ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ
ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΤΜΗΜΑ ΕΠΙΣΤΗΜΗΣ ΥΠΟΛΟΓΙΣΤΩΝ

Αλγόριθμοι χρήσης προτύπων αλλαγής προσήμων για τον εντοπισμό περιοχών FOE και AOR στην περίπτωση ιδίας κίνησης

Στογιαννίδης Αναστάσιος

Μεταπτυχιακή Εργασία

Ηράκλειο, Νοέμβριος 1995

ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ
ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΤΜΗΜΑ ΕΠΙΣΤΗΜΗΣ ΥΠΟΛΟΓΙΣΤΩΝ

Αλγόριθμοι χρήσης προτύπων αλλαγής προσήμων για τον εντοπισμό περιοχών FOE και AOR στην περίπτωση ιδίας κίνησης

Εργασία που υποβλήθηκε από τον
ΣΤΟΓΙΑΝΝΙΔΗ ΑΝΑΣΤΑΣΙΟ
ως μερική εκπλήρωση των απαιτήσεων
για την απόκτηση
ΜΕΤΑΠΤΥΧΙΑΚΟΥ ΔΙΠΛΩΜΑΤΟΣ ΕΙΔΙΚΕΥΣΗΣ

Συγγραφέας:

Στογιαννίδης Αναστάσιος
Τμήμα Επιστήμης Υπολογιστών

Εισηγητική Επιτροπή:

Στέλιος Ορφανουδάκης, Καθηγητής, Επόπτης

Απόστολος Τραγανίτης, Αναπληρωτής Καθηγητής, Μέλος

Πάνος Τραχανιάς, Επίκουρος Καθηγητής, Μέλος

Δεκτή:

Πάνος Κωνσταντόπουλος, Αναπληρωτής Καθηγητής
Πρόεδρος Επιτροπής Μεταπτυχιακών Σπουδών

Ηράκλειο, Νοέμβριος 1995

Ευχαριστίες

Η σύνθεση και κατάθεση της παρούσας εργασίας δε θα ήταν εφικτή χωρίς την πρακτική βοήθεια και την ψυχολογική συμπαράσταση πολλών ατόμων. Νιώθω την ανάγκη και οφείλω να τους ευχαριστήσω όλους θερμά. Στις ευχαριστίες αυτές, όμως, θα περιοριστώ για ευνόητους λόγους σε όσους σχετίστηκαν άμεσα με τον πρακτικό μόνο τομέα ενίσχυσης μου.

Οι αλγόριθμοι που παρουσιάζονται στο κείμενο στηρίζονται σε μια θεωρία που θεμελιώθηκε και μου υποδείχτηκε από τη Dr. Cornelia Fermüller και τον καθηγητή Γιάννη Αλοίμονο. Τους ευχαριστώ τόσο για την αρχική προσανατόλισή μου όσο και για τις μετέπειτα υποδείξεις τους.

Ευχαριστώ, επίσης, τον κύριο Στέλιο Ορφανουδάκη, επιβλέποντα καθηγητή της εργασίας μου, για τις συμβουλές, τη συνεργασία και την υποστήριξή του κατά τη διάρκεια της μεταπτυχιακής μου θητείας και ιδιαίτερα κατά τους τελευταίους μήνες εκπόνησης της εργασίας.

Πολύτιμη, επίσης, υπήρξε η συμβολή των καθηγητών Πάνου Τραχανιά και Απόστολου Τραγανίτη. Αμφότεροι ως μέλη της εξεταστικής επιτροπής και ειδικότερα ο πρώτος ως γνώστης του αντικειμένου της εργασίας συνετέλεσαν στην πιο ολοκληρωμένη συμπλήρωση του κειμένου.

Επιπλέον, οι συνεργάτες μου στο εργαστήριο Μηχανικής Ορασης και Ρομποτικής, Μανόλης Λουράκης και Αντώνης Αργυρός, δικαιούνται ένα ολόψυχο και μεγάλο ευχαριστώ, καθώς πάντοτε επεδείκνυαν ιδιαίτερη προθυμία στο να με βοηθούν και να με συμβουλεύουν για οποιοδήποτε θέμα προέκυπτε.

Οσον αφορά στη συγγραφή της εργασίας, οφείλω να αποδώσω τα εύσημα και να εκφράσω τις ευχαριστίες μου τόσο στον Γιώργο Ξουρή που διοχέτευσε το καλλιτεχνικό του ταπεραμέντο στο σχεδιασμό πολλών σχημάτων του κειμένου, όσο και στο Σταύρο Τρυπάκη που έφερε σε πέρας με ιδιαίτερη επιτυχία τις αποστολές καταγραφής επιμήκων μαθηματικών τύπων και μετάφρασης της περίληψης από την ελληνική στην αγγλική γλώσσα.

Τέλος, κατά τη διάρκεια της προετοιμασίας μου για την παρουσίαση της εργασίας δέχτηκα την αμέριστη συνδρομή πολλών συναδέλφων. Εκτιμώ βαθύτατα γι' αυτό και ευχαριστώ, λοιπόν, τους Γιάννη Τζίτζικα, Γιάννη Σουρλατζή, Βαρβάρα Μαστή, Μιράντα Κυπριώτη, Αθηνά Τραψιώτη, Θοδωρή Χατσιούλη, Δέσποινα Βαμβακά, Δημήτρη Δασκαλάκη και Χριστίνα Γκριτζάπη.

Αλγόριθμοι χρήσης προτύπων αλλαγής προσήμων για τον εντοπισμό περιοχών FOE και AOR στην περίπτωση ιδίας κίνησης

Στογιαννίδης Αναστάσιος

Μεταπτυχιακή Εργασία

Τμήμα Επιστήμης Υπολογιστών
Πανεπιστήμιο Κρήτης

ΠΕΡΙΛΗΨΗ

Ενα μεγάλου ενδιαφέροντος πεδίο εφαρμογών της Μηχανικής Ορασης αποτελεί η πλοήγηση αυτόνομων μηχανικών συστημάτων, τα οποία κινούνται σε ποικίλα περιβάλλοντα προσπαθώντας να επιτύχουν τους προδιαγεγραμμένους τους στόχους. Ανάμεσα στις ικανότητες που τέτοια συστήματα καλούνται να επιδείξουν περιλαμβάνονται ο υπολογισμός ιδίας κίνησης, η αποφυγή εμποδίων, η αναγνώριση ανεξάρτητης κίνησης, ο προγραμματισμός δράσεων για την επίτευξη στόχου και η χαρτογράφηση χώρων. Ο υπολογισμός ιδίας κίνησης, ειδικότερα, αποτελεί μία από τις περισσότερο θεμελιώδεις ικανότητες ενός αυτόνομου συστήματος, καθώς ο συνεχής έλεγχος των κινήσεων του όχι απλά χρησιμεύει για να αναφέρεται στο περιβάλλον του αλλά και επιβάλλεται για την ευχερέστερη σύνθεση ανώτερων λειτουργιών. Στα πλαίσια αυτής ακριβώς της εργασίας κατασκευάστηκαν και μελετήθηκαν τέσσερις αλγόριθμοι προσεγγιστικού προσδιορισμού των παραμέτρων ιδίας κίνησης του παρατηρητή από τις διαδοχικές εικόνες που καταγράφει η κάμερά του σε στατικά περιβάλλοντα. Οι παράμετροι αυτές στη γενική περίπτωση κίνησης στο τριδιάστατο χώρο είναι έξι, από τις οποίες τρεις εκφράζουν τη μεταφορά και τρεις την περιστροφή του αυτόνομου συστήματος.

Η προσέγγιση που ακολουθήθηκε βασίζεται στην εξαγωγή πληροφοριών από το πεδίο κάθετης ροής, το οποίο ουσιαστικά αναπαριστά τη μετατόπιση των σημείων των ακμών της σκηνής στις κάθετες προς τις ακμές διευθύνσεις. Παρά τη μη πλήρη πληροφορία κίνησης που εμπεριέχεται στο πεδίο κάθετης ροής, η χρήση του προτιμήθηκε έναντι εκείνου της οπτικής ροής (που αντανακλά την ολοκληρωμένη μετατόπιση των σημείων), διότι ο υπολογισμός του πρώτου εξάγεται με απλούστερες μεθόδους. Οι αλλαγές στα πρόσημα των μεταφορικών και περιστροφικών συνιστωσών των διανυσμάτων κάθετης ροής δημιουργούν γεωμετρικούς περιορισμούς για τη θέση στην εικόνα δύο ιδιαίτερων σημείων, του Focus of Expansion (FOE) και του Axis of Rotation (AOR), τα οποία χαρακτηρίζουν και χαρακτηρίζονται από τις παραμέτρους κίνησης. Συγκεκριμένα, το FOE είναι το σημείο που το διάνυσμα μεταφοράς τέμνει το επίπεδο προβολής (εικόνα), ενώ το AOR είναι το σημείο συνάντησης του επιπέδου προβολής με το διάνυσμα περιστροφής.

Οι αλγόριθμοι που κατασκευάστηκαν στοχεύουν στην οριοθέτηση μικρών περιοχών στην εικόνα όπου μπορούν να βρίσκονται τα FOE και AOR. Η εκμετάλλευση των γεωμετρικών ιδιοτήτων της κάθετης ροής συντελείται με τη θέσπιση απλών κριτηρίων που ελέγχουν την εμφάνιση συγκεκριμένων προτύπων στην εικόνα. Η απλότητα των κριτηρίων είναι ιδιαίτερα σημαντική, διότι οδηγεί σε γρήγορους υπολογισμούς, οι οποίοι είναι επιβεβλημένοι για τη γρήγορη απόκριση του συστήματος. Τα κριτήρια ελέγχου προτύπων εφαρμόζονται σε τρεις ομάδες διανυσμάτων κάθετης ροής, στα διανύσματα συγκεκριμένου κάθε φορά μέτρου και διεύθυνσης και στα διανύσματα που ανήκουν στα επονομαζόμενα coaxis και copoint σύνολα που αποτελούν δύο ειδικές κλάσεις διανυσμάτων του επίπεδου χώρου.

Η λειτουργία των αλγορίθμων εξετάστηκε σε δοκιμές με συνθετικά -κυρίως- και πραγματικά δεδομένα. Η εύρεση περιοχών όπου τα FOE και AOR ανήκουν ήταν επιτυχής. Το μέγεθος αυτών των περιοχών ήταν αρκετά μικρό στην περίπτωση των συνθετικών εικόνων, αλλά παρουσιαζόταν αυξημένο στην περίπτωση των πραγματικών εικόνων, αύξηση που οφείλεται κυρίως στη μείωση της διαθέσιμης ωφέλιμης πληροφορίας που εξάγεται από το πεδίο κάθετης ροής.

Επόπτης : Στέλιος Ορφανουδάκης, Καθηγητής
Τμήμα Επιστήμης Υπολογιστών, Πανεπιστήμιο Κρήτης

Algorithms using sign–alternation patterns to locate FOE and AOR areas in the case of egomotion

Stogiannidis Anastasios

Master of Science Thesis

Department of Computer Science
University of Crete

ABSTRACT

A very promising field in the domain of machine vision applications is the navigation of autonomous mechanical systems. These systems move in various environments trying to achieve a set of prescribed goals, and their capabilities include egomotion estimation, obstacle avoidance, independent motion detection, motion planning and space mapping. In particular, egomotion estimation constitutes one of the most principal tasks of an autonomous system, since the continuous control of its own motion, not only is used for self–reference with respect to the environment, but also is indispensable for making easier the composition of higher level functions. The contribution of this work consists in the implementation and study of four algorithms which compute approximately egomotion parameters from successive images, which the moving observer–camera records in static environments. There are, in the general case of three–dimension motion, six egomotion parameters, three of which represent the translation of the autonomous system, while the rest three represent its rotation.

The approach followed is based on extracting information from the normal flow field, which actually represents the displacement of the points of the edges in the scene, with respect to the direction perpendicular to the edges. Despite the fact that the motion information contained in the the normal flow field is not complete, the normal flow field has been preferred to the optical flow one which represents the complete displacement of the points. The reason for this choice is that the computation of the normal flow field is acheived by much easier methods. The sign alternations of the translative and rotative components of the normal flow vectors yield a set of geometrical constraints to

the position on the image of two special points, FOE and AOR, which characterize the movement parameters, and are characterized by them. Particularly, FOE and AOR are the intersections of the image plane with the vectors which represent respectively the translation and the rotation of the observer.

The goal of the designed algorithms is to specify small areas in the image that contain FOE and AOR. Exploiting the geometrical properties of normal flow consists in specifying simple criteria which test the appearance of specific patterns in the image. The simplicity of these criteria is of particular importance, since it allows fast computation, critical to the rapid response of the system. The patterns–search criteria are applied to three groups of normal flow vectors, namely, vectors of constant direction and constant norm, as well as vectors belonging to the so–called coaxis and copoint sets, constituting two particular classes of plane vectors.

The performance of the algorithms has been tested with synthetic – mainly – as well as real data. The regions where FOE and AOR belong have been successfully discovered. The size of these regions was small enough in the case of synthetic images, however, it appeared larger with real ones. This increase is mainly due to the diminution of available useful information extracted from the normal–flow field.

Supervisor : Stelios Orphanoudakis, Professor
Department of Computer Science, University of Crete

Περιεχόμενα

Κατάλογος Σχημάτων

Κεφάλαιο 1

Εισαγωγή

1.1 Γενική Θεώρηση

Η όραση αποτελεί τη σημαντικότερη των αισθήσεων του ανθρώπου. Αυτή είναι υπεύθυνη για τη συλλογή τη πλειονότητας των πληροφοριών που εκείνος κατέχει. Η κατανόηση των κανόνων που τη διέπουν καθώς και η σχέση της με τη λογική και τη σκέψη αποτέλεσε όχι μόνο αντικείμενο θεωρητικής, φιλοσοφικής αναζήτησης, αλλά και πεδίο έντονων επιστημονικών ερευνών. Παράλληλα, όμως, με τη μεταβολή των αντιλήψεων για τις αρχές που ορίζουν την όραση τόσο ως φυσικό φαινόμενο όσο και ως γνωστική λειτουργία, αναθεωρείται με το πέρασμα των χρόνων και η λογική χρησιμοποίησής της από τον άνθρωπο και τα υπολογιστικά συστήματα.

Οι ολοένα μεγαλύτερες απαιτήσεις του ανθρώπινου είδους για νέες τεχνολογικές εξελίξεις που θα διευκολύνουν τη ζωή του προώθησαν (και προωθούν) τη βελτίωση της ποιότητας των υπολογιστικών συστημάτων και διεύρυναν (και διευρύνουν) το φάσμα εφαρμογών τους. Ανάμεσα σ' αυτές συγκαταλέγονται και εκείνες που αφορούν στη ρομποτική και ειδικότερα στην κατασκευή μηχανών που προσομοιώνουν κάποιο είδος "έννοης" συμπεριφοράς που συναρτάται από τους ειδικότερους στόχους κάθε εφαρμογής. Η Μηχανική Οραση διαδραματίζει κεφαλαιώδη ρόλο (και) σε αυτόν τον τομέα.

Ενα μεγάλο πεδίο ερευνών αλλά και πρακτικών εφαρμογών της ρομποτικής είναι εκείνο της κατασκευής αυτόνομων μηχανικών συστημάτων τα οποία θα είνα ικανά να κινούνται σε δυναμικά περιβάλλοντα αλληλεπιδρώντας με αυτά ανάλογα με τις πληροφορίες που θα συλλέγουν. Οι κατασκευαστές των αυτόνομων αυτών συστημάτων οφείλουν να τα εφοδιάζουν με ισχυρούς μηχανισμούς και λειτουργίες αντίληψης, νόησης και δράσης, ώστε να είναι ικανά να ανταποκρίνονται σε πραγματικό χρόνο στις συχνές και απρόβλεπτες μεταβολές των χώρων δράσης τους [19]. Η έρευνα στη Μηχανική Οραση απαλλαγμένη πλέον από τους αφορισμούς παρελθόντων χρόνων, κατά τους οποίους το έργο της δε χρησίμευε παρά μόνο στη λειτουργία της αντίληψης, εξαπλώνεται την τελευταία δεκαπενταετία στη λειτουργία τόσο της νόησης όσο και της δράσης με βασικό κριτήριο την ποιοτικότερη και αποτελεσματικότερη χρησιμοποίηση των διαθέσιμων πόρων, των ανακατασκευαστικών μεθόδων και των ικανοτήτων δράσης του υπολογιστικού συστήματος.

Η βασιζόμενη στα δεδομένα, που καταγράφουν οι αισθητήρες του, κίνηση του συστήματος στο χώρο αναφέρεται και ως πλοήγηση. Οπτική πλοήγηση είναι η διαδικασία ελέγχου των κινήσεων μέσω της ανάλυσης των λαμβανομένων εικόνων. Η ορθή (οπτική) πλοήγηση ενός τέτοιου συστήματος περιλαμβάνει τη θεμελίωση ικανοτήτων που θα του επιτρέπουν να κινείται απρόσκοπτα στο χώρο. Τέτοιες ικανότητες αποτελούν η εξακρίβωση της ιδίας κίνησης, η αποφυγή εμποδίων, η αναγνώριση της κίνησης ανεξάρτητων αντικειμένων, η παρακολούθηση κινούμενων αντικειμένων, η διαδικασία επιστροφής στη βάση, η χαρτογράφηση χώρων κ.α.

Οι ικανότητες αυτές μπορούν να ταξινομηθούν ανάλογα με την πολυπλοκότητα των μαθηματικών που τα μοντελοποιούν σε ένα ιεραρχικό μοντέλο αξιολόγησης, σύμφωνα με το οποίο η επίλυση των προβλημάτων που αφορούν στις υψηλότερες βαθμίδες επιχειρείται έπειτα από την αντίστοιχη θεώρηση και εξαγωγή αποτελεσμάτων από τις χαμηλότερες, απλούστερες αλλά και βασικότερες βαθμίδες [11]. Π.χ. η αναγνώριση της κίνησης ενός ανεξάρτητου κινούμενου αντικειμένου στην περίπτωση μη στατικού παρατηρητή ευχεραίνεται σημαντικά, εάν πρώτα το σύστημα-παρατηρητής αναγνωρίσει τις δικές του παραμέτρους κίνησης και τις αφαιρέσει από τα γενικά διανύσματα κίνησης όλων των σημείων του χώρου, έτσι ώστε να απομονωθεί η κίνηση που οφείλεται στο άλλο αντικείμενο.

1.2 Το πρόβλημα ιδίας κίνησης

Μ' αυτήν τη λογική, η ικανότητα υπολογισμού της ιδίας κίνησης αποδεικνύεται να είναι θεμελιωδέστατη για ένα αυτόνομο σύστημα, μια και ο ακριβής έλεγχος των κινήσεων και της θέσης του μέσα στο χώρο είναι απαραίτητος για όλες τις υψηλόβαθμες ικανότητες του. Συνεπώς, απαιτούνται εύρωστοι και ευέλικτοι αλγόριθμοι που να δίνουν λύση στο πρόβλημα του υπολογισμού της ιδίας κίνησης, τον προσδιορισμό, δηλαδή, των παραμέτρων που εκφράζουν την αλλαγή θέσης ενός παρατηρητή.

