανό ζεύγος πλαισίων, εφαρμόσαμε το κριτήριο της νόρμας σε όλες τις 8 αντίστοιχες πεντάδες. Έτσι το συνολικό πλήθος πειραμάτων ήταν $\binom{50}{2} \times 8 = 9800$. Τα αποτελέσματα ήταν στατιστικά αποδεκτά. Στην εικόνα 4.1 παρουσιάζουμε ιστογράμματα που απεικονίζουν τις τιμές του κριτηρίου σε σχέση με το πλήθος των πειραμάτων. Όπως μπορεί να γίνει αντιληπτό, στην εικόνα 4.1-α, όπου εφαρμόσαμε το κριτήριο της νόρμας στις πραγματικά αντίστοιχες πεντάδες, οι τιμές του ήταν καλά περιορισμένες σε ένα συγκεκριμένο πεδίο τιμών, του οποίου το άνω όριο μας παρέχει το επιθυμητό κατώφλι. Στην εικόνα 4.1-β, όπου μόνο τέσσερα σημεία της κάθε πεντάδας είναι σε πραγματική αντιστοίχιση, οι τιμές απλώνονται σε ένα πλουσιότερο πεδίο τιμών, και μόνο ένα μικρό δείγμα από αυτές ικανοποιεί το κατώφλι της εικόνας 4.1-α. Στις εικόνες 4.1-γ έως 4.1-ε, χαλαρώνουμε την πραγματική αντιστοίχιση κάθε σημείου προοδευτικά, με προφανή ολίσθηση των τιμών του κριτηρίου, μέχρι την περίπτωση της εικόνας 4.1-ζ, όπου δεν υπάρχει κανένα αντίστοιχο σημείο στις υπο-μελέτη πεντάδες. Με τον τρόπο αυτό, επιβεβαιώνεται ότι η τεχνική της νόρμας του πίνακα συνδιασποράς είναι ένα αποδεκτό ποσοτικό κριτήριο σύγκρισης της χωρικής κατανομής των σημείων, που ανήκουν σε πεντάδες καταγραφόμενες από διαφορετικές οπτικές γωνίες.

4.2 Αποτελέσματα σε πραγματικά περιβάλλοντα

Για την καλύτερη αξιολόγηση της προτεινόμενης μεθόδου, θα χρειαζόταν η εφαρμογή της σε ένα πραγματικό σύστημα. Ωστόσο, θεωρούμε ότι οι πειραματισμοί μας σε στατικές ακολουθίες από εσωτερικούς διάδρομους κτιρίων, παρέχουν σημαντική πληροφορία για την απόδοση και τα χαρακτηριστικά της μεθόδου. Περιγράφουμε μια πλειάδα πειραμάτων, στην προσπάθεια μας να αξιολογήσουμε τη μέθοδο. Στο πρώτο από αυτά, περιγράφουμε αναλυτικά τα βήματα της μεθόδου, προκειμένου να δώσουμε μια καθολική εικόνα στον αναγνώστη, σχετικά με την εκμάθηση και αναγνώριση ενός ορόσημου.

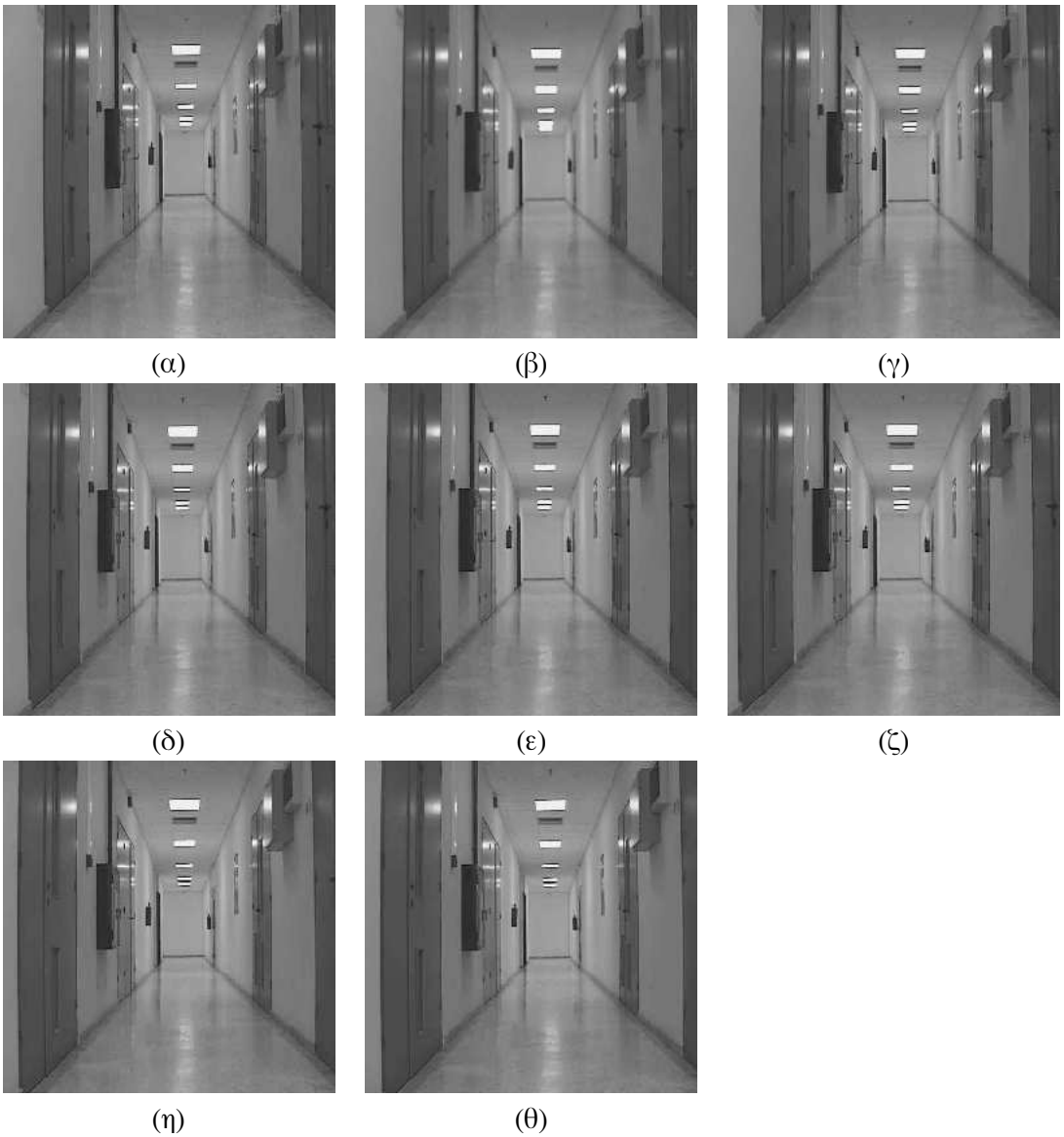
4.2.1 Παράδειγμα Ι

Αρχικά περιγράφουμε όλα τα βήματα του σταδίου εκμάθησης ενός ορόσημου και στη συνέχεια εξίσου αναλυτικά το στάδιο της αναγνώρισης.

[2]οι δύο τοίχοι, το πάτωμα, το ταβάνι, και το εσωτερικό επίπεδο

4.2.1.1 Περίοδος εκμάθησης

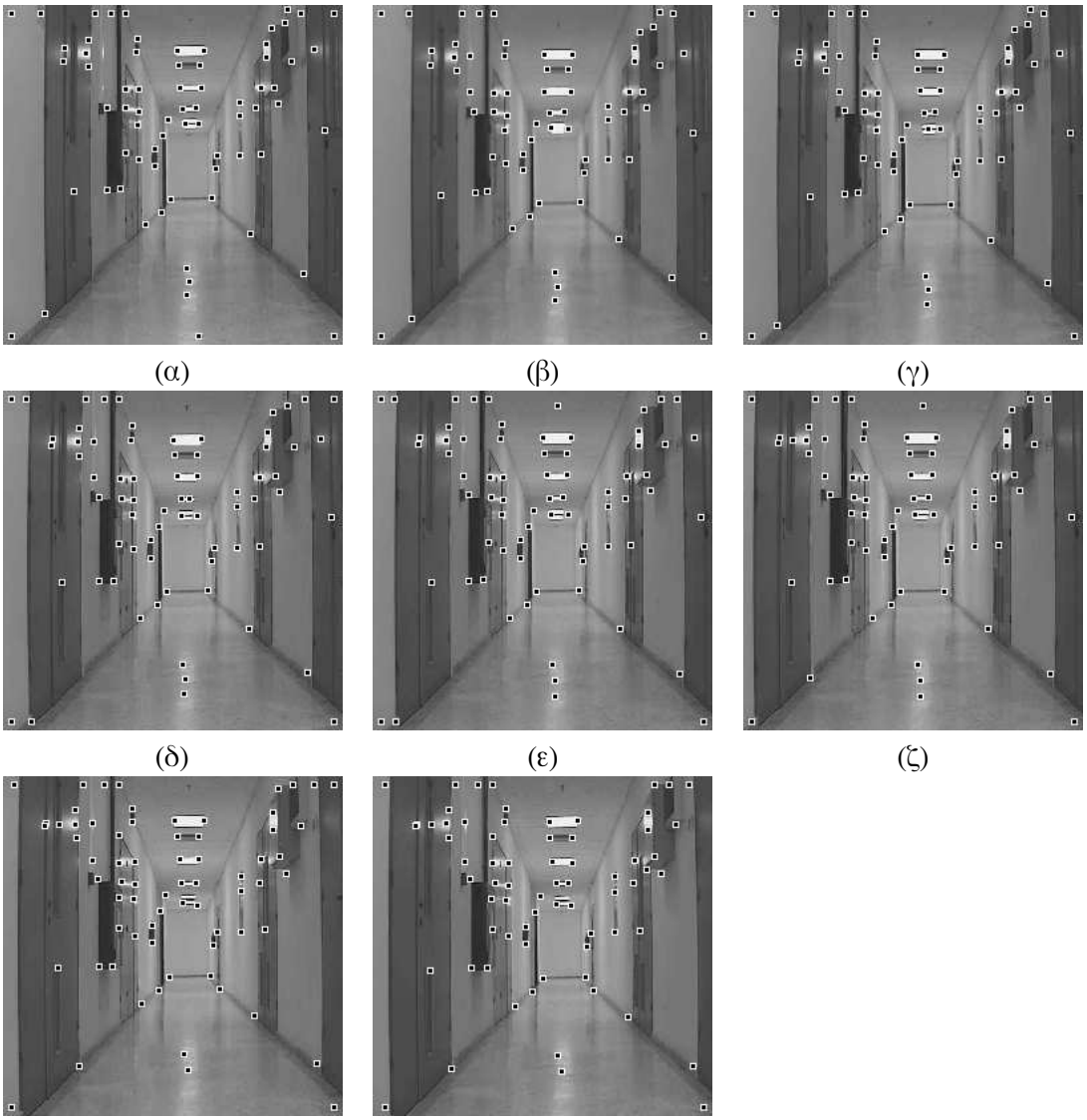
Ξεκινάμε το παράδειγμά μας, παραθέτοντας την ακολουθία εισόδου, στην εικόνα 4.2, από την οποία θα προσπαθήσουμε να ‘‘μάθουμε’’ ένα ορόσημο. Το επόμενο βήμα, είναι



Σχήμα 4.2: Η ακολουθία εισόδου στο πρώτο παράδειγμα

η εφαρμογή του ανιχνευτή χαρακτηριστικών, στην ακολουθία εισόδου, το αποτέλεσμα της οποίας εμφανίζεται στην εικόνα 4.3. Κατόπιν, εφαρμόζοντας το μηχανισμό εστίασης της προσοχής[3] στην πρώτη εικόνα, καταλήγουμε στην περιοχή αναζήτησης, που απεικονίζεται στην εικόνα 4.4, με τα εσωτερικά της χαρακτηριστικά τεθειμένα. Ανιχνεύουμε τις αντίστοιχες πεντάδες, τις ταξινομούμε με βάση τη σωρευτική τους απόσταση και παραθέτουμε τις δώδεκα πιο εξαιρετικές πεντάδες, στην εικόνα 4.5 με τις αντίστοιχες

