

ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ  
ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ  
ΤΜΗΜΑ ΕΠΙΣΤΗΜΗΣ ΥΠΟΛΟΓΙΣΤΩΝ

ΕΞΑΓΩΓΗ ΧΑΡΑΚΤΗΡΙΣΤΙΚΩΝ  
ΤΟΥ ΤΡΙΣΔΙΑΣΤΑΤΟΥ ΧΩΡΟΥ  
ΑΠΟ ΔΙΣΔΙΑΣΤΑΤΕΣ ΕΙΚΟΝΕΣ

Χαράλαμπος Μπαλτζάκης  
Μεταπτυχιακή Εργασία

Ηράκλειο, Οκτώβριος 1999

ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ  
ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ  
ΤΜΗΜΑ ΕΠΙΣΤΗΜΗΣ ΥΠΟΛΟΓΙΣΤΩΝ

ΕΞΑΓΩΓΗ ΧΑΡΑΚΤΗΡΙΣΤΙΚΩΝ ΤΟΥ ΤΡΙΣΔΙΑΣΤΑΤΟΥ ΧΩΡΟΥ  
ΑΠΟ ΔΙΣΔΙΑΣΤΑΤΕΣ ΕΙΚΟΝΕΣ

Εργασία που υποβλήθηκε από τον Χαράλαμπο Μπαλτζάκη  
ως μερική απαίτηση για την απόκτηση  
ΜΕΤΑΠΤΥΧΙΑΚΟΥ ΔΙΠΛΩΜΑΤΟΣ ΕΙΔΙΚΕΥΣΗΣ

Συγγραφέας

Χαράλαμπος Μπαλτζάκης  
Τμήμα Επιστήμης Υπολογιστών

Επιτροπή

Πάνος Τραχανιάς, αναπληρωτής καθηγητής, επόπτης

Γεώργιος Τζιρίτας, αναπληρωτής καθηγητής, μέλος

Γεώργιος Γεωργακόπουλος, επίκουρος καθηγητής, μέλος

Δεκτή

Πάνος Κωνσταντόπουλος  
πρόεδρος Επιτροπής Μεταπτυχιακών Σπουδών

Ηράκλειο, Οκτώβριος 1999

## ΕΥΧΑΡΙΣΤΙΕΣ

Θα ήθελα να ευχαριστήσω τον επόπτη καθηγητή κ. Πάνο Τραχανιά για τις πολύτιμες συμβουλές και την αμέριστη βοήθεια που μου παρείχε καθ' όλη τη διάρκεια εκπόνησης αυτής της εργασίας.

Επίσης, ευχαριστώ τα μέλη της εξεταστικής επιτροπής κ Γεώργιο Τζιρίτα και κ. Γεώργιο Γεωργακόπουλο για τις εύστοχες υποδείξεις τους.

Η μητέρα μου Φανή και η αδελφή μου Λύδια συμπαραστάθηκαν με κάθε τρόπο καθ' όλη τη διάρκεια της εργασίας μου και τους ευχαριστώ για αυτό.

Ευχαριστώ ακόμα το τμήμα επιστήμης υπολογιστών του πανεπιστημίου Κρήτης και το Ινστιτούτο Πληροφορικής του Ιδρύματος Τεχνολογίας και έρευνας για την υλικοτεχνική υποδομή που μου παρείχε..

στη μνήμη του πατέρα μου

Γεωργίου Μπαλτζάκη

# ΕΞΑΓΩΓΗ ΧΑΡΑΚΤΗΡΙΣΤΙΚΩΝ ΤΟΥ ΤΡΙΣΔΙΑΣΤΑΤΟΥ ΧΩΡΟΥ ΑΠΟ ΔΙΣΔΙΑΣΤΑΤΕΣ ΕΙΚΟΝΕΣ

Χαράλαμπος Μπαλτζάκης  
Μεταπτυχιακή Εργασία

Τμήμα Επιστήμης Υπολογιστών  
Πανεπιστήμιο Κρήτης

## Περίληψη

Σε αυτή την εργασία ασχολούμαστε με την εξαγωγή δύο συγκεκριμένων τρισδιάστατων χαρακτηριστικών: τις τρισδιάστατες διευθύνσεις ομάδων ευθυγράμμων τμημάτων, τις οποίες εξάγουμε από τα σημεία φυγής τους, και τις διευθύνσεις και θέσεις επιπέδων στο χώρο τις οποίες εξάγουμε με τη χρήση σειρών από εικόνες στις οποίες αυτά απεικονίζονται.

Η εξαγωγή σημείων φυγής από δισδιάστατες προοπτικές προβολές είναι ένα σημαντικό πρόβλημα στη μηχανική όραση. Η ανάλυση των εξαγχθέντων σημείων φυγής μπορεί να βοηθήσει πολύ στην κατανόηση της δομής ενός τρισδιάστατου χώρου ή αντικειμένου. Γνώση ενός σημείου φυγής, σε συνδυασμό με γνώση του εστιακού σημείου δίνει αρκετή πληροφορία για την ανάκτηση τρισδιάστατης διεύθυνσης γραμμών. Σε αυτή την εργασία μελετούνται μερικοί από τους υπάρχοντες αλγορίθμους, αλλά και προτείνονται παραλλαγές τους με συγκεκριμένα πλεονεκτήματα έναντι των αρχικών. Επιχειρείται επίσης και μια νέα, εντελώς διαφορετική προσέγγιση στο θέμα, με τη χρήση συναρτήσεων αξιολόγησης ικανών να υπολογίζουν από ακατέργαστη, ή σχεδόν ακατέργαστη, πληροφορία από εικόνες την πιθανότητα κάποιο σημείο να είναι σημείο φυγής.

Εκτός από τις τρισδιάστατες κατευθύνσεις γραμμών, σημαντική επίσης, για προβλήματα όπως η πλοήγηση, η αυτο-οριοθέτηση (self-localization) και προσάραξη (docking), είναι και η γνώση των παραμέτρων των επιπέδων που μπορεί να υπάρχουν μέσα σε μία σκηνή. Σε αυτή την εργασία μελετάμε και προτείνουμε μεθόδους για τον υπολογισμό συγκεκριμένων παραμέτρων επιπέδων, από ακολουθίες εικόνων, με αντοχή σε σφάλματα. Παράλληλα με τον υπολογισμό των παραμέτρων του επιπέδου, οι αλγόριθμοι που μελετάμε εμπεριέχουν και περιορισμένη δυνατότητα εύρεσης της τρισδιάστατης κίνησης της κάμερας.

Επόπτης: Πάνος Τραχανιάς  
Αναπληρωτής Καθηγητής Επιστήμης Υπολογιστών  
Πανεπιστήμιο Κρήτης

# EXTRACTION OF 3D SPACE FEATURES FROM 2D IMAGE DATA

Haris Baltzakis  
Master of Science Thesis

Department Of Computer Science  
University of Crete

## Abstract

In this thesis we study the extraction of 3d world features from image data. More specifically, we consider the extraction of groups of parallel line segments by the computation of their vanishing points, and the extraction of orientations and positions of planes in the 3d-space.

A basic problem in computer vision is the calculation of 3D vanishing points (VP) out of 2D perspective projections. Vanishing point analysis provides strong cues for inferring the 3D structure of a scene from only a single view. Knowledge of a VP, provided that we know the focal length of the camera, provides enough information to determine the 3D orientation of the corresponding group of space lines. In this work, we study some of the existing methods and we propose modifications exhibiting certain advantages with respect to the original methods. We also propose a novel approach for VP detection, based on the definition of vanishing point likelihood functions (VPLF). These functions estimate, for each point on the image plane, the likelihood that this point is a vanishing point and are directly applicable to the raw image data.

In addition to knowledge of 3d line orientations, knowledge of the position and the orientation of planes that exist in a 3d scene is very important for many machine vision problems such as navigation, self-localization and docking. In this thesis, we study and propose methods for robustly computing the plane parameters out of given image sequences. All methods studied in this thesis also include the (limited) ability to determine unknown parameters of the 3d camera motion.

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

# ΠΕΡΙΕΧΟΜΕΝΑ

Περίληψη.....	i
Abstract.....	ii
ΠΕΡΙΕΧΟΜΕΝΑ.....	iii
ΚΑΤΑΛΟΓΟΣ ΕΙΚΟΝΩΝ.....	v
1. ΜΑΘΗΜΑΤΙΚΟ ΥΠΟΒΑΘΡΟ.....	5
1.1. Προοπτική προβολή – μοντέλο κάμερας.....	5
1.2. Σημείο φυγής – ευθεία φυγής.....	6
1.3. Ακολουθία από εικόνες – κίνηση σημείου.....	8
1.3.1. Περιστροφική κίνηση.....	8
1.3.2. Μεταφορική κίνηση.....	12
1.3.3. Φαινόμενη κίνηση σημείου.....	12
1.3.4. Ροή πάνω στην εικόνα.....	13
1.3.5. Οπτική ροή.....	14
2. ΥΠΟΛΟΓΙΣΜΟΣ ΣΗΜΕΙΟΥ ΦΥΓΗΣ.....	17
2.1. Διατύπωση του προβλήματος.....	17
2.2. Πρακτικοί αλγόριθμοι.....	17
2.2.1. Εξαγωγή σημείων φυγής με ανίχνευση γραμμών.....	18
2.2.1.1. Προεπεξεργασία.....	18
2.2.1.2. Μετασχηματισμός Hough.....	21
2.2.1.3. Ανίχνευση ευθύγραμμων τμημάτων.....	22
2.2.1.4. Εύρεση σημείου φυγής από τις κατευθύνσεις των ευθυγράμμων τμημάτων με απευθείας εύρεση τομών στο επίπεδο.....	24
2.2.1.5. Εύρεση σημείου φυγής με προβολή στον μοναδιαίο κύκλο.....	26
2.2.2. Απευθείας εύρεση σημείου φυγής.....	33
2.2.2.1. Αξιολόγηση με τη συνάρτηση VPLF.....	34
2.2.2.2. Αξιολόγηση με ολοκλήρωση τον χώρο Hough (συνάρτηση VPLFSIN).....	39
2.2.2.3. Αξιολόγηση με προβολές της βαθμίδας της φωτεινότητας (συνάρτηση VPLFGRAD).....	42
2.2.2.4. Χρήση της συνάρτησης αξιολόγησης για την εύρεση των σημείων φυγής.....	44
2.3. Συζήτηση.....	46
3. ΥΠΟΛΟΓΙΣΜΟΣ ΠΑΡΑΜΕΤΡΩΝ ΕΠΙΠΕΔΟΥ.....	51
3.1. Διατύπωση του προβλήματος.....	51
3.2. Πρακτικοί αλγόριθμοι.....	53

3.2.1. Εξίσωση επιπέδου – παράμετροι .....	53
3.2.2. Εύρεση παραμέτρων επιπέδου ώστε να συμπίπτουν δύο εικόνες.....	55
3.2.2.1. Αντιστοιχίσεις σημείων από μία κάμερα σε μία άλλη .....	55
3.2.2.2. Εύρεση παραμέτρων επιπέδου .....	58
3.2.3. Εύρεση παραμέτρων επιπέδου με υπολογισμό «ανάμεσα σε δύο επίπεδα» .....	61
3.2.4. Εύρεση με χρήση του πίνακα ροής .....	64
3.2.4.1. Εύρεση του flow matrix .....	65
3.2.4.2. Εύρεση των παραμέτρων κίνησης και των παραμέτρων επιπέδου .....	67
3.3. Συζήτηση .....	70
4. ΣΥΜΠΕΡΑΣΜΑΤΑ.....	73
BIBΛΙΟΓΡΑΦΙΑ.....	77
4.1. Βιβλία .....	77
4.2. προεπεξεργασία .....	77
4.3. μετασχηματισμός Hough .....	77
4.4. σημείο φυγής - χρήση .....	78
4.5. σημείου φυγής - υπολογισμός.....	78
4.6. οπτική ροή – υπολογισμός.....	79
4.7. πλοήγηση .....	79
4.8. εξαγωγή μερικής 3d πληροφορίας .....	79
4.9. εξαγωγή κίνησης.....	80
4.10. εξαγωγή κίνησης και 3d πληροφορίας.....	80
4.11. κίνηση ώστε να έχουμε βέλτιστη εξαγωγή πληροφορίας.....	80
4.12. μη βαθμονομημένη κάμερα .....	81
4.13. διάφορα.....	81
ΕΥΡΕΤΗΡΙΟ .....	83



# ΚΑΤΑΛΟΓΟΣ ΕΙΚΟΝΩΝ

Εικόνα 1-1. Το μοντέλο της κάμερας «μάτι της βελόνας» (pinhole camera model) .....	5
Εικόνα 1-2. Σημείο φυγής .....	7
Εικόνα 1-3. Το μοντέλο για την κίνηση της κάμερας .....	8
Εικόνα 1-4 Μια σφαίρα εντελώς λεία και χωρίς καμία υφή ( $\alpha$ ), όταν εκτελεί περιστροφική κίνηση γύρω από τον εαυτό της, παρότι προκαλεί ένα πεδίο ροής ( $\beta$ ), δεν προκαλεί καθόλου οπτική ροή ( $\gamma$ ). .....	14
Εικόνα 2-1 Μια τεχνητή εικόνα κατά τα στάδια της προεπεξεργασίας .....	20
Εικόνα 2-2 Μια εικόνα ενός διαδρόμου πριν και μετά τα στάδια της προεπεξεργασίας .....	20
Εικόνα 2-3. Μερικές απλές περιπτώσεις μετασχηματισμού Hough .....	23
Εικόνα 2-4.Ο Μετασχηματισμός Hough για την τεχνητή σκηνή του διαδρόμου στην Εικόνα 2-1. 24	
Εικόνα 2-5. Οι κυριότερες κατευθύνσεις και το σημείο φυγής για την Εικόνα 2-1 .....	25
Εικόνα 2-6. το επίπεδο που ορίζεται από μια διεύθυνση ( $\rho, \theta$ ) πάνω στην εικόνα και την αρχή των αξόνων. ....	27
Εικόνα 2-7. Ένα σημείο πάνω στη τομή του επιπέδου με κάθετο διάνυσμα $\hat{V}^{mn}$ και στη μοναδιαία σφαίρα με κέντρο την αρχή των αξόνων. ....	28
Εικόνα 2-8. Η μοναδιαία σφαίρα όπως έχει διαμορφωθεί μετά την προβολή της πληροφορίας από τον πίνακα Hough για τη σκηνή στη Εικόνα 2-1 .....	30
Εικόνα 2-9. Η μοναδιαία σφαίρα όπως έχει διαμορφωθεί μετά την προβολή της πληροφορίας από τον πίνακα Hough για τη σκηνή στην Εικόνα 2-2 .....	30
Εικόνα 2-10. Η βαθμίδα της φωτεινότητας για τη σκηνή στην Εικόνα 2-1 .....	32
Εικόνα 2-11. Η μοναδιαία σφαίρα όπως έχει διαμορφωθεί μετά την προβολή της πληροφορίας από τη βαθμίδα της φωτεινότητας για την Εικόνα 2-1 και την Εικόνα 2-2 .....	33
Εικόνα 2-12. Η συνάρτηση συνεισφοράς $F_v(e_i)$ .....	34
Εικόνα 2-13. Μια απλή εικόνα ενός κύβου και οι τιμές της συνάρτησης VPLF πάνω στη μοναδιαία σφαίρα .....	36
Εικόνα 2-14. Οι κατανομές στα σημεία A, B, C, D στην .....	37
Εικόνα 2-15. Οι τιμές της VPLF για την Εικόνα 2-13 .....	37
Εικόνα 2-16. Αξιολόγηση σημείων φυγής με τη συνάρτηση VPLF για την Εικόνα 2-1 και την Εικόνα 2-2. Η μέθοδος βρήκε επιτυχώς τα κυρίαρχα σημεία φυγής της εικόνας (σημεία A, A', B, B', C) .....	38
Εικόνα 2-17. Μια αεροφωτογραφία. Αξιολόγηση σημείων φυγής με τη συνάρτηση VPLF. Οι τιμές της VPLF έχουν προβληθεί στη μοναδιαία σφαίρα. Η μέθοδος βρήκε επιτυχώς τα κυρίαρχα σημεία φυγής της εικόνας (σημεία A, A', B, B') .....	38

Εικόνα 2-18. Ευθύγραμμα τμήματα που όταν προεκταθούν τέμνονται σε ένα κοινό σημείο μετασχηματίζονται σε σημείο στο ιστόγραμμα Hough που κείτονται πάνων σε μία ημιτονοειδή γραμμή .....	40
Εικόνα 2-19. Η μοναδιαία σφαίρα όπως έχει διαμορφωθεί μετά την προβολή των τιμών της συνάρτησης VPLFSIN για την Εικόνα 2-1 και την Εικόνα 2-2. ....	41
Εικόνα 2-20. Η μοναδιαία σφαίρα όπως έχει διαμορφωθεί μετά την προβολή των τιμών της συνάρτησης VPLFGRAD για την Εικόνα 2-1 και την Εικόνα 2-2.....	43
Εικόνα 2-21. Οι τιμές της VPLF κατά τη διαδικασία διαίρεσης των κελιών για την Εικόνα 2-13. ....	45
Εικόνα 2-22. Γραφική παράσταση του σφάλματος και των τιμών $\phi-\theta$ του σημείου φυγής στο κέντρο του διαδρόμου στην Εικόνα 2-1 κατά τη διάρκεια σύγκλισης του αλγορίθμου simulated annealing. Χρήση της συνάρτησης αξιολόγησης VPLFGRAD.....	46
Εικόνα 3-1. Το επίπεδο που εκφράζεται από την εξίσωση $aX+bY+cZ+d=0$ αν $a^2+\beta^2+c^2=1$ και $d>0$ .....	54
Εικόνα 3-2. Ένα επίπεδο όπως φαίνεται από έξι διαφορετικές θέσεις.....	57
Εικόνα 3-3. Εφαρμογή του αλγορίθμου υπολογισμού μιας εικόνας από μία άλλη. ....	58
Εικόνα 3-4. Το σφάλμα κατά τη διάρκεια της σύγκλισης του αλγορίθμου. ....	60
Εικόνα 3-5. Οι παράμετροι $a, b, c, d$ του επιπέδου κατά τη διάρκεια της σύγκλισης.....	60
Εικόνα 3-6. Η μεταφορική κίνηση της κάμερας από τη θέση $\alpha$ στη θέση $\beta$ κατά τη διάρκεια της σύγκλισης.....	61
Εικόνα 3-7. Το ζητούμενο επίπεδο βρίσκεται ανάμεσα σε δύο παράλληλα με αυτό επίπεδα που απέχουν μεταξύ τους κατά $dd$ . ....	63
Εικόνα 3-8. Οι παράμετροι $a,b,c,d$ του επιπέδου και η απόσταση $dd$ κατά τη διάρκεια της σύγκλισης του αλγορίθμου. ....	64
Εικόνα 3-9. Η μεταφορική κίνηση της κάμερας από τη θέση $\alpha$ στη θέση $\beta$ κατά τη διάρκεια της σύγκλισης.....	64

## ΕΙΣΑΓΩΓΗ

Όταν σημεία ή γεωμετρικά χαρακτηριστικά από τον τρισδιάστατο κόσμο προβάλλονται στον αμφιβληστροειδή χιτώνα του ματιού μας ή στο φακό μίας κάμερας, πολύ μεγάλο μέρος της πληροφορίας για τη δομή και τη θέση των χαρακτηριστικών αυτών χάνεται. Συστηματική μελέτη των προβολών σε συνάρτηση με γνώση εκ των προτέρων για τη δομή του τρισδιάστατου χώρου και του μηχανισμού προβολής μπορούν να οδηγήσουν σε ανάκτηση μεγάλου μέρους από τη χαμένη πληροφορία.

Ο αμφιβληστροειδής χιτώνας του ματιού σε συνδυασμό με τον ανθρώπινο εγκέφαλο αποτελεί το πιο εξελιγμένο, μέχρι σήμερα γνωστό σύστημα όρασης. Συνάμα αποτελεί και μια μεγάλη πρόκληση για μεγάλη μερίδα επιστημόνων από όλο τον κόσμο, που προσπαθεί να κατασκευάσει μηχανές ικανές να αντιλαμβάνονται το ίδιο καλά τον χώρο μέσα στον οποίο βρίσκονται.

Η ενασχόληση αυτή του ανθρώπου με τη δημιουργία τέτοιων μηχανών δημιούργησε ένα καινούργιο κλάδο της επιστήμης υπολογιστών, τη μηχανική όραση. Η εργασία αυτή εντάσσεται σε αυτόν τον κλάδο.

Το πρόβλημα της μηχανικής όρασης είναι αποδεδειγμένα ένα δύσκολο πρόβλημα. Η αντιμετώπιση του σαν ολότητα, η πλήρης ανακατασκευή δηλαδή του τρισδιάστατου χώρου είναι πρακτικά ανέφικτη προς το παρόν και, όταν επιχειρείται [37,38,39], γίνεται συνήθως χωρίς ικανοποιητική ακρίβεια και κάτω από πολλές παραδοχές. Μια τακτική, που εφαρμόζεται και στη μηχανική όραση, είναι ο επιμερισμός του προβλήματος και η ενασχόληση με τη μελέτη και την κατασκευή αξιόπιστων αλγορίθμων που να επιλύουν μικρά αλλά σημαντικά σε πρακτικές εφαρμογές προβλήματα. Τέτοιες εφαρμογές περιλαμβάνουν την εξαγωγή της κίνησης μιας κάμερας στο χώρο [34,35,36], την καθοδήγηση ενός ρομποτικού συστήματος να κάνει απλές κινήσεις στο χώρο [25,26,27,28,40], τη βαθμονόμηση μίας κάμερας [10,14] ή την εξαγωγή συγκεκριμένων τρισδιάστατων χαρακτηριστικών [11,12,13,29,30,31,32,33] όπως γωνίες, κατευθύνσεις, απλά σχήματα κ.λ.π.

Σε αυτή την εργασία ασχολούμαστε με την εξαγωγή δύο συγκεκριμένων τρισδιάστατων χαρακτηριστικών: τις τρισδιάστατες διευθύνσεις ομάδων ευθυγράμμων τμημάτων, τις οποίες εξάγουμε από τα σημεία φυγής τους, και τις διευθύνσεις και θέσεις επιπέδων στο χώρο τις οποίες εξάγουμε με τη χρήση σειρών από εικόνες στις οποίες αυτά απεικονίζονται.

Η εξαγωγή σημείων φυγής από διςδιάστατες προοπτικές προβολές είναι ένα

σημαντικό πρόβλημα στη μηχανική όραση. Η ανάλυση των εξαγχθέντων σημείων φυγής μπορεί να βοηθήσει πολύ στην κατανόηση της δομής ενός τρισδιάστατου χώρου ή αντικειμένου. Γνώση ενός σημείου φυγής, σε συνδυασμό με γνώση του εστιακού σημείου δίνει αρκετή πληροφορία για την ανάκτηση τρισδιάστατης διεύθυνσης γραμμών. Ακόμα και όταν δεν γνωρίζουμε το εστιακό μήκος, μπορούμε να χρησιμοποιήσουμε το σημείο φυγής για να τμηματοποιήσουμε τις ακμές μιας εικόνας σε τμήματα με ίδια κατεύθυνση (π.χ. όταν θέλουμε να βρούμε όλα τα κατακόρυφα τμήματα σε μια εικόνα). Πρακτικοί αλγόριθμοι που έχουν προταθεί κατά καιρούς για τον υπολογισμό σημείων φυγής στηρίζονται στη γνώση των ευθυγράμμων τμημάτων που απαρτίζουν μια εικόνα. Σε αυτή την εργασία μελετάμε αυτούς τους αλγορίθμους αλλά και προτείνουμε κάποιες διαφορετικές προσεγγίσεις. Οι προσεγγίσεις αυτές παρουσιάζουν ένα μεγάλο πλεονέκτημα σε σχέση με τις υπάρχοντες μεθόδους μιας και δεν προϋποθέτουν γνώση των ευθυγράμμων τμημάτων.

Εκτός από τις τρισδιάστατες κατευθύνσεις γραμμών, σημαντική επίσης, για προβλήματα πλοήγησης, είναι και η γνώση των παραμέτρων των επιπέδων που μπορεί να υπάρχουν μέσα σε μία εικόνα. Οι παράμετροι ενός επιπέδου μπορούν να εξαχθούν εύκολα αν είναι γνωστά τρία σημεία που βρίσκονται πάνω σε αυτό. Σε αυτή την εργασία μελετάμε και προτείνουμε μεθόδους για υπολογισμό παραμέτρων επιπέδων με αντοχή σε σφάλματα. Παράλληλα με τον υπολογισμό των παραμέτρων του επιπέδου, αν δε διαθέτουμε πλήρη γνώση εκ των προτέρων σχετικά με την κίνηση της κάμερας, οι αλγόριθμοι που μελετούνται εμπεριέχουν, μέχρι κάποιο βαθμό, την ικανότητα να εξάγουν και πληροφορία για τη τρισδιάστατη κίνηση της κάμερας

Στο κεφάλαιο 2 δίνεται το πολύ βασικό θεωρητικό υπόβαθρο, απαραίτητο για την κατανόηση των αλγορίθμων που παρουσιάζονται στα επόμενα δύο κεφάλαια. Αρχικά, δίνονται στοιχεία για το προοπτικό προβολικό μοντέλο στο οποίο στηρίζεται αυτή η εργασία. Ακολουθεί ο ορισμός και οι βασικές σχέσεις για τα σημεία φυγής, και εξηγείται γιατί τα σημεία φυγής είναι κομβικά στοιχεία στην κατανόηση εικόνων που έχουν προκύψει σύμφωνα με το προβολικό μοντέλο που επιλέξαμε. Για τις περιπτώσεις που θα χρειαστεί να θεωρήσουμε παρακάτω ένα μοντέλο κινούμενης κάμερας, ή ένα μοντέλο από δύο κάμερες (stereo) που «βλέπουν» ταυτόχρονα τον ίδιο τρισδιάστατο κόσμο, δίνονται οι βασικές σχέσεις στις οποίες υπόκεινται σημεία της εικόνας καθώς μεταβαίνουμε από μια εικόνα ενός κόσμου σε μια άλλη του ίδιου κόσμου.

Στο κεφάλαιο 3 ασχολούμαστε με τον υπολογισμό σημείων φυγής. Ο υπολογισμός

σημείων φυγής έχει κατά καιρούς απασχολήσει πολλούς επιστήμονες οι οποίοι έχουν προτείνει μερικούς πολύ έξυπνους αλγόριθμους για αυτό το σκοπό [15,16,17,18,19,20]. Ο λόγος είναι ότι η γνώση των σημείων φυγής μας δίνει ένα αρκετά καλό όπλο για την αντιμετώπιση πολλών πρακτικών προβλημάτων όπως είναι η βαθμονόμηση μιας κάμερας [10,14], η ανίχνευση ευθυγράμμων τμημάτων σε συγκεκριμένες τρισδιάστατες διευθύνσεις [12] ή η προσαρμογή των εικόνων σε τρισδιάστατα μοντέλα που είναι γνωστά από πριν [11,13]. Εδώ εξετάζονται μερικοί από τους υπάρχοντες αλγόριθμους, αλλά και προτείνονται παραλλαγές τους με συγκεκριμένα πλεονεκτήματα έναντι των αρχικών. Επιχειρείται επίσης και μια εντελώς διαφορετική προσέγγιση στο θέμα με τη χρήση κάποιων συναρτήσεων αξιολόγησης ικανών να υπολογίζουν από ακατέργαστη, ή σχεδόν ακατέργαστη, πληροφορία από εικόνες την πιθανότητα κάποιο σημείο να είναι σημείο φυγής [21].

Στο κεφάλαιο 4 επιχειρούμε μια προσέγγιση στο πρόβλημα της εύρεσης της τρισδιάστατης θέσης και διεύθυνσης επιπέδων που απεικονίζονται σε εικόνες. Οι προσεγγίσεις που έχουν προταθεί κατά καιρούς είναι πάρα πολλές. Συνήθως για την αντιμετώπιση του προβλήματος γίνεται χρήση δύο τουλάχιστον εικόνων από διαφορετική θέση. Σε αυτή την περίπτωση είναι συνήθως απαραίτητη η επίλυση του προβλήματος της αντιστοίχισης σημείων (correspondence problem) [22,23,24], η εύρεση, δηλαδή, της θέσης της προβολής ενός τρισδιάστατου σημείου σε μια εικόνα όταν είναι γνωστή η θέση της προβολής του σε μια άλλη εικόνα. Επικεντρώνουμε το ενδιαφέρον μας σε δύο διαφορετικές προσεγγίσεις: α) στην επαναληπτική εύρεση των παραμέτρων του επιπέδου ώστε να έχουμε βέλτιστες αντιστοιχίσεις ανάμεσα σε εικόνες τις οποίες υπολογίζουμε αναλυτικά, και β) στην επαναληπτική εύρεση των παραμέτρων του επιπέδου ώστε αυτό να ταιριάζει όσο το δυνατόν περισσότερο σε τρισδιάστατα σημεία, τα οποία υπολογίζονται με βάση γνωστές αντιστοιχίσεις από εικόνα σε εικόνα. Παραθέτουμε επίσης, για λόγους σύγκρισης των αποτελεσμάτων, και μια, γνωστή από τη βιβλιογραφία μέθοδο, τη μέθοδο με τον πίνακα ροής [3].

Τέλος, στο κεφάλαιο 5 παρατίθενται επιγραμματικά τα συμπεράσματα που αποκοιμήθηκαν από την ενασχόληση με το αντικείμενο αυτής της εργασίας.

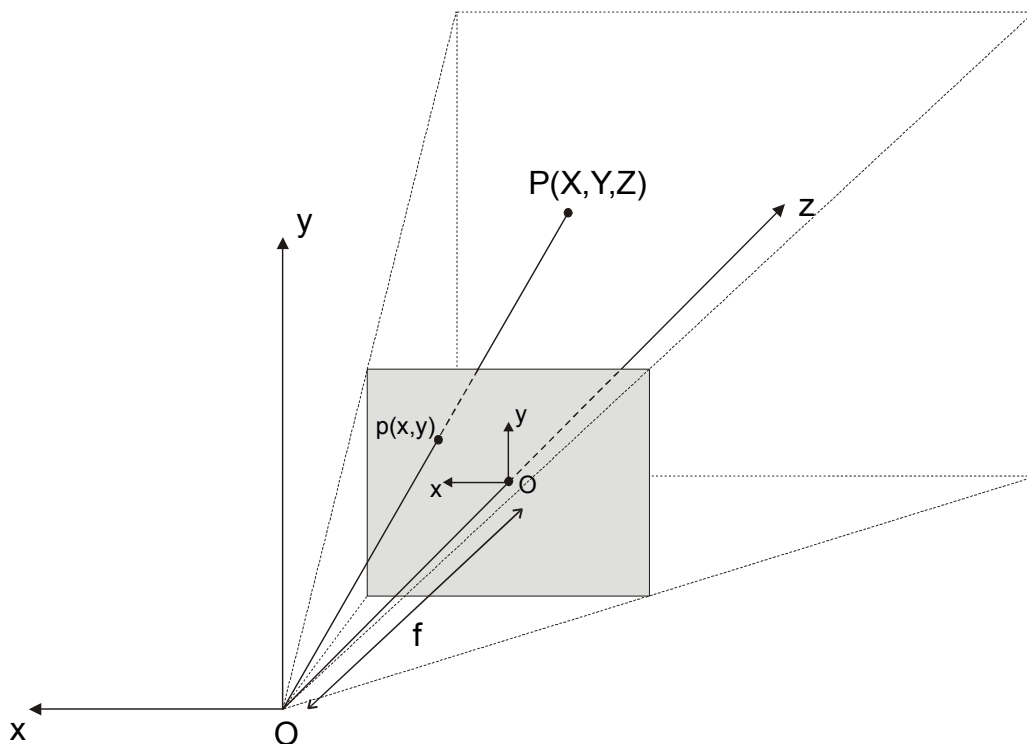


# 1. ΜΑΘΗΜΑΤΙΚΟ ΥΠΟΒΑΘΡΟ

## 1.1. Προοπτική προβολή – μοντέλο κάμερας

Οι καλλιτέχνες της αναγέννησης, από τον 15ο κιάλας αιώνα, στην προσπάθειά τους να φτιάξουν εικόνες-ζωγραφιές που να δείχνουν όσο το δυνατόν πιο αληθοφανείς (να δίνουν όσο ο δυνατόν πιο ρεαλιστική συναίσθηση του βάθους), ανακάλυψαν και ερεύνησαν αυτό που λέμε προοπτική. Τα περισσότερα προβλήματα της προοπτικής λύθηκαν κατά τη διάρκεια του 19ου αιώνα όταν ανατύχθηκε το απαιτούμενο θεωρητικό υπόβαθρο.

Ο μετασχηματισμός, σύμφωνα με τους κανόνες της προοπτικής, των σημείων του τρισδιάστατου χώρου σε σημεία πάνω στο δισδιάστατο επίπεδο προβολής (καμβάς του καλλιτέχνη, αμφιβληστροειδής χιτώνας του ματιού), ονομάζεται προοπτικός μετασχηματισμός.



Εικόνα 1-1. Το μοντέλο της κάμερας «μάτι της βελόνας» (pinhole camera model)

Σήμερα οι επιστήμονες της μηχανικής όρασης χρησιμοποιούν το προοπτικό προβολικό μοντέλο ως ένα απλό και συνάμα κοντινό στην πραγματικότητα μοντέλο για

προσομοίωση της διαδικασίας προβολής τρισδιάστατων αντικειμένων στο φακό της κάμερας. Το μοντέλο αυτό της κάμερας, το οποίο συχνά ονομάζεται «μάτι της βελόνας» (pinhole), φαίνεται στην Εικόνα 1-1.

Η αρχή των αξόνων  $O$ , ορίζεται στη θέση της οπτικής εστίας, της εστίας, δηλαδή, της κεντρικής προβολής που υλοποιεί τον προοπτικό μετασχηματισμό. Ο άξονας  $Z$  ονομάζεται οπτικός άξονας (optical axis). Το επίπεδο προβολής είναι κάθετο στον οπτικό άξονα και βρίσκεται σε συγκεκριμένη απόσταση  $f$  από την αρχή των αξόνων. Η απόσταση αυτή ονομάζεται εστιακό μήκος (focal length).

Η εικόνα που αποτυπώνεται στη κάμερα βρίσκεται στο επίπεδο προβολής με το κέντρο της πάνω στο σημείο τομής με τον οπτικό άξονα. Έχει δε, συγκεκριμένες διαστάσεις οι οποίες σε συνάρτηση με το εστιακό μήκος ορίζουν έναν πυραμιδοειδή όγκο ο οποίος ονομάζεται όγκος όρασης (view volume). Οτιδήποτε βρίσκεται έξω από όγκο όρασης δεν αποτυπώνεται στην κάμερα.