Το μοντέλο κίνησης που εξετάζεται είναι εκείνο της ***συμπαγούς κίνησης*** (*rigid motion*). Συμπαγής κίνηση σημαίνει ότι όλα τα σημεία που αποτελούν κάποιο κινούμενο σώμα κινούνται με την ίδια ταχύτητα, τόσο σε μέτρο όσο και φορά. Η παραδοχή αυτή απλοποιεί αισθητά τα μαθηματικά του μοντέλου κίνησης χωρίς μάλιστα να χαρακτηρίζει σπάνιες και εξεζητημένες καταστάσεις, αφού τα περισσότερα αντικείμενα είναι εκ της κατασκευής τους συμπαγή. Επίσης, το περιβάλλον στις περιπτώσεις υπολογισμού ιδίας κίνησης θεωρείται στατικό, τίποτα δηλαδή δεν κινείται εκτός από την κάμερα-παρατηρητή. Τότε, η κίνηση αυτή του παρατηρητή ισοδυναμεί με μια ακριβώς αντίθετη προς τη δική του μετατόπιση όλων των σημείων του χώρου. Αν και η συνθήκη αυτή δε φαίνεται ιδιαίτερα ρεαλιστική, υπάρχουν περιπτώσεις εφαρμογών όπου πραγματικά το μόνο αντικείμενο που κινείται στο χώρο είναι ο παρατηρητής–κάμερα, π.χ. βραδυνή εποπτεία εμπορικών καταστημάτων, ενδοοικιακή μετακίνηση ατόμων με ειδικές ανάγκες κ.λπ. Ωστόσο, αν κάποιο μικρό αντικείμενο κινείται με μικρή ταχύτητα ανεξάρτητα από τον παρατηρητή, θεωρείται ότι η κίνηση αυτή δεν επηρεάζει ουσιαστικά τό μέγεθος που αναπαριστά τη μετατόπιση των σημείων, επομένως δεν επηρεάζεται και η μέθοδος υπολογισμού της ιδίας κίνησης που περιγράφεται παρακάτω. Σε περιπτώσεις, όμως, όπου πολλά αντικείμενα μετακινούνται ανεξάρτητα από τον παρατηρητή, η μέθοδος αυτή, προφανώς, δε θα αποφέρει ικανοποιητικά αποτελέσματα.

Ο υπολογισμός της ιδίας κίνησης ανάγεται πρακτικά στον υπολογισμό των έξι παραμέτρων που χαρακτηρίζουν οποιαδήποτε συμπαγή κίνηση κάποιου αντικειμένου. Οι τρεις από αυτές τις παραμέτρους εκφράζουν τη μεταφορά και οι υπόλοιπες τρεις εκφράζουν την περιστροφή του κινούμενου παρατηρητή. Το πρόβλημα υπολογισμού των παραμέτρων έχει εξεταστεί αρκετές φορές στο παρελθόν [32] [10] [20] [31] [9] [11] [25]. Ενδεικτικά, οι Horn και Weldon ασχολούνται με τον προσδιορισμό της μεταφορικής κίνησης [32], οι Aloimonos και Duric, οι Sinclair,Blake και Murray και οι Burger και Bhanu ασχολούνται με την περίπτωση στην οποία η περιστροφή κυμαίνεται σε γνωστά μικρά όρια [10] [20] [31], ενώ η Fermuller εξετάζει τη γενική περίπτωση κίνησης [11].

1.3 Προσέγγιση της παρούσας εργασίας

Υιοθετώντας την προσέγγιση της Fermuller, στην παρούσα εργασία ο προσδιορισμός των έξι παραμέτρων επιδιώκεται μέσω της αναζήτησης και οριοθέτησης περιοχών όπου ανήκουν δύο ιδιαίτερα σημεία της εικόνας, το ***Εστιακό Σημείο Εξάπλωσης*** (***ΕΣΕ***) (*Focus of expansion*) (*FOE*) και το ***Σημείο Αξονα Περιστροφής*** (***ΣΑΠ***) (*Axis of Rotation*) (*AOR*). Το FOE είναι το σημείο που το διάνυσμα που εκφράζει τη μεταφορική κίνηση τέμνει την εικόνα, ενώ το AOR είναι το σημείο που το διάνυσμα της περιστροφικής συνιστώσας της γενικής κίνησης τέμνει την εικόνα. Για την ανίχνευση πληροφοριών που οδηγούν στην εύρεση των FOE και AOR χρησιμοποιείται το πεδίο κάθετης ροής αντί για εκείνο της οπτικής, μια και η εξαγωγή του πρώτου είναι γρηγορότερη και ευκολότερη. Η επεξεργασία των διανυσμάτων του πεδίου κάθετης ροής δεν ακολουθεί την πάγια τακτική της αποπεριστροφής (derotation) και απομεταφοράς (detranslation) [31], αλλά εστιάζει την προσοχή της στην ανακάλυψη περιοχών της εικόνας όπου συναντώνται συγκεκριμένα πρότυπα αλλαγής προσήμων των διανυσμάτων του πεδίου. Τα πρότυπα αυτά καθορίζονται από τις παραμέτρους κίνησης, επομένως παρέχουν πληροφορίες για τη θέση των FOE και AOR. Η ανίχνευση των προτύπων γίνεται με όσο το δυνατόν απλούστερα κριτήρια, ώστε να μην επιβαρύνεται ο χρόνος απόκρισης του συστήματος. Το τελικό αποτέλεσμα των αλγορίθμων είναι η οριοθέτηση μιας μικρής περιοχής στην εικόνα όπου ανήκουν τα FOE και AOR.

1.4 Οργάνωση κειμένου

Στο δεύτερο κεφάλαιο παρουσιάζεται μια γενική περιγραφή των εννοιών της Μηχανικής Ορασης και του ανακατασκευαστικού και τελεολογικού μοντέλου όρασης. Στο τρίτο κεφάλαιο συνοψίζονται επιγραμματικά οι κυριότερες μέθοδοι υπολογισμού κίνησης, επισημαίνονται κάποια μειονεκτήματά τους, και ακολουθεί μία γενική ανάλυση της ιδέας των αλγορίθμων που υλοποιήθηκαν. Στο τέταρτο κεφάλαιο ορίζονται και εξηγούνται οι έννοιες των πεδίων οπτικής και κάθετης ροής, ορίζεται το σύστημα συντεταγμένων μας και εξάγονται οι εξισώσεις του πεδίου οπτικής ροής. Οι εξισώσεις του πεδίου κάθετης ροής αναλύονται σχολαστικά στο πέμπτο κεφάλαιο. Παρουσιάζεται η αλυσιδωτή σχέση της τριδιάστατης κίνησης με το πεδίο κάθετης ροής, της ροής με τα σημεία FOE και AOR και των σημείων με τα επαναλαμβανόμενα πρότυπα εναλλαγής προσήμων των διανυσμάτων. Οι υλοποιήσεις των αλγορίθμων που βασίζονται σε αυτές τις ιδέες καταθέτονται συνοπτικά στο προτελευταίο κεφάλαιο, το οποίο ολοκληρώνεται με την παρουσίαση των αποτελεσμάτων. Τέλος, στο έβδομο κεφάλαιο της εργασίας παρουσιάζονται τα τελικά συμπεράσματα και κάποιες σκέψεις για την επέκταση και συμπλήρωση των αλγορίθμων.

Κεφάλαιο 2

Μηχανική Ορaση

Μια από τις σημαντικότερες ικανότητες του ανθρώπου είναι εκείνη της όρασης, η αίσθηση η οποία είναι κατά το μεγαλύτερο ίσως βαθμό υπεύθυνη για την αντίληψη του κόσμου που μας περιβάλλει καθώς και για τη δράση μας μέσα σε αυτόν. Συνειδητοποιώντας από αιώνες τη χρησιμότητα και βαρύτητα της όρασης για τη ζωή του, ο άνθρωπος την κατέστησε αντικείμενο επιστημονικών παρατηρήσεων και φιλοσοφικών προβληματισμών με βασικούς στόχους να κατορθώσει να συλλάβει τους μηχανισμούς λειτουργίας της, να ανακαλύψει και να ερμηνεύσει τους νόμους που τη διέπουν και να ανιχνεύσει τη σχέση της με τη νόηση και τη συνείδηση, τις κυριότερες λειτουργίες που τον ξεχωρίζουν ανάμεσα στα υπόλοιπα είδη του ζωικού βασιλείου.

Φυσιολογική απόρροια της έντονης, διαρκούς και κατά πολλούς τομείς τελέσφορης αυτής αναζήτησης αποτέλεσε και η σταδιακή κατασκευή τεχνητών αντικειμένων και μηχανημάτων, τα οποία διεπόμενα από αρχές που σέβονται τους νόμους της όρασης επιτυγχάνουν τη βελτίωση (φακοί οπτικών παθήσεων), συμπλήρωση (φακοί μακρινών αποστάσεων), αλλά και ολική αντικατάσταση του ανθρώπινου οπτικού συστήματος (κινηματογραφικές, φωτογραφικές μηχανές). Στην τελευταία αυτή κατηγορία ανήκουν και οι ηλεκτρονικές κάμερες που χρησιμοποιούνται για την πρόσληψη παραστάσεων από το εξωτερικό περιβάλλον, όπως επίσης και για την παρεπόμενη εισαγωγή των παραστάσεων αυτών στα ποικίλα υπολογιστικά συστήματα.

Η αλματώδης ανάπτυξη των υπολογιστικών συστημάτων και η εντυπωσιακή διεύρυνση του φάσματος χρήσης τους απαίτησαν, προκάλεσαν και εν τέλει βοηθήθηκαν πέρα των άλλων και από την διαμόρφωση των ιδεών στον τομέα "Υπολογιστική Οραση": εξελιγμένες συσκευές, καινούριοι αλγόριθμοι χρησιμοποίησης τους, νέες φιλοσοφίες απαίτησής τους.

2.1 Κλασική υπολογιστική όραση

Η ***υπολογιστική ή μηχανική όραση*** (*computer or machine vision*), σύμφωνα με έναν από τους πολλούς παρεμφερείς ορισμούς της, αποτελεί τον επιστημονικό τομέα που έχει ως αντικείμενό του την κατασκευή σαφών, νοηματικών περιγραφών της δομής, της κατασκευής και των ιδιοτήτων του τριδιάστατου χώρου χρησιμοποιώντας διδιάστατες εικόνες.

Η υπολογιστική όραση γενικά θεωρείται κομμάτι της ***τεχνητής νοημοσύνης*** (*Artificial Intelligence (AI)*). Η Τεχνητή Νοημοσύνη έχει ως σκοπό της τη μελέτη της έννοιας συμπεριφοράς και την προσομοίωσή της με τη συνδρομή υπολογιστικών μέσων. Για να επιτευχθεί ο μεγαλεπίβολος αυτός στόχος, οι ερευνητές της Τεχνητής Νοημοσύνης επικέντρωσαν την προσοχή τους και κατέβαλαν τις δυνάμεις τους στην εύρεση μεθοδολογιών πρόσληψης και αναπαράστασης πληροφοριών γενικής σκοπιμότητας, έτσι ώστε να δημιουργήσουν μοντέλα "σοφότερα" και ικανότερα. Η τακτική αυτή, όπως ήταν επόμενο, υιοθετήθηκε και ακολουθήθηκε για αρκετά χρόνια και από τους ασχολούμενους με τη Μηχανική Οραση.

Αποτέλεσμα αυτής της ευθυγράμμισης στη μεθοδολογία των δύο επιστημών ήταν η επικράτηση της φιλοσοφικής γραμμής που όριζε ότι μέσω της Μηχανικής Ορασης προσπαθούμε να αναπαραστήσουμε το περιβάλλον με όση μεγαλύτερη λεπτομέρεια γίνεται. Η θεωρία αυτή, κύριος εκφραστής της οποίας υπήρξε ο Marr, όριζε ότι η όραση αποτελεί μια διαδικασία ανακατασκευής, εξού και ο όρος ***ανακατασκευαστική όραση*** (*reconstructive vision*), διαμέσου της οποίας αντιμετωπίζεται το πρόβλημα της δημιουργίας αναπαραστάσεων αυξανόμενης αφαιρετικότητας, ξεκινώντας από τα πρωταρχικά στοιχεία των διδιάστατων εικόνων και καταλήγοντας στην πληρέστερη δυνατή περιγραφή των αντικειμένων, έτσι ώστε να καταστεί εφικτή η αποτελεσματική απόκριση του συστήματος σε ερωτήσεις που εξετάζουν το σύνολο των ιδιοτήτων αυτών [19].

Η εκτεταμένη ενασχόληση των ερευνητών με τα προβλήματα αναγνώρισης και αναπαράστασης των ιδιοτήτων των αντικειμένων του περιβάλλοντος οδήγησε στη σύλληψη και υλοποίηση έξυπνων και γρήγορων μεθόδων συλλογής και σχηματοποίησης των πληροφοριών που οι αισθητήρες παρέχουν στο σύστημα, καθώς επίσης και στην εμφάνιση και θεμελίωση νέων σχετικών περιοχών έρευνας, όπως η μηχανική μάθηση και οι γνωσιολογικές βάσεις δεδομένων.

Ομως, παρά τα σημαντικά αυτά επιτεύγματα, σταδιακά ενισχύθηκε η αντίληψη ότι η Μηχανική Οραση δεν θα πρέπει να αποτελεί ένα πεδίο ερευνών αποκομμένο από τους στόχους της Τεχνητής Νοημοσύνης. Κριτικοί της Τεχνητής Νοημοσύνης τόνιζαν ότι η μελέτη της νοημοσύνης δεν επιτρέπεται να διαχωρίζεται από τη μελέτη των μεθόδων επικοινωνίας μιας οντότητας με το περιβάλλον του. Ψυχολόγοι, μετά από πειράματα, αποφαίνονταν ότι στην πραγματικότητα ακόμα και στον άνθρωπο οι διαδικασίες της όρασης και της κατανόησης δεν μπορούν να θεωρούνται η μία έξω από την άλλη [33]. Η Μηχανική Οραση έπρεπε να σταματήσει να ασχολείται αποκλειστικά με την ικανοποιητικότερη αναπαράσταση της πληροφορίας και όφειλε να επεκταθεί σε ζητήματα λήψης αποφάσεων ώστε να διευκολυνθεί το ίδιο της το έργο [4] [5].

2.2 Τελεολογική - Ενεργός Οραση

Τις απόψεις των κριτικών της Τεχνητής Νοημοσύνης ήρθαν να ενισχύσουν εργασίες, που αποδείκνυαν ότι πολλά από τα προβλήματα που θεωρούνταν μη γραμμικά ή ill-posed είναι δυνατό να επιλυθούν με την προϋπόθεση ότι η κάμερα-παρατηρητής κινείται, αφού οι ελεγχόμενες αυτές μετακινήσεις εμπλουτίζουν το σύστημα με νέες, χρήσιμες και αποσαφηνιστικές πληροφορίες για το περιβάλλον [6]. Αλλά δεν είναι μόνο η διαπίστωση ότι ένας μετακινούμενος, και άρα ενεργός μέσω και της όρασης, παρατηρητής διαθέτει περισσότερα πλεονεκτήματα από έναν στατικό ανάλογο. Με έναυσμα τα αποτελέσματα μελετών σε διάφορους ζωικούς οργανισμούς, αλλά και αναλογιζόμενοι την περίπτωση του ανθρώπου, στο θεωρητικό τομέα προέκυψαν νέα ερωτήματα σχετικά με το πραγματικό μέγεθος και είδος της πληροφορίας που απαιτείται για να αλληλεπιδρά σωστά ένα έννοο σύστημα με το περιβάλλον του.

Από την κοινή εμπειρία γνωρίζουμε ότι το ανθρώπινο οπτικό σύστημα λειτουργεί με μεθόδους περισσότερο ποιοτικές παρά ποσοτικές. Π.χ. ένας άνθρωπος σχεδόν ποτέ δεν υπολογίζει ακριβώς, διότι δεν μπορεί και διότι έμαθε να μη του χρειάζεται, την απόστασή του από τα περιβάλλοντα αντικείμενα. Το μόνο που κάνει είναι να εκτιμά προσεγγιστικά την απόσταση αυτή, να τα κατατάσει σε γενικές κατηγορίες (δίπλα,κοντά,μακριά) ή ακόμα να αναφέρεται σε αυτά με βάση τη σχετική τους θέση ως προς άλλα αντικείμενα του ίδιου χώρου. Παρόμοιες αναφορές και συγκρίσεις χρησιμοποιούνται συνέχεια από την ανθρώπινη όραση για την ανάλυση των περισσοτέρων χαρακτηριστικών των αντικειμένων, για την ταχύτητα, το σχήμα, την υφή. Επίσης, πολύ συνηθισμένες είναι οι λειτουργίες της επιλεκτικής αποδοχής πληροφοριών και του εστιασμού της προσοχής. Σε μια διαδικασία αναζήτησης ο άνθρωπος επικεντρώνει την προσοχή του αποκλειστικά σχεδόν στα αντικείμενα που τον ενδιαφέρουν. Αγνοεί το πλήθος των αδιάφορων οπτικών -και άλλων- ερεθισμάτων που δέχεται και εκμεταλλεύεται μονάχα εκείνα που του παρέχουν πληροφόρηση ουσιαστική, άμεσα συνδεδεμένη με το σκοπό που έχει εκ των προτέρων θέσει.

Ακολουθώντας μια τέτοια χαλαρωτική, λιγότερο απαιτητική, αλλά όχι αναποτελεσματική λογική και στη Μηχανική Οραση, την τελευταία δεκαπενταετία μορφοποιήθηκε μια νέα θεώρηση του ρόλου και των μεθόδων της. Βασικά στοιχεία τής ***Ποιοτικής ή Τελεολογικής***, όπως ονομάζεται, Ορασης (*Qualitative or Purposive Vision*) είναι [6] [4] [5]:

- ο καθορισμός συγκεκριμένων έργων ή στόχων που το σύστημα καλείται να επιτύχει.
- η συλλογή εκείνων μόνο των πληροφοριών που θεωρούνται χρήσιμες σε σχέση με το πεδίο εφαρμογής ή ειδικότερα το συγκεκριμένο έργο.
- η ενεργοποίηση των αισθητήρων εκείνων που παρέχουν την κατάλληλη πληροφορία για την επίτευξη ενός στόχου.
- η μετάφραση των πρωτόγονων δεδομένων σε αναπαραστάσεις υπό το σκεπτικό της αποτελεσματικότερης εξυπηρέτησης των ειδικών αλγορίθμων επεξεργασίας τους.
- η δημιουργία αλγορίθμων ποιοτικών και όχι ποσοτικών όπου η ακρίβεια στις μετρήσεις και τα ενδιάμεσα αποτελέσματα δεν είναι σημαίνουσα.

