ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ
ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΤΜΗΜΑ ΕΠΙΣΤΗΜΗΣ ΥΠΟΛΟΓΙΣΤΩΝ

Πλοήγηση ρομποτικού συστήματος σε εσωτερικούς χώρους με χρήση οπτικής πληροφορίας

Σάββας Γ. Βελισσάρης

Μεταπτυχιακή Εργασία

Ηράκλειο, Δεκέμβριος 1996

ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ
ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΤΜΗΜΑ ΕΠΙΣΤΗΜΗΣ ΥΠΟΛΟΓΙΣΤΩΝ

Πλοήγηση ρομποτικού συστήματος σε εσωτερικούς χώρους με χρήση οπτικής πληροφορίας

Εργασία που υποβλήθηκε από τον
Σάββα Γ. Βελισσάρη
ως μερική εκπλήρωση των απαιτήσεων
για την απόκτηση
ΜΕΤΑΠΤΥΧΙΑΚΟΥ ΔΙΠΛΩΜΑΤΟΣ ΕΙΔΙΚΕΥΣΗΣ

Συγγραφέας:

Σάββας Γ. Βελισσάρης
Τμήμα Επιστήμης Υπολογιστών

Εισηγητική Επιτροπή:

Πάνος Τραχανιάς
Επίκουρος Καθηγητής, Επόπτης

Στέλιος Ορφανουδάκης
Καθηγητής, Μέλος

Νίκος Αλβέρτος
Επίκουρος Καθηγητής, Μέλος

Δεκτή:

Πανος Κωνσταντόπουλος
Αναπληρωτής Καθηγητής
Πρόεδρος Επιτροπής Μεταπτυχιακών Σπουδών

Ηράκλειο, Δεκέμβριος 1996

Ευχαριστίες

Στο σημείο αυτό θα ήθελα να ευχαριστήσω όλους όσους με τον ένα ή τον άλλο τρόπο με στήριξαν κατά τη διάρκεια της προσπάθειας για την πραγματοποίηση αυτής της εργασίας.

Πρώτα απ' όλα, οφείλω ένα μεγάλο ευχαριστώ στον επόπτη μου, τον Πάνο Τραχανιά, υπό την καθοδήγηση του οποίου πραγματοποιήθηκε αυτή η εργασία. Θέλω να τον ευχαριστήσω για την επιστημονική του καθοδήγηση, τις συμβουλές του και τις ώρες που με μεγάλη υπομονή και διάθεση μου αφιέρωσε, πάντα άψογος στο επίπεδο των προσωπικών σχέσεων. Εύχομαι να έχω την τύχη να συνεργαστώ με ανθρώπους σαν τον Πάνο Τραχανιά και στο μέλλον.

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

Σημαντική συμπαράσταση είχα και απο πολλούς φίλους και συναδέλφους φοιτητές, που παρά το δικό τους φόρτο εργασίας με υπομονή και όρεξη ήταν πάντα διαθέσιμοι για σχόλια και παρατηρήσεις σε πολλά στάδια της εργασίας. Ιδιαίτερα θέλω να ευχαριστήσω τον Θόδωρο Γκαράβελο, με τον οποίο είχα μία απρόσκοπτη και εποικοδομητική συνεργασία, τον Μιχάλη Αρταβάνη για τις χρήσιμες παρατηρήσεις του πάνω στο κείμενο και τον Μάριο Συντιχάκη για την ανυπολόγιστη βοήθειά του σε τεχνικά θέματα της συγγραφής.

Θέλω τέλος να ευχαριστήσω τους γονείς μου, Γιώργο και Δέσποινα, για την εμπιστοσύνη τους, το κουράγιο και τη στήριξη που μου πρόσφεραν όλα αυτά τα χρόνια. Σ' αυτούς θέλω να αφιερώσω αυτή την εργασία.

Στους γονείς μου

Πλοήγηση robot σε εσωτερικούς χώρους με χρήση οπτικής πληροφορίας

Σάββας Βελισσάρης
Μεταπτυχιακή Εργασία

Τμήμα Επιστήμης Υπολογιστών
Πανεπιστήμιο Κρήτης

Περίληψη

Η πλοήγηση ενός αυτόνομου συστήματος αποτελεί ένα από τα πιο απαιτητικά προβλήματα της ρομποτικής. Μία από τις προσεγγίσεις στο πρόβλημα αυτό βασίζεται στη χρήση οροσήμων, δηλαδή χαρακτηριστικών σημείων του χώρου που μπορούν σχετικά εύκολα να αναγνωριστούν από το αυτόνομο σύστημα και να χρησιμοποιηθούν για την αυτο-οριοθέτησή του (self-localization) μέσα στο χώρο. Ωστόσο, η πολυπλοκότητα του προβλήματος του ορισμού και της αναγνώρισης των οροσήμων, 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 σε εσωτερικούς χώρους, με χρήση οπτικών οροσήμων, όπου περιορίζουμε την αναγκαιότητα ύπαρξης προηγούμενης γνώσης. Αρχικά εξάγεται από την εικόνα που λαμβάνει το σύστημα μία ποιοτική περιγραφή του χώρου. Η περιγραφή αυτή παράγεται από το διαχωρισμό της εικόνας σε τμήματα που αντιστοιχούν στους τοίχους, το πάτωμα και την οροφή του χώρου κίνησης. Με την περιγραφή αυτή, καθορίζονται τα τμήματα της εικόνας όπου θα αναζητηθούν τα ορόσημα. Τα ορόσημα που χρησιμοποιεί το αυτόνομο σύστημα, δεν είναι προκαθορισμένα ή προσχεδιασμένα, αλλά εξάγονται από το σύστημα με μία διαδικασία εστίασης της προσοχής σε χαρακτηριστικά σημεία της εικόνας. Αυτό γίνεται με την χρήση συναρτήσεων που εκφράζουν την ανομοιογένεια της εικόνας σε κάποια περιοχή.

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 μνήμη του συστήματος, μαζί με πληροφορίες για τη σχετική θέση τους στο χώρο. Το αποτέλεσμα είναι η δημιουργία ενός τοπολογικού χάρτη του χώρου, βασισμένου στα ορόσημα που εξάχθηκαν. Κατά τη διάρκεια της πλοήγησης, το αυτόνομο σύστημα συγκρίνει τα ορόσημα που συναντά με αυτά που είναι αποθηκευμένα στη μνήμη του. Η αναγνώριση ενός οροσήμου του επιτρέπει να εντοπίσει τη θέση του στο χώρο και να καθορίσει την επόμενη κίνησή του.

Στα πλαίσια της εργασίας υλοποιήθηκαν η εξαγωγή της ποιοτικής περιγραφής του χώρου, με βάση την οπτική πληροφορία, καθώς και η εξαγωγή και αναγνώριση των οροσήμων. Η αξιολόγηση της μεθόδου έγινε με την εφαρμογή της σε μία ακολουθία εικόνων από το εσωτερικό ενός κτιρίου.

Επόπτης: Πάνος Τραχανιάς,
Επίκουρος καθηγητής,
Πανεπιστήμιο Κρήτης.

Indoor navigation of a robot using visual information

Savvas G. Velissaris
Master's Thesis

Department of Computer Science
University of Crete, Greece

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 objects in space which can be easily detected by the robot and used for its self-localization in space. However, the complexity of landmark definition 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, in which we minimize the need for previous knowledge. Initially, a qualitative description of space is extracted from the images acquired by the robot. This description is derived from the segmentation of the image in areas which correspond to the walls, the floor and the ceiling of the workspace. The areas of the image where the landmarks will be searched for are defined by this description. The landmarks used by the robot are not predefined or predesigned, but rather they are extracted by the robot using a focus of attention procedure on salient image areas. This is accomplished by using functions that represent the *business* of an image area.

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 memory, along with information about their position in space. As a result, a topological map of the workspace is created, based on the extracted landmarks. During navigation, the robot compares the landmarks that it encounters with the ones stored in memory. The identification of a landmark allows the robot to approximately localize itself and select its next movement.

In this thesis we implemented the extraction of the qualitative description of space,

based on visual information and the extraction and recognition of landmarks. The method has been evaluated through its application to an image sequence obtained from the corridors of a building.

Advisor: Panos Trahanias
Assistant Professor,
University of Crete.

Περιεχόμενα

Ευχαριστίες **i**

Περίληψη **iv**

Abstract **vii**

Περιεχόμενα **1**

Κατάλογος σχημάτων **3**

1 Εισαγωγή **5**
1.1 Γενική εισαγωγή 5
1.1.1 Ορισμός του προβλήματος 6
1.1.2 Οργάνωση της γραπτής εργασίας 7
1.2 Βιβλιογραφική ανασκόπηση 8
1.2.1 Τρισδιάστατη δομή και προγραμματισμός κίνησης 8
1.2.1.1 Εξαγωγή δομής από το πεδίο οπτικής ροής 9
1.2.1.2 Δομή γνωστή εκ των προτέρων 9
1.2.1.3 Προγραμματισμός κίνησης 11
1.2.2 Πλοήγηση με ορόσημα 12
1.2.3 Συμπεράσματα 14

2 Μέθοδος πλοήγησης με χρήση οπτικών οροσήμων **17**
2.1 Εισαγωγή 17
2.2 Παραδοχές 19
2.3 Γενική επισκόπηση της μεθόδου 21
2.4 Ποιοτική τμηματοποίηση χώρου 22
2.4.1 Στατική τμηματοποίηση χώρου 23
2.4.2 Δυναμική τμηματοποίηση χώρου 24
2.4.3 Έλεγχος επιτυχίας 25
2.5 Ορόσημα 26

2.5.1 Εξαγωγή οροσήμων 26
2.5.2 Αναγνώριση οροσήμων 28
2.5.2.1 Μετασχηματισμός θέσης παρατηρητή (ΜΘΠ) . . . 29
2.5.2.2 Διγραμμική παρεμβολή 35
2.5.2.3 Σύγκριση οροσήμων 35
2.6 Στάδιο εκμάθησης 36
2.7 Πλοήγηση 38

3 Υλοποίηση **41**
3.1 Εισαγωγή 41
3.2 Στατική τμηματοποίηση χώρου 41
3.2.1 Εξαγωγή ακμών 42
3.2.2 Λέπτυνση ακμών 42
3.2.3 Αφαίρεση μικρών αντικειμένων 42
3.2.4 Εύρεση ευθειών 43
3.2.5 Τμηματοποίηση εικόνας 44
3.3 Δυναμική τμηματοποίηση χώρου 45
3.3.1 Χρήση προηγούμενης γνώσης 46
3.3.2 Παραγώγιση 47
3.3.3 Από τα σημεία στην ευθεία 48
3.3.3.1 Least Median of Squares 48
3.3.3.2 Η χρήση του LMedS στη μέθοδο μας 51
3.3.4 Τμηματοποίηση εικόνας 52
3.4 Ορόσημα 52
3.4.1 Συναρτήσεις saliency map 53
3.4.2 Αλγόριθμος αύξησης περιοχής 54

4 Αποτελέσματα **57**
4.1 Εισαγωγή 57
4.2 Αποτέλεσμα 1 58
4.3 Αποτέλεσμα 2 63

5 Συμπεράσματα - Μελλοντική εργασία **67**

Βιβλιογραφία **69**

Κατάλογος Σχημάτων

2.1 Χώρος κίνησης 19
2.2 Αναμενόμενη όψη 20
2.3 Σημείο διαφυγής 20
2.4 Πλοήγηση με χρήση οπτικών οροσήμων 22
2.5 Τμηματοποίηση χώρου 23
2.6 Χρήση προηγούμενης εικόνας 25
2.7 Θέση οροσήμου - πληροφορία για τη θέση του παρατηρητή 28
2.8 Φαινομενικό πλάτος διαδρόμου 29
2.9 Κάτοψη χώρου 30
2.10 Διαφορετικές θέσεις παρατηρητή 32
2.11 Μέτρηση βάθους από το φαινομενικό πλάτος του χώρου 33
2.12 Έλεγχος παραλληλίας αξόνων 34
2.13 Υπολογισμός φωτεινότητας σημείων στο ακέραιο πλέγμα 36
2.14 Διαδικασία εκμάθησης 37
2.15 Τοπολογικός χάρτης 37
2.16 Διαδικασία πλοήγησης 38
2.17 Ενημέρωση γνώσης 39

3.1 Φίλτρο Sobel - κάθετη και οριζόντια παραγώγιση 42
3.2 Διαδοχικές εφαρμογές του AHT 44
3.3 Τελικό αποτέλεσμα 45
3.4 Τριπλό σημείο τομής 45
3.5 Απαραίτητη προηγούμενη πληροφορία 46
3.6 Παραγώγιση με sub-pixel υπολογισμό 47
3.7 Σύγκριση μεθόδου ελαχίστων τετραγώνων (LS) και Least Median of Squares (LMedS) 50
3.8 Αλγόριθμος αύξησης περιοχής 55

4.1 Αρχική επεξεργασία εικόνας 58
4.2 Μετασχηματισμός Hough 59
4.3 Δυναμική τμηματοποίηση 60

4.4 Saliency map και εξαγωγή οροσήμου (1η εικόνα) 61
4.5 Saliency map και εξαγωγή οροσήμου (2η εικόνα) 62
4.6 Ορόσημα 63
4.7 Στατική και δυναμική τμηματοποίηση 64
4.8 Saliency map 65

Κεφάλαιο 1

Εισαγωγή

1.1 Γενική εισαγωγή

Ένας από τους σημαντικότερους στόχους της ρομποτικής είναι η κατασκευή αυτόνομων μηχανών με δυνατότητες να εκτελούν σύνθετες εργασίες, με αποτελεσματικότητα αλλά και ευελιξία σχετικά με οποιαδήποτε πιθανή κατάσταση. Είναι προφανές πως μία τέτοια πολύπλοκη συμπεριφορά, που δεν διαφέρει πολύ από την έλλογη συμπεριφορά πολλών βιολογικών οργανισμών, δεν μπορεί να επιτευχθεί χωρίς πρώτα να ερευνηθούν και να αναπτυχθούν οι απλούστερες λειτουργίες στις οποίες στηρίζεται.

Στην παρούσα εργασία θα ασχοληθούμε με μία από τις πιο σπουδαίες λειτουργίες μιας αυτόνομης μηχανής, την αυτόνομη πλοήγηση (autonomous navigation). Με τον όρο αυτό εννοούμε την ικανότητα κάποιου συστήματος να κινείται με επιτυχία (π.χ. αποφυγή εμποδίων, επιλογή κατάλληλου μονοπατιού, επίτευξη στόχου κίνησης, κλπ) μέσα στο περιβάλλον του, χρησιμοποιώντας πληροφορίες που συλλέγει μέσω αισθητήρων. Το πόσο σημαντική είναι για ένα αυτόνομο σύστημα η αυτόνομη πλοήγηση, φαίνεται από τη ανάπτυξη αυτής της ικανότητας στους περισσότερους από τους ζωντανούς οργανισμούς. Όλες σχεδόν οι μορφές ζωής χρησιμοποιούν την πλοήγηση για βασικές λειτουργίες τους, όπως η αναζήτηση τροφής, η επιστροφή στην κατοικία, η μετανάστευση, αλλά και η εξερεύνηση καινούριων περιοχών του περιβάλλοντός τους.

Ιδιαίτερα σημαντικό είναι όλες οι παραπάνω λειτουργίες να είναι δυνατές σε ποικίλα και μεταβαλλόμενα περιβάλλοντα. Η πλοήγηση δεν πρέπει να εμποδίζεται από αλλαγές, ή να αποτυγχάνει σε καινούριες καταστάσεις. Από αυτό γίνεται σαφές ότι η αυτόνομη πλοήγηση είναι γενικότερη έννοια από αυτή του προγραμματισμού κίνησης (motion planning). Ο προγραμματισμός κίνησης αναφέρεται στο σχεδιασμό της ακολουθίας κινήσεων που πρέπει να πραγματοποιηθούν, έχοντας δεδομένη τη θέση του συστήματος και γνωρίζοντας εκ των

προτέρων το χώρο όπου θα πραγματοποιηθεί η κίνηση (π.χ. χάρτες). Η αυτόνομη πλοήγηση αντίθετα, αντιμετωπίζει ταυτόχρονα τα προβλήματα της εκτίμησης της ιδίας θέσης και κίνησης (egomotion estimation) του συστήματος, της εξαγωγής πληροφοριών από το περιβάλλον (τμηματοποίηση του χώρου, εκτίμηση αποστάσεων, αναγνώριση εμποδίων ή γνωστών χαρακτηριστικών) και τελικά την αξιολόγηση και χρήση αυτών των πληροφοριών για την πραγματοποίηση της κίνησης.

1.1.1 Ορισμός του προβλήματος

Ο σχεδιασμός συστημάτων αυτόνομης πλοήγησης διαχωρίζεται σε δύο κυρίως κατηγορίες, ανάλογα με τα περιβάλλοντα εργασίας των συστημάτων : πλοήγηση σε εξωτερικούς (outdoor) και εσωτερικούς (indoor) χώρους. Η πλοήγηση σε εσωτερικούς χώρους, με την οποία ασχολείται αυτή η εργασία, μπορεί να χρησιμοποιηθεί σε πολλές εφαρμογές, όπως εργασία αυτόματων συστημάτων σε δημόσιες υπηρεσίες, νοσοκομεία, εργοστάσια, επίβλεψη και ασφάλεια κτιρίων ή παροχή βοήθειας σε άτομα με ειδικές ανάγκες.

Οι ικανότητες που θα πρέπει να διαθέτουν τα συστήματα αυτά προκειμένου να ανταποκρίνονται σε αντικειμενικές απαιτήσεις, δεν είναι ούτε λίγες, ούτε εύκολο να επιτευχθούν. Θα πρέπει να υπάρχει η ικανότητα αντίληψης της κίνησης του ίδιου του συστήματος, όπως και η ικανότητα αντίληψης ανεξάρτητης κίνησης (independent motion estimation) [Arg96, Par96]. Απαραίτητες είναι επίσης οι ικανότητες αποφυγής εμποδίων (obstacle avoidance) και κίνησης από σημείο σε σημείο (homing). Ένα σύστημα αυτόματης πλοήγησης θα έπρεπε ακόμα να είναι ικανό να διαθέτει κάποια μορφή αντίληψης του περιβάλλοντος (space perception) καθώς και την ικανότητα να ”απομνημονεύει” κάποια χαρακτηριστικά του χώρου και να χρησιμοποιεί τεχνικές εξαγωγής συμπερασμάτων (reasoning) ώστε να διατηρεί την ευελιξία του σε μία πληθώρα διαφορετικών καταστάσεων. Όλα τα παραπάνω αποτελούν σήμερα ανοιχτά θέματα στην έρευνα και το σχεδιασμό συστημάτων αυτόνομης πλοήγησης.

Όπως αναφέρθηκε, η εξαγωγή των απαραίτητων πληροφοριών από το περιβάλλον γίνεται με τη χρήση αισθητήρων. Η φύση έχει να παρουσιάσει μία μεγάλη ποικιλία δυνατοτήτων γι’ αυτό το σκοπό, όπως τη χρήση της όρασης, της ακοής, της αφής, της όσφρησης, ή ακόμη και τη χρήση υπερήχων. Ο άνθρωπος έχει επίσης χρησιμοποιήσει για την πλοήγηση οχημάτων εκπομπές ραδιοκυμάτων. Στην εργασία αυτή θα θεωρήσουμε ότι γίνεται χρήση οπτικών αισθητήρων. Η μεγάλη διάδοση της όρασης στη φύση και η μεγάλη ποσότητα πληροφοριών που μπορεί να αντληθεί από αυτή, έχουν φέρει δικαιολογημένα την όραση στην πρώτη θέση ανάμεσα στις προτιμήσεις των ερευνητών συστημάτων αυτόνομης πλοήγησης. Θα επικεντρώσουμε λοιπόν το ενδιαφέρον μας

στην **οπτική πλοήγηση** (**visual navigation**), την εξαγωγή δηλαδή πληροφοριών από μία ακολουθία εικόνων, κατάλληλων για τον έλεγχο των κινήσεων ενός μηχανικού συστήματος, προς την επίτευξη ενός συγκεκριμένου στόχου κίνησης μέσα στο χώρο.

Σκοπός της παρούσης εργασίας είναι η περιγραφή και αξιολόγηση μιας μεθόδου οπτικής πλοήγησης που δεν προϋποθέτει λεπτομερή προηγούμενη γνώση για το περιβάλλον. Η μόνη προϋπόθεση είναι η κίνηση του αυτόνομου συστήματος να πραγματοποιείται σε εσωτερικούς χώρους που έχουν τη δομή ορθογώνιου παραλληλεπιπέδου. Η απαίτηση αυτή ικανοποιεί την πλειοψηφία των εσωτερικών χώρων κτιρίων και οικιών, κι έτσι δεν περιορίζει αισθητά τις δυνατότητες εφαρμογής της μεθόδου. Πέρα από αυτή τη γενική παραδοχή, 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. Η αξιολόγηση της μεθόδου γίνεται με την εφαρμογή της σε μία ακολουθία εικόνων από τους διαδρόμους ενός κτιρίου.

1.1.2 Οργάνωση της γραπτής εργασίας

Μετά τη συνοπτική περιγραφή του περιεχομένου της εργασίας αυτής, στη συνέχεια του **κεφαλαίου 1** το πρόβλημα εντάσσεται στην ευρύτερη ερευνητική περιοχή. Παρουσιάζονται και αξιολογούνται διάφορες προσεγγίσεις που παρουσιάζουν ομοιότητες αλλά και διαφορές με την προτεινόμενη μέθοδο.

Στο **κεφάλαιο 2** γίνεται αναλυτική περιγραφή της προτεινόμενης προσέγγισης στο πρόβλημα της αυτόνομης πλοήγησης. Παρουσιάζονται οι τρόποι με τους οποίους το αυτόνομο σύστημα συλλέγει πληροφορίες για τη δομή του χώρου και τα αντικείμενα που υπάρχουν σ' αυτόν, καθώς και η χρήση αυτών των πληροφοριών για την πλοήγηση.

Στο **κεφάλαιο 3** παρουσιάζονται οι συγκεκριμένες τεχνικές επεξεργασίας της εικόνας που χρησιμοποιήθηκαν για την υλοποίηση της μεθόδου.

Στο **κεφάλαιο 4** παρουσιάζονται αποτελέσματα που προέκυψαν από την εφαρμογή της μεθόδου σε μία ακολουθία εικόνων.

Στο **κεφάλαιο 5** παρατίθενται τα συμπεράσματα καθώς και οι πιθανές επεκτάσεις της εργασίας αυτής.

Τέλος, ακολουθεί η βιβλιογραφία που χρησιμοποιήθηκε.

1.2 Βιβλιογραφική ανασκόπηση

Στο τμήμα αυτό της εργασίας γίνεται μία ανασκόπηση των μέχρι τώρα ερευνητικών προσπαθειών που έχουν γίνει για τη σχεδίαση συστημάτων αυτόματης πλοήγησης και των επιμέρους χαρακτηριστικών τους. Η ανασκόπηση επικεντρώνεται στην οπτική πλοήγηση σε εσωτερικούς χώρους, που είναι και το θέμα της συγκεκριμένης εργασίας.

Η κλασική προσέγγιση μέχρι τώρα στο πρόβλημα της οπτικής πλοήγησης συνίσταται στο να το διαχωρίζει σε δύο ξεχωριστά προβλήματα : το πρώτο έχει να κάνει με την κατασκευή μιας τρισδιάστατης, γεωμετρικής περιγραφής του χώρου μέσω μιας ακολουθίας εικόνων που παράγονται κατά την κίνηση του αυτόνομου συστήματος μέσα στο χώρο. Το δεύτερο πρόβλημα είναι η χρήση αυτής της περιγραφής για τον υπολογισμό μιας ακολουθίας κινήσεων που να επιτελεί τον σκοπό του συστήματος. Το πρώτο πρόβλημα είναι γνωστό σαν δομή από κίνηση (structure from motion) και το δεύτερο σαν πρόβλημα προγραμματισμού κίνησης (motion planning).

Μία εναλλακτική προσέγγιση είναι βασισμένη στην έννοια των οροσήμων (landmarks). Με τον όρο ορόσημο εννοούμε κάποιο ιδιαίτερο χαρακτηριστικό του κόσμου, που ξεχωρίζει από το υπόλοιπο περιβάλλον και είναι εύκολο να αναγνωριστεί από το αυτόνομο σύστημα. Στην προσέγγιση αυτή, η γεωμετρική περιγραφή του χώρου δεν είναι απαραίτητη. Οι σχετικές θέσεις των οροσήμων 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 και τις ενέργειες που πρέπει να εκτελέσει για την ολοκλήρωση μιας κίνησης.

Οι τρόποι με τους οποίους έχουν αντιμετωπισθεί τα παραπάνω θέματα παρουσιάζονται αναλυτικά παρακάτω.