Κάθε σημείο  $P(X,Y,Z)$  του τρισδιάστατου χώρου είναι συγραμμικό με την απεικόνισή του  $p(x,y)$  στο επίπεδο προβολής και το εστιακό σημείο. Έτσι ισχύουν οι σχέσεις:

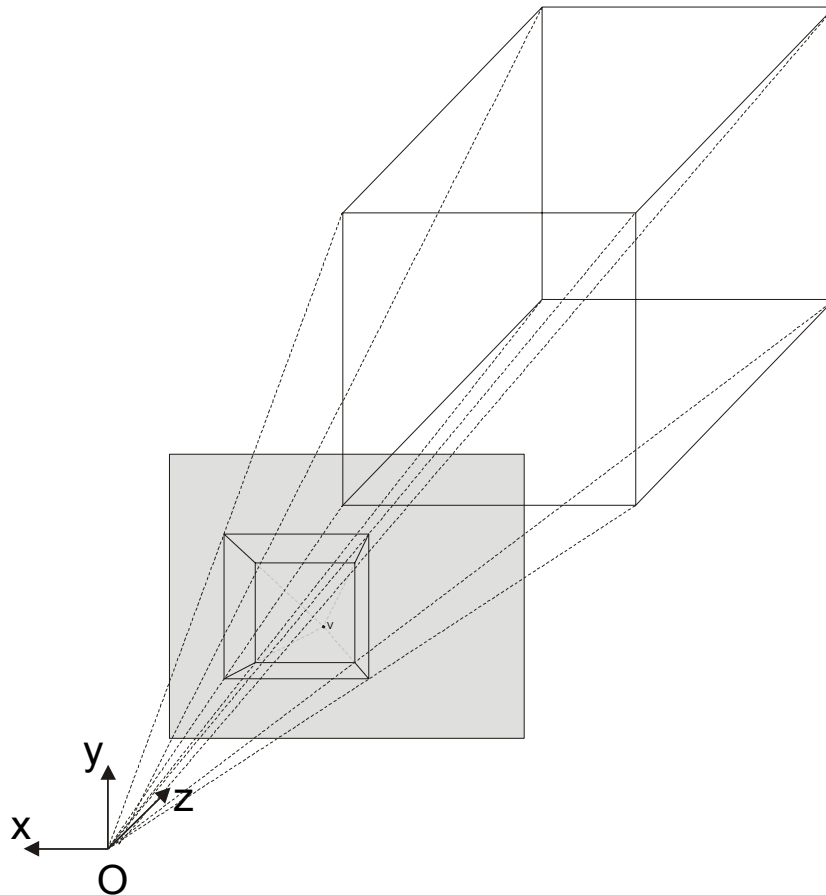
$$\frac{X}{Z} = \frac{x}{f} \quad \text{Σχέση 1-1}$$

$$\frac{Y}{Z} = \frac{y}{f} \quad \text{Σχέση 1-2}$$

## 1.2. Σημείο φυγής – ευθεία φυγής

Παράλληλες γραμμές, μετά από τον προοπτικό μετασχηματισμό, δεν παραμένουν παράλληλες αλλά τέμνονται σε ένα σημείο που ονομάζεται σημείο φυγής (vanishing point). Στην Εικόνα 1-2 φαίνεται ένα από τα τρία σημεία φυγής ενός κύβου. Τα άλλα δύο σημεία φυγής, λόγω του ότι αντιστοιχούν σε γραμμές παράλληλες με το επίπεδο προβολής βρίσκονται στο άπειρο. Δύο ή περισσότερα σημεία φυγής από γραμμές που βρίσκονται σε παράλληλα επίπεδα είναι συνευθειακά και ορίζουν μια ευθεία γνωστή ως ευθεία φυγής (vanishing line).





Εικόνα 1-2. Σημείο φυγής

Το γεγονός αυτό, ή άρση της παραλληλίας των γραμμών δηλαδή, που δεν συμβαδίζει με τους κανόνες την παραδοσιακής ευκλείδειας γεωμετρίας, συνιστά το κύριο χαρακτηριστικό του προοπτικού μετασχηματισμού και αποτελεί τη βάση για ένα καινούργιο κλάδο της γεωμετρίας, την προβολική γεωμετρία.

Έστω ότι έχουμε μια δέσμη από παράλληλες ευθείες. Αν  $v(v_x, v_y)$  είναι το σημείο φυγής που αντιστοιχεί στη δέσμη και  $f$  το εστιακό μήκος, τότε το διάνυσμα κατεύθυνσης της δέσμης θα είναι όσο το διάνυσμα κατεύθυνσης οποιουδήποτε σημείου πάνω στην ευθεία που ανήκει στη δέσμη και περνάει από την αρχή των αξόνων. Δηλαδή:

$$p = \text{Κατεύθυνση} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \text{Κατεύθυνση} \begin{pmatrix} X/Z \\ Y/Z \\ 1 \end{pmatrix} = \text{Κατεύθυνση} \begin{pmatrix} v_x / f \\ v_y / f \\ 1 \end{pmatrix} \quad \text{Σχέση 1-3}$$

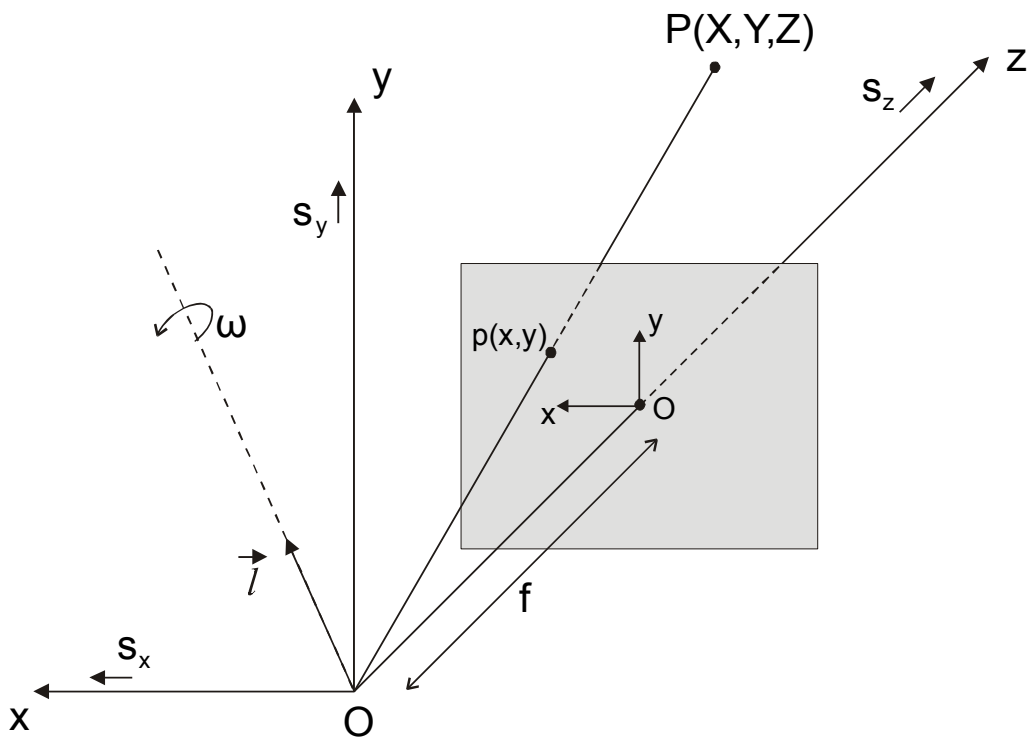
Δηλαδή, γνώση ενός σημείου φυγής, σε συνδυασμό με γνώση του εστιακού μήκους μετατρέπεται εύκολα σε γνώση για την τρισδιάστατη κατεύθυνση των ευθειών που

αντιστοιχούν στο σημείο φυγής. Αντίστοιχα, γνώση μίας ευθείας φυγής σε συνδυασμό με γνώση του εστιακού μήκους μετατρέπεται εύκολα σε γνώση για την τρισδιάστατη κατεύθυνση του παράλληλων επιπέδων που αντιστοιχούν στην ευθεία φυγής.

### 1.3. Ακολουθία από εικόνες – κίνηση σημείου

Σε μερικές περιπτώσεις είναι διαθέσιμη μια ολόκληρη ακολουθία από εικόνες (image sequence). Οι εικόνες αυτές προκύπτουν από την κίνηση της κάμερας μέσα στο χώρο και απεικονίζουν την ίδια τρισδιάστατη σκηνή από διαφορετικές θέσεις.

Οποιαδήποτε κίνηση και να κάνει η κάμερα μέσα στο χώρο, αυτή μπορεί να περιγραφεί πάντα σαν άθροισμα δύο κινήσεων: μιας μεταφορικής κίνησης σε σχέση με το στατικό περιβάλλον της και μιας περιστροφικής κίνησης γύρω από έναν άξονα που περνάει από την οπτική της εστία.



Εικόνα 1-3. Το μοντέλο για την κίνηση της κάμερας

#### 1.3.1. Περιστροφική κίνηση

Έστω ότι έχουμε μία κάμερα που περιστρέφεται γύρω από ένα άξονα  $\vec{l} = [l_x, l_y, l_z]^T$  που περνάει από το εστιακό της σημείο όπως φαίνεται στην Εικόνα 1-3. Οι άξονες του

συστήματος της κάμερας, τη χρονική στιγμή  $t_0$ , ορίζονται από τα μοναδιαία διανύσματα  $\bar{x}^{t_0}$ ,  $\bar{y}^{t_0}$ ,  $\bar{z}^{t_0}$ :

$$\bar{x}^{t_0} = \begin{bmatrix} x_x^{t_0} \\ x_y^{t_0} \\ x_z^{t_0} \end{bmatrix} \quad \bar{y}^{t_0} = \begin{bmatrix} y_x^{t_0} \\ y_y^{t_0} \\ y_z^{t_0} \end{bmatrix} \quad \bar{z}^{t_0} = \begin{bmatrix} z_x^{t_0} \\ z_y^{t_0} \\ z_z^{t_0} \end{bmatrix} \quad \text{Σχέση 1-4}$$

Έστω ότι μετά από παρέλευση χρόνου  $\Delta t$ , τη χρονική στιγμή  $t_1=t_0+\Delta t$ , λόγω της περιστροφικής της κίνησης, το σύστημα αξόνων της κάμερας θα έχει περιστραφεί γύρω από τον άξονα  $\vec{l} = [l_x, l_y, l_z]^T$  κατά γωνία  $\Omega$ . Η καινούργια βάση για το σύστημα αξόνων της κάμερας, όπως αυτή «φαίνεται» από το παλιό σύστημα αξόνων, είναι τα μοναδιαία διανύσματα  $\bar{x}^{t_1}$ ,  $\bar{y}^{t_1}$ ,  $\bar{z}^{t_1}$ .

$$\bar{x}^{t_1} = \begin{bmatrix} x_x^{t_1} \\ x_y^{t_1} \\ x_z^{t_1} \end{bmatrix} \quad \bar{y}^{t_1} = \begin{bmatrix} y_x^{t_1} \\ y_y^{t_1} \\ y_z^{t_1} \end{bmatrix} \quad \bar{z}^{t_1} = \begin{bmatrix} z_x^{t_1} \\ z_y^{t_1} \\ z_z^{t_1} \end{bmatrix} \quad \text{Σχέση 1-5}$$

Ο γραμμικός μετασχηματισμός που αντιστοιχεί συντεταγμένες από το σύστημα συντεταγμένων που ορίζουν τα μοναδιαία διανύσματα  $\bar{x}^{t_0}$ ,  $\bar{y}^{t_0}$ ,  $\bar{z}^{t_0}$  στο σύστημα συντεταγμένων που ορίζουν τα μοναδιαία διανύσματα  $\bar{x}^{t_1}$ ,  $\bar{y}^{t_1}$ ,  $\bar{z}^{t_1}$  θα είναι ένας  $3 \times 3$  πίνακας  $R^{t_0 \rightarrow t_1}$  που ονομάζεται πίνακας περιστροφής (rotation matrix) και περιέχει σαν στήλες τα μοναδιαία διανύσματα των αξόνων της κάμερας τη χρονική στιγμή  $t_1$  όπως αυτά εκφράζονται στη βάση  $\{\bar{x}^{t_0}, \bar{y}^{t_0}, \bar{z}^{t_0}\}$ .

$$R^{t_0 \rightarrow t_1} = \left[ \bar{x}^{t_1} \mid \bar{y}^{t_1} \mid \bar{z}^{t_1} \right] = \begin{bmatrix} x_x^{t_1} & y_x^{t_1} & z_x^{t_1} \\ x_y^{t_1} & y_y^{t_1} & z_y^{t_1} \\ x_z^{t_1} & y_z^{t_1} & z_z^{t_1} \end{bmatrix} \quad \text{Σχέση 1-6}$$

Αν P είναι ένα σημείο στο χώρο και οι συντεταγμένες του όπως αυτές φαίνονται από την κάμερα τις χρονικές στιγμές  $t_0$  και  $t_1$  είναι  $(X^{t_0}, Y^{t_0}, Z^{t_0})$  και  $(X^{t_1}, Y^{t_1}, Z^{t_1})$  αντίστοιχα, θα ισχύει:

$$\begin{bmatrix} X^{t_1} \\ Y^{t_1} \\ Z^{t_1} \end{bmatrix} = R^{t_0 \rightarrow t_1} \begin{bmatrix} X^{t_0} \\ Y^{t_0} \\ Z^{t_0} \end{bmatrix} \quad \text{Σχέση 1-7}$$

Επειδή το σύστημα αξόνων της κάμερας είναι ορθογώνιο, δηλαδή το εσωτερικό γινόμενο ανά δύο των μοναδιαίων της διανυσμάτων είναι μηδέν, θα ισχύει:

$$\left(R^{t_0 \rightarrow t_1}\right)^T \cdot R^{t_0 \rightarrow t_1} = I \quad \text{Σχέση 1-8}$$

όπου I ο μοναδιαίος πίνακας. Δηλαδή ο  $R^{t_0 \rightarrow t_1}$  είναι ορθογώνιος πίνακας και ισχύει:

$$\left(R^{t_0 \rightarrow t_1}\right)^{-1} = \left(R^{t_0 \rightarrow t_1}\right)^T \quad \text{Σχέση 1-9}$$

Συνεπώς, η Σχέση 1-7 μπορεί να γραφτεί και με τη μορφή:

$$\begin{bmatrix} X^{t_0} \\ Y^{t_0} \\ Z^{t_0} \end{bmatrix} = \left(R^{t_0 \rightarrow t_1}\right)^T \begin{bmatrix} X^{t_1} \\ Y^{t_1} \\ Z^{t_1} \end{bmatrix} \quad \text{Σχέση 1-10}$$

Για τις συντεταγμένες  $(X^{t_0}, Y^{t_0}, Z^{t_0})$  και  $(X^{t_1}, Y^{t_1}, Z^{t_1})$  θα ισχύει, όπως αποδεικνύεται με απλή γεωμετρία, η φόρμουλα Rodrigues [3]:

$$\begin{bmatrix} X^{t_1} \\ Y^{t_1} \\ Z^{t_1} \end{bmatrix} = \begin{bmatrix} X^{t_0} \\ Y^{t_0} \\ Z^{t_0} \end{bmatrix} \cos \Omega + \hat{l} \times \begin{bmatrix} X^{t_0} \\ Y^{t_0} \\ Z^{t_0} \end{bmatrix} \sin \Omega + (1 - \cos \Omega) (\hat{l} \cdot \begin{bmatrix} X^{t_0} \\ Y^{t_0} \\ Z^{t_0} \end{bmatrix}) \hat{l} \quad \text{Σχέση 1-11}$$

Γράφοντας τη παραπάνω σχέση σε μορφή πινάκων και αντιστοιχίζοντας με τη Σχέση 1-7 έχουμε για τον  $R^{t_0 \rightarrow t_1}$ :

$$R^{t_1 \rightarrow 0} = \begin{bmatrix} \cos \Omega + l_x^2 (1 - \cos \Omega) & l_x l_y (1 - \cos \Omega) - l_z \sin \Omega & l_x l_z (1 - \cos \Omega) - l_y \sin \Omega \\ l_y l_x (1 - \cos \Omega) + l_z \sin \Omega & \cos \Omega + l_y^2 (1 - \cos \Omega) & l_y l_z (1 - \cos \Omega) - l_x \sin \Omega \\ l_z l_x (1 - \cos \Omega) - l_y \sin \Omega & l_z l_y (1 - \cos \Omega) - l_x \sin \Omega & \cos \Omega + l_z^2 (1 - \cos \Omega) \end{bmatrix}$$

Σχέση 1-12

Δηλαδή, και για τα μοναδιαία διανύσματα  $\bar{x}^{t_1}$ ,  $\bar{y}^{t_1}$ ,  $\bar{z}^{t_1}$  ισχύει:

$$\bar{x}^{t_1} = \begin{bmatrix} \cos \Omega + l_x^2 (1 - \cos \Omega) \\ l_y l_x (1 - \cos \Omega) + l_z \sin \Omega \\ l_z l_x (1 - \cos \Omega) - l_y \sin \Omega \end{bmatrix} \quad \text{Σχέση 1-13}$$

$$\bar{y}^{t_1} = \begin{bmatrix} l_x l_y (1 - \cos \Omega) - l_z \sin \Omega \\ \cos \Omega + l_y^2 (1 - \cos \Omega) \\ l_z l_y (1 - \cos \Omega) - l_x \sin \Omega \end{bmatrix} \quad \text{Σχέση 1-14}$$

$$\bar{z}^{t_1} = \begin{bmatrix} l_x l_z (1 - \cos \Omega) - l_y \sin \Omega \\ l_y l_z (1 - \cos \Omega) - l_x \sin \Omega \\ \cos \Omega + l_z^2 (1 - \cos \Omega) \end{bmatrix} \quad \text{Σχέση 1-15}$$

Μια προϋπόθεση που τίθεται συχνά είναι η θέση και η διεύθυνση της κάμερας από εικόνα σε εικόνα να αλλάζει πολύ λίγο (μεγάλο sampling rate ή μικρή ταχύτητα κίνησης). Σε αυτή την περίπτωση, αν δηλαδή η γωνία  $\Omega$  έχει πολύ μικρή τιμή, έστω  $\Delta\Omega$ , ο πίνακας  $R^{t_0 \rightarrow t_1}$  τείνει να γίνει ίσος με τον μοναδιαίο πίνακα. Σε αυτή τη περίπτωση, αποδεικνύεται [3] ότι η Σχέση 1-11 παίρνει τη μορφή:

$$R^{t_0 \rightarrow t_1} = I + W \Delta\Omega + O(\Delta\Omega^2) \text{ με } W = \begin{bmatrix} 0 & -l_z & l_y \\ l_z & 0 & -l_x \\ -l_y & l_x & 0 \end{bmatrix} \quad \text{Σχέση 1-16}$$

Αν η κάμερα κινείται με σταθερή γωνιακή ταχύτητα  $\omega = \Delta\omega / \Delta t$  γύρω από τον άξονα  $\vec{l}$  τότε ορίζεται η περιστροφική ταχύτητα  $\vec{\omega}$  ως:

$$\vec{\omega} = \lim_{\Delta t \rightarrow 0} \frac{\Delta\Omega}{\Delta t} \vec{l} = \lim_{\Delta t \rightarrow 0} \omega \vec{l} \quad \text{Σχέση 1-17}$$

Σε αυτή τη περίπτωση, όπως βγαίνει από τη Σχέση 1-16, αν αντικαταστήσουμε όπου  $\Delta\Omega$  τη γωνία  $\omega\Delta t$  και πάρουμε το όριο  $\Delta t \rightarrow 0$ , ένα σημείο με σταθερό ως προς το σύστημα της κάμερας διάνυσμα θέσης  $\vec{F}$  θα κινείται λόγω της σταθερής περιστροφικής ταχύτητας  $\vec{\omega}$  με ταχύτητα  $\vec{q}$  που δίνεται από τη σχέση:

$$\vec{q} = \vec{\omega} \times \vec{F} \quad \text{Σχέση 1-18}$$

### 1.3.2. Μεταφορική κίνηση

Έστω ότι, εκτός από τη περιστροφική κίνηση που περιγράψαμε παραπάνω, η κάμερα κινείται και μεταφορικά, και το εστιακό της σημείο αλλάζει από τη χρονική στιγμή  $t_0$  στη χρονική στιγμή  $t_1$  κατά  $\overset{P}{h} = [h_x, h_y, h_z]^T$ , με το  $\overset{P}{h}$  εκφρασμένο στο σύστημα συντεταγμένων της κάμερας τη χρονική στιγμή  $t_0$ .

Αν  $O^{t_0} = (O_x^{t_0}, O_y^{t_0}, O_z^{t_0})$  είναι το εστιακό σημείο της κάμερας τη χρονική στιγμή  $t_0$  και  $O^{t_1} = (O_x^{t_1}, O_y^{t_1}, O_z^{t_1})$  είναι το εστιακό σημείο της κάμερας τη χρονική στιγμή  $t_1$ , τότε:

$$\begin{bmatrix} O_x^{t_1} \\ O_y^{t_1} \\ O_z^{t_1} \end{bmatrix} = \begin{bmatrix} h_x \\ h_y \\ h_z \end{bmatrix} \quad \text{Σχέση 1-19}$$

Αν το χρονικό διάστημα  $\Delta t$  είναι αρκετά μικρό τότε ορίζεται η στιγμιαία μεταφορική ταχύτητα ως:

$$\overset{P}{s} = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \overset{P}{h} \quad \text{Σχέση 1-20}$$

### 1.3.3. Φαινόμενη κίνηση σημείου

Έστω ότι μια κάμερα κάνει ταυτόχρονα μεταφορική κίνηση με ταχύτητα  $\overset{P}{s} = [s_x, s_y, s_z]$  και περιστροφική κίνηση με ταχύτητα  $\overset{P}{\omega} = \omega [l_x, l_y, l_z]^T$ . Ένα σημείο με σταθερό ως προς το σύστημα της κάμερας διάνυσμα θέσης  $\overset{F}{r}$  θα κινείται στο χώρο με ταχύτητα που θα προκύπτει ως το άθροισμα της μεταφορικής ταχύτητας  $\overset{P}{s}$  και της ταχύτητας  $\overset{w}{q}$  λόγω της περιστροφικής κίνησης  $\overset{P}{\omega}$ . Η κατάσταση αυτή είναι ισοδύναμη με την κατάσταση που θα προέκυπτε αν η κάμερα ήταν σταθερή και η τρισδιάστατη σκηνή κινιόταν με μεταφορική και περιστροφική ταχύτητα  $-\overset{P}{s}$  και  $-\overset{P}{\omega}$ , αντίστοιχα.

Έτσι, από το τοπικό σύστημα της κάμερας, κάθε σταθερό σημείο (X, Y, Z) του τρισδιάστατου χώρου θα φαίνεται να κινείται με ταχύτητα:

$$\overset{P}{V} = -\overset{P}{s} - \overset{P}{q} \quad \text{Σχέση 1-21}$$

ή, αλλιώς, αντικαθιστώντας από τη Σχέση 1-18:

$$\dot{V} = -\dot{s} - \dot{\omega} \times \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad \text{Σχέση 1-22}$$

#### 1.3.4. Ροή πάνω στην εικόνα

Η κίνηση της κάμερας θα έχει ως αποτέλεσμα η απεικόνιση κάθε σημείου του τρισδιάστατου χώρου στο επίπεδο προβολής να κινείται με ταχύτητα  $(u_r, v_r)$  ως προς το σύστημα συντεταγμένων  $Oxy$  του επιπέδου προβολής. Δηλαδή θα είναι:

$$u_r = \frac{dx}{dt} \quad \text{Σχέση 1-23}$$

$$v_r = \frac{dy}{dt} \quad \text{Σχέση 1-24}$$

Ουσιαστικά, το διάνυσμα  $(u_r, v_r)$  θα είναι η προβολή του διανύσματος κίνησης  $\dot{V}$  πάνω στην εικόνα

Αντικαθιστώντας από τη Σχέση 1-1 και τη Σχέση 1-2 έχουμε:

$$u_r = f \frac{d\left(\frac{X}{Z}\right)}{dt} \quad \text{Σχέση 1-25}$$

$$v_r = f \frac{d\left(\frac{Y}{Z}\right)}{dt} \quad \text{Σχέση 1-26}$$

Χρησιμοποιώντας τη Σχέση 1-22 καταλήγουμε τελικά:

$$u_r = \frac{-s_x f + x s_z}{Z} + \omega_x \frac{xy}{f} - \omega_y \left(\frac{x^2}{f} + f\right) + \omega_z y \quad \text{Σχέση 1-27}$$

$$v_r = \frac{-s_y f + y s_z}{Z} + \omega_x \left(\frac{y^2}{f} + f\right) - \omega_y \frac{xy}{f} - \omega_z x \quad \text{Σχέση 1-28}$$

Αν η κίνηση της κάμερας είναι μόνο μεταφορική ( $\omega=0$ ), τότε οι παραπάνω σχέσεις παίρνουν την απλοποιημένη μορφή:

$$u_r = \frac{-s_x f + x s_z}{Z}$$

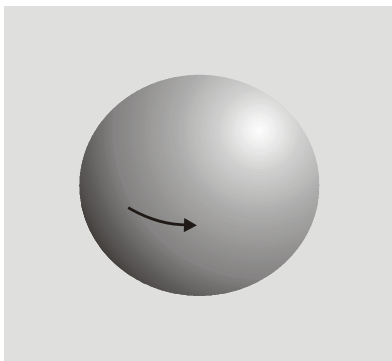
Σχέση 1-29

$$v_r = \frac{-s_y f + y s_z}{Z}$$

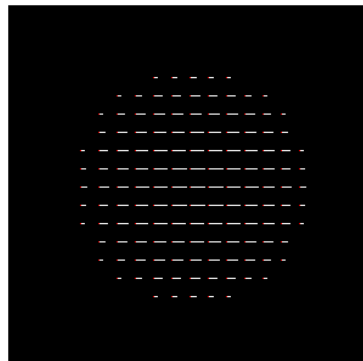
Σχέση 1-30

### 1.3.5. Οπτική ροή

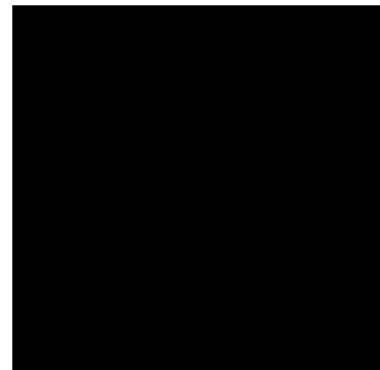
Οι εξισώσεις που παρουσιάστηκαν στη προηγούμενη παράγραφο συσχετίζουν τη τρισδιάστατη κίνηση ενός σημείου με την προβολή της κίνησής του πάνω στο επίπεδο προβολής. Η προβολή αυτή της κίνησης είναι ένα καθαρά θεωρητικό μέγεθος μιας και δεν μπορεί να μετρηθεί. Στη θέση της χρησιμοποιείται συνήθως αυτό που ονομάζουμε οπτική ροή. Η οπτική ροή είναι η φαινομενική ταχύτητα της κίνησης των φωτεινών σχηματισμών πάνω στην εικόνα (αυτό δηλαδή που βλέπουμε). Η οπτική ροή δεν συμπίπτει πάντα με την πραγματική δισδιάστατη ροή. Για παράδειγμα, η απεικόνιση μιας σφαίρας όπως αυτή στην Εικόνα 1-4, δεν αλλάζει καθόλου καθώς αυτή περιστρέφεται γύρω από τον εαυτό της. Το αποτέλεσμα είναι το πεδίο της οπτικής ροής να είναι παντού μηδενικό.



α. Αρχική εικόνα



β. Πραγματική ροή



γ. Οπτική ροή

Εικόνα 1-4 Μια σφαίρα εντελώς λεία και χωρίς καμία υφή (α), όταν εκτελεί περιστροφική κίνηση γύρω από τον εαυτό της, παρότι προκαλεί ένα πεδίο ροής (β), δεν προκαλεί καθόλου οπτική ροή (γ).



Έστω ότι  $(x,y)$  είναι ένα σημείο σε μία εικόνα που αντιστοιχεί στη χρονική στιγμή  $t$  και έστω ότι  $E(x,y,t)$  είναι η φωτεινότητά του. Έστω ότι το σημείο  $(x,y)$ , στην αμέσως επόμενη εικόνα, τη χρονική στιγμή  $t+\delta t$ , έχει μετακινηθεί κατά  $(\delta x, \delta y)$ . Η θέση του θα είναι  $(x+\delta x, y+\delta y)$  και η φωτεινότητά του  $E(x+\delta x, y+\delta y, t+\delta t)$ . Αν  $(u,v)$  είναι η οπτική ροή στο σημείο  $(x,y)$  τη χρονική στιγμή  $t$  τότε θα ισχύει για τη μεταβολή του σημείου  $(x,y)$  τη χρονική στιγμή  $t$ :

$$(\delta x, \delta y) = (u\delta t, v\delta t) \quad \text{Σχέση 1-31}$$

Θεωρώντας ότι η φωτεινότητα στο σημείο  $(x,y)$  τη χρονική στιγμή  $t$  θα είναι ίδια με τη φωτεινότητά στο σημείο  $(x+\delta x, y+\delta y)$  τη χρονική στιγμή  $t+\delta t$ , καταλήγουμε στη σχέση:

$$E(x + u\delta t, y + v\delta t, t + \delta t) = E(x, y, t) \quad \text{Σχέση 1-32}$$

Θεωρώντας ότι η φωτεινότητα μεταβάλλεται ομαλά κοντά στο σημείο  $(x,y)$  αναπτύσσουμε τη παραπάνω σχέση κατά Taylor και έχουμε:

$$E(x, y, t) + \delta x \frac{\partial E}{\partial x} + \delta y \frac{\partial E}{\partial y} + \delta t \frac{\partial E}{\partial t} + e = E(x, y, t) \quad \text{Σχέση 1-33}$$

όπου ο όρος  $e$  περιλαμβάνει τις παραγώγους από δεύτερης τάξης και πάνω. Θεωρώντας επιπλέον ότι η οπτική ροή  $u,v$  είναι συνεχής, μπορούμε να πάρουμε το όριο της παραπάνω σχέσης για  $\delta t \rightarrow 0$  και έχουμε:

$$\frac{\partial E}{\partial x} \frac{dx}{dt} + \frac{\partial E}{\partial y} \frac{dy}{dt} + \frac{\partial E}{\partial t} = 0 \quad \text{Σχέση 1-34}$$

Η Σχέση 1-34 στην οποία καταλήξαμε είναι γνωστή ως σχέση περιορισμού της οπτικής ροής (optical flow constraint equation). Η σχέση αυτή, σε συνδυασμό με άλλες παραδοχές, όπως η συνέχεια του πεδίου της οπτικής ροής, μπορεί να χρησιμοποιηθεί για τον υπολογισμό της οπτικής ροής.



## 2. ΥΠΟΛΟΓΙΣΜΟΣ ΣΗΜΕΙΟΥ ΦΥΓΗΣ

### 2.1. Διατύπωση του προβλήματος

Ένα από τα βασικά προβλήματα στην μηχανική όραση είναι η εξαγωγή σημείων φυγής από δισδιάστατες προοπτικές προβολές. Η ανάλυση των εξαγχθέντων σημείων φυγής μπορεί να βοηθήσει πολύ στην κατανόηση της δομής ενός τρισδιάστατου χώρου ή αντικειμένου.

Γνώση ενός σημείου φυγής, σε συνδυασμό με γνώση του εστιακού σημείου δίνει αρκετή πληροφορία για την ανάκτηση τρισδιάστατης διεύθυνσης γραμμών (Σχέση 1-3). Αντίστοιχα, γνώση μίας ευθείας φυγής σε συνδυασμό με γνώση του εστιακού μήκους μετατρέπεται εύκολα σε γνώση για την τρισδιάστατη διεύθυνση των παράλληλων επιπέδων που αντιστοιχούν στην ευθεία φυγής.

Ακόμα και όταν δεν γνωρίζουμε το εστιακό μήκος, μπορούμε να χρησιμοποιήσουμε το σημείο φυγής για να τμηματοποιήσουμε τις ακμές μιας εικόνας σε τμήματα με ίδια κατεύθυνση., όπως για παράδειγμα όταν θέλουμε να βρούμε όλα τα κατακόρυφα τμήματα σε μια εικόνα.

Οι McGlone et al. [12] χρησιμοποίησαν γνώση σημείων φυγής για να ανιχνεύσουν κτήρια και ανθρώπινες κατασκευές σε εξωτερικούς χώρους. Οι Daniilidis et al. [10] και οι Carpile, et al. [14] χρησιμοποίησαν τα σημεία φυγής για να επιτύχουν βαθμονόμηση (calibration) του συστήματος της κάμερας. Οι Parodi et al. [11] χρησιμοποίησαν σημεία φυγής για να εξάγουν τρισδιάστατη πληροφορία για τον χώρο. Οι Collins et al. [13] χρησιμοποίησαν ευθείες φυγής που υπολόγισαν από σημεία φυγής για να ταιριάξουν εικόνες επίπεδων αντικειμένων, είτε μεταξύ τους, είτε με γνωστά εκ των προτέρων μοντέλα.

### 2.2. Πρακτικοί αλγόριθμοι

Κατά τη διάρκεια των τελευταίων ετών πολλοί επιστήμονες έχουν ασχοληθεί με την έρευνα για την ανάπτυξη αξιόπιστων αλγορίθμων για την εύρεση των σημείων φυγής. Σε αυτές τις προσεγγίσεις υπάρχει εγγενής η παραδοχή ότι «ισχυρά σημεία τομής ευθειών»

αντιστοιχούν σε σημεία φυγής.

Πρακτικοί αλγόριθμοι που έχουν προταθεί κατά καιρούς αποτελούνται από τρία διαφορετικά στάδια: (α) εξαγωγή ευθυγράμμων τμημάτων, (β) τμηματοποίηση αυτών σε ομάδες με κοινό σημείο φυγής και (γ) τελική εκτίμηση των σημείων φυγής.

Προσεγγίσεις με χρήση του μετασχηματισμού Hough όπως του Kender [18], του Barnard [15] και των Quan et al. [20] προτείνουν την αθροιστική προβολή των ευθυγράμμων τμημάτων μιας εικόνας σε διακριτά τμηματοποιημένους χώρους, και την ανίχνευση των σημείων φυγής σαν κορυφές στο τελικό ιστόγραμμα.

Οι Magee et al. [19] πρότειναν την ανίχνευση των σημείων φυγής απευθείας από τα εξαχθέντα ευθύγραμμα τμήματα χρησιμοποιώντας υπολογισμούς με εξωτερικά γινόμενα. Ο Collins et al. [16] χρησιμοποίησε συνδυασμό των μεθόδων που περιγράφηκαν από τους Barnard [15] και Magee et al [19] και παρουσίασε μια τεχνική που συνδυάζει την υψηλή ταχύτητα υπολογισμών του Barnard με την ακρίβεια των αποτελεσμάτων των Magee et al. Οι Collins et al. [17] βελτίωσαν τις τεχνικές με ιστόγραμμα προτείνοντας μια μέθοδο βασισμένη σε στατιστική ανάλυση για την ανίχνευση των σημείων φυγής από τα ιστογράμματα.

Στα πλαίσια αυτής της εργασίας μελετήθηκαν κάποιες μέθοδοι βασισμένοι στον μετασχηματισμό Hough για την ανίχνευση των ευθυγράμμων τμημάτων. Επίσης, ερευνήθηκε το ενδεχόμενο ύπαρξης μιας αξιόπιστης τεχνικής που να μην βασίζεται στην εξαγωγή ευθυγράμμων τμημάτων και παρουσιάζουμε τρεις διαφορετικές, νέες προσεγγίσεις στο θέμα αυτό.

## 2.2.1. Εξαγωγή σημείων φυγής με ανίχνευση γραμμών

### 2.2.1.1. Προεπεξεργασία

Το πρώτο στάδιο σε αυτή τη διεργασία είναι η εξαγωγή των ακμών (edge detection). Στην εργασία αυτή χρησιμοποιήθηκαν δύο διαγώνιες μάσκες Roberts της μορφής.