2.3 Πλοήγηση

Ενα από τα πιο ενδιαφέροντα πεδία εφαρμογών της Μηχανικής Ορασης είναι εκείνο της ***πλοήγησης*** (*navigation*) κινούμενων μηχανών (mobile robots). Τα mobile robots αποτελούν (ή τουλάχιστον προορίζονται για να αποτελέσουν) "έξυπνες" μηχανές που κινούνται σε δυναμικά περιβάλλοντα προσπαθώντας να επιτύχουν τους προδιαγεγραμμένους τους στόχους. Για να αποκτήσουν ικανοποιητική αυτονομία ώστε να καταφέρνουν να δρουν αποτελεσματικά σε μη στατικά περιβάλλοντα που μεταβάλλονται συχνά και απρόβλεπτα, πρέπει να εφοδιάζονται με ισχυρές μεθόδους ***αντίληψης*** (*perception*), ***νόησης*** (*reasoning*) και ***δράσης*** (*acting*), λειτουργίες που η ορθή υποστήριξή τους βασίζεται στο σωστό σχεδιασμό και συντονισμό τόσο των μηχανικών μερών, όσο και της αρχιτεκτονικής αλλά και των αλγορίθμων που χρησιμοποιούνται από το σύστημα [19].

Ανάμεσα στις οπτικές ικανότητες, που ένα τέτοιο σύστημα καλείται να επιδείξει, περιλαμβάνονται ο προσδιορισμός της ιδίας κίνησης, η αλλαγή γωνίας λήψης της κάμερας, η εκτίμηση του βάθους των σημείων της σκηνής, η εκτίμηση της κίνησης ανεξάρτητα κινούμενων αντικειμένων, η αποφυγή εμποδίων, ο καθορισμός στόχου, η παρακολούθηση κινούμενων αντικειμένων, η κατάστρωση σχεδίου μελλοντικών κινήσεων, η διαδικασία επιστροφής στη βάση και η χαρτογράφηση μιας περιοχής. Η κατάταξη των ικανοτήτων αυτών σε ένα μοντέλο το οποίο θα διαπνέεται από τη λογική της τελεολογικής όρασης οδηγεί στη σύνθεση μιας ιεραρχίας, τα επίπεδα της οποίας καθορίζονται από την πολυπλοκότητα των μαθηματικών μοντέλων που συνοδεύουν κάθε ικανότητα [11]. Μ' αυτόν τον τρόπο, δεδομένου ενός προβλήματος συμπεριφοράς του αυτόνομου συστήματος στο χώρο, εξετάζεται αρχικά ο έλεγχος των μαθηματικά απλούστερων αλλά συνάμα λειτουργικά βασικότερων ικανοτήτων του συστήματος και στη συνέχεια ακολουθεί η δραστηριοποίηση των συνθετότερων ικανοτήτων, η οποία βασίζεται τόσο στις χαμηλότερου επιπέδου αναλύσεις όσο και στην ανάκτηση από το περιβάλλον των απαραίτητων για τον εκάστοτε στόχο πληροφοριών.

Π.χ., η ικανότητα αποφυγής εμποδίων βασίζεται σε κείνη του προσδιορισμού των σχετικών βαθών μιας σκηνής και τη γνώση ότι απλά ο χρόνος πιθανής πρόσκρουσης είναι αρκετά πάνω από το χρόνο αντίδρασης του συστήματος. Και η ικανότητα εκτίμησης ανεξάρτητης κίνησης στηρίζεται σε κείνη του υπολογισμού της ιδίας κίνησης, μια και για να "συνειδητοποιήσει" ένα σύστημα την ανεξάρτητη μετατόπιση κάποιων αντικειμένων επιβάλλεται η απαλοιφή των δικών του παραμέτρων από το γενικό πεδίο κίνησης. Σύμφωνα με αυτή τη διαβάθμιση η ικανότητα προσδιορισμού της ιδίας κίνησης κρίνεται θεμελιωδέστατη για την πλοήγηση του αυτόνομου μηχανικού συστήματος, αφού σε κάθε περίπτωση η δυνατότητα ελέγχου των κινήσεων και αναφοράς στο περιβάλλον όχι απλά χρησιμεύει αλλά επιβάλλεται.

Κεφάλαιο 3

Προσδιορισμός κίνησης

Το πρόβλημα υπολογισμού ιδίας κίνησης, γενικά υπάγεται σε εκείνο του υπολογισμού της τριδιάστατης κίνησης ενός αντικειμένου μέσω των πληροφοριών που αποκτούνται από διδιάστατες κάμερες. Οι τεχνικές που χρησιμοποιούνταν για τον προσδιορισμό της τριδιάστατης κίνησης μέχρι πριν μερικά χρόνια ακολουθούσαν τη φιλοσοφία της ανακατασκευαστικής όρασης. Οπως θα δειχτεί παρακάτω, όμως, το τελεολογικό μοντέλο όρασης μπορεί να εφαρμοστεί στο συγκεκριμένο πρόβλημα της ιδίας κίνησης, δημιουργώντας έτσι πιο γρήγορους και ευέλικτους αλγορίθμους [11].

3.1 Τεχνικές υπολογισμού κίνησης

Παλιότερες προσεγγίσεις της οπτικής πλοήγησης μηχανικών συστημάτων, επηρεασμένες από την ανακατασκευαστική θεώρηση της όρασης του Marr, ακολουθούν μια διαδικασία δύο βημάτων στην προσπάθεια ανακάλυψης της τριδιάστατης κίνησης από δυναμικές εικόνες [29]. Στο πρώτο βήμα, επιχειρείται να υπολογιστούν οι ακριβείς μετατοπίσεις των εικόνων στα διαδοχικά frames είτε με τη μορφή αντιστοίχησης χαρακτηριστικών είτε με τη μορφή πυκνών πεδίων κίνησης (optical flow). Στην περίπτωση που η μετατόπιση είναι μεγάλη, συγκεκριμένα χαρακτηριστικά των εικόνων απομονώνονται και παρακολουθούνται σε μια ακολουθία πολλών εικόνων (correspondence problem). Στην αντίθετη περίπτωση, δηλαδή όταν η μετατόπιση είναι μικρή, οι κινήσεις των εικόνων θεωρούνται ως μια τριδιάστατη συνάρτηση των δύο χωρικών και της μιας χρονικής μεταβλητής. Με βάση τις χωροχρονικές παραγώγους καθώς και παραδοχές που υιοθετούνται σχετικά με τη συνέχεια αυτής της συνάρτησης υπολογίζεται η οπτική ροή ή αλλιώς η ταχύτητα των σημείων στο επίπεδο προβολής. Στη συνέχεια, στο δεύτερο βήμα της διαδικασίας, ξεκινώντας από τις εξισώσεις που συνδέουν τις παραμέτρους κίνησης με τη διδιάστατη κίνηση των σημείων στο επίπεδο προβολής επιχειρείται ο ακριβής προσδιορισμός της πραγματικής τριδιάστατης κίνησης καθώς και η κατασκευή μιας περιγραφής για τα αντικείμενα του χώρου (structure from motion). Μέσα στα γενικά πλαίσια της τήρησης των δύο αυτών κύριων βημάτων, έχουν δημιουργηθεί και μελετηθεί πολλές μέθοδοι, οι οποίες διαφέρουν στη χρησιμοποίηση επιμέρους παραμέτρων, όπως ο τύπος των παρακολουθητέων χαρακτηριστικών στη διαδικασία αντιστοίχησης, το πλήθος των εικόνων, το μοντέλο προβολής κ.α [17] [18] [26] [14].

3.1.1 Το πρόβλημα αντιστοίχησης

Το ***πρόβλημα αντιστοίχησης*** (*correspondence problem*) συνίσταται στην εύρεση ενός ταιριάσματος μεταξύ τμημάτων μιας εικόνας και των αντίστοιχών τους σε επακόλουθες εικόνες [3]. Η αντιστοίχηση επιλύεται συνήθως σε δύο βήματα. Αρχικά, επιλέγονται πρότυπα χαρακτηριστικά από την εικόνα καθώς και ένα μέτρο ομοιότητας για την παρακολούθησή τους στα επόμενα frames. Η αντιστοίχηση στη συνέχεια υπολογίζεται ελαχιστοποιώντας μία συνάρτηση ως προς όλα τα δυνατά ταιριάσματα. Η επιλογή του είδους των χαρακτηριστικών των οποίων θα εξεταστεί η πορεία στο χρόνο εξαρτάται από τη διαθεσιμότητά τους στην εικόνα και την αξιοπιστία των μετρήσεών τους. Οταν σημεία δε συναντώνται σε μεγάλες ποσότητες, τότε χρησιμοποιούνται άλλα χαρακτηριστικά, όπως γραμμές, περιγράμματα, μεγάλες κουκίδες, γωνίες, χαρακτηριστικά δηλαδή που γενικά ξεχωρίζουν από το περιβάλλον τους [21] [28].

3.1.2 Μέθοδοι υπολογισμού οπτικής ροής

Οπτική ροή, όπως εξηγείται αναλυτικότερα στο επόμενο κεφάλαιο, είναι η φαινόμενη κίνηση των σημείων του χώρου, όπως αυτή παρουσιάζεται στις διαδοχικές εικόνες που καταγράφει η κάμερα ενός κινούμενου παρατηρητή. Ο σωστός υπολογισμός της παρέχει πολύτιμες πληροφορίες για την κίνηση του παρατηρητή, καθώς και για την κίνηση ανεξάρτητα κινούμενων αντικειμένων. Οι μέθοδοι που εφαρμόζονται για τον υπολογισμό της οπτικής ροής μπορούν να χωριστούν σε τέσσερις κατηγορίες.

- ***Τεχνικές παραγώγισης*** (*Differential techniques*). Στις τεχνικές παραγώγισης υπολογίζεται η ταχύτητα από τις χωροχρονικές παραγώγους της έντασης της εικόνας ή από φιλτραρισμένες εκδόσεις της εικόνας (χρησιμοποιώντας κατωδιαβατά (low-pass) ή ζωνοπερατά (band-pass) φίλτρα) [8] [22] [30].

- ***Ταίριασμα περιοχών*** (*Region-based matching*). Σύμφωνα με αυτές τις προσεγγίσεις, η ταχύτητα v ορίζεται ως η μετατόπιση $d = (d_x, d_y)$ που παράγει το καλύτερο ταίριασμα μεταξύ περιοχών των εικόνων σε διαφορετικές χρονικές στιγμές. Η εύρεση του καλύτερου ταιριάσματος αναλογεί στη μεγιστοποίηση ως προς d ενός μεγέθους ομοιότητας των εικόνων [24] [27].

- ***Μέθοδοι συχνοτήτων*** (*Frequency-based, Energy-based*). Σ' αυτή την ομάδα τεχνικών η ταχύτητα υπολογίζεται μέσω του πεδίου συχνοτήτων. Οι χωροχρονικές συχνότητες μιας εικόνας συνδέονται με τις συνιστώσες της κίνησης μέσω περιορισμών-εξισώσεων, η λύση των οποίων εκτιμάται ότι προσφέρει μια καλή προσέγγιση της οπτικής ροής [12].

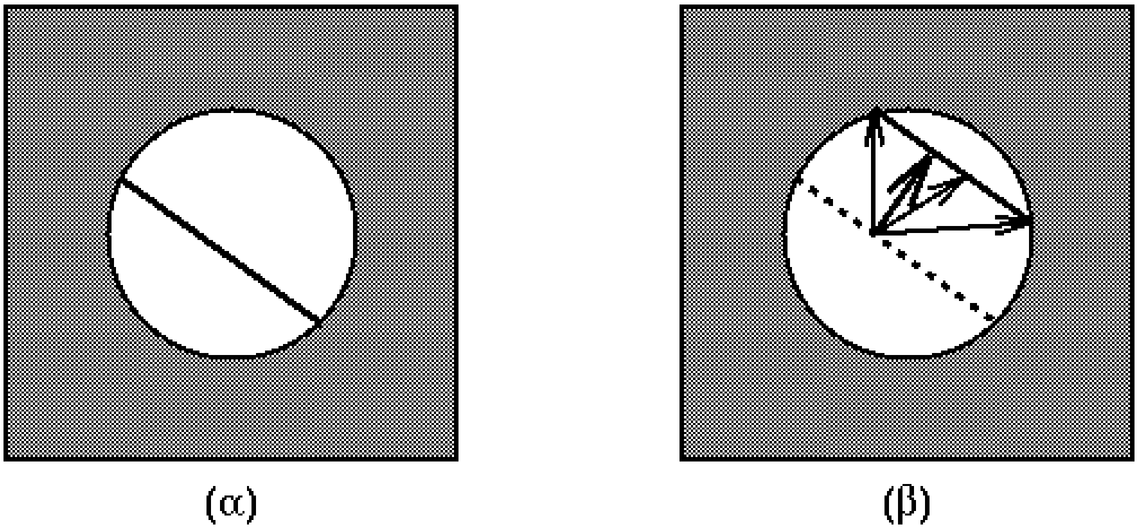
- ***Φασικές τεχνικές*** (*Phase-based techniques*). Σ' αυτήν την κλάση μεθόδων η ταχύτητα ορίζεται χρησιμοποιώντας πληροφορίες από τη συμπεριφορά φάσης των εξόδων κατωδιαβατών φίλτρων (low-pass filtering) [16] [7].

3.1.3 Προβλήματα στις μεθόδους υπολογισμού αντιστοίχησης και οπτικής ροής

Παρά το γεγονός ότι η έρευνα και ανάλυση τέτοιων μεθόδων συνοδεύεται από πολλά πειράματα, φαίνεται ότι οι τεχνικές αυτές δεν είναι κατάλληλες για να χρησιμοποιηθούν ως βάση ενός ολοκληρωμένου συστήματος που θα δρα ευέλικτα και ικανοποιητικά σε ένα γενικό περιβάλλον. Η εκτίμηση της κίνησης, όπως αντιμετωπίζεται στις μεθόδους αυτές, περιλαμβάνει βήματα των οποίων η μαθηματική μοντελοποίησή τους καταλήγει σε μη γραμμικές εξισώσεις ή σε αριθμό εξισώσεων που περιέχουν περισσότερους από αυτές αγνώστους. Η συνήθης τακτική παράκαμψης τέτοιων προβλημάτων είναι η εισαγωγή νέων παραδοχών σχετικά με το περιβάλλον και τη συνέχεια της κίνησης, παραδοχές όμως που θέτουν περιορισμούς στην ευελιξία του μηχανικού συστήματος.

Στο πρόβλημα αντιστοίχησης π.χ. εισάγονται παραδοχές όπως η ομοιότητα της κίνησης των χαρακτηριστικών μιας γειτονιάς σημείων της εικόνας, παραδοχή που σε παραδείγματα σκηνών με μεγάλες διαφορές βάθους ατονεί και κρίνεται ανεφάρμοστη.

Στην περίπτωση της οπτικής ροής, επίσης, συναντάται το περίφημο ***πρόβλημα ανοίγματος*** (*apperture problem*) (σχ. 3.1). Εστω ότι παρακολουθείται τοπικά η κίνηση κάποιας ακμής. Οταν δεν υπάρχουν ασυνέχειες και η ένταση των σημείων της παραμένει σταθερή στο χρόνο, τότε ο προσδιορισμός της πλήρους κίνησης των σημείων με τοπικές μόνο μετρήσεις καθίσταται



Σχήμα 3.1: Το πρόβλημα ανοίγματος

Τοπικά, μόνο η προβολή στην κάθετη προς την ακμή διεύθυνση μπορεί να υπολογιστεί ικανοποιητικά. Για την εξακρίβωση της πραγματικής κίνησης ενός σημείου της ακμής απαιτούνται επιπλέον γενικότερες παρατηρήσεις (π.χ. παραδοχές smoothness, εύρεση αντιστοιχιών σε δύο εικόνες, εντοπισμός ασυνεχειών κλπ.)

προβληματικός, μια και το μόνο μέγεθος που μπορούμε να παρατηρήσουμε με ασφάλεια είναι η προβολή του διανύσματος κίνησης στην κάθετη προς την ακμή διεύθυνση. Για να υπερπηδηθεί αυτό το εμπόδιο ακολουθείται μια διαδικασία ανίχνευσης σε ολόκληρη τη σκηνή στοιχείων που θα διασαφηνίσουν τη σύγχυση για την ακριβή ταυτότητα των κινήσεων (γωνίες, διαφορές στο χρώμα, smoothness). Ομως, μ' αυτόν τον τρόπο οι αλγόριθμοι επιβαρύνονται και η χρησιμοποίησή τους καθίσταται προβληματική σε εφαρμογές πραγματικού χρόνου.

3.2 Συλλογιστική των υλοποιημένων αλγορίθμων

Από τη στιγμή που ο υπολογισμός της οπτικής ροής -του πιο παραστατικού ίσως μεγέθους αναπαράστασης της τριδιάστατης κίνησης στο επίπεδο προβολής της κάμερας- απαιτεί γενικούς υπολογισμούς ευαίσθητους σε λάθη [15] [10], η προσπάθεια προσδιορισμού της τριδιάστατης κίνησης στρέφεται προς τον εντοπισμό άλλων μεγεθών που είναι ικανά να παρέχουν πληροφορίες για την κίνηση. Ενα τέτοιο μέγεθος, όπως φάνηκε και στο πρόβλημα ανοίγματος, μπορεί να είναι η κάθετη προς τη διεύθυνσή της μετατόπιση κάποιας ακμής. Το πεδίο αυτό ***κάθετης ροής***, όπως ονομάζεται, φυσικά περιέχει λιγότερη πληροφορία απ' ό,τι εκείνο της οπτικής ροής, δεν παύει, όμως, να εξαρτάται και αυτό από την τριδιάστατη κίνηση του παρατηρητή και άρα να αντανακλά μέρος -σεβαστό όπως αποδεικνύεται- αυτής της εξάρτησης στη διδιάστατη γεωμετρία του στο επίπεδο προβολής.

Αν αναλογιστεί κανείς τις περιπτώσεις απλών κινήσεων, όπου μηδενίζονται οι τιμές κάποιων παραμέτρων, και πιο ειδικά εκείνες της καθαρής μεταφορικής και της καθαρής περιστροφικής κίνησης, διαπιστώνει ότι ο υπολογισμός των μη μηδενικών παραμέτρων είναι ευκολότερος, αφού οι μαθηματικές εξισώσεις που συνδέουν την τριδιάστατη κίνηση με τη διδιάστατη προβολή της απλοποιούνται [32].

Επίσης, η εξέταση των εξισώσεων που δίνουν την ταχύτητα των σημείων της σκηνής σε μια γενική κίνηση φανερώνει μια συμμετρία ως προς τις δύο συντεταγμένες του διανύσματος της ταχύτητας: τόσο ως προς τον X, όσο και προς τον Y άξονα, οι μεταφορικές συνιστώσες της κίνησης είναι γραμμικές ως προς x και y, αντίθετα με τις περιστροφικές όπου οι όροι x και y παρουσιάζονται υψωμένοι στο τετράγωνο [13] [11]. Επιπροσθέτως, η κρίσιμη τιμή του βάθους Z ενός σημείου που μετατοπίζεται σχετικά κατά την κίνηση του παρατηρητή υπεισέρχεται μόνο στις εξισώσεις της μεταφορικής συνιστώσας.