[3]συνδιάζοντας ουσιαστικά την πληροφορία του χάρτη προεξοχής και του ανιχνευτή χαρακτηριστικών



Σχήμα 4.3: Ανιχνευθέντα χαρακτηριστικά στην αρχική ακολουθία

Σχήμα 4.4: Περιοχή αναζήτησης ορόσημων

(1) (2) (3)

(4) (5) (6)

(7) (8) (9)

(10) (11) (12)

Σχήμα 4.5: Οι δώδεκα πιο εξαιρετικές πεντάδες

τιμές της σωρευτικής απόστασης των στον πίνακα 4.1. Όπως γίνεται άμεσα αντιληπτό,

Αριθμός πεντάδας	**Σωρευτική απόσταση**
1	24.022816
2	12.991822
3	11.542354
4	11.496682
5	10.915649
6	10.679019
7	10.182075
8	10.161320
9	9.959856
10	9.840040
11	9.714285
12	9.327250

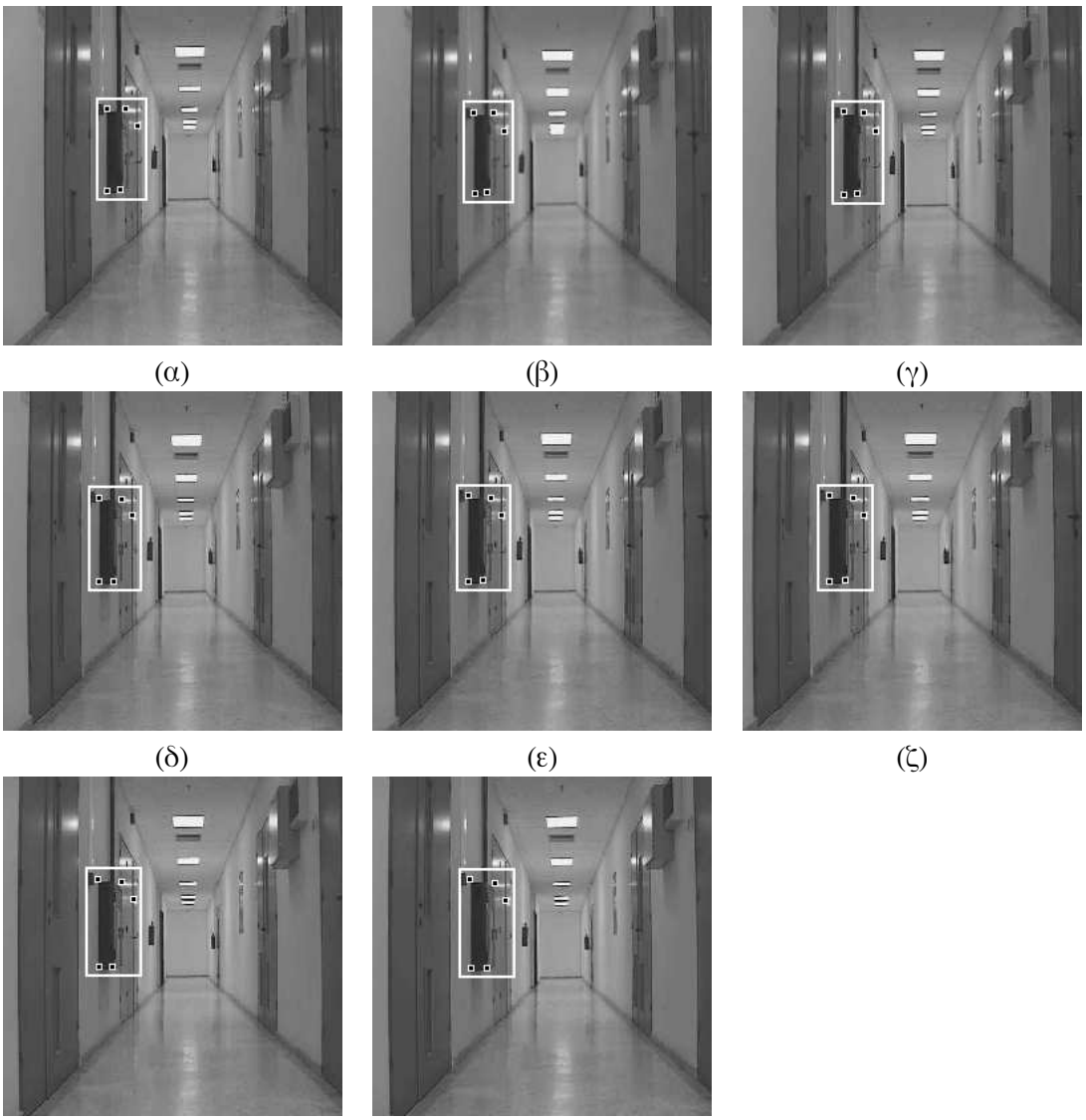
Πίνακας 4.1: Οι τιμές της σωρευτικής απόστασης για τις δώδεκα πιο εξαιρετικές πεντάδες

η πρώτη πεντάδα *διαφέρει* αισθητά από τις υπόλοιπες, οπότε και προσωρινά την επιλέγουμε. Η πεντάδα αυτή, για λόγους αυτονομίας[4] επανεμφανίζεται στην εικόνα 4.6. Το επόμενο βήμα είναι να ελέγξουμε αν διατηρείται στα επόμενα πλαίσια, προκειμένου

Σχήμα 4.6: Η πιο εξαιρετική πεντάδα

να υπολογίσουμε το χρόνο ζωής της και να παρατηρήσουμε τη συνέπεια της στο χώρο και στο χρόνο. Η εφαρμογή της διαδικασίας υπολογισμού του χρόνου ζωής, καταλήγει στο συμπέρασμα ότι ο χρόνος ζωής για τη συγκεκριμένη πεντάδα είναι οκτώ, μια και διατηρείται σε οκτώ συνολικά διαδοχικά πλαίσια, όπως παρατηρούμε στην εικόνα 4.7. Στον πίνακα 4.2 εμφανίζονται οι τιμές της αναλλοίωτης, της νόρμας του πίνακα συνδιασποράς της πιο εξαιρετικής πεντάδας και της νόρμας διαφοράς στα διαδοχικά πλαίσια όπου εμφανίζεται. Μελετώντας τον πίνακα, παρατηρούμε την διατήρηση των τιμών της αναλλοίωτης, όπως είχαμε τονίσει για τις ομοεπίπεδες πεντάδες, της νόρμας

[4]και για να δικαιολογήσει το όνομά της



Σχήμα 4.7: Διατήρηση της πεντάδας στην αρχική ακολουθία.

Αριθμός πλαισίου	Τιμή αναλλοίωτης	Τιμή νόρμας πεντάδας	Τιμή νόρμας διαφοράς
α	11.744196	1101.049927	13.129150
β	11.755899	1087.920776	13.129150
γ	11.737316	1121.781250	33.860474
δ	11.666253	1168.986816	47.20556
ε	11.652237	1191.930176	22.94335
ζ	11.634658	1192.859253	0.92907
η	11.642653	1227.172729	34.313477
θ	11.593908	1237.499634	10.32690

Πίνακας 4.2: Οι τιμές της αναλλοίωτης και της νόρμας για κάθε αντίστοιχη πεντάδα.

και της νόρμας διαφοράς, που υπακούει με ευλάβεια τα ιστογράμματα της εικόνας 4.1.

Περατώνοντας τη διαδικασία εκμάθησης, συνθέτουμε το ορόσημο, που αποτελείται από το υπο-ορόσημο[5] της εικόνας 4.8 και το αποθηκεύουμε με την μορφή της πλειάδας περιγραφής του πίνακα 4.3.

Σχήμα 4.8: Το υπο-ορόσημο που αποτελεί το ορόσημο της ακολουθίας

Πλειάδα περιγραφής ''Εκμάθηση 1.1''							
I	N	L_0	P_1	P_2	P_3	P_4	P_5
11.666	1168.986	0	(81,89)	(80,72)	(143,83)	(93,97)	(143,72)

Πίνακας 4.3: Η πλειάδα περιγραφής του υπο-ορόσημου στη βιβλιοθήκη

4.2.1.2 Περίοδος αναγνώρισης

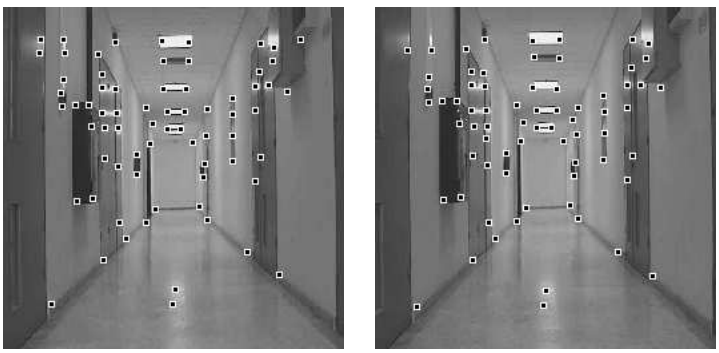
Στη συνέχεια αναλύουμε τα βήματα της μεθόδου στο στάδιο της αναγνώρισης. Αρχικά παραθέτουμε, στην εικόνα 4.9, τα δύο διαδοχικά πλαίσια, στα οποία θα αναζητήσουμε, αν υπάρχει, κάποιο γνωστό ή άγνωστο ορόσημο. Όμοια με τη φάση εκμάθησης,

(α) (β)

Σχήμα 4.9: Ζεύγος εικόνων εισόδου στη διαδικασία αναγνώρισης

[5]εμφανίζεται στο μεσαίο πλαίσιο της ακολουθίας εισόδου, αφού μας ενδιαφέρει η σχετική θέση του

εφαρμόζουμε τον ανιχνευτή χαρακτηριστικών στο αρχικό ζεύγος, παρατηρώντας τα αποτελέσματα της εικόνας 4.10. Κατόπιν, εφαρμόζουμε τους μηχανισμούς εστίασης της