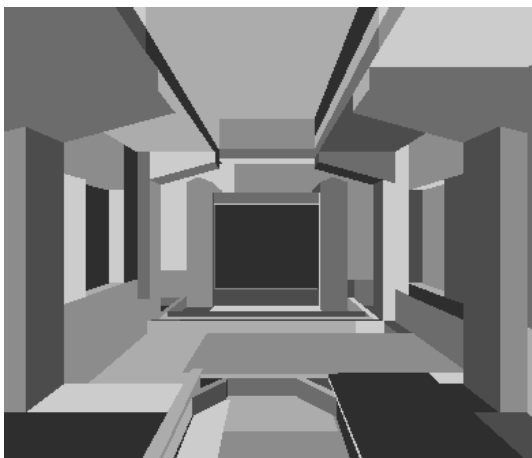
-1	0	0	0	0	-1
0	1	0	0	0	0
0	0	0	0	0	0

Οι μάσκες αυτής της μορφής δεν αποκρίνονται καλά στις περιπτώσεις που οι εικόνες εμπεριέχουν μεγάλα ποσοστά θορύβου. Παρόλα αυτά έδωσαν αρκετά καλά αποτελέσματα

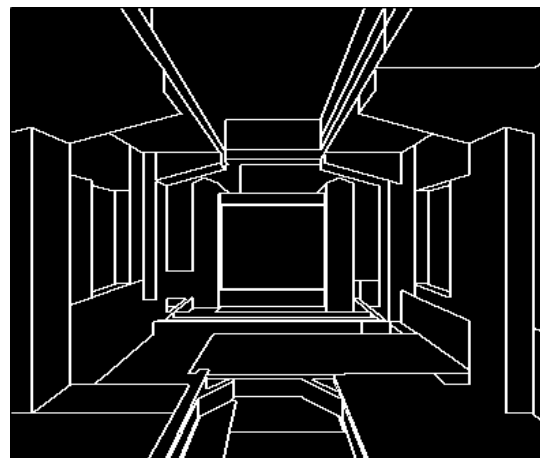
για τις εικόνες που χρησιμοποιήσαμε στην εργασίας μας μιας και αυτές ήταν αρκετά καλής ποιότητας.

Για τη βελτίωση των αποτελεσμάτων χρησιμοποιήθηκε στις τελικές ακμές και ένας αλγόριθμος σκελετοποίησης. Ο αλγόριθμος που χρησιμοποιήθηκε είναι μια παραλλαγή του αλγορίθμου που παρουσιάστηκε από τους Chu et al. [6] και βελτιώθηκε από τους Lam et al [7]. Επιγραμματικά ο αλγόριθμος σκελετοποίησης που χρησιμοποιήθηκε αποτελείται από τα εξής βήματα:

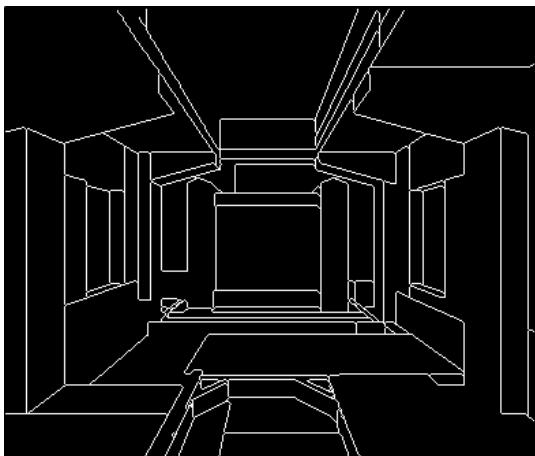
- i. Ανίχνευση συνοριακών γραμμών (contours)
- ii. Χωρισμός των σημείων που απαρτίζουν τις συνοριακές γραμμές σε trace points (ένας τουλάχιστον 4-γείτονας είναι λευκός) και corner points (ακριβώς ένας 8-γείτονας είναι λευκός)
- iii. Εξέταση με τη σειρά των σημείων κάθε συνοριακής γραμμής ξεχωριστά και απαλοιφή των trace-points που δεν είναι ούτε break points (ακριβώς δύο 8-γείτονες μαύροι που ο ένας δεν ανήκει στην 8-γειτονιά του άλλου) ούτε end-points (ακριβώς ένας 8-γείτονας μαύρος)
- iv. Εξέταση με τη σειρά των σημείων κάθε συνοριακής γραμμής ξεχωριστά και απαλοιφή των corner-points που η απαλοιφή τους δημιουργεί κάποιο end-point η απαλοιφή τους σύμφωνα με κάποιο κριτήριο (βασισμένο σε 16 διαφορετικές μάσκες-πρωτότυπα [7]), κάνει πιο ομαλή (λεία) τη συνοριακή γραμμή.
- v. Εφόσον έγινε έστω και μία απαλοιφή σημείου στα προηγούμενα δύο βήματα επανάληψη της διαδικασίας (επιστροφή στο πρώτο βήμα).



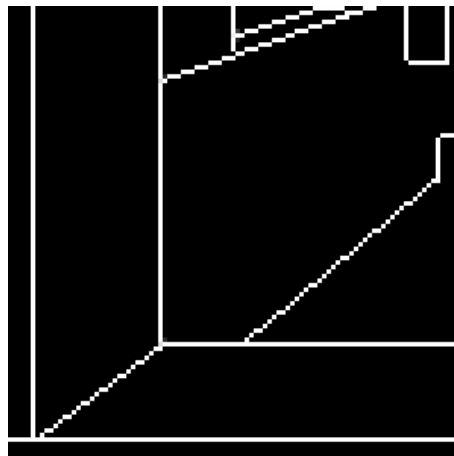
*α. Αρχική εικόνα*



*β. Εικόνα μετά την εξαγωγή ακμών*



*γ. Εικόνα μετά τη διαδικασία της σκελετοποίησης*

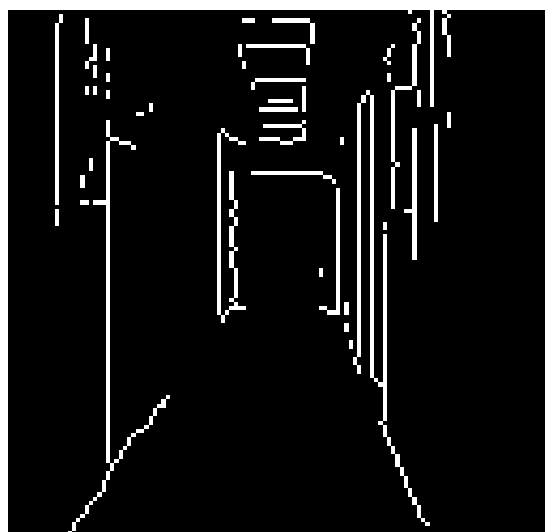


*δ. Λεπτομέρεια από το κάτω αριστερό άκρο της σκελετοποιημένης εικόνας*

*Εικόνα 2-1 Μια τεχνητή εικόνα κατά τα στάδια της προεπεξεργασίας.*



*α. Αρχική εικόνα.*



*β. Μετά τα στάδια της προεπεξεργασίας.*

*Εικόνα 2-2 Μια εικόνα ενός διαδρόμου πριν και μετά τα στάδια της προεπεξεργασίας.*

Στην Εικόνα 2-1 απεικονίζονται οι φάσεις της προεπεξεργασίας στις οποία υπόκειται μια τεχνητή εικόνα ενός διαδρόμου. Ομοίως, στην Εικόνα 2-2 βλέπουμε το αποτέλεσμα της προεπεξεργασίας σε μια πραγματική εικόνα ενός διαδρόμου του κεντρικού κτηρίου του Ιδρύματος Έρευνας και Τεχνολογίας.

### 2.2.1.2. Μετασχηματισμός Hough

Το επόμενο στάδιο είναι ο μετασχηματισμός Hough για γραμμές. Ο μετασχηματισμός Hough, στην απλή του μορφή, επιτρέπει τον μετασχηματισμό μίας ευθείας του επιπέδου των καρτεσιανών συντεταγμένων σε ένα σημείο του επιπέδου των πολικών συντεταγμένων. Μία ευθεία στο επίπεδο των καρτεσιανών συντεταγμένων μπορεί να οριστεί με τη σχέση

$$\rho = x \cos \theta + y \sin \theta \quad \text{Σχέση 2-1}$$

όπου  $\rho$  η απόσταση από την αρχή των αξόνων και  $\theta$  η γωνία που σχηματίζει με τον άξονα  $x$ . Αν γνωρίζουμε τα  $\rho$ ,  $\theta$  τότε μπορούμε να ορίζουμε πλήρως την ευθεία.

Εξ ορισμού, ο μετασχηματισμός Hough της ευθείας που δίνεται από τη Σχέση 2-1 είναι το σημείο  $(\rho, \theta)$  του επιπέδου των πολικών συντεταγμένων. Αντίστροφα, ένα σημείο  $(x_0, y_0)$  του επιπέδου των καρτεσιανών συντεταγμένων μετασχηματίζεται στην καμπύλη:

$$\rho = x_0 \cos \theta + y_0 \sin \theta \quad \text{Σχέση 2-2}$$

στο επίπεδο των πολικών συντεταγμένων. Αν θέσουμε:

$$\tan \varphi = \frac{x_0}{y_0} \quad \text{Σχέση 2-3}$$

η Σχέση 2-2 μετασχηματίζεται σε:

$$\rho \frac{\cos \varphi}{y_0} = \sin \varphi \cos \theta + \cos \varphi \sin \theta \quad \text{Σχέση 2-4}$$

ή αλλιώς:

$$\rho = a \sin(\theta + \phi) \quad \text{Σχέση 2-5}$$

με:

$$a = \frac{y_0}{\cos \varphi} \quad \text{Σχέση 2-6}$$

Δηλαδή ο μετασχηματισμός Hough ενός σημείου είναι μια ημιτονοειδής καμπύλη στο επίπεδο των πολικών συντεταγμένων.

Οι Duda και Hart προσάρμοσαν το μετασχηματισμό Hough και έδωσαν μια μέθοδο υλοποίησής του για δυαδικές εικόνες. Σύμφωνα με τη μεθοδό τους, κάθε μαύρο pixel της εικόνας (στην οποία έχει γίνει εξαγωγή ακμών και σκελετοποίηση) μετασχηματίζεται σε μία ημιτονοειδή καμπύλη στο επίπεδο  $\rho$ - $\theta$ . Το επίπεδο αυτό υλοποιείται με ένα ιστόγραμμα διαστάσεων  $M \times N$  που μπορούμε να τον ονομάσουμε ιστόγραμμα Hough. Έτσι, κάθε σημείο της ημιτονοειδούς καμπύλης αυξάνει κατά ένα την τιμή στην κοντινότερη προς αυτό θέση του ιστογράμματος Hough. Στο τέλος αυτής της διαδικασίας το ιστόγραμμα θα αντιπροσωπεύει το μετασχηματισμό Hough της δυαδικής εικόνας. Κάθε γραμμή στην αρχική εικόνα θα αντιστοιχίζεται σε μία κορυφή στον μετασχηματισμό Hough.

Στην Εικόνα 2-3 φαίνονται μερικές απλές περιπτώσεις μετασχηματισμού Hough.

### 2.2.1.3. Ανίχνευση ευθύγραμμων τμημάτων

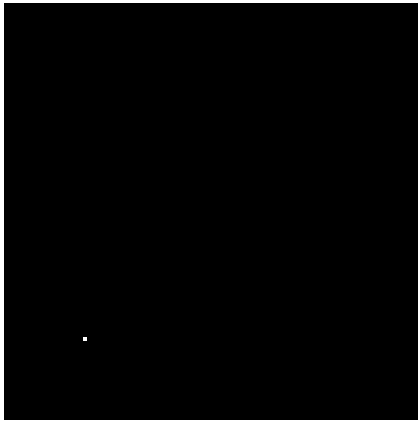
Για να ανιχνεύσουμε τα ευθύγραμμα τμήματα στο ιστόγραμμα Hough ελέγχουμε για καταχωρίσεις που να ικανοποιούν τα παρακάτω δύο κριτήρια

- η τιμή τους να υπερβαίνει ένα προκαθορισμένο κατώφλι.
- να είναι κορυφές (να μην υπάρχει κάποιο άλλο σημείο στη γειτονιά τους που να έχει μεγαλύτερη τιμή).

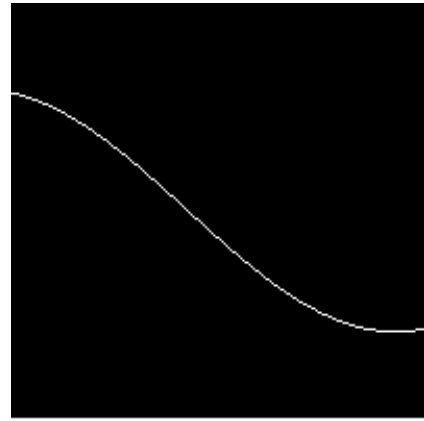
Κάθε σημείο που ικανοποιεί τα παραπάνω κριτήρια αξιολογείται με βάση τη μέση τιμή που έχει ο μετασχηματισμός Hough σε μια πολύ μικρή γειτονιά του και του προσάπτεται ένα βάρος (αυτό για να αποφύγουμε σφάλματα κβαντισμού – πολλές φορές μια τιμή που θα έπρεπε να έχει κάποιο κελί του ιστογράμματος Hough είναι μοιρασμένο σε μερικά γειτονικά κελιά). Αν κάποιες από τις κατευθύνσεις βρίσκονται πολύ κοντά (έχουν απόσταση μικρότερη από ένα προκαθορισμένο κατώφλι) τότε αυτές ενώνονται σε μία με κέντρο το πιο ισχυρό σημείο και βάρος το άθροισμα των βαρών.

Με βάση την παραπάνω αξιολόγηση οι κατευθύνσεις ταξινομούνται κατά φθίνουσα σειρά και επιλέγεται ένας προκαθορισμένος αριθμός από τις πιο σημαντικές. Στην Εικόνα 2-4 φαίνεται ο μετασχηματισμός Hough για την τεχνητή σκηνή του διαδρόμου στην Εικόνα 2-1 καθώς και οι 20 πιο σημαντικές κατευθύνσεις.



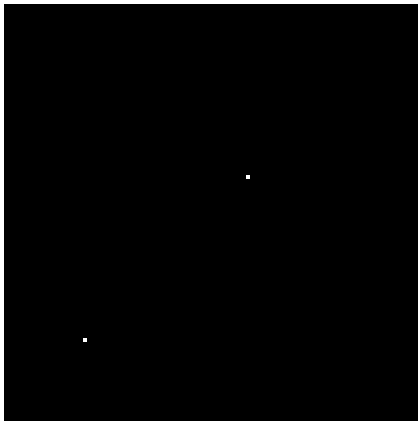


Αρχική εικόνα

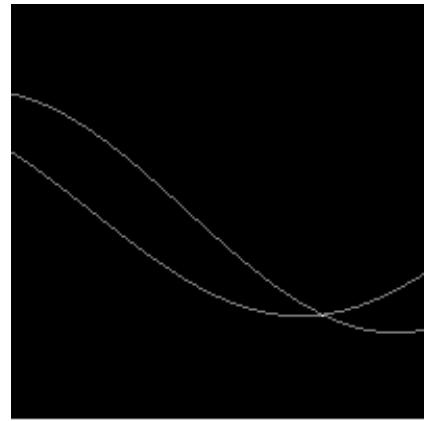


Μετασχηματισμός Hough

α. Ένα σημείο στο επίπεδο των καρτεσιανών συντεταγμένων μετασχηματίζεται σε μια ημιτονοειδή καμπύλη στο επίπεδο των πολικών συντεταγμένων.

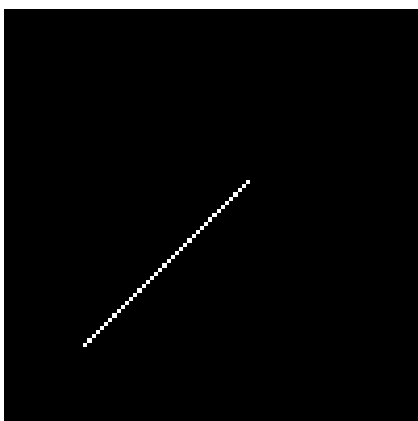


Αρχική εικόνα

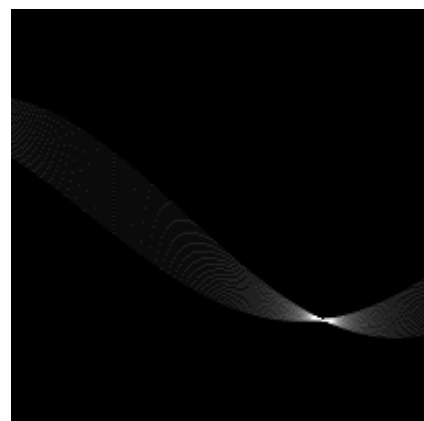


Μετασχηματισμός Hough

β. Δύο σημεία στο επίπεδο των καρτεσιανών συντεταγμένων μετασχηματίζονται σε δύο ημιτονοειδείς καμπύλες στο επίπεδο των πολικών συντεταγμένων. Το σημείο που τέμνονται θα είναι το σημείο στο οποίο θα μετασχηματιζόταν η ευθεία που ορίζεται από τα δύο σημεία.



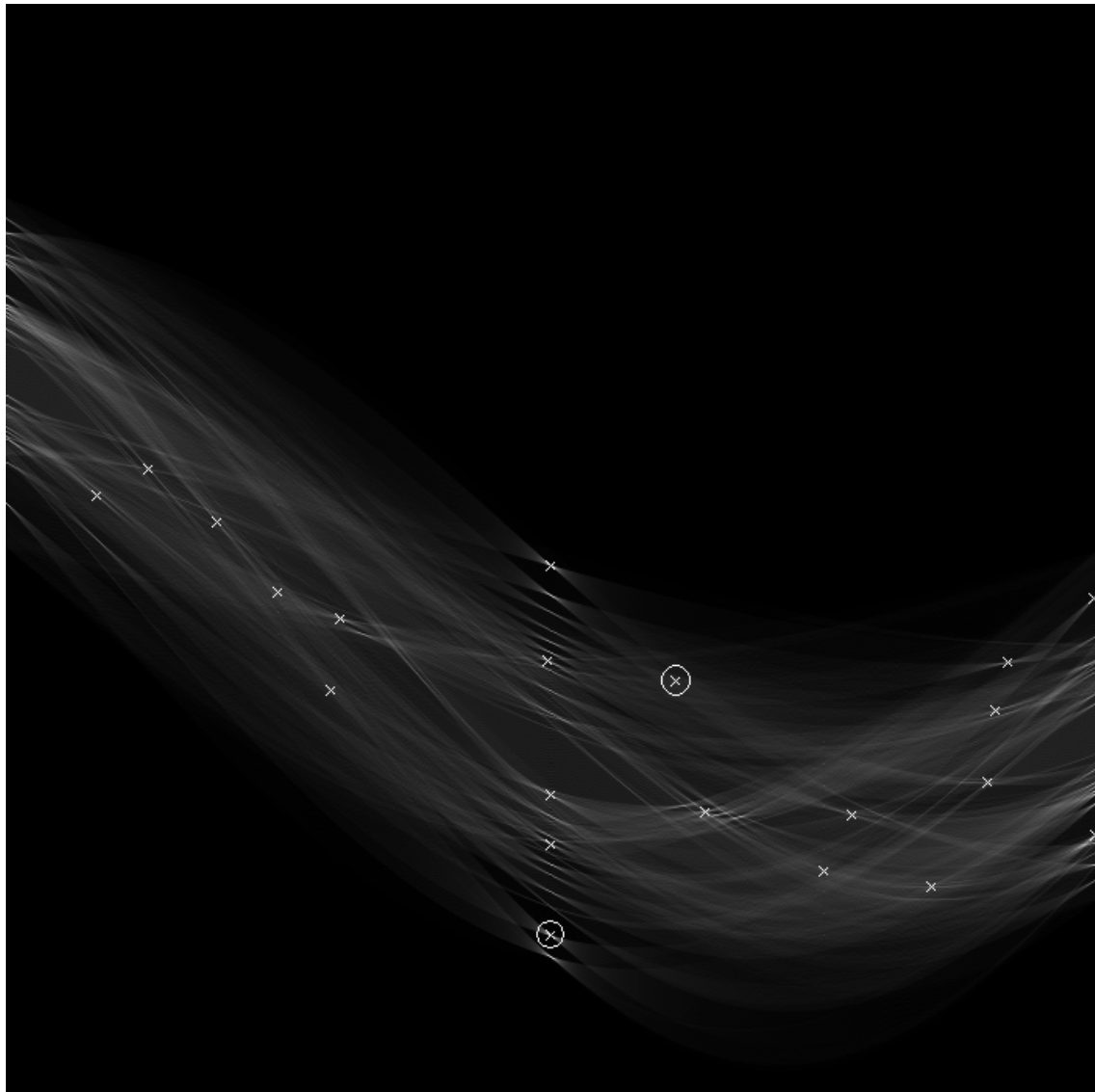
Αρχική εικόνα



Μετασχηματισμός Hough

γ. Μία ευθεία στο επίπεδο των καρτεσιανών συντεταγμένων μετασχηματίζεται σε ένα σημείο στο επίπεδο των πολικών συντεταγμένων.

Εικόνα 2-3. Μερικές απλές περιπτώσεις μετασχηματισμού Hough



*Εικόνα 2-4. Ο Μετασχηματισμός Hough για την τεχνητή σκηνή του διαδρόμου στην Εικόνα 2-1.*

*Οι 20 κυριότερες ευρεθείσες διευθύνσεις είναι σημειωμένες με «x» οι κατευθύνσεις που είναι σημειωμένες με ⊗ έχουν προκύψει από συγχώνευση.*

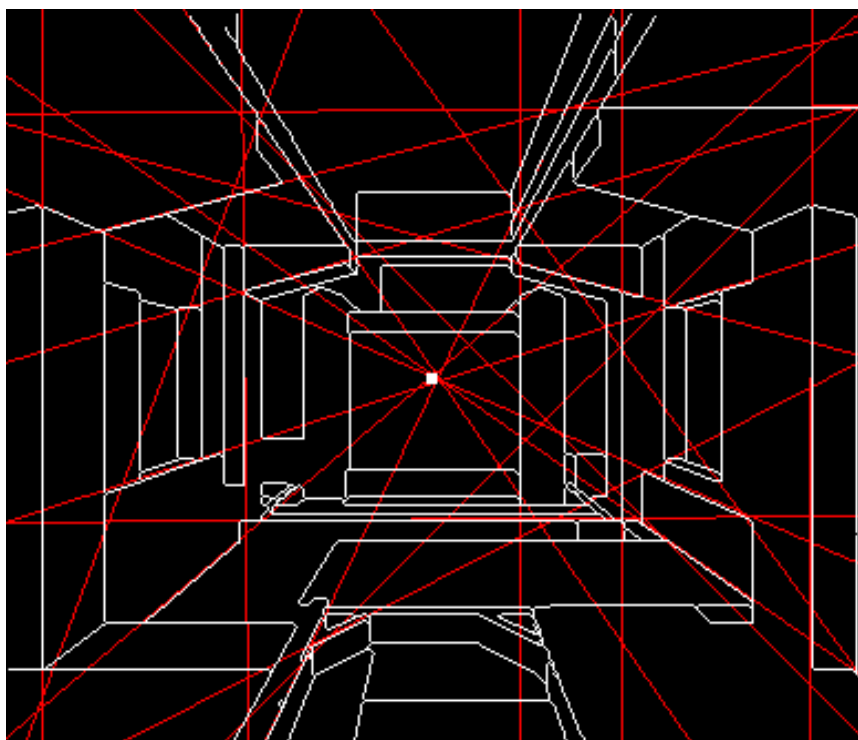
#### **2.2.1.4. Εύρεση σημείου φυγής από τις κατευθύνσεις των ευθυγράμμων τμημάτων με απευθείας εύρεση τομών στο επίπεδο**

Αφού έχουν βρεθεί οι διευθύνσεις των κυρίαρχων ευθυγράμμων τμημάτων της εικόνας, οι αντίστοιχες ευθείες εξετάζονται ανά δύο και σχηματίζεται μία λίστα με σημεία τομών. Οι ευρεθείσες τομές ομαδοποιούνται με τη χρήση ενός αλγορίθμου ιεραρχικής τμηματοποίησης, παραλλαγή του γνωστού αλγορίθμου average linkage [2]. Ο αλγόριθμος έχει ως εξής:

- i. Κάθε σημείο τομής αρχικά αποτελεί από μόνο του μία ομάδα με κέντρο το ίδιο το σημείο.
- ii. Βρίσκονται οι ομάδες που τα κέντρα τους απέχουν τη μικρότερη απόσταση. Αν η απόσταση είναι μικρότερη από ένα κατώφλι, οι ομάδες ενώνονται και στην προκύπτουσα ομάδα κέντρο θεωρείται η μέση τιμή των σημείων που την αποτελούν.
- iii. Αν στο προηγούμενο βήμα έγινε ένωση (η μικρότερη απόσταση ήταν κάτω από το κατώφλι) τότε έχουμε ξανά επανάληψη του αλγορίθμου από το βήμα ii.

Τα κέντρα των ομάδων μετά την τμηματοποίηση θα είναι και τα πιθανά σημεία φυγής

Στην Εικόνα 2-5 φαίνεται η λειτουργία του αλγορίθμου. Το σημείο φυγής που υπολογίστηκε προέκυψε από την ομαδοποίηση των 15 σημείων τομής που σχηματίζουν μεταξύ τους οι έξι γραμμές που τέμνονται κοντά σε αυτό. Ο αλγόριθμος υπολόγισε το κυρίαρχο σημείο φυγής της εικόνας στη θέση  $(x,y)=[194.582,166.624]$ . Στην εικόνα φαίνονται και οι κυρίαρχες διευθύνσεις όπως αυτές εξάχθηκαν από τον πίνακα Hough στην Εικόνα 2-4. Το σημείο φυγής που υπολογίστηκε φαίνεται σημειωμένο με ένα άσπρο τετράγωνο περίπου στο κέντρο της εικόνας



*Εικόνα 2-5. Οι κυριότερες κατευθύνσεις και το σημείο φυγής για την Εικόνα 2-1.*

### 2.2.1.5. Εύρεση σημείου φυγής με προβολή στον μοναδιαίο κύκλο.

Οι Magee et al. [19] παρουσίασαν μία μέθοδο (μέθοδος cross-product) που αποτελεί παραλλαγή της παραπάνω τεχνικής. Χρησιμοποιούν πράξεις με εξωτερικά γινόμενα για να υπολογίζουν τις προβολές των τομών των ευθυγράμμων τμημάτων πάνω στη μοναδιαία σφαίρα που έχει κέντρο την αρχή των αξόνων. Η μοναδιαία σφαίρα έχει παραμετροποιηθεί με χρήση πολικών συντεταγμένων και έχει δημιουργηθεί ένα ιστόγραμμα. Για κάθε τομή που υπολογίζεται, η τιμή στην αντίστοιχη θέση του ιστογράμματος αυξάνεται κατά ένα. Με αυτό τον τρόπο, όταν έχουν υπολογιστεί όλες οι τομές, οι κορυφές του ιστογράμματος θα αντιστοιχούν στα πιθανά σημεία φυγής.

Τη μοναδιαία σφαίρα τη χρησιμοποίησε για πρώτη φορά ο Barnard [15]. Στην δικιά του υλοποίηση η ανίχνευση των ευθυγράμμων τμημάτων γίνεται με μια τεχνική που συνδυάζει ανίχνευση συνοριακών γραμμών (contours), σπάσιμό τους σε μικρά ευθύγραμμα κομμάτια με κάποιο αναδρομικό αλγόριθμο, και επανασύνδεσή τους σε μεγάλα ευθύγραμμα τμήματα. Για κάθε ευθύγραμμο τμήμα βρίσκεται το επίπεδο που ορίζεται από αυτό και την αρχή των αξόνων (interpretation plane). Το επίπεδο αυτό τέμνει τη μοναδιαία σφαίρα σε ένα μέγιστο κύκλο με κέντρο την αρχή των αξόνων. Τα σημεία του ιστογράμματος της μοναδιαίας σφαίρας από τα οποία περνάει ο κάθε μέγιστος κύκλος σημαδεύονται και αυξάνεται η τιμή τους κατά ένα. Πιθανά σημεία φυγής θα είναι κορυφές του τελικού ιστογράμματος (σημεία τομής πολλών μέγιστων κύκλων μαζί).

Οι Collins et al. [16] πρόβαλλαν με τον ίδιο τρόπο όπως ο Barnard [15] τα ευθύγραμμα τμήματα στη μοναδιαία σφαίρα αλλά για κάθε τμήμα της, εκτός από το άθροισμα των μέγιστων κύκλων που έχουν περάσει από αυτό, αποθηκεύουν και τις κατευθύνσεις των ευθυγράμμων τμημάτων από τα οποία προέρχεται ο κάθε μέγιστος κύκλος. Με αυτό τον τρόπο, μετά την τελική ανίχνευση των σημείων φυγής, μπορούν να ξέρουν από ποια ευθύγραμμα τμήματα προέρχεται το κάθε ένα, και να ξαναυπολογίσουν με ακρίβεια τη θέση του, βρίσκοντας τομές.

Παρακάτω υλοποιούμε μια μέθοδο, παραλλαγή της μεθόδου του Barnard [15]. Αντί να ανιχνεύσουμε όμως ευθύγραμμα τμήματα προβάλουμε στη μοναδιαία σφαίρα την πληροφορία που εμπεριέχει ο μετασχηματισμός Hough της εικόνας.

Έστω ότι  $H[m,n]$  είναι η τιμή του  $[m,n]$  στοιχείου του ιστογράμματος Hough και έστω ότι η  $[m,n]$  θέση στο ιστόγραμμα Hough αντιστοιχεί στη διεύθυνση  $(\rho^{mn}, \theta^{mn})$  στο καρτεσιανό επίπεδο της εικόνας. Όπως φαίνεται στην Εικόνα 2-6 η διεύθυνση αυτή σε

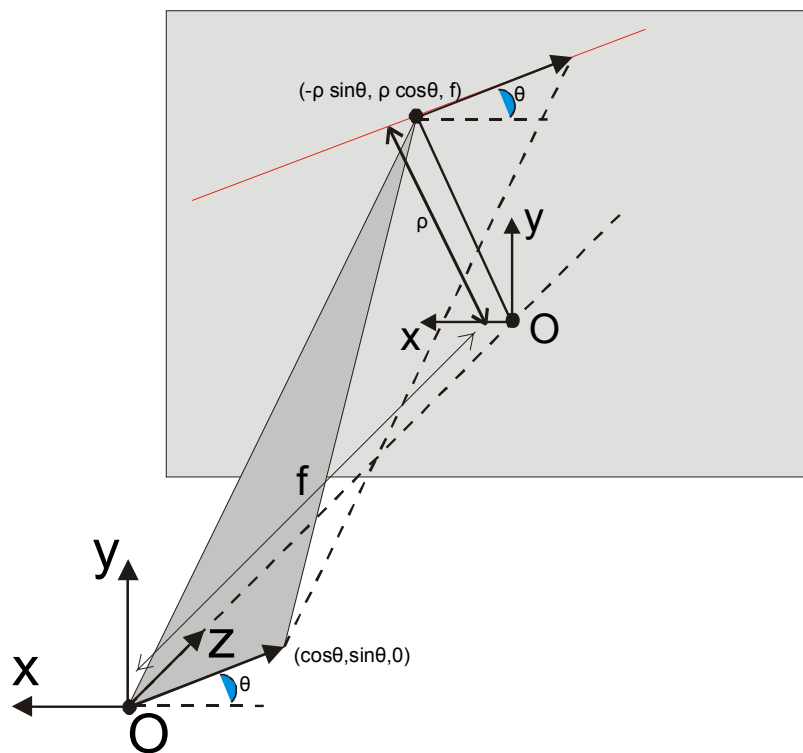
συνδυασμό με την αρχή των αξόνων ορίζει ένα επίπεδο (interpretation plane) με κάθετο διάνυσμα

$$\mathcal{P}^{mn} = \begin{bmatrix} v_x^{mn} \\ v_y^{mn} \\ v_z^{mn} \end{bmatrix} = \begin{bmatrix} \cos(\theta) \\ \sin(\theta) \\ 0 \end{bmatrix} \times \begin{bmatrix} -\rho \sin(\theta) \\ \rho \cos(\theta) \\ f \end{bmatrix} \quad \text{αν} \quad \theta \in [0, \pi/2] \quad \text{Σχέση 2-7}$$

ή

$$\mathcal{P}^{mn} = \begin{bmatrix} v_x^{mn} \\ v_y^{mn} \\ v_z^{mn} \end{bmatrix} = \begin{bmatrix} \cos(\theta) \\ \sin(\theta) \\ 0 \end{bmatrix} \times \begin{bmatrix} \rho \cos(\theta) \\ \rho \sin(\theta) \\ f \end{bmatrix} \quad \text{αν} \quad \theta \in (\pi/2, \pi] \quad \text{Σχέση 2-8}$$

όπου  $f$  είναι το εστιακό μήκος.



Εικόνα 2-6. το επίπεδο που ορίζεται από μια διεύθυνση  $(\rho, \theta)$  πάνω στην εικόνα και την αρχή των αξόνων.

Αν  $s^{mn-i} : (s_x^{mn-i}, s_y^{mn-i}, s_z^{mn-i})$  είναι ένα σημείο πάνω στον μέγιστο κύκλο που ορίζεται από την τομή του επιπέδου με κάθετο διάνυσμα  $\mathcal{P}^{mn}$  και της μοναδιαίας σφαίρας με κέντρο την αρχή των αξόνων, τότε από την Εικόνα 2-7 ισχύουν οι σχέσεις:

$$\tan \phi = \frac{s_x^{mn-i}}{s_z^{mn-i}}$$

Σχέση 2-9

$$\tan \theta = \frac{s_y^{mn-i}}{\sqrt{s_x^{mn-i^2} + s_z^{mn-i^2}}}$$

Σχέση 2-10

επειδή το σημείο  $s^{mn-i}$  ανήκει στη μοναδιαία σφαίρα θα είναι

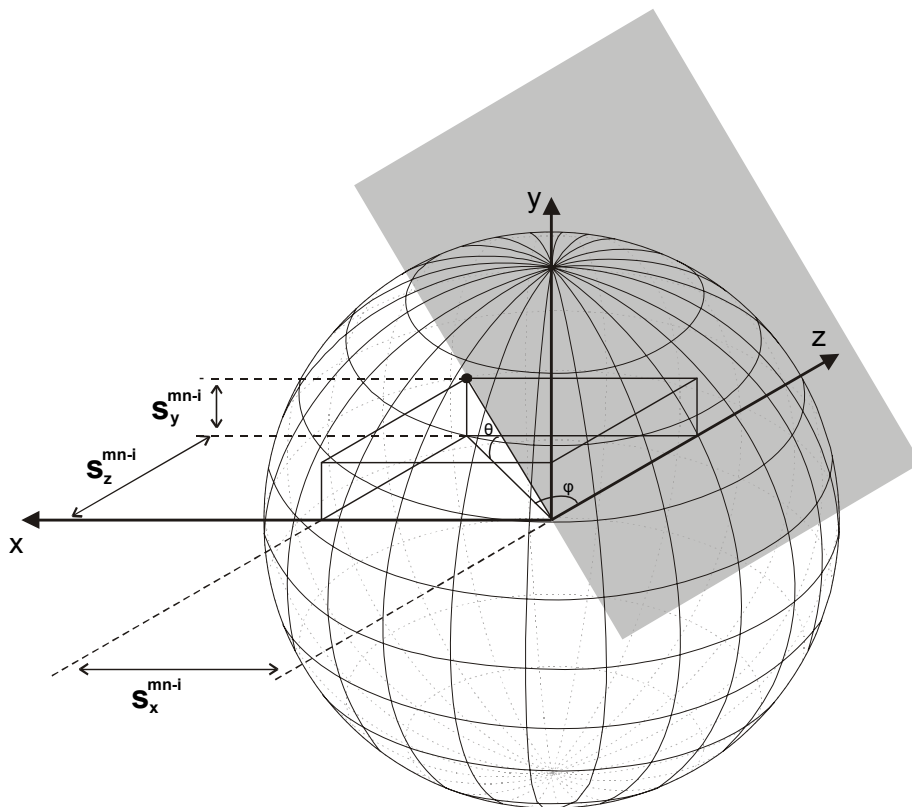
$$s_x^{mn-i^2} + s_y^{mn-i^2} + s_z^{mn-i^2} = 1$$

Σχέση 2-11

και επειδή το σημείο  $s^{mn-i}$  ανήκει στο επίπεδο με κάθετο διάνυσμα  $\mathcal{V}^{mn}$  θα είναι ακόμα:

$$v_x^{mn} s_x^{mn-i} + v_y^{mn} s_y^{mn-i} + v_z^{mn} s_z^{mn-i} = 0$$

Σχέση 2-12



Εικόνα 2-7. Ένα σημείο πάνω στη τομή του επιπέδου με κάθετο διάνυσμα  $\mathcal{V}^{mn}$  και στη μοναδιαία σφαίρα με κέντρο την αρχή των αξόνων.