Συνδυάζοντας τις τρεις παραπάνω παρατηρήσεις, η Fermuller εξετάζει τον τρόπο με τον οποίο οι διαφορετικές τιμές στη μεταφορική και την περιστροφική συνιστώσα μιας συμπαγούς κίνησης επηρεάζουν το πεδίο κάθετης ροής [11]. Αποδεικνύεται ότι υπάρχει μια άρρηκτη σχέση μεταξύ της διαμόρφωσης των τιμών του πεδίου κάθετης ροής και της "αντιπαλότητας" των δύο συνιστωσών. Εξετάζοντας τις γεωμετρικές κατανομές των τιμών του πεδίου κάθετης ροής εξάγονται ουσιώδη συμπεράσματα για τη θέση των χαρακτηριστικών σημείων FOE και AOR (βλ. κεφάλαιο 4) που ορίζονται από τις παραμέτρους κίνησης, άρα και για τις ίδιες τις παραμέτρους.

Οι αλγόριθμοι που υλοποιήθηκαν στην παρούσα εργασία στηρίζονται σε αυτήν ακριβώς τη θεωρία της Fermuller. Διαπνέονται από τις αρχές του τελεολογικού μοντέλου όρασης, καθώς δε δίνεται ιδιαίτερη έμφαση στις συγκεκριμένες τιμές του πεδίου κάθετης ροής, παρά μόνο τούτο ανιχνεύεται για τον εντοπισμό σ'αυτό ειδικών προτύπων. Οι θέσεις των προτύπων στην εικόνα συνηγορούν στον αποκλεισμό ή την πρόκριση περιοχών, οι οποίες είναι υποψήφιες κατοχής των κρίσιμων σημείων FOE και AOR. Τα πρότυπα αυτά αναπαριστούν εναλλαγές περιοχών με θετικό και περιοχών με αρνητικό πρόσημο, όμως, η ιδέα γενικεύεται και εκμεταλλευόμαστε οποιαδήποτε αλλαγή στην ανάμειξη των προσήμων μιας περιοχής (π.χ. από μεικτά θετικά και αρνητικά σε μόνο θετικά). Η επαναληπτική αναζήτηση διαφορετικών τέτοιων προτύπων μειώνει αισθητά την πιθανότητα λαθεμένης εκτίμησης.

Πρέπει να σημειωθεί εδώ ότι, αν και η όλη διαδικασία στοχεύει στην ανακάλυψη των σημείων FOE και AOR, ο ακριβής υπολογισμός της μοναδικής θέσης των σημείων αποτελεί ένα δύσκολο πρόβλημα, κυρίως λόγω λαθών στη ψηφιοποίηση, αναξιόπιστων διανυσμάτων μετατόπισης και θορύβου στην εικόνα. (Αξίζει να τονιστεί ακόμη ότι φαίνεται να υπάρχουν ψυχολογικές αποδείξεις ότι και οι άνθρωποι δυσκολεύονται να υπολογίσουν την ακριβή κατεύθυνση του προορισμού τους σε συγκρίσιμες περιπτώσεις.) Η λύση στο πρόβλημα είναι να μην ερευνάται η ακριβής θέση των FOE και AOR, αλλά με την προϋπόθεση αυτά να βρίσκονται μέσα στο επίπεδο προβολής να παράγονται δύο περιοχές δύο διαστάσεων στην εικόνα μέσα στις οποίες τα FOE και AOR να περικλείονται. Η παραδοχή πτώσης των FOE και AOR μέσα στην εικόνα δεν είναι εξωπραγματική. Πέραν του γεγονότος ότι μπορούν να χρησιμοποιηθούν ερυγώνιοι φακοί που "αντικρύζουν" μεγαλύτερο οπτικό πεδίο της σκηνής, δύο συνήθεις τακτικές κίνησης των mobile περιορίζουν τη θέση των FOE και AOR. Συγκεκριμένα και όσον αφορά στο AOR, η περιστροφή κυμαίνεται συνήθως μέσα σε μικρά, γνωστά όρια, τα οποία μεταφρασμένα σε συντεταγμένες του επιπέδου προβολής δεν βρίσκονται έξω από τις εικόνες που καταγράφουν οι κάμερες. Οσον αφορά στο FOE, αυτό κατά μεγάλη πιθανότητα θα βρίσκεται μέσα στην εικόνα, καθώς τα mobile (κατ' αναλογία με τον άνθρωπο) κινούνται προς την κατεύθυνση που στρέφονται οι κάμερές τους. Οι περιοχές εντοπισμού, λοιπόν, των FOE και AOR ονομάζονται αντίστοιχα fuzzy-FOE και fuzzy-AOR. Παρά την απώλεια σε ακρίβεια, οι fuzzy περιοχές αποτελούν πρακτικά εργαλεία στην ανάλυση δυναμικών σκηνικών, μια και υπάρχουν διαδικασίες, οι οποίες δεχόμενες ως είσοδο τις περιοχές του FOE και του AOR εκτελούν μια γρήγορη και εξαντλητική αναζήτηση στο χώρο των πιθανών παραμέτρων κίνησης που ορίζονται από αυτές τις περιοχές για να βρουν την εξάδα (ή τις εξάδες) που ανταποκρίνεται καλύτερα στα δεδομένα του πεδίου κάθετης ροής [1] [2] [11]. Οι διαδικασίες αυτές αφαιρούν επαναληπτικά κάποια μεταφορική και κάποια περιστροφική κίνηση που ορίζεται από το χώρο εύρεσης του FOE και του AOR, και κρατούν ως θετικά αποτελέσματα εκείνες τις τιμές των σημείων για τις οποίες οι derotation και detranslation τεχνικές ελαχιστοποιούν κάποια συνάρτηση κόστους.

Κεφάλαιο 4

Οπτική και κάθετη ροή

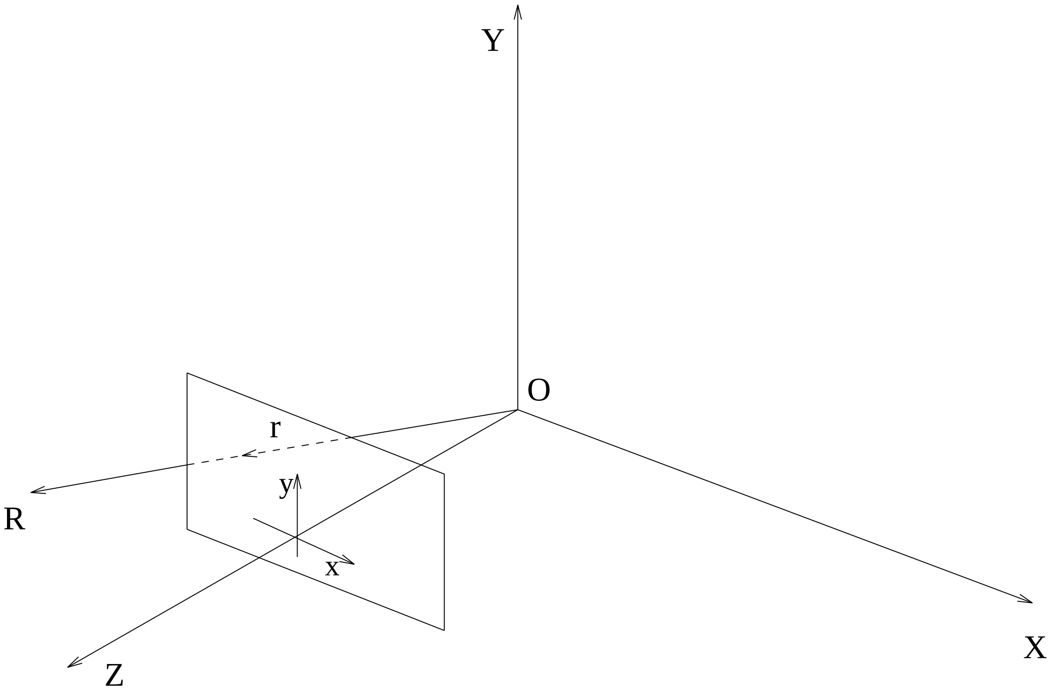
Στο κεφάλαιο αυτό περιγράφεται ένα τυπικό σύστημα τριών αξόνων για την αναφορά του κόσμου, ερμηνεύονται και συγκρίνονται συνοπτικά οι έννοιες των πεδίων ***κίνησης***, ***κάθετης κίνησης***, ***οπτικής ροής*** και ***κάθετης ροής*** και, τέλος, παράγονται οι εξισώσεις του πεδίου οπτικής ροής που συνεπάγονται από το σύστημα που ορίσαμε.

4.1 Αναφορικό σύστημα αξόνων

Θεωρούμε ότι το οπτικό μας σύστημα αποτελείται από μία κάμερα η οποία κινείται στο χώρο. Τα σημεία του τριδιάστατου χώρου αντιστοιχίζονται σε σημεία της διδιάστατης εικόνας που καταγράφει η κάμερα μέσω προοπτικής προβολής. Η αρχή O των αξόνων του δεξιόστροφου συστήματος συντεταγμένων μας OXYZ τοποθετείται στο εστιακό σημείο (nodal point) της κάμερας (σχ. 4.1). Το επίπεδο προβολής-εικόνα είναι παράλληλο με το επίπεδο OXY και τέμνει τον άξονα Z στο σημείο $(0, 0, f)$. Το μέτρο f ονομάζεται εστιακή απόσταση (focal length) της κάμερας. Το τυχαίο σημείο R (R_1, R_2, R_3) του τριδιάστατου χώρου αντιστοιχίζεται στο r (r_1, r_2, r_3) της εικόνας, όπου r είναι το σημείο που το διάνυσμα $\vec{OR}$ τέμνει το επίπεδο προβολής. Προφανώς ισχύει $r_3 = f$, ενώ από ομοιότητα τριγώνων εξάγεται ότι $r_1 = \frac{R_1 f}{R_3}$ και $r_2 = \frac{R_2 f}{R_3}$, όπου R_3 το βάθος του σημείου R.

4.2 Πεδίο κίνησης και οπτική ροή

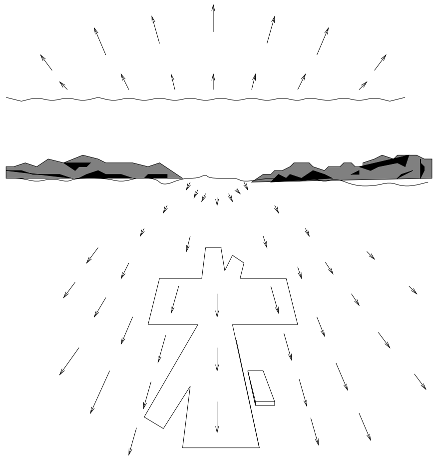
Το ***πεδίο κίνησης*** (*motion field*) ενός παρατηρητή καθώς αυτός κινείται με κάποια ταχύτητα μέσα σε μία αμετάβλητη σκηνή, αποτελεί ουσιαστικά μία εκχώρηση διανυσμάτων στα σημεία της σκηνής, έτσι ώστε κάθε διάνυσμα να αντικατοπτρίζει τη σχετική κίνηση του



Σχήμα 4.1: Το σύστημα αναφοράς

αντίστοιχου σημείου πάνω στη διδιάστατη επιφάνεια προβολής (σχ. 4.2). Στην περίπτωση ενός παρατηρητή-κάμερας η σχετική κίνηση αυτή των σημείων (που βέβαια ποσοτικά είναι αντίθετη από την κίνηση του παρατηρητή τόσο κατά το μεταφορικό όσο και κατά το περιστροφικό κομμάτι της) προκαλεί μία διαφορά στην ένταση της φωτεινότητας των σημείων των διαδοχικών εικόνων που καταγράφει η κάμερα. Οι στιγμιαίες αλλαγές της φωτεινότητας στο επίπεδο προβολής-εικόνα αναλύονται για να δημιουργηθεί το ***πεδίο οπτικής ροής*** (*optic flow field*), ένα διδιάστατο πεδίο διανυσμάτων που αντανακλά τη μετατόπιση των σημείων σε δύο διαδοχικές χρονικές στιγμές.

Τα δύο πεδία, αν και συγγενεύουν αρκετά, δεν ταυτίζονται, μια και η "υφή" τους είναι διαφορετική. Το μεν πεδίο κίνησης αναπαριστά ρεαλιστικές αλλαγές στο χώρο, το δε πεδίο οπτικής ροής αναπαριστά φαινομενικές διαφορές σε εντάσεις φωτεινότητας. Αν και συνήθως μεταβολές του πεδίου κίνησης προκαλούν ανάλογες μεταβολές στο πεδίο οπτικής ροής, η αλήθεια είναι ότι τα δύο είδη πεδίων δεν ακολουθούν υποχρεωτικά το ένα το άλλο. Αναλογιστείτε το παράδειγμα μιας δίχως texture, τέλειας σφαίρας η οποία φωτίζεται από κάποια πηγή φωτός [23]. Οταν η σφαίρα περιστρέφεται γύρω από κάποια διάμετρό της, το πεδίο κίνησης όπως είναι φυσικό "ενεργοποιείται" και τα διανύσματα κίνησης παίρνουν συγκεκριμένες τιμές ανάλογα με την ταχύτητα της σφαίρας. Οι εντάσεις στη φωτεινότητα των



Σχήμα 4.2: Πεδίο κίνησης

Το πεδίο κίνησης ενός πιλότου που κοιτάει ευθεία μπροστά σε μια επίπεδη πτήση. Το πεδίο κίνησης μιας κάμερας ή ενός παρατηρητή, ο οποίος κινείται σε σχέση με το περιβάλλον αποτελεί μια εκχώρηση διανυσμάτων σε σημεία της εικόνας κατά την οποία κάθε διάνυσμα αναπαριστά την κίνηση στην εικόνα του αντίστοιχου σκηνικού σημείου.

σημείων, όμως, δε μεταβάλλονται μια και η τοπογραφία του χώρου παραμένει απαράλλαχτη. Αντίθετα, αν υποθέσουμε ότι η σφαίρα παραμένει ακίνητη, οπότε και τα διανύσματα κίνησης είναι μηδενικά, αλλά η πηγή φωτός αλλάζει γωνία εκπομπής, τότε προφανώς οι νέες φωτοσκιάσεις δημιουργούν ένα οπτικό πεδίο ροής που αντανακλά αυτή την αλλαγή γωνίας.

Ακραίες περιπτώσεις, όπως αυτή της σφαίρας, δε θα μας απασχολήσουν στην παρούσα μελέτη. Απ' την άλλη, για να είναι αποδοτικός ο υπολογισμός του οπτικού πεδίου, εισάγουμε την εξής -όχι εξωπραγματική- παραδοχή: δεδομένης μίας σκηνής, η τιμή της έντασης φωτεινότητας ενός σημείου του κόσμου παραμένει σταθερή στις απεικονίσεις του στο επίπεδο προβολής. Αν I είναι η συνάρτηση έντασης της φωτεινότητας, τότε η παραδοχή αυτή εκφράζεται μαθηματικά ως [8]:

$$I(x, y, t) = I(x + dx, y + dy, t + dt)$$

Θεωρώντας $dx = udt$ και $dy = \upsilon dt$ όπου (u, υ) το διάνυσμα οπτικής ροής, ουσιαστικά η στιγμιαία ταχύτητα, του σημείου (x, y), αναπτύσσουμε το δεύτερο σκέλος της εξίσωσης με σειρά Taylor και προκύπτει η εξίσωση:

$$I(x, y, t) = I(x, y, t) + \frac{\partial I}{\partial x}dx + \frac{\partial I}{\partial y}dy + \frac{\partial I}{\partial t}dt + e$$

όπου e το άθροισμα των υπόλοιπων όρων και $\frac{\partial I}{\partial x} = I_x$, $\frac{\partial I}{\partial y} = I_y$, $\frac{\partial I}{\partial t} = I_t$ οι μερικοί παράγωγοι της I ως προς το χώρο και το χρόνο. Οταν $dt \rightarrow 0$, τότε $e \rightarrow 0$, οπότε διαιρώντας με

dt καταλήγουμε στη γνωστή ***εξίσωση περιορισμού της οπτικής ροής*** (*optical flow constraint equation*), η οποία επιτρέπει, όπως δείχθηκε σχηματικά στο προηγούμενο κεφάλαιο (σχ. 3.1), τον υπολογισμό μόνο της προβολής της οπτικής ροής στο διάνυσμα της κλίσης της έντασης (intensity gradient) [8].

$$0 = \frac{dx}{dt}I_x + \frac{dy}{dt}I_y + I_t \Rightarrow$$

$$0 = uI_x + \upsilon I_y + I_t \tag{4.1}$$

ή μεταγράφοντάς την με μορφή εσωτερικού γινομένου

$$-I_t = (I_x, I_y) \cdot (u, \upsilon) \tag{4.2}$$

4.3 Κάθετη ροή

Το ***πεδίο κάθετης ροής*** (*normal flow field*) αποτελείται από τις προβολές των διανυσμάτων οπτικής ροής στη διεύθυνση της έντασης, η οποία διεύθυνση για ένα σημείο μιας ακμής είναι κάθετη προς την ακμή. Το πεδίο αυτό συχνά συγχέεται με το ***πεδίο κάθετης κίνησης*** (*normal motion field*), τις κάθετες συνιστώσες δηλαδή των προβολών των διανυσμάτων κίνησης στο επίπεδο προβολής. Αν και φυσικά μια τέτοια ταύτιση αποτελεί νοηματικό σφάλμα, αποδεικνύεται στην πράξη ότι για συγκεκριμένες συνθήκες το πεδίο κάθετης ροής αποτελεί ικανοποιητική προσέγγιση του πεδίου κάθετης κίνησης και επομένως, μπορεί να χρησιμοποιηθεί για την εξαγωγή συμπερασμάτων που αφορούν στην τριδιάστατη κίνηση.