1.2.1 Τρισδιάστατη δομή και προγραμματισμός κίνησης

Χαρακτηριστικό αυτής της προσέγγισης στην πλοήγηση είναι ότι στοχεύει σε μία ακριβή αναπαράσταση του χώρου και τη χρήση καθαρά μετρικών μεθόδων για τον προγραμματισμό κίνησης. Η λεπτομερειακή 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.

1.2.1.1 Εξαγωγή δομής από το πεδίο οπτικής ροής

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

- Αρχικά υπολογίζεται η μετατόπιση των προβολών κάποιων σημείων μεταξύ συνεχόμενων εικόνων, η οποία οφείλεται στην κίνηση του παρατηρητή. Η μετατόπιση υπολογίζεται μέσω της αντιστοίχισης (correspodence) τέτοιων σημείων (ή ομάδων σημείων όπως γωνίες και ευθύγραμμα τμήματα). Έτσι προκύπτει μία συνάρτηση δύο χωρικών μεταβλητών (οι δύο διαστάσεις της εικόνας) και μίας χρονικής (η χρονική ακολουθία των εικόνων). Οι μερικές παράγωγοι αυτής της συνάρτησης (σε συνδυασμό με κάποιες υποθέσεις για την κίνηση) δίνουν το διάνυσμα της ταχύτητας κάθε σημείου στο επίπεδο της εικόνας (image plane). Το διανυσματικό πεδίο που προκύπτει είναι γνωστό σαν πεδίο οπτικής ροής (optical flow field)[TL94].

- Στη συνέχεια, υπολογίζεται η τρισδιάστατη ταχύτητα και η δομή της σκηνής, μέσω των εξισώσεων που συσχετίζουν την τρισδιάστατη ταχύτητα των σημείων της σκηνής με τη διδιάστατη ταχύτητα των αντίστοιχων προβολών τους πάνω στην εικόνα.

Το πρόβλημα της δομής από κίνηση έχει μελετηθεί διεξοδικά σε θεωρητικό επίπεδο [Ull79, SA89]. Έτσι είναι γνωστά τα προβλήματα που παρουσιάζει η παραπάνω προσέγγιση. Το κυριότερο αφορά την υπολογιστική πολυπλοκότητα που παρουσιάζει το πρόβλημα της αντιστοίχισης, αφού πρέπει να εφαρμοστούν κριτήρια ομοιότητας σε μεγάλα τμήματα της εικόνας. Ακόμα χειρότερα, μπορεί κάποια χαρακτηριστικά της μιας εικόνας να μην υπάρχουν στην επόμενη, όπως στην περίπτωση παρεμβολής κάποιου αντικειμένου. Τέλος, οι επιπλέον περιορισμοί που απαιτούνται για την εξασφάλιση της μοναδικότητας της λύσης (π.χ. ομαλότητα του πεδίου ροής) συνήθως δεν ανταποκρίνονται στην πραγματικότητα και οδηγούν τις τεχνικές αυτές σε κακές αποδόσεις.

1.2.1.2 Δομή γνωστή εκ των προτέρων

Με βάση τα παραπάνω μειονεκτήματα, αρκετοί ερευνητές έχουν προσεγγίσει το πρόβλημα διαφορετικά. Μία τέτοια προσέγγιση είναι να θεωρήσουμε ότι κάποιο μοντέλο του κόσμου είναι γνωστό εκ των προτέρων (a priori knowledge). Έτσι η βασική στρατηγική συνίσταται στο να αντιληφθεί το σύστημα το περιβάλλον

με τη χρήση των αισθητήρων και να συγκριθεί η πληροφορία αυτή με την προϋπάρχουσα γνώση για τον κόσμο. Εάν επιτευχθεί ακριβές ταίριασμα (matching), καταλήγουμε με κάποια ακρίβεια στην εκτίμηση της θέσης του συστήματος και κατ' επέκταση σε μια περιγραφή του άμεσου περιβάλλοντος. Το πρόβλημα της περιγραφής του περιβάλλοντος σε αυτή την προσέγγιση ταυτίζεται με το πρόβλημα του εντοπισμού της θέσης του συστήματος (self-location), αφού ξέροντας την θέση και έχοντας μία προηγούμενη περιγραφή του κόσμου, μπορούμε να βγάλουμε συμπεράσματα σχετικά με 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 καλύτερα τον κόσμο.

Ένας τρόπος αναπαράστασης του μοντέλου του κόσμου είναι με χάρτες [LPM84, VLS90, Sug88] ή άλλες απλές γεωμετρικές αναπαραστάσεις [Lat91]. Σε αυτές τις εργασίες έχουν αναπτυχθεί μέθοδοι για την εκτίμηση της θέσης του αυτόνομου συστήματος, μέσα στο γνωστό μοντέλο του κόσμου[TL92, TK96]. Μετά από αυτή την εκτίμηση, το πρόβλημα της πλοήγησης αντιμετωπίζεται με μετρικές μεθόδους.

Μία ανάλογη τεχνική χρησιμοποιήθηκε από τους Kak et.al.[AKLA89, KK92] στη δημιουργία του αυτόνομου συστήματος πλοήγησης FINALE. Εδώ είναι γνωστή μία τρισδιάστατη περιγραφή του χώρου, από την οποία μπορεί να προκύψει η όψη του κόσμου, όπως θα φαινόταν από μία υποθετική θέση του αυτόνομου συστήματος. Το αυτόνομο σύστημα κάνει μία εκτίμηση της θέσης του με βάση την προηγούμενη και τις εντολές κίνησης που ακολουθήθηκαν ενδιάμεσα (η εκτίμηση δεν μπορεί να είναι ακριβής λόγω των ατελειών στα συστήματα κίνησης). Με βάση αυτή την εκτίμηση και το γνωστό μοντέλο του κόσμου, προκύπτει μία αναμενόμενη όψη του περιβάλλοντος. Στη συνέχεια αυτή η εικόνα συγκρίνεται με την πραγματική εικόνα που συλλαμβάνουν οι κάμερες του συστήματος. Το ταίριασμα των εικόνων έχει σαν αποτέλεσμα να πάρουμε μία εκτίμηση της θέσης του συστήματος με μειωμένη αβεβαιότητα.

Οι Tsubouchi et.al. [TY87] χρησιμοποιούν επίσης την παραπάνω προσέγγιση, βασιζόμενοι όμως κυρίως στο χρώμα των εικόνων, υποστηρίζοντας πως αυτό το χαρακτηριστικό δεν επηρεάζεται από τη φωτεινότητα ή ενδεχόμενη σκιά στις εικόνες.

Τέλος, σημαντικές πληροφορίες για την τρισδιάστατη δομή του κόσμου μπορούν να αντληθούν από το σημείο διαφυγής [PP96] και από άλλες ευθείες γραμμές της εικόνας [CA95], προσέγγιση που χρησιμοποιείται και στην παρούσα εργασία. Οι πληροφορίες αυτές συλλέγονται με απλές τεχνικές ανάλυσης εικόνων, όπως ανίχνευση ακμών, μετασχηματισμός Hough, και μπορούν να

οδηγήσουν σε μία διαχώριση του χώρου σε περιοχές, ποιοτικά διαφορετικές μεταξύ τους (τοίχοι, πάτωμα κλπ).

1.2.1.3 Προγραμματισμός κίνησης

Έχοντας δεδομένη μία γεωμετρική περιγραφή του χώρου, ο προγραμματισμός της κίνησης του αυτόνομου συστήματος από μία αρχική θέση σε μία άλλη, συνίσταται στον υπολογισμό ενός ελεύθερου μονοπατιού που να συνδέει τις δύο θέσεις [HA92, Lat91]. Το πρόβλημα αυτό είναι γνωστό σαν "piano movers problem" ή "generalised movers problem"[SS83]. Οι μέθοδοι που έχουν προταθεί μπορούν να καταταγούν στις γενικές κατηγορίες της υπολογιστικής γεωμετρίας [PS85, Nil69, LPW79, RW87], του δυναμικού πεδίου [Til90] και του μαθηματικού προγραμματισμού [FT87].

Στις μεθόδους υπολογιστικής γεωμετρίας, τα πιθανά μονοπάτια απεικονίζονται σε ένα γράφο ή δίκτυο, κι έτσι ο υπολογισμός μιας ακολουθίας κινήσεων ανάγεται σε ένα πρόβλημα αναζήτησης σε γράφους [Nil69, COY84, Can88]. Πολλές τεχνικές αναζήτησης σε γράφους, όπως 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 περιβάλλει, ενώ ο στόχος ασκεί αντίστοιχες ελκτικές δυνάμεις[Til90, Kha86]. Η ένταση των δυνάμεων αυτών μεταβάλλεται ανάλογα με την απόσταση από την πηγή τους. Παίρνοντας τη συνισταμένη των δυνάμεων αυτών για κάθε σημείο του χώρου, δημιουργείται ένα διανυσματικό πεδίο, από το οποίο ένα σύστημα μπορεί να εξάγει τη σωστή κατεύθυνση κίνησης υπολογίζοντας την κλίση (gradient) του πεδίου στο σημείο στο οποίο βρίσκεται. Τέλος, στη μέθοδο του μαθηματικού προγραμματισμού, ο υπολογισμός μιας ακολουθίας κινήσεων ανάγεται σε πρόβλημα βελτιστοποίησης κάποιας συνάρτησης, που προκύπτει από το περιβάλλον και τα αντικείμενα που υπάρχουν μέσα σ' αυτό [FT87].

Το βασικό πλεονέκτημα των παραπάνω προσεγγίσεων είναι ότι τα μονοπάτια που υπολογίζονται είναι βέλτιστα σύμφωνα με κάποιο κριτήριο, όπως ελάχιστο μήκος διαδρομής, απόσταση από εμπόδια κ.ά. Έχουν όμως τα βασικά μειονεκτήματα της αναγκαιότητας ύπαρξης μιας ακριβούς αναπαράστασης του κόσμου και του μεγάλου υπολογιστικού κόστους, που κάνει τις παραπάνω προσεγγίσεις μη ρεαλιστικές για εφαρμογή σε περίπλοκα, άγνωστα ή δυναμικά μεταβαλλόμενα περιβάλλοντα. Επιπλέον, 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[TL92, TK96].

Τα μειονεκτήματα αυτά έχουν οδηγήσει σε άλλες προσεγγίσεις που δεν απαιτούν μία ακριβή και λεπτομερή αναπαράσταση του κόσμου. Στις προσεγγίσεις αυτές, που υιοθετούν τις αρχές της θεωρίας της ενεργούς και τελεολογικής όρασης (active and purposive vision) [AB87, Alo90, Bal91], δίνεται προτεραιότητα στα πιο "χαρακτηριστικά" σημεία του περιβάλλοντος (τα προαναφερθέντα ορόσημα). Μετά τον εντοπισμό τους από το αυτόνομο σύστημα, τα ορόσημα μπορούν να χρησιμοποιηθούν για τη μείωση της αβεβαιότητας της θέσης του συστήματος [LL94] και για την πλοήγησή του, που πλέον γίνεται ποιοτικά (π.χ. στο ορόσημο Α στρίψε δεξιά). Τα ζητήματα ωστόσο που προκύπτουν από αυτή την προσέγγιση είναι πολλά και αφορούν τον τρόπο επιλογής των κατάλληλων οροσήμων, το είδος τους και τους τρόπους αναγνώρισης από το αυτόνομο σύστημα.

1.2.2 Πλοήγηση με ορόσημα

Με τον όρο ορόσημο (landmark) εννοούμε κάποιο ιδιαίτερο χαρακτηριστικό του κόσμου, το οποίο, λόγω της ιδιαιτερότητάς του μας δίνει τη δυνατότητα να αντλήσουμε πληροφορίες σχετικές με τα προβλήματα που αναφέρθηκαν παραπάνω. Αναγνωρίζοντας ένα ορόσημο, μπορούμε να χρησιμοποιήσουμε την πληροφορία, είτε για να τμηματοποιήσουμε μερικώς τον κόσμο, είτε για να βοηθηθούμε στον προσδιορισμό της θέσης του αυτόνομου συστήματος και κατ' επέκταση στον προγραμματισμό κίνησης. Ορόσημα, στην περίπτωση που ασχολούμαστε με οπτικούς αισθητήρες, μπορούν να είναι μία μεγάλη ποικιλία από οπτικά ερεθίσματα, όπως χαρακτηριστικά αντικείμενα, απλές γωνίες ή ακμές, ή ακόμα και ειδικά αντικείμενα που έχουν τοποθετηθεί στο περιβάλλον του συστήματος σκόπιμα, ώστε να το βοηθήσουν στην πλοήγηση. Γίνεται φανερό λοιπόν, ότι η επιλογή του κατάλληλου οροσήμου δεν είναι απλή υπόθεση[GI96]. Αρκετά δύσκολο είναι επίσης το πρόβλημα του εντοπισμού των οροσήμων στις εικόνες που λαμβάνονται μέσω των οπτικών αισθητήρων. Είναι δυνατόν να διαχωρίσουμε τα οπτικά ορόσημα, όπως έχουν εμφανιστεί στη βιβλιογραφία, στις ακόλουθες κατηγορίες :

- Προσχεδιασμένα ορόσημα (predesigned landmarks). Στην προσέγγιση αυτή σαν ορόσημα χρησιμοποιούνται προκαθορισμένα πρότυπα (patterns) όπως bar codes [TK94] ή άλλα απλά γεωμετρικά σχήματα [LL95, LL94, MA95]. Τα πρότυπα αυτά σχεδιάζονται έτσι ώστε να είναι εύκολα αναγνωρίσιμα

και να ξεχωρίζουν από τα υπόλοιπα αντικείμενα του περιβάλλοντος.

Οι Taylor et.al.[TK94] χρησιμοποιούν σαν ορόσημα bar codes, που αποτελούνται το καθένα από πέντε κάθετες λωρίδες, που μπορεί να έχουν ένα από τρία διαφορετικά χρώματα. Η αναγνώριση ενός οροσήμου γίνεται εύκολα με την εξαγωγή των κάθετων ακμών της εικόνας και τα χρώματα χρησιμοποιούνται για τον διαχωρισμό των διαφορετικών οροσήμων. Στην εργασία αυτή προτείνεται ένας αλγόριθμος ”εξερεύνησης” του χώρου, όπου τα γνωστά ορόσημα και ένας αισθητήρας υπερήχων χρησιμοποιούνται για τη δημιουργία ενός γράφου που αναπαριστά τα εμπόδια του περιβάλλοντος. Ο γράφος αυτός μπορεί να χρησιμοποιηθεί στη συνέχεια για την πλοήγηση του αυτόνομου συστήματος.

Οι Lazanas et.al.[LL94] 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, όπως γνωστό είναι και το 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 και ατέλειες του μηχανικού ελέγχου του συστήματος.

Οι Magge et.al.[MA95] ασχολούνται με το πρόβλημα του εντοπισμού της θέσης του αυτόνομου συστήματος. 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.

Το βασικό μειονέκτημα των παραπάνω προσεγγίσεων είναι προφανώς η αδυναμία εφαρμογής των σε άγνωστα περιβάλλοντα.

- Επιλεγμένα ορόσημα (selected landmarks). Σε αυτή την προσέγγιση τα ορόσημα επιλέγονται μέσα από το περιβάλλον από το σχεδιαστή του αυτόνομου συστήματος [NB88]. Έτσι το σύστημα μπορεί να χρησιμοποιήσει για την πλοήγηση αντικείμενα που υπάρχουν ήδη στο χώρο, ο οποίος όμως πρέπει να είναι γνωστός εκ των προτέρων. Το κριτήριο επιλογής των κατάλληλων οροσήμων είναι και πάλι το πόσο ”ξεχωριστά” (distinctive) είναι, ώστε να μπορούν να περιγράψουν το περιβάλλον ικανοποιητικά.

Για παράδειγμα, οι Yeh et.al[YK95], από ένα καθορισμένο σύνολο οπτικών χαρακτηριστικών, επιλέγουν ένα ορόσημο έτσι ώστε 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. Σαν ορόσημο επιλέγεται το χαρακτηριστικό εκείνο που ελαχιστοποιεί το κριτήριο του Bayes.

- Ευθείες γραμμές(straight lines). Σαν ορόσημα μπορούν να θεωρηθούν και οι ευθείες γραμμές [Sug88, KP92, YK95], αφού κυριαρχούν στους εσωτερικούς χώρους και μπορούν να χρησιμοποιηθούν για την περιγραφή του περιβάλλοντος. Η χρήση των ευθειών σε ορισμένες περιπτώσεις έχει καλά αποτελέσματα, σε χώρους όπως διάδρομοι ή γραφεία. Δεν μπορούμε όμως να θεωρήσουμε την προσέγγιση αυτή κατάλληλη για γενικότερους χώρους, λόγω της εγγενούς ασάφειας που υπάρχει στην περιγραφή με ευθείες γραμμές.

- Πανοραμική άποψη(panoramic view). Η προσέγγιση αυτή βασίζεται στην αναπαράσταση πανοραμικής άποψης του χώρου, δηλ. στην κράτηση σε μια μνήμη οπτικής πληροφορίας όλης της διαδρομής που διανύθηκε κατά την εκμάθηση [ZBT91, Zhe89]. Τον ρόλο των οροσήμων εδώ παίζουν ολόκληρα τμήματα από την παραπάνω αναπαράσταση. Το αυτόνομο σύστημα σε αυτή την προσέγγιση, χρησιμοποιεί για την πλοήγηση κάποιες συνολικές απόψεις "χαρακτηριστικών" τμημάτων της διαδρομής, και όχι επιμέρους αντικείμενα όπως στις προηγούμενες προσεγγίσεις.

1.2.3 Συμπεράσματα

Από τα παραπάνω γίνεται φανερό ότι το πρόβλημα της αυτόνομης πλοήγησης έχει αντιμετωπισθεί αποσπασματικά, και οι ερευνητές που έχουν ασχοληθεί με αυτό έχουν κάνει πολλές παραδοχές για να στηρίξουν την προσπάθεια τους. Ο λόγος είναι η μεγάλη δυσκολία που παρουσιάζει το επιχείρημα να αποδώσουμε σε ένα μηχανικό σύστημα πραγματικές ικανότητες αντίληψης του χώρου και λήψης αποφάσεων για την πλοήγηση μέσα σ' αυτόν.

Μία συχνή παραδοχή είναι η εκ των προτέρων γνώση του χώρου κίνησης του αυτόνομου συστήματος. Αυτή η παραδοχή περιορίζει αισθητά τις δυνατότητες χρησιμοποίησης των αντίστοιχων αυτόνομων συστημάτων μόνο σε περιβάλλοντα επακριβώς γνωστά και όχι σε περιβάλλοντα άγνωστα ή μεταβαλλόμενα. Επίσης, οι αντίστοιχοι αλγόριθμοι προγραμματισμού κίνησης είναι χρονοβόροι και μπορεί να παρουσιάσουν προβλήματα όταν υπάρχει ατελής μηχανικός

έλεγχος του συστήματος. Όλα αυτά δείχνουν ότι η ακριβής αναπαράσταση του περιβάλλοντος αποτελεί μία προσέγγιση με μειονεκτήματα. Είναι επίσης μία προσέγγιση μη ρεαλιστική, αφού θα απαιτούσε από το αυτόνομο σύστημα εξελιγμένες ικανότητες αντίληψης που είναι πολύ δύσκολο να επιτευχθούν στην πράξη.

Οι πιο ελπιδοφόρες προσπάθειες είναι αυτές που στοχεύουν μάλλον σε μια *ποιοτική* περιγραφή του χώρου, και η χρήση των οροσήμων σαν μέσον για μια τέτοια περιγραφή έχει φανεί ότι υπόσχεται πολλά. Παραμένει όμως ανοιχτό το πρόβλημα του καθορισμού του είδους των οροσήμων καθώς και της εύρεσης ικανοποιητικών τεχνικών αναγνώρισης τους. Είναι φανερό ότι αν ο τελικός μας στόχος είναι ένα μηχανικό σύστημα ικανό για κίνηση μέσα σε πολύπλοκα και δυναμικά μεταβαλλόμενα περιβάλλοντα, η χρήση προσχεδιασμένων ή επιλεγμένων οροσήμων δεν αρκεί, αφού περιορίζει την χρήση του συστήματος σε συγκεκριμένους χώρους εργασίας.

Κεφάλαιο 2

Μέθοδος πλοήγησης με χρήση οπτικών οροσήμων

2.1 Εισαγωγή

Στην εργασία αυτή προτείνουμε μία προσέγγιση στο πρόβλημα της αυτόνομης οπτικής πλοήγησης σε εσωτερικούς χώρους με χρήση οπτικών οροσήμων. Η προσέγγιση μας εντάσσεται στη θεωρία της ενεργούς και τελεολογικής όρασης [AB87, Alo90], σύμφωνα με την οποία οι διαδικασίες όρασης που διαθέτει ένα σύστημα 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. Στην προσέγγισή μας, η περιγραφή του περιβάλλοντος είναι ένας συνδυασμός μιας *ποιοτικής* περιγραφής της δομής του χώρου και των *οροσήμων* που υπάρχουν σε αυτόν. Με τον όρο "ποιοτική περιγραφή του χώρου", εννοούμε μία διαχώρισή του σε περιοχές με διαφορετική ποιοτική πληροφορία (τοίχοι, πάτωμα, οροφή). Η πληροφορία αυτή θα χρησιμοποιηθεί για τον καλύτερο εντοπισμό πιθανών θέσεων των οροσήμων.

Τα ίδια τα ορόσημα δεν είναι προεπιλεγμένα ή προκαθορισμένα. Προσπαθώντας να διατηρήσουμε την ευελιξία του συστήματος και την ικανότητα κίνησης του μέσα σε άγνωστα περιβάλλοντα, προτείνουμε μία μέθοδο *εστίασης της προσοχής* (focus of attention) σε *προέχοντα* (salient) σημεία της εικόνας. Γι' αυτό το λόγο υιοθετούμε την έννοια του *saliency map*[CF93], όπου με χρήση διάφορων κριτηρίων εντοπίζουμε στην εικόνα τα πιο "χαρακτηριστικά" σημεία. Στα σημεία αυτά γίνεται επιλεκτική επεξεργασία στη συνέχεια. Με τον τρόπο αυτό αφήνουμε στο αυτόνομο σύστημα την επιλογή των οροσήμων, δίνοντας του την ευχέρεια να προσαρμόζεται σε αλλαγές στο περιβάλλον του.

Η μέθοδος πλοήγησης που προτείνεται περιλαμβάνει δύο στάδια, που κάνουν χρήση της ποιοτικής περιγραφής του χώρου και της επιλογής οροσήμων : το στάδιο της **εκμάθησης** και το στάδιο της **πλοήγησης**.

1. Κατά το στάδιο της εκμάθησης, το αυτόνομο σύστημα εξάγει πληροφορίες από το περιβάλλον. Εστιάζοντας την προσοχή του στα προέχοντα σημεία της εικόνας, εντοπίζει σημεία όπου μπορεί να υπάρχει ένα υποψήφιο ορόσημο. Με περαιτέρω επεξεργασία αυτών των σημείων, καθώς και τη χρήση της περιγραφής της δομής του χώρου που προαναφέρθηκε, αποσπάται ένα κομμάτι από την εικόνα που χαρακτηρίζεται σαν ορόσημο. 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 και η αντιστοίχιση των οροσήμων πάνω σ' αυτόν. Με τη βοήθεια ενός δασκάλου, που επιβλέπει τη διαδικασία της εκμάθησης, μπορεί να γίνει και αντιστοίχιση των οροσήμων με *σημαντική* (semantic) πληροφορία για τη θέση τους στο χώρο (π.χ. το ορόσημο Α βρίσκεται μέσα στο εργαστήριο). Έτσι οι εντολές πλοήγησης μπορούν να δίνονται με ευκολότερο και σαφέστερο τρόπο.

2. Στο στάδιο της πλοήγησης, πραγματοποιούνται σε κάθε εικόνα οι ίδιες εργασίες ποιοτικής αναπαράστασης του χώρου και εστίασης της προσοχής, όπως και στην επιλογή των οροσήμων κατά το στάδιο της εκμάθησης. Μετά την εξαγωγή του οροσήμου, το αυτόνομο σύστημα προσπαθεί να το *αναγνωρίσει*, συγκρίνοντας το με τα ήδη αποθηκευμένα ορόσημα. Ενδεχόμενη ταύτιση (matching), επιτρέπει στο σύστημα να εντοπίσει τη θέση του στο χώρο, και να κινηθεί.

Τα δύο στάδια δεν είναι εντελώς ανεξάρτητα. Θεωρητικά, η διαδικασία της εκμάθησης θα πραγματοποιηθεί μόνο μία φορά, και οι πληροφορίες που θα αποθηκευτούν θα χρησιμοποιούνται κάθε φορά που το αυτόνομο σύστημα εκτελεί μία εντολή πλοήγησης (navigation task). Αν όμως κατά τη διάρκεια της πλοήγησης συναντηθεί ένα καινούριο ορόσημο, 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 μπορεί να προσαρμόζεται σε καινούρια δεδομένα και αλλαγές του περιβάλλοντος.