Λύνοντας το σύστημα των παραπάνω τεσσάρων εξισώσεων καταλήγουμε στην εξίσωση του μέγιστου κύκλου που βρίσκεται πάνω στη μοναδιαία σφαίρα με κέντρο την αρχή των αξόνων.

$$\theta = \arctan \left( \frac{s_x^{mn-i} \tan(\phi) + s_z^{mn-i}}{s_y^{mn-i} \sqrt{\tan(\phi)^2 + 1}} \right) \quad \text{Σχέση 2-13}$$

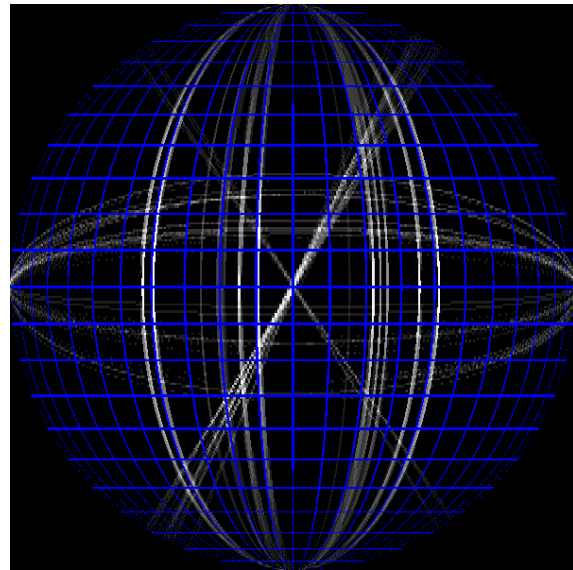
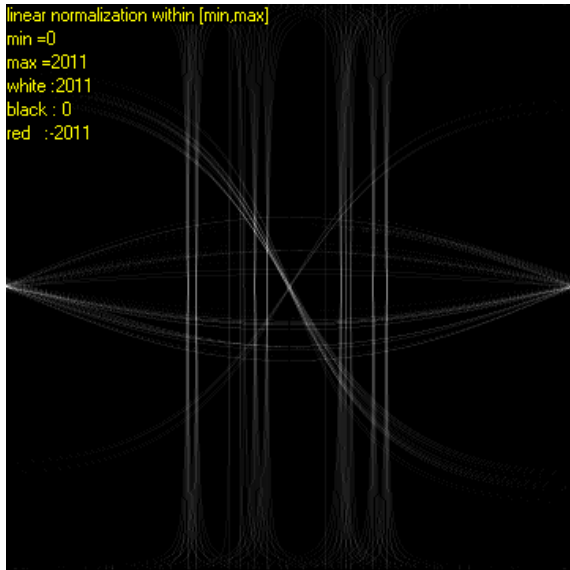
ή αλλιώς:

$$\phi = \arctan \left( \frac{s_x^{mn-i} s_z^{mn-i} + s_y^{mn-i} \tan(\theta) \sqrt{s_x^{mn-i^2} - s_y^{mn-i^2} \tan(\theta)^2 + s_z^{mn-i^2}}}{s_x^{mn-i^2} - s_y^{mn-i^2} \tan(\theta)^2} \right) \quad \text{Σχέση 2-14}$$

Κάθε καταχώριση στον πίνακα Hough που η τιμή της υπερβαίνει ένα προκαθορισμένο κατώφλι, προβάλλεται στη μοναδιαία σφαίρα με ένα βάρος που είναι η ίδια η τιμή της καταχώρισης. Τα σημεία φυγής ανιχνεύονται σαν κορυφές στο ιστόγραμμα  $\phi$ - $\theta$  της μοναδιαίας σφαίρας.

Στην Εικόνα 2-8 βλέπουμε το αποτέλεσμα του αλγορίθμου όπως αυτός εφαρμόστηκε για τη σκηνή στην Εικόνα 2-1. Το σημείο φυγής στο κέντρο της εικόνας ( $\theta=\pi/2$ ,  $\phi=0$ ) καθώς και το σημείο φυγής για της οριζόντιες γραμμές ( $\theta=\pi/2$ ,  $\phi=\pm\pi/2$ ) ανιχνεύθηκαν σωστά. Το σημείο φυγής για της κατακόρυφες γραμμές ( $\theta=0$  και  $\theta=\pi$ ) δεν μπορούσε να ανιχνευθεί. Ο λόγος είναι ότι η άνιση τμηματοποίηση της σφαίρας εμποδίζει την ανίχνευση σημείων κοντά στους πόλους της σφαίρας γιατί εκεί τα τμήματα είναι πολύ μικρά και τα ίχνη που αφήνουν οι μέγιστοι κύκλοι κατανέμονται σε πολλά τμήματα (δεν τέμνονται όπως φαίνεται στην Εικόνα 2-8α)

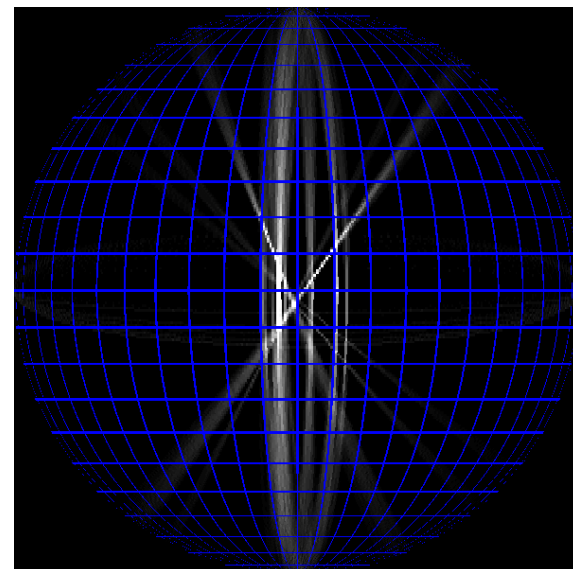
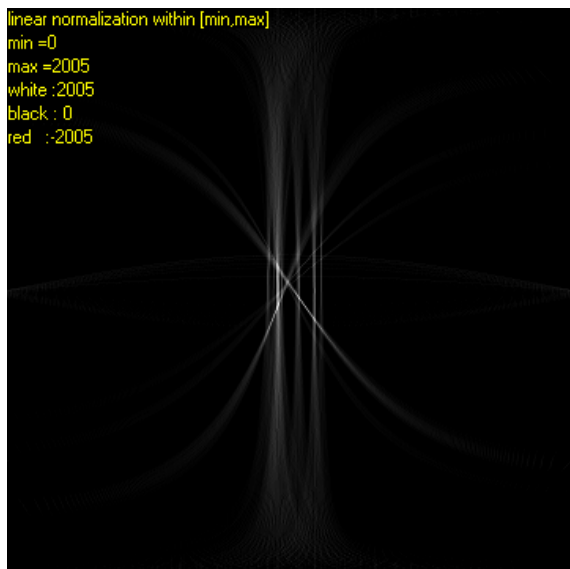
Στην Εικόνα 2-9 βλέπουμε την εφαρμογή του αλγορίθμου για την Εικόνα 2-2. Ο αλγόριθμος απέτυχε να βρει σημεία φυγής λόγω των πολλών συγκεντρωμένων κατακόρυφων ευθυγράμμων τμημάτων στην αρχική εικόνα, τα οποία ο μετασχηματισμός Hough δεν μπόρεσε να διακρίνει.



α. Οι τιμές του  $\varphi$ - $\theta$  ιστογράμματος

β. Οι τιμές πάνω στη μοναδιαία σφαίρα

Εικόνα 2-8. Η μοναδιαία σφαίρα όπως έχει διαμορφωθεί μετά την προβολή της πληροφορίας από τον πίνακα Hough για τη σκηνή στη Εικόνα 2-1.



α. Οι τιμές του  $\varphi$ - $\theta$  ιστογράμματος

β. Οι τιμές πάνω στη μοναδιαία σφαίρα

Εικόνα 2-9. Η μοναδιαία σφαίρα όπως έχει διαμορφωθεί μετά την προβολή της πληροφορίας από τον πίνακα Hough για τη σκηνή στην Εικόνα 2-2.



Μια παραλλαγή του παραπάνω αλγορίθμου που μελετήθηκε επίσης είναι με χρήση των διανυσμάτων της βαθμίδας της φωτεινότητας στη θέση των κατευθύνσεων του πίνακα Hough.

Αν  $I[x,y]$  είναι η φωτεινότητα της εικόνας στο σημείο  $[x,y]$  τότε θα ισχύει για τη βαθμίδα της φωτεινότητας  $\nabla I$  στο σημείο  $[x,y]$ :

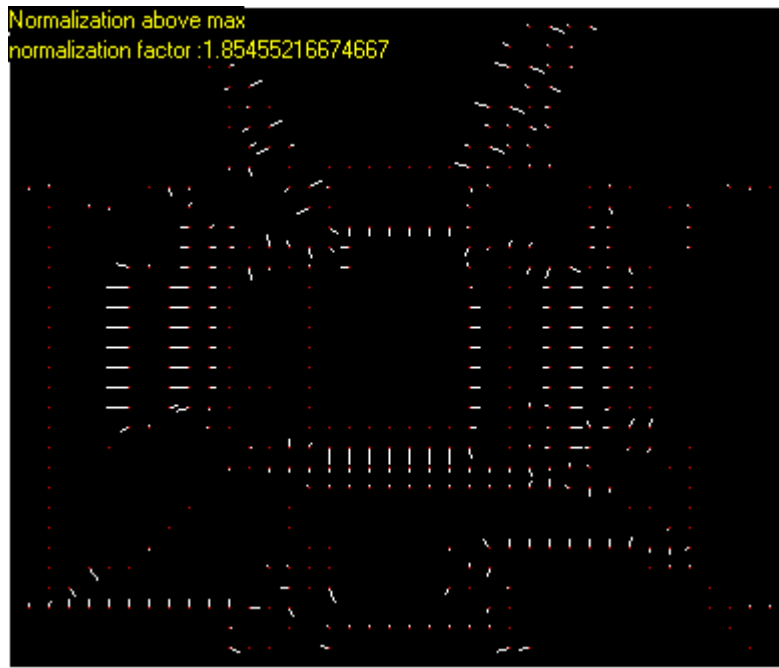
$$\nabla I[x,y] = \begin{bmatrix} \frac{\partial I[x,y]}{\partial x} \\ \frac{\partial I[x,y]}{\partial y} \end{bmatrix}$$

Σχέση 2-15

Η βαθμίδα  $\nabla I[x,y]$  μπορεί να υπολογιστεί πολύ εύκολα με χρήση μασκών όπως οι  $7 \times 7$  μάσκες τύπου Prewitt που χρησιμοποιήθηκαν στην υλοποίησή μας. Οι μάσκες αυτές έχουν τη μορφή:

-1	-1	-1	0	1	1	1	-1	-1	-1	-1	-1	-1	-1
-1	-1	-1	0	1	1	1	-1	-1	-1	-1	-1	-1	-1
-1	-1	-1	0	1	1	1	-1	-1	-1	-1	-1	-1	-1
-1	-1	-1	0	1	1	1	0	0	0	0	0	0	0
-1	-1	-1	0	1	1	1	1	1	1	1	1	1	1
-1	-1	-1	0	1	1	1	1	1	1	1	1	1	1
-1	-1	-1	0	1	1	1	1	1	1	1	1	1	1

Στην Εικόνα 2-10 φαίνεται η βαθμίδα της φωτεινότητας όπως αυτή έχει προκύψει για την τεχνητή σκηνή στην Εικόνα 2-1. Πριν την εφαρμογή των παραπάνω μασκών, η εικόνα πέρασε από ένα φίλτρο μέσης τιμής για να είναι πιο ομαλές οι μεταβάσεις στα επίπεδα της φωτεινότητας και να έχουμε καλύτερο υπολογισμό της βαθμίδας της.



Εικόνα 2-10. Η βαθμίδα της φωτεινότητας για τη σκηνή στην Εικόνα 2-1.

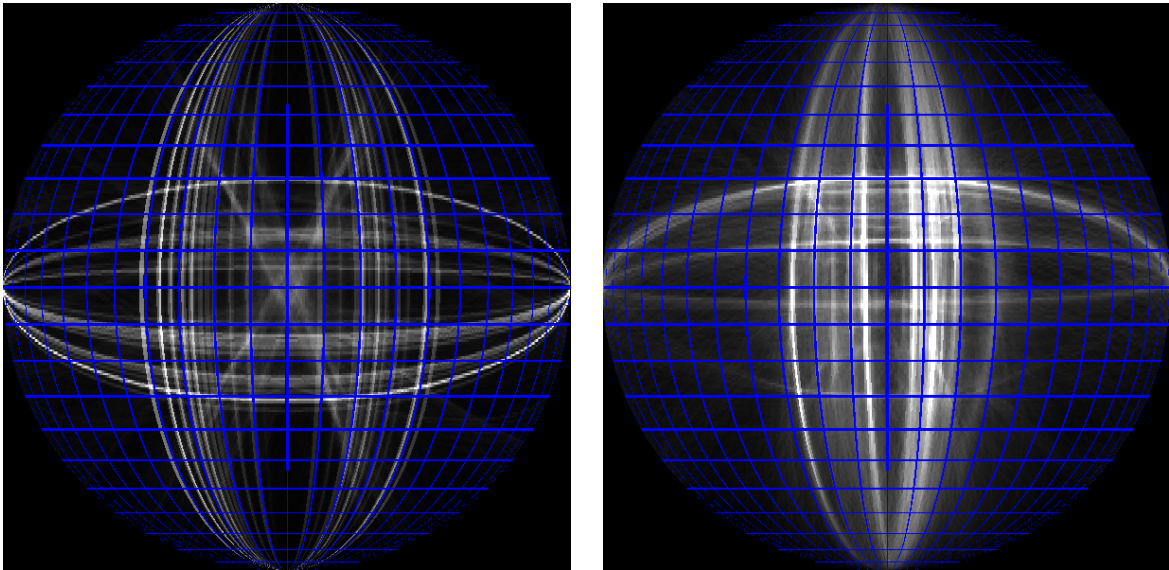
Επειδή η βαθμίδα της φωτεινότητας τείνει να είναι κάθετη στις ακμές της εικόνας, μας δίνει μια ένδειξη για τις διευθύνεις των τμημάτων που υπάρχουν σε αυτή. Κάθε σημείο  $[x,y]$  «ψηφίζει» μια διεύθυνση  $k[x,y]$  την οποία υπολογίζει από τη βαθμίδα της φωτεινότητας ως:

$$k[x,y] = \left[ -\frac{\partial I[x,y]}{\partial y}, \frac{\partial I[x,y]}{\partial x} \right] \quad \text{Σχέση 2-16}$$

Η παραπάνω διεύθυνση, σε συνδυασμό με την αρχή των αξόνων, ορίζει το επίπεδο το οποία τέμνει τη μοναδιαία σφαίρα στο μέγιστο κύκλο που δίνεται από τη Σχέση 2-13 (ή εναλλακτικά, για επίπεδα που έχουν μικρή γωνία με το επίπεδο  $y-z$ , από τη Σχέση 2-14). Το ιστόγραμμα της μοναδιαίας σφαίρας συμπληρώνεται από τα «ίχνη» των μέγιστων κύκλων όπως παραπάνω και τα σημεία φυγής υπολογίζονται με τον ίδιο τρόπο ως κορυφές του ιστογράμματος.

Στην Εικόνα 2-11 φαίνεται το αποτέλεσμα του αλγορίθμου όπως αυτός εφαρμόστηκε για τη σκηνή στην Εικόνα 2-1. Ο αλγόριθμος ανίχνευσε κορυφές για τα σημεία φυγής στο άπειρο αλλά δεν κατάφερε να βρει το σημείο φυγής στο κέντρο της εικόνας. Αυτό οφείλεται στη κακή εκτίμηση της βαθμίδας της φωτεινότητας για ευθύγραμμα τμήματα που δεν είναι εντελώς κατακόρυφα ή οριζόντια. Στην Εικόνα 2-11 φαίνεται το αποτέλεσμα του αλγορίθμου για την Εικόνα 2-2. Αν και στα σημεία φυγής στο

άπειρο ανιχνεύθηκαν κορυφές, η τιμή τους ήταν πολύ μικρή σε σχέση με κορυφές μέσα στην εικόνα που προέκυπταν από τομές ζευγαριών ευθυγράμμων τμημάτων



α. Οι τιμές του  $\varphi$ - $\theta$  ιστογράμματος πάνω στη μοναδιαία σφαίρα για την Εικόνα 2-1.      β. Οι τιμές του  $\varphi$ - $\theta$  ιστογράμματος πάνω στη μοναδιαία σφαίρα για την Εικόνα 2-2.

*Εικόνα 2-11. Η μοναδιαία σφαίρα όπως έχει διαμορφωθεί μετά την προβολή της πληροφορίας από τη βαθμίδα της φωτεινότητας για την Εικόνα 2-1 και την Εικόνα 2-2.*

### 2.2.2. Απευθείας εύρεση σημείου φυγής

Το βασικό πλεονέκτημα τέτοιων απευθείας μεθόδων είναι ότι δεν βασίζονται στην ανίχνευση ευθυγράμμων τμημάτων αλλά εφαρμόζονται κατευθείαν στην αρχική εικόνα. Όλα γίνονται σε ένα βήμα και έτσι δεν έχουμε καθόλου σφάλματα να προωθούνται από βήμα σε βήμα. Ένα ακόμα χαρακτηριστικό τους είναι ότι μπορούν να εφαρμοστούν και σε περιπτώσεις όπου δεν θέλουμε να υπολογίσουμε όλα τα σημεία φυγής αλλά μόνο μερικά για τα οποία έχουμε ενδείξεις για τη θέση τους. (π.χ. θέλουμε να υπολογίσουμε σημεία φυγής στο άπειρο μόνο). Θα μπορούσαν ακόμα οι τεχνικές αυτές να εφαρμοστούν και σε περιπτώσεις όπου ξέρουμε τα σημεία φυγής και θέλουμε απλά να τα αξιολογήσουμε ή να υπολογίσουμε με περισσότερη ακρίβεια τη θέση τους.

Η μέθοδος που αναπτύξαμε βασίζεται στον ορισμό μιας συνάρτησης αξιολόγησης που μας δίνει την πιθανότητα ένα σημείο  $v(v_x, v_y)$  να είναι σημείο φυγής. Αν και η μέθοδος είναι ενιαία και ακολουθεί σαφή βήματα για την εύρεση σημείων φυγής, μελετήσαμε τρεις διαφορετικές συναρτήσεις αξιολόγησης που παρουσιάζονται στη συνέχεια.

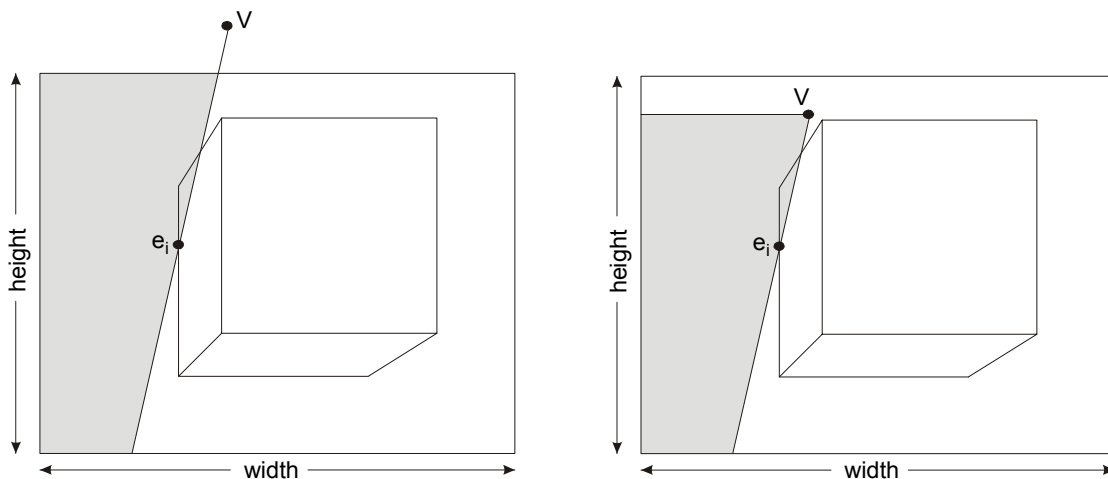
### 2.2.2.1. Αξιολόγηση με τη συνάρτηση VPLF

Έστω ότι  $e_i$  είναι ένα σημείο ακμής, με  $1 \leq e_i \leq n$  και έστω ότι  $v:(v_x, v_y)$  είναι ένα ενδεχόμενο σημείο φυγής. Ορίζουμε την συνάρτηση συνεισφοράς του σημείου ακμής  $e_i$  για το  $v$  να είναι σημείο φυγής ως:

$$F_v(e_i) = \frac{E_v(e_i)}{w \cdot h} \quad \text{Σχέση 2-17}$$

όπου  $E_v(e_i)$  είναι το εμβαδόν που φαίνεται σκιασμένο στην Εικόνα 1-1 και  $w, h$  το πλάτος και το ύψος της εικόνας αντίστοιχα.

Αφού η ποσότητα  $w \cdot h$  ισούται με το εμβαδόν όλης της εικόνας, δηλαδή τη μέγιστη τιμή που μπορεί να πάρει το εμβαδόν  $E_v(e_i)$ , εύκολα συμπεραίνει κάποιος ότι  $0 \leq F_v(e_i) \leq 1$ .



περίπτωση α. το ενδεχόμενο σημείο φυγής  $v$  είναι έξω από τα όρια της εικόνας  
περίπτωση β. το ενδεχόμενο σημείο φυγής  $v$  είναι μέσα στην εικόνα.

Εικόνα 2-12. Η συνάρτηση συνεισφοράς  $F_v(e_i)$ .

Η συνάρτηση συνεισφοράς  $F_v(e_i)$  είναι πρακτικά ένας τρόπος να εκφραστεί η κατεύθυνση της γραμμής που ενώνει το πιθανό σημείο φυγής  $v$  και το σημείο ακμής  $e_i$ . Εύκολα καταλαβαίνει κανείς ότι σημεία ακμών που ανήκουν σε ευθύγραμμα τμήματα που δείχνουν προς το πιθανό σημείο φυγής τείνουν να παρουσιάζουν παραπλήσιες τιμές της συνάρτησης συνεισφοράς τους. Η πραγματική κατεύθυνση (γωνία) της γραμμής που ενώνει το πιθανό σημείο φυγής  $v$  και το σημείο ακμής  $e_i$  δεν θα μπορούσε να έχει χρησιμοποιηθεί στη θέση της συνάρτησης συνεισφοράς, γιατί, για σημεία  $v$  μακριά από το κέντρο της εικόνας η κατεύθυνση τείνει να γίνει ανεξάρτητη του σημείου ακμής  $e_i$  και

επηρεάζεται από τη θέση του  $\nu$  και μόνο. Η συνάρτηση συνεισφοράς μπορεί να θεωρηθεί σαν μια σταθμισμένη εκδοχή της πραγματικής κατεύθυνσης. Παρατηρεί κανείς ακόμα ότι παρουσιάζει συνέχεια στα άκρα της εικόνας και έχει την ιδιότητα να παίρνει την ίδια τιμή όταν υπολογίζεται για σημεία που ανήκουν σε ευθύγραμμα τμήματα που «βλέπουν» στο σημείο φυγής. Αυτό ακριβώς το χαρακτηριστικό χρησιμοποιείται στη συνέχεια για την ανίχνευση πιθανών σημείων φυγής.

Ο στόχος είναι να ανιχνευθούν πιθανά σημεία φυγής σαν σημεία που «βλέπουν» μεγάλες ομάδες (που αντιστοιχούν σε ευθύγραμμα τμήματα που «δείχνουν» προς το υποψήφιο σημείο φυγής) από σημεία ακμών με παρόμοιες τιμές της συνάρτησης  $F_\nu(e_i)$ . Για να επιτύχουμε αυτό το σκοπό υπολογίζουμε τη κατανομή  $DF_\nu(x)$  των τιμών της  $F_\nu(e_i)$ . Αφού  $0 \leq F_\nu(e_i) \leq 1$ , η παράμετρος  $x$  της κατανομής  $DF_\nu(x)$  θα παίρνει τιμές στο διάστημα  $[0, 1]$ .

Τα ευθύγραμμα τμήματα που «δείχνουν» προς το υποψήφιο σημείο φυγής  $\nu$  ανιχνεύονται σαν κορυφές στο ιστόγραμμα της  $DF_\nu(x)$ . Ιστογράμματα που παρουσιάζουν δύο ή περισσότερες υψηλές και καθαρές κορυφές αντιστοιχούν σε πιθανά σημεία φυγής (Εικόνα 2-14).

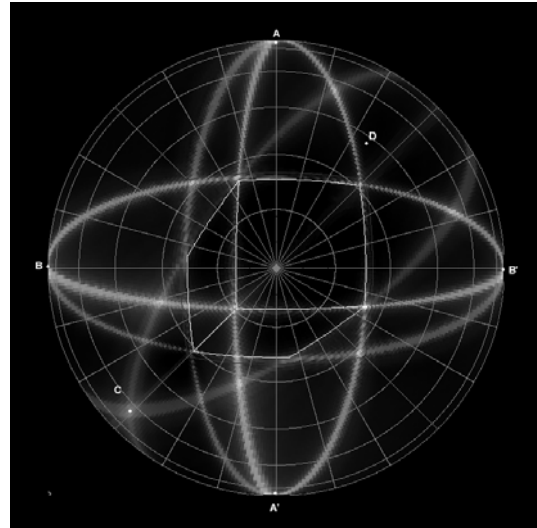
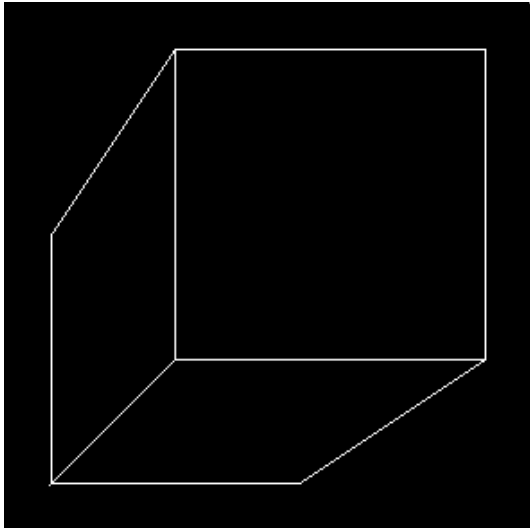
Το εμβαδόν της περιοχής κάτω από το ιστόγραμμα της  $DF_\nu(x)$  είναι πάντα αριθμητικά ίσο με τον αριθμό των σημείων ακμής, συνεπώς, το εμβαδόν του τετραγωνισμένο ιστογράμματος θα τείνει να είναι μεγαλύτερο για ιστογράμματα που παρουσιάζουν μεγάλες και καθαρές κορυφές. Ορίζουμε λοιπόν την συνάρτηση αξιολόγησης  $VPLF_\nu$  ως:

$$VPLF_\nu = \int_{x=0}^1 DF_\nu^2(x) \cdot dx \quad \text{Σχέση 2-18}$$

Μεγαλύτερες τιμές της  $VPLF_\nu$  θα αντιστοιχούν σε σημεία πιο πιθανά να είναι σημεία φυγής

Στην Εικόνα 2-13 βλέπουμε την εικόνα ενός κύβου και την προβολή των τιμών της συνάρτησης αξιολόγησης πάνω στη μοναδιαία σφαίρα. Το ιστόγραμμα με τις τιμές της συνάρτησης αξιολόγησης φαίνεται στην Εικόνα 2-15.

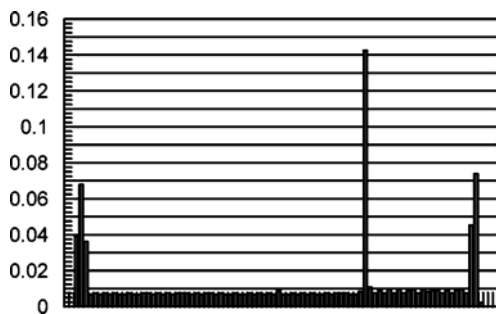
Στην Εικόνα 2-16 και την Εικόνα 2-17 βλέπουμε εφαρμογές του αλγορίθμου σε εικόνες διαφορετικής φύσης.



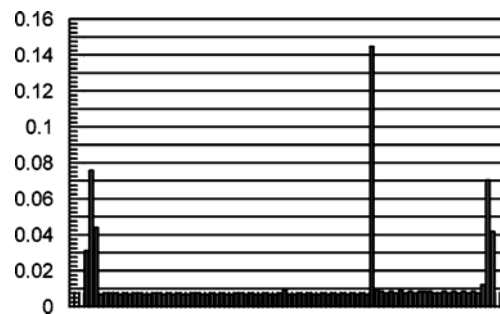
Εικόνα 2-13. Μια απλή εικόνα ενός κύβου και οι τιμές της συνάρτησης VPLF πάνω στη μοναδιαία σφαίρα.

Παραμετροποίηση  $\varphi$ - $\theta$ . Επιλέγουμε πιθανά σημεία φυγής από την επιφάνεια της σφαίρας και υπολογίζουμε για κάθε ένα από αυτά τη VPLF. Τιμές της VPLF φαίνονται στην εικόνα σαν αποχρώσεις του γκρι. Για ευκολία, οι ακμές της εικόνας έχουν προβληθεί επίσης στη σφαίρα.

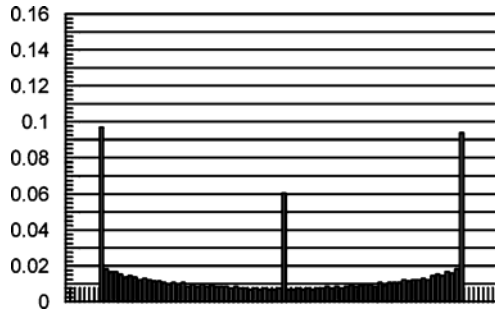
Ο πόλος της σφαίρας ( $\theta=0$ ) είναι στο κέντρο της εικόνας. Η παράμετρος  $\varphi$  παίρνει τιμές ωρολογιακά, αρχίζοντας από την αριστερή πλευρά της εικόνας.



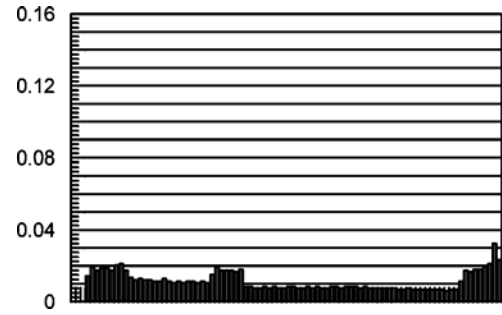
α. Η κατανομή DF για το σημείο φυγής A στην Εικόνα 2-13



β. Η κατανομή DF για το σημείο φυγής B στην Εικόνα 2-13

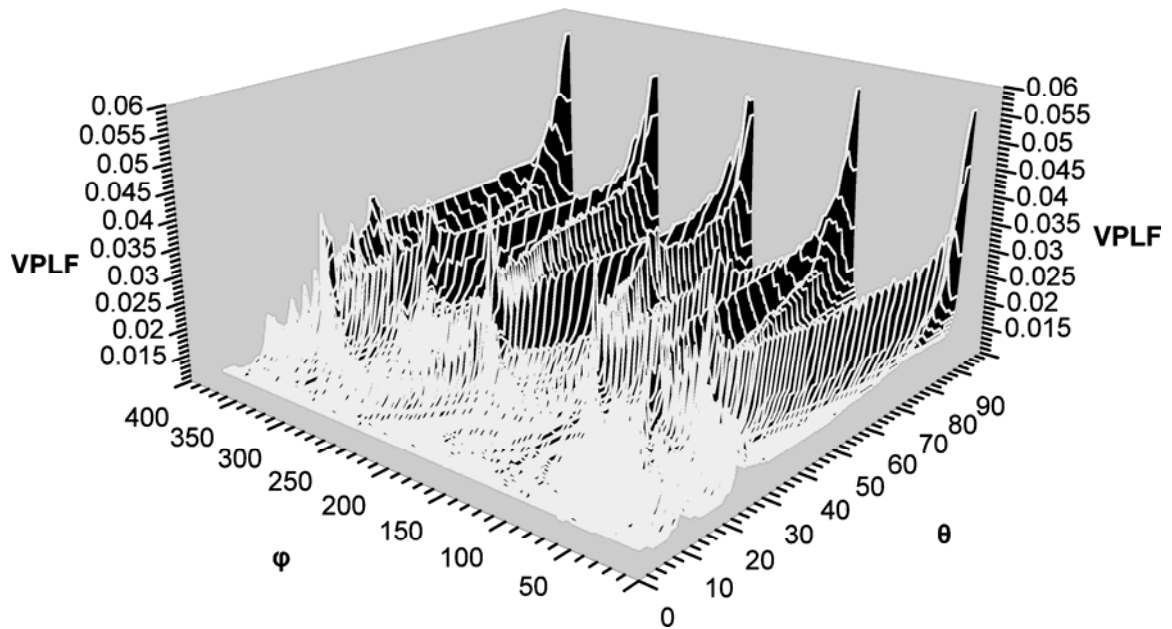


γ. Η κατανομή  $DF$  για το σημείο φυγής  $C$   
στην Εικόνα 2-13

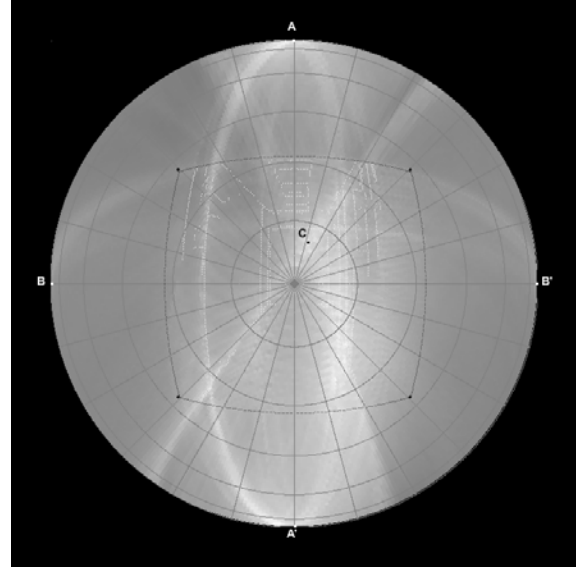
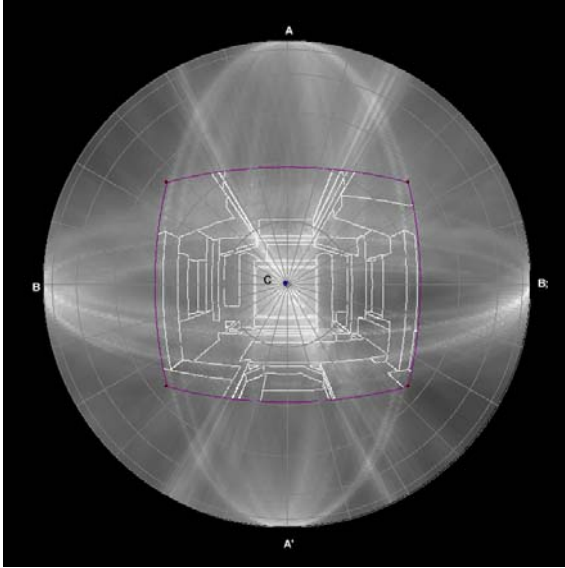


δ. Η κατανομή  $DF$  για το τυχαίο σημείο  $D$   
στην Εικόνα 2-13

Εικόνα 2-14. Οι κατανομές στα σημεία  $A, B, C, D$  στην



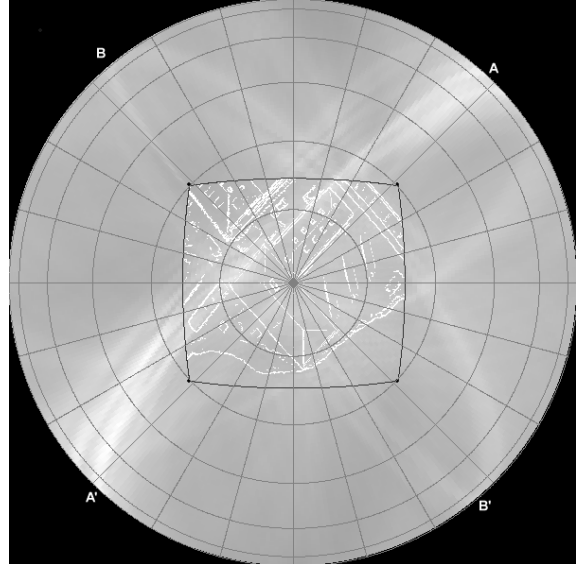
Εικόνα 2-15. Οι τιμές της  $VPLF$  για την Εικόνα 2-13



α. Η τιμές της VPLF για την Εικόνα 2-1.