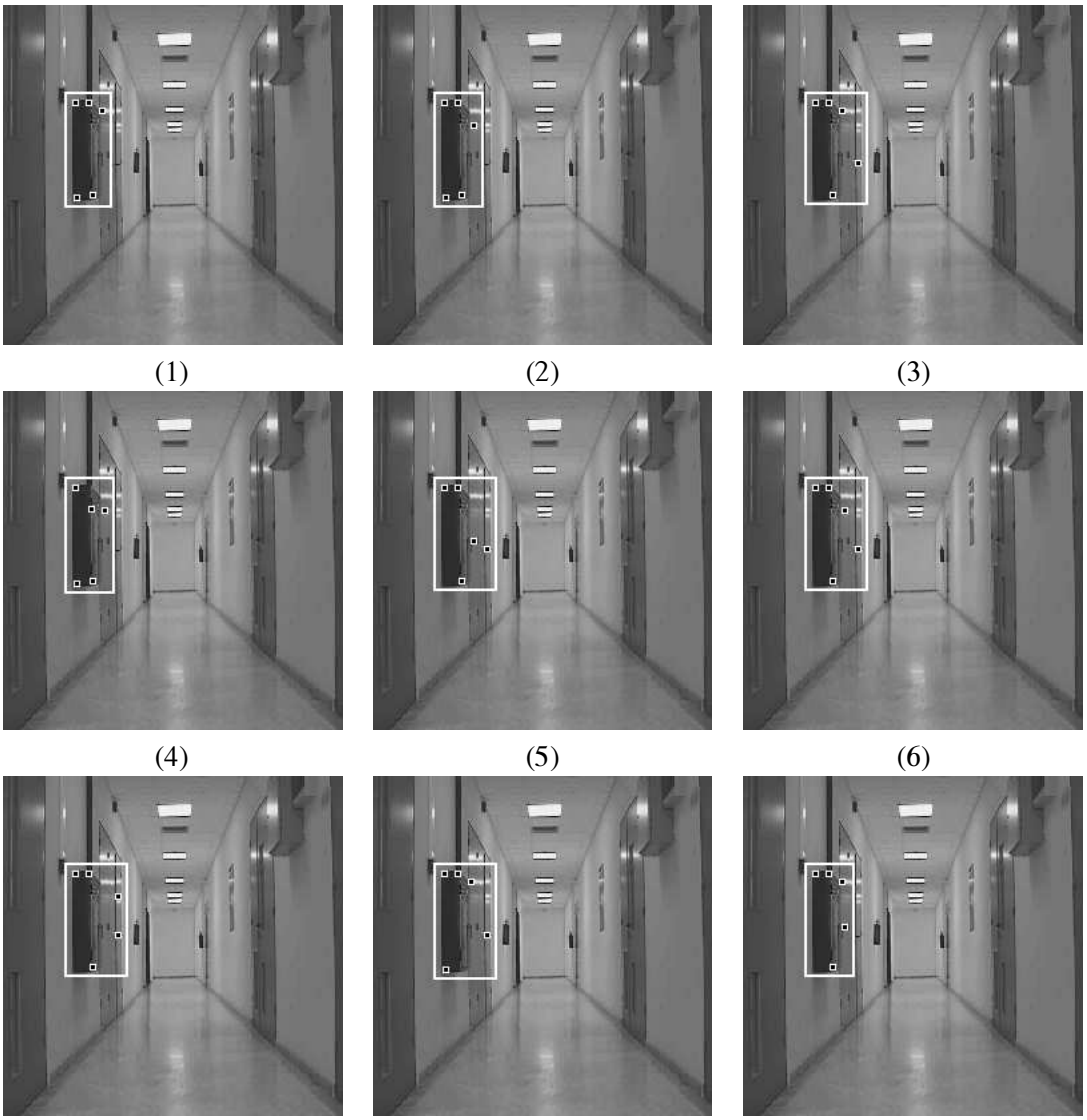
(α) (β)

Σχήμα 4.10: Ανιχνευθέντα χαρακτηριστικά στο ζεύγος εισόδου

προσοχής στην πρώτη εικόνα, προκειμένου να ανακαλύψουμε τις περιοχές αναζήτησης των ορόσημων. Συνδυάζοντας την πληροφορία του χάρτη προεξοχής και του ανιχνευτή χαρακτηριστικών, καταλήγουμε στην περιοχή, με τα εσωτερικά της χαρακτηριστικά τεθειμένα, που απεικονίζεται στην εικόνα 4.11. Έπειτα, εφαρμόζοντας τους μηχανισμούς

Σχήμα 4.11: Περιοχή αναζήτησης ορόσημων

εύρεσης αντίστοιχων πεντάδων καταλήγουμε σε ένα σύνολο από ζεύγη πεντάδων. Στο σύνολο αυτό, εφαρμόζουμε την τεχνική του εκτιμητή εξαιρετικών πεντάδων. Ταξινομώντας τις πεντάδες, με βάση τη σωρευτική απόστασή τους, καταλήγουμε σε μια ιεραρχία, από την οποία παραθέτουμε τα πρώτα εννέα στοιχεία της στην εικόνα 4.12, για να παρουσιάσουμε τις πιθανές επιλογές εξαιρετικών πεντάδων. Οι αντίστοιχες τιμές της σωρευτικής απόστασης, για κάθε μία πεντάδα εκ των εννέα, παρατίθενται στον πίνακα 4.4. Όπως γίνεται άμεσα αντιληπτό, οι δύο πρώτες πεντάδες *διαφέρουν* αισθητά από τις προηγούμενες. Με βάση την παρατήρηση αυτή, επιλέγουμε μόνο τις δύο πρώτες πεντάδες, ως υποψήφιες να αποτελέσουν υπο-ορόσημα. Οι πεντάδες αυτές επανεμφανίζονται στην εικόνα 4.13. Στη συνέχεια, όπως έχουμε αναφέρει, κατασκευάζουμε τις



Σχήμα 4.12: Οι εννέα πιο εξαιρετικές πεντάδες

Αριθμός πεντάδας	**Σωρευτική απόσταση**
1	212.837860
2	184.536346
3	146.331741
4	140.293228
5	138.958206
6	134.524780
7	128.650085
8	125.757790
9	123.613831

Πίνακας 4.4: Οι τιμές της σωρευτικής απόστασης για τις εννέα πιο εξαιρετικές πεντάδες

Σχήμα 4.13: Οι δύο επιλεχθείσες, πιο εξαιρετικές πεντάδες

πλειάδες περιγραφής για κάθε ένα από τα υπο-ορόσημα, προκειμένου να ελέγξουμε αν υπάρχουν ή όχι στη βιβλιοθήκη. Οι πλειάδες αυτές φαίνονται στους πίνακες 4.5 και 4.6. Έπειτα, συγκρίνουμε την πλειάδα περιγραφής ''Αναγνώριση 1.1'' με την

Πλειάδα περιγραφής ''Αναγνώριση 1.1''							
I	N	L_0	P_1	P_2	P_3	P_4	P_5
11.584373	1484.900024	0	(79,74)	(73,54)	(145,55)	(73,64)	(143,67)

Πίνακας 4.5: Η πλειάδα περιγραφής της πρώτης πεντάδας

Πλειάδα περιγραφής ''Αναγνώριση 1.2''							
I	N	L_0	P_1	P_2	P_3	P_4	P_5
11.316676	1355.899902	0	(73,54)	(145,55)	(73,64)	(143,67)	(90,76)

Πίνακας 4.6: Η πλειάδα περιγραφής της δεύτερης πεντάδας

πλειάδα ''Εκμάθηση 1.1'' της βιβλιοθήκης, για να δούμε αν το μηχανικό σύστημα έχει εκπαιδευτεί με αυτή. Η σύγκριση επιτελείται, όπως έχουμε αναφέρει, με βάση το κριτήριο του k-nearest-neighbor, το οποίο για την συγκεκριμένη περίπτωση δίνει την τιμή, 315.913208, η οποία είναι μέσα στα αποδεκτά μας όρια. Άρα, σε πρώτο βήμα, αντιστοιχούμε την πεντάδα ''Αναγνώριση 1.1'', με την πλειάδα ''Εκμάθηση 1.1'', η οποία μιας και αποτελεί μοναδικό συστατικό ενός ορόσημου, οδηγεί στην αναγνώριση ενός ορόσημου. Απομένει τώρα να διαπιστώσουμε αν η πλειάδα ''Αναγνώριση 1.2'', μπορεί να αντιστοιχηθεί στην βιβλιοθήκη. Η τιμή του k-nearest-neighbor στη σύγκριση της πλειάδας ''Αναγνώριση 1.2'' με την πλειάδα ''Εκμάθηση 1.1'', είναι 186.913405, η οποία αποτελεί ένδειξη αντιστοίχισης της πλειάδας ''Αναγνώριση 1.2'' με την ''Εκμάθηση 1.1''. Μιας και η τιμή του k-nearest-neighbor στη δεύτερη αντιστοίχιση, είναι μικρότερη της πρώτης, αποφασίζουμε να επιλέξουμε ως καλύτερη πλειάδα αντιστοίχισης

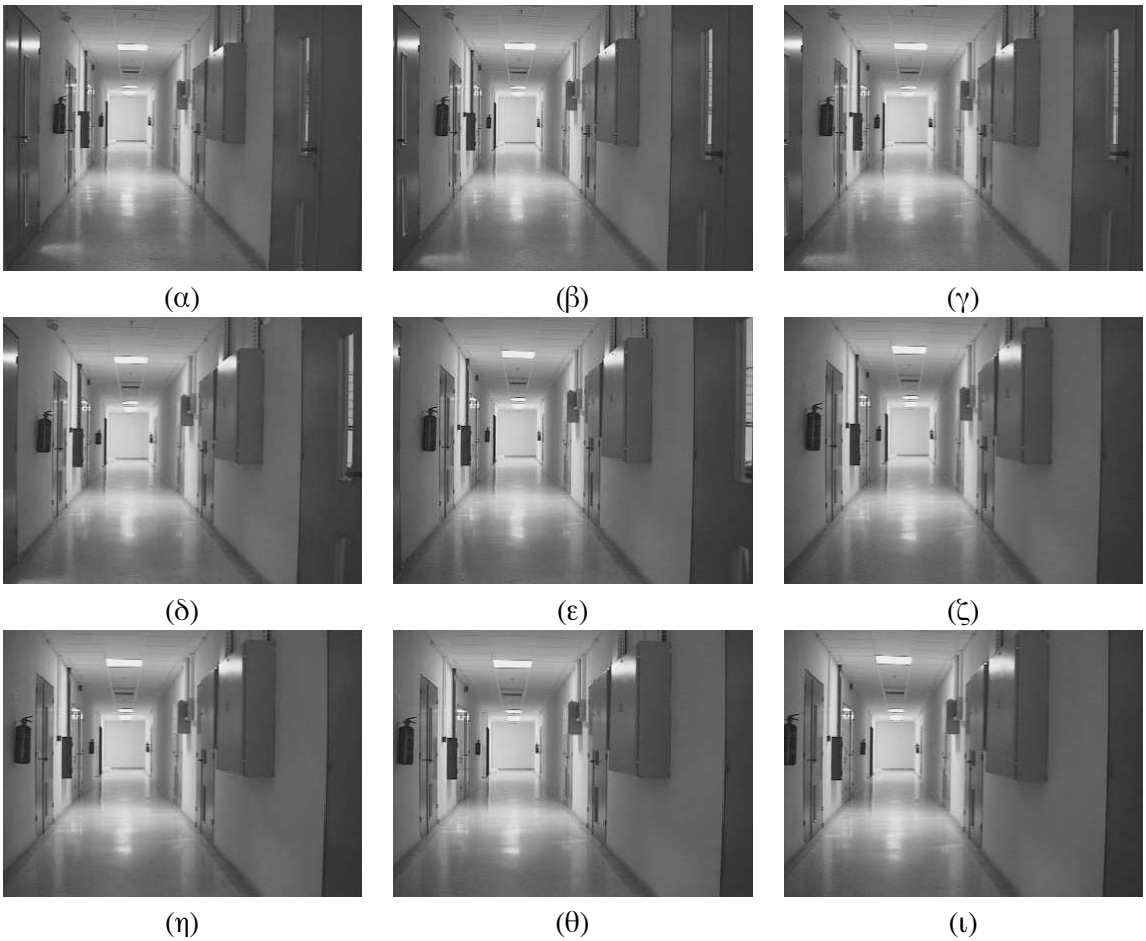
της ‘‘Εκμάθηση 1.1’’, την ‘‘Αναγνώριση 1.2’’, απορρίπτοντας την πλειάδα ‘‘Αναγνώριση 1.1’’. Εδώ τελειώνει η περιήγηση σε μία περίπτωση που η μέθοδος αντιμετώπισε.

4.2.2 Παράδειγμα ΙΙ

Παρουσιάσαμε το πρώτο παράδειγμα αναλυτικά, για να περιγράψουμε αναλυτικά την εξέλιξη της μεθόδου σε μια δεδομένη ακολουθία πλαισίων. Στη συνέχεια, παραθέτουμε ένα δεύτερο, πιο συνοπτικό αυτή τη φορά, παράδειγμα, προκειμένου να παρουσιάσουμε την δυνατότητα της μεθόδου στην αναγνώριση υπο-ορόσημων, ιδωμένα, κάτω από διαφορετικές οπτικές γωνίες.