Σημαντικό ρόλο στην προσέγγισή μας έπαιξε και η εισαγωγή της διάστασης του χρόνου στις παραπάνω μεθόδους. Εφόσον η πλοήγηση του αυτόνομου συστήματος γίνεται με χρήση οπτικής πληροφορίας, τα δεδομένα του θα είναι μία ακολουθία εικόνων και μπορούμε να απαιτήσουμε "συμβατότητα" μεταξύ

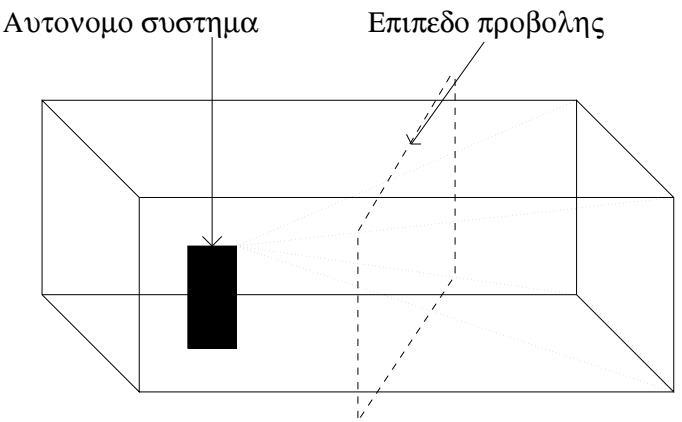
χρονικά διαδοχικών εικόνων (περιμένουμε π.χ. ένα ορόσημο να βρίσκεται σε γειτονική θέση και να έχει παρόμοια χαρακτηριστικά σε δύο διαδοχικές εικόνες). Με τον τρόπο αυτό μειώνεται το υπολογιστικό κόστος της μεθόδου, αφού μπορούμε να εστιάσουμε την προσοχή μας σε περιοχές της εικόνας όπου *αναμένουμε* τον εντοπισμό ενός οροσήμου και να μειώσουμε τους σχετικούς υπολογισμούς. Επιπλέον, τα αποτελέσματα που προκύπτουν αποκτούν μεγαλύτερη αξιοπιστία όταν συμφωνούν με αυτά των προηγούμενων εικόνων.

Ιδιαίτερη βαρύτητα έχει δωθεί σε ορισμένα επιμέρους τμήματα της μεθόδου που θα περιγραφούν αναλυτικά παρακάτω. Ένα από αυτά είναι ο συγκερασμός των διάφορων τεχνικών επεξεργασίας εικόνας που χρησιμοποιούνται, για τον προσανατολισμό τους στο δικό μας στόχο. Μεγάλη σημασία δώθηκε και στην ποιοτική τμηματοποίηση του χώρου, από την οποία προκύπτουν πολλές σημαντικές και απαραίτητες πληροφορίες. Ιδιαίτερη προσοχή βέβαια αποδώθηκε και στην εξαγωγή και αναγνώριση των οροσήμων.

2.2 Παραδοχές

Στο τμήμα αυτό παρουσιάζονται και αναλύονται οι παραδοχές που έχουν γίνει στο σχεδιασμό και την υλοποίηση της προτεινόμενης μεθόδου.

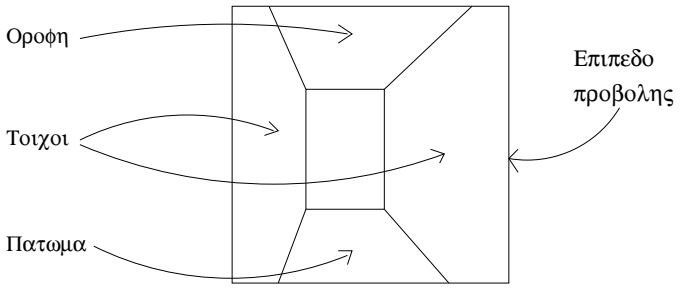
Έχει ήδη αναφερθεί η παραδοχή της κίνησης του αυτόνομου συστήματος σε εσωτερικούς χώρους. Μία επιπλέον προϋπόθεση σχετικά με τους χώρους κίνησης είναι να αποτελούνται από ορθογώνια παραλληλεπίπεδα (σχήμα 2.1). Η προϋπόθεση αυτή ικανοποιείται από την πλειοψηφία των εσωτερικών χώρων, όπως δωμάτια ή διάδρομοι.



Σχήμα 2.1: Χώρος κίνησης

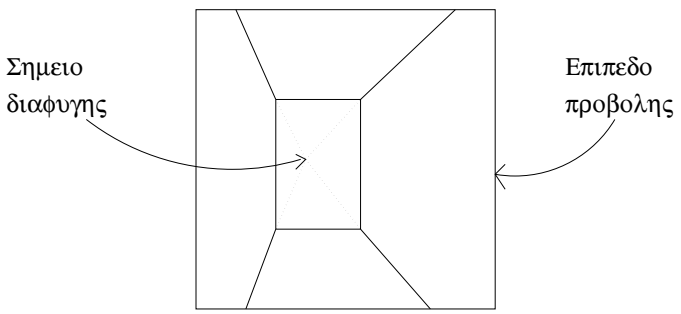
Ο λόγος για αυτή την απαίτηση είναι η δυνατότητα εύκολης εξαγωγής

ποιοτικών πληροφοριών για το χώρο. Στο σχήμα 2.2 φαίνεται η αναμενόμενη όψη του χώρου από το αυτόνομο σύστημα, αν κάνουμε την υπόθεση ότι ο χώρος αποτελείται από ορθογώνια παραλληλεπίπεδα. Βλέπουμε ότι έχουμε τη δυνατότητα να χωρίσουμε την εικόνα σε κατάλληλες περιοχές, ποιοτικά διαφορετικές μεταξύ τους. Ο διαχωρισμός αυτός θα επιτρέψει στο αυτόνομο σύστημα να εντοπίζει ορόσημα πάντα στις περιοχές που θέλουμε.



Σχήμα 2.2: Αναμενόμενη όψη

Από αυτό το σχήμα γίνεται φανερό ότι ο διαμερισμός του χώρου βασίζεται στην εξαγωγή των βασικών ακμών που σχηματίζουν οι τοίχοι με την οροφή, το πάτωμα και μεταξύ τους. Απαιτείται λοιπόν κάποια διαφορά φωτεινότητας μεταξύ τους ώστε η εξαγωγή αυτή να είναι δυνατή. Είναι λογικό να υποθέσουμε ότι αυτή η διαφορά υπάρχει, είτε λόγω διαφορετικού χρωματισμού, είτε λόγω του διαφορετικού προσανατολισμού των επιπέδων. Οπωσδήποτε η ”ευαισθησία” του αυτόνομου συστήματος στις διαφορές φωτεινοτήτων είναι εύκολο να προσαρμοστεί στα πραγματικά δεδομένα.



Σχήμα 2.3: Σημείο διαφυγής

Η προϋπόθεση της καθετότητας μεταξύ των επιπέδων που σχηματίζουν το χώρο εγγυάται την ύπαρξη του *σημείου διαφυγής* (vanishing point). Το σημείο διαφυγής είναι το σημείο όπου συγκλίνουν όλες οι βασικές ακμές των τοίχων

με την οροφή και το πάτωμα (σχήμα 2.3). Η εύρεσή του είναι απαραίτητη, αφού παρέχει έμμεσα πολλές πληροφορίες για τα τρισδιάστατα χαρακτηριστικά του χώρου (κυρίως το βάθος).

Μία παραδοχή είναι ότι η είσοδος στο σύστημα θα είναι μία ακολουθία από διαδοχικές στο χρόνο εικόνες, που θα διαφέρουν λίγο μεταξύ τους. Το αυτόνομο σύστημα θα μπορεί έτσι να χρησιμοποιήσει την πληροφορία μιας εικόνας για να πάρει ευκολότερα πιο αξιόπιστα αποτελέσματα από την επόμενη. Η παραδοχή αυτή δεν είναι περιοριστική, αφού απλώς βελτιώνει τα αποτελέσματα που παίρνουμε. Η μέθοδος είναι ικανή να λειτουργήσει και στην περίπτωση που δεν υπάρχουν προηγούμενα αποτελέσματα, όπως στην εκκίνηση του συστήματος, ή στην περίπτωση που η χρήση των προηγούμενων αποτελεσμάτων αποτύχει (π.χ. αν οι διαδοχικές εικόνες ”απέχουν” πολύ μεταξύ τους).

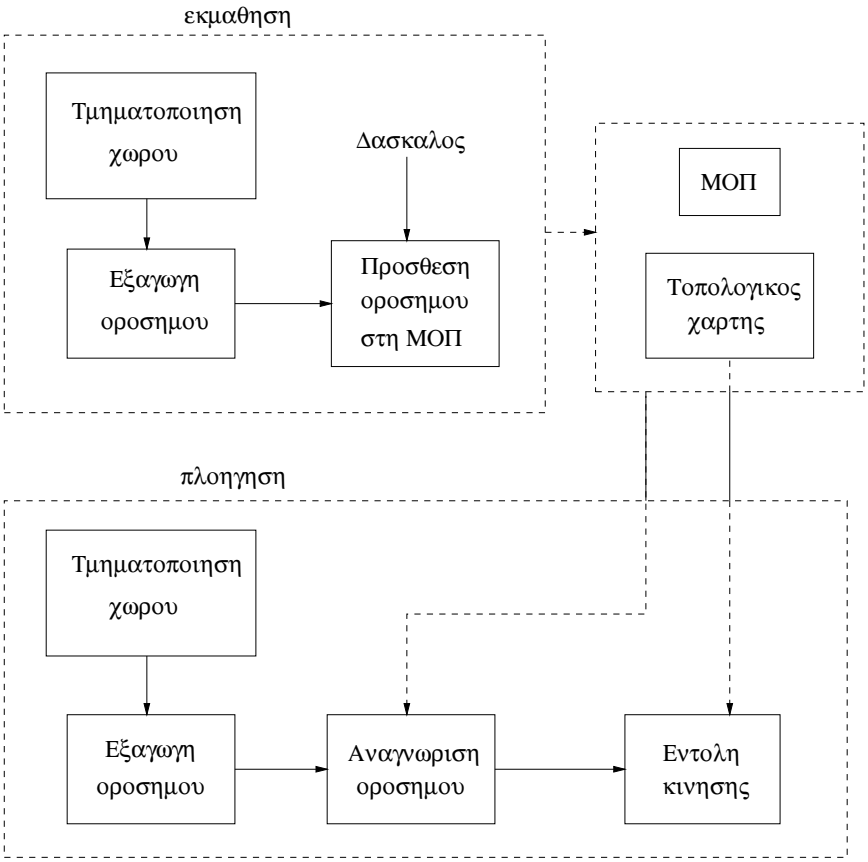
Τέλος, μία προφανής παραδοχή είναι η ύπαρξη οροσήμων στο περιβάλλον κίνησης του αυτόνομου συστήματος. Η ύπαρξη δηλαδή περιοχών ή αντικειμένων στο χώρο που να ξεχωρίζουν και να μπορούν να χρησιμοποιηθούν σαν μέσα περιγραφής και δείκτες πλοήγησης. Είναι και εδώ λογικό να θεωρήσουμε ότι τέτοια αντικείμενα υπάρχουν. Δεν περιμένουμε άλλωστε από το αυτόνομο σύστημα να μπορεί να κινηθεί μέσα σε ένα εντελώς ομοιόμορφο περιβάλλον, όπου ακόμα και ένας άνθρωπος θα μπορούσε να *χάσει* τον προσανατολισμό του.

2.3 Γενική επισκόπηση της μεθόδου

Με τις παραδοχές αυτές, μπορούμε να προχωρήσουμε στην περιγραφή της μεθόδου. Τα κύρια σημεία της φαίνονται στο σχήμα 2.4.

Όπως αναφέρθηκε, η μέθοδος περιλαμβάνει δύο βασικά στάδια : την εκμάθηση και την πλοήγηση. Βασικά τμήματα και των δύο είναι η ποιοτική τμηματοποίηση του χώρου και η εξαγωγή των οροσήμων. Με τον όρο ποιοτική τμηματοποίηση χώρου αναφερόμαστε στον διαχωρισμό της εικόνας στις περιοχές που αναφέρθηκαν παραπάνω (σχήμα 2.2) και την εξαγωγή άλλων πληροφοριών για το χώρο (σημείο διαφυγής). Όλες αυτές οι πληροφορίες χρησιμοποιούνται για την εξαγωγή των οροσήμων, που πραγματοποιείται με την εστίαση της προσοχής σε προέχοντα σημεία σε συγκεκριμένες περιοχές της εικόνας.

Στο στάδιο της εκμάθησης, τα εξαγώμενα ορόσημα, καθώς και άλλες πληροφορίες που σχετίζονται με αυτά, αποθηκεύονται σε μία **Μνήμη Οπτικής Πληροφορίας** (ΜΟΠ), με αποτέλεσμα τη δημιουργία ενός τοπολογικού χάρτη του χώρου. Στο στάδιο της πλοήγησης, το αυτόνομο σύστημα προσπαθεί να αναγνωρίσει κάθε ορόσημο που συναντά, αναζητώντας το στη ΜΟΠ. Ενδε-



Σχήμα 2.4: Πλοήγηση με χρήση οπτικών οροσήμων

χόμενη αναγνώριση, με τη βοήθεια του τοπολογικού χάρτη, του επιτρέπει να καθορίσει την κίνησή του.

Στη συνέχεια θα παρουσιάσουμε τα παραπάνω στάδια αναλυτικά.

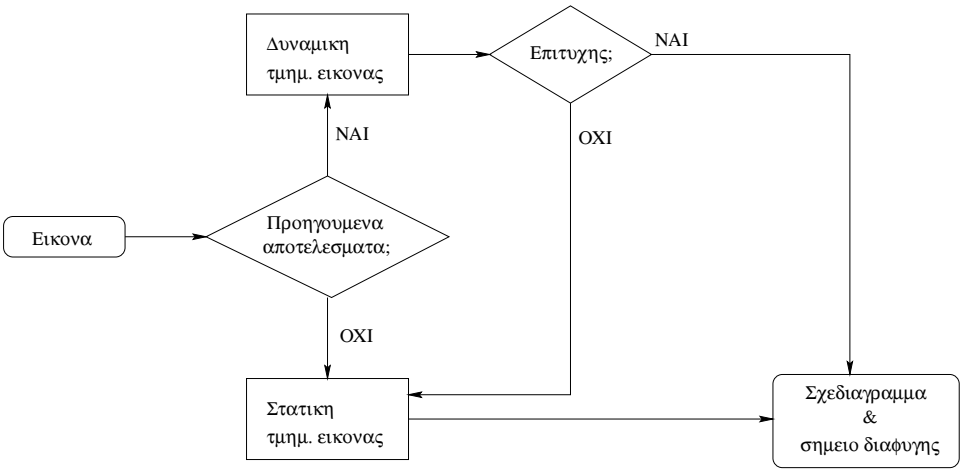
2.4 Ποιοτική τμηματοποίηση χώρου

Με τον όρο ποιοτική τμηματοποίηση χώρου εννοούμε την επεξεργασία της εικόνας για να πάρουμε κάποιο αποτέλεσμα σαν του σχήματος 2.2. Οι πληροφορίες που θέλουμε να αντλήσουμε περιορίζονται στο διαχωρισμό των κυρίων μερών της εικόνας. Ο λόγος που θέλουμε το διαχωρισμό της εικόνας σε τμήματα είναι η κατάλληλη επιλογή των σημείων που θα αναζητήσουμε τα ορόσημα. Θέλουμε π.χ. να αποφύγουμε την επιλογή οροσήμων στην οροφή, γιατί τα φώτα είναι άλλοτε ανοιχτά και άλλοτε κλειστά και δεν θα ήταν αξιόπιστα σαν ορόσημα. Οι περιοχές που επιλέγονται για την αναζήτηση οροσήμων περιορίζονται στους τοίχους, όπου και είναι πιο πιθανό να βρεθούν αντικείμενα κατάλληλα

για ορόσημα.

Από την διαχώριση της εικόνας προκύπτουν και άλλες πληροφορίες, όπως το σημείο διαφυγής, που αναφέρονται σε μετρικές σχέσεις του χώρου. Οι πληροφορίες αυτές θα χρησιμοποιηθούν στη συνέχεια στην εξαγωγή των οροσήμων.

Ο τρόπος με τον οποίο πραγματοποιείται αυτή η ποιοτική τμηματοποίηση παρουσιάζεται στο σχήμα 2.5.



Σχήμα 2.5: Τμηματοποίηση χώρου

Η τμηματοποίηση του χώρου μπορεί να πραγματοποιηθεί με δύο τρόπους, ”δυναμικά” ή ”στατικά”. Με τους όρους ”δυναμική” και ”στατική” τμηματοποίηση εννοούμε τη χρήση ή όχι των αποτελεσμάτων που πήραμε από την προηγούμενη εικόνα στη χρονική ακολουθία. Όπως έχει ήδη αναφερθεί, η στατική τμηματοποίηση είναι δυνατή σε όλες τις εικόνες, ενώ η δυναμική τμηματοποίηση παρουσιάζει μειωμένο υπολογιστικό κόστος. Οι δύο δυνατότητες παρουσιάζονται αναλυτικά παρακάτω. Και στις δύο περιπτώσεις χρησιμοποιούνται απλές τεχνικές επεξεργασίας εικόνας.

2.4.1 Στατική τμηματοποίηση χώρου

Στη συγκεκριμένη μέθοδο τμηματοποίησης χώρου ξεκινάμε χωρίς καμμιά πληροφορία για τις ενδεχόμενες θέσεις των βασικών ακμών που αναζητούμε. Η μέθοδος αποτελείται από μία ακολουθία τεχνικών επεξεργασίας της εικόνας :

1. Γίνεται **εξαγωγή των ακμών** από την εικόνα (**edge detection**), χρησιμοποιώντας το φίλτρο του **Sobel**.

2. Ακολουθεί η **λέπτυνση** (**thinning**) των ακμών που προέκυψαν.

3. Αφαιρούνται οι μικρές ακμές, που αντιστοιχούν σε μικρότερα, ασήμαντα σε αυτό το στάδιο αντικείμενα.

4. Εφαρμόζεται ο **απλός μετασχηματισμός Hough** (**simple Hough transform**) για τον εντοπισμό ευθειών γραμμών στην εικόνα.

5. Στα αποτελέσματα που παίρνουμε εφαρμόζεται ο **επαναληπτικός μετασχηματισμός Hough** (**adaptive Hough transform**) για καλύτερα αποτελέσματα.

6. Τα αποτελέσματα (ευθείες) που προκύπτουν συνδυάζονται για την εξαγωγή του τελικού αποτελέσματος.

2.4.2 Δυναμική τμηματοποίηση χώρου

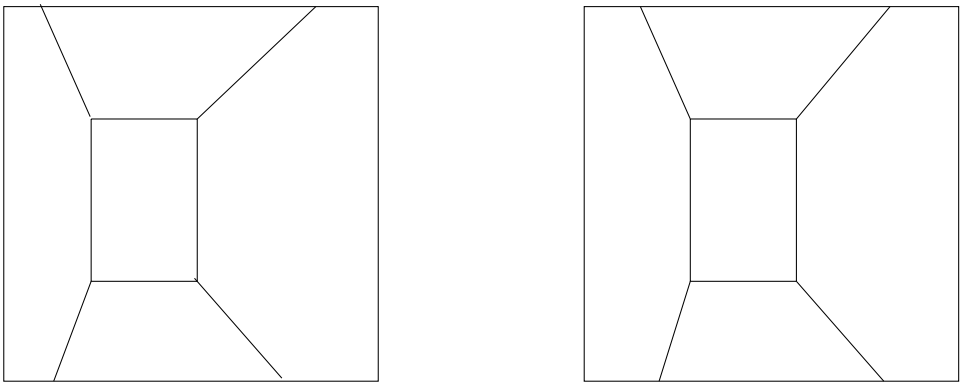
Με τη μέθοδο δυναμικής τμηματοποίησης χώρου στοχεύουμε στη χρήση των αποτελεσμάτων της αμέσως προηγούμενης εικόνας. Υποθέτοντας ότι το αυτόνομο σύστημα δεν έχει μετακινηθεί σημαντικά μεταξύ των χρονικών στιγμών κατά τις οποίες λήφθηκαν οι δύο εικόνες, μπορούμε να περιορίσουμε σημαντικά τα τμήματα της δεύτερης εικόνας στα οποία θα αναζητήσουμε τις βασικές ακμές. Στο σχήμα 2.6-α, βλέπουμε το αποτέλεσμα της τμηματοποίησης χώρου της πρώτης εικόνας, και στο σχήμα 2.6-β το -όχι ακόμα γνωστό- αποτέλεσμα της τμηματοποίησης της δεύτερης.

Από το σχήμα σχήμα 2.6-γ, γίνεται φανερό ότι μπορούμε να περιοριστούμε σε τμήματα της εικόνας που βρίσκονται κοντά στις βασικές ακμές που γνωρίζουμε. Αναλυτικά η προσέγγιση έχει ως εξής :

1. Αρχικά ορίζουμε μία περιοχή που να περικλείει την ακμή της πρώτης εικόνας, με την ελπίδα ότι η περιοχή αυτή θα περικλείει την ίδια ακμή και στη δεύτερη εικόνα (σχήμα 2.6-γ).

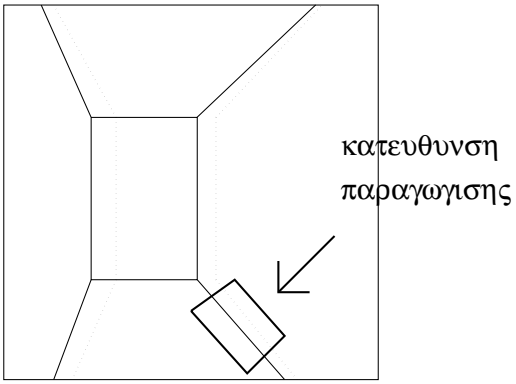
2. Στη συνέχεια παίρνουμε την **παράγωγο** σε κάθε σημείο (pixel) της περιοχής αυτής στη δεύτερη εικόνα. Η παραγώγιση γίνεται σε κατεύθυνση κάθετη στην ακμή της πρώτης εικόνας (σχήμα 2.6-γ), για καλύτερα αποτελέσματα, χρησιμοποιώντας **υπο-σημειακούς υπολογισμούς** (**sub-pixel computations**). Το αποτέλεσμα είναι ένα σύνολο από σημεία (σχήμα 2.6-δ)

3. Σε αυτά τα σημεία προσαρμόζεται μία ευθεία, που είναι και το τελικό αποτέλεσμα για την ακμή στη δεύτερη εικόνα. Για το σκοπό αυτό υλοποιήθηκε μία παραλλαγή του γνωστού αλγορίθμου **Least Median of Squares**.



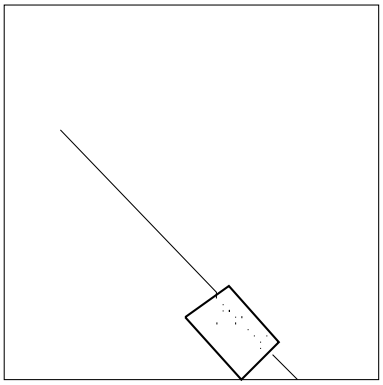
α) 1η εικονα

β) 2η εικονα



γ) 1η &2η εικονα

Τμημα παραγωγισης



δ) Χαραξη γραμμης

Σχήμα 2.6: Χρήση προηγούμενης εικόνας

Η διαδικασία επαναλαμβάνεται για όλες τις βασικές ακμές της εικόνας. Οι ακμές που βρίσκονται και πάλι συνδυάζονται για να δώσουν το τελικό αποτέλεσμα.

2.4.3 Έλεγχος επιτυχίας

Το αποτέλεσμα και των δύο τρόπων ποιοτικής τμηματοποίησης μπορεί να είναι λανθασμένο. Η δυναμική τμηματοποίηση μπορεί να αποτύχει, εάν οι διαδοχικές εικόνες διαφέρουν αρκετά μεταξύ τους και οι ακμές που ζητούμε δεν βρίσκονται στις περιοχές όπου τις αναζητούμε. Η αποτυχία αυτή μπορεί να γίνει αντιληπτή ελέγχοντας το πλήθος των σημείων που προέκυψαν από την παραγώγιση : πολύ μικρός αριθμός σημείων σημαίνει ότι αποτύχαμε να εντοπίσουμε την

ακμή. Η στατική τμηματοποίηση θα μπορούσε επίσης να αποτύχει, αν στην εικόνα υπάρχουν πολλές ακμές που θα μπορούσαν να παραπλανήσουν τον μετασχηματισμό Hough. Σαν κριτήριο επιτυχίας και για τους δύο τρόπους τμηματοποίησης μπορεί να χρησιμοποιηθεί η σύγκριση με τα αποτελέσματα της προηγούμενης στη χρονική ακολουθία εικόνας, με την οποία περιμένουμε μικρές διαφορές στην τελική τμηματοποίηση (π.χ. στην κλίση και το μέγεθος των ακμών).