α. Η τιμές της VPLF για την Εικόνα 2-1.

Εικόνα 2-16. Αξιολόγηση σημείων φυγής με τη συνάρτηση VPLF για την Εικόνα 2-1 και την Εικόνα 2-2. Η μέθοδος βρήκε επιτυχώς τα κυρίαρχα σημεία φυγής της εικόνας (σημεία A,A',B,B',C).



Εικόνα 2-17. Μια αεροφωτογραφία. Αξιολόγηση σημείων φυγής με τη συνάρτηση VPLF. Οι τιμές της VPLF έχουν προβληθεί στη μοναδιαία σφαίρα. Η μέθοδος βρήκε επιτυχώς τα κυρίαρχα σημεία φυγής της εικόνας (σημεία A,A',B,B').



### 2.2.2.2. Αξιολόγηση με ολοκλήρωση τον χώρο Hough (συνάρτηση VPLFSIN)

Έστω ότι έχουμε στο καρτεσιανό επίπεδο ευθύγραμμα τμήματα που ανήκουν σε μία δέσμη ευθειών που διέρχονται από ένα κοινό σημείο  $v(v_x, v_y)$ . Κάθε ευθεία της δέσμης θα μπορεί να γραφτεί παραμετρικά ως προς τη γωνία  $\theta$  που σχηματίζει με τον άξονα των  $x$ :

$$\theta = (x - v_x) \cos \theta + (y - v_y) \sin \theta \quad \text{Σχέση 2-19}$$

ή αλλιώς:

$$v_x \cos \theta + v_y \sin \theta = x \cos \theta + y \sin \theta \quad \text{Σχέση 2-20}$$

Η ευθεία στην παραπάνω σχέση μετασχηματίζεται σύμφωνα με τον ορισμό του μετασχηματισμού Hough στο σημείο  $(v_x \cos \theta + v_y \sin \theta, \theta)$  στο επίπεδο των πολικών συντεταγμένων. Παραμετρικά ως προς τη γωνία  $\theta$  ολόκληρη η δέσμη των ευθειών που περνά από το σημείο  $v(v_x, v_y)$  θα εικονίζεται στο επίπεδο των πολικών συντεταγμένων με τη γραμμή:

$$\rho = v_x \cos \theta + v_y \sin \theta \quad \text{Σχέση 2-21}$$

αν θέσουμε:

$$\tan \varphi = \frac{v_x}{v_y} \quad \text{Σχέση 2-22}$$

και

$$a = \frac{v_y}{\cos \varphi} \quad \text{Σχέση 2-23}$$

η γραμμή στη Σχέση 2-21 μπορεί να γραφτεί και με τη μορφή:

$$\rho = a \sin(\theta + \varphi) \quad \text{Σχέση 2-24}$$

Είναι δηλαδή μια γραμμή με ημιτονοειδή μορφή (μάλιστα είναι η ίδια ακριβώς με την γραμμή στη Σχέση 2-5, τη γραμμή δηλαδή που θα απεικονιζόταν το σημείο  $v$  αν

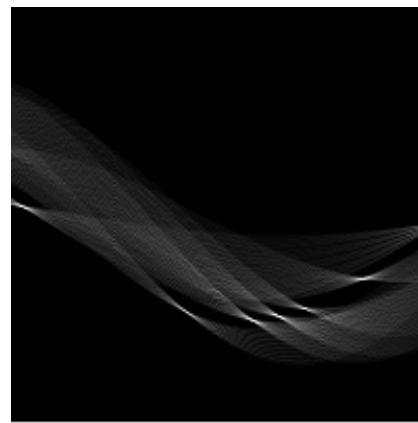
υπήρχε μέσα στην εικόνα)

Σκεπτόμενος αντίστροφα, συμπεραίνει κανείς εύκολα ότι σημεία στο ιστόγραμμα Hough τα οποία βρίσκονται πάνω σε μία ημιτονοειδή καμπύλη θα αντιστοιχούν σε ευθύγραμμα τμήματα στην αρχική εικόνα που έχουν κατευθύνσεις τέτοιες ώστε όταν προεκταθούν να τέμνονται σε ένα σταθερό σημείο, ένα πιθανό σημείο φυγής.

Στην Εικόνα 2-18 φαίνεται ο μετασχηματισμός Hough μιας εικόνας που περιέχει ευθύγραμμα τμήματα που όταν προεκταθούν τέμνονται σε ένα κοινό σημείο.



Αρχική εικόνα



Μετασχηματισμός Hough

Εικόνα 2-18. Ευθύγραμμα τμήματα που όταν προεκταθούν τέμνονται σε ένα κοινό σημείο μετασχηματίζονται σε σημείο στο ιστόγραμμα Hough που κείτονται πάνω σε μία ημιτονοειδή γραμμή

Αν  $v(v_x, v_y)$  είναι ένα πιθανό σημείο φυγής τότε θα πρέπει να υπάρχουν στην εικόνα ευθύγραμμα τμήματα που όταν προεκταθούν να τέμνονται σε αυτό. Η ημιτονοειδής γραμμή στη Σχέση 2-24 θα πρέπει να «περνάει» από τις απεικονίσεις αυτών των ευθυγράμμων τμημάτων στο ιστόγραμμα Hough.

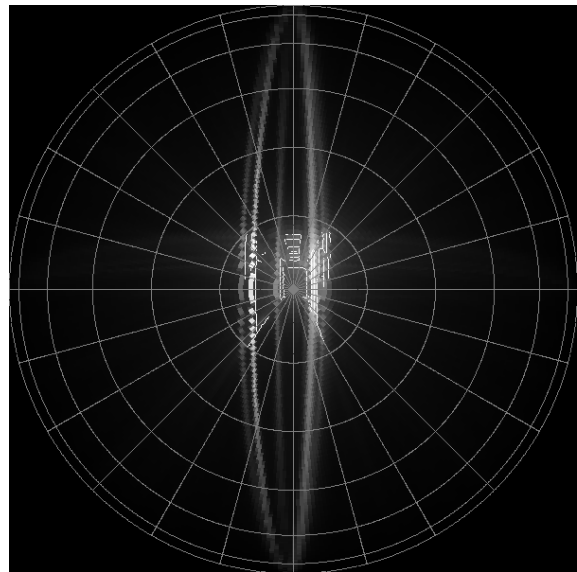
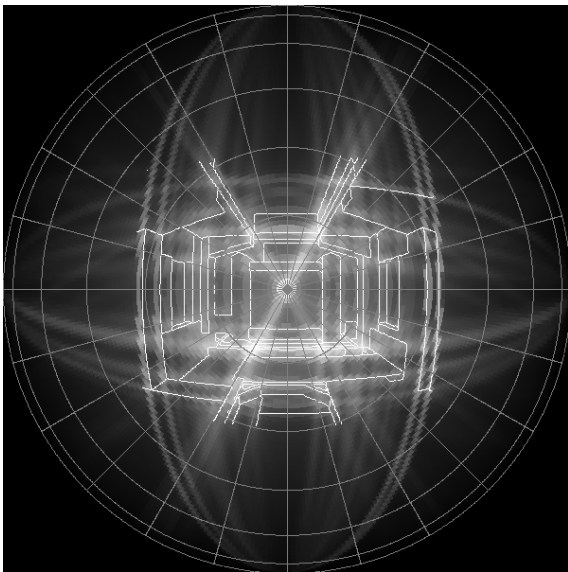
Έστω ότι  $H[m, n]$  είναι η τιμή του  $[n, m]$  στοιχείου του ιστογράμματος Hough. Έστω ότι η γραμμή που προκύπτει από τη Σχέση 2-24 για το σημείο  $v(v_x, v_y)$  διακριτοποιείται στο ιστόγραμμα Hough σε  $N$  σημεία και έστω ότι  $[v_m^k, v_n^k]$  είναι η θέση του  $k$ -οστού σημείου ( $1 < k < N$ ). Ορίζουμε τη συνάρτηση αξιολόγησης  $VPLFSIN_v$  ως:

$$VPLFSIN_v = \sum_{k=1}^N (H[v_m^k, v_n^k])^2 \quad \text{Σχέση 2-25}$$

Μεγάλες τιμές της  $VPLFSIN_v$  θα αντιστοιχούν σε ημιτονοειδής γραμμές στο

ιστόγραμμα Hough που «περνάνε» από σημεία-απεικονίσεις πολλών και έντονων ευθυγράμμων τμημάτων, πιθανών δηλαδή σημείων φυγής.

Στην Εικόνα 2-19 βλέπουμε την εφαρμογή της συνάρτησης αξιολόγησης για την Εικόνα 2-1 και την Εικόνα 2-2. Για την Εικόνα 2-1 ο αλγόριθμος κατάφερε να δώσει σαν κορυφή το σημείο φυγής στο κέντρο της εικόνας χωρίς να καταφέρει να δώσει τα σημεία φυγής στα άπειρο. Για την Εικόνα 2-2 ο αλγόριθμος δεν κατάφερε να βρει κάποιες κορυφές που να αντιστοιχούν σε σωστό σημείο φυγής. Η αιτία για την αποτυχία του αλγορίθμου είναι ότι ο μετασχηματισμός Hough δεν κατάφερε να υπολογίσει με ικανοποιητική ακρίβεια τις κατευθύνσεις των ευθυγράμμων τμημάτων.



*α. Οι τιμές της VPLFSIN πάνω στη μοναδιαία σφαίρα για την Εικόνα 2-1.*

*β. Οι τιμές της VPLFSIN πάνω στη μοναδιαία σφαίρα για την Εικόνα 2-2.*

*Εικόνα 2-19. Η μοναδιαία σφαίρα όπως έχει διαμορφωθεί μετά την προβολή των τιμών της συνάρτησης VPLFSIN για την Εικόνα 2-1 και την Εικόνα 2-2.*

### 2.2.2.3. Αξιολόγηση με προβολές της βαθμίδας της φωτεινότητας (συνάρτηση VPLFGRAD)

Αν το σημείο  $[x,y]$  ανήκει σε ένα ευθύγραμμο τμήμα (σε κάποια ακμή της εικόνας) που συγκλίνει προς ένα σημείο φυγής  $v(v_x, v_y)$  τότε η βαθμίδα της φωτεινότητας  $\overset{\rho}{\nabla} I[x,y]$  θα τείνει να είναι κάθετη στην ευθεία που ενώνει το σημείο  $[x,y]$  με το σημείο φυγής  $v(v_x, v_y)$ . Αλλιώς, το κάθετο διάνυσμα  $\overset{k}{\rho}$  στη βαθμίδα της φωτεινότητας που δίνεται από τη Σχέση 2-16 θα δείχνει προς το σημείο φυγής  $v(v_x, v_y)$ .

Αντίστροφα, αν το σημείο  $v(v_x, v_y)$  είναι ένα σημείο φυγής, θα υπάρχουν αρκετά σημεία στην εικόνα που θα ανήκουν σε ευθύγραμμα τμήματα που θα συγκλίνουν προς αυτό. Για τα περισσότερα από τα σημεία αυτά, η βαθμίδα της φωτεινότητάς τους θα είναι κάθετη στο ευθύγραμμο τμήμα που τα ενώνει με το σημείο φυγής

Η κατεύθυνσή του ευθυγράμμου τμήματος που ενώνει το σημείο  $[x,y]$  με το σημείο  $v(v_x, v_y)$  είναι:

$$\overset{\rho}{l}[x,y] = \left[ \frac{v_x - x}{\sqrt{(v_x - x)^2 + (v_y - y)^2}}, \frac{v_y - y}{\sqrt{(v_x - x)^2 + (v_y - y)^2}} \right]^T \quad \text{Σχέση 2-26}$$

Η κατεύθυνσή του διανύσματός της βαθμίδας της φωτεινότητάς δίνεται ως:

$$\overset{\rho}{m}[x,y] = \frac{\overset{\rho}{\nabla} I[x,y]}{|\overset{\rho}{\nabla} I[x,y]|} \quad \text{Σχέση 2-27}$$

ή αλλιώς:

$$\overset{\rho}{m}[x,y] = \left[ \frac{\frac{\partial I[x,y]}{\partial x}}{\sqrt{\left(\frac{\partial I[x,y]}{\partial x}\right)^2 + \left(\frac{\partial I[x,y]}{\partial y}\right)^2}}, \frac{\frac{\partial I[x,y]}{\partial y}}{\sqrt{\left(\frac{\partial I[x,y]}{\partial x}\right)^2 + \left(\frac{\partial I[x,y]}{\partial y}\right)^2}} \right] \quad \text{Σχέση 2-28}$$

Τα διανύσματα  $\overset{\rho}{l}$  και  $\overset{\rho}{m}$  θα πρέπει να είναι κάθετα (ή περίπου κάθετα) δηλαδή θα πρέπει το εσωτερικό τους γινόμενο να είναι πολύ κοντά στο μηδέν.

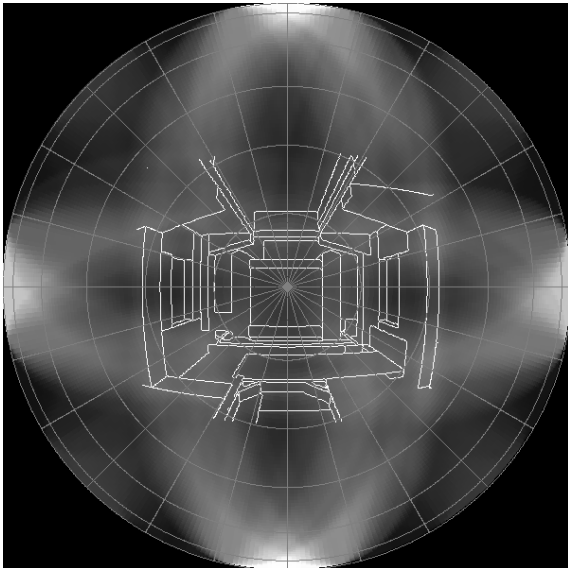
Ορίζουμε τη συνάρτηση αξιολόγησης  $VPLFGRAD_v$  που μας δίνει την πιθανότητα ένα σημείο  $v(v_x, v_y)$  να είναι σημείο φυγής ως:

$VPLFGRAD_v = \text{Άθροισμα όλων των σημείων } [x,y] \text{ της εικόνας για τα}$   
 οποία η τιμή του εσωτερικού γινομένου  $\vec{l} \cdot \vec{m}$  είναι  
 μικρότερη από ένα κατώφλι.

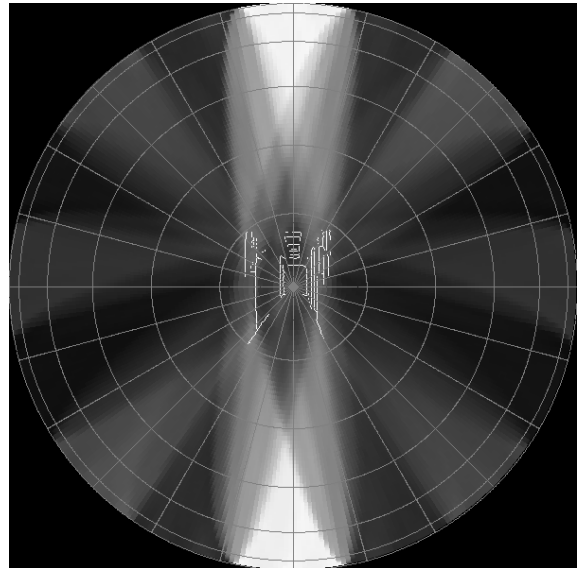
Σχέση 2-29

Μεγαλύτερες τιμές της  $VPLFGRAD_v$  θα αντιστοιχούν σε μεγαλύτερη πιθανότητα το σημείο  $v(v_x, v_y)$  να είναι σημείο φυγής.

Στην Εικόνα 2-20 βλέπουμε την εφαρμογή της συνάρτησης αξιολόγησης  $VPLFGRAD$  για την Εικόνα 2-1 και την Εικόνα 2-2. Ο αλγόριθμος κατάφερε να δώσει σημεία φυγής στο άπειρο χωρίς να καταφέρει να δώσει τα σημεία φυγής στα κέντρα των εικόνων. Η αιτία, όπως εξηγήθηκε και στην παράγραφο 2.2.1.5, είναι ότι, για τα μη κατακόρυφα ή οριζόντια τμήματα, οι εκτιμήσεις της βαθμίδας της φωτεινότητας δεν ήταν σταθερές.



α. Οι τιμές της  $VPLFGRAD$  πάνω στη μοναδιαία σφαίρα για την Εικόνα 2-1.



β. Οι τιμές της  $VPLFGRAD$  πάνω στη μοναδιαία σφαίρα για την Εικόνα 2-2.

Εικόνα 2-20. Η μοναδιαία σφαίρα όπως έχει διαμορφωθεί μετά την προβολή των τιμών της συνάρτησης  $VPLFGRAD$  για την Εικόνα 2-1 και την Εικόνα 2-2.

#### 2.2.2.4. Χρήση της συνάρτησης αξιολόγησης για την εύρεση των σημείων φυγής

Ο πιο απλός τρόπος να υπολογίσουμε σημεία φυγής είναι να σχηματίσουμε ένα σύνολο από πιθανά σημεία φυγής και να εξετάσουμε κάθε ένα από αυτά χρησιμοποιώντας μια συνάρτηση αξιολόγησης. Σημεία που παρουσιάζουν μεγάλες τιμές της συνάρτησης αξιολόγησης θα είναι πιθανά σημεία φυγής.

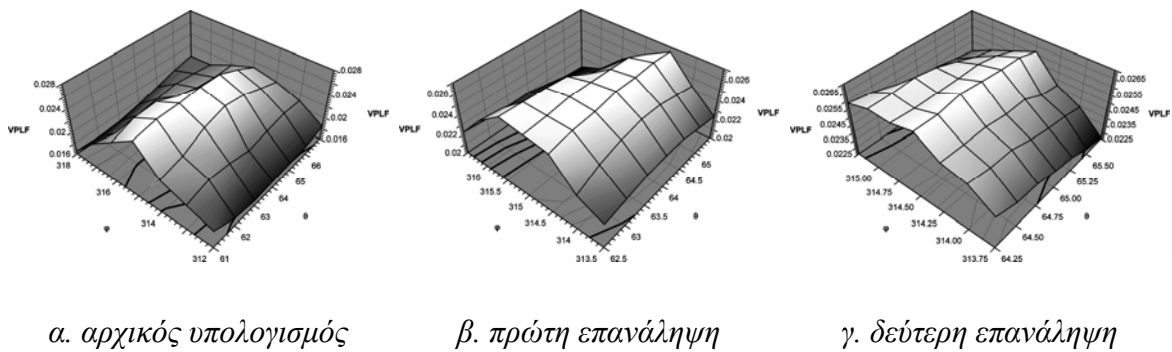
Στην υλοποίηση που έγινε στα πλαίσια αυτής της εργασίας, σχηματίζουμε το σύνολο δοκιμής επιλέγοντας σημεία από τη μοναδιαία σφαίρα και προβάλλοντας τα στο επίπεδο. Η μοναδιαία σφαίρα έχει παραμετροποιηθεί με χρήση πολικών συντεταγμένων (παραμετροποίηση  $\varphi$ - $\theta$ ). Έχει χωριστεί δηλαδή σε συγκεκριμένο αριθμό κομματιών, για κάθε ένα από τα οποία υπολογίζεται ξεχωριστά η συνάρτηση αξιολόγησης. Η χρήση της μοναδιαίας σφαίρας βοηθά στο να επιλεγούν σημεία τόσο στο άπειρο όσο και κοντά στο κέντρο της εικόνας.

Η εστιακή απόσταση χρησιμοποιείται για να προβληθούν τα πιθανά σημεία φυγής από την μοναδιαία σφαίρα στο επίπεδο, αλλά δεν υπάρχει καμία ανάγκη για ακριβή γνώση του. Χρήση ενός εστιακού μήκους διαφορετικού από το πραγματικό είναι ισοδύναμο με χρήση μιας σφαίρας διαφορετική από τη μοναδιαία και δεν έχει καμία σημασία στην ακρίβεια των υπολογισμών των σημείων φυγής όταν αυτά εκφράζονται με συντεταγμένες του επιπέδου προβολής.

Αφού υπολογιστεί η συνάρτηση αξιολόγησης για κάθε κομμάτι της μοναδιαίας σφαίρας, τα σημεία φυγής ανιχνεύονται σαν κορυφές στο εξαγόμενο  $\varphi$ - $\theta$  ιστόγραμμα. Οι κορυφές ανιχνεύονται σαν κελιά του ιστογράμματος που παρουσιάζουν τιμές της συνάρτησης αξιολόγησης μεγαλύτερες από ένα κατώφλι και ταυτόχρονα μεγαλύτερες από τις αντίστοιχες τιμές των κελιών στην 8-γειτονιά τους.

Για να επιτύχουμε ένα συνδυασμό ταχύτητας υπολογισμών και ακρίβειας, υλοποιήσαμε μια ιεραρχική προσέγγιση της μεθόδου. Αρχικά η μοναδιαία σφαίρα παραμετροποιείται σε αρκετά μεγάλα κομμάτια, και μετά από την πρώτη εκτίμηση των σημείων φυγής, τα κομμάτια κοντά στα σημεία φυγής διαιρούνται και η διαδικασία υπολογισμού της συνάρτησης αξιολόγησης επαναλαμβάνεται. Η διαδικασία διαίρεσης των κομματιών επαναλαμβάνεται δύο φορές. Κάθε κομμάτι, σε κάθε επανάληψη διαιρείται σε τέσσερα υπό-κομμάτια.

Στην Εικόνα 2-21 (α) φαίνονται με λεπτομέρεια οι τιμές της συνάρτησης αξιολόγησης VPLF για τιμές  $\varphi$ - $\theta$  κοντά στο σημείο φυγής C (Εικόνα 2-13). Οι τιμές της συνάρτησης αξιολόγησης VPLF υπολογίζονται σε ένα «δύο φορές πιο πυκνό» πλέγμα κοντά στο υπολογισμένο σημείο φυγής (β), και υπολογίζεται ξανά μια πιο καλή εκτίμηση του σημείο φυγής. Η διαδικασία επαναλαμβάνεται για μία ακόμα φορά (γ). Τα αριθμητικά αποτελέσματα για τα τρία σημεία φυγής στην Εικόνα 2-13 παρουσιάζονται στον Πίνακα 3-1.



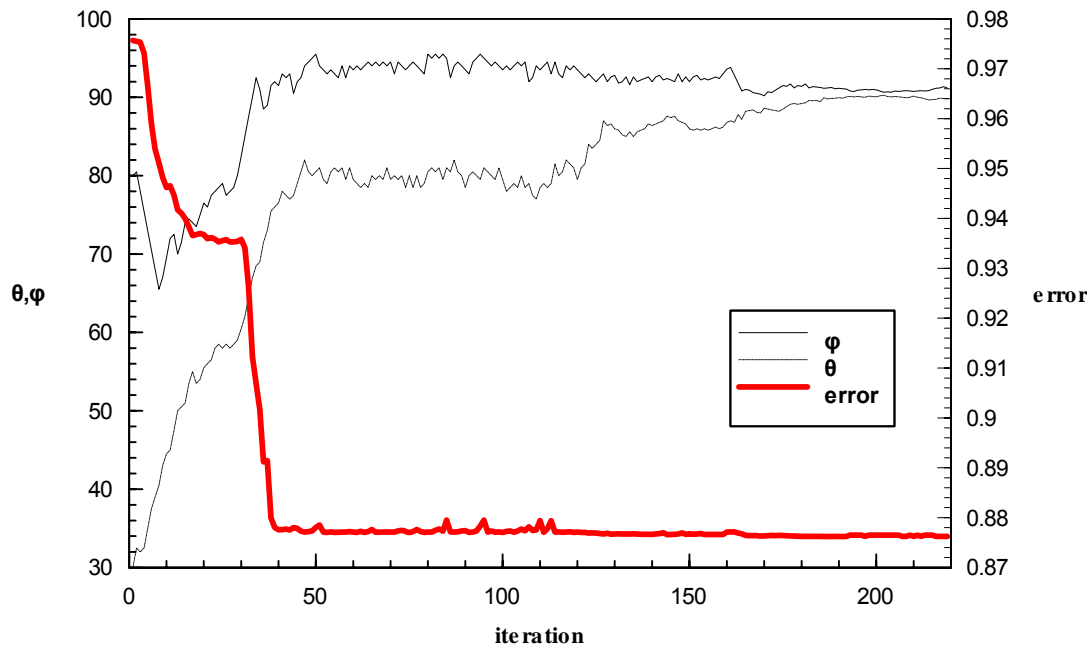
*Εικόνα 2-21. Οι τιμές της VPLF κατά τη διαδικασία διαίρεσης των κελιών για την Εικόνα 2-13.*

**Πίνακας 3-1**

Πραγματικό σημείο φυγής ( $\varphi, \theta$ )	Αρχική εκτίμηση του σημείου φυγής με τη συνάρτηση αξιολόγησης VPLF		Εκτίμηση μετά την πρώτη επανάληψη του αλγορίθμου διαίρεσης των κελιών		Εκτίμηση μετά την δεύτερη επανάληψη του αλγορίθμου διαίρεσης των κελιών	
	Σημείο φυγής ( $\varphi, \theta$ )	Σφάλμα (μοίρες)	Σημείο φυγής ( $\varphi, \theta$ )	Σφάλμα (μοίρες)	Σημείο φυγής ( $\varphi, \theta$ )	Σφάλμα (μοίρες)
(90,90)	(90,88)	2	(90,90)	0	(90,90)	0
(0,90)	(356,88)	4,471	(0,90)	0	(0,90)	0.25
(314.5,64.7)	(315,64)	1.546	(314,64)	0.875	(314,65)	0.517

Αν ψάχνουμε ένα μόνο σημείο φυγής, για το οποίο γνωρίζουμε προσεγγιστικά τη θέση του, μπορεί να χρησιμοποιηθεί και επαναληπτικός αλγόριθμος για την εύρεσή του, σημείου φυγής. Στα πλαίσια αυτής της εργασίας αναπτύχθηκε ένας επαναληπτικός αλγόριθμος τύπου simulated annealing που, σε συνδυασμό με τις συναρτήσεις

αξιολόγησης, παρουσίασε αρκετά καλά αποτελέσματα. Στην Εικόνα 2-22 φαίνεται η γραφική παράσταση της θέσης στη μοναδιαία σφαίρα και του σφάλματος (1-VPLFGRAD) κατά τη διαδικασία εκτέλεσης του επαναληπτικού αλγορίθμου με χρήση της συνάρτησης VPLFGRAD για την Εικόνα 2-1



Εικόνα 2-22. Γραφική παράσταση του σφάλματος και των τιμών  $\varphi$ - $\theta$  του σημείου φυγής στο κέντρο του διαδρόμου στην Εικόνα 2-1 κατά τη διάρκεια σύγκλισης του αλγορίθμου *simulated annealing*. Χρήση της συνάρτησης αξιολόγησης VPLFGRAD.

### 2.3. Συζήτηση

Στο πρώτο μέρος της παρούσας εργασίας μελετήθηκαν και αναπτύχθηκαν μερικοί αλγόριθμοι για την εύρεση σημείων φυγής από δισδιάστατες εικόνες. Η εύρεση σημείων φυγής είναι ένα πολύ δύσκολο πρόβλημα. Αν στην εικόνα δεν υπάρχουν ευδιάκριτες ακμές που να υπερισχύουν των υπόλοιπων χαρακτηριστικών η εύρεση των σημείων φυγής καθίσταται σχεδόν αδύνατη. Αυτό συμβαίνει γιατί όλοι οι αλγόριθμοι που έχουν προταθεί στη βιβλιογραφία, καθώς και όλοι όσοι μελετούνται εδώ, στηρίζονται, με διάφορους τρόπους, είτε στις ακμές της εικόνας είτε στη βαθμίδα της φωτεινότητας της εικόνας.

Κάθε αλγόριθμος από αυτούς που παρουσιάστηκαν παραπάνω έχει τα πλεονεκτήματα και τα μειονεκτήματά του. Ανάλογα με το είδος της εικόνας κάποιοι από τους αλγόριθμους μπορεί να έχουν καλή απόδοση ή να δυσκολεύονται πολύ στο να



υπολογίσουν σωστά και με ακρίβεια τα σημεία φυγής, ή μπορεί ακόμη να είναι και εντελώς ακατάλληλοι

Μελετήθηκαν δύο κατηγορίες αλγορίθμων. Στη πρώτη κατηγορία ανήκουν οι περισσότεροι από τους αλγόριθμους που υπάρχουν στη βιβλιογραφία καθώς και οι παραλλαγές τους που παρουσιάστηκαν στην παράγραφο 2.2.1. Αυτοί οι αλγόριθμοι έχουν ένα κοινό χαρακτηριστικό: θεωρούν δεδομένο ότι από την εικόνα έχει είδη εξαχθεί πληροφορία που αφορά τις δισδιάστατες διευθύνσεις των ευθυγράμμων τμημάτων που σχηματίζονται.

Για την ανίχνευση των διευθύνσεων των ευθυγράμμων τμημάτων χρησιμοποιήθηκε ο μετασχηματισμός Hough για γραμμές. Ο μετασχηματισμός Hough χρησιμοποιήθηκε (α) για να μας δώσει τις δισδιάστατες διευθύνσεις των κυρίαρχων ευθυγράμμων τμημάτων της εικόνας (από τα οποία στην παράγραφο 2.2.1.4 υπολογίσαμε τα σημεία φυγής ως τα σημεία τομής τους), και, (β) για να μας δώσει γενική πληροφορία σχετικά με τις διευθύνσεις που υπάρχουν στην εικόνα. Την πληροφορία αυτή, στην παράγραφο 2.2.1.5, την προβάλαμε στη μοναδιαία σφαίρα, όπως ο Barnard [15], όπου έγινε ο υπολογισμός των σημείων φυγής.

Ο μετασχηματισμός Hough αποδείχτηκε ένα ισχυρό εργαλείο στην προσπάθεια για εύρεση δισδιάστατων κατευθύνσεων. Όμως, παρά τις καλές του επιδόσεις σε ταχύτητα και σε ακρίβεια υπολογισμών, παρουσιάζει μερικά εγγενή χαρακτηριστικά που τον καθιστούν μάλλον δύσχρηστο σε εφαρμογές που απαιτείται χειρισμός χωρίς ανθρώπινη παρέμβαση. Αυτό συμβαίνει γιατί το βέλτιστο μέγεθος του μετασχηματισμού Hough εξαρτάται από την πολυπλοκότητα της εικόνας. Για εικόνες με μεγάλη λεπτομέρεια και πολλά ευθύγραμμα τμήματα που βρίσκονται πολύ κοντά μεταξύ τους είναι απαραίτητο ο πίνακας Hough να έχει πολύ μεγάλη διάσταση ώστε να μπορεί να διακρίνει της μικρές αλλαγές στις διευθύνσεις τους. Αν όμως, η διάσταση γίνει πάρα πολύ μεγάλη δημιουργούνται προβλήματα στην ανίχνευση των κορυφών αφού, ημίτονα που θα έπρεπε να τέμνονται στο ίδιο κελί καταλήγουν να τέμνονται σε παραπλήσια κελιά (αυτό συμβαίνει γιατί, λόγω του κβαντισμού της εικόνας, οι κατευθύνσεις των ευθυγράμμων τμημάτων δε είναι απόλυτα σταθερές).

Στη δεύτερη κατηγορία αλγορίθμων που μελετήθηκαν ανήκουν αυτοί που υπολογίζουν σημεία φυγής απευθείας από την εικόνα, χωρίς να χρειάζονται γνώση σχετικά με τις δισδιάστατες διευθύνσεις των ευθυγράμμων τμημάτων που σχηματίζονται. Οι

αλγόριθμοι αυτοί στηρίζονται σε μια συνάρτηση αξιολόγησης η οποία ορίζεται για κάθε σημείο του επιπέδου προβολής και μας δίνει ένα μέτρο της πιθανότητας το σημείο αυτό να είναι σημείο φυγής. Μελετήθηκαν τρεις διαφορετικές συναρτήσεις αξιολόγησης: (α) αξιολόγηση με υπολογισμό εμβαδού (β) αξιολόγηση με ολοκλήρωση στο χώρο Hough, και (γ) αξιολόγηση με προβολές της βαθμίδας της φωτεινότητας.

Η μέθοδος με αξιολόγηση εμβαδού παρουσίασε κατά πολύ καλύτερα αποτελέσματα από τις άλλες μεθόδους, κυρίως για τους παρακάτω λόγους:

α. Στην περίπτωση της αξιολόγησης με ολοκλήρωση στο χώρο του Hough μεγάλη σημασία στην ποιότητα των αποτελεσμάτων έπαιξε η επιλογή τόσο του μεγέθους του πίνακα Hough όσο και του κατωφλίου το οποίο έπρεπε να υπερβαίνουν οι τιμές του πίνακα Hough για να γίνεται προβολή τους στη μοναδιαία σφαίρα. Οι βέλτιστες τιμές και των δύο αυτών παραμέτρων ήταν συνάρτηση της πολυπλοκότητας της εικόνας και λανθασμένη επιλογή τους οδηγούσε σε μη ικανοποιητικά αποτελέσματα.

β. Για την περίπτωση της αξιολόγησης με προβολές της βαθμίδας της φωτεινότητας, απαιτείται πολύ καλός υπολογισμός της βαθμίδας σε κάθε σημείο ακμής της εικόνας. Οι μάσκες Prewitt που χρησιμοποιήθηκαν δεν έδιναν πολύ καλά αποτελέσματα για σημεία σε ακμές που δεν ήταν απόλυτα κατακόρυφες ή οριζόντιες με αποτέλεσμα τα αποτελέσματα να μην είναι ικανοποιητικά.