4.2.2.1 Περίοδος εκμάθησης

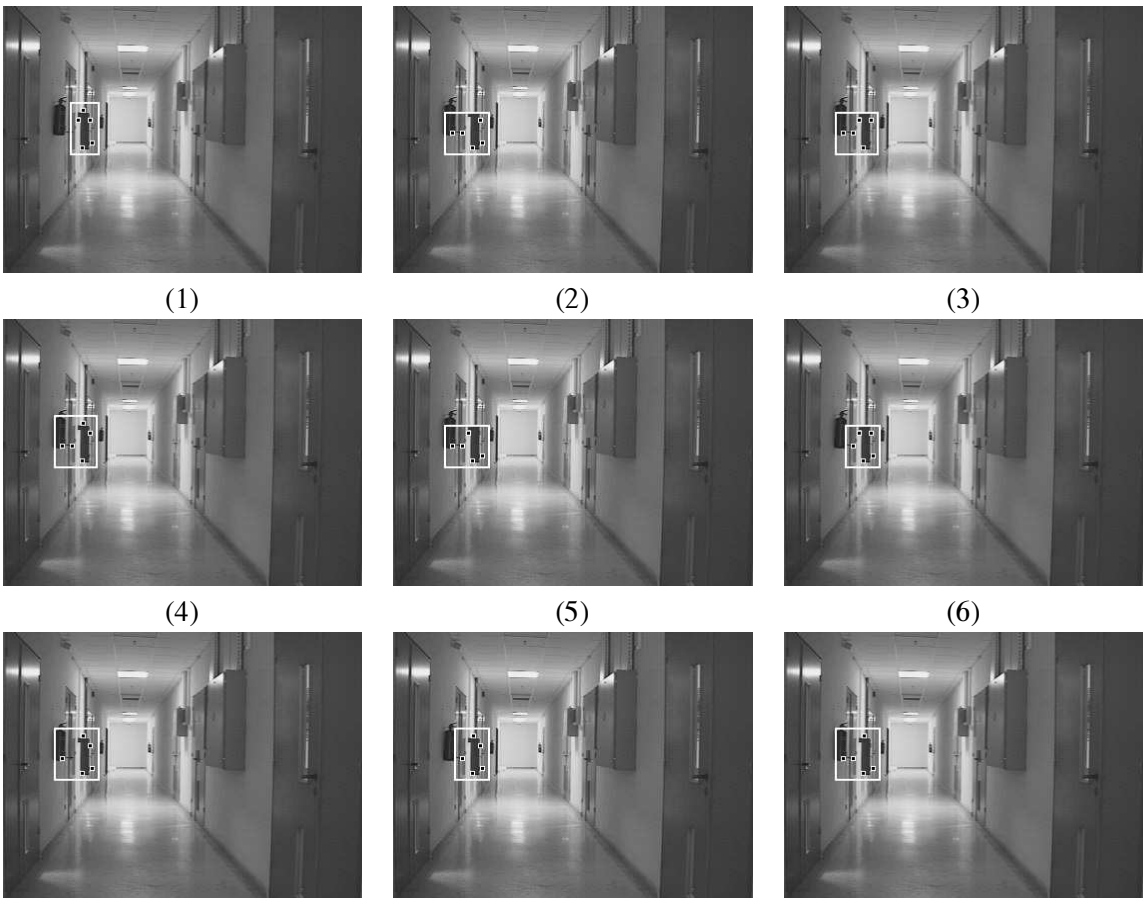
Ξεκινάμε το παράδειγμά μας, παραθέτοντας τα δεδομένα εισόδου του προγράμματος. Η ακολουθία των πλαισίων στην οποία θα προσπαθήσουμε να ανακαλύψουμε το ορόσημο, απεικονίζεται στην εικόνα 4.14. Στις περιοχές αναζήτησης, που προκύπτουν,



Σχήμα 4.14: Ακολουθία εικόνων στο δεύτερο παράδειγμα

ανιχνεύουμε αντίστοιχες πεντάδες, και τις ταξινομούμε βάσει της σωρευτικής απόστασής

των. Από το προκύπτον σύνολο, παραθέτουμε τα πρώτα εννέα στοιχεία στην εικόνα 4.15, και τις αντίστοιχες τιμές της σωρευτικής απόστασής των στον πίνακα 4.7. Όπως



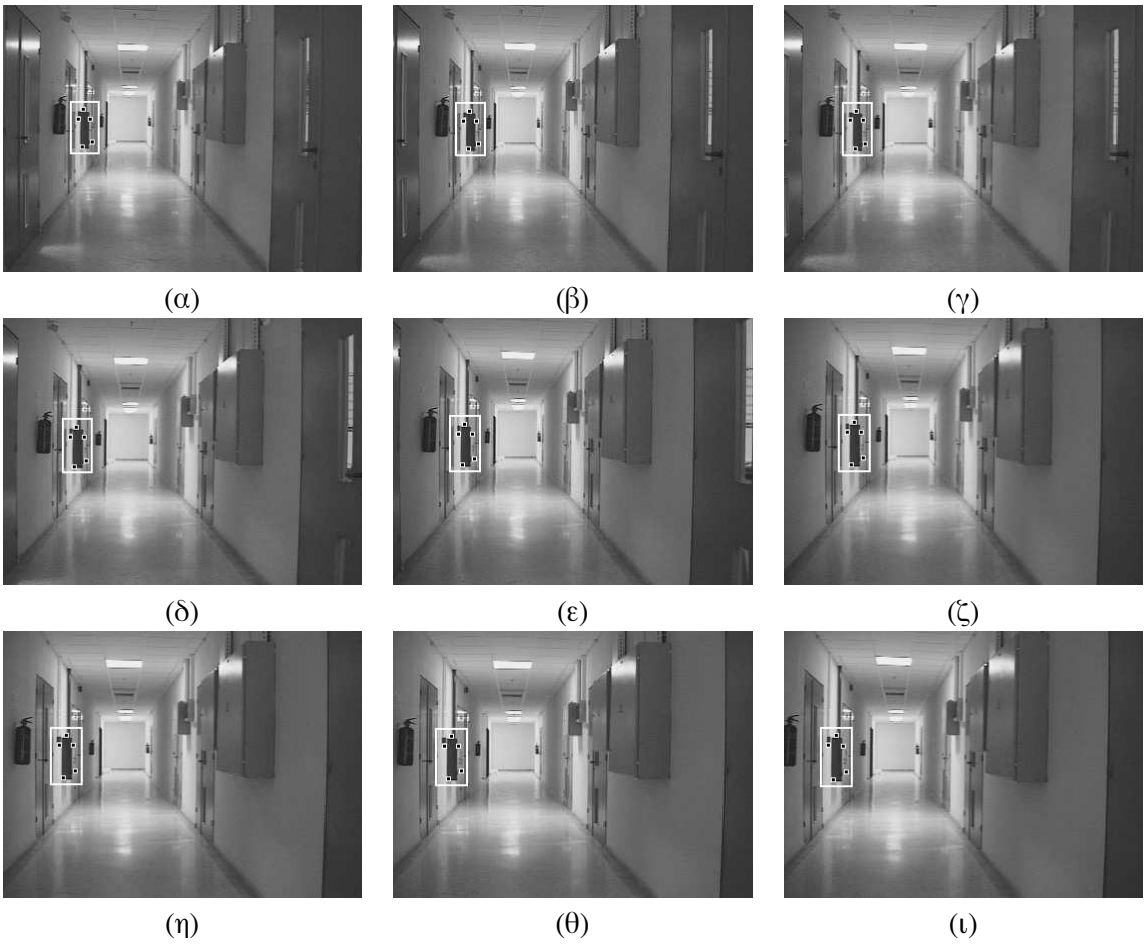
Σχήμα 4.15: Οι εννέα πιο εξαιρετικές πεντάδες

Αριθμός πεντάδας	**Σωρευτική απόσταση**
1	2.337794
2	1.940947
3	1.883315
4	1.580842
5	1.576593
6	1.568442
7	1.468993
8	1.400645
9	1.334796

Πίνακας 4.7: Οι τιμές της σωρευτικής απόστασης για τις εννέα πιο εξαιρετικές πεντάδες

γίνεται άμεσα αντιληπτό, η πρώτη πεντάδα *διαφέρει* αισθητά από τις προηγούμενες. Στη συνέχεια, υπολογίζουμε το χρόνο ζωής της πρώτης μόνο πεντάδα, που όπως

φαίνεται στην εικόνα 4.16, είναι εννέα. Στον πίνακα 4.8 εμφανίζονται οι τιμές της



(α) (β) (γ)

(δ) (ε) (ζ)

(η) (θ) (ι)

Σχήμα 4.16: Διατήρηση της πεντάδας στην αρχική ακολουθία.

αναλλοίωτης, της νόρμας του πίνακα συνδιασποράς της απόμακρης πεντάδας και της νόρμας διαφοράς στα διαδοχικά πλαίσια όπου εμφανίζεται. Από τον πίνακα 4.8,

Αριθμός πλαισίου	**Τιμή αναλλοίωτης**	**Τιμή νόρμας πεντάδας**	**Τιμή νόρμας διαφοράς**
α	12.985630	293.449982	13.492401
β	12.992584	306.942383	13.492401
γ	12.993261	305.767975	1.174408
δ	12.972743	322.271118	16.503143
ε	12.991507	343.523315	21.252197
ζ	12.996494	365.902679	22.379364
η	12.996074	367.340607	1.437927
θ	12.996093	362.720337	4.620270
ι	12.995901	395.246338	32.526001

Πίνακας 4.8: Τιμές αναλλοίωτης, νόρμας και νόρμας διαφοράς για τις αντίστοιχες πεντάδες

φαίνεται, και σε αυτό το παράδειγμα, η διατήρηση των τιμών της αναλλοίωτης, της νόρμας και της διαφοράς της νόρμας, η οποία είναι εντός των επιθυμητών ορίων.

Περατώνοντας τη διαδικασία εκμάθησης, συνθέτουμε και αποθηκεύουμε το ορόσημο, το οποίο, αποτελείται από ένα υπο-ορόσημο, που απεικονίζεται στην εικόνα 4.17. Η

Σχήμα 4.17: Το υπο-ορόσημο που αποτελεί το ορόσημο της ακολουθίας

πλειάδα περιγραφής του υπο-ορόσημου παρουσιάζεται στον πίνακα 4.9.