2.5 Ορόσημα

Τα ορόσημα έχουν θεμελιώδη σημασία στην προσέγγισή μας. Αποτελούν τη βάση, τόσο για την περιγραφή της τοπολογικής δομής του περιβάλλοντος, όσο και για την διαδικασία της πλοήγησης. Η σωστός καθορισμός τους επομένως, όπως και η διαχείριση και χρήση τους, αποτελούν ένα σημαντικό κομμάτι της μεθόδου.

Διακρίνουμε δύο βασικές λειτουργίες σχετικές με τα ορόσημα. Η πρώτη είναι η **εξαγωγή** τους από τις εικόνες που λαμβάνει το αυτόνομο σύστημα κατά το στάδιο της εκμάθησης. Για τον εντοπισμό και την επιλογή των οροσήμων προτείνεται η χρήση του **saliency map**, που αναλύεται στη συνέχεια. Η δεύτερη λειτουργία είναι η **αναγνώριση**, που πραγματοποιείται κατά το στάδιο της πλοήγησης.

2.5.1 Εξαγωγή οροσήμων

Η διαδικασία του εντοπισμού μη προσχεδιασμένων οροσήμων, απαιτεί ένα τρόπο εστίασης στα μέρη εκείνα της εικόνας που είναι δυνατόν να περιέχουν αντικείμενα που μπορούν να χρησιμοποιηθούν σαν ορόσημα. Με άλλα λόγια, χρειαζόμαστε ένα μέτρο του κατά πόσο ένα τμήμα της εικόνας διαφέρει αρκετά από τα υπόλοιπα, ώστε να μπορεί να χρησιμοποιηθεί σαν ορόσημο. Στην προσέγγισή μας, το μέτρο αυτό υλοποιείται με το **saliency map**, ένα χάρτη της εικόνας, όπου οι περιοχές της εικόνας που "ξεχωρίζουν" έχουν μεγαλύτερη τιμή. Ο χάρτης αυτός υπολογίζεται σε κάθε σημείο της εικόνας, ως ένα σταθμισμένο άθροισμα κατάλληλων συναρτήσεων f_i, που εκφράζουν με ποσοτικό τρόπο την *ανομοιογένεια* της εικόνας στην περιοχή του σημείου. Έτσι, η τιμή του saliency map σε κάποιο σημείο p είναι

$$SM = \sum_i w_i f_i(p) \tag{2.1}$$

όπου w_i το βάρος της συνάρτησης f_i στο τελικό άθροισμα. Οι συναρτήσεις που χρησιμοποιήσαμε είναι μία συνάρτηση συσχετίσεως (σε δύο μορφές),

η εντροπία της εικόνας, μία συνάρτηση *κατωφλίου σημείων*(pixel threshold) και μία συνάρτηση που μετρά τη διαφορά μεταξύ διαδοχικών εικόνων. Οι συναρτήσεις αυτές περιγράφονται αναλυτικά στο κεφάλαιο 3.

Για τον εντοπισμό των περιοχών της εικόνας όπου θα ψάξουμε για ορόσημα, χρησιμοποιούμε το αποτέλεσμα της τμηματοποίησης του χώρου της εικόνας. Αυτό που κάνουμε είναι να ελέγχουμε μόνο τα μέρη εκείνα της εικόνας που αντιστοιχούν στους τοίχους, αφού θέλουμε μόνο τα αντικείμενα που θα βρεθούν εκεί να εξυπηρετήσουν σαν ορόσημα.

Μετά τον εντοπισμό των περιοχών της εικόνας που παρουσιάζουν ενδιαφέρον, ελέγχουμε το saliency map στις ίδιες περιοχές, ψάχνοντας για σημεία με μεγάλη τιμή. Γίνεται επιλογή των σημείων με τις k μεγαλύτερες τιμές (όπου k ένας σχετικά μικρός αριθμός). Για να αποφασίσουμε εάν οι τιμές αυτές αντιστοιχούν πράγματι σε ένα αντικείμενο, παίρνουμε το σημείο με τη μεγαλύτερη τιμή και εκτελούμε τα ακόλουθα βήματα :

1. Μετράμε τον αριθμό των άλλων σημείων με μεγάλες τιμές που βρίσκονται ”αρκετά κοντά” για να είναι μέρος ενδεχόμενου αντικειμένου.

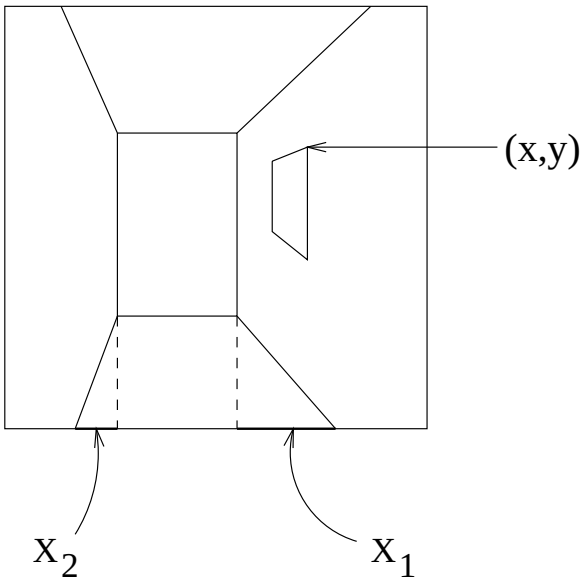
2. Αν ο αριθμός αυτός δεν είναι ικανοποιητικός, απορρίπτουμε το σημείο και επαναλαμβάνουμε από την αρχή. Ο λόγος για αυτό είναι ότι περιμένουμε από ένα αντικείμενο που ξεχωρίζει από τον άμεσο περίγυρό του να συγκεντρώνει πολλά από τα σημεία με τις μεγαλύτερες τιμές.

3. Αν ο αριθμός αυτός είναι ικανοποιητικός, αποφασίζουμε ότι τα σημεία αυτά πράγματι αντιστοιχούν σε κάποιο αντικείμενο.

4. Για να εντοπίσουμε ένα σημείο που να βρίσκεται *εντός* του αντικειμένου, παίρνουμε το κέντρο βάρους όλων αυτών των σημείων.

Με αφετηρία αυτό το σημείο, ένας αλγόριθμος *αύξησης περιοχής* (region growing), αναλαμβάνει να εντοπίσει το τμήμα εκείνο της εικόνας που καταλαμβάνει το αντικείμενο. Το κομμάτι που αποσπάται από την εικόνα αποθηκεύεται στη μνήμη. Αυτό το κομμάτι είναι και το οπτικό ορόσημο που θα χρησιμοποιηθεί κατά τη διάρκεια της πλοήγησης. Μαζί με την εικόνα αποθηκεύονται και άλλες πληροφορίες, που σχετίζονται με τη θέση του οροσήμου, τόσο πάνω στην εικόνα, όσο και στον τρισδιάστατο χώρο. Αυτές είναι :

- Η θέση του οροσήμου πάνω στην εικόνα. Συγκεκριμένα κρατούνται οι συντεταγμένες μίας συγκεκριμένης γωνίας του, το μήκος και το πλάτος του (σχήμα 2.7).



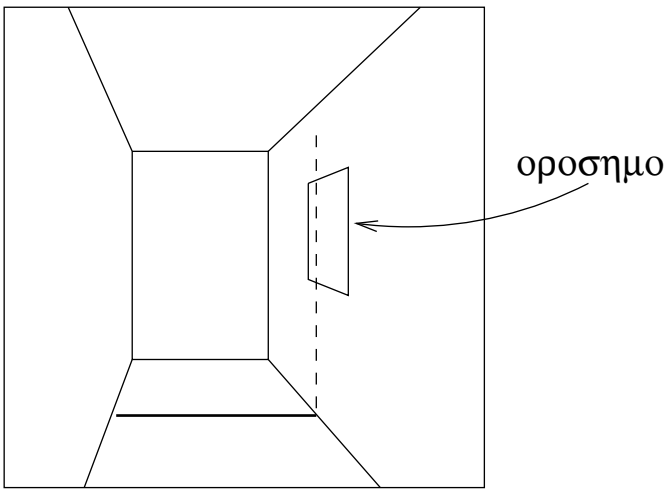
Σχήμα 2.7: Θέση οροσήμου - πληροφορία για τη θέση του παρατηρητή

- Ο λόγος ($t = \frac{x_2}{x_1}$) των μηκών των τμημάτων που ορίζονται από τα ίχνη των δύο κάτω βασικών ακμών και τις προβολές των κατακόρυφων ακμών στο κάτω μέρος της εικόνας (σχήμα 2.7). Τα μήκη των τμημάτων αυτών παρέχουν πληροφορία για τη σχετική θέση του παρατηρητή ως προς τους δύο απέναντι τοίχους.

- Το φαινομενικό πλάτος του πατώματος, στο ίδιο βάθος με τα σημεία (pixels) του οροσήμου πάνω στην εικόνα (σχήμα 2.8). Η μέτρηση αυτή παρέχει πληροφορία ως προς το βάθος της θέσης του οροσήμου στον τρισδιάστατο χώρο, σε σχέση με τον παρατηρητή.

2.5.2 Αναγνώριση οροσήμων

Κατά τη φάση της πλοήγησης, ακολουθείται πρώτα μια διεργασία, πανομοιότυπη με αυτή που ακολουθείται κατά τη φάση της εξαγωγής, ώστε να εντοπισθούν πρώτα πιθανά ορόσημα. Στη συνέχεια, η αναγνώριση ενός οροσήμου συνίσταται στην επιτυχή αντιστοίχισή του με ένα από τα ορόσημα που βρίσκονται αποθηκευμένα στη μνήμη του συστήματος, από τη διαδικασία της εκμάθησης. Η σύγκριση των δύο οροσήμων γίνεται με τη χρήση ενός συντελεστή ομοιότητας. Πριν όμως εφαρμοστεί αυτός ο συντελεστής, πρέπει να ξεπεραστεί ένα άλλο πρόβλημα : το αποθηκευμένο ορόσημο και το ορόσημο που θέλουμε να αναγνωρίσουμε μπορεί να έχουν εξαχθεί από εικόνες που πάρθηκαν από



Σχήμα 2.8: Φαινομενικό πλάτος διαδρόμου

διαφορετικές θέσεις του παρατηρητή.

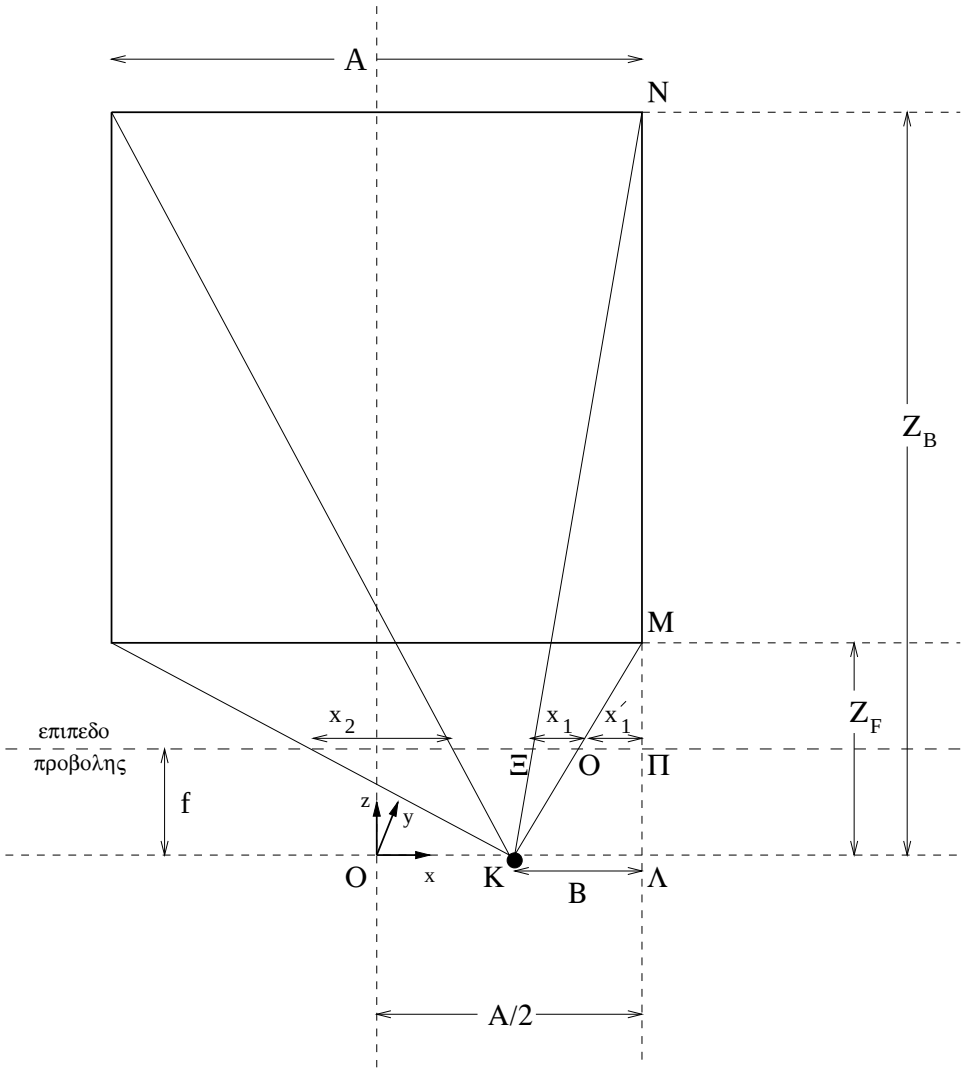
2.5.2.1 Μετασχηματισμός θέσης παρατηρητή (ΜΘΠ)

Είναι απαραίτητος λοιπόν πρώτα κάποιος τρισδιάστατος μετασχηματισμός του ενός οροσήμου, ώστε να γίνει δυνατή η σύγκριση. Το ορόσημο που υφίσταται αυτή την επεξεργασία στη προσέγγισή μας, είναι το ορόσημο που είναι αποθηκευμένο στην μνήμη του αυτόνομου συστήματος. Οι πληροφορίες που είναι απαραίτητες για τον μετασχηματισμό βρίσκονται στη μνήμη μαζί με το ορόσημο. Ο μετασχηματισμός που πραγματοποιείται έχει ως αποτέλεσμα το ορόσημο, όπως θα φαινόταν από την οπτική γωνία του νέου οροσήμου, στο συγκεκριμένο βάθος που βρίσκεται.

Η διατύπωση του ΜΘΠ ακολουθεί απλούς γεωμετρικούς υπολογισμούς οι οποίοι αναλύονται στη συνέχεια. Στο σχήμα 2.9 φαίνεται μία κάτοψη του χώρου[1], με το αυτόνομο σύστημα να βρίσκεται στο σημείο Κ. Τα τμήματα x_1 και x_2 είναι τα αντίστοιχα τμήματα του σχήματος 2.7 (υπολογισμένα σε pixels). Αρχικά, οι μετρήσεις αυτές συνδέονται με ποσότητες του *πραγματικού* χώρου κίνησης : Το πλάτος του χώρου Α, την απόσταση του αυτόνομου συστήματος από τον δεξιό τοίχο, τα βάθη Z_F και Z_B, καθώς και την εστιακή απόσταση (focal length) της κάμερας f. Οι ποσότητες αυτές δεν είναι απαραίτητο να είναι γνωστές, αφού θα απλοποιηθούν στη συνέχεια.

Από τα όμοια τρίγωνα ΚΛΝ και ΞΠΝ έχουμε :

[1]Στο σχήμα έχει γίνει υπέρθεση ενός καθολικού συστήματος συντεταγμένων Oxyz πάνω στην κάτοψη του χώρου.



Σχήμα 2.9: Κάτοψη χώρου

$$x_1 + x_1^{'} = B\frac{Z_B - f}{Z_B} \tag{2.2}$$

Ομοίως, από τα όμοια τρίγωνα ΚΛΜ και ΟΠΜ έχουμε :

$$x_1^{'} = B\frac{Z_F - f}{Z_F} \tag{2.3}$$

Αφαιρώντας κατά μέλη τις (2.2) και (2.3) παίρνουμε :

$$x_1 = Bf(\frac{1}{Z_F} - \frac{1}{Z_B}) \tag{2.4}$$

Με όμοιο τρόπο, έχουμε για το x_2 :

$$x_2 = (\mathrm{A} - B) f (\frac{1}{Z_F} - \frac{1}{Z_B}) \tag{2.5}$$

Διαιρώντας τις (2.4) και (2.5) βρίσκουμε τον λόγο $\frac{x_2}{x_1}$:

$$\frac{x_2}{x_1} = \frac{\mathrm{B}}{\mathrm{A} - \mathrm{B}} \tag{2.6}$$

Αν συμβολίσουμε το λόγο B/A με λ, παίρνουμε

$$\frac{x_2}{x_1} = \frac{\lambda}{1 - \lambda} \tag{2.7}$$

και τελικά

$$\lambda = \frac{x_1}{x_1 + x_2} \tag{2.8}$$

Αυτό που έχουμε καταφέρει είναι να υπολογίσουμε την απόσταση του αυτόνομου συστήματος από δεξιά, ως ποσοστό του συνολικού πλάτους του χώρου κίνησης, χρησιμοποιώντας μόνο μετρήσεις πάνω στην εικόνα (λόγος t=$\frac{x_2}{x_1}$). Η μετάφραση αυτής της γνώσης σε συντεταγμένες ως προς ένα καθολικό σύστημα συντεταγμένων είναι εύκολη. Θεωρούμε ένα τέτοιο καθολικό τέτοιο σύστημα συντεταγμένων $Oxyz$ στο μέσο του χώρου κίνησης, όπως φαίνεται στο σχήμα 2.9. Αν συμβολίσουμε με x_0 την x συντεταγμένη της θέσης του αυτόνομου συστήματος, θα έχουμε

$$\begin{aligned} x_0 &= \tfrac{A}{2} - B \\ &= \tfrac{A}{2} - \lambda A \\ &= \frac{A}{2} - \frac{x_1}{x_1 + x_2} A \end{aligned} \tag{2.9}$$

Ας εξετάσουμε τώρα την ύπαρξη ενός οροσήμου, το οποίο παρατηρεί το αυτόνομο σύστημα από την παραπάνω θέση x_0. Αν θεωρήσουμε ότι κάποιο σημείο που ανήκει στο ορόσημο βρίσκεται σε βάθος Z και έχει X_W τετμημένη (σχήμα 2.10), τότε θα ισχύει

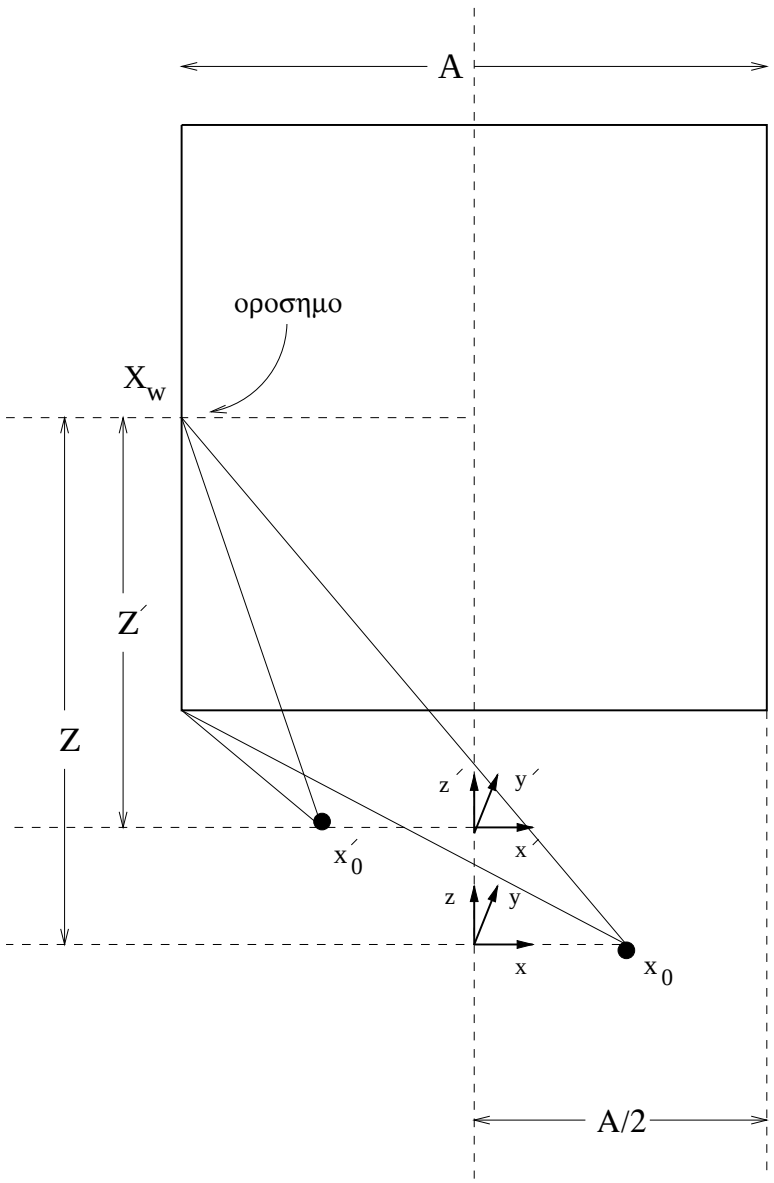
$$X_W = \frac{xZ}{f} - x_0 \tag{2.10}$$

όπου x η τετμημένη του σημείου πάνω στην εικόνα. Κατά τη διάρκεια της πλοήγησης, το αυτόνομο σύστημα μπορεί να συναντήσει το ορόσημο από κάποια άλλη θέση $x_0^{'}$ (σχήμα 2.10). Στο καινούριο σύστημα συντεταγμένων, το βάθος $Z^{'}$ του οροσήμου μπορεί να διαφέρει, όπως επίσης και η τετμημένη του $x^{'}$, όπως θα φαινόταν στην καινούρια εικόνα. Η τετμημένη όμως X_W θα είναι η ίδια

$$X_W = \frac{x^{'} Z^{'}}{f} - x_0^{'} \tag{2.11}$$

με το $x_0^{'}$ να βρίσκεται όπως και το x_0 :

$$x_0^{'} = \frac{A}{2} - \frac{x_1^{'}}{x_1^{'} + x_2^{'}} A \tag{2.12}$$

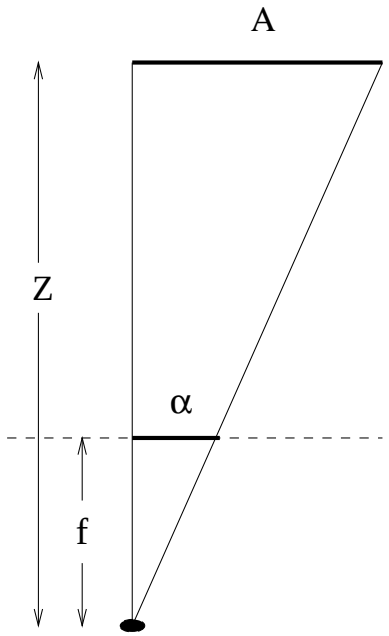


Σχήμα 2.10: Διαφορετικές θέσεις παρατηρητή

Εξισώνοντας τις (2.10) και (2.11 έχουμε)

$$x^{'} = \frac{xZ}{Z^{'}} - \frac{(x_0 - x_0^{'})f}{Z^{'}} \tag{2.13}$$

που μας δίνει την τετμημένη στην καινούρια εικόνα όπου θα βρίσκαμε το ορόσημο, κοιτώντας το από την καινούρια θέση. Αυτό που μένει είναι να εξαλείψουμε τις άγνωστες ποσότητες που εμφανίζονται στην εξίσωση (βάθη Z και $Z^{'}$ και θέσεις του παρατηρητή x_0 και $x_0^{'}$ που εκφράζονται σύμφωνα με την εξίσωση (2.9). Για να γίνει αυτό θα χρησιμοποιήσουμε τη μέτρηση πάνω στην εικόνα του πλάτους του χώρου, που έχει αποθηκευτεί μαζί με το ορόσημο (σχήμα 2.8). Αν συμβολίσουμε με α το πλάτος αυτό σε pixels, θα έχουμε (σχήμα 2.11) :



Σχήμα 2.11: Μέτρηση βάθους από το φαινομενικό πλάτος του χώρου

$$\mathrm{Z} = \frac{Af}{\alpha} \tag{2.14}$$

και ομοίως για το Z$^{'}$. Αντικαθιστώντας στην (2.13) τα Z, $Z^{'}$, x_0 και $x_0^{'}$, από τις (2.14), (2.9), (2.12), όλες οι άγνωστες ποσότητες απλοποιούνται τελικά και παίρνουμε

$$x^{'} = \frac{\alpha^{'}}{\alpha}x - \alpha^{'}[\frac{x_1^{'}}{x_1^{'} + x_2^{'}} - \frac{x_1}{x_1 + x_2}] \tag{2.15}$$

ή, θέτοντας $t = \frac{x_2}{x_1}$ και αντίστοιχα $t^{'} = \frac{x_2^{'}}{x_1^{'}}$

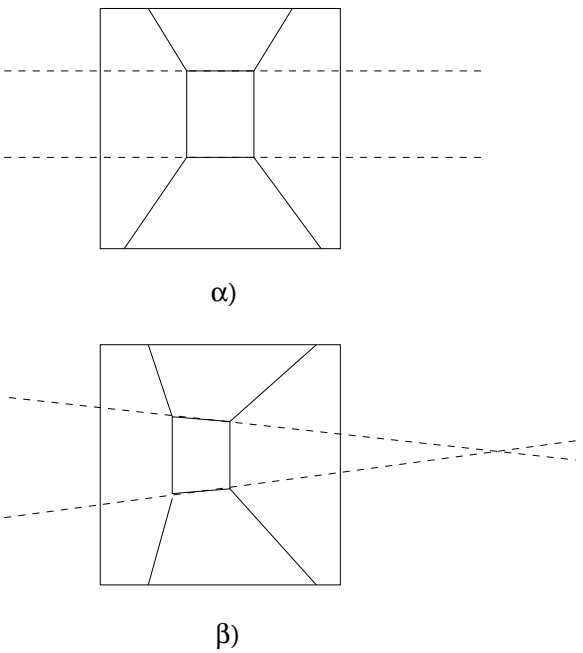
$$x^{'} = \frac{\alpha^{'}}{\alpha} x - \alpha^{'} [\frac{1}{1+t^{'}} - \frac{1}{1+t}] \tag{2.16}$$

Ακολουθώντας την ίδια μέθοδο και για τον άξονα y, βρίσκουμε

$$y^{'} = \frac{\alpha^{'}}{\alpha} y \tag{2.17}$$

Έχοντας τις εξισώσεις (2.16) και (2.17) ο ΜΘΠ είναι εύκολος. Αν ένα σημείο του οροσήμου βρίσκεται στη θέση (x, y) στην εικόνα του οροσήμου που βρίσκεται στη μνήμη, τότε στην καινούρια εικόνα θα βρίσκεται στη θέση $(x^{'}, y^{'})$. Εφαρμόζοντας τις εξισώσεις για κάθε σημείο του αποθηκευμένου οροσήμου, παίρνουμε την εικόνα του από την καινούρια θέση του παρατηρητή.