Και για τις τρεις διαφορετικές μεθόδους με συνάρτηση αξιολόγησης που μελετήθηκαν, τα υποψήφια σημεία φυγής επιλεγόταν από τη μοναδιαία σφαίρα και προβαλλόταν στο επίπεδο. Αυτό έδινε τη δυνατότητα να επιλέγονται σημεία φυγής τόσο κοντά στο κέντρο της εικόνας όσο και μακριά από αυτό ή στο άπειρο. Για τη προβολή σημείων από τη μοναδιαία σφαίρα στο επίπεδο προβολής είναι απαραίτητη η γνώση της εστιακής απόστασης. Όμως, ακόμα και αν δεν τη γνωρίζουμε ακριβώς, αυτό δεν παίζει καμία σημασία στα τελικά αποτελέσματα. Χρήση μιας εστιακής απόστασης διαφορετικής από τη πραγματική ισοδυναμεί με χρήση μιας σφαίρας διαφορετικής από τη μοναδιαία και δεν έχει καμία σημασία στην ακρίβεια των αποτελεσμάτων, όταν αυτά εκφράζονται σε συντεταγμένες της εικόνας. Η γνώση της πραγματικής εστιακής απόστασης είναι χρήσιμη μόνο όταν θέλουμε να μετατρέψουμε τα σημεία φυγής, όπως αυτά έχουν ανιχνευθεί πάνω στην εικόνα, σε πραγματικές τρισδιάστατες διευθύνσεις.

Ένα μειονέκτημα που παρουσιάζει η χρήση της μοναδιαίας σφαίρας (με τη φ-θ παραμετροποίηση που χρησιμοποιήθηκε), σε σχέση με επιλογή σημείων κατευθείαν από

το επίπεδο προβολής είναι ότι τα κομμάτια στα οποία χωρίζεται δεν έχουν ίσο εμβαδόν. Το μειονέκτημα αυτό είναι πιο εμφανές στις μεθόδους που απαιτούν προβολή μεγίστων κύκλων στη μοναδιαία σφαίρα (οι παραλλαγές της μεθόδου του Barnard που μελετήθηκαν στην παράγραφο 2.2.1.5). Σε αυτές τις μεθόδους είναι σχεδόν αδύνατον να υπολογιστούν σημεία φυγής κοντά στους πόλους της σφαίρας. Οι Quan et al. [20], προτείνουν μια πιο καλή προσέγγιση με χρήση ιεραρχικής παραμετροποίησης της σφαίρας

Στη μέθοδο που χρησιμοποιεί συνάρτηση αξιολόγησης είναι απαραίτητος ο καθορισμός μιας τεχνικής, με βάση την οποία θα επιλέγονται τιμές από τη μοναδιαία σφαίρα και να δοκιμάζονται. Η χρήση κάποιας επαναληπτικής τεχνικής δεν ενδείκνυται για δύο λόγους: α) δεν μπορεί να υλοποιηθεί εύκολα μιας και οι συναρτήσεις αξιολόγησης παρουσιάζουν πολλά τοπικά ακρότατα με αποτέλεσμα να κολλάνε τον αλγόριθμο και β) η χρήση μιας τέτοιας τεχνικής μας περιορίζει στην εύρεση ενός μόνο σημείου φυγής. Αν όπως υπάρχει κάποια γνώση εκ των προτέρων σχετικά με τη θέση των σημείων φυγής, τότε μπορεί να χρησιμοποιηθεί κάποια επαναληπτική τεχνική της μορφής που παρουσιάζεται στη παράγραφο 2.2.2.4, η οποία δεν στηρίζεται στη χρήση παραγώγων της συνάρτησης αξιολόγησης μιας και ο υπολογισμός τους είναι πολύ δύσκολος και δεν μπορεί να γίνει αναλυτικά.

Αντίθετα με την επαναληπτική μέθοδο, η ιεραρχική προσέγγιση στο πρόβλημα της εύρεσης των κορυφών της συνάρτησης αξιολόγησης που μελετήσαμε, παρουσίασε πολύ καλά αποτελέσματα. Παραμετροποιώντας τη σφαίρα σε ολοένα και μικρότερα κομμάτια επιτεύχθηκε ένας πολύ καλός συνδυασμός ταχύτητας και ακρίβειας υπολογισμών.



# 3. ΥΠΟΛΟΓΙΣΜΟΣ ΠΑΡΑΜΕΤΡΩΝ ΕΠΙΠΕΔΟΥ

## 3.1. Διατύπωση του προβλήματος

Το πρόβλημα του υπολογισμού των παραμέτρων ενός επιπέδου είναι ένα πρόβλημα που η επίλυσή του μπορεί να βρει σημαντικές εφαρμογές όπως π.χ. η πλοήγηση ενός ρομποτικού συστήματος σε ένα χώρο με τοίχους, η καθοδήγηση ενός βραχίονα ώστε να τοποθετήσει ένα αντικείμενο σε κάποια συγκεκριμένη θέση (docking) ή η περιστροφή ενός αντικειμένου με βάση οπτική πληροφορία, ώστε να ευθυγραμμιστεί η επιφάνειά του με κάποιο εργαλείο σε μια γραμμή παραγωγής. Όταν λέμε υπολογισμό των παραμέτρων ενός επιπέδου εννοούμε τον υπολογισμό της απόστασής του από κάποιο σημείο (π.χ. την αρχή των αξόνων κάποιου συστήματος συντεταγμένων ή την οπτική εστία της κάμερας) και τη τρισδιάστατη διεύθυνσή του.

Έχοντας υπόψη ότι ένα επίπεδο ορίζεται πλήρως από τρία σημεία, θα μπορούσε κάποιος πολύ εύκολα να σκεφτεί ότι αν μπορούσαμε να βρούμε τις τρισδιάστατες συντεταγμένες τριών συνεπίπεδων σημείων, αυτομάτως θα μπορούσαμε να βρούμε και τις παραμέτρους του επιπέδου πάνω στο οποίο βρίσκονται. Ακόμα όμως και αν υποθέσουμε ότι μπορούμε να βρούμε με κάποιο τρόπο τρία σημεία που θα μπορούσαν να είναι συνεπίπεδα [44] παραμένει το πρόβλημα της εύρεσης των τρισδιάστατων συντεταγμένων τους από τις δισδιάστατες προβολές τους.

Με τον υπολογισμό τρισδιάστατων συντεταγμένων σημείων από δισδιάστατες προβολές (το πρόβλημα της ανακατασκευής) έχουν ασχοληθεί κατά καιρούς πολλοί επιστήμονες και έχουν προταθεί αρκετοί αλγόριθμοι [37,38,39]. Συνήθως, για τη λύση του προβλήματος της εύρεσης τρισδιάστατων συντεταγμένων απαιτούνται δύο τουλάχιστον προβολές των σημείων αυτών (δύο εικόνες στα οποία να απεικονίζονται) και:

- Ακριβής γνώση των θέσεων των σημείων αυτών πάνω και στις δύο εικόνες (αντιστοίχιση σημείων).
- Ακριβής γνώση των παραμέτρων της κάμερας
- Ακριβής γνώση της μεταβολής της θέσης της κάμερας.

Στη πραγματικότητα, η ικανοποίηση οποιασδήποτε από τις τρεις παραπάνω απαιτήσεις συνεπάγεται από μόνη της και την επίλυση ενός ξεχωριστού δύσκολου προβλήματος.

Έστω ότι η παράμετροι της κάμερας είναι γνωστοί ή μπορούν με κάποια μέθοδο [10,14] να εξαχθούν με ικανοποιητική ακρίβεια. Επειδή το πρόβλημα της εύρεσης της κίνησης της κάμερας μπορεί να λυθεί αν ξέρουμε μερικές αντιστοιχίσεις σημείων [37,38,39], ουσιαστικά, οι τρεις παραπάνω απαιτήσεις ,ανάγονται στην απαίτηση για γνώση ενός συγκεκριμένου, μικρού αριθμού αντιστοιχίσεων σημείων από εικόνα σε εικόνα.

Θα μπορούσε να σχεδιαστεί ένας αλγόριθμος ώστε να ψάχνει, με ικανοποιητικά ίσως αποτελέσματα, για τις αντιστοιχίσεις των απαραίτητων σημείων σε εμφανή, αναλλοίωτα από εικόνα σε εικόνα, χαρακτηριστικά, όπως σημεία τομής ακμών ή κορυφές πολυγώνων. Στην πραγματικότητα, όμως, δεν υπάρχει κάποιος γενικός κανόνας ως προς το που να ψάξει κάποιος για σωστές αντιστοιχίσεις. Άλλωστε, η ακρίβεια που χρειαζόμαστε είναι τόσο μεγάλη που, ακόμα και αν ήμασταν σίγουροι ότι μπορούμε να βρούμε σωστές αντιστοιχίσεις, σφάλματα κβαντισμού στις εικόνες μπορεί να οδηγήσουν σε λανθασμένα αποτελέσματα

Συνήθως, είναι προτιμότερο να βρούμε μεγαλύτερο αριθμό αντιστοιχίσεων από όσες χρειαζόμαστε, και, είτε εκτελώντας τον αλγόριθμο πολλές φορές να πάρουμε πολλά αποτελέσματα και να τα αξιολογήσουμε σε ένα δεύτερο στάδιο, είτε, πιο σωστά, να σχεδιάσουμε τον αλγόριθμο έτσι ώστε να τις χρησιμοποιεί όλες. Σε αυτή τη περίπτωση, ακόμα και αν οι αντιστοιχίσεις δεν είναι αρκετά αξιόπιστες, μεγάλο πλήθος από αυτές, στατιστικά, θα οδηγεί σε ικανοποιητικά αποτελέσματα. Συχνά, η οπτική ροή, όπως αυτή υπολογίζεται με βάση τις σχέσεις που δόθηκαν στην παράγραφο 1.3.5, είναι ικανή να δώσει τις αντιστοιχίσεις με αρκετά καλά αποτελέσματα.

Στην περίπτωση του προβλήματος της εύρεσης των παραμέτρων του επιπέδου θα μπορούσαμε να βρούμε πολλές αντιστοιχίσεις σημείων, από αυτές να βρούμε πολλές τρισδιάστατες συντεταγμένες και με πολύ απλό τρόπο να βρούμε το βέλτιστο επίπεδο πάνω στο οποίο θα βρίσκονται.

Στα πλαίσια αυτής της εργασίας παρουσιάζουμε μια κάπως διαφορετική προσέγγιση, η οποία είναι επαναληπτική και στηρίζεται στην ελαχιστοποίησης κάποιας συνάρτησης σφάλματος. Η επαναληπτικότητα έχει το πλεονέκτημα ότι μπορούμε από επανάληψη σε

επανάληψη να μεταβάλουμε, εκτός από τις παραμέτρους του επιπέδου, και τις παραμέτρους κίνησης της κάμερας και έτσι το σφάλμα να εκφράζει τόσο το σφάλμα λόγω λάθους εκτίμησης της κίνησης της κάμερας όσο και το σφάλμα λόγω λάθους εκτίμησης του επιπέδου. Αν στις εικόνες που θεωρούμε ως είσοδο υπάρχει αρκετή πληροφορία για την εκτίμηση όλων των ελεύθερων παραμέτρων, ο αλγόριθμος συγκλίνει δίνοντας αρκετά καλά αποτελέσματα.

Ο παραπάνω τρόπος είναι ουσιαστικά ένας τρόπος πλήρους ανακατασκευής του χώρου μιας και συμπεριλαμβάνει τον υπολογισμό τρισδιάστατων συντεταγμένων. Από τις αντιστοιχίσεις σημείων βρίσκουμε τρισδιάστατες συντεταγμένες και από εκεί βρίσκουμε το σωστό επίπεδο και τις παραμέτρους κίνησης.

Πράττοντας αντίστροφα, έχοντας γνωστές τις παραμέτρους κίνησης και τις παραμέτρους του επιπέδου, μπορούμε, όπως αποδεικνύουμε, να βρούμε τις αντιστοιχίσεις σημείων από εικόνα σε εικόνα. Αν δεν ξέρουμε ακριβώς κάποιες από αυτές τις παραμέτρους αλλά μπορούμε να κάνουμε υποθέσεις για αυτές, μπορούμε εξετάζοντας απλά την ορθότητα των εξεταζόμενων αντιστοιχίσεων να αξιολογήσουμε τις υποθέσεις που κάναμε. Πάνω σε αυτή την αρχή υλοποιήθηκε ένας επαναληπτικός αλγόριθμος που υπολογίζει τις παραμέτρους του επιπέδου χωρίς να χρειάζεται αντιστοιχίσεις σημείων.

Για να συμπληρωθεί η εικόνα πάνω στο αντικείμενο αναλύουμε ένα ακόμα υπάρχοντα αλγόριθμο υπολογισμού των παραμέτρων επιπέδου και κίνησης που κάνει χρήση του πίνακα ροής (flow matrix), ενός πίνακα που περιγράφει το γραμμικό μετασχηματισμό που υφίστανται οι προβολές σημείων πάνω στο φακό της κάμερας καθώς αυτή κινείται.

## 3.2. Πρακτικοί αλγόριθμοι

### 3.2.1. Εξίσωση επιπέδου – παράμετροι

Τρισδιάστατα σημεία (X,Y,Z) που ανήκουν σε ένα επίπεδο ικανοποιούν γενικά μια εξίσωση επιπέδου της μορφής

$$aX + bY + cZ + d = 0$$

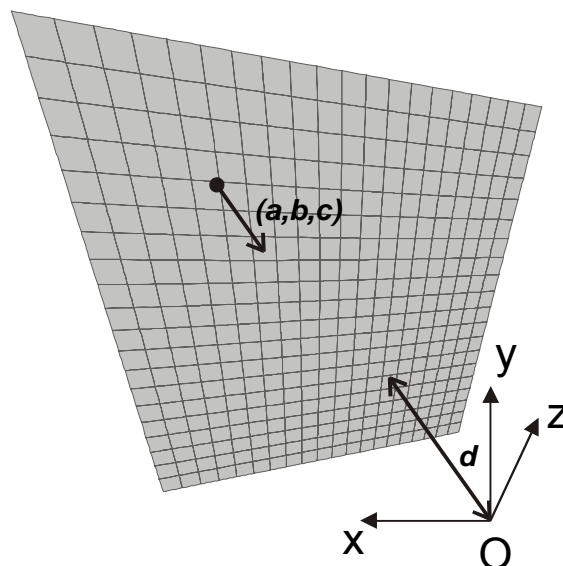
Σχέση 3-1

Αν  $a^2 + b^2 + c^2 = 1$  και  $d > 0$  τότε το διάνυσμα  $[a,b,c]^T$  είναι μοναδιαίο διάνυσμα

κάθετο στο επίπεδο με φορά προς το ημιεπίπεδο που συμπεριλαμβάνει την αρχή των αξόνων (διάνυσμα κατεύθυνσης επιπέδου) και η τιμή της παραμέτρου  $d$  εκφράζει την απόσταση του επιπέδου από την αρχή των αξόνων (Εικόνα 3-1). Μπορούμε πάντα, πολλαπλασιάζοντας με τον κατάλληλο αριθμό, να φέρουμε την εξίσωση του επιπέδου σε αυτή τη μορφή. Το επίπεδο με εξίσωση αυτή στη Σχέση 3-1 θα αναφέρεται σαν επίπεδο  $(a,b,c,d)$ .

Παρακάτω, θεωρούμε πάντα ότι το διάνυσμα  $[a,b,c]^T$  είναι μοναδιαίο και ότι η παράμετρος  $d$  είναι μεγαλύτερη από μηδέν. Όταν κατά τη διάρκεια εκτέλεσης κάποιου αλγόριθμου από αυτούς που παρουσιάζονται γίνεται αλλαγή σε κάποια από τις παραμέτρους  $a,b,c,d$  θεωρούμε ότι οι παράμετροι  $a,b,c,d$  κανονικοποιούνται αμέσως έτσι ώστε να ισχύει αυτή η θεώρηση. Με αυτό το τρόπο, η ζητούμενη εξίσωση του επιπέδου γίνεται μοναδική.

Αν το επίπεδο δεν περνάει από την αρχή των αξόνων, οι παράμετροι είναι ουσιαστικά μόνο τρεις (μπορούμε να διαιρέσουμε με  $d$ ). Αν το επίπεδο περνάει από την αρχή των αξόνων οι παράμετροι είναι πάλι ουσιαστικά μόνο τρεις ( $d=0$ ). Ο αλγόριθμος υπολογισμού του επιπέδου θα μπορούσε να ψάχνει μόνο για τρεις πραγματικές παραμέτρους και μια λογική (αν περνάει από την αρχή των αξόνων ή όχι). Για λόγους ευκολίας θεωρούμε τη γενική εξίσωση στις υλοποιήσεις που ακολουθούν.



Εικόνα 3-1. Το επίπεδο που εκφράζεται από την εξίσωση  $aX+bY+cZ+d=0$  αν  $a^2+\beta^2+c^2=1$  και  $d>0$



Μπορούμε πάντα, πολλαπλασιάζοντας με τον κατάλληλο αριθμό, να φέρουμε την εξίσωση του επιπέδου σε αυτή τη μορφή.

### 3.2.2. Εύρεση παραμέτρων επιπέδου ώστε να συμπίπτουν δύο εικόνες.

#### 3.2.2.1. Αντιστοιχίσεις σημείων από μία κάμερα σε μία άλλη

Έστω ότι έχουμε δύο κάμερες για τις οποίες είναι γνωστή η σχετική θέση της μίας ως προς την άλλη. Έστω επίσης ότι  $O^{t_1} = (O_x^{t_1}, O_y^{t_1}, O_z^{t_1})$  είναι η θέση της οπτικής εστίας και  $\bar{x}^{t_1}, \bar{y}^{t_1}, \bar{z}^{t_1}$  είναι τα μοναδιαία διανύσματα του συστήματος συντεταγμένων  $O^{t_1}x^{t_1}y^{t_1}z^{t_1}$  της δεύτερης κάμερας ως προς το σύστημα συντεταγμένων  $O^{t_0}x^{t_0}y^{t_0}z^{t_0}$  της πρώτης κάμερας.

Κάθε σημείο  $P$  που έχει συντεταγμένες  $(X^{t_0}, Y^{t_0}, Z^{t_0})$  στο σύστημα  $O^{t_0}x^{t_0}y^{t_0}z^{t_0}$  της πρώτης κάμερας θα έχει συντεταγμένες  $(X^{t_1}, Y^{t_1}, Z^{t_1})$  στο σύστημα  $O^{t_1}x^{t_1}y^{t_1}z^{t_1}$  της δεύτερης κάμερας που η τιμή τους θα δίνεται από τη σχέση:

$$\begin{bmatrix} X^{t_1} \\ Y^{t_1} \\ Z^{t_1} \end{bmatrix} = - \begin{bmatrix} O_x^{t_1} \\ O_y^{t_1} \\ O_z^{t_1} \end{bmatrix} + R^{t_0 \rightarrow t_1} \begin{bmatrix} X^{t_0} \\ Y^{t_0} \\ Z^{t_0} \end{bmatrix} \quad \text{Σχέση 3-2}$$

Έστω ότι το σημείο  $P$  ανήκει σε ένα επίπεδο  $(a,b,c,d)$  στο σύστημα αξόνων  $O^{t_0}x^{t_0}y^{t_0}z^{t_0}$ . Από την εξίσωση του επιπέδου και τις εξισώσεις της προοπτικής προβολής (Σχέση 1-1 και Σχέση 1-2) έχουμε ότι για κάθε σημείο με συντεταγμένες  $(x^{t_0}, y^{t_0})$  στο επίπεδο προβολής της πρώτης κάμερας, οι τριςδιάστατες συντεταγμένες του θα είναι:

$$\begin{bmatrix} X^{t_0} \\ Y^{t_0} \\ Z^{t_0} \end{bmatrix} = \frac{-1}{ax^{t_0} + by^{t_0} + cf} \begin{bmatrix} x^{t_0} \\ y^{t_0} \\ d \end{bmatrix} \quad \text{Σχέση 3-3}$$

Αντικαθιστώντας στη Σχέση 1-33 έχουμε τις τριςδιάστατες συντεταγμένες του ως προς το σύστημα της δεύτερης εικόνας:

$$\begin{bmatrix} X^{t_1} \\ Y^{t_1} \\ Z^{t_1} \end{bmatrix} = - \begin{bmatrix} O_x^{t_1} \\ O_y^{t_1} \\ O_z^{t_1} \end{bmatrix} - \frac{I}{ax^{t_0} + by^{t_0} + cf} R^{t_0 \rightarrow t_1} \begin{bmatrix} x^{t_0} \\ y^{t_0} \\ d \end{bmatrix} \quad \text{Σχέση 3-4}$$

Μπορούμε εύκολα να αντικαταστήσουμε τις συντεταγμένες  $(X^{t_1}, Y^{t_1}, Z^{t_1})$  που μας δίνει η Σχέση 3-4 στις εξισώσεις της προοπτικής προβολής (Σχέση 1-1 και Σχέση 1-2) και να βρούμε τις συντεταγμένες  $(x^{t_1}, y^{t_1})$  στο επίπεδο προβολής της δεύτερης κάμερας.

Δηλαδή, όταν γνωρίζουμε τη σχετική θέση μίας κάμερας ως προς μία άλλη και την εξίσωση ενός επιπέδου, σημεία από το οποίο απεικονίζονται και στις δύο κάμερες, μπορούμε να υπολογίζουμε τις αντιστοιχίες κοινών σημείων από τη μια κάμερα στην άλλη. Υπολογίζοντας τις αντιστοιχίες σημείων μπορούμε να κατασκευάσουμε την εικόνα ενός επιπέδου όπως αυτή θα φαινόταν αν βρισκόμασταν σε μια άλλη γνωστή θέση.

Στην Εικόνα 3-2 φαίνονται πέντε διαφορετικές εικόνες ενός τοίχου. Οι εικόνες προκύπτουν από τις διαδοχικές θέσεις στις οποίες μεταβαίνει μια κάμερα που έχει αριστερά της ένα τοίχο. Αρχικά η κάμερα βρίσκεται σε απόσταση 200 μονάδων από τον τοίχο και σχηματίζει με αυτόν γωνία  $45^\circ$ . Ο τοίχος έχει εξίσωση  $x=-50$  και η κάμερα βρίσκεται στη θέση  $(150,0,0)$  ως προς το γενικό σύστημα συντεταγμένων.

Αρχικά η κάμερα βρίσκεται στη θέση  $(150,0,)$   
σε απόσταση 200 μονάδων από το επίπεδο  
που έχει εξίσωση  $x=-50$ .

Η γωνία που σχηματίζει με αυτό είναι  $45^\circ$ .

Η εστιακή απόσταση είναι  $f=10$ .

Κάθε εικόνα έχει διαστάσεις  $256 \times 256$ .

Ο λόγος εικονοστοιχείων ανά μονάδα μήκους  
είναι 35,3.



α. Αρχική εικόνα



β. Η κάμερα κινείται μεταφορικά, στη  
διεύθυνση του οπτικού της άξονα κατά 3  
μονάδες



γ. Η κάμερα περιστρέφεται γύρω από τον  $y$   
άξονά της κατά  $3^\circ$

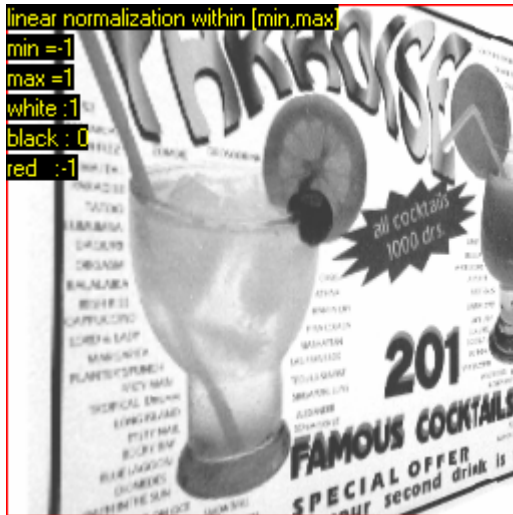


δ. Η κάμερα κινείται μεταφορικά, στη  
διεύθυνση του οπτικού της άξονα κατά 3  
μονάδες



ε. Η κάμερα περιστρέφεται γύρω από τον  $y$   
άξονά της κατά  $3^\circ$

Εικόνα 3-2. Ένα επίπεδο όπως φαίνεται από έξι διαφορετικές θέσεις.



Η εικόνα στη θέση α όπως υπολογίστηκε από την εικόνα στη θέση β. Η εικόνα στη θέση ε όπως υπολογίστηκε από την εικόνα στη θέση ε.

Εικόνα 3-3. Εφαρμογή του αλγορίθμου υπολογισμού μιας εικόνας από μία άλλη.

### 3.2.2.2. Εύρεση παραμέτρων επιπέδου

Έστω ότι δεν ξέρουμε τις παραμέτρους του επιπέδου, αλλά μπορούμε να κάνουμε υποθέσεις για αυτές. Ένας εύκολος τρόπος να αξιολογήσουμε τις υποθέσεις που κάνουμε για το επίπεδο είναι να βρούμε τις αντιστοιχίες των σημείων, ανάμεσα σε γνωστές εικόνες, σύμφωνα με τις σχέσεις που περιγράφηκαν στη προηγούμενη παράγραφο και να ελέγξουμε αν οι αντιστοιχίες αυτές είναι σωστές ή όχι.

Στηριζόμενοι στο παραπάνω συμπέρασμα υλοποιήσαμε έναν επαναληπτικό αλγόριθμο παραλλαγή του αλγορίθμου simulated annealing. Ο αλγόριθμος στηρίζεται στα εξής βήματα.

1. Αρχικοποίηση των παραμέτρων  $a$ ,  $b$ ,  $c$ ,  $d$  του επιπέδου καθώς και των άγνωστων παραμέτρων κίνησης της κάμερας (τυχαία ή με βάση κάποια προηγούμενη γνώση)
2. Υπολογισμός των αντιστοιχιών σημείων όπως αυτές προκύπτουν από τις σχέσεις στην παράγραφο 3.2.2.1 για τις παραμέτρους  $a$ ,  $b$ ,  $c$ ,  $d$ .

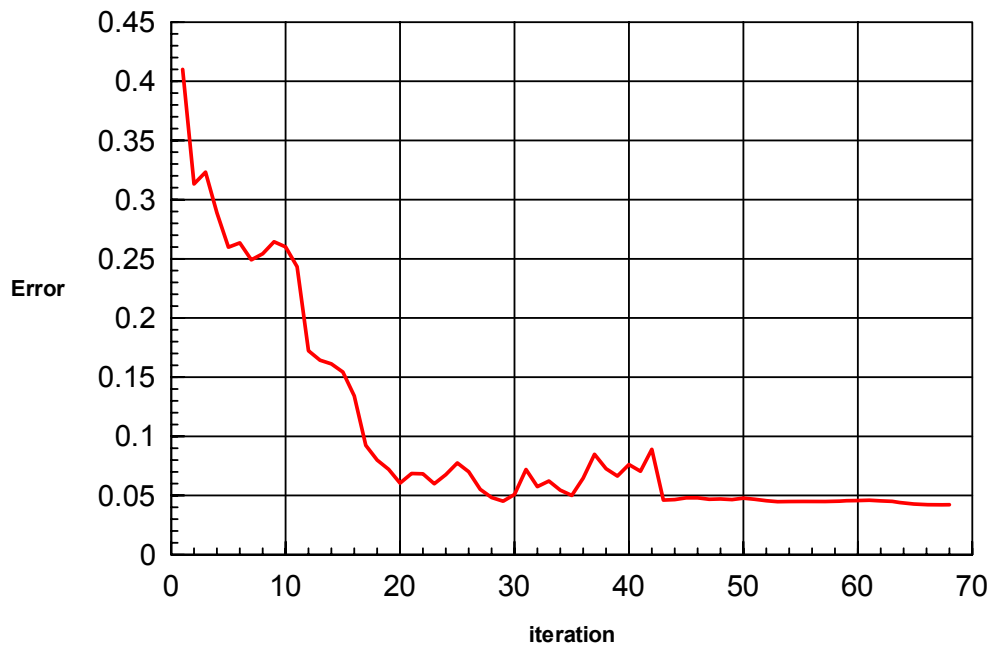
3. Αξιολόγηση των αντιστοιχιών και εξαγωγή του σφάλματος<sup>1</sup>
4. Με βάση το σφάλμα αλλαγή κάθε φορά μιας από τις παραμέτρους  $a$ ,  $b$ ,  $c$ ,  $d$  του επιπέδου ή μιας από τις άγνωστες παραμέτρους κίνησης προς τη σωστή κατεύθυνση.
5. Επανάληψη του αλγορίθμου από το βήμα 2 μέχρι να ικανοποιηθεί το κριτήριο λήξης του αλγορίθμου (να μην υπάρχουν πια μεγάλες αλλαγές στο σφάλμα).

Η αλλαγή των τιμών  $a$ ,  $b$ ,  $c$ ,  $d$ , καθώς και των άγνωστων παραμέτρων κίνησης γίνεται με ένα σταθερό βήμα προς τα πάνω ή προς τα κάτω με μια πιθανότητα που προκύπτει ανάλογα με την προηγούμενη αλλαγή που έγινε και το πόσο επιτυχής ήταν. Σταδιακά, μειώνεται μια παράμετρος θερμοκρασίας και οι αλλαγές τείνουν να γίνονται όλο και πιο τυχαίες (να έχουν μέση τιμή όλο και πιο κοντά στο μηδέν), μέχρι τον τερματισμό του αλγορίθμου όταν η μέση τιμή των αλλαγών γίνει πολύ μικρή.

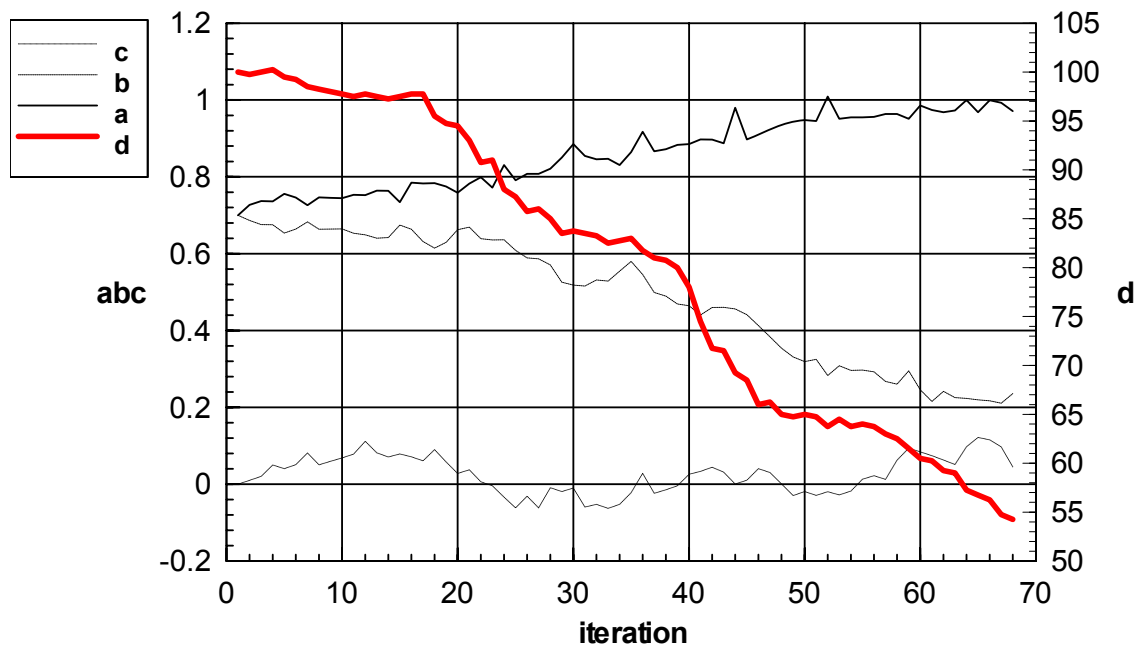
Στην Εικόνα 3-4 φαίνεται το σφάλμα του κατά τη διάρκεια της σύγκλισης του αλγορίθμου όταν ζητήθηκε να βρει το επίπεδο που φαίνεται στις τεχνητές σκηνές στην Εικόνα 3-2. Οι τιμές των παραμέτρων  $a$ ,  $b$ ,  $c$ ,  $d$  του επιπέδου κατά τη διάρκεια της σύγκλισης φαίνονται στην Εικόνα 3-5. Ο αλγόριθμος μετά από 70 περίπου επαναλήψεις κατάφερε, με ικανοποιητική ακρίβεια να συγκλίνει κοντά στις σωστές τιμές ( $a=1$ ,  $b=0$ ,  $c=0$ ,  $d=200$ ). Οι παράμετροι μεταφορικής κίνησης της κάμερας από την προβολή  $\alpha$  στη προβολή  $\beta$  φαίνονται στην Εικόνα 3-6.

---

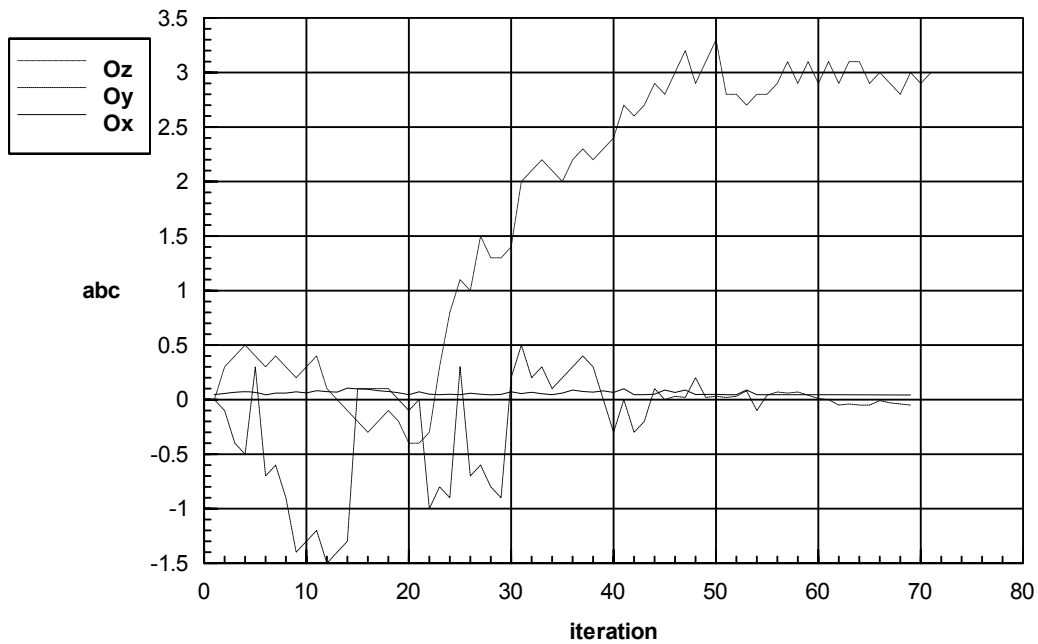
<sup>1</sup> Για την αξιολόγηση των αντιστοιχιών, αυτές εφαρμόζονται σε μία εικόνα και στη συνέχεια υπολογίζεται, με τον τρόπο που περιγράφηκε στη παράγραφο 3.2.2.1, μια από τις άλλες εικόνες. Η εικόνα που προκύπτει συγκρίνεται με την πραγματική και το ποσοστό λάθους που προκύπτει είναι το σφάλμα.



Εικόνα 3-4. Το σφάλμα κατά τη διάρκεια της σύγκλισης του αλγορίθμου.



Εικόνα 3-5. Οι παράμετροι  $a$ ,  $b$ ,  $c$ ,  $d$  του επιπέδου κατά τη διάρκεια της σύγκλισης.



Εικόνα 3-6. Η μεταφορική κίνηση της κάμερας από τη θέση α στη θέση β κατά τη διάρκεια της σύγκλισης.