Πλειάδα περιγραφής ‘‘Εκμάθηση 2.1’’							
I	N	L_0	P_1	P_2	P_3	P_4	P_5
12.991507	343.523315	1	(107,71)	(148,69)	(117,65)	(117,79)	(142,81)

Πίνακας 4.9: Η πλειάδα περιγραφής του υπο-ορόσημου στη βιβλιοθήκη

4.2.2.2 Περίοδος αναγνώρισης

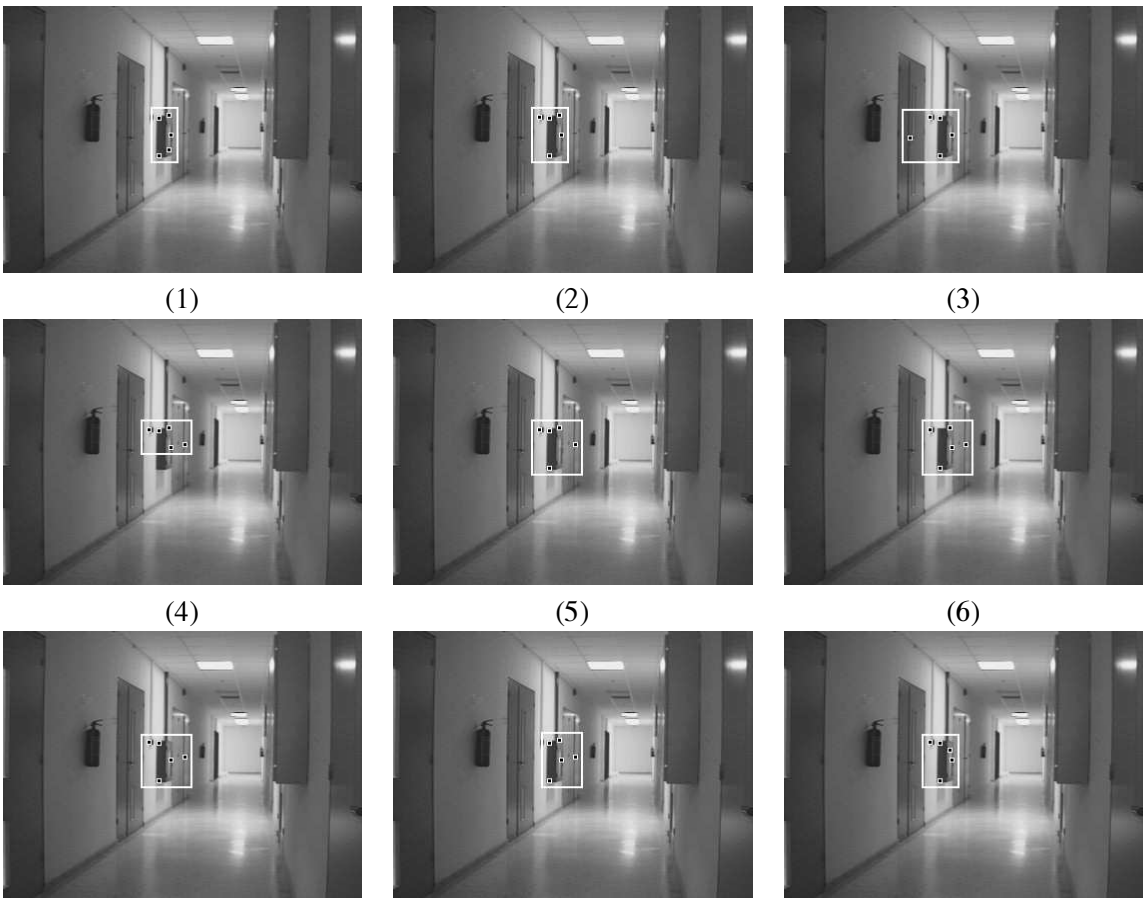
Στη συνέχεια αναλύουμε τα βήματα της μεθόδου στο στάδιο της αναγνώρισης. Η ακολουθία στην οποία θα προσπαθήσουμε να ανακαλύψουμε το ορόσημο, απεικονίζεται στην εικόνα 4.18. Κατόπιν, στην περιοχή αναζήτησης, ανιχνεύουμε τις αντίστοιχες

(α) (β)

Σχήμα 4.18: Ακολουθία εισόδου στο δεύτερο παράδειγμα

πεντάδες και από αυτές, με χρήσης του εκτιμητή των στατιστικά εξαιρετικών πεντάδων,

παραθέτουμε τις εννέα πιο εξαιρετικές πεντάδες στην εικόνα 4.19 και τις σωρευτικές αποστάσεις των στον πίνακα 4.10. Όπως γίνεται άμεσα αντιληπτό, μόνο η πρώτη



Σχήμα 4.19: Οι εννέα πιο εξαιρετικές πεντάδες

Αριθμός πεντάδας	**Σωρευτική απόσταση**
1	1.454170
2	1.145593
3	1.001361
4	0.978853
5	0.884949
6	0.852614
7	0.826797
8	0.717207
9	0.693897

Πίνακας 4.10: Οι σωρευτικές αποστάσεις των εννέα πιο εξαιρετικών πεντάδων

πεντάδα *διαφέρει* αισθητά από τις προηγούμενες. Με βάση την παρατήρηση αυτή, επιλέγουμε μόνο την πρώτη πεντάδα, ως υποψήφια να αποτελέσει υπο-ορόσημο. Η πιο εξαιρετική πεντάδα εμφανίζεται στην εικόνα 4.20. Στη συνέχεια, κατασκευάζουμε

Σχήμα 4.20: Η επιλεχθείσα, πιο εξαιρετική πεντάδα

την πλειάδα περιγραφής για το υπο-ορόσημο, προκειμένου να το αναζητήσουμε στη βιβλιοθήκη. Η πλειάδα αυτή φαίνεται στον πίνακα 4.11. Κατόπιν, συγκρίνουμε

Πλειάδα περιγραφής ‘‘Αναγνώριση 2.1’’							
I	N	L_0	P_1	P_2	P_3	P_4	P_5
12.842331	349.600006	1	(113,56)	(151,156)	(110,166)	(130,168)	(145,166)

Πίνακας 4.11: Η πλειάδα περιγραφής της πρώτης πεντάδας

την πλειάδα περιγραφής ‘‘Αναγνώριση 2.1’’ με την πλειάδα ‘‘Εκμάθηση 2.1’’ της βιβλιοθήκης, για να δούμε αν το μηχανικό σύστημα έχει εκπαιδευτεί με αυτή. Η σύγκριση επιτελείται, όπως έχουμε αναφέρει, με βάση το κριτήριο του k-nearest-neighbor, το οποίο για την συγκεκριμένη περίπτωση δίνει την τιμή, 6.078522, η οποία είναι μέσα στα αποδεκτά μας όρια. Άρα αντιστοιχούμε την πλειάδα ‘‘Αναγνώριση 2.1’’, με την πλειάδα ‘‘Εκμάθηση 2.1’’, η οποία μιας και αποτελεί μοναδικό συστατικό ενός ορόσημου, οδηγεί στην αναγνώριση ενός ορόσημου. Εδώ, όπως αναφέραμε και στην αρχή του παραδείγματος, είναι εμφανής η ισχύς της μεθόδου στη δυνατότητα αναγνώρισης ορόσημων, ιδωμένων από διαφορετικές οπτικές γωνίες, αποδεικνύοντας την ευρωστία και αποτελεσματικότητα της σε τέτοιου είδους καταστάσεις.

4.2.3 Παράδειγμα ΙΙΙ

Παρουσιάζουμε, ένα πιο συνοπτικό παράδειγμα, όπου ένα ορόσημο αποτελείται από περισσότερα του ενός υπο-ορόσημα.

4.2.3.1 Περίοδος εκμάθησης

Όπως και στα προηγούμενα παραδείγματα, ξεκινούμε με την περιγραφή των δεδομένων εισόδου. Η ακολουθία των πλαισίων στην οποία θα προσπαθήσουμε να ανακαλύψουμε

το ορόσημο, απεικονίζεται στην εικόνα 4.21. Κατόπιν, εφαρμόζουμε τους μηχανισμούς

(α) (β) (γ)

(δ) (ε) (ζ)

Σχήμα 4.21: Ακολουθία εικόνων στο τρίτο παράδειγμα

εστίασης της προσοχής στην πρώτη εικόνα και τελικά επιλέγουμε τις δύο πιο εξαιρετικές πεντάδες που απεικονίζονται στην εικόνα 4.22. Στη συνέχεια, υπολογίζουμε το χρόνο

(1) (2)

Σχήμα 4.22: Οι δύο πιο εξαιρετικές πεντάδες που επιλέγουμε

ζωής των επιλεχθέντων πεντάδων ο οποίος, όπως παρατηρούμε στις εικόνες 4.23 και 4.24, είναι έξι, μιας και διατηρούνται σε έξι διαδοχικά πλαίσια. Στους πίνακες 4.12

(α) (β) (γ)

(δ) (ε) (ζ)

Σχήμα 4.23: Διατήρηση της πρώτης πεντάδας στην αρχική ακολουθία.

(α) (β) (γ)

(δ) (ε) (ζ)

Σχήμα 4.24: Διατήρηση της δεύτερης πεντάδας στην αρχική ακολουθία.

και 4.13 εμφανίζονται οι τιμές της αναλλοίωτης, της νόρμας του πίνακα συνδιασποράς και της νόρμας διαφοράς για την πρώτη και τη δεύτερη πεντάδα αντίστοιχα, σε κάθε πλαίσιο της ακολουθίας. Η διατήρηση των τιμών της αναλλοίωτης, της νόρμας και της

Αριθμός πλαισίου	Τιμή αναλλοίωτης	Τιμή νόρμας πεντάδας	Τιμή νόρμας διαφοράς
α	11.912922	320.250000	0.989624
β	11.915797	319.260376	0.989624
γ	12.018351	328.688843	9.428467
δ	12.148144	332.335510	3.646667
ε	12.145685	353.305206	20.969696
ζ	12.200800	357.122833	3.817627

Πίνακας 4.12: Αναλλοίωτες, νόρμες και νόρμες διαφοράς για την πρώτη πεντάδα

Αριθμός πλαισίου	Τιμή αναλλοίωτης	Τιμή νόρμας πεντάδας	Τιμή νόρμας διαφοράς
α	12.716562	398.200012	1.547089
β	12.725794	396.652924	1.547089
γ	12.769866	413.957825	17.304901
δ	12.765260	425.971741	12.013916
ε	12.517303	443.469513	17.497772
ζ	12.558254	447.084290	3.614777

Πίνακας 4.13: Αναλλοίωτες, νόρμες και νόρμες διαφοράς για την δεύτερη πεντάδα

διαφοράς της νόρμας, που είναι εντός των επιθυμητών ορίων, είναι και εδώ εμφανής. Περατώνοντας τη διαδικασία εκμάθησης, συνθέτουμε και αποθηκεύουμε το ορόσημο, το οποίο, αποτελείται από τα δύο υπο-ορόσημα της εικόνας 4.25. Οι πλειάδες περιγραφής

Σχήμα 4.25: Τα δύο υπο-ορόσημα, που συνθέτουν το ορόσημο

των υπο-ορόσημων παρουσιάζονται στους πίνακες 4.14 και 4.15.

Πλειάδα περιγραφής ‘‘Εκμάθηση 3.1’’							
I	N	L_0	P_1	P_2	P_3	P_4	P_5
12.018351	328.688843	3	(97,144)	(52,141)	(72,146)	(70,159)	(55,152)

Πίνακας 4.14: Η πλειάδα περιγραφής του πρώτου υπο-ορόσημου στη βιβλιοθήκη

Πλειάδα περιγραφής ‘‘Εκμάθηση 3.2’’							
I	N	L_0	P_1	P_2	P_3	P_4	P_5
12.769866	413.957825	3	(97,144)	(72,146)	(102,154)	(70,159)	(55,152)

Πίνακας 4.15: Η πλειάδα περιγραφής του δεύτερου υπο-ορόσημου στη βιβλιοθήκη

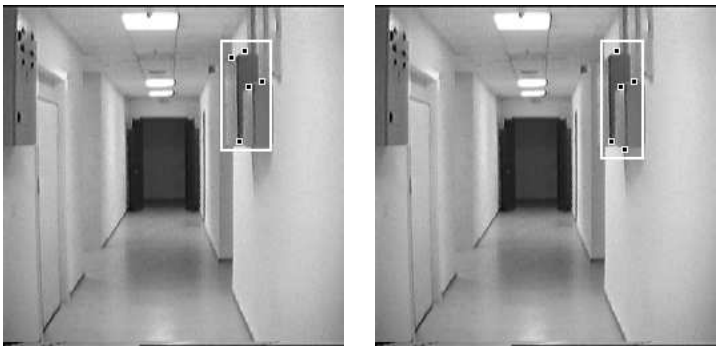
4.2.3.2 Περίοδος αναγνώρισης

Στην περίοδο της αναγνώρισης, η ακολουθία των πλαισίων εισόδου απεικονίζεται στην εικόνα 4.26. Εκτελώντας τα ενδιάμεσα βήματα, καταλήγουμε στις δύο πιο εξαιρετικές