Προϋπόθεση για την ισχύ όλων των παραπάνω εξισώσεων, είναι ο άξονας z της κάμερας του αυτόνομου συστήματος να είναι παράλληλος με τον άξονα z του χώρου. Η προϋπόθεση αυτή δεν περιορίζει την εφαρμογή του παραπάνω μετασχηματισμού, αφού η παραλληλία των αξόνων μπορεί να επιτευχθεί με μία απλή περιστροφή του αυτόνομου συστήματος[2]. Ο έλεγχος της παραλληλίας ή όχι των αξόνων γίνεται εύκολα από το διάγραμμα της μερικής τμηματοποίησης του χώρου (σχήμα 2.12).



Σχήμα 2.12: Έλεγχος παραλληλίας αξόνων

[2] Ένα καλό παράδειγμα ενεργούς όρασης, κατά το οποίο ο παρατηρητής έχει πλήρη έλεγχο των παραμέτρων λήψης των εικόνων.

Στην περίπτωση παραλληλίας, οι δύο οριζόντιες ακμές θα είναι παράλληλες (σχήμα 2.12-α), επομένως δεν θα ανιχνεύεται σημείο διαφυγής στον άξονα x. Εάν δεν υπάρχει παραλληλία, οι δύο ακμές θα συγκλίνουν (σχήμα 2.12-β). Η θέση του σημείου σύγκλισης δείχνει την κατεύθυνση προς την οποία πρέπει να περιστραφεί η κάμερα του αυτόνομου συστήματος για να επιτευχθεί παραλληλία.

2.5.2.2 Διγραμμική παρεμβολή

Ο τρισδιάστατος μετασχηματισμός του οροσήμου, όπως αναλύθηκε παραπάνω, δεν είναι αρκετός. Το αποτέλεσμα του μετασχηματισμού είναι μάλλον ένα σύνολο σημείων στο δισδιάστατο χώρο, παρά το τμήμα μιας εικόνας. Ο λόγος γι' αυτό είναι η συγκεκριμένη μορφή των εξισώσεων του τρισδιάστατου μετασχηματισμού. Έτσι, τα αποτελέσματα που παίρνουμε αφορούν σημεία στο *πραγματικό* δισδιάστατο χώρο και όχι σε ένα ακέραιο δισδιάστατο πλέγμα.

Για να αντιμετωπίσουμε το πρόβλημα, χρησιμοποιούμε διγραμμική παρεμβολή, ώστε να υπολογιστούν οι τιμές της συνάρτησης εικόνας πάνω στα pixels. Οι φωτεινότητα ενός σημείου του ακέραιου πλέγματος, προσεγγίζεται από τις τιμές των σημείων που βρίσκονται "κοντά" του. Στο σχήμα 2.13 φαίνεται ότι τα σημεία που λαμβάνονται υπ' όψιν είναι αυτά που βρίσκονται μέσα στο 2x2 παράθυρο γύρω από το σημείο του πλέγματος. Η φωτεινότητα Φ_i ενός σημείου σε αυτή την περιοχή παίρνει ένα βάρος w_i, ανάλογα με την απόσταση του d_i από το σημείο του πλέγματος. Αφού $0 \leq d_i \leq \sqrt{2}$, ορίζουμε

$$w_i = \sqrt{2} - d_i \tag{2.18}$$

Έτσι, η φωτεινότητα που αποδίδουμε στο σημείο του πλέγματος είναι

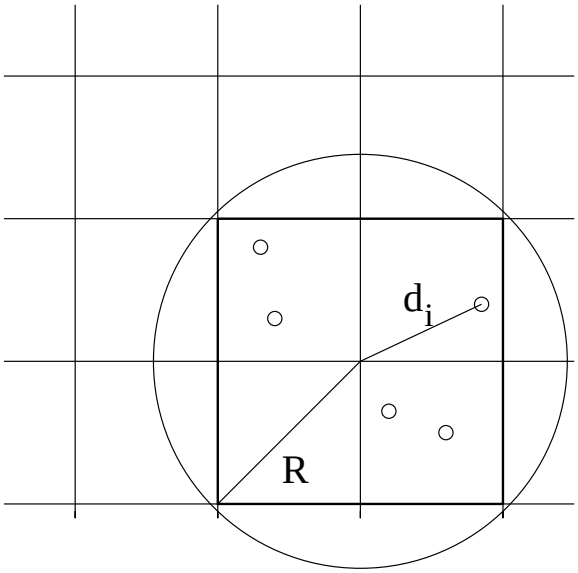
$$\mathbf{\Phi} = \sum_i \frac{w_i \mathbf{\Phi}_i}{\sum w_i} \tag{2.19}$$

2.5.2.3 Σύγκριση οροσήμων

Μετά και τη διγραμμική παρεμβολή, τα δύο ορόσημα είναι έτοιμα για σύγκριση. Ο συντελεστής που χρησιμοποιείται δίνεται από τον τύπο [Pap91]

$$Coef = \frac{< I_p T > - < I_p >< T >}{\sigma_{I_p} \sigma_T} \tag{2.20}$$

όπου I_p το τμήμα της εικόνας (image patch) που καταλαμβάνει το ένα ορόσημο, T το αποθηκευμένο στη μνήμη ορόσημο (template) και $I_p T$ το γινόμενο των φωτεινοτήτων αντίστοιχων pixels των δύο οροσήμων. Η αντιστοίχηση των



Σχήμα 2.13: Υπολογισμός φωτεινότητας σημείων στο ακέραιο πλέγμα

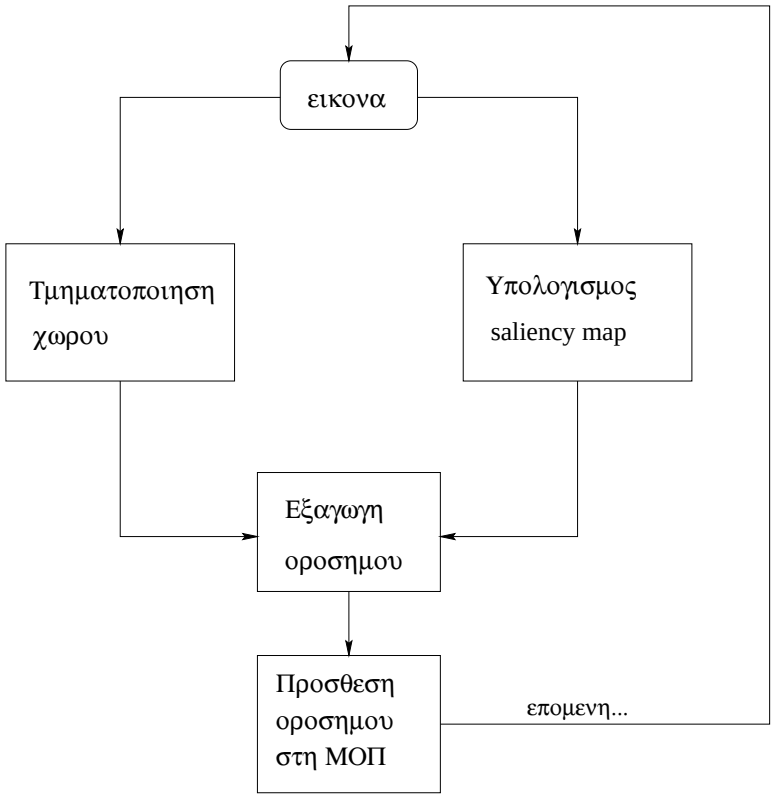
δύο προτύπων, και συνεπώς των σημείων τους, γίνεται ως εξής : Το ένα πρότυπο τοποθετείται ως προς το άλλο σύμφωνα με όλους τους δυνατούς τρόπους (ολισθαίνει) και υπολογίζεται η μέγιστη ομοιότητα, η οποία και κρατιέται ως τιμή για τον συντελεστή σύγκρισης. Ο συμβολισμός <> δηλώνει μέση τιμή και σ είναι η διασπορά. Ο $Coef$, όπως έχει ορισθεί, παίρνει τιμές μεταξύ 0 και 1. Μία τιμή κοντά στο 1 υποδηλώνει πλήρη ομοιότητα μεταξύ των δύο οροσήμων, ενώ μια τιμή κοντά στο 0 υποδηλώνει ανομοιότητα.

Έχοντας περιγράψει τις μεθόδους τμηματοποίησης της δομής του χώρου, καθώς και τον χειρισμό των οροσήμων, είμαστε έτοιμοι να παρουσιάσουμε τα δύο βασικά στάδια της μεθόδου που αναφέρθηκαν στην αρχή του κεφαλαίου : την εκμάθηση και την πλοήγηση.

2.6 Στάδιο εκμάθησης

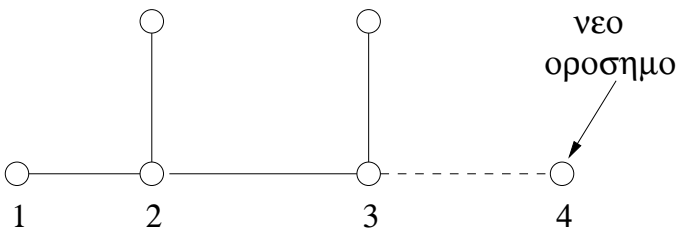
Σε αυτό το στάδιο, το αυτόνομο σύστημα συλλέγει όλες τις πληροφορίες που του είναι απαραίτητες για την πλοήγηση. Χρησιμοποιώντας τις πληροφορίες που προκύπτουν από την τμηματοποίηση του χώρου (διαχώριση σε ποιοτικές περιοχές), το σύστημα εξάγει ορόσημα από το περιβάλλον και τα αποθηκεύει για μελλοντική χρήση. Ένα γενικό διάγραμμα της διαδικασίας της εκμάθησης παρουσιάζεται στο σχήμα 2.14.

Τα ορόσημα αποθηκεύονται σε μια **Μνήμη Οπτικής Πληροφορίας** (ΜΟΠ).



Σχήμα 2.14: Διαδικασία εκμάθησης

Εκεί κρατούνται για κάθε ορόσημο η εικόνα του (το τμήμα της εικόνας που αναγνωρίστηκε σαν ορόσημο), τα χαρακτηριστικά του και η θέση του. Η σχετική θέση του οροσήμου ως προς τα υπόλοιπα προσδιορίζεται από το αυτόνομο σύστημα πάνω σε ένα τοπολογικό χάρτη (σχήμα 2.15).

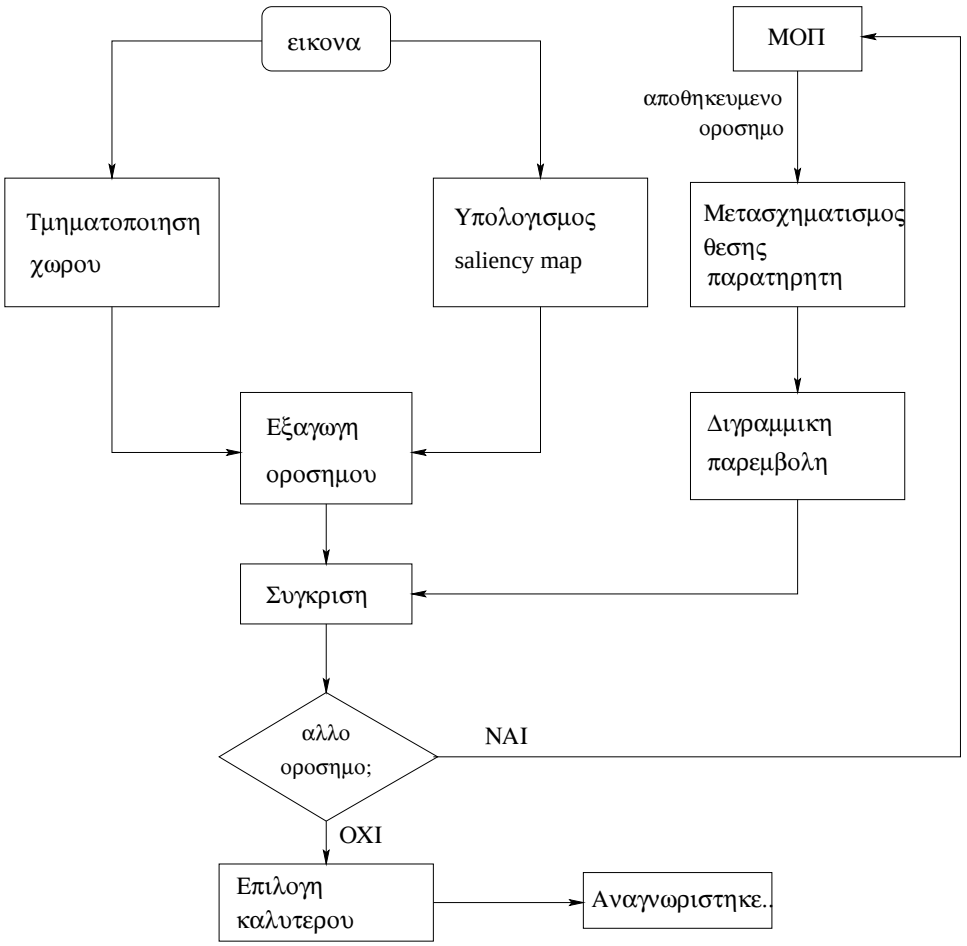


Σχήμα 2.15: Τοπολογικός χάρτης

Με αυτό τον τρόπο, η ΜΟΠ αποτελεί για το αυτόνομο σύστημα μία μορφή περιγραφής του περιβάλλοντος του. Στη μνήμη αυτή μπορεί να βρει περιγραφές των αντικειμένων που θα συναντήσει και τις θέσεις τους, με βάση τις οποίες μπορεί να προγραμματίσει την κίνησή του. Για παράδειγμα, για τη μετακίνηση

από τη θέση 1 του σχήματος στη θέση 4, το αυτόνομο σύστημα γνωρίζει ότι πρέπει να περάσει από τα ορόσημα στις ενδιάμεσες θέσεις 2 και 3. Η παρουσία ενός δασκάλου κατά το στάδιο της εκμάθησης επιτρέπει ακόμα τον *σημαντικό* (semantic) προσδιορισμό των οροσήμων μέσα στο χώρο (το ορόσημο 1 βρίσκεται στην κουζίνα, το ορόσημο 4 στο εργαστήριο κλπ.). Έτσι, οι εντολές πλοήγησης μπορούν να δίνονται με πιο εύκολο τρόπο (”πήγαινε στο εργαστήριο” αντί ”πήγαινε στο ορόσημο 4”). Με τις παραπάνω πληροφορίες στη ΜΟΠ, το αυτόνομο σύστημα είναι έτοιμο για πλοήγηση.

2.7 Πλοήγηση



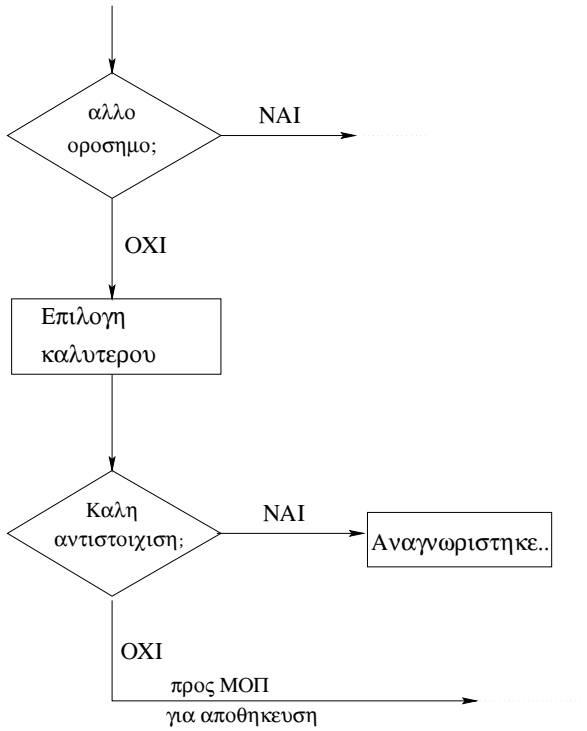
Σχήμα 2.16: Διαδικασία πλοήγησης

Η διαδικασία της πλοήγησης ακολουθεί, μέχρι ένα σημείο, τη διαδικασία της εκμάθησης : σε κάθε εικόνα, γίνεται προσπάθεια εντοπισμού και εξαγωγής

οροσήμων, χρησιμοποιώντας πάλι την ποιοτική τμηματοποίηση του χώρου και το saliency map (σχήμα 2.16).

Από τη στιγμή όμως που εντοπισθεί ένα ορόσημο, αρχίζει η διαδικασία της αναγνώρισής του. Το ορόσημο που εντοπίστηκε συγκρίνεται με τη σειρά με κάθε ένα από τα ορόσημα που είναι αποθηκευμένα στη μνήμη. Η σύγκριση απαιτεί τους μετασχηματισμούς που περιγράφηκαν στο τμήμα της αναγνώρισης των οροσήμων. Το άγνωστο ορόσημο αναγνωρίζεται τελικά σαν το αποθηκευμένο στη μνήμη ορόσημο με το οποίο ”ταυτίζεται” περισσότερο.

Μπορεί να γίνει ακόμη μια προσθήκη στη μέθοδο της πλοήγησης, ώστε να μπορεί το αυτόνομο σύστημα να ανταποκρίνεται σε αλλαγές του περιβάλλοντος, που μπορεί να σημαίνουν την παρουσία καινούριων οροσήμων στο χώρο κίνησης. Αυτό μπορεί να πραγματοποιηθεί με τον έλεγχο του βαθμού ομοιότητας του οροσήμου που επιλέχθηκε από τη ΜΟΠ ως καλύτερη αντιστοίχιση του υπό εξέταση οροσήμου. Αν αυτός ο βαθμός ομοιότητας δεν κριθεί ”ικανοποιητικός”, μπορούμε να θεωρήσουμε ότι συναντήσαμε ένα καινούριο ορόσημο, που δεν υπάρχει στην μνήμη οπτικής πληροφορίας (σχήμα 2.17).



Σχήμα 2.17: Ενημέρωση γνώσης

Σε αυτή την περίπτωση επιστρέφουμε στη διαδικασία της εκμάθησης. Το

ορόσημο αποθηκεύεται στη ΜΟΠ με όλες τις συνοδεύουσες πληροφορίες για μελλοντική χρήση. Με αυτό τον τρόπο, η γνώση του αυτόνομου συστήματος για το περιβάλλον ενημερώνεται, με κάθε καινούρια αλλαγή.

Κεφάλαιο 3

Υλοποίηση

3.1 Εισαγωγή

Στο κεφάλαιο αυτό θα περιγράψουμε την υλοποίηση της προτεινόμενης μεθόδου. Θα περιγράψουμε τους αλγορίθμους και τις τεχνικές επεξεργασίας εικόνας που χρησιμοποιήθηκαν, καθώς και τον τρόπο εφαρμογής τους στην παρούσα εργασία. Το κεφάλαιο περιλαμβάνει τον τρόπο υλοποίησης της τμηματοποίησης χώρου, στατικής και δυναμικής, και τον τρόπο υλοποίησης του εντοπισμού και της αναγνώρισης των οροσήμων.

3.2 Στατική τμηματοποίηση χώρου

Η στατική τμηματοποίηση του χώρου αποτελείται από μία ακολουθία απλών τεχνικών επεξεργασίας εικόνας. Στην υλοποίηση της μεθόδου που έγινε, η εικόνα που δίνεται σαν είσοδος αρκεί να περιέχει μόνο αποχρώσεις του γκρι (grey_scaled). Αν η αρχική εικόνα είναι έγχρωμη, μετατρέπεται σε gray_scaled χρησιμοποιώντας τον τύπο

$$L = \sqrt{\frac{R^2 + G^2 + B^2}{3}}$$

όπου R,G,B οι φωτεινότητες του σημείου της έγχρωμης εικόνας στα τρία βασικά χρώματα και L η καινούργια φωτεινότητα του σημείου στην gray_scaled εικόνα. Οι τεχνικές επεξεργασίας εικόνας που χρησιμοποιούνται είναι κατά σειρά η εξαγωγή και λέπτυνση των ακμών, η αφαίρεση των μικρών αντικειμένων και η χρήση του μετασχηματισμού Hough για εύρεση ευθειών. Ακολουθεί η τελική τμηματοποίηση της εικόνας.

3.2.1 Εξαγωγή ακμών

Το πρώτο βήμα στην στατική τμηματοποίηση είναι η εξαγωγή των ακμών της εικόνας. Η εξαγωγή περιλαμβάνει αρχικά τον υπολογισμό των τιμών της παραγώγου σε κάθε σημείο της εικόνας. Για την παραγώγιση, μετά από πειραματισμό με εικόνες από εσωτερικούς διαδρόμους του ITE, επιλέχθηκε το γνωστό φίλτρο του Sobel (σχήμα 3.1). Η εικόνα στη συνέχεια μετατρέπεται σε δυαδική, με τη χρήση ενός κατωφλίου. Η τιμή που χρησιμοποιείται στην εφαρμογή μας είναι 100. Έτσι, όσα σημεία έδωσαν τιμή της παραγώγου μεγαλύτερη από 100, θεωρούνται σημεία ακμών, ενώ τα υπόλοιπα διαγράφονται.

−1	−2	−1
0	0	0
1	2	1

−1	0	1
−2	0	2
−1	0	1

Σχήμα 3.1: Φίλτρο Sobel - κάθετη και οριζόντια παραγώγιση

3.2.2 Λέπτυνση ακμών

Για τη λέπτυνση (thinning) των ακμών, χρησιμοποιήθηκε ο αλγόριθμος των Naccache και Shinghal [NS84]. Ο αλγόριθμος βασίζεται στην έννοια του *σκελετού* μιας περιοχής [Blu67]. Η δυαδική εικόνα σαρώνεται επαναληπτικά και αφαιρούνται τα σημεία των ακμών που δεν αποτελούν μέρος του σκελετού των ακμών. Το αν ένα σημείο είναι μέρος ή όχι του σκελετού, αποφασίζεται με το ταίριασμα (matching) της γειτονιάς 8 σημείων (8-neighborhood) σημείου που εξετάζουμε με πρότυπα (patterns) πολύ γνωστά σε αλγόριθμους λέπτυνσης.

Η λέπτυνση των ακμών είναι αναγκαία σ' αυτό το στάδιο γιατί η εικόνα που προκύπτει θα δοθεί αργότερα σαν είσοδος στο μετασχηματισμό Hough, για εντοπισμό ευθειών. Η λέπτυνση μειώνει τον αριθμό των σημείων στην εικόνα και επιτρέπει στον Hough καλύτερα αποτελέσματα.