Πιο αναλυτικά, έστω $I(x, y, t)$ η ένταση της εικόνας, $\vec{\upsilon} = (u, \upsilon)$ το πεδίο οπτικής ροής και $\vec{\overline{\upsilon}} = (\overline{u}, \overline{\upsilon})$ το πεδίο κίνησης σ' ένα σημείο (x, y). Το μοναδιαίο διάνυσμα στη διεύθυνση της κλίσης της έντασης είναι $\vec{n} = (I_x, I_y)/\sqrt{(I_x^2, I_y^2)}$. Το πεδίο κάθετης κίνησης έχει μέτρο προφανώς ίσο με:

$$\overline{u}_n = \vec{\overline{\upsilon}} \cdot \vec{n} \Rightarrow$$

$$\overline{u}_n = (u, v) \cdot (\frac{1}{\sqrt{(I_x^2, I_y^2)}}(I_x, I_y)) \Rightarrow$$

$$\overline{u}_n = (\frac{dx}{dt}, \frac{dy}{dt}) \cdot \frac{\bigtriangledown I}{\|\bigtriangledown I\|} \Rightarrow$$

$$\overline{u}_n = \frac{1}{\|\bigtriangledown I\|}(I_x\frac{dx}{dt} + I_y\frac{dy}{dt}) \tag{4.3}$$

Απ' την άλλη, εφόσον το πεδίο κάθετης κίνησης ορίζεται ως η προβολή του πεδίου οπτικής ροής, διαιρώντας στη (4.2) με το μέτρο του (I_x, I_y) για να προκύψει το μοναδιαίο διάνυσμα στην κατεύθυνση αυτή, έχουμε [13]

$$\frac{-I_t}{\sqrt{I_x^2+I_y^2}} = \frac{1}{\sqrt{I_x^2+I_y^2}}((I_x, I_y)\cdot(u, \upsilon)) = u_n \Rightarrow$$

$$u_n = -\frac{1}{\|\bigtriangledown I\|} I_t \tag{4.4}$$

Από τις (4.3) και (4.4) προκύπτει ότι

$$\overline{u}_n - u_n = \frac{1}{\|\bigtriangledown I\|}\frac{dI}{dt} \tag{4.5}$$

Η (4.5) δηλώνει ότι τα δύο πεδία είναι περίπου ίσα όταν η κλίση της έντασης $\bigtriangledown I$ κυμαίνεται σε υψηλά επίπεδα. Αυτό μας προϊδεάζει να περιμένουμε ότι το πεδίο κάθετης ροής θα μας εφοδιάζει με αξιόπιστες πληροφορίες για την τριδιάστατη κίνηση μόνο όταν η τιμή της κλίσης της έντασης είναι αρκετά μεγάλη, γι' αυτό και στον υπολογισμό του πεδίου κάθετης ροής λαμβάνονται υπόψη μόνο τα σημεία για τα οποία το μέγεθος $\sqrt{I_x^2+I_y^2}$ ξεπερνάει κάποιο κατώφλι.

Στο σχήμα (4.4) φαίνονται τα σημεία όπου παρουσιάστηκαν διανύσματα κάθετης ροής από τη μετακίνηση του παρατηρητή στο σκηνικό του σχήματος (4.3). Η παραγώγιση ως προς το χώρο πραγματοποιήθηκε με την εφαρμογή φίλτρων Sobel στην 3×3 γειτονιά κάθε pixel, ενώ εκείνη ως προς το χρόνο εκφράστηκε με τη μέση τιμή της διαφοράς της έντασης της 3×3 γειτονιάς κάθε pixel (βλ. Παράρτημα Α).

4.4 Υπολογισμός εξισώσεων οπτικής ροής

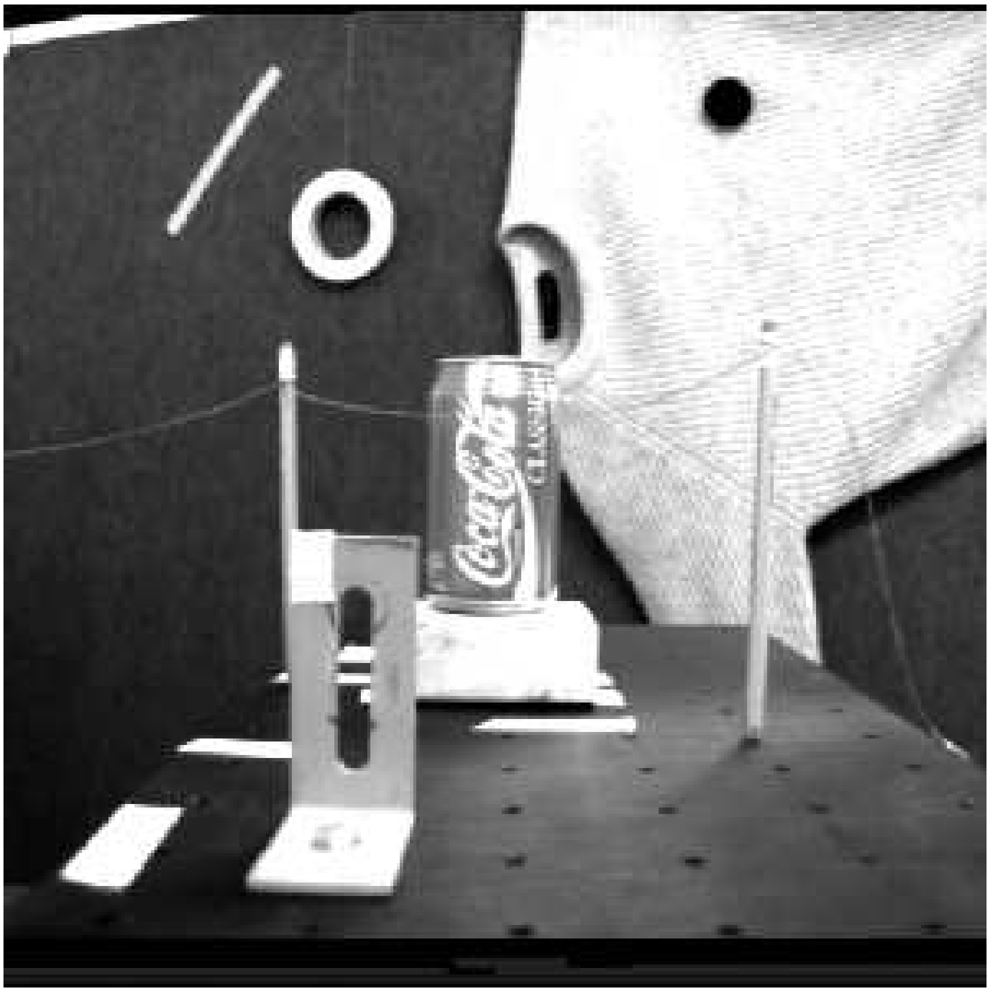
Η κάμερα-παρατηρητής κινείται σ' ένα κατά τα άλλα στατικό περιβάλλον. Η κάμερα ακολουθεί ένα συνδυασμό περιστροφική ς και μεταφορικής κίνησης. Συμβολίζουμε με ω το διάνυσμα που εκφράζει την περιστροφική κίνηση και με t το διάνυσμα που εκφράζει την μεταφορική κίνηση. Εστω ότι οι στιγμιαίες συντεταγμένες ενός σημείου P στο χώρο είναι (X, Y, Z). Η στιγμιαία ταχύτητα V του σημείου P δίνεται από τη σχέση

$$V = -t - \omega \times r \tag{4.6}$$

Θέτοντας $t = (U, V, W)$ και $\omega = (\alpha, \beta, \gamma)$ η εξίσωση (4.6) μπορεί να ξαναγραφεί χρησιμοποιώντας τις συντεταγμένες των διανυσμάτων ως [13]

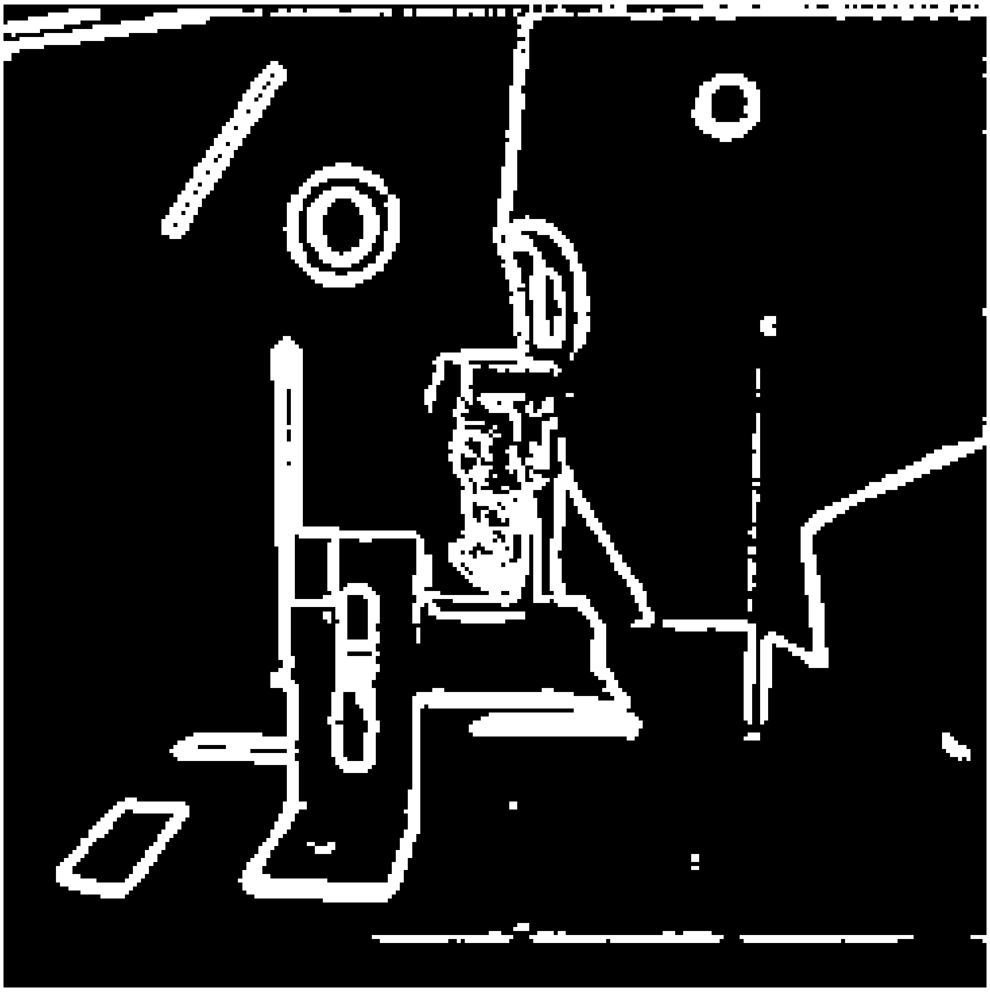
$$\dot{X} = -U - \beta Z + \gamma Y,$$

$$\dot{Y} = -V - \gamma X + \alpha Z,$$



Σχήμα 4.3: Εικόνα κάμερας

Σκηνικό κάποια χρονική στιγμή $t1$. Στη συνέχεια ο παρατηρητής εκτέλεσε μεταφορική και περιστροφική κίνηση και κατέγραψε νέα εικόνα η οποία διαφέρει ελάχιστα από την πρώτη.



Σχήμα 4.4: Παράδειγμα πεδίου κάθετης ροής

Διανύσματα κάθετης ροής βρέθηκαν γύρω από τις πιο έντονες ακμές των εικόνων, "ερεθισμένα" προφανώς από την αλλαγή φωτεινότητας λόγω της κίνησης της κάμερας.

$$\dot{Z} = -W - \alpha Y + \beta X,$$

όπου η τελεία δηλώνει την παραγώγιση ως προς το χρόνο.

Η οπτική ροή σε κάθε σημείο της εικόνας είναι η στιγμιαία ταχύτητα της έντασης φωτεινότητας στο σημείο εκείνο. Εστω (x, y) ένα σημείο στο επίπεδο προβολής. Υποθέτοντας προοπτική προβολή οι συντεταγμένες του σημείου $p(x, y)$, της προβολής δηλαδή του P στην εικόνα δίνονται από τις σχέσεις

$$x = \frac{Xf}{Z} \text{ και } y = \frac{Yf}{Z}$$

Η οπτική ροή (u, υ) σε ένα σημείο (x, y) της εικόνας είναι

$$u = \dot{x} \text{ και } \upsilon = \dot{y}$$

Παραγωγίζοντας τις εξισώσεις για x και y ως προς το χρόνο και χρησιμοποιώντας τις παραγώγους των X, Y και Z, προκύπτουν οι παρακάτω εξισώσεις για την οπτική ροή:

$$u = \frac{\dot{X}}{Z} - \frac{X\dot{Z}}{Z^2} = (-\frac{Uf}{Z} - \beta f + \gamma y) - x(-\frac{W}{Z} - \frac{\alpha y}{f} + \beta x f),$$

$$\upsilon = \frac{\dot{Y}}{Z} - \frac{Y\dot{Z}}{Z^2} = (-\frac{Vf}{Z} - \gamma x + \alpha f) - y(-\frac{W}{Z} - \frac{\alpha y}{f} + \frac{\beta x}{f}).$$

Ξαναγράφουμε τις παραπάνω εξισώσεις στη μορφή

$$u = u_{trans} + u_{rot}, \upsilon = \upsilon_{trans} + \upsilon_{rot},$$

όπου $(u_{trans}, \upsilon_{trans})$ δηλώνει το μεταφορικό κομμάτι και $(u_{rot}, \upsilon_{rot})$ δηλώνει το περιστροφικό κομμάτι της κίνησης. Τελικά καταλήγουμε στις εξισώσεις:

$$u_{trans} = \frac{-Uf + xW}{Z}, u_{rot} = \alpha\frac{xy}{f} - \beta(\frac{x^2}{f} + f) + \gamma y, \tag{4.7}$$

$$\upsilon_{trans} = \frac{-Vf + yW}{Z}, \upsilon_{rot} = \alpha(\frac{y^2}{f} + f) - \beta\frac{xy}{f} - \gamma x. \tag{4.8}$$

Παρατηρούμε ότι η οπτική ροή εξαρτάται από τις έξι παραμέτρους κίνησης. Οι μεταφορικές συνιστώσες είναι γραμμικές ως προς x και y, ενώ οι περιστροφικές όχι. Επίσης, η τιμή του βάθους Z υπεισέρχεται μόνο στη μεταφορική συνιστώσα.

4.5 Παραδείγματα απλών κινήσεων

4.5.1 Μεταφορική

Εστω ότι ο παρατηρητής εκτελεί μόνο μεταφορική κίνηση. Τότε, όπως φαίνεται και από τη μορφή που λαμβάνουν οι εξισώσεις κίνησης μηδενίζοντας τις παραμέτρους περιστροφής,

$$u = u_{trans} = \frac{-Uf + xW}{Z} = (-x_0 + x)\frac{W}{Z}$$

$$\upsilon = \upsilon_{trans} = \frac{-Vf + yW}{Z} = (-y_0 + y)\frac{W}{Z}$$

, όπου $(x_0, y_0) = (\frac{Uf}{W}, \frac{Vf}{W})$. Η κίνηση αυτή δημιουργεί ένα πεδίο οπτικής ροής με το χαρακτηριστικό ότι όλα τα διανύσματα βρίσκονται πάνω σε φορείς που περνάνε από το συγκεκριμένο σημείο (x_0, y_0) του επίπεδου προβολής. Το σημείο αυτό αποτελεί την τομή του διανύσματος μεταφορικής κίνησης (U, V, W) με το επίπεδο προβολής (σχ. 4.5). Η οπτική ροή σε κείνο το σημείο είναι 0. Αν ο παρατηρητής πλησιάζει το σκηνικό, τότε όλα τα διανύσματα "ξεπηδούν" από το σημείο αυτό, ενώ αν απομακρύνεται τα διανύσματα κατευθύνονται προς εκείνο. Στις αντίστοιχες περιπτώσεις το σημείο ονομάζεται ***Εστιακό Σημείο Εξάπλωσης*** (***ΕΣΕ***) (*Focus of expansion*) (*FOE*) ή ***Εστιακό Σημείο Συρρίκνωσης*** (***ΕΣΣ***) (*Focus of contraction*) (*FOC*). Το μήκος των διανυσμάτων εξαρτάται φυσικά από τη θέση των αντίστοιχων σημείων στο τριδιάστατο χώρο.

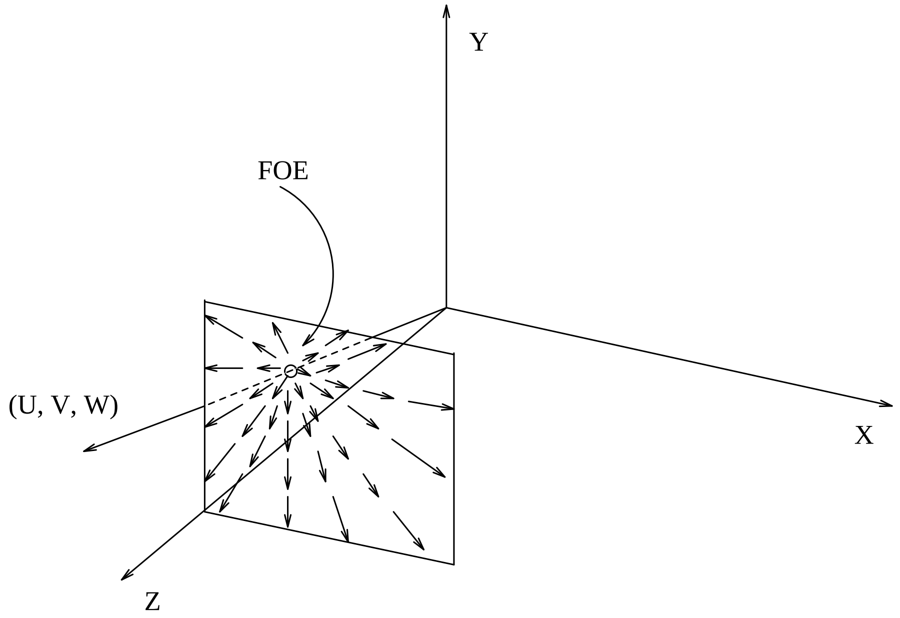
4.5.2 Περιστροφική

Στην περίπτωση που ο παρατηρητής εκτελεί μόνο περιστροφική κίνηση, τότε οι εξισώσεις κίνησης λαμβάνουν την εξής μορφή

$$u = u_{rot} = \alpha\frac{xy}{f} - \beta(\frac{x^2}{f} + f) + \gamma y$$

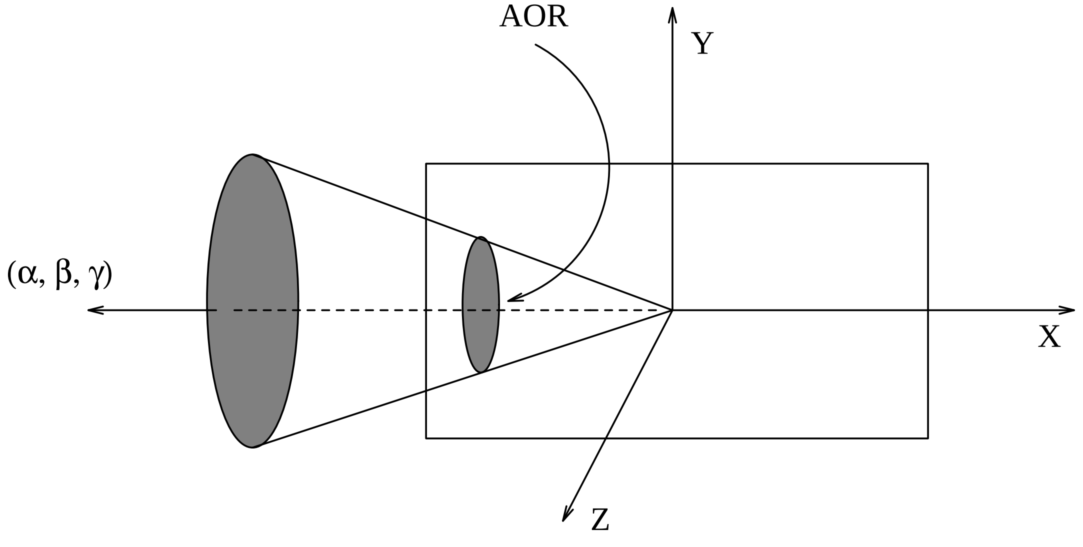
$$\upsilon = \upsilon_{rot} = \alpha(\frac{y^2}{f} + f) - \beta\frac{xy}{f} - \gamma x$$

Τα σημεία του χώρου κινούνται σε κυκλικές περιφέρειες σε επίπεδα τα οποία είναι κάθετα στον άξονα περιστροφής, δηλαδή το διάνυσμα α, β, γ. Η απεικόνιση αυτών των κυκλικών περιφερειών μέσω της προοπτικής προβολής βρίσκεται να είναι η τομή του επιπέδου προβολής και ενός κώνου που ορίζεται από τον άξονα περιστροφής και την κυκλική περιφέρεια (σχ. 4.6). Η κωνική αυτή τομή λαμβάνει ποικίλες μορφές -υπερβολή, παραβολή, έλλειψη, κύκλος- ανάλογα με τη θέση του άξονα περιστροφής σε σχέση με το επίπεδο προβολής,



Σχήμα 4.5: Περίπτωση μεταφορικής κίνησης

Οι φορείς των προβολών των διανυσμάτων κίνησης στην εικόνα διέρχονται όλοι από το σημείο FOE ($\frac{Uf}{W}$, $\frac{Vf}{W}$).