(α) (β)

Σχήμα 4.26: Ακολουθία εικόνων στο τρίτο παράδειγμα

πεντάδες της εικόνας 4.27, των οποίων οι αντίστοιχες πλειάδες περιγραφής παρατίθενται στους πίνακες 4.16 και 4.17. Κατόπιν, συγκρίνουμε αρχικά την πλειάδα περιγραφής ‘‘Αναγνώριση 3.1’’ με τις πλειάδες ‘‘Εκμάθηση 3.1’’ και ‘‘Εκμάθηση 3.2’’ της βιβλιοθήκης, για να δούμε αν το μηχανικό σύστημα έχει εκπαιδευτεί με αυτή. Η σύγκριση επιτελείτε μέσω του κριτηρίου του k-nearest-neighbor, το οποίο δίνει σαν ελάχιστη τιμή, για τις δύο συγκρίσεις, την 393.812257, η οποία παράγεται όταν συγκρίνουμε την πλειάδα ‘‘Αναγνώριση 3.1’’ με την πλειάδα ‘‘Εκμάθηση 3.1’’. Άρα αντιστοιχούμε την πλειάδα ‘‘Εκμάθηση 3.1’’, με την πλειάδα ‘‘Αναγνώριση 3.1’’. Στη συνέχεια, παρατηρούμε, με βάση τον αριθμό του κατόχου της πλειάδας ‘‘Εκμάθηση 3.1’’, ότι ενδέχεται να υπάρχει επιπλέον ένα υπο-ορόσημο στο πλαίσιο αναγνώρισης. Ελέγχουμε αν η πλειάδα ‘‘Αναγνώριση 3.2’’ μπορεί να αντιστοιχιθεί με την πλειάδα ‘‘Εκμάθηση



(1) (2)

Σχήμα 4.27: Οι δύο πιο εξαιρετικές πεντάδες που επιλέγουμε

Πλειάδα περιγραφής ‘‘Αναγνώριση 3.1’’							
I	N	L_0	P_1	P_2	P_3	P_4	P_5
11.382854	1031.099	3	(102,177)	(34,181)	(61,184)	(108,187)	(57,194)

Πίνακας 4.16: Η πλειάδα περιγραφής για την πρώτη πεντάδα

Πλειάδα περιγραφής ‘‘Αναγνώριση 3.2’’							
I	N	L_0	P_1	P_2	P_3	P_4	P_5
12.918646	722.500061	3	(102,177)	(34,181)	(61,184)	(39,171)	(57,194)

Πίνακας 4.17: Η πλειάδα περιγραφής για τη δεύτερη πεντάδα

3.2’’. Συγκρίνοντας την πλειάδα ‘‘Αναγνώριση 3.2’’ με τις πλειάδες ‘‘Εκμάθηση 3.1’’ και ‘‘Εκμάθηση 3.2’’, μέσω του k-nearest-neighbor, λαμβάνουμε σαν ελάχιστη τιμή την 308.542267, η οποία παράγεται κατά την σύγκριση της πλειάδας ‘‘Αναγνώριση 3.2’ με την πλειάδα ‘‘Εκμάθηση 3.2’’. Έτσι, καταλήγουμε στο συμπέρασμα της αναγνώρισης ενός ορόσημου.

Κεφάλαιο 5

Επίλογος

Στην εργασία αυτή, παρουσιάστηκε μια προσέγγιση για την αυτόνομη πλοήγηση ενός μηχανικού συστήματος μέσω της αυτοματοποιημένης εξαγωγής ορόσημων κατά τη διάρκεια μιας αρχικής περιόδου εκμάθησης και της ικανότητας αναγνώρισης των κατά τη διάρκεια πλοηγητικών αποστολών. Η μέθοδος αυτή βασίζεται στην έννοια της τοπολογικής πλοήγησης και χρησιμοποιεί μεταθετικά αναλλοίωτες ποσότητες για τους σκοπούς της εκμάθησης και της αναγνώρισης. Η μόνη της παραδοχή είναι η ύπαρξη επίπεδων επιφανειών στο χώρο εργασίας. Πειραματικά αποτελέσματα έδειξαν ότι η μέθοδος διευκολύνει την επακριβή αναγνώριση, αποφεύγοντας περιορισμούς που επιβάλλονται από γεωμετρικές μεθόδους ή υπερβολικά επιβαρυντικούς υπολογισμούς που χρησιμοποιούνται στην επίλυση της σημείο προς σημείο αντιστοίχισης. Τα αποτελέσματα από αυτή την εργασία ήταν αρκετά ενθαρρυντικά σχετικά με την προτεινόμενη μέθοδο οπτικής πλοήγησης και στην περίπτωση των πραγματικών δεδομένων.

Η εφαρμογή της μεθόδου για την εξαγωγή μη προκαθορισμένων ορόσημων έδειξε ότι μπορούμε να ελπίζουμε σε καλά αποτελέσματα στην κατεύθυνση αυτή. Η αποφυγή της προεπιλογής των ορόσημων φανέρωσε ότι μπορούμε να βελτιώσουμε την ευελιξία και την δυναμική προσαρμογή των αυτόνομων συστημάτων.

Οι χρόνοι εκτέλεσης της μεθόδου στις ακολουθίες εικόνων που εφαρμόστηκε υπήρξαν ικανοποιητικοί, αναλογικά με την δυσκολία του προβλήματος το οποίο αντιμετωπίζει. Σημαντικό ρόλο στην βελτίωση της ταχύτητας εκτέλεσης είχαν η χρήση των ποιοτικών κριτηρίων επιλογής χαρακτηριστικών καθώς και η συνεισφορά της βιβλιοθήκης στη φάση της αναγνώρισης.

5.1 Μελλοντικές κατευθύνσεις

Οπωσδήποτε η συγκεκριμένη εργασία δεν εξαντλεί το πρόβλημα της οπτικής πλοήγησης. Ένα επόμενο βήμα θα ήταν να αντιμετωπιστούν οι περιπτώσεις όπου δεν ισχύει η παραδοχή της ύπαρξης επίπεδων επιφανειών στο χώρο κίνησης. Στη συγκεκριμένη

περίπτωση θα πρέπει να μελετήσουμε το δύσκολο πρόβλημα εύρεσης αναλλοίωτων οι οποίες υπόκεινται σε πιο χαλαρούς περιορισμούς. Το πρόβλημα είναι δύσκολο γιατί, όπως απέδειξαν οι Burns et al. [130], δεν υπάρχουν προβολικές αναλλοίωτες που να μην υπόκεινται σε κάποιο περιορισμό. Οπότε θα πρέπει να ερευνήσουμε ποιο είναι το σύνολο των περιορισμών που δεν περιορίζει την αυτονομία του μηχανικού συστήματος και το καθιστά λειτουργικό και ευέλικτο σε γενικής φύσεως περιβάλλοντα. Μια άμεση εφαρμογή της επέκτασης θα ήταν η εφαρμογή της μεθόδου σε εξωτερικά περιβάλλοντα, όπου οι σχηματισμοί είναι πιο πολύπλοκοι και η δομή του χώρου είναι αρκετά ανομοιογενής.

Επίσης θα πρέπει να διερευνήσουμε, κατά πόσο οι συναρτήσεις που συνθέτουν το χάρτη προεξοχής είναι επαρκείς για τον χαρακτηρισμό των σημαντικών περιοχών μιας σκηνής. Ενδεχομένως, να απαιτείται ο εμπλουτισμός του χάρτη με γεωμετρικής φύσεως πληροφορία, η οποία, να αντιμετωπίζει τα προβλήματα μιας καθαρά ποσοτικής αντιμετώπισης.

Προκειμένου να αντιμετωπιστούν οι έντονες αριθμητικές αστάθειες των γεωμετρικών αναλλοίωτων, το κριτήριο αντιστοίχισης θα μπορούσε να συνδυαστεί και με άλλες προϋποθέσεις. Για παράδειγμα, ο Morrin et. al. [131] αναφέρει πως μία απαραίτητη, αλλά όχι ικανή, συνθήκη για την επιβεβαίωση της επιτυχής αντιστοίχισης ανάμεσα σε δυο σύνολα σημείων είναι ότι θα πρέπει να βρίσκονται σε αντιστοίχιση και τα κυρτά τους περιβλήματα, μιας και αυτά διατηρούνται κάτω από προοπτικούς μετασχηματισμούς [132].

Επίσης, ο ανιχνευτής χαρακτηριστικών μπορεί να βελτιστοποιηθεί με την εισαγωγή αυτόματων μηχανισμών απόρριψης *κίβδηλων (spurrious)* χαρακτηριστικών. Μια τέτοια προσέγγιση αναπτύσσεται στο [131], όπου τα κίβδηλα χαρακτηριστικά είναι αυτά που δεν έχουν αντίστοιχα σε μια ακολουθία εικόνων[1]. Με τον τρόπο αυτό, μειώνουμε το πλήθος των χρήσιμων χαρακτηριστικών, οδηγώντας σε απλοποίηση της διαδικασίας συλλογής και των υπολογισμών αλλά και σε ελλάττωση της χρονικής και χωρικής πολυπλοκότητας της μεθόδου.

Ένα αρκετά ενδιαφέρον πρόβλημα είναι η μελέτη της αβεβαιότητας που εισάγεται λόγω μερικής ή ολικής απόκρυψης (occlusion) των υπο-ορόσημων. Για την αντιμετώπιση της θα μπορούσαν να χρησιμοποιηθούν κάποια μέτρα εμπιστοσύνης (confidence measures) [111] με τη βοήθεια των οποίων μπορούμε να προβλέψουμε τα πιθανά τοπολογικά σενάρια.

Τέλος, ένας τρόπος επέκτασης της μεθόδου θα ήταν η χρήση και άλλων αισθητήρων από το αυτόνομο σύστημα (π.χ. αισθητήρες απόστασης) εκτός από τους οπτικούς που χρησιμοποιήθηκαν στην εργασία αυτή. Ο συνδυασμός των πληροφοριών από διάφορους αισθητήρες (sensor fusion) θα επιτρέψει στο σύστημα την εξαγωγή ασφαλέστερων συμπερασμάτων και τη βελτίωση της ταχύτητας εκτέλεσής του.

[1]Τα χαρακτηριστικά αυτά είναι γνωστά σαν *outliers*

Παράρτημα Α

Ευρετήριο όρων

Α

Άμεση διαχωρισιμότητα	Distinctiveness
Αισθητήρας υπερήχων	Sonar
Ανομοιότητα	Dissimilarity
Αντίληψη του χώρου	Space perception
Αντιστοίχιση	Correspodence
Αυτόνομη πλοήγηση	Autonomous navigation

Β

Βαθμονομημένος	Calibrated
Βιβλιοθήκη ορόσημων	Landmark library

Γ

Γράφος ορατότητας	Visibility graph

Δ

Διάνυσμα μετατόπισης	Displacement vector
Δια χειρός	Manually
Διάγραμμα Voronoi	Voronoi diagram
Διαχωρισμός θέσης	Cell decomposition
Διμερή συμμετρία	Bilateral symmetry
Δομή από κίνηση	Structure from motion

Ε

Εκ των προτέρων γνώση	A priori knowledge
Εκτίμηση παραμέτρων ανεξάρτητης κίνησης	Independent motion estimation
Εκτίμηση παραμέτρων κίνησης	Egomotion estimation
Εκτιμητής εξαιρετικού	Outlier estimator
Ελάχιστο περίβλημα	Minimum bounding box
Ενεργής και τελεολογικής όραση	Active and purposive vision
Εξαγωγή συμπερασμάτων	Reasoning
Εξωτερικός λόγος	Cross-ratio
Επίπεδο εικόνας	Image plane
Επιλεγμένο ορόσημο	Selected landmark
Εστίαση της προσοχής	Focus of attention
Ευθεία γραμμή	Straight line

Η

Ημι-συνευθειακός	Quasi-collinear

Κ

Κλίση	Gradient

Μ

Μαθηματικός προγραμματισμός	Mathematical programming
Μετακίνηση από σημείο σε σημείο	Homing
Μη-παθολογικός	Nonsingular

Ο

Ομαλός	Smoothed
Οπτικό πρότυπο	Visual pattern
Οπτική πλοήγηση	Visual navigation
Οριοθέτηση	Localization
Ορόσημο	Landmark

Π

Πανοραμική άποψη	Panoramic view
Πεδίο οπτικής ροής	Optical flow field
Πλειάδα περιγραφής	Descriptor tupple
Πλοήγηση προς κάποιο στόχο	Targeted navigation
Πλοηγητική αποστολή	Navigational task
Πολυμορφικός	Generic
Προέχων	Salient
Προγραμματισμός κίνησης	Motion planning
Προσχεδιασμένο ορόσημο	Predesigned landmark
Πρόβλημα αντιστοίχισης	Correspondence problem
Πρότυπο	Pattern

Σ

Σημείο διαφυγής	Vanishing point
Σκελετός	Skeleton
Συνεχής εντοπισμός	Tracking
Σωρευτική απόσταση	aggregate distance

Τ

Ταύτιση φορμών	Template matching

Χ

Χώρος ακτινών	Ray space

Βιβλιογραφία

[1] U. Nehmzow. Animal and robot navigation. *Robotics And Autonomous Systems*, 15(15):71--81, July 1995.