3.2.3 Αφαίρεση μικρών αντικειμένων

Πριν ο μετασχηματισμός Hough αναλάβει τον εντοπισμό των ακμών, αφαιρούμε από την εικόνα όλες τις "μικρές" ακμές. Λέγοντας "μικρές" εννοούμε ότι η

μεγαλύτερη απόσταση μεταξύ δύο σημείων τους είναι μικρότερη από κάποιο κατώφλι, που δίνεται σαν παράμετρος στο πρόγραμμα. Η τυπική τιμή αυτού του κατωφλίου είναι μεταξύ 20 και 25 pixels. Με αυτόν τον τρόπο αφαιρούνται ακμές που θα μπορούσαν να επηρεάσουν τον Hough και να τον οδηγήσουν σε λανθασμένα αποτελέσματα. Δεν υπάρχει κίνδυνος να αφαιρεθεί και κάποια από τις βασικές ακμές, αφού περιμένουμε αυτές να έχουν μήκος σημαντικά μεγαλύτερο από 20 pixels.

3.2.4 Εύρεση ευθειών

Το αποτέλεσμα του προηγούμενου σταδίου είναι μία δυαδική εικόνα, που περιλαμβάνει τα σημεία των ακμών της αρχικής εικόνας. Σ' αυτήν την εικόνα εφαρμόζεται ο γνωστός μετασχηματισμός Hough για την εύρεση ευθειών.

Η εφαρμογή του Hough εδώ δεν έχει σαν στόχο την εύρεση όλων των ευθειών που αντιστοιχούν στις βασικές ακμές που ψάχνουμε. Κάτι τέτοιο θα ήταν αδύνατο, λόγω της εφαρμογής του μετασχηματισμού σε ολόκληρη την εικόνα, η οποία περιέχει πολλά σημεία, σε βαθμό που τα αποτελέσματα που θα έδινε δεν θα ήταν αξιόπιστα.

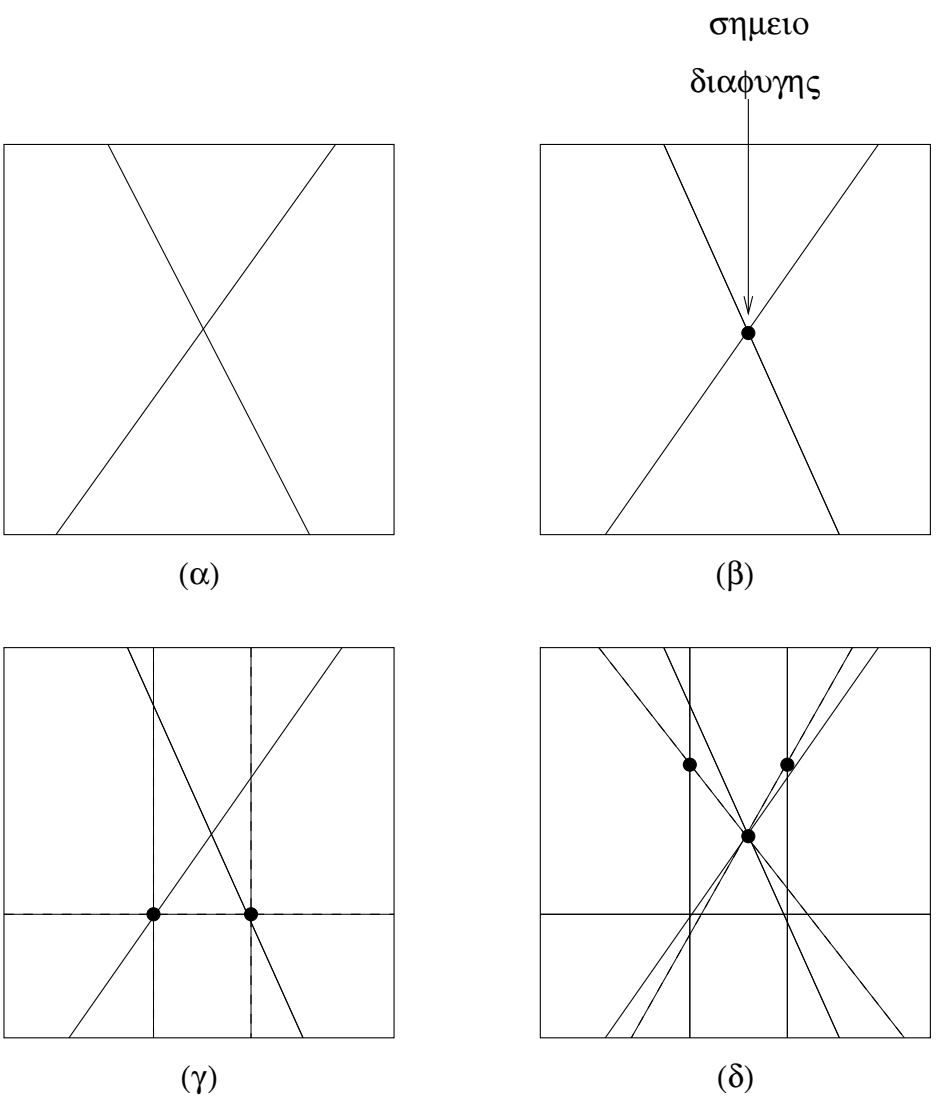
Για τον λόγο αυτό η εύρεση των ευθειών γίνεται επιλεκτικά. Ο μετασχηματισμός Hough χρησιμοποιείται μόνο μία φορά σε ένα συγκεκριμένο εύρος των παραμέτρων ρ και θ, για να πάρουμε μία *προσέγγιση* των δύο ακμών που σχηματίζουν οι τοίχοι με το πάτωμα (σχήμα 3.2-α). Από εδώ και πέρα, η εύρεση των ευθειών γίνεται σταδιακά :

1. Έχοντας τις προσεγγίσεις των δύο ακμών του σχήματος 3.2-α, εφαρμόζουμε τον επαναληπτικό μετασχηματισμό Hough (Adaptive Hough Transform (AHT)[1])[IK87], για τον ακριβή εντοπισμό των ακμών. Το σημείο τομής τους μας δίνει αμέσως το σημείο διαφυγής, το οποίο θα χρησιμοποιήσουμε στη συνέχεια (σχήμα 3.2-β).

2. Από τα σημεία της εικόνας που ανήκουν στις ευθείες που βρήκαμε, επιλέγουμε για κάθε ακμή το σημείο με την μεγαλύτερη y συντεταγμένη (σχήμα 3.2-γ). Τα σημεία αυτά μας υποδεικνύουν τις περιοχές όπου θα πρέπει να αναζητήσουμε τις επόμενες ακμές. Εφαρμόζουμε και πάλι τον AHT για τον εντοπισμό τους.

[1]Ο AHT υπολογίζει με μεγαλύτερη ακρίβεια τις παραμέτρους ρ και θ μίας γραμμής, με την επαναληπτική εφαρμογή του σε όλο και πιο *λεπτομερείς* (fine) υποδιαιρέσεις των παραμέτρων όπου έχει ανιχνευτεί μία γραμμή.

3. Όπως και πριν, βρίσκουμε τα σημεία της εικόνας με τη μεγαλύτερη y συντεταγμένη που ανήκουν στις κάθετες ακμές. Τα σημεία αυτά, μαζί με το σημείο διαφυγής, αρκούν για τον εντοπισμό και των υπόλοιπων ακμών (σχήμα 3.2-δ).

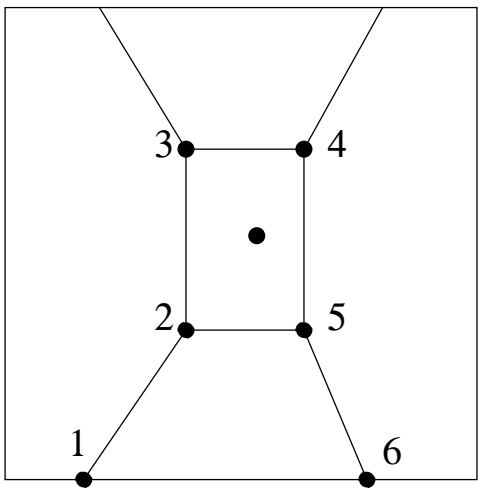


Σχήμα 3.2: Διαδοχικές εφαρμογές του ΑΗΤ

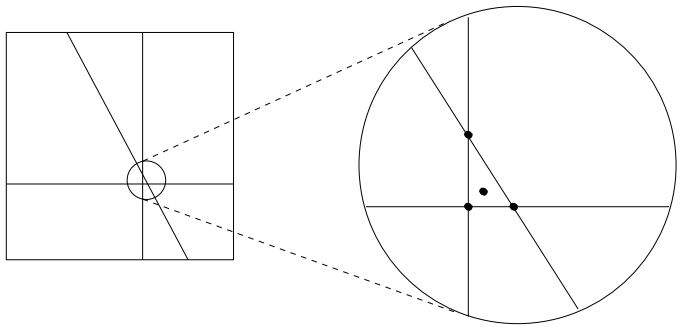
3.2.5 Τμηματοποίηση εικόνας

Η τμηματοποίηση του χώρου στην ουσία γίνεται βήμα προς βήμα στο προηγούμενο στάδιο. Τα σημεία 1 έως 6 του σχήματος 3.3 βρίσκονται με τον υπολογισμό

σημείων τομής μεταξύ των κατάλληλων ακμών. Στην περίπτωση που κάποιο σημείο πρέπει να ανήκει σε τρεις από τις ακμές (σχήμα 3.4), βρίσκουμε που τέμνονται οι ακμές ανά δύο, και παίρνουμε το μέσο όρο των τριών σημείων τομής. Το μόνο που μένει μετά, είναι η ένωση των σημείων και η έξοδος του αποτελέσματος, που είναι η εικόνα που παριστά την τμηματοποίηση.



Σχήμα 3.3: Τελικό αποτέλεσμα



Σχήμα 3.4: Τριπλό σημείο τομής

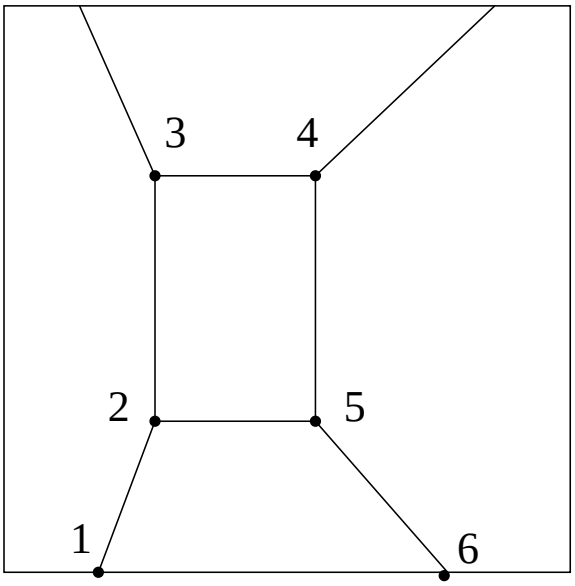
3.3 Δυναμική τμηματοποίηση χώρου

Όπως αναφέρθηκε και στο κεφάλαιο 2, η δυναμική τμηματοποίηση του χώρου πραγματοποιείται με παραγώγιση σε τμήματα της εικόνας όπου περιμένουμε να συναντήσουμε τις βασικές ακμές. Η παραγώγιση έχει σαν αποτέλεσμα ένα

σύνολο σημείων (για κάθε περιοχή), από το οποίο εντοπίζεται η αντίστοιχη ακμή με ένα αλγόριθμο χάραξης γραμμής. Το πρώτο βήμα όμως σε αυτή την προσέγγιση είναι ο καθορισμός των κατάλληλων περιοχών στην εικόνα, όπου θα ψάξουμε για ακμές.

3.3.1 Χρήση προηγούμενης γνώσης

Η πληροφορία που χρειαζόμαστε από την προηγούμενη στην ακολουθία εικόνα, είναι η θέση των βασικών ακμών σε αυτή. Για το σκοπό αυτό, το μόνο που χρειάζεται να γνωρίζουμε είναι τα έξι σημεία που φαίνονται στο σχήμα 3.5.



Σχήμα 3.5: Απαραίτητη προηγούμενη πληροφορία

Τα σημεία αυτά ορίζουν απ' ευθείας τις έξι από τις οκτώ βασικές ακμές της τμηματοποίηση. Οι υπόλοιπες δύο ακμές (μεταξύ οροφής και τοίχων) βρίσκονται εύκολα από τα σημεία 3 και 4 και το σημείου διαφυγής, που με την σειρά του βρίσκεται από τις ακμές μεταξύ τοίχων και δαπέδου (σημεία 1-2 και 5-6).

Έχοντας την πληροφορία για την θέση των ακμών στην προηγούμενη εικόνα, ορίζεται για κάθε μία η περιοχή της καινούριας εικόνας όπου θα αναζητηθεί. Η περιοχή αυτή είναι ένα παραλληλόγραμμο κατά μήκος της ακμής, με τον ίδιο προσανατολισμό με αυτή. Το πλάτος του παραλληλόγραμμου καθορίζεται αρκετά μεγάλο, ώστε να διασφαλιστεί το γεγονός ότι θα συμπεριλαμβάνει την ακμή, σε περίπτωση που οι διαδοχικές εικόνες διαφέρουν αισθητά μεταξύ τους. Το μήκος του είναι επίσης αρκετά μεγάλο, ώστε να συμπεριλάβει όσο το δυνατόν περισσότερα σημεία της ακμής. Τα μεγέθη αυτά διαφέρουν ανάλογα με την ακμή που θέλουμε να εντοπίσουμε, και κυμαίνονται μεταξύ 60-80 pixels για το

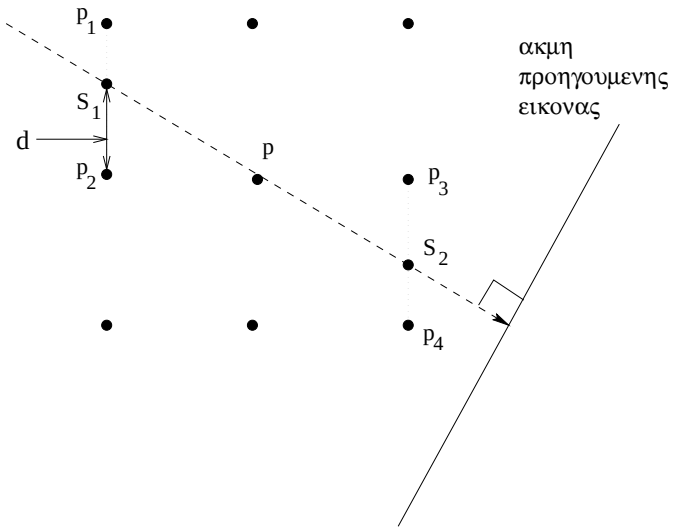
πλάτος και 80-140 pixels για το μήκος. Εάν ξέρουμε ότι δεν υπάρχει μεγάλη διαφορά μεταξύ των εικόνων, μπορούμε να μειώσουμε το μέγεθος της περιοχής και να επιταχύνουμε έτσι ακόμη περισσότερο τη μέθοδο.

3.3.2 Παραγώγιση

Σε κάθε σημείο των περιοχών που ορίστηκαν παίρνουμε την παράγωγο, για να εντοπίσουμε τα σημεία που ανήκουν στη ζητούμενη ακμή. Ο τύπος της παραγώγισης είναι απλός : είναι η απόλυτη διαφορά των φωτεινοτήτων των γειτονικών του σημείων. Στο σχήμα 3.6, η τιμή d_p της οριζόντιας παραγώγου στο pixel p είναι

$$d_p = |f_{p_2} - f_{p_3}|$$

όπου f_{p_i} η φωτεινότητα του pixel p_i.



Σχήμα 3.6: Παραγώγιση με sub-pixel υπολογισμό

Για καλύτερα αποτελέσματα, η παραγώγιση γίνεται σε κατεύθυνση κάθετη στην κλίση της ακμής στην προηγούμενη εικόνα (σχήμα 3.6), αφού περιμένουμε η ακμή στην καινούρια εικόνα να έχει παρόμοια κλίση. Αυτό περιπλέκει τον υπολογισμό της παραγώγου, αν η κλίση αυτή δεν είναι οριζόντια ή κάθετη, αφού δεν μπορούν να χρησιμοποιηθούν απ' ευθείας οι φωτεινότητες των γειτονικών pixels. Σε αυτή την περίπτωση, καταφεύγουμε στον υπολογισμό των φωτεινοτήτων των *υποσημείων* (subpixels) s_1 και s_2. Οι φωτεινότητα του s_1 υπολογίζεται από τις φωτεινότητες των pixels p_1 και p_2 ως εξής :

$$f(s_1) = dp_1 + (1-d)p_2$$

όπου d η απόσταση μεταξύ των s_1 και p_2, που υπολογίζεται εύκολα από τη γνωστή κλίση της κατεύθυνσης παραγώγισης. Στην ουσία υπολογίζουμε το σταθμισμένο άθροισμα των φωτεινοτήτων των δύο pixels. Ανάλογα, λόγω συμμετρίας, θα ισχύει

$$f(s_2) = dp_4 + (1-d)p_3$$

Η τιμή της παραγώγου d_p στο σημείο p θα δίνεται τώρα από την

$$d_p = |(f_{s_1}) - (f_{s_1})|$$

Στη συνέχεια, εφαρμόζουμε ένα κατώφλι(threshold) στις τιμές της παραγώγου που υπολογίστηκαν. Η τιμή του κατωφλίου καθορίζεται στην εφαρμογή μας περίπου στο 20[2]. Τα pixels με τιμές πάνω από το κατώφλι θεωρούνται σαν σημεία που ανήκουν στην ακμή και κρατούνται στη μνήμη. Το αποτέλεσμα της παραγώγισης δηλαδή είναι τελικά ένα σύνολο από σημεία.

3.3.3 Από τα σημεία στην ευθεία

Τα σημεία που προκύπτουν από την παραγώγιση σε κάθε περιοχή περιμένουμε ότι θα ανήκουν στην ακμή που αναζητάμε (τα περισσότερα τουλάχιστον). Το επόμενο βήμα που πρέπει να γίνει είναι να χαράξουμε μία ευθεία που να προσεγγίζει τα σημεία αυτά όσο ”καλύτερα” γίνεται. Για το σκοπό αυτό χρησιμοποιήθηκε ο αλγόριθμος **Least Median of Squares** (**LMedS**, μέθοδος ελάχιστου διαμέσου τετραγώνων)[Rou84]. Ακολουθεί μία παρουσίαση του αλγορίθμου και αναλύεται ο τρόπος με τον οποίο εφαρμόστηκε στην εργασία μας.

3.3.3.1 Least Median of Squares

Ο αλγόριθμος LMedS, που προτάθηκε αρχικά από τον Rousseau [Rou84], ανήκει στη γενική κατηγορία των αλγορίθμων για **robust regression**. Οι αλγόριθμοι αυτοί αναφέρονται στον υπολογισμό των παραμέτρων ενός μοντέλου από ανακριβή δεδομένα. Συγκεκριμένα, στη γενική περίπτωση ενός γραμμικού μοντέλου που περιγράφεται από τη σχέση

$$y_i = x_{i1}\theta_1 + \cdots + x_{ip}\theta_p + e_i$$

[2]Θυμίζουμε ότι αναφέρεται σε διαφορά φωτεινοτήτων.

το πρόβλημα είναι ο υπολογισμός των παραμέτρων $\theta_k, k = 1, \cdots, p$, από τις μετρήσεις y_i και τις βοηθητικές μεταβλητές x_{ik}. Ο όρος e_i παριστάνει το λάθος σε κάθε μέτρηση. Αν έχουμε τις *εκτιμήσεις* $\hat{\theta_1}, \cdots, \hat{\theta_p}$, τότε μπορούμε να έχουμε μία *πρόβλεψη* των μετρήσεων :

$$\hat{y_1} = x_{i1}\hat{\theta_1} + \cdots + x_{ip}\hat{\theta_p} + e_i$$

Έτσι, μπορούμε να ορίσουμε τις αποκλίσεις μεταξύ των τιμών που προβλέπει το μοντέλο μας και των παρατηρούμενων μετρήσεων :

$$r_i = y_i - \hat{y_i}$$

Για την εκτίμηση των παραμέτρων $\theta_k, k = 1, \cdots, p$, η κλασική προσέγγιση είναι η χρήση του αλγορίθμου *ελάχιστων τετραγώνων*, ή *least squares method*(LS). Ο LS επικεντρώνεται σε ένα πρόβλημα ελαχιστοποίησης :

$$Minimize \sum_{i=1}^{n} r_i^2$$

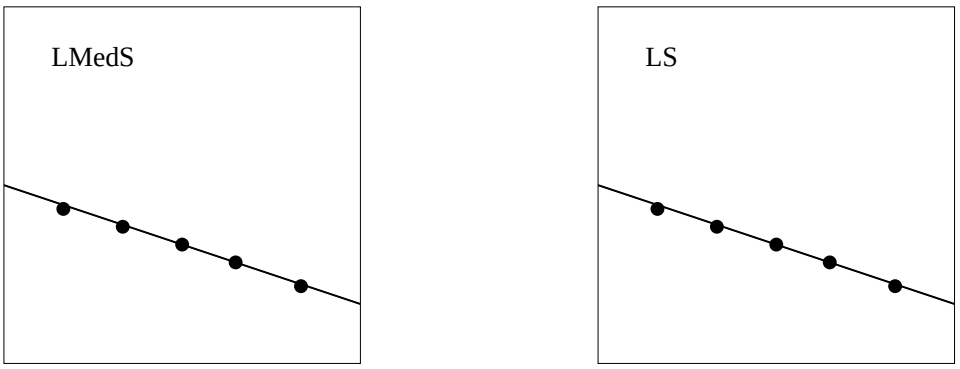
και είναι ευρέως διαδεδομένος, λόγω του χαμηλού υπολογιστικού του κόστους. Το κύριο χαρακτηριστικό του αλγορίθμου αυτού είναι ότι βρίσκει ένα μοντέλο που ταιριάζει καλύτερα σε *όλες* τις μετρήσεις. Αυτός είναι ο λόγος που ο LS δίνει αναξιόπιστα αποτελέσματα σε περιπτώσεις που κάποιες μετρήσεις διαφέρουν αισθητά από το μοντέλο που ταιριάζει στις υπόλοιπες. Τέτοιες περιπτώσεις παρουσιάζονται π.χ. όταν υπάρχει θόρυβος, ή όταν στις παρατηρήσεις εισέρχονται και κάποιες μετρήσεις που *δεν* ανήκουν στο μοντέλο, λόγω της φύσεως του προβλήματος[3]. Οι αλγόριθμοι robust regression αντιμετωπίζουν αυτό ακριβώς το πρόβλημα.

Ο LMedS συγκεκριμένα, χρησιμοποιεί ένα διαφορετικό κριτήριο αξιολόγησης της εκτίμησης των παραμέτρων του μοντέλου :

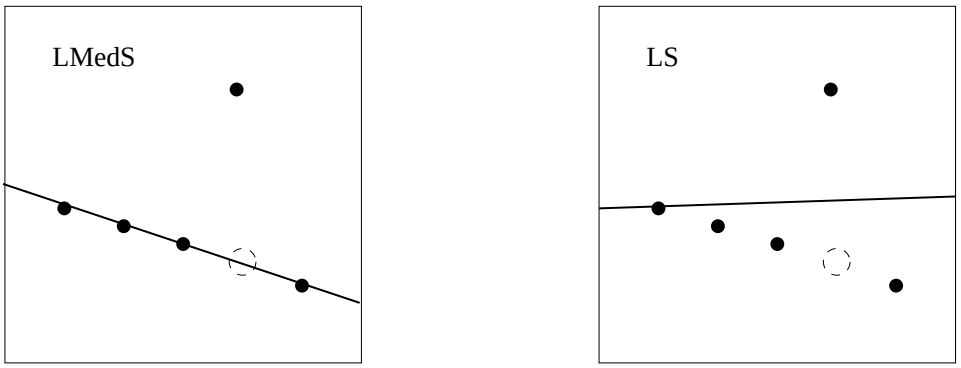
$$Minimize\{median\, r_i^2\},\ i = 1, \cdots, n$$

όπου median η *διάμεσος* των υπολογισμένων αποκλίσεων (αν δηλαδή ταξινομήσουμε τις αποκλίσεις σε ένα πίνακα n στοιχείων, τότε ο median είναι η τιμή που βρίσκεται στη θέση n/2). Διασθητικά, ο LMedS προσπαθεί να βρει ένα μοντέλο που να ταιριάζει καλύτερα στην *πλειοψηφία* των μετρήσεων. Έτσι το τελικό αποτέλεσμα δεν επηρεάζεται καθόλου από την ύπαρξη κάποιων μετρήσεων διαταραγμένων από θόρυβο. Το ποσοστό των μετρήσεων αυτό μπορεί να φτάσει και το 50%. Μία οπτική σύγκριση των τρόπων με τους οποίους χειρίζονται τέτοιες μετρήσεις οι LS και LMedS φαίνεται στο σχήμα 3.7.