Σχήμα 4.6: Περίπτωση περιστροφικής κίνησης

ουσιαστικά δηλαδή ανάλογα με τις τιμές των παραμέτρων α, β, γ. Το σημείο που το διάνυσμα περιστροφής (α, β, γ) τέμνει το επίπεδο προβολής ονομάζεται ***Σημείο Αξονα Περιστροφής*** (***ΣΑΠ***) (*Axis of Rotation*) (*AOR*). Οι συντεταγμένες του AOR στην εικόνα είναι $x = \frac{\alpha f}{\gamma}$ και $y = \frac{\beta f}{\gamma}$.

Κεφάλαιο 5

Εκμετάλλευση γεωμετρικών ιδιοτήτων κάθετης ροής

Αν και η πληροφορία που μας παρέχει το πεδίο κάθετης ροής είναι μικρότερη σε αξία από εκείνη του πεδίου οπτικής ροής, προτιμείται σε αυτήν την προσέγγιση ο υπολογισμός αυτού του πεδίου, ο οποίος είναι ευκολότερος και απαιτεί λιγότερες υποθέσεις για το σκηνικό. Στο κεφάλαιο αυτό θα δειχθεί πώς, χρησιμοποιώντας τις γεωμετρικές ιδιότητες του πεδίου, εξάγονται χρήσιμες πληροφορίες για τη θέση των κρίσιμων σημείων FOE και AOR [11].

5.1 Περιοχές σταθερής κάθετης ροής

Οπως έχει προαναφερθεί, ένα διάνυσμα του πεδίου κάθετης ροής αποτελεί την κάθετη προβολή του αντίστοιχου διανύσματος οπτικής ροής στη διεύθυνση του gradient. Ετσι, αν είναι $\mathbf{u}$ η οπτική ροή στο σημείο (x,y) και $\mathbf{n} = (\mathbf{n_x}, \mathbf{n_y})$ το μοναδιαίο διάνυσμα στην κατεύθυνση του gradient, το διάνυσμα κάθετης ροής είναι :

$$\mathbf{u_n} = (\mathbf{u} \cdot \mathbf{n})\mathbf{n}$$

Αντικαθιστώντας στην εξίσωση αυτή τις παραστάσεις των συνιστωσών του $\mathbf{u}$, η εξίσωση του μέτρου u_n του $\mathbf{u_n}$ παίρνει την εξής μορφή:

$$u_n = u_{rot}n_x + u_{trans}n_x + \mathfrak{v}_{rot}n_y + \mathfrak{v}_{trans}n_y \overset{(4.7)(4.8)}{\Longrightarrow}$$

$$u_n = \frac{W}{Z}(x-x_0)n_x + (\alpha\frac{xy}{f} - \beta(\frac{x^2}{f}+f)+\gamma y)n_x + \frac{W}{Z}(y-y_0)n_y + (\alpha(\frac{y^2}{f}+f) - \beta\frac{xy}{f} - \gamma x)n_y \tag{5.1}$$

όπου $(x_0, y_0) = (\frac{Uf}{W}, \frac{Vf}{W})$ είναι το FOE.

Εχοντας ως δεδομένο το πεδίο κάθετης ροής δύο διαδοχικών εικόνων, μας απασχολεί το ερώτημα αν από τη θέση των διανυσμάτων κάθετης ροής με συγκεκριμένη τιμή σε μέτρο και φορά μπορούμε να εξαγάγουμε συμπεράσματα για τη θέση του FOE και του AOR. Αναδιατάσσοντας την εξίσωση (5.1), έτσι ώστε να διαχωρίσουμε το γραμμικό (και στην περίπτωσή μας μεταφορικό) από το μη γραμμικό (περιστροφικό) κομμάτι της παράγεται η:

$$(u_n - (\alpha\frac{xy}{f} - \beta(\frac{x^2}{f}+f)+\gamma y)n_x - (\alpha(\frac{y^2}{f}+f)-\beta\frac{xy}{f}-\gamma x)n_y = \frac{W}{Z}((x-x_0)n_x+(y-y_0)n_y) \tag{5.2}$$

Η ποσότητα $(\alpha\frac{xy}{f} - \beta(\frac{x^2}{f}+f)+\gamma y)n_x - (\alpha(\frac{y^2}{f}+f)-\beta\frac{xy}{f}-\gamma x)n_y$ αποτελεί ουσιαστικά την προβολή της περιστροφικής συνιστώσας της κίνησης στο διάνυσμα (n_x, n_y). Αντίθετα, η ποσότητα $\frac{W}{Z}((x-x_0)n_x+(y-y_0)n_y)$ αποτελεί την προβολή της μεταφορικής συνιστώσας της κίνησης στο διάνυσμα (n_x, n_y).

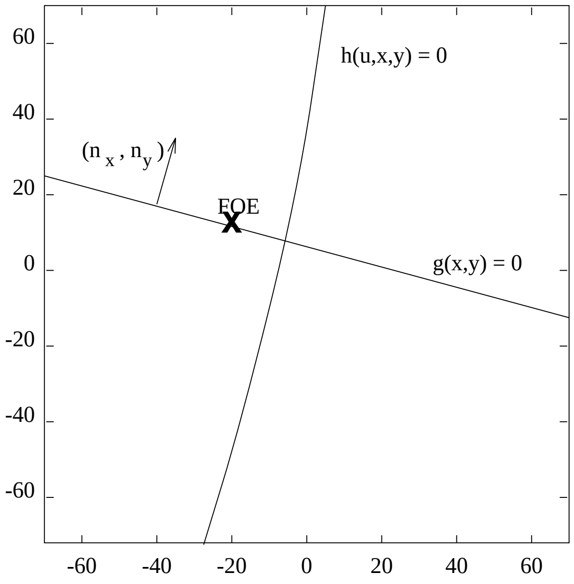
Δεχόμενοι τη ρεαλιστική υπόθεση ότι το βάθος είναι πάντα θετικό, δηλαδή $Z > 0$, και υποθέτοντας ότι ο παρατηρητής πλησιάζει τη σκηνή, και άρα $W > 0$, προκύπτει η παρακάτω ανίσωση:

$$(u_n - (\alpha\frac{xy}{f} - \beta(\frac{x^2}{f}+f)+\gamma y)n_x - (\alpha(\frac{y^2}{f}+f)-\beta\frac{xy}{f}-\gamma x)n_y)((x-x_0)n_x+(y-y_0)n_y) > 0$$

ή πιο σύντομα

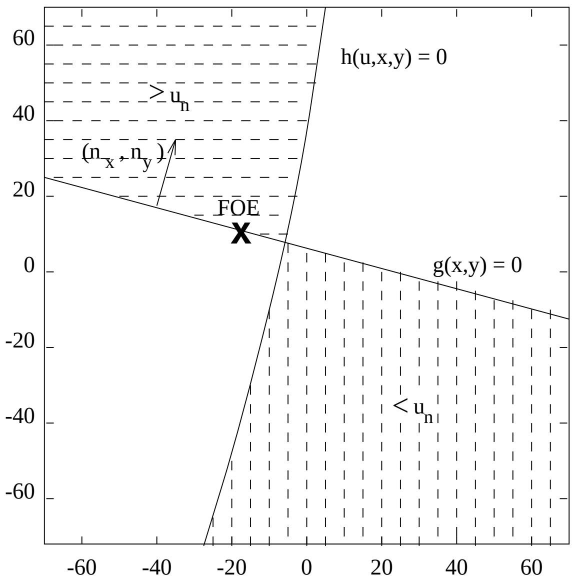
$$h(u_n, \alpha, \beta, \gamma; x, y)g(x_0, y_0; x, y) > 0 \tag{5.3}$$

όπου $h(u_n, \alpha, \beta, \gamma, x, y) = u_n - (\alpha\frac{xy}{f} - \beta(\frac{x^2}{f}+f)+\gamma y)n_x - (\alpha(\frac{y^2}{f}+f)-\beta\frac{xy}{f}-\gamma x)n_y$ και $g(x_0, y_0, x, y) = (x-x_0)n_x+(y-y_0)n_y$. Τόσο η εξίσωση $h(x, y) = 0$, η οποία περιγράφει μια υπερβολή , όσο και η εξίσωση $g(x, y) = 0$, η οποία περιγράφει μια ευθεία, χωρίζουν την εικόνα σε δύο περιοχές όπου στη μία λαμβάνουν θετικές τιμές ενώ στην άλλη αρνητικές. Χαρακτηριστικά στοιχεία της g είναι ότι διέρχεται από το σημείο FOE, αφού οι συντεταγμένες του, (x_0, y_0), μηδενίζουν την εξίσωση $g = 0$, ενώ συγχρόνως έχει διεύθυνση κάθετη στο μοναδιαίο διάνυσμα (n_x, n_y), αφού οι συντελεστές του x και y στην εξίσωση $g = 0$ είναι n_x και n_y αντίστοιχα. Επίσης, παρατηρούμε ότι η g είναι ανεξάρτητη του μέτρου u_n των διανυσμάτων που εξετάζουμε. Αν, λοιπόν, βρίσκαμε τη θέση δύο τέτοιων ευθειών για δύο διαφορετικές ομάδες διανυσμάτων κάθετης ροής, θα ανακαλύπταμε αμέσως και τη θέση του FOE, που δεν είναι άλλη από την τομή των ευθειών. Επομένως, το πρόβλημα μετατίθεται στο να ανακαλυφθεί ένας τρόπος ώστε να εντοπίζονται οι προσεγγιστικές έστω θέσεις τέτοιων ευθειών στο επίπεδο προβολής, δεδομένου ότι αν και γνωρίζουμε τη διεύθυνσή τους από τον ορισμό της αναζήτησης σταθερής κάθετης ροής, αγνοούμε την ακριβή τους τοποθέτηση στην εικόνα. Μια πιθανή θέση των h και g στο επίπεδο προβολής φαίνεται στο σχήμα 5.1.



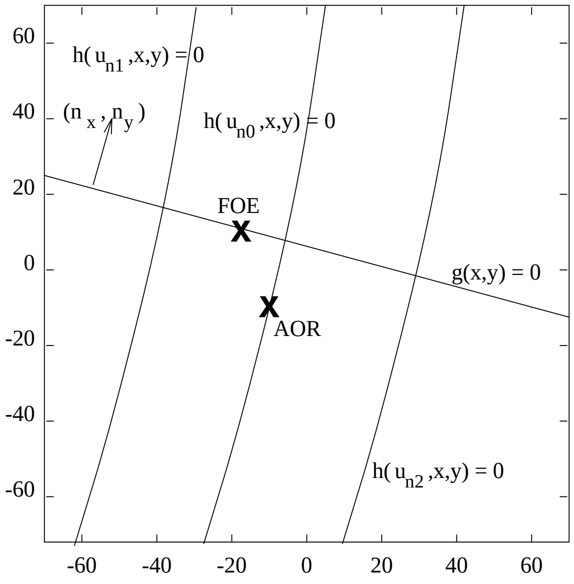
Σχήμα 5.1: Καμπύλες χαρακτηρισμού συγκεκριμένης κάθετης ροής

Οι περιοχές κάθετης ροής συγκεκριμένου μέτρου και διεύθυνσης χαρακτηρίζονται από μια ευθεία και μια υπερβολή. Η ευθεία είναι κάθετη στη διεύθυνση των διανυσμάτων και διέρχεται από το FOE, ενώ η υπερβολή διέρχεται από το AOR, όταν η κάθετη ροή είναι 0.



Σχήμα 5.2: Κατανομή μέτρων των διανυσμάτων κάθετης ροής.

Το επίπεδο προβολής διαχωρίζεται σε περιοχές ανάλογα με τις τιμές του μέτρου των διανυσμάτων κάθετης ροής με διεύθυνση (n_x, n_y). Στην περιοχή με την οριζόντια γράμμωση τα διανύσματα έχουν μέτρο μόνο μεγαλύτερο από u_n. Στην περιοχή με την κατακόρυφη γράμμωση μόνο μικρότερη του u_n. Στις υπόλοιπες περιοχές υπάρχουν διανύσματα με μέτρο μεγαλύτερο, μικρότερο, αλλά και ίσο με u_n.

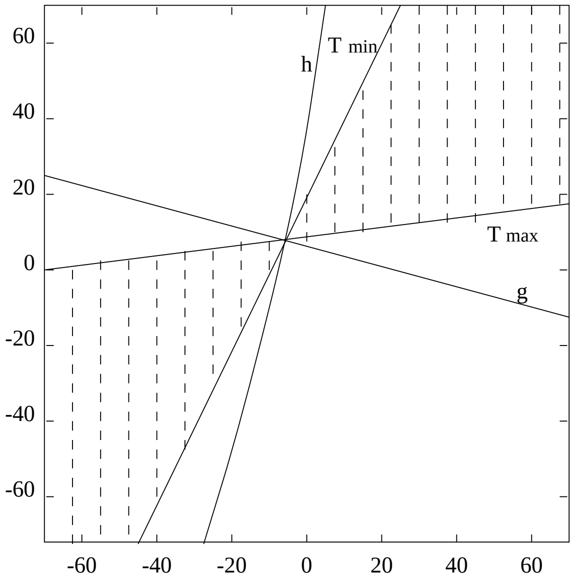


Σχήμα 5.3: Περιοχές κάθετης ροής παράλληλων διανυσμάτων

Οι υπερβολές $h(u_{n0},x,y), h(u_{n1},x,y), h(u_{n2},x,y)$ αντιστοιχούν σε παράλληλα διανύσματα κάθετης ροής μέτρου u_{n0}, u_{n1}, u_{n2}. Το μήκος u_{n0} ισούται με 0, γι' αυτό και η υπερβολή περνάει από το AOR. Η ευθεία $g = 0$ είναι ανεξάρτητη από το μήκος των διανυσμάτων και αφού αυτά είναι παράλληλα μεταξύ τους εκείνη είναι η ίδια για όλα.

Διερευνώντας τις τιμές των h και g στις τέσσερις περιοχές που σχηματίζονται, εξάγουμε χρήσιμα συμπεράσματα για την τιμή του μέτρου u_n του διανύσματος κάθετης ροής στο σημείο (x,y). Αν $h(x,y) > 0$ τότε $u_n > (\alpha\frac{xy}{f} - \beta(\frac{x^2}{f} + f) + \gamma y)n_x + (\alpha(\frac{y^2}{f} + f) - \beta\frac{xy}{f} - \gamma x)n_y \Rightarrow u_n > u_{rot}$, δηλαδή το μέτρο του διανύσματος είναι μεγαλύτερο από τη συνιστώσα περιστροφής της κάθετης κίνησης. Αν $g(x,y) < 0$ τότε και $0 > \frac{W}{Z}g(x,y) \Rightarrow 0 > u_{trans}$, οπότε η συνιστώσα μεταφοράς της κάθετης κίνησης είναι μικρότερη του μηδενός. Προσθέτοντας κατά μέλη τις δύο παραπάνω ανισότητες προκύπτει $u_n + 0 > u_{rot} + u_{trans}$ και επομένως το μέτρο της κάθετης κίνησης είναι μικρότερο από u_n, όπου $g(x,y) < 0$ και $h(x,y) > 0$. Με την ίδια λογική συμπεραίνεται ότι στις περιοχές όπου $g(x,y) > 0$ και $h(x,y) < 0$ το μέτρο της κάθετης κίνησης είναι μεγαλύτερο από u_n. Συνεπώς, τα διανύσματα κάθετης κίνησης με μέτρο u_n και διεύθυνση (n_x, n_y) , θα βρίσκονται αυστηρά μόνο στις υπόλοιπες δύο περιοχές (χωρίς αυτό να σημαίνει, βέβαια, ότι όλα τα διανύσματα αυτών των περιοχών έχουν μέτρο u_n) (σχήμα 5.2).

Μία επιπλέον ενδιαφέρουσα παρατήρηση, που θα βοηθήσει στην εξαγωγή πληροφορίας για τη θέση του FOE, είναι και η εξής: αν θεωρήσουμε παράλληλα διανύσματα κάθετης κίνησης, δηλαδή διανύσματα της μορφής $k(u_n, \upsilon_n)$, οι περιοχές στο επίπεδο προβολής



Σχήμα 5.4: Θέση υπερβολών λόγω περιορισμών βάθους

Οι περιορισμοί στις τιμές του Z περιορίζουν την περιοχή (γραμμοσκιασμένη επιφάνεια) που μπορεί να βρίσκονται τα διανύσματα κάθετης ροής συγκεκριμένου μέτρου και διεύθυνσης. Το στένεμα της κατανομής των διανυσμάτων πάνω και κάτω από την ευθεία μας δίνει ένα μέτρο προσεγγιστικού προσδιορισμού της θέσης της.

οριοθετούνται και πάλι από μία και την αυτήν ευθεία, αφού η διεύθυνση των διανυσμάτων είναι σταθερή, ενώ οι υπερβολές που σχηματίζονται διαφέρουν κατά τους γραμμικούς μόνο όρους τους, και άρα είναι κατά κάποιον τρόπο "παράλληλες" η μία με την άλλη (σχήμα 5.3).