### 3.2.3. Εύρεση παραμέτρων επιπέδου με υπολογισμό «ανάμεσα σε δύο επίπεδα»

Έστω ότι έχουμε δύο εικόνες που προέρχονται από δύο όμοιες κάμερες με συστήματα αξόνων  $O^{t_0} x^{t_0} y^{t_0} z^{t_0}$  και  $O^{t_1} x^{t_1} y^{t_1} z^{t_1}$ . Έστω επίσης ότι ένα σημείο με συντεταγμένες στο σύστημα αξόνων της πρώτης κάμερας  $(X^{t_0}, Y^{t_0}, Z^{t_0})$  και  $(X^{t_1}, Y^{t_1}, Z^{t_1})$  στο σύστημα αξόνων της πρώτης και της δεύτερης κάμερας αντίστοιχα.

Οι προβολή του σημείου αυτού στην εικόνα της πρώτης κάμερας θα έχει συντεταγμένες εικόνας:

$$(x^{t_0}, y^{t_0}) = \left( f \frac{X^{t_0}}{Z^{t_0}}, f \frac{Y^{t_0}}{Z^{t_0}} \right) \quad \text{Σχέση 3-5}$$

για τη δεύτερη κάμερα οι συντεταγμένες εικόνας θα είναι:

$$(x^{t_1}, y^{t_1}) = \left( f \frac{X^{t_1}}{Z^{t_1}}, f \frac{Y^{t_1}}{Z^{t_1}} \right) \quad \text{Σχέση 3-6}$$

Λύνοντας ως προς  $X^{t_0}, Y^{t_0}, X^{t_1}, Y^{t_1}$  και αντικαθιστώντας στην Σχέση 3-2 έχουμε:

$$\begin{bmatrix} x^{t_1} \frac{Z^{t_1}}{f} \\ y^{t_1} \frac{Z^{t_1}}{f} \\ Z^{t_1} \end{bmatrix} = - \begin{bmatrix} O_x^{t_1} \\ O_y^{t_1} \\ O_z^{t_1} \end{bmatrix} + R^{t_0 \rightarrow t_1} \begin{bmatrix} x^{t_0} \frac{Z^{t_0}}{f} \\ y^{t_0} \frac{Z^{t_0}}{f} \\ Z^{t_0} \end{bmatrix} \quad \text{Σχέση 3-7}$$

Το παραπάνω γραμμικό σύστημα μπορεί να μας δώσει δύο ζευγάρια λύσεων ως προς  $Z^{t_0}$  και  $Z^{t_1}$ . Το πρώτο ζευγάρι ουσιαστικά προκύπτει από την επίλυση με βάση τις οριζόντιες συντεταγμένες των αντιστοιχίσεων και το δεύτερο με βάση με τις κατακόρυφες. Ανάλογα αν η μετατόπιση (η διαφορά στις συντεταγμένες εικόνας των αντιστοιχισμένων σημείων) είναι μεγαλύτερη στον οριζόντιο ή στον κατακόρυφο άξονα επιλέγουμε το κατάλληλο ζευγάρι (θα δώσει μεγαλύτερη ακρίβεια).

Αντικαθιστώντας την τιμή  $Z^{t_0}$  στη Σχέση 3-5 βρίσκουμε και τις άλλες δύο συντεταγμένες  $X^{t_0}$ ,  $Y^{t_0}$

Έχοντας τις τρισδιάστατες συντεταγμένες για τα σημεία του επιπέδου μπορούμε να βρούμε το επίπεδο που ταιριάζει περισσότερο σε αυτά. Για την εύρεση των παραμέτρων του επιπέδου χρησιμοποιούμε ένα αλγόριθμο που αποτελείται από τα παρακάτω βήματα

1. Αρχικοποίηση των παραμέτρων  $a$ ,  $b$ ,  $c$ ,  $d$  του επιπέδου (τυχαία ή με βάση κάποια προηγούμενη γνώση)
2. Αρχικοποίηση των άγνωστων παραμέτρων κίνησης της κάμερας.
3. Αρχικοποίηση μιας παραμέτρου  $dd$  με μία μεγάλη τιμή.
4. Υπολογισμός του σφάλματος σαν το ποσοστό των σημείων  $(X^{t_0}, Y^{t_0}, Z^{t_0})$  τα οποία δε βρίσκονται ανάμεσα στα επίπεδα  $(a, b, c, d - dd/2)$  και  $(a, b, c, d + dd/2)$ . Για τα σημεία αυτά θα ισχύει:

$$|aX^{t_0} + bY^{t_0} + cZ^{t_0} + d| \geq dd/2 \quad \text{Σχέση 3-8}$$

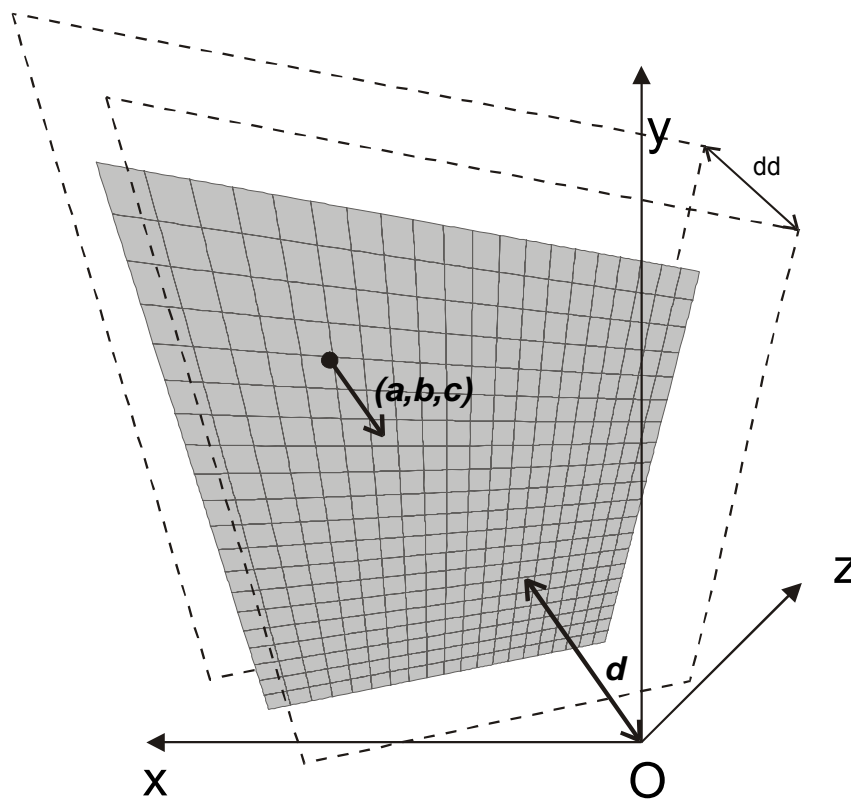
5. Αν το σφάλμα είναι κάτω από ένα κατώφλι μείωσε την παράμετρο  $dd$  πολλαπλασιάζοντάς τη με ένα σταθερό συντελεστή (μικρότερο της μονάδας).
6. Αν το σφάλμα είναι πάνω από ένα κατώφλι αύξησε την παράμετρο  $dd$  πολλαπλασιάζοντάς τη με ένα σταθερό συντελεστή (μεγαλύτερο της μονάδας)



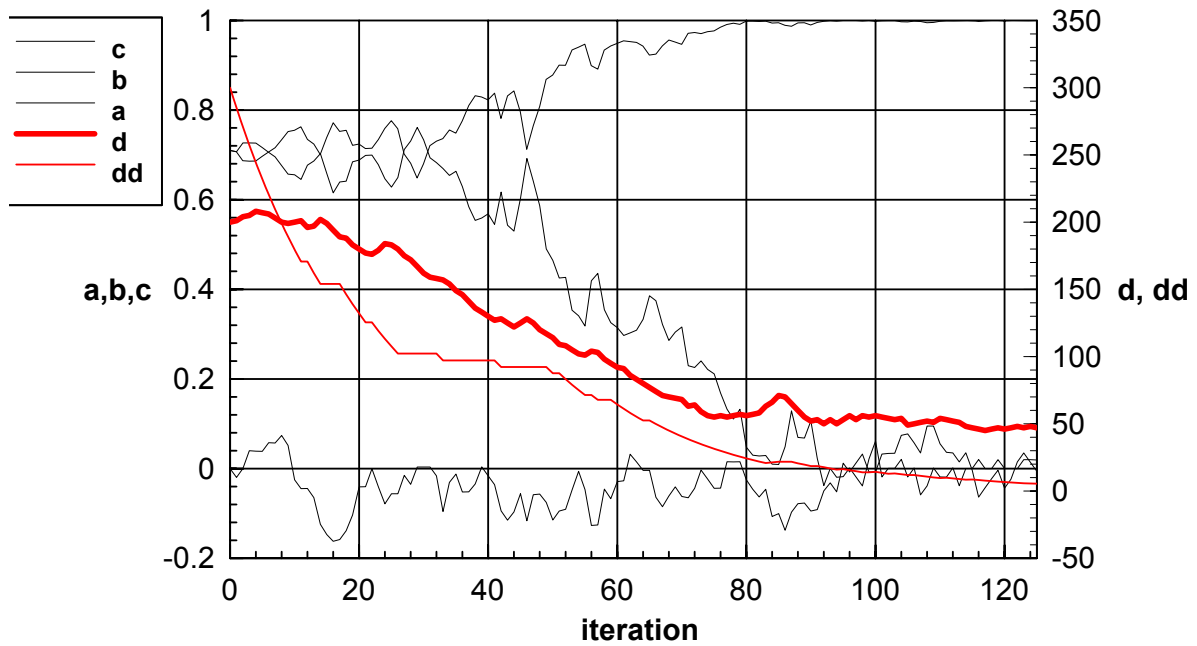
7. Αν το σφάλμα είναι μικρότερο από ένα κατώφλι και η παράμετρος  $dd$  είναι επίσης μικρότερη από ένα κατώφλι, έξοδος από τον αλγόριθμο.
8. Με βάση το σφάλμα αλλαγή μίας από τις παραμέτρους  $a, b, c, d$  του επιπέδου ή μιας από τις άγνωστες παραμέτρους κίνησης της κάμερας προς τη σωστή κατεύθυνση και επανάληψη του αλγορίθμου από το βήμα 4.

Η αλλαγή των τιμών  $a, b, c, d$  καθώς και των άγνωστων παραμέτρων κίνησης της κάμερας στο βήμα 8 γίνεται με ένα σταθερό βήμα προς τα πάνω ή προς τα κάτω με μια πιθανότητα που προκύπτει ανάλογα με την προηγούμενη αλλαγή που έγινε και το πόσο επιτυχής ήταν.

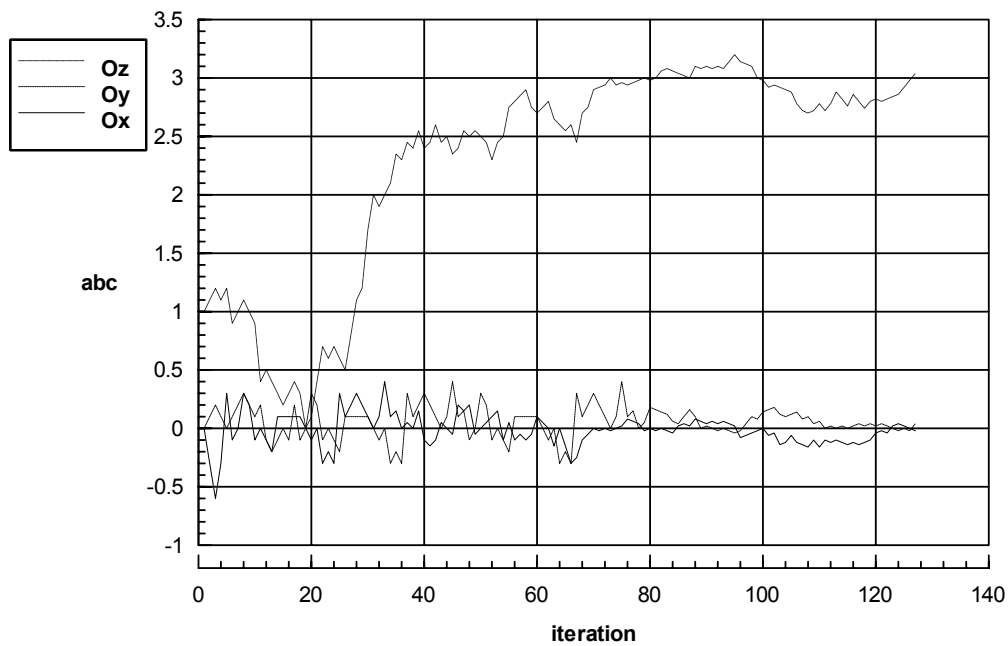
Στην Εικόνα 3-8 φαίνονται τιμές των παραμέτρων  $a, b, c, d$  του επιπέδου, καθώς και η τιμή της παραμέτρου  $dd$ , κατά τη διάρκεια της σύγκλισης του αλγορίθμου, όταν ζητήθηκε να βρει το επίπεδο που φαίνεται στις τεχνητές σκηνές στην Εικόνα 3-2. Ο αλγόριθμος, μετά από 130 περίπου επαναλήψεις, κατάφερε, με ικανοποιητική ακρίβεια να συγκλίνει στις σωστές τιμές ( $a=1, b=0, c=0, d=50$ ). Οι παράμετροι μεταφορικής κίνησης της κάμερας από την προβολή  $\alpha$  στη προβολή  $\beta$  φαίνονται στην Εικόνα 3-9.



Εικόνα 3-7. Το ζητούμενο επίπεδο βρίσκεται ανάμεσα σε δύο παράλληλα με αυτό επίπεδα που απέχουν μεταξύ τους κατά  $dd$ .



Εικόνα 3-8. Οι παράμετροι  $a, b, c, d$  του επιπέδου και η απόσταση  $dd$  κατά τη διάρκεια της σύγκλισης του αλγορίθμου.



Εικόνα 3-9. Η μεταφορική κίνηση της κάμερας από τη θέση  $\alpha$  στη θέση  $\beta$  κατά τη διάρκεια της σύγκλισης.

### 3.2.4. Εύρεση με χρήση του πίνακα ροής

Όταν μία κάμερα κινείται στο χώρο σημεία που βρίσκονται πάνω σε ένα επίπεδο φαίνονται να κινούνται πάνω στο επίπεδο προβολής της κάμερας. Ο μετασχηματισμός

αυτός των σημείων είναι τέτοιος ώστε: (i) συνευθειακά σημεία παραμένουν συνευθειακά (ii) συντρέχουσες γραμμές (concurrent lines) παραμένουν συντρέχουσες. (iii) οι τομές (incidences) διατηρούνται. Άρα, ο μετασχηματισμός της εικόνας μπορεί να χαρακτηριστεί σαν collineation και έχει τη μορφή:

$$\begin{bmatrix} x^{t_1} \\ y^{t_1} \\ f \end{bmatrix} = \pm N \begin{bmatrix} A^T & \\ & \end{bmatrix} \begin{bmatrix} x^{t_0} \\ y^{t_0} \\ f \end{bmatrix} \quad \text{Σχέση 3-9}$$

όπου  $A$  είναι ένας  $3 \times 3$  πίνακας και  $(x^{t_0}, y^{t_0}, f)$  και  $(x^{t_1}, y^{t_1}, f)$  είναι συντεταγμένες πάνω στην εικόνα αντιστοιχισμένων σημείων τη χρονική στιγμή  $t=t_0$  και  $t=t_1$  αντίστοιχα.

Από τη παραπάνω σχέση φαίνεται ότι ο πίνακας  $A$  μπορεί να υπολογιστεί από αντιστοιχίσεις σημείων. Επειδή μπορεί να υπολογιστεί μόνο ως παράμετρος μίας σταθεράς δεχόμαστε την παραδοχή ότι  $\det A = 1$ .

Κάθε χρονική στιγμή ορίζεται ο πίνακας ροής  $W$  ως:

$$W = \lim_{\Delta t \rightarrow 0} \frac{A(\Delta t) - I}{\Delta t} \quad \text{Σχέση 3-10}$$

Η ροή πάνω στην εικόνα μπορεί να υπολογιστεί από τον πίνακα ροής [3] ως:

$$\begin{bmatrix} u \\ v \\ f \end{bmatrix} = W^T \begin{bmatrix} x \\ y \\ f \end{bmatrix} - \left( \begin{bmatrix} x \\ y \\ f \end{bmatrix} \times W^T \begin{bmatrix} x \\ y \\ f \end{bmatrix} \right) \begin{bmatrix} x \\ y \\ f \end{bmatrix} \quad \text{Σχέση 3-11}$$

### 3.2.4.1. Εύρεση του flow matrix

Αν θέσουμε:

$$A = W_{11} - W_{33} \quad D = W_{22} - W_{33} \quad U = W_{31}$$

$$B = W_{21} \quad E = -W_{13} \quad V = W_{32}$$

$$C = W_{12} \quad F = -W_{23}$$

και αντικαταστήσουμε στη Σχέση 3-11 έχουμε μετά από πράξεις:

$$u = fU + Ax + By + \frac{1}{f}(Ex + Fy)x \quad \text{Σχέση 3-12}$$

$$v = fV + Cx + Dy + \frac{1}{f}(Ex + Fy)y \quad \text{Σχέση 3-13}$$

Αν ξέρουμε τη ροή σε τέσσερα σημεία της εικόνας μπορούμε, λύνοντας το γραμμικό σύστημα των 8 εξισώσεων που θα προκύψει από τις παραπάνω εξισώσεις, να βρούμε τα  $A, B, C, D, E, F, U, V$  και από αυτά τον πίνακα  $W$  ως:

$$W = \begin{bmatrix} 2A - D & C & -E \\ B & (-A + 2D)/3 & -F \\ U & V & -(A + D) \end{bmatrix} \quad \text{Σχέση 3-14}$$

Αν ξέρουμε τη ροή σε παραπάνω από τέσσερα σημεία, έστω  $N$  τότε μπορούμε να βρούμε το βέλτιστο πίνακα  $W$  που τα περιγράφει χρησιμοποιώντας τη μέθοδο των ελαχίστων τετραγώνων. Αρκεί σε αυτή τη περίπτωση να ελαχιστοποιήσουμε το άθροισμα:

$$J = \sum_{n=1}^N \left( \left( fU + Ax_n + By_n + \frac{1}{f}(Ex_n + Fy_n)x_n - v_n \right)^2 + \left( fV + Cx_n + Dy_n + \frac{1}{f}(Ex_n + Fy_n)y_n - u_n \right)^2 \right) \quad \text{Σχέση 3-15}$$

Λύνοντας τις εξισώσεις:

$$\frac{\partial J}{\partial A} = 0 \quad \frac{\partial J}{\partial D} = 0 \quad \frac{\partial J}{\partial U} = 0$$

$$\frac{\partial J}{\partial B} = 0 \quad \frac{\partial J}{\partial E} = 0 \quad \frac{\partial J}{\partial V} = 0$$

$$\frac{\partial J}{\partial C} = 0 \quad \frac{\partial J}{\partial F} = 0$$

καταλήγουμε στο παρακάτω σύστημα 8 γραμμικών εξισώσεων:

$$fU = \bar{x}A + \bar{y}B + \frac{1}{f}(\bar{x}^2E + \bar{xy}F) = \bar{u} \quad \text{Σχέση 3-16}$$

$$fV = \bar{x}C + \bar{y}D + \frac{1}{f}(\bar{xy}E + \bar{y}^2F) = \bar{v} \quad \begin{array}{l} \text{Σχέση} \\ 3-17 \end{array}$$

$$fxU + x^2A + xyB + \frac{1}{f}(x^3E + x^2yF) = ux \quad \begin{array}{l} \text{Σχέση} \\ 3-18 \end{array}$$

$$f\bar{y}U + \bar{xy}A + \bar{y}^2B + \frac{1}{f}(\bar{x}^2\bar{y}E + \bar{xy}^2F) = \bar{u}\bar{y} \quad \begin{array}{l} \text{Σχέση} \\ 3-19 \end{array}$$

$$f\bar{x}V + \bar{x}^2C + \bar{xy}D + \frac{1}{f}(\bar{x}^2\bar{y}E + \bar{xy}^2F) = \bar{v}\bar{x} \quad \begin{array}{l} \text{Σχέση} \\ 3-20 \end{array}$$

$$f\bar{y}V + \bar{xy}A + \bar{y}^2B + \frac{1}{f}(\bar{x}^3E + \bar{x}^2\bar{y}F) = \bar{u}\bar{x} \quad \begin{array}{l} \text{Σχέση} \\ 3-21 \end{array}$$

$$\bar{x}^2U + \bar{xy}V + \frac{1}{f}(\bar{x}^3A + \bar{x}^2\bar{y}(B+C) + \bar{xy}^2D) + \frac{1}{f^2}(\bar{x}^4 + \bar{x}^2\bar{y}^2E + \bar{x}^3\bar{y} + \bar{xy}^3F) = \frac{1}{f}\bar{u}\bar{x}^2 + \bar{v}\bar{xy} \quad \begin{array}{l} \text{Σχέση} \\ 3-22 \end{array}$$

$$\bar{xy}U + \bar{y}^2V + \frac{1}{f}(\bar{xy}^2A + \bar{xy}^2(B+C) + \bar{y}^3D) + \frac{1}{f^2}(\bar{x}^3\bar{y} + \bar{xy}^3E + \bar{x}^2\bar{y}^2 + \bar{y}^4F) = \frac{1}{f}\bar{u}\bar{xy} + \bar{v}\bar{y}^2 \quad \begin{array}{l} \text{Σχέση} \\ 3-23 \end{array}$$

Λύνοντας το παραπάνω σύστημα ως προς  $A, B, C, D, E, F, U, V$  και αντικαθιστώντας στη Σχέση 3-14 έχουμε τον πίνακα ροής  $W$

### 3.2.4.2. Εύρεση των παραμέτρων κίνησης και των παραμέτρων επιπέδου

Έστω η κάμερα κινείται στο χώρο με περιστροφική και μεταφορική ταχύτητα  $\vec{\omega}$  και  $\vec{v}$  αντίστοιχα.

Αν  $\|\vec{v}\| = l$  (ή  $\|\vec{v}\| = 0$ ) τότε μπορούμε να υπολογίζουμε τις παραμέτρους κίνησης της κάμερας και τις παραμέτρους  $a, b, c, d$  από τον πίνακα ροής  $W$ , σύμφωνα με την ακόλουθη διαδικασία

- Υπολογίζουμε τους πίνακες  $W_\sigma$  και  $W_\alpha$  και το διάνυσμα  $\vec{w}$  έτσι ώστε:

$$W_\sigma = \frac{1}{2}(W + W^T) \quad \text{Σχέση 3-24}$$

$$W_\alpha = \frac{1}{2}(W - W^T) \quad \text{Σχέση 3-25}$$

$$W_\alpha = \overset{P}{W} \times I \quad \text{Σχέση 3-26}$$

Το διάνυσμα  $\overset{P}{W}$  θα είναι το διάνυσμα  $[W_{\alpha 32}, W_{\alpha 13}, W_{\alpha 21}]^T$ , το διάνυσμα δηλαδή που έχει σαν στοιχεία τα μη μηδενικά στοιχεία του αντισυμμετρικού πίνακα  $W_\alpha$ .

- Αν  $W_\sigma = O$  τότε η μεταφορική ταχύτητα  $\overset{P}{h}$  θα είναι μηδέν, για την περιστροφική ταχύτητα θα είναι  $\overset{P}{\omega} = \overset{P}{W}$  και οι παράμετροι του επιπέδου δεν μπορούν να υπολογιστούν.
- Αν  $W_\sigma \neq O$ ,  $\sigma_1, \sigma_2, \sigma_3$ , είναι οι ιδιοτιμές του  $W_\sigma$  με  $\sigma_1 \geq \sigma_2 \geq \sigma_3$  και  $\overset{P}{u}_1, \overset{P}{u}_2, \overset{P}{u}_3$  τα αντίστοιχα ιδιοδιανύσματα τότε:

Το διάνυσμα  $\overset{P}{p}$  που δείχνει κάθετα από την αρχή των αξόνων στο επίπεδο  $(a, b, c, d)$  θα είναι:

$$\overset{P}{p} = \sqrt{\sigma_1 - \sigma_3} (\pm \sqrt{\sigma_1 - \sigma_2} \overset{P}{u}_1 + \sqrt{\sigma_2 - \sigma_3} \overset{P}{u}_3) \quad \text{Σχέση 3-27}$$

και οι παράμετροι  $a, b, c, d$  του επιπέδου θα είναι:

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \frac{\overset{P}{p}}{\|\overset{P}{p}\|} \quad \text{και} \quad d = \|\overset{P}{p}\| \quad \text{Σχέση 3-28}$$

Η μεταφορική ταχύτητα  $\overset{P}{v}$  θα δίνεται από τη σχέση:

$$\overset{P}{v} = \frac{1}{\sqrt{\sigma_1 - \sigma_3}} (\pm \sqrt{\sigma_1 - \sigma_2} \overset{P}{u}_1 - \sqrt{\sigma_2 - \sigma_3} \overset{P}{u}_3) \quad \text{Σχέση 3-29}$$

Η περιστροφική ταχύτητα  $\overset{P}{\omega}$  θα δίνεται από τη σχέση:

$$\overset{P}{\omega} = \overset{P}{W} - \frac{1}{2} \overset{P}{p} \times \overset{P}{v} \quad \text{Σχέση 3-30}$$

Λύσεις επίσης προκύπτουν με ταυτόχρονη αλλαγή των προσήμων στα  $\overset{P}{p}$  και  $\overset{P}{v}$ .

Στον πίνακα 4-1 δίνουμε μερικά αποτελέσματα της μεθόδου με τον πίνακα ροής σε

σύγκριση με τις επαναληπτικές μεθόδους που παρουσιάστηκαν στις παραγράφους 3.2.2 και 3.2.3. Για όλες τις περιπτώσεις θεωρήθηκε δεδομένο το μέτρο της ταχύτητας της κίνησης της κάμερας καθώς και οι αντιστοιχίσεις σημείων από εικόνα σε εικόνα. Για τις επαναληπτικές μεθόδους η αρχικοποίηση έγινε ως:

$$a=-1, b=0, c=0, d=200, (v_x, v_y, v_z)=(0,0,1), (\omega_x, \omega_y, \omega_z)=(0,0,0)$$

Παρατηρούμε στα αποτελέσματα ότι η μέθοδος με τον πίνακα ροής δίνει καλά αποτελέσματα όταν δεν υπάρχει καθόλου περιστροφική κίνηση ή όταν αυτή είναι πολύ μικρή. Αντίθετα, σε περιπτώσεις όπου έχουμε μεγάλη περιστροφική κίνηση, τα αποτελέσματα απέχουν πολύ από τα πραγματικά. Αυτό οφείλεται στο γεγονός ότι στη μέθοδο με τον πίνακα ροής χρησιμοποιούνται εξισώσεις στιγμιαίας ταχύτητας η οποία δεν μπορεί να οριστεί σωστά όταν από εικόνα σε εικόνα η μεταβολή της θέσης της κάμερας δεν είναι αρκετά μικρή.

Αντιθέτως, τα αποτελέσματα των επαναληπτικών μεθόδων δεν επηρεάζονται σχεδόν καθόλου από το μέγεθος της μεταβολής της θέσης της κάμερας. Αυτό συμβαίνει λόγω των γενικών εξισώσεων κίνησης που χρησιμοποιήθηκαν για τους απαραίτητους υπολογισμούς. Αν και η ακρίβεια των αποτελεσμάτων είναι ενδεικτική, μιας και εξαρτάται από τις παραμέτρους των επαναληπτικών αλγορίθμων (βήμα, τρόπος μεταβολής της θερμοκρασίας, κριτήριο λήξης κλπ) μπορεί κάποιος να παρατηρήσει μια ελαφριά υπεροχή του αλγορίθμου με υπολογισμό «ανάμεσα σε δύο επίπεδα» έναντι του αλγορίθμου «ώστε να συμπίπτουν δύο εικόνες». Εν μέρη, αυτό μπορεί να αιτιολογηθεί λόγω της επιπρόσθετης γνώσης των αντιστοιχίσεων των σημείων που χρησιμοποιεί η συγκεκριμένη μέθοδος.

**Πίνακας 4.1**

πραγματικές τιμές	αλγόριθμος «ώστε να συμπίπτουν δύο εικόνες»	αλγόριθμος «ανάμεσα σε δύο επίπεδα»	αλγόριθμος με χρήση του πίνακα ροής
a=-0.702 b= 0 c=-0.702 d=200 ( $v_x, v_y, v_z$ )=(0,0,3) ( $\omega_x, \omega_y, \omega_z$ )=(0,0,0)	a=-0.70 b=0.22 c=-0.68	a=-0.72 b=-0.10 c=-0.70	a=-0.71 b=-0.11 c=-0.70
a=-0.702 b= 0 c=-0.702 d=200 ( $v_x, v_y, v_z$ )=(0,0,3) ( $\omega_x, \omega_y, \omega_z$ )=(1°,1°,1°)	a=-0.78 b= 0.03 c=-0.63	a=-0.72 b= 0.02 c=-0.69	a=-0.69 b= 0.21 c=-0.69
a=-0.702 b=0 c=-0.702 d=200 ( $v_x, v_y, v_z$ )=(0,0,3) ( $\omega_x, \omega_y, \omega_z$ )=(10°,10°,10°)	Ο αλγόριθμος δεν κατάφερε να συγκλίνει	a=-0.72 b= 0.11 c=-0.70	a=-0.94 , b=0.31 c=-0.13

### 3.3. Συζήτηση

Στο κεφάλαιο αυτό μελετήσαμε το πρόβλημα της εύρεσης των παραμέτρων επιπέδων που απεικονίζονται σε δισδιάστατες εικόνες. Η ταυτόχρονη εύρεση όλων των παραμέτρων ενός επιπέδου είναι ένα δύσκολο πρόβλημα που εμπεριέχει στοιχεία από το, πολύ γνωστό στην μηχανική όραση, πρόβλημα της ανακατασκευής. Πράγματι, με οποιοδήποτε τρόπο και να δοκιμάσει κάποιος να προσεγγίσει το πρόβλημα της εύρεσης των παραμέτρων ενός επιπέδου, αναγκαστικά θα χρειαστεί, με τον ένα ή με τον άλλο τρόπο, να περάσει από το στάδιο της εύρεσης τρισδιάστατων συντεταγμένων. Άλλωστε, όταν ξέρουμε την εξίσωση ενός επιπέδου στο χώρο, ξέρουμε και τις τρισδιάστατες συντεταγμένες των σημείων που το απαρτίζουν.

Στην εργασία αυτή μελετήσαμε μεθόδους που χρησιμοποιούν παραπάνω από μια εικόνες και αναζητήσαμε πληροφορία στη μεταβολή των σημείων του επιπέδου πάνω στο φακό της κάμερας καθώς αυτή κινείται στο χώρο. Θεωρήσαμε ότι η εικόνα του επιπέδου είναι τέτοια ώστε η κίνηση της κάμερας να προκαλεί μεταβολή των σημείων (να υπάρχει δηλαδή έντονη υφή πάνω στην επιφάνεια του επιπέδου). Αν υπάρχει υφή πάνω στο επίπεδο, θεωρητικά, είναι εύκολο να υπολογισθεί η οπτική ροή αρκετά αξιόπιστα αφού όλες οι απαιτήσεις για τον υπολογισμό της, όπως παρουσιάστηκαν στην παράγραφο 1.3.5,



ικανοποιούνται. Η οπτική ροή είναι ίσως η σίγουρος τρόπος για να βρει κάποιος αντιστοιχίες σημείων σε τέτοιου είδους εικόνες.

Επικεντρώσαμε το ενδιαφέρον μας σε δύο διαφορετικές προσεγγίσεις: α) στην επαναληπτική εύρεση των παραμέτρων του επιπέδου ώστε να έχουμε βέλτιστες αντιστοιχίσεις ανάμεσα σε εικόνες τις οποίες υπολογίζουμε, και β) στην επαναληπτική εύρεση των παραμέτρων του επιπέδου ώστε αυτό να ταιριάζει όσο το δυνατόν περισσότερο σε τρισδιάστατα σημεία, τα οποία υπολογίζονται με βάση γνωστές αντιστοιχίσεις από εικόνα σε εικόνα.

Σύμφωνα με την πρώτη προσέγγιση, θεωρούμε ότι ξέρουμε τις παραμέτρους του επιπέδου και με βάση αυτές ελέγχουμε αν οι αντιστοιχίσεις σημείων ανάμεσα στις διαθέσιμες εικόνες είναι σωστές. Τις αντιστοιχίσεις αυτές μπορούμε, όπως αποδεικνύουμε, να τις βρίσκουμε αναλυτικά από τις παραμέτρους του επιπέδου και έτσι δεν χρειάζεται να τις υπολογίζουμε με οπτική ροή ή κάποιο άλλο τρόπο. Επαναληπτικά μεταβάλλουμε τις παραμέτρους του επιπέδου ώστε να μειώσουμε όσο το δυνατόν περισσότερο το σφάλμα ανάμεσα στις υπολογισμένες αντιστοιχίσεις και στις πραγματικές. Η μέθοδος παρουσίαζε πολύ καλά αποτελέσματα αν οι εικόνες είχαν παρθεί από θέσεις που δεν απείχαν πολύ μεταξύ τους. Σε διαφορετική περίπτωση έπρεπε να γίνει αρχικοποίηση των παραμέτρων του επιπέδου και της κίνησης με βάση κάποια γνώση ώστε οι τιμές τους να μην απέχουν πολύ από τις πραγματικές.

Σύμφωνα με τη δεύτερη προσέγγιση, υπολογίζουμε επαναληπτικά τις παραμέτρους έχοντας σας δεδομένες τις αντιστοιχίες σημείων. Τις αντιστοιχίες αυτές τις εξάγουμε από την οπτική ροή την οποία και πρέπει να υπολογίζουμε. Από τις αντιστοιχίες υπολογίζουμε τρισδιάστατες συντεταγμένες από τις τρισδιάστατες συντεταγμένες υπολογίζουμε τις παραμέτρους του επιπέδου. Ακόμα και όταν οι εικόνες βρίσκονται πολύ μακριά μεταξύ τους ή ακόμα και αν η αρχικές τιμές των παραμέτρων του επιπέδου ήταν πολύ μακριά από τις πραγματικές, ο αλγόριθμος δεν παρουσίαζε κανένα πρόβλημα στη σύγκλισή του. Αυτό οφείλεται κατά πολύ μεγάλο βαθμό στο γεγονός ότι μεγάλο μέρος της λύσης του προβλήματος έχει μετατεθεί στην εύρεση της οπτικής ροής, η οποία για εικόνες που απέχουν πολύ η μία από την άλλη είναι πολύ δύσκολο υπολογιστεί. Στα πειράματα που έγιναν, και στα αποτελέσματα που παρουσιάστηκαν η οπτική ροή θεωρήθηκε δεδομένη.

Μια σημαντική παράμετρος η οποία, αν δεν είναι είδη γνωστή, θα πρέπει να υπολογιστεί επίσης, είναι η κίνηση της κάμερας. Και οι δυο προσεγγίσεις που

παρουσιάστηκαν μελετήθηκαν με τέτοιο τρόπο ώστε να εμπεριέχουν σε κάποιο βαθμό και τη δυνατότητα να υπολογίζουν την κίνηση της κάμερας. Αν όπως η αρχικοποίηση δεν γίνει αρκετά κοντά στις πραγματικές τιμές οι επαναληπτικοί αλγόριθμοι δεν μπορούν να συγκλίνουν.