[3]Το πρόβλημα αυτό παρουσιάζεται και στη δική μας περίπτωση.



(α) Αξιοπιστες μετρησεις



(β) Εισαγωγη θορυβου σε μια μετρηση

Σχήμα 3.7: Σύγκριση μεθόδου ελαχίστων τετραγώνων (LS) και Least Median of Squares (LMedS)

Το υπολογιστικό κόστος του LMedS θεωρείται μεγάλο [PMR91]. Αυτό συμβαίνει γιατί το πρόβλημα ελαχιστοποίησης που περιλαμβάνει είναι μη γραμμικό και επιλύεται με αναζήτηση στο χώρου των λύσεων (οι παράμετροι του μοντέλου θ_k). Στην δική μας χρήση του αλγορίθμου ωστόσο, το υπολογιστικό κόστος ήταν μικρό για τους εξής λόγους :

1. Όπως θα δούμε παρακάτω, οι παράμετροι που πρέπει να εκτιμηθούν είναι λίγες.

2. Οι μετρήσεις στην εφαρμογή μας αντιστοιχούν στα σημεία που προέκυψαν από την παραγώγιση. Με δεδομένο ότι οι περιοχές που εξετάζουμε είναι μικρές, τα σημεία είναι επίσης σχετικά λίγα.

3. Η αναζήτηση στο χώρο λύσεων δεν γίνεται εντελώς τυχαία, αλλά με τρόπο που να εγγυάται ότι κάθε υποψήφια λύση μπορεί να είναι η βέλτιστη που

ψάχνουμε.

Τα σημεία αυτά θα γίνουν πιο εύκολα κατανοητά στην περιγραφή της χρήσης του αλγορίθμου που ακολουθεί.

3.3.3.2 Η χρήση του LMedS στη μέθοδο μας

Πρέπει αρχικά να τονίσουμε ότι ο LMedS ταιριάζει με φυσικό τρόπο στο πρόβλημα μας, της εύρεσης μιας ευθείας που να ταιριάζει στα σημεία που προέκυψαν από την παραγώγιση. Αυτό συμβαίνει γιατί :

1. Αυτό που ψάχνουμε είναι μία ευθεία, θέλουμε δηλαδή να εκτιμήσουμε τα α και β της εξίσωσης

$$y_i = \alpha x_i + \beta$$

   που είναι ένα γραμμικό μοντέλο, με τα α και β να είναι οι δύο παράμετροι θ_1 και θ_2 που θέλουμε να υπολογίσουμε. Για κάθε σημείο (x_i, y_i) μπορούμε να θεωρήσουμε την τεταγμένη y_i σαν την μέτρηση και την τετμημένη x_i σαν βοηθητική μεταβλητή. Σαν απόκλιση r_i του μοντέλου μας και της παρατηρούμενης μέτρησης ορίζουμε την $απόσταση$ του σημείου από την ευθεία που προκύπτει.

2. Η περιοχή στην οποία γίνεται η παραγώγιση σχεδόν σίγουρα θα περιέχει και άλλες, μικρότερες ακμές. Τα σημεία που θα προκύψουν από αυτές τις ακμές μπορούν να θεωρηθούν σαν ανακριβείς μετρήσεις, αφού δεν θα βρίσκονται στην ίδια ευθεία με αυτή που ψάχνουμε (με άλλα λόγια δεν θα ταιριάζουν στο ζητούμενο μοντέλο). Είναι φανερό ότι δεν πρέπει αυτά τα σημεία να επηρεάσουν το αποτέλεσμα. Κάνοντας την λογική υπόθεση ότι αυτά τα σημεία αποτελούν μειοψηφία στο σύνολο των σημείων που έχουμε, ο LMedS δεν θα τα λάβει καθόλου υπ' όψιν για την χάραξη της ευθείας.

Το πρόβλημα της αναζήτησης στο χώρο λύσεων στη συγκεκριμένη εφαρμογή του αλγορίθμου, δηλαδή του καθορισμού των υποψήφιων ευθειών, είναι απλό. Τα περισσότερα σημεία θα είναι λίγο-πολύ συνευθειακά, και επομένως πάνω ή πολύ κοντά στη ζητούμενη ευθεία. Μπορούμε λοιπόν να πάρουμε τυχαία ένα ζεύγος σημείων και να βρούμε την ευθεία που ορίζεται από αυτά. Με αρκετές δοκιμές τέτοιων ευθειών, μπορούμε να περιμένουμε πως θα πετύχουμε τη βέλτιστη λύση, ή κάποια πολύ κοντά σ' αυτή.

Μετά από τα παραπάνω, μπορούμε να συνοψίσουμε τη χρήση του LMedS :

```
/* χοντας το σύνολο σημείων που
   προέκυψε από την παραγώγιση ... */
Επανέλαβε 50 φορές {
    Πάρε τυχαία δύο σημεία.
    Βρες την ευθεία που ορίζεται από αυτά.
    Για κάθε σημείο,
       υπολόγισε την απόσταση του από την ευθεία.
    Βρες την διάμεσο (median) των υπολογισμένων αποστάσεων.
}
Επέλεξε τον καλύτερο από τους medians.
Δώσε σαν λύση την αντίστοιχη ευθεία.
```

Η επιλογή του διαμέσου των αποστάσεων μπορεί να γίνει γρήγορα ($O(n)$), χωρίς να χρειάζεται ταξινόμησή τους. Ο έλεγχος 50 ευθειών είναι αρκετός για μία καλή προσέγγιση της βέλτιστης λύσης, όπως θα φανεί στα αποτελέσματα.

3.3.4 Τμηματοποίηση εικόνας

Εφαρμόζοντας την παραπάνω διαδικασία για όλες τις ακμές, καταλήγουμε στην εύρεση αντίστοιχων ευθειών. Η τελική τμηματοποίηση της εικόνας προκύπτει από τα σημεία τομής αυτών των ευθειών. Σε περίπτωση που τρεις ευθείες θα έπρεπε να συναντώνται στο ίδιο σημείο (π.χ. σημείο 2 στο σχήμα 3.5) και αυτό δεν συμβαίνει, το τριπλό σημείο τομής εξάγεται με τον συνυπολογισμό και των τριών ευθειών (μέσος όρος των τριών σημείων τομής). Η χρήση της μεθόδου έδειξε ότι οι διαφορές που προκύπτουν είναι σχεδόν αμελητέες.

3.4 Ορόσημα[4]

Μετά την ποιοτική τμηματοποίηση της εικόνας, αρχίζει η αναζήτηση των οροσήμων στα τμήματα εκείνα της εικόνας που χαρακτηρίστηκαν σαν τοίχοι. Όπως αναφέρθηκε στο κεφάλαιο 2, τα ορόσημα επιλέγονται από το αυτόνομο σύστημα χρησιμοποιώντας το saliency map, όπου εντοπίζονται τα πιο χαρακτηριστικά σημεία των τμημάτων που αναφέρθηκαν παραπάνω. Στη συνέχεια, ένας αλγόριθμος αύξησης περιοχής αναλαμβάνει να εντοπίσει ολόκληρο το ορόσημο. Στη συνέχεια θα περιγράψουμε τις συναρτήσεις που χρησιμοποιήθηκαν για τον υπολογισμό του saliency map και τον αλγόριθμο αύξησης περιοχής.

[4]Η υλοποίηση του τμήματος αυτού έγινε από τον Θόδωρο Γκαράβελο, στα πλαίσια της πτυχιακής του εργασίας

3.4.1 Συναρτήσεις saliency map

Οι συναρτήσεις που χρησιμοποιήσαμε έχουν την ιδιότητα να μετρούν την ανομοιογένεια περιοχών της εικόνας. Χρησιμοποιήσαμε μία συνάρτηση συσχετίσεως με δύο μορφές, αφού μετά από πειραματισμούς φάνηκε ότι αποδίδουν καλά και οι δύο. Αναλυτικά, όλες οι συναρτήσεις είναι οι εξής :

1. Μία συνάρτηση R_N συσχετίσεως μικρών περιοχών της εικόνας με τύπο

$$R_N = \frac{C_{MN}}{\sigma_M * \sigma_N} \tag{3.1}$$

με

$$C_{MN} = \mu_{MN} - \mu_M * \mu_N$$

$$\sigma_N = \mu_{NN} - \mu_N * \mu_N$$

$$\sigma_M = \mu_{MM} - \mu_M * \mu_M$$

όπου μ_N η μέση τιμή στην περιοχή N, μ_{NN} η μέση τιμή των τετραγώνων των τιμών της περιοχής N και σ_N η διασπορά στο παράθυρο N. Είναι φανερό ότι παρόμοιες περιοχές θα δίνουν μεγαλύτερη τιμή στην R_N, ενώ περιοχές που διαφέρουν θα δίνουν μικρότερες τιμές, γι' αυτό η συνάρτηση χρησιμοποιείται ανεστραμμένη στον υπολογισμό του saliency map.

2. Άλλη μια συνάρτηση συσχετίσεως που δίνεται από τον τύπο

$$\rho(x,y) = \frac{\sum\sum_{i,j\in S}[I(x,y) - I(x+i,y+j)]^2}{\#pixels\ in\ S}$$

$$C = \sum \sum_{i,j\in W} \rho(i,j) \tag{3.2}$$

όπου W μία περιοχή της εικόνας και S μία περιοχή του ίδιου μεγέθους με κέντρο το (x,y). Η τιμή αυτής της συνάρτησης είναι μεγάλη ανάμεσα σε περιοχές που διαφέρουν μεταξύ τους.

3. Την εντροπία H(s) της εικόνας, που δίνεται από τον τύπο

$$H(h) = -\sum_i P(h_i) \log_2 P(h_i) \tag{3.3}$$

όπου $P(h_i)$ η πιθανότητα ένα pixel να έχει τιμή i. Η εντροπία στην προσέγγισή μας αποτελεί ένα μέτρο της ανομοιότητας, αφού έχει μεγάλες τιμές σε περιοχές με διαφορετικές φωτεινότητες.

4. Το *κατώφλι σημείων* (pixel threshold) είναι ένα μέτρο ανομοιότητας. Αυτό που μετρά είναι ο αριθμός των διαφορετικών φωτεινοτήτων μέσα σε μία περιοχή της εικόνας. Ένας αριθμός μεγαλύτερος από κάποιο κατώφλι υποδεικνύει αυξημένη ανομοιογένεια στη συγκεκριμένη περιοχή.

5. Ένα τελευταίο μέτρο προκύπτει από τη διαφορά μεταξύ διαδοχικών εικόνων. Αυτό που υπολογίζεται είναι αρχικά η απόλυτη διαφορά μεταξύ των φωτεινοτήτων των εικόνων :

$$f_i = \mid f_{1i} - f_{2i} \mid$$

Στη συνέχεια, σε μικρές περιοχές της εικόνας υπολογίζεται η διασπορά των διαφορών αυτών :

$$D = \sum_i \frac{(f_i - \mu)^2}{N} \tag{3.4}$$

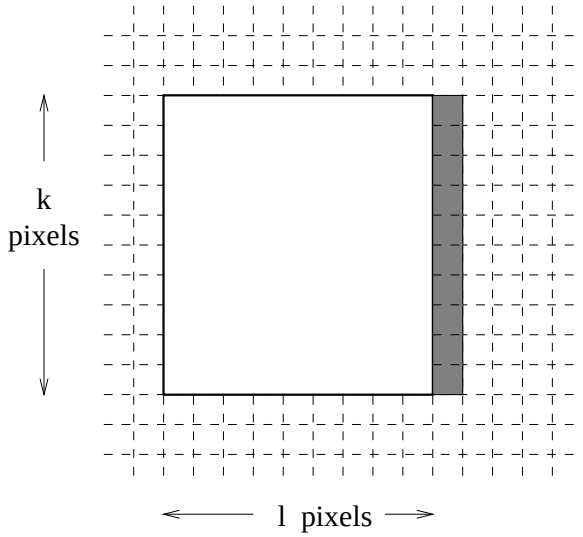
όπου N το πλήθος των σημείων της περιοχής και μ ο μέσος των διαφορών φωτεινοτήτων f_i. Όσο πιο μεγάλη η διασπορά των φωτεινοτήτων, τόσο μεγαλύτερη η διαφορά των διαδοχικών εικόνων και κατ' επέκταση και η ανομοιογένεια της περιοχής.

Οι τιμές των παραπάνω συναρτήσεων αθροίζονται για την τιμή του saliency map σε κάθε σημείο, αφού πρώτα σταθμιστούν κατάλληλα, ώστε να συμμετέχουν στο αποτέλεσμα με την ίδια βαρύτητα.

3.4.2 Αλγόριθμος αύξησης περιοχής

Μετά τη δημιουργία του saliency map και την επιλογή ενός σημείου με μεγάλη τιμή, ο αλγόριθμος αύξησης περιοχής αναλαμβάνει να εντοπίσει ολόκληρο το τμήμα της εικόνας που περιέχει το ορόσημο. Οι υποθέσεις που γίνονται σχετικά με το ορόσημο, είναι ότι έχει μέγεθος τουλάχιστον 7x7 pixels στην εικόνα και

ότι το αρχικό σημείο που προέκυψε από το saliency map βρίσκεται μέσα στην περιοχή της εικόνας που καταλαμβάνει το ορόσημο. Έτσι, ξεκινώντας από το παράθυρο με μέγεθος 7x7 pixels που περιβάλλει το αρχικό σημείο στην grey-scaled εικόνα, ο αλγόριθμος προσθέτει διαδοχικές γραμμές και στήλες, μέχρι το ορόσημο να συμπεριληφθεί στο παράθυρο. Η συνθήκη τερματισμού εξαρτάται από τις φωτεινότητες των υποψήφιων σημείων. Εάν η μέση τιμή των σημείων διαφέρει από τη μέση τιμή των σημείων που έχουν ήδη συμπεριληφθεί στο ορόσημο περισσότερο από κάποιο ποσοστό, η επέκταση προς την συγκεκριμένη κατεύθυνση σταματά.



Σχήμα 3.8: Αλγόριθμος αύξησης περιοχής

Στο σχήμα 3.8, τα καινούρια σημεία θα συμπεριληφθούν στο ορόσημο μόνο εάν

$$\frac{\mid Mean - NewMean \mid}{100} < p \tag{3.5}$$

όπου

$$Mean = \frac{\sum f_{i,j}}{kl} \tag{3.6}$$

ο μέσος όρος των φωτεινοτήτων στο τμήμα της εικόνας που έχουμε ήδη συμπεριλάβει στο ορόσημο (το άσπρο τετράγωνο στο σχήμα),

$$NewMean = \frac{\sum f_{i,j}}{l} \tag{3.7}$$

ο μέσος όρος των φωτεινοτήτων των υποψήφιων σημείων (το γραμμοσκιασμένο κομμάτι στο σχήμα) και p το ποσοστό επιτρεπόμενης διαφοράς μεταξύ των δύο μέσων όρων για να συνεχιστεί η επέκταση.

Όταν ολοκληρωθεί η επέκταση, το τμήμα της εικόνας αποθηκεύεται στη μνήμη, μαζί με τις διαστάσεις του και τις συντεταγμένες της επάνω αριστερά γωνίας του.

Κεφάλαιο 4

Αποτελέσματα

4.1 Εισαγωγή

Για την καλύτερη αξιολόγηση της προτεινόμενης μεθόδου, θα χρειαζόταν η εφαρμογή της σε ένα αυτόνομο σύστημα, σε διαφορετικούς χώρους, σε συνθήκες πραγματικού χρόνου. Αυτό ωστόσο θα απαιτούσε πολύ προσπάθεια για την προσαρμογή της μεθόδου στο αυτόνομο σύστημα. Επιλέξαμε λοιπόν να αξιολογήσουμε τη μέθοδο εφαρμόζοντάς την σε μία ακολουθία εικόνων από τον διάδρομο ενός κτιρίου[1]. Η ακολουθία αυτή προέκυψε από την προς τα εμπρός κίνηση μιας κάμερας μέσα στο διάδρομο. Με αυτό τον τρόπο προσομοιώθηκε η όψη του περιβάλλοντος από το αυτόνομο σύστημα, καθώς αυτό κινείται. Διαδοχικές εικόνες στην ακολουθία διαφέρουν λίγο μεταξύ τους, όπως περιμένουμε να συμβαίνει και στην πραγματικότητα, αφού θεωρούμε ότι το αυτόνομο σύστημα θα λαμβάνει συνεχώς εικόνες κατά τη διάρκεια της πλοήγησης.

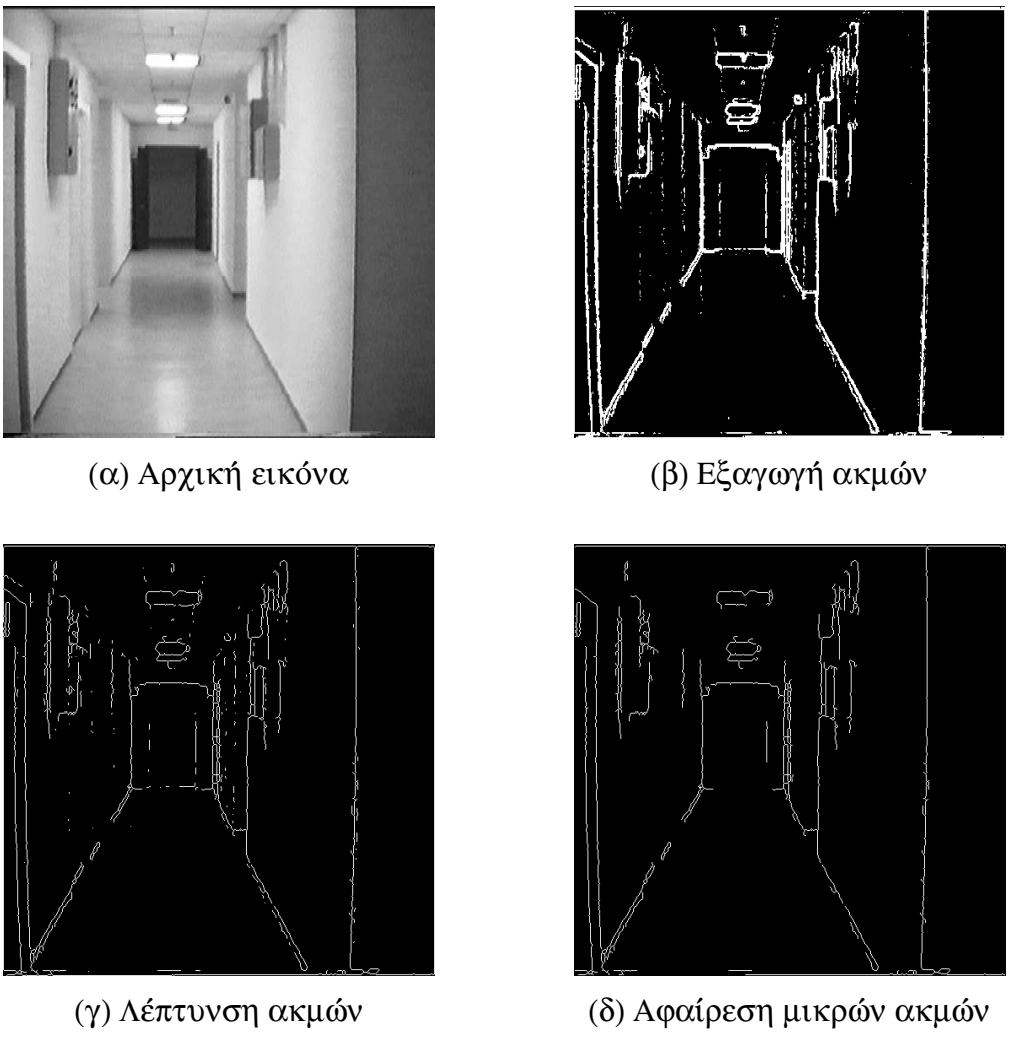
Το βασικό μειονέκτημα της χρήσης μίας τεχνητής ακολουθίας εικόνων, σαν δεδομένα για την εφαρμογή της μεθόδου, είναι ότι δεν έχουμε τη δυνατότητα να ελέγξουμε τη διαφοροποίηση που θα προκαλούν οι (πιθανώς σκόπιμες) κινήσεις του αυτόνομου συστήματος στην όψη του περιβάλλοντος που θα λαμβάνει. Κάθε κίνηση του αυτόνομου συστήματος που θα αλλάζει τη θέση ή τον προσανατολισμό του, θα έχει σαν αποτέλεσμα να καταγράφεται μία διαφορετική εικόνα από τις κάμερες του. Οι εικόνες που θα προκύπτουν έτσι σίγουρα θα διαφέρουν, λίγο ή πολύ, από τη δικιά μας ακολουθία εικόνων.

Δεν έχουμε τρόπο λοιπόν να ελέγξουμε την επίδραση των κινήσεων του συστήματος στα οπτικά δεδομένα που θα χρησιμοποιήσει. Έτσι, δεν μπορούμε να δούμε την πραγματική εφαρμογή της μεθόδου κατά τη διάρκεια της πλοή-

[1] Ίδρυμα Τεχνολογίας και Έρευνας

γησης. Μπορούμε όμως, χρησιμοποιώντας την τεχνητή ακολουθία εικόνων, να πάρουμε αποτελέσματα σχετικά με τα κυριότερα σημεία της μεθόδου : την τμηματοποίηση του χώρου από την εικόνα, τον εντοπισμό και την αναγνώριση των οροσήμων. Τα αποτελέσματά αυτά παρουσιάζονται αναλυτικά παρακάτω, με την εφαρμογή της μεθόδου σε μερικές από τις εικόνες της ακολουθίας που χρησιμοποιήθηκε.

4.2 Αποτέλεσμα 1



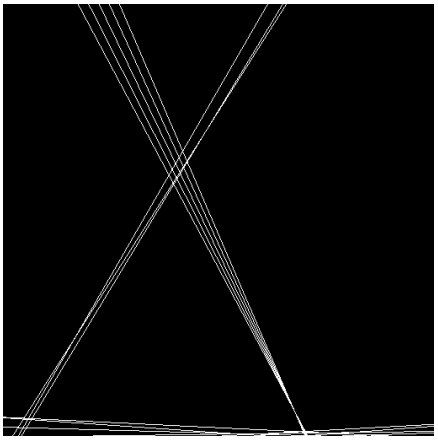
(α) Αρχική εικόνα (β) Εξαγωγή ακμών

(γ) Λέπτυνση ακμών (δ) Αφαίρεση μικρών ακμών

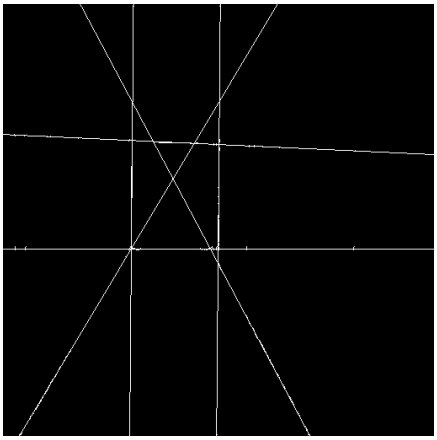
Σχήμα 4.1: Αρχική επεξεργασία εικόνας

Στο σχήμα 4.1 φαίνονται τα αρχικά στάδια επεξεργασίας της εικόνας στην

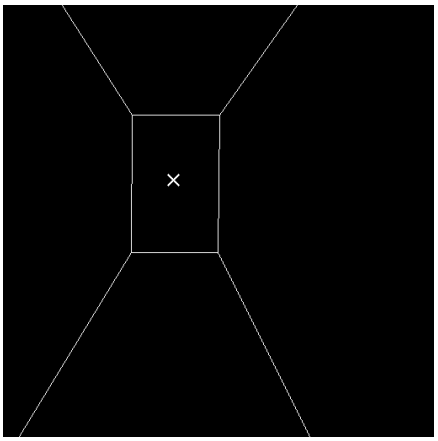
περίπτωση της στατικής τμηματοποίησης. Το τελικό αποτέλεσμα (σχήμα 4.1-δ), μετά την αφαίρεση των μικρών ακμών δίνεται σαν είσοδος στο μετασχηματισμό Hough. Στο σχήμα 4.2-α βλέπουμε το αποτέλεσμα του μετασχηματισμού Hough σε όλη την εικόνα. Διακρίνονται καθαρά οι προσεγγίσεις για δύο από τις βασικές ακμές. Στο σχήμα 4.2-β βλέπουμε το αποτέλεσμα των διαδοχικών εφαρμογών του AHT και στο σχήμα 4.2-γ έχουμε το τελικό αποτέλεσμα της τμηματοποίησης.