5.2 Εισαγωγή ορίων βάθους

Εστω ότι αναζητούμε, όπως και πριν, όλα τα διανύσματα κάθετης ροής με μέτρο u_n και διεύθυνση (n_x, n_y). Είδαμε ότι αυτά τα διανύσματα βρίσκονται αποκλειστικά σε δύο περιοχές οι οποίες οριοθετούνται από μια ευθεία και μια υπερβολή, καμπύλες που ορίζονται από τις έξι -άγνωστές μας- παραμέτρους κίνησης καθώς βέβαια και από τις συγκεκριμένες τιμές του μέτρου και διεύθυνσης. Στις συνηθισμένες, όμως, πρακτικές εφαρμογές λογικό είναι να υποθέσουμε ότι η σκηνή χαρακτηρίζεται τόσο από ένα ελάχιστο (Z_{min}) όσο και από ένα μέγιστο βάθος (Z_{max}). Αντικαθιστώντας, λοιπόν, αυτές τις δύο τιμές στην εξίσωση (5.1), προκύπτουν οι εξισώσεις δύο νέων υπερβολών, οι οποίες τέμνονται στην ευθεία $g = 0$ και οι οποίες ορίζουν μια περιοχή στο επίπεδο προβολής μέσα στην οποία **πρακτικά** ανήκουν όλα τα διανύσματα κάθετης ροής με τις επιθυμητές τιμές σε μέτρο και διεύθυνση (σχήμα

5.4). Παρατηρώντας προσεκτικά το σχήμα διαπιστώνουμε ότι παρότι δεν ξέρουμε την ακριβή θέση της ευθείας g (παρά μόνο τη διεύθυνσή της), αποκτούμε πλέον ένα ισχυρό μέσο προσεγγιστικού υπολογισμού της: το σχήμα που περιβάλλει όλα τα διανύσματα κάθετης ροής με συγκεκριμένο μέτρο και διεύθυνση στενεύει χαρακτηριστικά στην περιοχή γύρω από τη ζητούμενη ευθεία.

5.3 Ομοσημειακά και ομοαξονικά διανύσματα

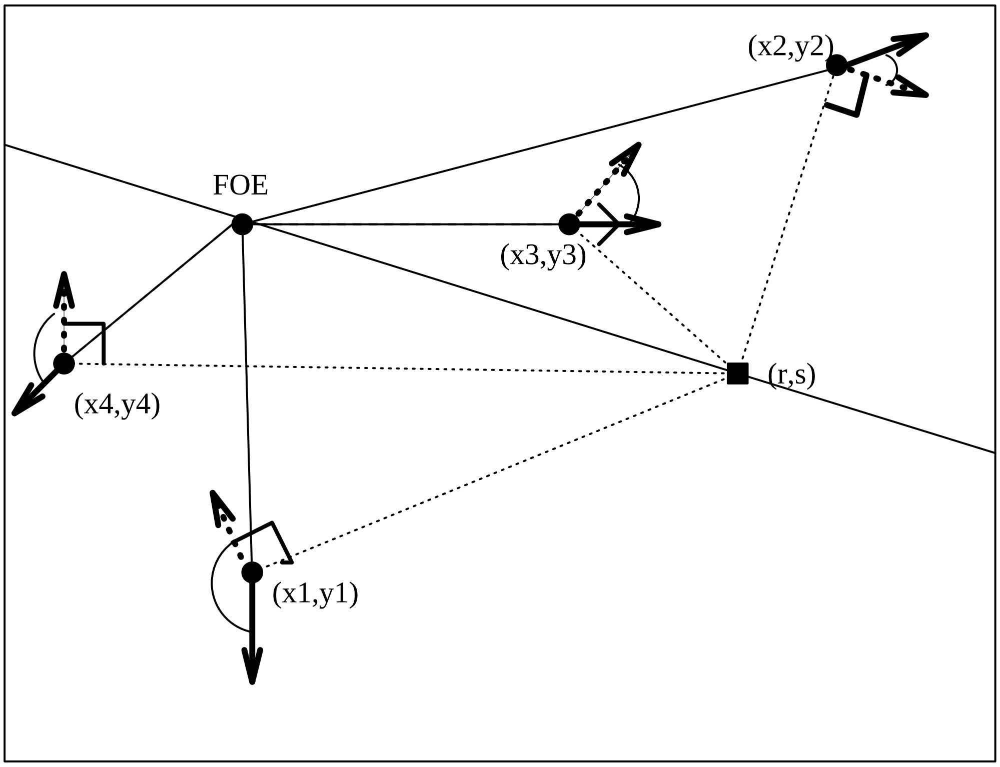
Η ανάγκη για την εύρεση περισσότερων περιορισμών για τη θέση των FOE και AOR οδηγεί στην προσπάθεια εύρεσης ομαδοποιήσεων των διανυσμάτων κάθετης ροής οι οποίες να μη χρησιμοποιούν ως μοναδικό κριτήριο την ταύτιση σε μέτρο και φορά, αλλά που να επιτρέπουν σε μια ομάδα να περιέχει διανύσματα διαφορετικής διεύθυνσης τα οποία να συσχετίζονται με κάποιο σταθερό τρόπο από τα ζητούμενα σημεία FOE και AOR.

Ας αναλογιστούμε π.χ. την περίπτωση μιας καθαρά μεταφορικής κίνησης. Οπως φαίνεται στο σχήμα (4.5) όλα τα διανύσματα οπτικής ροής "ξεπηδούν" από το FOE. Αν θεωρήσουμε ένα τυχαίο σημείο (r, s) της εικόνας, τότε παρατηρούμε ότι αν για κάθε σημείο (x, y) της εικόνας προβάλλουμε το αντίστοιχο διάνυσμα οπτικής ροής στο διάνυσμα με διεύθυνση $(-y + s, x - r)$, που είναι κάθετη προς εκείνη της ευθείας που ενώνει τα (x, y) και (r, s), τότε τα μέτρα των προβολών είναι θετικά πάνω και αρνητικά κάτω από την ευθεία που διέρχεται από τα FOE και (r, s) (σχ. 5.5). Επομένως, μέσω του (r, s) βρίσκουμε έναν ακόμα περιορισμό για το FOE: θα βρίσκεται πάνω στην ευθεία που περνάει από το (r, s) και χωρίζει τα διανύσματα κάθετης ροής με διευθύνσεις $(-y + s, x - r)$ σε ομάδες με θετικό και αρνητικό μέτρο.

Με βάση αυτή τη λογική γενικεύονται και εξετάζονται στις επόμενες παραγράφους δύο νέες μέθοδοι διαχωρισμού του πεδίου κάθετης ροής σε ομάδες, οι οποίες παρέχουν πληροφορίες για τη θέση των FOE και AOR.

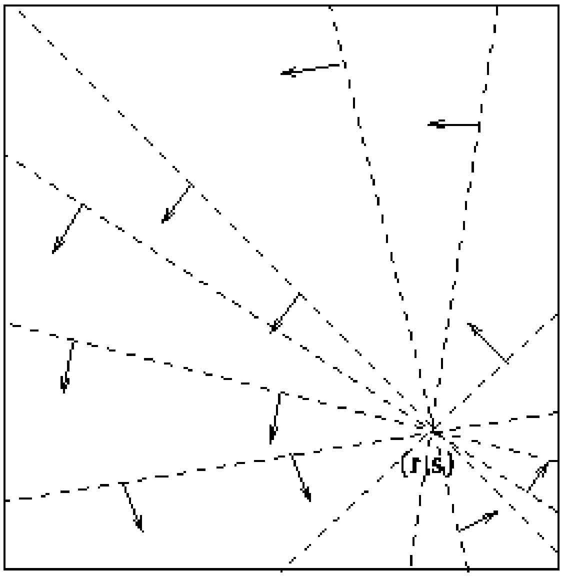
5.3.1 Ομοσημειακά διανύσματα

Τα ***ομοσημειακά διανύσματα*** (*copoint vectors*) ορίζονται με βάση κάποιο σημείο του διδιάστατου χώρου. Συγκεκριμένα, το διάνυσμα κάθετης ροής $\mathbf{u}$ σ' ένα σημείο (x, y) θα ανήκει στο σύνολο των διανυσμάτων copoint ως προς κάποιο σημείο αναφοράς (r, s) εάν και μόνο εάν η διεύθυνση του $\mathbf{u}$ είναι κάθετη στην ευθεία που ενώνει τα (x, y) και (r, s), δηλαδή ισούται με $(-y + s, x - r)$ (σχήμα 5.6). Το μέτρο του μοναδιαίου (r, s) διανύσματος copoint στο



Σχήμα 5.5: Παράδειγμα ομαδοποίησης διανυσμάτων διαφορετικών διευθύνσεων

Στα τέσσερα σημεία (x, y) τα διανύσματα με τη στέρεη γραμμή είναι η οπτική ροή που εκφράζει τη μεταφορική κίνηση. Τα διανύσματα με τη διακεκομμένη γραμμή ορίζονται ως κάθετα στην ευθεία που ενώνει τα (r, s) και (x, y). Σε όλα τα σημεία πάνω από την ευθεία που διέρχεται από το FOE και το τυχαίο (r, s) η γωνία που σχηματίζουν τα διανύσματα είναι οξεία, άρα η προβολή του διανύσματος οπτικής ροής στο δεύτερο διάνυσμα (δηλαδή η κάθετη ροή) είναι θετική. Αντίθετα, κάτω από την ευθεία όλες οι γωνίες είναι αμβλείες, επομένως όλες οι προβολές είναι αρνητικές.



Σχήμα 5.6: Διανύσματα (r,s) copoint

Στο σχήμα φαίνονται τα διανύσματα (r, s) copoint. Σε κάθε σημείο (x, y) τα διανύσματα έχουν διεύθυνση $(-y + s, x - r)$.

σημείο (x, y) έχει μέτρο προφανώς,

$$\frac{(-y+s, x-r)}{\sqrt{(x-r)^2 + (y-s)^2}}$$

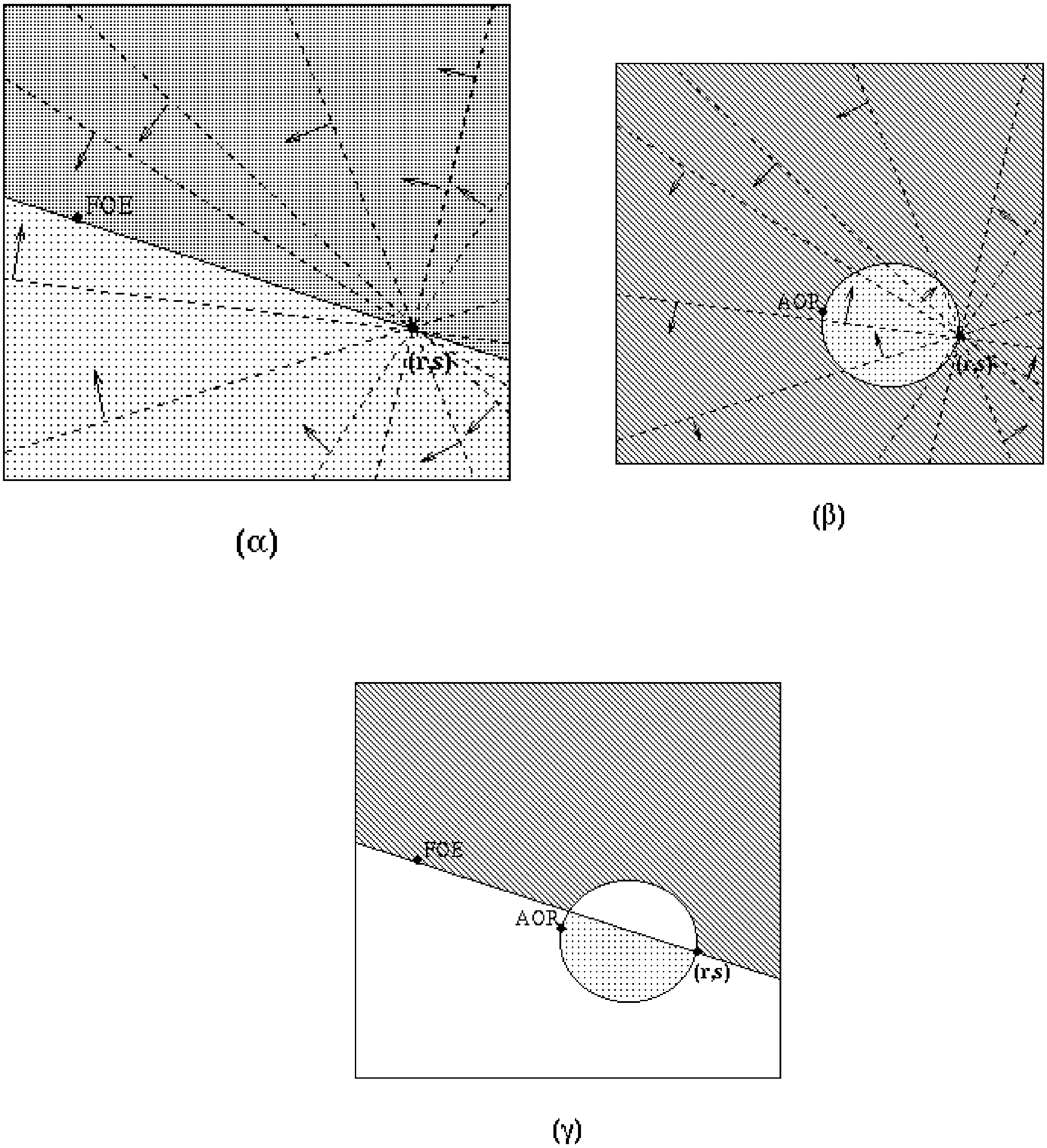
Με διαδικασία ανάλογη με την απλή περίπτωση σταθερού μέτρου και διεύθυνσης (βλ. Παράρτημα Β), προκύπτει η παρακάτω ανίσωση για τα (r, s) διανύσματα copoint με μέτρο $u_n(x, y)$:

$$m(r, s, x_0, y_0; x, y)(u_n(x, y) - n(r, s, \alpha, \beta, \gamma; x, y)) > 0, \text{όπου}$$

$$\begin{aligned} m(r, s, x_0, y_0; x, y) &= y(x_0 - r) - x(y_0 - s) - x_0 s + y_0 r \\ n(r, s, \alpha, \beta, \gamma; x, y) &= -x^2(\beta s + \gamma f) - y^2(\alpha r + \gamma f) + (xy(\alpha s + \beta r) + xf(\alpha f + \gamma r) + \\ &\quad yf(\beta f + \gamma s) - f^2(\alpha r - \beta s) \end{aligned}$$

, όπου (x_0, y_0) το FOE. Η ευθεία $m(r, s, x_0, y_0; x, y)$ διαχωρίζει τις αρνητικές από τις θετικές μεταφορικές συνιστώσες των διανυσμάτων. Η καμπύλη από την άλλη, όταν το μέτρο $u_n(x, y)$ ισούται με 0, διαχωρίζει τις περιστροφικές αρνητικές και θετικές συνιστώσες (σχήμα 5.7).

Η μέθοδος εντοπισμού του FOE που απορρέει από μια τοπογραφία όπως η παραπάνω είναι απλή. Τόσο το FOE ($\frac{Uf}{W}, \frac{Vf}{W}$) όσο και το (r, s) μηδενίζουν την εξίσωση της ευθείας m, άρα η ευθεία διέρχεται από τα σημεία αυτά. Εφόσον η ακριβής θέση του (r, s) είναι γνωστή, ψάχνουμε σε ευθείες με διάφορες διευθύνσεις αλλά που πάντα περνούν από το (r, s) να



Σχήμα 5.7: Πρόσημο διανυσμάτων copoint

Τα διανύσματα (r, s) copoint λόγω μεταφοράς σχηματίζουν δύο ημιεπίπεδα όπου στο ένα μαζεύονται όλα τα αρνητικά (αραιή περιοχή με κουκίδες) και στο άλλο όλα τα θετικά (πυκνή περιοχή με κουκίδες). Η γραμμή διαχωρισμού διέρχεται από τα FOE και (r, s) (σχήμα (α)). Τα διανύσματα (r, s) copoint λόγω περιστροφής είναι αρνητικά όταν βρίσκονται μέσα σε μια δευτέρας τάξης καμπύλη (αραιή περιοχή με κουκίδες), η οποία διέρχεται από τα AOR και (r, s), και είναι θετικά σε όλες τις άλλες θέσεις (αραιή περιοχή με κουκίδες) (σχήμα (β)). Στη γενική περίπτωση συμπαγούς κίνησης τα διανύσματα κάθετης ροής που ανήκουν στο σύνολο των διανυσμάτων (r, s) copoint κατανέμονται σε τέσσερις περιοχές (σχήμα (γ)). Οπου $u_{trans} > 0$ και $u_{rot} > 0$ (αραιή περιοχή με κουκίδες) η ροή είναι αρνητική, όπου $u_{trans} < 0$ και $u_{rot} < 0$ (πυκνή περιοχή με κουκίδες) η ροή είναι θετική, όπου, τέλος, $u_{trans} < 0$ και$u_{rot} > 0$ ή $u_{trans} > 0$ και $u_{rot} < 0$ (περιοχές δίχως κουκίδες) αναμειγνύονται θετικά, αρνητικά και μηδενικά διανύσματα. Θετική θεωρούμε τη φορά των διανυσμάτων που "κινούνται" αντίθετα από τους δείκτες των ρολογιών.

βρούμε ποια ανταποκρίνεται καλύτερα στο μοντέλο διαχωρισμού θετικών και αρνητικών τιμών. Επανάληψη της όλης διαδικασίας για διαφορετικά σημεία αναφοράς (r, s) και συγκερασμός των αποτελεσμάτων οδηγεί στην τοποθέτηση του FOE στην τομή δύο ή περισσότερων ζωνών[1].