Παράλληλα με τις δύο αυτές προσεγγίσεις μελετήσαμε και μια ακόμα, γνωστή από τη βιβλιογραφία, μέθοδο: τη μέθοδο με τον πίνακα ροής. Ο πίνακας ροής είναι ένας πίνακας που εμπεριέχει την πληροφορία για τον δισδιάστατο μετασχηματισμό που υφίστανται τα σημεία σε μια εικόνα ενός επιπέδου καθώς η κάμερα κινείται στο χώρο. Δείχνουμε με ποιο τρόπο ο πίνακας αυτός μπορεί να υπολογιστεί από την οπτική ροή και με ποιο τρόπο μπορούμε να εξάγουμε την κίνηση της κάμερας και τις παραμέτρους του επιπέδου από αυτόν. Στη μελέτη που κάναμε για τον πίνακα ροής θεωρήσαμε μόνο ένα ζευγάρι από εικόνες πράγμα που μας εμπόδιζε να υπολογίσουμε όλες τις ζητούμενες παραμέτρους ταυτόχρονα. Θεωρήθηκε ως προϋπόθεση ότι η μεταφορική κίνηση της κάμερας είχε μέτρο μονάδα. Παράλληλα, θεωρήθηκαν, για αυτό τον αλγόριθμο, οι απλοποιημένες εξισώσεις για την κίνηση της κάμερας, οι εξισώσεις δηλαδή στιγμιαίας ταχύτητας που ισχύουν μόνο αν η απόσταση που διανύει η κάμερα και η αλλαγή στις γωνίες των αξόνων της από εικόνα σε εικόνα είναι πολύ μικρές. Όταν η περιστροφική κίνηση της κάμερας είναι αρκετά μεγάλη από εικόνα σε εικόνα ο συγκεκριμένος αλγόριθμος αδυνατεί να δώσει αξιόπιστα αποτελέσματα.

## 4. ΣΥΜΠΕΡΑΣΜΑΤΑ

Σε αυτή την εργασία ασχοληθήκαμε με την εξαγωγή δύο συγκεκριμένων τρισδιάστατων χαρακτηριστικών από δισδιάστατες εικόνες: τις τρισδιάστατες διευθύνσεις ομάδων ευθυγράμμων τμημάτων, όπως αυτές μπορούν να υπολογισθούν από τα σημεία φυγής τους, και τις παραμέτρους επιπέδων στο χώρο τις οποίες υπολογίσαμε χρησιμοποιώντας παραπάνω από μια εικόνες.

Η εξαγωγή σημείων φυγής από δισδιάστατες προοπτικές προβολές είναι ένα σημαντικό πρόβλημα στη μηχανική όραση. Η ανάλυση των εξαγχθέντων σημείων φυγής μπορεί να βοηθήσει πολύ στην κατανόηση της δομής ενός τρισδιάστατου χώρου ή αντικειμένου. Πρακτικοί αλγόριθμοι που έχουν προταθεί κατά καιρούς για τον υπολογισμό σημείων φυγής στηρίζονται στη γνώση των ευθυγράμμων τμημάτων που απαρτίζουν μια εικόνα. Σε αυτή την εργασία μελετήσαμε αυτούς τους αλγορίθμους αλλά και προτείναμε κάποιες διαφορετικές προσεγγίσεις. Οι προσεγγίσεις αυτές παρουσιάζουν ένα μεγάλο πλεονέκτημα σε σχέση με τις υπάρχοντες μεθόδους μιας και δεν προϋποθέτουν εξαγωγή ευθυγράμμων τμημάτων, μια διαδικασία που όπως φάνηκε στα αποτελέσματα που παραθέσαμε (τουλάχιστον με την προσέγγιση με μετασχηματισμό Hough που δοκιμάσαμε) δεν είναι πάντα πολύ αξιόπιστη. Πάντως, οποιοσδήποτε αλγόριθμος και να χρησιμοποιηθεί, αν στην εικόνα δεν υπάρχουν ευδιάκριτες ακμές που να υπερισχύουν των υπόλοιπων χαρακτηριστικών της, η εύρεση των σημείων φυγής καθίσταται σχεδόν αδύνατη. Αυτό συμβαίνει γιατί όλοι οι αλγόριθμοι στηρίζονται, με διάφορους τρόπους, είτε στις ακμές της εικόνας είτε στη βαθμίδα της φωτεινότητας της εικόνας.

Κάθε αλγόριθμος από αυτούς που μελετήθηκαν έχει τα πλεονεκτήματα και τα μειονεκτήματα του. Ανάλογα με το είδος της εικόνας κάποιοι από τους αλγορίθμους μπορεί να δυσκολεύονται πολύ στο να υπολογίσουν σωστά και με ακρίβεια τα σημεία φυγής ή μπορεί να είναι εντελώς ακατάλληλοι. Στους αλγορίθμους με ανίχνευση ευθυγράμμων τμημάτων, σημαντικό ρόλο παίζει το μέγεθος του μετασχηματισμού Hough που θα χρησιμοποιηθεί το οποίο γενικά εξαρτάται από το μέγεθος και την πολυπλοκότητα της εικόνας. Για τους αλγορίθμους που δεν χρησιμοποιούν ευθύγραμμα τμήματα, καλύτερα αποτελέσματα παρουσίασαν οι αλγόριθμοι που κάνανε χρήση συνάρτησης με το εμβαδόν (VPLF) αφού η συνάρτηση με ολοκλήρωση στο χώρο του Hough (VPLFSIN) παρουσίαζε τα εγγενή προβλήματα του μετασχηματισμού ενώ η συνάρτηση που έκανε

χρήση των προβολών της βαθμίδας της φωτεινότητας απαιτούσε πολύ καλή γνώση της βαθμίδας, πράγμα το οποίο δεν ήταν πάντα εφικτό (ιδιαίτερα σε γραμμές που δεν ήταν απόλυτα οριζόντιες ή κατακόρυφες)

Για την επιλογή πιθανών σημείων φυγής και την αξιολόγησή τους χρησιμοποιήσαμε τη μοναδιαία σφαίρα. Με αυτό τον τρόπο είχαμε τη δυνατότητα να επιλέγουμε υποψήφια σημεία φυγής τόσο κοντά στην εικόνα όσο και πολύ μακριά από αυτή (ακόμα και στο άπειρο). Για ανίχνευση σημείων που να παρουσιάζουν μεγάλες τιμές της συνάρτησης αξιολόγησης (πιθανά σημεία φυγής) μελετήσαμε δύο διαφορετικούς τρόπους: ένα επαναληπτικό αλγόριθμο τύπου simulated annealing και ένα αλγόριθμο που εξετάζει εξαντλητικά όλες τις τιμές από ένα ιστόγραμμα ( $\varphi$ - $\theta$ ), τα κελιά του οποίου, κατά τη διάρκεια εκτέλεσης ενός ιεραρχικού αλγορίθμου, μειώνονται συνεχώς σε μέγεθος γύρω από σημεία που παρουσιάζουν ενδιαφέρον (παρουσιάζουν μέγιστα-κορυφές της συνάρτησης αξιολόγησης). Και οι δύο αλγόριθμοι παρουσίασαν καλά και αναμενόμενα (ανάλογα με τη φύση τους) αποτελέσματα.

Στο 4<sup>ο</sup> κεφάλαιο της εργασίας μας ασχοληθήκαμε με τον υπολογισμό παραμέτρων επιπέδων στο χώρο, ένα πρόβλημα που, εν μέρη, εμπεριέχει το, γνωστό στη μηχανική όραση, πρόβλημα της ανακατασκευής. Μελετήσαμε μεθόδους που χρησιμοποιούν παραπάνω από μια εικόνες και αναζητήσαμε πληροφορία στη μεταβολή των σημείων του επιπέδου πάνω στο φακό της κάμερας καθώς αυτή κινείται στο χώρο. Θεωρήσαμε ότι η εικόνα του επιπέδου είναι τέτοια ώστε η κίνηση της κάμερας να προκαλεί μεταβολή των σημείων (να υπάρχει δηλαδή έντονη υφή πάνω στην επιφάνεια του επιπέδου).

Επικεντρώσαμε το ενδιαφέρον μας σε δύο διαφορετικές προσεγγίσεις: α) στην επαναληπτική εύρεση των παραμέτρων του επιπέδου ώστε να έχουμε βέλτιστες αντιστοιχίσεις ανάμεσα σε εικόνες τις οποίες υπολογίζουμε αναλυτικά, και β) στην επαναληπτική εύρεση των παραμέτρων του επιπέδου ώστε αυτό να ταιριάζει όσο το δυνατόν περισσότερο σε τρισδιάστατα σημεία, τα οποία υπολογίζονται με βάση γνωστές αντιστοιχίσεις από εικόνα σε εικόνα. Και οι δύο προσεγγίσεις μελετήθηκαν με τέτοιο τρόπο ώστε να εμπεριέχουν σε κάποιο βαθμό την δυνατότητα εύρεσης των παραμέτρων κίνησης της κάμερας και, παρουσίασαν καλά και αξιόπιστα αποτελέσματα.

Για την πρώτη προσέγγιση, καθοριστική σημασία στην επιτυχία του αλγορίθμου έπαιξε η σωστή αρχικοποίηση των παραμέτρων του επιπέδου και της κίνησης της κάμερας. Σε περιπτώσεις που οι αρχικές τιμές δεν ήταν πολύ κοντά στις πραγματικές

(ιδιαίτερα αυτές που αφορούσαν τη κίνηση της κάμερας) ο αλγόριθμος δεν κατάφερε να συγκλίνει. Για τη δεύτερη προσέγγιση η σωστή αρχικοποίηση δεν έπαιξε και είχε πολύ μεγάλη σημασία. Αυτό συνέβαινε γιατί μεγάλο μέρος των υπολογισμών είχε μετατεθεί στον υπολογισμό της οπτικής ροής, την γνώση της οποίας ο αλγόριθμος θεωρούσε δεδομένη.

Παράλληλα με τις δύο αυτές προσεγγίσεις μελετήσαμε και μια ακόμα, γνωστή από τη βιβλιογραφία, μέθοδο: τη μέθοδο με τον πίνακα ροής. Ο πίνακας ροής είναι ένας πίνακας που εμπεριέχει την πληροφορία για τον δισδιάστατο μετασχηματισμό που υφίστανται τα σημεία σε μια εικόνα ενός επιπέδου καθώς η κάμερα κινείται στο χώρο. Δείχνουμε με ποιο τρόπο ο πίνακας αυτός μπορεί να υπολογιστεί από την οπτική ροή και με ποιο τρόπο μπορούμε να εξάγουμε την κίνηση της κάμερας και τις παραμέτρους του επιπέδου από αυτόν. Στη μελέτη που κάναμε για τον πίνακα ροής θεωρήσαμε μόνο ένα ζευγάρι από εικόνες πράγμα που μας εμπόδιζε να υπολογίσουμε όλες τις ζητούμενες παραμέτρους ταυτόχρονα. Θεωρήθηκαν σαν προϋποθέσεις ότι η μεταφορική κίνηση της κάμερας είχε μέτρο μονάδα και ότι η απόσταση του επιπέδου από το εστιακό σημείο ήταν επίσης μονάδα. Παράλληλα, θεωρήθηκαν, για αυτό τον αλγόριθμο, οι απλοποιημένες εξισώσεις για την κίνηση της κάμερας, οι εξισώσεις δηλαδή στιγμιαίας ταχύτητας που ισχύουν μόνο αν η απόσταση που διανύει η κάμερα και η αλλαγή στις γωνίες των αξόνων της από εικόνα σε εικόνα είναι πολύ μικρές.



# ΒΙΒΛΙΟΓΡΑΦΙΑ

## 4.1. Βιβλία

1. Bow, S.T. Pattern Recognition and Image Preprocessing, Marcel Dekker. New York, 1992.
2. Cose, E., Johnosonbaugh, R., Jost, S. Pattern Recognition and Image Analysis, Prentice Hall PTR. New Jersey, 1996.
3. Kanatani, K. Geometric Computation for Machine Vision, Oxford University Press. Yew York, 1995.
4. Mundy, J.L, Zisserman, A. Projective geometry in machine vision. In Mundy, J.L, Zisserman, A., editors, Geometric Invariance in Computer Vision, pages 463-534. MIT Press, 1992.
5. Αργυρός, Α.. Διδακτορική διατριβή. Τμήμα επιστήμης υπολογιστών, Πανεπιστήμιο Κρήτης, 1999.

## 4.2. προεπεξεργασία

6. Chu, Y.K., Suen, C.Y. An alternate smoothing and stripping algorithm for thinning digital binary patterns. *Signal Processing*, vol. 11, no. 3, pp. 207-222. 1986.
7. Lam, L., Suen, C.Y. A Dynamic Shape Preserving Thinning algorithm. *Signal Processing*, **22**, 199-208. 1991.

## 4.3. μετασχηματισμός Hough

8. Ballard, D.H. Generalizing the Hough transform to detect arbitrary shapes. *Pattern Recognition*, vol. 13(2), pp. 111-122. 1981.
9. Illingworth, J., Kittler, J. A survey of the Hough transform. *Computer Vision, Graphics and Image Processing*, Vol. 44, pp. 87-116. 1988.

#### 4.4. σημείο φυγής - χρήση

10. K. Daniilidis and J. Ernst. Active Intrinsic Calibration Using Vanishing Points, *Pattern Recognition Letters*, vol. 17(11), pp. 1179-1189. 1996.
11. Parodi, P., Picciolo, G. 3D shape reconstruction by using vanishing points, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 18, No. 2. 1996.
12. McGlone, J.C., Shufelt, J.A. Incorporating vanishing Point Geometry into a Building Extraction System, *Proceedings of the SPIE: Integrating Photogrammetric Techniques with Scene Analysis and Machine Vision*, September 1993, vol. 1944, pp. 273-284. 1993.
13. Collins, R. T., Beveridge J. R., Matching perspective views of coplanar structures using projective unwrapping and similarity matching. *IEEE Conference on Computer Vision and Pattern Recognition*, New York City, June 1993. 1993.
14. Carpile, B., Torre, T. Using vanishing points for camera calibration. *International Journal of Computer Vision*, vol. 4, pp. 127-140. 1990

#### 4.5. σημείου φυγής - υπολογισμός

15. Barnard, S.T. Interpreting Perspective Images. *AI Journal*, vol. 21, no. 44, pp. 435-462. 1983.
16. Collins, R.T., Weiss, R.S. An efficient and accurate method for computing vanishing points. *Topical meeting on image understanding and machine vision*, Optical Society of America, Technical Digest Series, vol. 14, pp. 92-94. 1989.
17. Collins, R.T., Weiss, R.S. Vanishing point calculation as a statistical inference on the unit sphere. *Intl. Conf. Computer Vision (ICCV'90)*, Osaka, Japan, pp.400-403. 1990.
18. Kender, J.R. Shape from texture: An aggregation transform that maps a class of textures into surface orientation. *Proceedings IJCAI 6*, Tokyo, Japan, pp. 475-480. 1979.
19. Magee, M.J., Aggrwal, J.K. Determining vanishing points from perspective images. *Computer Vision, Graphics and Image Processing*, vol. 26, pp. 256-267. 1984.
20. Quan, L., Mohr, R. Determining perspective structures using hierarchical Hough transform. *Pattern Recognition Letters*, vol. 9, pp. 279-286. 1989.
21. Baltzakis, H., Trahanias, P. The VPLF method for computing vanishing points, submitted in *Image and Vision Computing*.



## 4.6. οπτική ροή – υπολογισμός

22. Baron, J. L., Fleet, D. J., Beauchemin, S. S. Performance of optical flow techniques. *International Journal of Computer Vision*, 12:1, 43-77. 1994.
23. Buxton, B., Blake, A. Computation of optic flow from the motion of edge features in image sequences. *Image Vis. Comput.* 2:59-74.1984.
24. Simoncelli, E.P., Adelson, E.B., Heeger, D.J. Probability distributions of optical flow. *IEEE Conference on Computer Vision and Pattern Recognition*, Maui, Hawaii. June 1991.

## 4.7. πλοήγηση

25. Madsen, C.B., Christensen. H.I., A viewpoint planning strategy for determining true angles on polyhedral objects by camera alignment. *IEEE Transactions Pattern Analysis and Mashine Intelligence*, vol. 19, no. 2, February 1997.
26. Dev, A., Krose, B., Groen, F. Navigation of a mobile robot on the temporal development of the optic flow. *Proceedings IROS'97, IEEE* , pp. 558-563. 1997.
27. Santos-Victor, J., Sandini, G., Gurotto, F., Garibaldi, S. Divergent stereo in autonomous navigation. From bees to robots. *International Journal of Computer Vision*, 14, 159-177 (1995)
28. Dev A., Krose, B.J.A., Groen F.C.A. “Where are you driving to ?” Heading Direction for a Mobile Robot from Optical Flow. *Proc. Of the IEEE Int. Conf on Robotics and Automation*, Lenven, pp 558-563. May 1997.

## 4.8. εξαγωγή μερικής 3δ πληροφορίας

29. Sipola, R., Buxton, H. Surface orientation and time to contact from image divergence and deformation. *Proceedings. 2nd European. Conference. Computer. Vision., Ligure, Italy*, pp. 187-202. 1992.
30. Tan, T.N., Baker, K.D., Sullivan G.D. Model independent recovery of object orientations. *IEEE Transactions of Robotics and Automation*, vol. 13, no. 4. August 1997
31. Madsen, C.B. Direct computation of true angle probabilities from single perspective views of polyhedral objects. *IEEE Transactions Pattern Analysis and Machine Intelligence*, July 1997, submitted for review.

32. Sinclair D., Blake, A. Quantitive planar region detection. *International Journal of Computer Vision* 18 (1), 77-91. 1996.
33. Coelho, C., Straforini, M., Campani, M. Using geometrical rules and a priori knowledge for the understanding of indoor scenes. *Robust90*, University of Genoa, Italy. 1990.

#### 4.9. εξαγωγή κίνησης

34. Brouce, A.R., Horn, B.K., *Passive Navigation*. *Computer Graphics and Image Processing*, 21, pp. 3-20. 1983.
35. Longuet-Higgins H.C. A computer algorithm for reconstructing a scene from two projections. *Nature* 293 (10), pp 133-135. 1981.
36. Negahdaripour, S., Lee, S. Motion recovery from image sequences using only first order optic flow information. *International Journal of Computer Vision*, 9 (3), 163-184. 1992.

#### 4.10. εξαγωγή κίνησης και 3d πληροφορίας

37. Zhengyou Zhang. Estimating Motion and Structure from Correspondences of Line Segments between Two Perspective Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, no. 12. December 1995.
38. Huang, T.S., Netravali, A. Motion and Structure from Feature Correspondences: A review. *Proceedings of the IEEE*, vol. 82, no. 2. February 1994.
39. Wend, J., Huang, T.S. Nerendra Ahuja. Motion and Structure from Two Perspective Views: Algorithm, Error Analysis, and Error Estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 5. May 1989

#### 4.11. κίνηση ώστε να έχουμε βέλτιστη εξαγωγή πληροφορίας

40. Kutulakos, K.N., Dyer, C. R. Global surface reconstruction by purposive control of observer motion. *Proceedings. IEEE conference on Computer Vision and Pattern Recognition*, June 1994, pp 331-338. 1994.

## 4.12. μη βαθμονομημένη κάμερα

41. Faugeras, O.D. What can be seen in three dimensions with an uncalibrated rig. Proceedings (Gioulio Sandini, editor). 2nd European Conference on Computer Vision, pages 563-578. Springer-Verlag, Lecture Notes in Computer Science 588.1992.
42. Hartley, R., Gupta, R., Chang, T. Stereo from uncalibrated cameras. Proceedings. CVPR92 Urbana-Champaign, Illinois, Pages 761-764. 1992.
43. Mohr, R., Arbogast E. It can be done without camera calibration. Pattern Recognition Letters, 12: 39-43. 1990.

## 4.13. διάφορα

44. Kanatani, K. Computational cross ratio for computer vision, CVGIP - Image Understanding, vol. 60, no. 3, 371-381. 1994.
45. Faugeras, O.D., Robert, L. What can two Images tell us about a third one. International Journal of Computer Vision, 18, 5-19. 1996.
46. Madsen, C.B., Christensen, H.I. Modeling and testing the stability of edge segments: length and orientation. Proceedings : Ninth Scandinavian conference on Image Analysis. Uppsala, Sweden, pp. 1011-1019. June 1995.
47. Burns, J.B. , Hanson, A.R. Riseman E.M. Extracting straight lines. IEEE Transactions on Pattern analysis and Machine Intelligence, vol. 8, No. 4, pp. 425-456. 1986.



# ΕΥΡΕΤΗΡΙΟ

calibration.....	17, 78, 81
concurrent lines .....	65
contour.....	19, 26
corner point .....	19
cross-product.....	26
edge detection.....	18
end-point .....	19
flow matrix .....	53, 65
focal length.....	ii, 6
Hough μετασχηματισμός.....	18, 21, 22, 23, 24, 25, 26, 29, 30, 31, 39, 40, 41, 47, 48, 73, 77, 78
ιστόγραμμα.....	22, 26, 40
πίνακας .....	25, 29, 30, 31, 48
image sequence .....	ii, 8, 79, 80
interpretation plane.....	26, 27
pinhole.....	5, 6
simulated annealing.....	45, 46, 58, 74
stereo .....	2, 79
texture.....	78
trace point.....	19
vanishing line .....	6
vanishing point.....	ii, 6, 78
view volume .....	6
VPLF .....	ii, 34, 36, 37, 38, 45, 73, 78
VPLFGRAD.....	42, 43, 46
VPLFSIN.....	39, 41, 73
ακμή .....	2, 17, 18, 19, 22, 32, 34, 35, 36, 42, 46, 48, 73
ακμή	
εξαγωγή ακμών .....	18, 19, 22
σημείο ακμής.....	34, 35
ακρίβεια.....	1, 18, 26, 33, 41, 44, 47, 48, 52, 59, 62, 63, 69, 73
αλγόριθμος i, 2, 17, 18, 19, 24, 25, 26, 29, 31, 32, 35, 41, 43, 45, 46, 47, 49, 51, 52, 53, 54, 58, 59,	
60, 62, 63, 64, 69, 70, 71, 72, 73, 74, 75	
αλγόριθμος	

average linkage.....	24
αναδρομικός.....	26
επαναληπτικός.....	45, 52, 53, 58, 72, 74
ιεραρχικός.....	24, 44, 49
σύγκλιση.....	42, 46, 53, 59, 60, 61, 63, 64, 70, 75
αμφιβληστροειδής χιτώνας.....	1, 5
αναδρομικός αλγόριθμος.....	26
ανακατασκευή.....	1, 51, 53, 70, 74
αναλλοίωτο χαρακτηριστικό.....	52
ανίχνευση.....	3, 18, 26, 29, 33, 47, 73, 74
ανίχνευση	
γραμμών.....	18
ευθυγράμμων τμημάτων.....	3, 18, 26, 33, 73
σημείων φυγής.....	18, 26
συνοριακών γραμμών.....	26
αντιστοιχία σημείων.....	3, 51, 52, 53, 56, 58, 59, 62, 65, 69, 71, 74
αξιολόγηση.....	i, 3, 22, 33, 35, 44, 45, 48, 49, 52, 53, 58, 74
άξονας.....	6, 7, 8, 9, 10, 11, 21, 26, 27, 28, 29, 32, 39, 54, 57, 61, 62, 72, 75
άξονας	
αρχή των αξόνων.....	6, 7, 21, 26, 27, 28, 29, 32, 51, 54, 68
οπτικός.....	6
σύστημα αξόνων.....	9, 10, 55, 61
απόσταση.....	6, 21, 22, 25, 51, 54, 56, 57, 64, 72, 75
αρχή των αξόνων.....	6, 7, 21, 26, 27, 28, 29, 32, 51, 54, 68
αρχικοποίηση.....	69, 71, 72, 74
βαθμίδα φωτεινότητας.....	31, 32, 33, 42, 43, 46, 48, 73, 74
βαθμονόμηση κάμερας.....	1, 3, 17
βάρος.....	22, 29
βήμα.....	19, 25, 33, 58, 59, 62, 63, 69
γειτονιά.....	19
γειτονιά	
4-γειτονιά.....	19
8-γειτονιά.....	19, 44
γινόμενο	
εξωτερικό.....	18, 26
εσωτερικό.....	10
γνώση εκ των προτέρων.....	1, 2, 17, 49

γραμμή	
διεύθυνση .....	i, 2, 17
συνοριακή.....	19, 26
συντρέχουσες γραμμές .....	65
γραμμικό σύστημα .....	62, 66
διάνυσμα .....	7, 9, 10, 11, 12, 13, 27, 28, 31, 42, 53, 54, 67, 68
διάνυσμα	
κάθετο .....	27, 28, 42
κατεύθυνσης.....	7, 54
μοναδιαίο.....	9, 10, 53, 55
διεύθυνση .....	i, 2, 3, 11, 17, 24, 25, 26, 27, 32, 47, 48, 57, 73
διεύθυνση	
γραμμών .....	i, 2, 17
δισδιάστατος.....	i, 2, 17, 46, 47, 51, 70, 73
εικόνα i, 2, 3, 5, 6, 8, 11, 13, 14, 15, 17, 18, 19, 20, 22, 23, 24, 25, 26, 27, 29, 31, 32, 33, 34, 35,	
36, 38, 40, 41, 42, 43, 44, 46, 47, 48, 51, 52, 53, 55, 56, 57, 58, 61, 65, 66, 69, 70, 71, 72, 73, 74,	
75	
εικόνα	
εικονοστοιχείο .....	57
συντεταγμένες .....	61, 62
εικονοστοιχείο.....	57
εμβαδόν .....	34, 35, 49, 73
εξαγωγή	
ακμών .....	18, 19, 22
ευθυγράμμων τμημάτων.....	18, 73
εξίσωση .....	14, 29, 53, 54, 55, 56, 57, 66, 69, 70, 72, 75
εξίσωση	
επιπέδου .....	53, 54, 55, 56, 70
προοπτικής προβολής.....	55, 56
εξωτερικά γινόμενα.....	18, 26
επανάληψη .....	19, 25, 44, 45, 52, 59, 63
επίπεδο i, 2, 3, 5, 6, 8, 13, 14, 17, 21, 22, 23, 24, 26, 27, 28, 31, 32, 39, 44, 48, 49, 51, 52, 53, 54,	
55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 68, 69, 70, 71, 72, 73, 74, 75	
επίπεδο	
διάνυσμα κατεύθυνσης.....	54
εξίσωση .....	53, 54, 55, 56, 70
επίπεδο αντικείμενο .....	17

παράμετροι .....	51, 53, 55, 58, 61, 67, 68, 71, 72, 75
πολικών συντεταγμένων.....	21, 22, 23, 39
προβολής .....	5, 6, 13, 14, 48, 49, 55, 56, 64
φωτεινότητας.....	31
εστιακό μήκος .....	2, 6, 7, 17, 27, 44, 48, 57
εστιακό σημείο .....	i, 2, 6, 12, 17, 75
εσωτερικό γινόμενο.....	10
ευθεία .....	6, 7, 8, 17, 21, 23, 24, 39, 42
ευθεία	
φυγής .....	6, 8, 17
ευθύγραμμο τμήμα .....	i, 2, 3, 18, 22, 24, 26, 29, 32, 33, 34, 35, 39, 40, 41, 42, 47, 73
ευθύγραμμο τμήμα	
ανίχνευση .....	3, 18, 26, 33, 73
ευθύγραμμο τμήμα	
εξαγωγή .....	18, 73
ημιεπίπεδο .....	54
ημιτονοειδής καμπύλη.....	22
θερμοκρασία.....	59, 69
ιεραρχικός αλγόριθμος .....	24, 44, 49
ιστόγραμμα.....	18, 22, 26, 29, 30, 32, 33, 35, 40, 41, 44, 74
ίχνος.....	29, 32
κάθετο διάνυσμα .....	27, 28, 42
κάμερα i, 2, 3, 5, 6, 8, 9, 10, 11, 12, 13, 17, 51, 52, 53, 55, 56, 57, 58, 59, 61, 62, 63, 64, 67, 69, 70,	
71, 72, 74, 75, 81	
κάμερα	
βαθμονόμηση .....	1, 3, 17
μοντέλο.....	5
όμοιες κάμερες .....	61
παράμετροι .....	51
φακός.....	6, 53, 70, 74
καμπύλη.....	21, 22, 23, 40
καμπύλη	
ημιτονοειδής.....	22
κανονικοποίηση.....	54
καρτεσιανές συντεταγμένες.....	21, 23
καρτεσιανό επίπεδο .....	26, 39
κατακόρυφα τμήματα.....	2, 17



κατακόρυφος .....	29, 48, 62, 74
κατεύθυνση .....	2, 7, 17, 34, 59, 63
κατεύθυνση	
σημαντικές κατευθύνσεις .....	22
κατώφλι .....	22, 25, 29, 43, 44, 62, 63
κεντρική προβολή .....	6
κίνηση	
μεταφορική .....	12, 57, 61, 64, 72, 75
παράμετροι .....	53, 58, 59, 62, 63, 67, 74
περιστροφική .....	8, 12, 14, 69, 72
κορυφή .....	18, 22, 26, 29, 32, 35, 41, 44, 47, 49, 52, 74
κορυφή	
πολυγώνου .....	52
κριτήριο .....	19, 22, 59, 69
κριτήριο	
λήξης .....	59, 69
κύκλος	
μέγιστος .....	26, 27, 29, 32
μάσκα .....	18, 19, 31, 48
μάσκα	
Prewitt .....	31, 48
Roberts .....	18
μάτι της βελόνας .....	5, 6
μέγιστος κύκλος .....	26, 27, 29, 32
μέθοδος .....	3, 18, 22, 26, 33, 38, 48, 49, 52, 66, 69, 71, 72, 75
μέθοδος ελαχίστων τετραγώνων .....	66
μέση τιμή .....	22, 25, 31, 59
μετασχηματισμός	
προοπτικός .....	5, 7
μετασχηματισμός Hough .....	18, 21, 22, 23, 24, 25, 26, 29, 30, 31, 39, 40, 41, 47, 48, 73, 77, 78
μετασχηματισμός Hough	
ιστόγραμμα .....	22, 26, 40
πίνακας .....	25, 29, 30, 31, 48
μεταφορική κίνηση .....	12, 61, 64, 72, 75
μεταφορική ταχύτητα .....	12, 67, 68
μηχανική όραση .....	i, 2, 5, 17, 70, 73, 74
μοναδιαία σφαίρα .....	26, 27, 28, 29, 30, 32, 33, 35, 36, 38, 41, 43, 44, 46, 47, 48, 49, 74

μοναδιαίο διάνυσμα.....	9, 10, 53, 55
μοντέλο.....	2, 3, 5, 8, 17
μοντέλο	
μοντέλο κάμερας.....	5
προβολικό.....	2, 5
όγκος όρασης.....	6
ολοκλήρωση.....	39, 48, 73
οπτική εστία.....	6, 51, 55
οπτική ροή.....	14, 15, 52, 70, 71, 72, 75, 79
οπτικός άξονας.....	6
οριζόντιες συντεταγμένες.....	62
παραδοχή.....	1, 15, 65
παραμετροποίηση.....	26, 36, 44
παράμετρος $i$ , 2, 3, 35, 36, 48, 51, 52, 53, 54, 55, 58, 59, 61, 62, 63, 64, 65, 67, 68, 69, 70, 71, 72, 73, 74, 75	
πιθανότητα.....	$i$ , 3, 33, 42, 43, 59, 63
πίνακας ροής.....	3, 53, 64, 65, 67, 68, 69, 70, 72, 75
πλοήγηση.....	$i$ , 2, 51, 79
πολικές συντεταγμένες.....	21, 22, 23, 26, 39, 44
ποσοστό.....	62
προβολή.....	5, 13, 14, 18, 26, 30, 33, 35, 41, 43, 48, 49, 59, 61, 63
προβολή	
κεντρική.....	6
προοπτική.....	$i$ , 2, 17, 73
προβολικό μοντέλο.....	2, 5
προεπεξεργασία.....	20, 77
προοπτική.....	5
προοπτική προβολή.....	$i$ , 2, 17, 73
προοπτικός μετασχηματισμός.....	5, 7
προϋπόθεση.....	11, 75
ροή	
οπτική.....	14, 15, 52, 70, 71, 72, 75, 79
ρομποτικό σύστημα.....	1, 51
σημείο $i$ , 2, 3, 5, 6, 7, 8, 9, 11, 12, 13, 14, 15, 17, 18, 19, 21, 22, 23, 24, 25, 26, 27, 28, 29, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 51, 52, 53, 55, 56, 58, 61, 62, 64, 65, 66, 69, 70, 71, 72, 73, 74, 75, 78	

σημείο	
ακμής.....	34, 35
αντιστοιχία σημείων.....	3, 51, 52, 53, 56, 58, 59, 62, 65, 69, 71, 74
εστιακό.....	i, 2, 6, 12, 17, 75
συνεπίπεδα σημεία.....	51
τομής.....	52
σημείο φυγής i, 2, 3, 6, 7, 17, 18, 25, 26, 29, 32, 33, 34, 35, 36, 37, 38, 40, 41, 42, 43, 44, 45, 47, 48, 49, 73, 74, 78	
σκελετοποίηση.....	19, 20, 22
στατικό περιβάλλον.....	8
στατιστική.....	18
στατιστική ανάλυση.....	18
σύγκλιση.....	42, 46, 53, 59, 60, 61, 63, 64, 70, 75
συγχώνευση.....	24
συνάρτηση	
αξιολόγησης.....	i, 3, 33, 35, 40, 41, 42, 43, 44, 45, 46, 48, 49, 74
συνεισφοράς.....	34
συνοριακή γραμμή.....	19, 26
συντεταγμένες.....	9, 10, 12, 13, 21, 23, 26, 44, 48, 51, 53, 55, 56, 61, 62, 65, 70, 71
συντεταγμένες	
εικόνας.....	61, 62
καρτεσιανές.....	21, 23
οριζόντιες.....	62
πολικές.....	21, 22, 23, 26, 39, 44
σύστημα συντεταγμένων.....	9, 12, 13, 51, 55, 56
τριδιάστατες.....	51, 52, 53, 55, 62, 70, 71
συντρέχουσες γραμμές.....	65
σύστημα.....	1, 9, 10, 11, 12, 13, 17, 29, 51, 55, 56, 61, 66, 67
σύστημα	
αξόνων.....	9, 10, 55, 61
γραμμικό.....	62, 66
όρασης.....	1
συντεταγμένων.....	9, 12, 13, 51, 55, 56
σφαίρα.....	14, 26, 27, 28, 29, 30, 32, 33, 36, 44, 48, 49
σφαίρα	
μοναδιαία.....	26, 27, 28, 29, 30, 32, 33, 35, 36, 38, 41, 43, 44, 46, 47, 48, 49, 74
σφάλμα.....	i, 2, 22, 33, 46, 52, 59, 60, 62, 63, 71

σφάλμα	
κβαντισμού .....	22, 52
σχετική θέση .....	55, 56
ταξινόμηση .....	22
ταχύτητα .....	11, 12, 13, 14, 18, 47
ταχύτητα	
μεταφορική .....	12, 67, 68
περιστροφική .....	11, 12, 68
τερματισμός .....	59
τεχνική .....	18, 26, 33, 49
τμηματοποίηση .....	2, 17, 18, 24, 25, 29
τοίχος .....	51
τομή .....	6, 24, 25, 26, 27, 28, 33, 47, 65
τριδιάστατες συντεταγμένες .....	51, 52, 53, 55, 62, 70, 71
τριδιάστατη πληροφορία .....	17
τριδιάστατος χώρος .....	i, 2, 5, 6, 12, 13, 17, 73
τριδιάστατων .....	i, 6, 51, 53, 70, 73
υλοποίηση .....	26, 44, 54
υπόθεση .....	53, 58
φακός της κάμερας .....	6, 53, 70, 74
φωτεινότητα .....	15, 31
φωτεινότητα	
βαθμίδα .....	31, 32, 33, 42, 43, 46, 48, 73, 74
επίπεδο .....	31
χαρακτηριστικό	
αναλλοίωτο .....	52
ψήφος .....	32