(α) Hough



(β) AHT



(γ) Τελικό αποτέλεσμα

Σχήμα 4.2: Μετασχηματισμός Hough

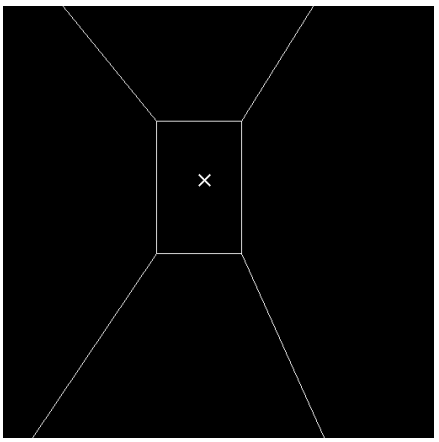
Η μέθοδος της δυναμικής τμηματοποίησης παρουσιάζεται στο σχήμα 4.3. Η εικόνα 4.3-α είναι η αμέσως επόμενη μετά την 4.1-α στην χρονική ακολουθία. Μπορούμε λοιπόν να καθορίσουμε τις περιοχές αναζήτησης των ακμών, με βάση την ήδη γνωστή τμηματοποίηση της προηγούμενης εικόνας (σχήμα 4.2-γ). Ο ορισμός των περιοχών φαίνεται στο σχήμα 4.3-β, όπου βλέπουμε επίσης το

αποτέλεσμα της παραγώγισης στις περιοχές αναζήτησης των ακμών. Στο σχήμα 4.3-γ έχουμε το αποτέλεσμα της εφαρμογής του LMedS στα σύνολα σημείων που προκύπτουν. Οι γραμμές που προκύπτουν συνδυάζονται για το τελικό αποτέλεσμα που φαίνεται στο σχήμα 4.3-δ.

(α) Αρχική εικόνα

(β) Παραγώγιση σε περιοχές

(γ) Εφαρμογή του LMedS



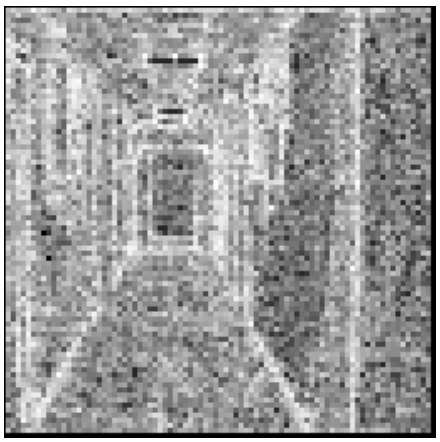
(δ) Τελικό αποτέλεσμα

Σχήμα 4.3: Δυναμική τμηματοποίηση

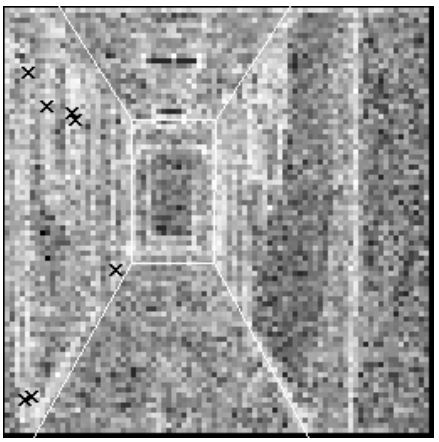
Στο σχήμα 4.4-α φαίνεται ο saliency map για την πρώτη εικόνα (σχήμα 4.1-α). Τα σημεία με τις μεγαλύτερες τιμές αναζητούνται στους τοίχους, σύμφωνα με την τμηματοποίηση που προηγήθηκε. Στο σχήμα 4.4-β αυτά τα σημεία συμβολίζονται με ”Χ” και βρίσκονται στο πάνω-αριστερά μέρος της εικόνας. Στο ίδιο σχήμα βλέπουμε και την αντίστοιχη τμηματοποίηση, με βάση την οποία έγινε η αναζήτηση.

Επιλέγοντας από αυτά τα σημεία εκείνα με τη μεγαλύτερη τιμή, ο αλγόριθμος

αύξησης περιοχής εντοπίζει το ορόσημο που φαίνεται στο σχήμα 4.4-γ. Το ορόσημο αυτό εξάγεται από την εικόνα και αποθηκεύεται στη μνήμη του αυτόνομου συστήματος (σχήμα 4.6-α).



(α) Saliency map



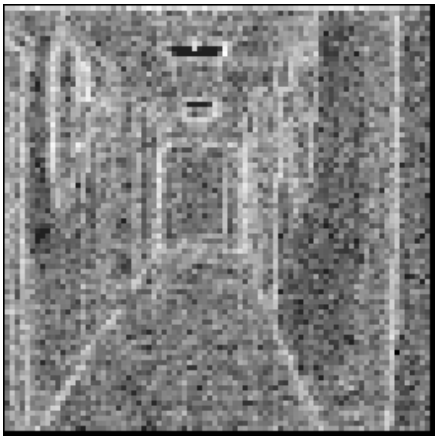
(β) Τα x στην εικόνα σημειώνουν τις μεγαλύτερες τιμές του saliency map (βλ. κείμενο)

(γ) Ορόσημο πρώτης εικόνας

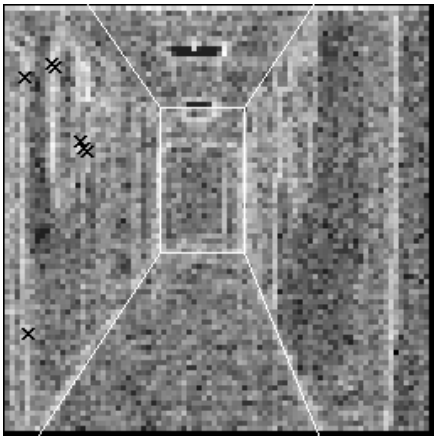
Σχήμα 4.4: Saliency map και εξαγωγή οροσήμου (1η εικόνα)

Στο σχήμα 4.5 βλέπουμε την ίδια διαδικασία για τον εντοπισμό και την εξαγωγή του οροσήμου στη δεύτερη εικόνα : Υπολογίζεται ο saliency map (σχήμα 4.5-α), εντοπίζονται οι μεγάλες τιμές (σχήμα 4.5-β) και τελικά το ορόσημο εντοπίζεται (σχήμα 4.5-γ) και εξάγεται από την εικόνα. Το εξαγώμενο ορόσημο φαίνεται στο σχήμα 4.6-β είναι το ίδιο με εκείνο της πρώτης εικόνας, με τη διαφορά ότι το βλέπουμε από διαφορετική οπτική γωνία.

Η αναγνώριση του οροσήμου, απαιτεί τη σύγκρισή του με το ορόσημο που



(α) Saliency map



(β) Μεγαλύτερες τιμές του saliency map

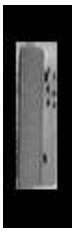
(γ) Ορόσημο δεύτερης εικόνας

Σχήμα 4.5: Saliency map και εξαγωγή οροσήμου (2η εικόνα)

βρίσκεται αποθηκευμένο στη μνήμη. Πριν από τη σύγκριση, εφαρμόζουμε στο ορόσημο αυτό το Μετασχηματισμό Θέσης Παρατηρητή και τη Διγραμμική παρεμβολή. Το αποτέλεσμα αυτών των μετασχηματισμών πάνω στο ορόσημο 4.6-α φαίνεται στο σχήμα 4.6-γ. Η τελική σύγκριση θα πραγματοποιηθεί μεταξύ των οροσήμων 4.6-β και 4.6-γ. Η αναγκαιότητα του Μετασχηματισμού Θέσης Παρατηρητή και της Διγραμμικής Παρεμβολής φαίνεται από τις τιμές που παίρνει ο συντελεστής σύγκρισης (εξίσωση 2.20) στις συγκρίσεις του αποθηκευμένου στη μνήμη ορόσημο με το καινούριο ορόσημο, μετασχηματισμένο ή όχι. Ο συντελεστής ομοιότητας του οροσήμου στο σχήμα 4.6-β και του μη μετασχηματισμένου οροσήμου (σχήμα 4.6-α) ήταν 0.7, ενώ το αποτέλεσμα της σύγκρισής του με το μετασχηματισμένο ορόσημο του σχήματος 4.6-γ ήταν 0.8, δείχνοντας

μεγαλύτερη ομοιότητα. Η διαφορά ανάμεσα στα δύο αποτελέσματα γίνεται μεγαλύτερη, όταν οι εικόνες από τις οποίες εξάχθηκαν τα ορόσημα απέχουν περισσότερο μεταξύ τους (στο παράδειγμα που παρουσιάστηκε οι εικόνες ήταν διαδοχικές).

(α) Ορόσημο πρώτης εικόνας



(β) Ορόσημο δεύτερης εικόνας

(γ) ΜΘΠ και Διγραμμική Παρεμβολή

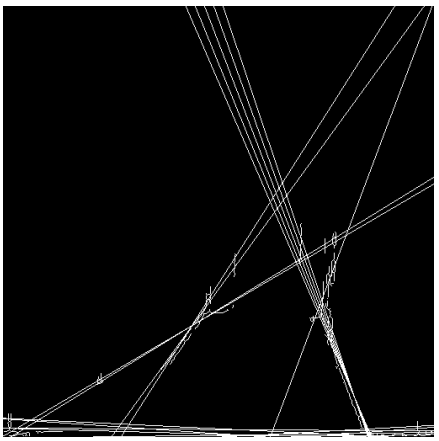
Σχήμα 4.6: Ορόσημα

4.3 Αποτέλεσμα 2

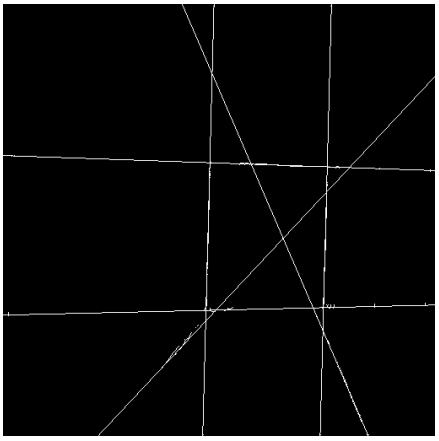
Παραθέτουμε ακόμα ένα παράδειγμα τμηματοποίησης του χώρου σε μία διαφορετική εικόνα, όπου δεν υπάρχει κάποιο ορόσημο. Στο σχήμα 4.7 (α, β, γ και δ) φαίνονται τα στάδια της στατικής τμηματοποίησης. Στο σχήμα 4.7 (ε και στ) παρουσιάζονται η δυναμική τμηματοποίηση με την παραγώγιση σε περιοχές και το τελικό αποτέλεσμα.

Στο σχήμα 4.8 βλέπουμε τον saliency map για τη συγκεκριμένη εικόνα. Οι μεγαλύτερες τιμές του δεν είναι αρκετά μεγάλες για να υποδηλώσουν την ύπαρξη κάποιου οροσήμου.

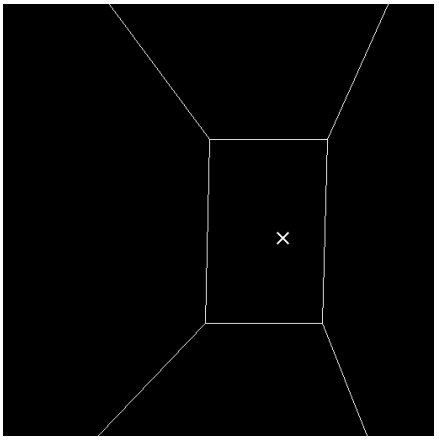
(α) Αρχική εικόνα



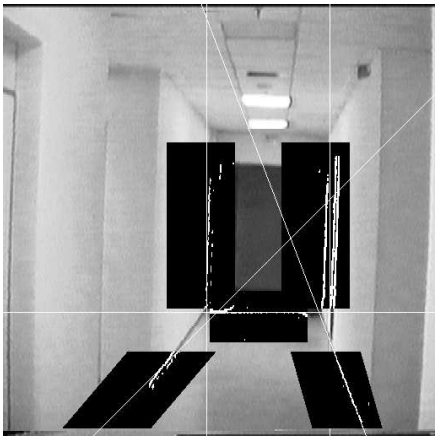
(β) Hough



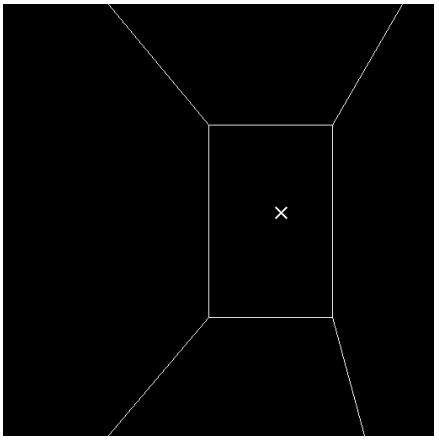
(γ) ΑΗΤ



(δ) Τμηματοποίηση

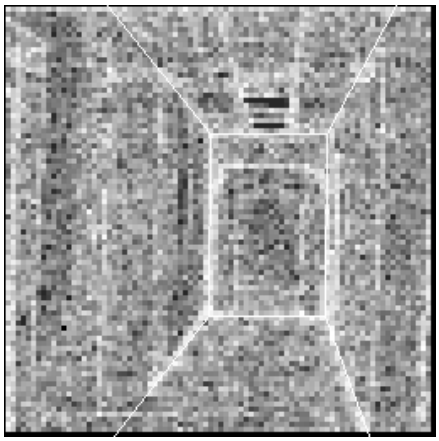


(ε) Παραγώγιση σε περιοχές



(στ) Τμηματοποίηση

Σχήμα 4.7: Στατική και δυναμική τμηματοποίηση



Σχήμα 4.8: Saliency map

Κεφάλαιο 5

Συμπεράσματα - Μελλοντική εργασία

Τα αποτελέσματα από αυτή την εργασία ήταν αρκετά ενθαρρυντικά σχετικά με την προτεινόμενη μέθοδο οπτικής πλοήγησης. Η τμηματοποίηση του χώρου, στατική και δυναμική, έδωσε αξιόπιστα αποτελέσματα στις περισσότερες περιπτώσεις. Η εφαρμογή της μεθόδου για την εξαγωγή μη προκαθορισμένων οροσήμων έδειξε ότι μπορούμε να ελπίζουμε σε καλά αποτελέσματα στην κατεύθυνση αυτή. Η αποφυγή της προεπιλογής των οροσήμων φανέρωσε ότι μπορούμε να βελτιώσουμε την ευελιξία και την προσαρμοσιμότητα των αυτόνομων συστημάτων. Επίσης, η μέθοδος αναγνώρισης των οροσήμων έδειξε ότι δουλεύει αρκετά αξιόπιστα, τουλάχιστον σε περιπτώσεις αντικειμένων που συναντούμε συχνά σε εσωτερικούς χώρους.

Ιδιαίτερα πρέπει να τονιστεί η αποφυγή εξαγωγής συμπερασμάτων σε υψηλό επίπεδο (high-level reasoning) σε ολόκληρη τη μέθοδο. Όλες οι πράξεις γίνονται σε επίπεδο εικόνας και όλες οι ποσότητες που είναι απαραίτητες για τους μετασχηματισμούς και την αναγνώριση των οροσήμων είναι μετρήσιμες πάνω στην εικόνα. Το γεγονός αυτό εγγυάται μεγαλύτερη αξιοπιστία για τα αποτελέσματα της μεθόδου.

Οι χρόνοι εκτέλεσης της μεθόδου στην ακολουθία εικόνων όπου εφαρμόστηκε υπήρξαν ικανοποιητικοί, αναλογικά με την δυσκολία του προβλήματος το οποίο αντιμετωπίζει. Σημαντικό ρόλο στην βελτίωση της ταχύτητας εκτέλεσης είχαν η χρήση προηγούμενης γνώσης για την τμηματοποίηση της εικόνας καθώς και η μη εξαντλητική αναζήτηση οροσήμων σε ολόκληρη την εικόνα, αλλά μόνο σε περιοχές της. Προς την κατεύθυνση της βελτίωσης της ταχύτητας μπορεί να προσφέρει η παράλληλη εκτέλεση τμημάτων της μεθόδου, όπως η τμηματοποίηση του χώρου και ο υπολογισμός του saliency map, που είναι ανεξάρτητα μεταξύ τους.

Περισσότερα συμπεράσματα μπορούν να εξαχθούν από την εφαρμογή της μεθόδου σε ένα πραγματικό αυτόνομο σύστημα, που είναι και ο αντικειμενικός προορισμός της. Είναι σίγουρο πως θα χρειαστεί να ξεπεραστούν πολλές δυσκολίες για να επιτευχθεί αυτός ο σκοπός. Η εφαρμογή της μεθόδου σε ένα πραγματικό σύστημα και οι πειραματισμοί με αυτό θα δείξουν με σαφή τρόπο τα πλεονεκτήματα ή μειονεκτήματα που παρουσιάζει η προτεινόμενη προσέγγιση στην αυτόνομη πλοήγηση.

Οπωσδήποτε η συγκεκριμένη εργασία,ακόμη και μετά την υλοποίηση και αξιολόγησή της σε ένα πραγματικό σύστημα, δεν εξαντλεί το πρόβλημα της οπτικής πλοήγησης. Κατ' αρχήν πρέπει να επεκταθεί η μέθοδος της τμηματοποίησης του χώρου, ώστε να συμπεριλάβει τις προβληματικές περιπτώσεις που δεν αντιμετωπίζει αυτή η εργασία (όπως η περίπτωση να μην είναι ορατές όλες οι ακμές). Ένα επόμενο βήμα θα ήταν να αντιμετωπιστούν οι περιπτώσεις όπου δεν ισχύουν όλες οι παραδοχές που έγιναν σχετικά με τους χώρους κίνησης. Με αυτό τον τρόπο το αυτόνομο σύστημα μπορεί να γίνει ακόμα πιο ευέλικτο.

Αρκετή ακόμα προσπάθεια και σκέψη πρέπει να αφιερωθεί στον τρόπο εξαγωγής οροσήμων. Οι συναρτήσεις που χρησιμοποιήθηκαν σ' αυτή την εργασία για τη δημιουργία του saliency map πιθανώς να μην είναι οι βέλτιστες, ή να πρέπει να συμπληρωθούν. Ενδιαφέρον είναι ακόμη να σκεφτούμε αν πρέπει να δεχόμαστε σαν ορόσημα οτιδήποτε μας δίνει η ανάλυση του saliency map, ή αν πρέπει να γίνεται κάποια επιλογή με βάση κάποιο κριτήριο καταλληλότητας των αντίστοιχων περιοχών της εικόνας.

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 αποτελεσμάτων και θα βελτιώσει την ταχύτητά του.

Τέλος, η προτεινόμενη μέθοδος πλοήγησης θα μπορούσε να συνδυαστεί με άλλες προσεγγίσεις (π.χ. χάρτες, προκαθορισμένα ορόσημα) προς την κατεύθυνση της βελτίωσης της ευελιξίας του αυτόνομου συστήματος, διατηρώντας την ποιότητα των αποτελεσμάτων.

Βιβλιογραφία

[AB87] J. Aloimonos and A. Badyopadhyay. Active vision. In *1991 IEEE 1st Int. Conf. on Computer Vision*, pages 35--54, June 1987.

[AKLA89] K. Andress A. Kak and C. Lopez-Albadia. Mobile robot self-location with the pseiki system. In *Working Notes, AAAI Spring Symposium on Robot Navigation, Cambridge*, March 1989.

[Alo90] J. Aloimonos. Purposive and qualitative vision. In *DARPA Image Understanding Workshop*, pages 816--828, 1990.

[Arg96] A. Argyros. Visual detection of independent 3d motion by a moving observer. *Ph.D.thesis. University of Crete, Computer Science Department*, 1996.

[Bal91] D.H. Ballard. Animate vision. *Artificial Intelligence*, 48:57--86, 1991.

[Blu67] H. Blum. A tranformation for extracting new descriptors of shape. *Models for the Perception of Speech and Visual Form*, 1967.

[CA95] Y.L. Chang and J.K. Aggarwalls. Representing and estimeting 3-d lines. *Pattern Recognition*, 28(8):1181--1190, 1995.

[Can88] J. F. Canny. The complexity of robot motion planning. *MIT Press, Cambridge, MA*, 1988.

[CF93] J. Clark and N. Ferrier. *Active Vision*, chapter 9, pages 137--154. Artificial Intelligence series. MIT Press, Cambridge, MA, 1993.

[COY84] M. Sharir C. O'Dunlaing and C. K. Yap. Generalised voronoi diagrams for moving a ladder. *Technical Report 32, Robotics Laboratory, Courant Institute,New York University*, 1984.

[FT87] 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.

[GI96] R. Greiner and R. Isukapalli. Learning to select useful landmarks. *IEEE Trans. Systems, Man, Cybern. Part B: Cybernetics*, 26(3):437--449, 1996.

[HA92] Yong K. Hwang and Narendra Ahuja. Gross motion planning - a survey. *ACM Computing Surveys*, 24(3):221--291, 1992.

[IK87] J. Illingworth and J. Kittler. The adaptive hough transform. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9(5):690--698, 1987.

[Kha86] O. Khatib. Real-time obstacle avoidance for manipulators and mobile robots. *Int. J. Robotics Res.*, 5(1):90--98, 1986.

[KK92] 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.

[KP92] 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.

[Lat91] J.C. Latombe. *Robot Motion Planning*. Kluver Academic Publishers, Boston, MA, 1991.

[LL94] Anthony Lazanas and Jean-Claude Latombe. Motion planning with uncertainty : A landmark approach. Technical report, Dept. of Computer Science, Stanford UIniv.,(to appear in Artificial Intellegence Journal), 1994.

[LL95] Anthony Lazanas and Jean-Claude Latombe. Landmark-based robot navigation. Technical report, Dept. of Computer Science, Stanford UIniv.,(to appear in Algorithmica), 1995.

[LPM84] T. Lozano-Perez and M.T. Mason. Automatic synthesis of fine-motion strategies for robots. *Int. J. Robotics Res.*, 3(1):3--24, 1984.

[LPW79] 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.

[MA95] Michael Magee and J.K. Aggarwalls. Robot self-location using visual reasoning relative to a single target object. *Pattern Recognition*, 28(2), 1995.

[NB88] H. Nasr and B. Bhanu. landmark recognition for autonomous mobile robots. *IEEE Intl. Conf. on Robotics and Automation*, pages 1218--1223, 1988.

[Nil69] N.J. Nilsson. A mobile automation, an application of artificial intelligence techniques. *Int. Joint Conf. Artificial Intelligence*, pages 509--520, 1969.

[NS84] N. J. Naccache and R. Shinghal. Spta : A proposed algorithm for thinning binary patterns. *IEEE Trans. Systems, Man, Cybern.*, 14(3):409--418, 1984.

[Pap91] A. Papoulis. *Probability, Random Variables and Stochastic Processes.* 3rd ed. McGraw Hill, New York, 1991.

[Par96] N. Paragyios. Adaptive detection and location of moving objects in image sequences. *Master's thesis. University of Crete, Computer Science Department*, 1996.

[PMR91] A. Mintz P. Meer and A. Rosenffeld. Robust reggression methods for comuter vision : A review. *International Journal of Computer Vision*, 6(1):59--70, 1991.

[PP96] P. Parodi and G. Piccioli. 3d shape reconstruction by using vanishing points. *IEEE Trans. on Pattern Anal. and Mach. Intell.*, 18(2):211--217, Feb 1996.

[PS85] F. P. Preparata and M. I. Shamos. Computational geometry : An introduction. *Springer-Verlag, New York, NY*, 1985.

[Rou84] P.J. Rousseuw. Least median of squares regression. *Journal of American Statistics Assocciation*, 79:871--880, 1984.

[RW87] 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.

[SA89] M.E. Spetsakis and J. Aloimonos. Optimal motion estimation. *IEEE Workshop on visual motion*, pages 229--237, 1989.

[SS83] 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.

[Sug88] K. Sugihara. Some location problems for robot navigation using a single camera. *Computer Vision, Graphics and Image Processing*, 42:112--129, Apr. 1988.

[Til90] 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.

[TK94] Camillo J. Taylor and David J. Kriegman. Vision-based motion planning and exploration algorithms for mobile robots. Technical report, Yale University, 1994.

[TK96] 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.

[TL92] 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.

[TL94] G. Tziritas and C. Labit. *Motion Analysis for Image Sequence Coding*. Elsevier, 1994.

[TY87] 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.

[Ull79] S. Ullman. The interpretation of structure from motion. *Royal Society, London*, 203:405--426, 1979.

[VLS90] S. Mukhopadhay V.J. Lumelsky and K. Sun. Dynamic path planning in sensor-based terrain acquisition. *IEEE Trans. on Robotics and Automation*, 6(4), Aug 1990.

[YK95] Erliang Yeh and David J. Kriegman. Toward selecting and recognizing natural landmarks. Technical report, Yale University, 1995.

[ZBT91] 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.

[Zhe89] J.Y. Zheng. Dynamic projection, panoramic representation of scenes for route recognition. *Ph.D.thesis in Osaka University*, Dec 1989.