[2] A. Argyros. Visual detection of independent 3d motion by a moving observer. *Ph.D.thesis. University of Crete, Computer Science Department*, 1996.

[3] N. Paragyios. Adaptive detection and location of moving objects in image sequences. *Master's thesis. University of Crete, Computer Science Department*, 1996.

[4] D. Marr. Vision: A computational investigation into the human representation and processing of visual information. In *W.H. Freeman*, 1982.

[5] G. Tziritas and C. Labit. *Motion Analysis for Image Sequence Coding*. Elsevier, 1994.

[6] S. Ullman. The interpretation of structure from motion. *Royal Society, London*, 203:405--426, 1979.

[7] G. Adiv. Determining 3-d motion and structure from optical flow generated by several moving objects. *PAMI*, 7(4):384--401, July 1985.

[8] M. Subbarao and A.M. Waxman. Closed form solutions to image flow equations for planar surfaces in motion. *CVGIP*, 36(2/3):208--228, November 1986.

[9] J.Y.A. Wang and E.H. Adelson. Representing moving images with layers. In *Vismod*, 1993.

[10] Y. Huang, K. Palaniappan, X. Zhuang, and J.E. Cavanaugh. Optic flow field segmentation and motion estimation using a robust genetic partitioning algorithm. *PAMI*, 17(12):1177--1190, December 1995.

[11] Michael Magee and J.K. Aggarwalls. Robot self-location using visual reasoning relative to a single target object. *Pattern Recognition*, 28(2), 1995.

[12] J.K. Aggarwal and N. Nandhakumar. On the computation of motion from sequences of images: A review. *PIEEE*, 76:917--935, 1988.

[13] S. Ullman. The interpretation of visual motion. In *MIT Press*, 1979.

[14] D. Murray and A. Basu. Motion tracking with an active camera. *PAMI*, 16(5):449--459, May 1994.

[15] B.K.P. Horn and B.G. Schunck. Determining optical flow. *AI*, 17:185--203, 1981.

[16] N. Cornelius and T. Kanade. Adapting optical-flow to measure object motion in reflectance and x-ray image sequences. In *Motion83*, pages 50--58, 1983.

[17] H.H. Nagel. On the estimation of optical flow: Relations between different approaches and some new results. *AI*, 33(3):299--324, November 1987.

[18] M.J. Black and P. Anandan. The robust estimation of multiple motions: Parametric and piecewise-smooth flow-fields. *CVIU*, 63(1):75--104, January 1996.

[19] J. Biemond, L. Looijenga, D. E. Boekee, and R. H. J. M. Plompen. A pel-recursive wiener-based displacement estimation algorithm. *Signal Processing*, 13:399--412, 1987.

[20] J. N. Driessen and J. Biemond. Motion field estimation for complex scenes. *Proceedings of the SPIE - The International Society for Optical Engineering*, 1605(pt.1):511--21, 1991.

[21] A.N. Netravali and J.D. Robbins. Motion-compensated television coding: Part 1. *Bell System Tech.*, 58(3):631--670, March 1979.

[22] J.A. Stuller and G. Krishnamurthy. Kalman filter formulation of low-level television image motion estimation. *CVGIP*, 21(2):169--204, 1983.

[23] G. Tziritas. Recursive and/or iterative estimation of the two-dimensional velocity field and reconstruction of three-dimensional motion. *Signal Processing*, 16:53--72, 1989.

[24] D.R. Walker and K.R. Rao. Improved pel-recursive motion compensation. *Commun*, 32:1128--1134, 1984.

[25] A. Mitiche and P. Bouthemy. Computation and analysis of image motion: A synopsis of current problems and methods. *IJCV*, 19(1):29--55, July 1996.

[26] T. Lozano-Perez and M.T. Mason. Automatic synthesis of fine-motion strategies for robots. *Int. J. Robotics Res.*, 3(1):3--24, 1984.

[27] V.J. Lumelsky, S. Mukhopadhay, and K. Sun. Dynamic path planning in sensor-based terrain acquisition. *IEEE Trans. on Robotics and Automation*, 6(4), Aug 1990.

[28] K. Sugihara. Some location problems for robot navigation using a single camera. *Computer Vision, Graphics and Image Processing*, 42:112--129, Apr. 1988.

[29] B. C. Bloom. Use of landmarks for mobile robot navigation. *Proc.* SPIE *Intelligent Robots and Computer Vision*, 579:351--355, 1985.

[30] Peter Kampmann and Gunther Schmidt. Indoor navigation of mobile robots by use of learned maps. Information Processing in Autonomous Mobile Robots *Springer-Verlag, Berlin, Heidelberg, New York, NY*, 1991.

[31] A. Zelinsky. Environment mapping with a mobile robot using sonar. In C. J. Barter and M. J. Brooks, editors, *Proceedings of the 2nd Australian Joint Artificial Intelligence Conference, AI '88*, volume 406 of *LNAI*, pages 363--378, Adelaide, Australia, nov 1988. Springer-Verlag.

[32] S. S. Iyengar, C. C. Jorgensen, S. V. N. Rao, and C. R. Weisbin. Robot navigation algorithms using learned spatial graphs. *Robotica*, 4:93--100, 1986.

[33] H. Takeda and J.-C. Latombe. Sensory uncertainty field for mobile robot navigation. *IEEE International Conference on Robotics and Automation, Nice, France*, pages 2465--2472, 1992.

[34] P.E. Trahanias and Y. Komninos. Robot motion planning: Multi-sensory uncertainty fields enhanced with obstacle avoidance. *IEEE/RSJ Intl. Conf. on Intell. Robots and Systems, IROS'96*, pages 1141--1148, Nov. 1996.

[35] A. Kak, K. Andress, and C. Lopez-Albadia. Mobile robot self-location with the pseiki system. In *Working Notes, AAAI Spring Symposium on Robot Navigation, Cambridge*, March 1989.

[36] Akio Kosaka and A.C. Kak. Fast vision-guided mobile robot navigation using model-based reasoning and prediction of uncertainties. *Computer Vision, Graphics and Image Processing-IU*, 56(3), March 1992.

[37] T. Tsubouchi and S. Yuta. Map assisted vision system of mobile robots for reckoning in a building environment. In *Proc. IEEE Int. Conf. Robotics and Automation, (Raleigh, North Carolina)*, March 1987.

[38] J.C. Latombe. *Robot Motion Planning.* Kluver Academic Publishers, Boston, MA, 1991.

[39] Yong K. Hwang and Narendra Ahuja. Gross motion planning - a survey. *ACM Computing Surveys*, 24(3):221--291, 1992.

[40] J. T. Schwartz and M. Sharir. On the 'piano movers' problem: II general techniques for computing topological properties of real algebraic manifolds. *Advances in Applied Mathematics, Academic Press 4*, pages 298--351, 1983.

[41] N.J. Nilsson. A mobile automation, an application of artificial intelligence techniques. *Int. Joint Conf. Artificial Intelligence*, pages 509--520, 1969.

[42] J. F. Canny. The complexity of robot motion planning. *MIT Press, Cambridge, MA*, 1988.

[43] F. Aurenhammer. Voronoi diagrams: A survey of a fundamental geometric data structure. *Surveys*, 23(3):345--405, September 1991.

[44] C. O'Dunlaing, M. Sharir, and C. K. Yap. Generalised voronoi diagrams for moving a ladder. *Technical Report 32, Robotics Laboratory, Courant Institute, New York University*, 1984.

[45] T. Lozano-Perez and M.A. Wesley. An algorithm for planning collision-free paths among polyhedral obstacles. *Commun. ACM*, 22(10):560--570, Oct. 1979.

[46] K.D. Rueb and K.C. Wong. Structuring free space as a hypergraph for roving robot path planning and navigation. *IEEE Trans. on Pattern Anal. and Mach. Int.*, 9(2):263--273, Mar. 1987.

[47] F. P. Preparata and M. I. Shamos. Computational geometry: An introduction. *Springer-Verlag, New York, NY*, 1985.

[48] R.B. Tilove. Local obstacle avoidance for mobile robots based on the method of artificial potentials. *Proc. of the IEEE Int. Conf. on Robotics and Automation, Cincinati, OH*, pages 566--571, 1990.

[49] O. Khatib. Real-time obstacle avoidance for manipulators and mobile robots. *Int. J. Robotics Res.*, 5(1):90--98, 1986.

[50] B. Faverjon and P. Tournassoud. A local based approach for path planning of manipulators with a high number of degrees of freedom. *Proc. of the IEEE Int. Conf. on Robotics and Automation, Railegh, NC*, pages 1152--1159, 1987.

[51] H. Takeda and J.-C. Latombe. Sensory uncertainty field for mobile robot navigation. *IEEE International Conference on Robotics and Automation, Nice, France*, pages 2465--2472, 1992.

[52] J. Aloimonos and A. Badyopadhyay. Active vision. In *1991 IEEE 1st Int. Conf. on Computer Vision*, pages 35--54, June 1987.

[53] J. Aloimonos. Purposive and qualitative vision. In *DARPA Image Understanding Workshop*, pages 816--828, 1990.

[54] D.H. Ballard. Animate vision. *Artificial Intelligence*, 48:57--86, 1991.

[55] J. Clark and N. Ferrier. *Active Vision*, chapter 9, pages 137--154. Artificial Intelligence series. MIT Press, Cambridge, MA, 1993.

[56] A. Lazanas and J.C. Latombe. Motion planning with uncertainty: A landmark approach. *AI*, 76:287--317, 1995.

[57] Shishir Shah and J.K. Aggarwalls. Modeling structured environments using robot vision. In *Proc. of Conf. on Computer Vision, Singapure*, Dec 1995.

[58] R. Greiner and R. Isukapalli. Learning to select useful landmarks. *IEEE Trans. Systems, Man, Cybern. Part B: Cybernetics*, 26(3):437--449, 1996.

[59] Camillo J. Taylor and David J. Kriegman. Vision-based motion planning and exploration algorithms for mobile robots. Technical report, Yale University, 1994.

[60] A. Lazanas and J.C. Latombe. Landmark-based robot navigation. In *AAAI-92*, pages 816--822, 1992.

[61] Michael Magee and J.K. Aggarwalls. Robot self-location using visual reasoning relative to a single target object. *Pattern Recognition*, 28(2), 1995.

[62] H. Nasr and B. Bhanu. Landmark recognition for autonomous mobile robots. *IEEE Intl. Conf. on Robotics and Automation*, pages 1218--1223, 1988.

[63] J.Y. Bouguet and P. Perona. Visual navigation using a single camera. In *ICCV95*, pages 645--652, 1995.

[64] Erliang Yeh and David J. Kriegman. Toward selecting and recognizing natural landmarks. Technical report, Yale University, 1995.

[65] Erliang Yeh and David J. Kriegman. Toward selecting and recognizing natural landmarks. *IEEE/RSJ Intl. Conf. on Intell. Robots and Systems, IROS'95*, Aug. 1995.

[66] Panos Trahanias, Manolis Lourakis, Antonis Argyros, and Stelios Orphanoudakis. Vision-based assistive navigation for robotic wheelchair platforms. Technical Report no. 178, ICS-FORTH, Heraklion, Crete, Greece, oct 1996.

[67] Panos Trahanias, Manolis Lourakis, Antonis Argyros, and Stelios Orphanoudakis. Navigational support for robotic wheelchair platforms: An approach that combines vision and range sensors. In *Proc. of ICPA 1997*, pages 20--25, Albuquerque, New Mexico, April 1997.

[68] Akio Kosaka and Juiyao Pan. Purdue experiments in model-based vision for hallway navigation. In *Proc. of Workshop on vision for Robots in IROS'95 Conference, Pittsburgh*, Oct 1992.

[69] Jiang Yu Zheng, Matthew Barth, and Saburo Tsuji. Autonomous landmark selection for route recognition by a mobile robot. In *Proc. of the 1991 IEEE International Conference on robotics and Automation, Sacramento, California*, April 1991.

[70] J.Y. Zheng. Dynamic projection, panoramic representation of scenes for route recognition. *Ph.D.thesis in Osaka University*, Dec 1989.

[71] J. Y. Zheng and S. Tsuji. Spatial representattion and analysis of temporal visual events. In *Proc. of the 1991 IEEE International Conference on Image Processing*, pages 775--779, September 1989.

[72] P.E. Trahanias, S. Velissaris, and T. Garavelos. Visual landmark extraction and recognition for autonomous robot navigation. *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems, IROS'97*, pages 7--11, September 1997.

[73] Kyoung Sig Roh and In So Kweon Wang Heon Lee. Obstacle detection and self-localization without camera calibration using projective invariants. *IEEE/RSJ Intl. Conf. on Intell. Robots and Systems, IROS'97*, pages 1030--1035, August 1997.

[74] Vassilios S. Tsonis, Konstantinos V. Chandrinos, and Panos E. Trahanias. Landmark-based navigation using projective invariants. *IEEE/RSJ Intl. Conf. on Intell. Robots and Systems, IROS'98*, October 1998.

[75] Konstantinos V. Chandrinos, Vassilios S. Tsonis, and Panos E. Trahanias. Automated landmark extraction and retrieval. In *Proceedings of the 6th International Symposium on Intelligent Robotic Systems, SIRS'98, Edinburgh*, pages 223--230, 21-23 July 1998.

[76] R. Mohr and B. Triggs. Projective geometry for image analysis. In *International Symposium of Photogrammetry and Remote Sensing*, Vienna, July 1996.

[77] O.D. Faugeras. Three-dimensional computer vision. *MIT Press*, 1993.

[78] J.L. Mundy and A. Zisserman. Geometric invariance in computer vision. In *MIT Press*, 1992.

[79] R.C. Gonzalez and P.A. Wintz. Digital image processing. In *Addison-Wesley*, 1982.

[80] L. G. Roberts. *Machine Perception of Three-Dimensional Solids, in: J.T. Tippet et. al., Optical and Electro-Optical Information Processing*, pages 159--197. MIT Press, Cambridge, MA, 1965.

[81] I. Weiss. Geometric invariants and object recognition. *IJCV*, 10(3):207--231, June 1993.

[82] L. Quan, P. Gros, and R. Mohr. Invariants of a pair of conics revisited. *IVC*, 10:319--323, 1992.

[83] R.O. Duda and P.E. Hart. Pattern classification and scene analysis. In *Wiley*, 1973.

[84] C. E. Springer Van Loan. *Geometry and Analysis of Projective Spaces.* W. H. Freeman and Company, San Francisco, 1964.

[85] M.H. Brill and E.B. Barrett. Closed-form extension of the anharmonic ratio to n-space. *CVGIP*, 23(1):92--98, 1983.

[86] J. Coehlo, C. Heller A. Mundy and D. Fosyth. An experimental evaluation of projective invariants. In *DARPA Image Understanding Workshop*, pages 273--294, Aug. 1991.

[87] D. Sinclair and A. Blake. Quantitative planar region detection. *IJCV*, 18(1):77--91, April 1996.

[88] M.A. Naimark and A.I.Stern. Theory of group representations. New York, 1982. Springer-Verlag.

[89] Charles W. Curtis and Irving Reiner. *Representation Theory of Finite Groups and Associative Algebras.* Number XI in Pure and Applied Mathematics. Interscience Publishers, 1962.

[90] P. Meer and R. Lenz. Projection and permutation invariant object recognition. pages 389--396, Tromsö, Norway, May 1993.

[91] W. Fulton and J. Harris. *Representation Theory.* Graduate Texts in Mathematics. Springer-Verlag, New York, 1991.

[92] Robin Hartshorne. *Algebraic Geometry*, volume 52 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, 1977.

[93] S. J. Maybank. The projection of two non-coplanar conics. In Geometric Invariance in Machine Vision (Mundy, J.L. and Zisserman, A. Eds), pages 105--119, Cambridge, Massachusetts, 1992. MIT Press.

[94] P. Meer, S. Ramakrishna, and R Lenz. Indexing power of the projective/permutation invariant of five coplanar points. volume 52 of *Application of Invariance in Computer Vision II, (Mundy, J.L. and Zisserman, A. Eds)*, pages 133--150, Ponta-Deglada, Azores, 1993. MIT Press.

[95] R.M. Haralick. Monocular vision using inverse perspective projection geometry: Analytic relations. In *CVPR*, pages 370--378, 1989.

[96] P. Meer, S. Ramakrishna, and R. Lenz. Correspondence of coplanar features through p2-invariant representations. In *ICPR94*, pages A:196--200, 1994.

[97] D.A. Forsyth, J.L. Mundy, A. Zisserman, C. Coelho, A. Heller, and C.A. Rothwell. Invariant descriptors for 3-d object recognition and pose. *PAMI*, 13(10):971--991, October 1991.

[98] C.A. Rothwell, A. Zisserman, C.I. Marinos, D.A. Forsyth, and J.L. Mundy. Relative motion and pose from arbitrary plane curves. *IVC*, 10:250--262, 1992.

[99] C. Rothwell and J. Stern. Understanding the shape properties of trihedral polyhedra. In *ECCV96*, pages I:175--185, 1996.

[100] C.A. Rothwell. Object recognition through invariant indexing. In *Oxford University Press*, 1995.

[101] C.A. Rothwell. Recognition using projective invariance. In *Ph. D.*, 1993.

[102] C.A. Rothwell and S. Utcke. Morse: A 3d object recognition system based on geometric invariants. In *ARPA94*, pages II:1393--1402, 1994.

[103] D.A. Forsyth, J.L. Mundy, A. Zisserman, and C.A. Rothwell. Recognizing rotationally symmetric surfaces from their outlines. *ECCV*, 92:639--647.

[104] C.A. Rothwell, D.A. Forsyth, A. Zisserman, and J.L. Mundy. Extracting projective structure from single perspective views of 3d point sets. In *ICCV93*, pages 573--582, 1993.

[105] J.L. Mundy, A. Liu, N. Pillow, A. Zisserman, S. Abdallah, S. Utcke, S. Nayar, and C.A. Rothwell. An experimental comparison of appearance and geometric model recognition. In *ORCV96*, page 247, 1996.

[106] C.A. Rothwell, A. Zisserman, D.A. Forsyth, and J.L. Mundy. Planar object recognition using projective shape representation. *IJCV*, 16(1):57--99, September 1995.

[107] S. Carlsson. Projectively invariant decomposition and recognition of planar shapes. *IJCV*, 17(2):193--209, February 1996.

[108] S. Carlsson, R. Mohr, T. Moons, L. Morin, C.A. Rothwell, M. VanDiest, L. VanGool, F. Veillon, and A. Zisserman. Semilocal projective invariants for the recognition of smooth plane-curves. *IJCV*, 19(3):211--236, August 1996.

[109] R. Lenz. Group theoretic methods in image processing. In *Springer-Verlag*, 1990.

[110] R. Lenz. Optimal filters for the detection of linear patterns in 2-D and higher dimensional images. *Pattern Recognition*, 20:163--172, 1987.

[111] S. Halkidis. Planar feature matching of views related through general projective transformations through the use of geometric invariants. *Master's thesis. University of Crete, Department of Computer Science*, 1998.

[112] G. S. Cox and G. de Jager. A survey of point pattern matching techniques and a new approach to point pattern recognition. In *South African Symposium on Communications and Signal Processing*, pages 243--8, 1993.

[113] S.M. Smith and J.M. Brady. Susan: A new approach to low-level image-processing. *IJCV*, 23(1):45--78, May 1997.

[114] C. Tomasi and J. Shi. Good features to track. In *CVPR94*, pages 593--600, 1994.

[115] C. Tomasi and J. Shi. Good features to track. Technical Report TR 93-1399, Cornell University, December 1993.

[116] C. Tomasi and T. Kanade. Detection and tracking of point features. Technical Report CMU-CS-91-132, Carnegie Mellon University, Computer Science Department, Pittsburgh, PA, April 1991.

[117] G. H. Golub and C. F. Van Loan. *Matrix Computations*, 2nd ed. Johns Hopkins University Press, Baltimore, 1989.

[118] K.I. Kanatani. Computational projective geometry. *CVGIP*, 54(3):333--348, November 1991.

[119] E. Biglieri and K. Yao. Some properties of singular value decomposition and their applications to digital signal processing. *Signal Processing*, 18:277--289, 1989.

[120] Huttenlocher, Klanderman, and Rucklidge. Comparing images using the hausdorff distance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15, 1993.

[121] A. Khotanzad and Y.H. Hong. Invariant image recognition by zernike moments. *PAMI*, 12(5):489--497, May 1990.

[122] S. Maitra. Moment invariants. *PIEEE*, 67:697--699, 1979.

[123] E. Persoon and K. Fu. Shape discrimination using fourier descriptors. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 8(3):388--391, May 1986.

[124] William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery. *Numerical Recipes in C, 2nd. edition*. Cambridge University Press, 1992.

[125] R. Smith and P. Cheeseman. On the representation and estimation of spatial uncertainty. *International Journal of Robotics Research*, 5(4):56--69, 1987.

[126] H. F. Durrant-Whyte. *Integration, Coordination, and Control of Multi-Sensor Robot Systems.* Kluwer Academic Publishers, Boston, 1987.

[127] R. J. Schalkoff. *Pattern Recognition: Statistical, Structural and Neural Approaches.* Wiley, New York, 1992.

[128] George E. Forsythe, Michael A. Malcolm, and Cleve B. Moler. *Computer Methods for Mathematical Computations.* Prentice-Hall, Englewood Cliffs, NJ 07632, USA, 1977.

[129] Christos Faloutsos. *Searching Multimedia Databases by Content.* Kluwer Academic Publishers, 1996.

[130] J.B. Burns, R. Weiss, and E.M. Riseman. View variation of point set and line segment features. *PAMI*, 15(1):51--68, January 1993.

[131] P. Mohr Morin, L. Brand. Indexing with projective invariants. In *Presented at the Syntactical and Structural Pattern Recognition. Workshop*, Nahariya, Israel, 1994.

[132] R.M. Haralick and L.G. Shapiro. Computer and robot vision. In *Addison-Wesley*, 1992.