5.3.2 Ομοαξονικά διανύσματα

Ενας ακόμη πιο γενικός τρόπος θεώρησης της εύρεσης διανυσμάτων κάθετης ροής με συγκεκριμένες αλλά όχι σταθερές τιμές σε μέτρο και διεύθυνση είναι εκείνος που χρησιμοποιεί τα ***ομοαξονικά διανύσματα*** (*coaxis vectors*), τα οποία ορίζονται ως προς μια τριάδα πραγματικών αριθμών (A, B, C) για τους οποίους ισχύει $A^2 + B^2 + C^2 = 1$. Κάθε τέτοιο διάνυσμα (A, B, C) ορίζει μια άπειρη κλάση από κώνους με άξονα (A, B, C) και κορυφή την αρχή των αξόνων. Η τομή των κώνων με το επίπεδο προβολής δημιουργεί κωνικές τομές που ονομάζονται ***γραμμές πεδίου A,B,C*** (*(A,B,C) field lines*). Οι γραμμές αυτές είναι ακριβώς οι γραμμές πάνω στις οποίες θα κινούνταν τα σημεία του επιπέδου προβολής αν ο παρατηρητής περιστρεφόταν γύρω από τον άξονα (A, B, C). Ενα διάνυσμα κάθετης ροής στο σημείο (x, y) ανήκει στην κλάση των (A, B, C) coaxis διανυσμάτων, εάν και μόνο εάν η διεύθυνσή του είναι ίδια με εκείνη του διανύσματος $\vec{M} = (M_x, M_y)$ (σχήμα 5.8), όπου[2]

$$(M_x, M_y) = ((-A(y^2 + f^2) + Bxy + Cxf), (Axy - B(x^2 + f^2) + Cyf))$$

Με την ίδια λογική όπως και στις δύο προηγούμενες περιπτώσεις, επιθυμούμε να βρούμε γεωμετρικές καμπύλες που να χαρακτηρίζουν και να χαρακτηρίζονται από τα FOE και AOR και που να διαχωρίζουν αρνητικές και θετικές τιμές των μεταφορικών και περιστροφικών συνιστωσών. Με ανάλογο τρόπο όπως και πριν (βλ. Παράρτημα B) προκύπτουν οι ανισότητες:

$$k(A, B, C, x_0, y_0; x, y)(u_n(x, y) - l(A, B, C, \alpha, \beta, \gamma; x, y)) > 0$$

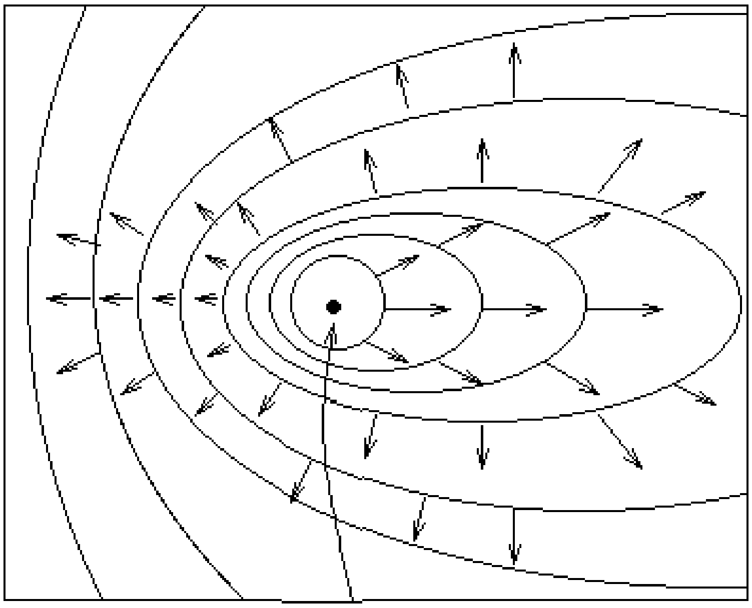
όπου,

$$\begin{aligned} k(A, B, C, x_0, y_0; x, y) &= x^2(Cf + By_0) + y^2(Cf + Ax_0) - xy(Ay_0 + Bx_0) - \\ &\quad xf(Af + Cx_0) - yf(Bf + Cy_0) + f^2(Ax_0 + By_0) \\ l(A, B, C, \alpha, \beta, \gamma; x, y) &= y(\alpha C - \gamma A) - x(\beta C - \gamma B) + \beta Af - \alpha Bf \end{aligned}$$

, όπου (x_0, y_0) το FOE. Αντίθετα με την περίπτωση των διανυσμάτων copoint, στην προκειμένη περίπτωση η ευθεία k είναι εκείνη που διαχωρίζει τις περιστροφικές συνιστώσες, ενώ οι

[1] Κάθε ζώνη εκφράζει μια περιοχή όπου βρίσκεται η αντίστοιχη ευθεία.

[2] Η διεύθυνση (M_x, M_y) στο σημείο (x, y) προκύπτει ως εξής: α) Γράφεται η εξίσωση του κύκλου που περιγράφει την περιστροφική κίνηση των σημείων του χώρου με άξονα περιστροφής το διάνυσμα (A, B, C). β) Υπολογίζεται η εξίσωση της προβολής του κύκλου στο επίπεδο προβολής. γ) Υπολογίζεται η εφαπτομένη της εξίσωσης της κωνικής τομής στο σημείο (x, y).



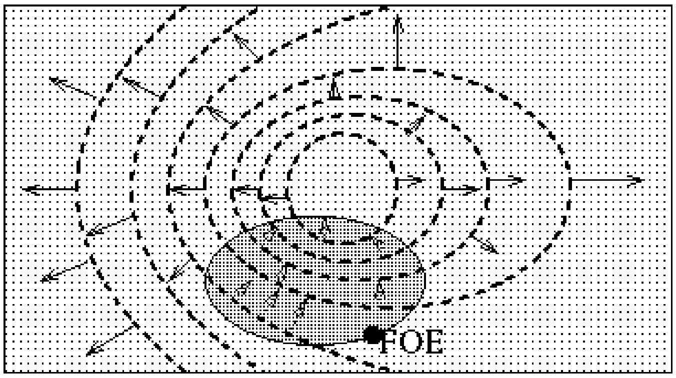
Σχήμα 5.8: Διανύσματα (A,B,C) coaxis

Σρο σχήμα φαίνονται τα διανύσματα (A,B,C) coaxis. Σε κάθε σημείο (x, y) οι διευθύνσεις των διανυσμάτων είναι $(M_x, M_y) = ((-A(y^2 + f^2) + Bxy + Cxf), (Axy - B(x^2 + f^2) + Cyf))$.

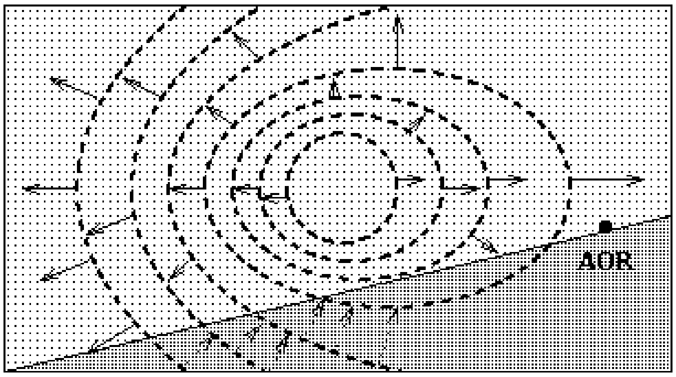
μεταφορικές, όταν το μέτρο $u_n(x, y)$ ισούται με 0, διαχωρίζονται από την καμπύλη l (σχήματα 5.9).

Παρατηρούμε ότι το FOE βρίσκεται πάνω στην καμπύλη k και το AOR βρίσκεται πάνω στην ευθεία l, καθώς οι συντεταγμένες τους, (x_0, y_0) και $(\frac{\alpha f}{\gamma}, \frac{\beta f}{\gamma})$ μηδενίζουν τις αντίστοιχες εξισώσεις. Αντίθετα, όμως, με τις περιπτώσεις της σταθερής κάθετης ροής και των διανυσμάτων (r, s) copoint, αγνοούμε τόσο τη διεύθυνση όσο και το ύψος της ευθείας l καθώς οι παράμετροι αυτές εξαρτώνται από τα διανύσματα (α, β, γ) και (A, B, C). Επομένως, για να μπορέσουμε να χρησιμοποιήσουμε τα διανύσματα coaxis (A, B, C) με ένα τρόπο παρόμοιο με εκείνο των διανυσμάτων copoint και των διανυσμάτων σταθερής κάθετης ροής, οφείλουμε να απλοποιήσουμε τις εξισώσεις τους. Αυτό γίνεται διαλέγοντας ειδικές τιμές για τα A, B και C. Συγκεκριμένα εξετάζουμε τη μορφή των εξισώσεων για τις εξής τρεις περιπτώσεις:

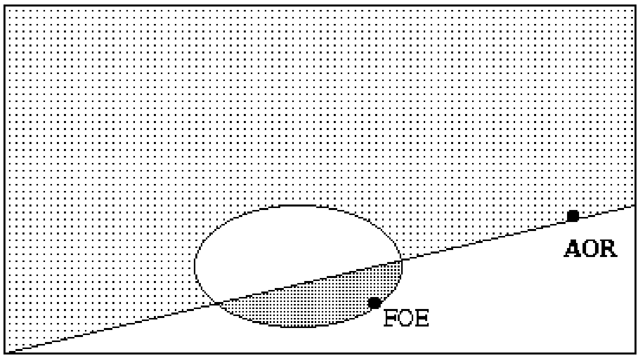
- $(A, B, C) = (1, 0, 0)$, οπότε τα (A, B, C) coaxis ονομάζονται α διανύσματα
- $(A, B, C) = (0, 1, 0)$, οπότε τα (A, B, C) coaxis ονομάζονται β διανύσματα



(α)



(β)



(γ)

Σχήμα 5.9: Πρόσημο διανυσμάτων coaxis

Τα διανύσματα (A, B, C) coaxis λόγω μεταφοράς είναι αρνητικά όταν βρίσκονται μέσα σε μια δευτέρας τάξης καμπύλη (πυκνή περιοχή με κουκίδες), η οποία διέρχεται από το FOE, και είναι θετικά σε όλες τις άλλες θέσεις (αραιή περιοχή με κουκίδες) (σχήμα (α)). Τα διανύσματα (A, B, C) coaxis λόγω περιστροφής σχηματίζουν δύο ημιεπίπεδα όπου στο ένα μαζεύονται όλα τα αρνητικά (πυκνή περιοχή με κουκίδες) και στο άλλο όλα τα θετικά (αρνητική περιοχή με κουκίδες). Η γραμμή διαχωρισμού διέρχεται από το AOR (σχήμα (β)). Στη γενική περίπτωση συμπαγούς κίνησης τα διανύσματα κάθετης ροής που ανήκουν στο σύνολο των διανυσμάτων (A, B, C) coaxis κατανέμονται σε τέσσερις περιοχές (σχήμα (γ)). Οπου $u_{trans} > 0$ και $u_{rot} > 0$ (αραιή περιοχή με κουκίδες) η ροή είναι θετική, όπου $u_{trans} < 0$ και $u_{rot} < 0$ (πυκνή περιοχή με κουκίδες) η ροή είναι αρνητική, όπου, τέλος, $u_{trans} < 0$ και$u_{rot} > 0$ ή $u_{trans} > 0$ και $u_{rot} < 0$ (περιοχές δίχως κουκίδες) αναμειγνύονται θετικά, αρνητικά και μηδενικά διανύσματα. Θετική θεωρούμε τη φορά των διανυσμάτων που "απομακρύνονται" από το σημείο $P(\frac{Af}{C}, \frac{Bf}{C})$.

- $(A, B, C) = (0, 0, 1)$, οπότε τα (A, B, C) coaxis ονομάζονται γ διανύσματα

α διανύσματα

Οι εξισώσεις της καμπύλης και της ευθείας μεταγράφονται ως:

$$k(A, B, C, x_0, y_0; x, y) = 0, A = 1, B = 0, C = 0 \Rightarrow$$

$$x^2(0f+0y_0)+y^2(0f+1x_0)-xy(1y_0+0x_0)-xf(1f+0x_0)-yf(0f+0y_0)+f^2(1x_0+0y_0) = 0 \Rightarrow$$

$$y^2x_0 - xyy_0 - xf^2 + f^2x_0 = 0$$

και

$$l(A, B, C, \alpha, \beta, \gamma; x, y) = 0, A = 1, B = 0, C = 0 \Rightarrow$$

$$y(\alpha 0 - \gamma 1) - x(\beta 0 - \gamma 0) + \beta 1 f - \alpha 0 f = 0 \Rightarrow$$

$$-y\gamma + \beta 1 f = 0 \Rightarrow$$

$$y = \frac{\beta f}{\gamma}$$

Παρατηρούμε ότι η καμπύλη k παίρνει τη μορφή υπερβολής. Επίσης, η ευθεία l αποκτά γνωστή πλέον διεύθυνση και συγκεκριμένα παράλληλη με τον άξονα X. Το σχήμα διαχωρισμού θετικών και αρνητικών διανυσμάτων παίρνει τη μορφή του σχ. 5.10α.

β διανύσματα

$$k(A, B, C, x_0, y_0; x, y) = 0, A = 0, B = 1, C = 0 \Rightarrow$$

$$x^2(0f+1y_0)+y^2(0f+0x_0)-xy(0y_0+1x_0)-xf(0f+0x_0)-yf(1f+0y_0)+f^2(0x_0+1y_0) = 0 \Rightarrow$$

$$x^2y_0 - xyx_0 - yf^2 + f^2y_0 = 0$$

και

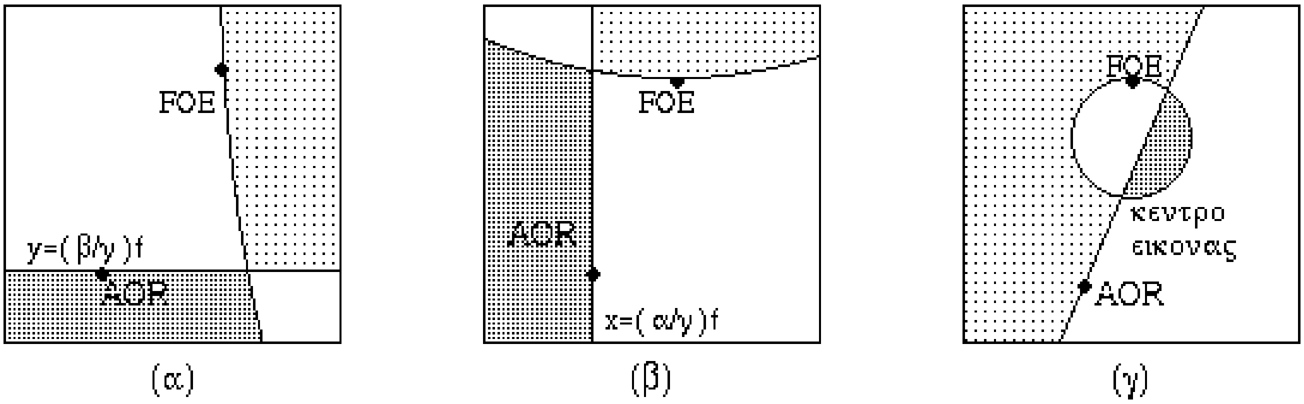
$$l(A, B, C, \alpha, \beta, \gamma; x, y) = 0, A = 0, B = 1, C = 0 \Rightarrow$$

$$y(\alpha 0 - \gamma 0) - x(\beta 0 - \gamma 1) + \beta 0 f - \alpha 1 f = 0 \Rightarrow$$

$$x\gamma - \alpha f = 0 \Rightarrow$$

$$x = \frac{\alpha f}{\gamma}$$

Οπως και πριν, η καμπύλη k παίρνει τη μορφή υπερβολής. Η ευθεία l αποκτά διεύθυνση παράλληλη με τον άξονα Y. Το σχήμα διαχωρισμού θετικών και αρνητικών διανυσμάτων παίρνει τη μορφή του σχ. 5.10β.



Σχήμα 5.10: Σχέση FOE και AOR με α, β, γ διανύσματα

Το (α) είναι η απλοποίηση του 5.9(γ), όταν $(A, B, C) = (1, 0, 0)$ (διανύσματα α). Το (β) είναι η αντίστοιχη απλοποίηση του **??**(γ), όταν $(A, B, C) = (0, 1, 0)$ (διανύσματα β). Το (γ), τέλος, είναι η απλοποίηση του **??**(γ), όταν $(A, B, C) = (0, 0, 1)$ (διανύσματα γ). Και οι τρεις ευθείες των σχημάτων διέρχονται από το AOR. Αντίθετα, οι καμπύλες διέρχονται από το FOE.

γ διανύσματα

$$k(A, B, C, x_0, y_0; x, y) = 0, A = 0, B = 0, C = 1 \Rightarrow$$

$$x^2(1f + 0y_0) + y^2(1f + 0x_0) - xy(0y_0 + 0x_0) - xf(0f + 1x_0) - yf(0f + 1y_0) + f^2(0x_0 + 0y_0) = 0 \Rightarrow$$

$$x^2 + y^2 - xx_0 - yy_0 = 0$$

και

$$l(A, B, C, \alpha, \beta, \gamma; x, y) = 0, A = 0, B = 0, C = 1 \Rightarrow$$

$$y(\alpha 1 - \gamma 0) - x(\beta 1 - \gamma 0) + \beta 0 f - \alpha 0 1 f = 0 \Rightarrow$$

$$y\alpha - x\beta = 0$$

Στην περίπτωση των γ διανυσμάτων η κωνική τομή k απλουστεύεται σε κύκλο. Τα FOE και το κέντρο της εικόνας βρίσκονται σε αντιδιαμετρικά σημεία[3]. Η ευθεία l αποκτά κλίση $\frac{\alpha}{\beta}$. Το μοντέλο διαχωρισμού θετικών και αρνητικών διανυσμάτων παίρνει τη μορφή του σχ. 5.10γ.

[3]Η αντιδιαμετρικότητα των σημείων (0,0) και FOE προκύπτει ως εξής: Θεωρούμε ένα σημείο P (x, y) της κυκλικής περιφέρειας. Θα πρέπει η γωνία OPF, όπου O το κέντρο της εικόνας και F το FOE, να είναι κάθετη, άρα $\vec{OP} \cdot \vec{PF} = 0 \Rightarrow (x, y) \cdot (x - x_0, y - y_0) = 0 \Rightarrow x^2 - xx_0 + y^2 - yy_0 = 0$. Αυτό όμως, όπως φαίνεται και από την εξίσωση της καμπύλης k, ισχύει, αφού το (x, y) είναι σημείο του κύκλου.

Κεφάλαιο 6

Παρουσίαση Αλγορίθμων

Στο κεφάλαιο αυτό αναφέρονται συνοπτικά οι αλγόριθμοι που υλοποιήθηκαν και εκτιμώνται τα αποτελέσματα τους. Θεωρούμε πάντοτε ως είσοδο των αλγορίθμων το πεδίο κάθετης ροής, το οποίο παράγεται από δύο διαδοχικές εικόνες που καταγράφηκαν από την κάμερα του κινούμενου παρατηρητή.

6.1 Εντοπισμός FOE μέσω σταθερής κάθετης ροής

Ανατρέχοντας στο προηγούμενο κεφάλαιο, στο κομμάτι που αναφέρεται στην προσπάθεια εντοπισμού των διανυσμάτων κάθετης ροής συγκεκριμένου μέτρου u_n και διεύθυνσης (n_x, n_y), προέκυψαν κλιμακωτά τα εξής συμπεράσματα:

1. υπάρχει μια ευθεία που χωρίζει τις θετικές από τις αρνητικές μεταφορικές συνιστώσες των διανυσμάτων[1]. Η ευθεία αυτή έχει γνωστή κλίση (κάθετη στη (n_x, n_y)) και διέρχεται από το σημείο FOE.

2. Υπάρχει μια υπερβολή η οποία χωρίζει τις θετικές από τις αρνητικές περιστροφικές συνιστώσες των διανυσμάτων[1]. Η υπερβολή αυτή όταν $u_n = 0$ περνάει από το AOR.

3. Τα διανύσματα με μέτρο ακριβώς u_n και διεύθυνση (n_x, n_y) μπορούν να βρεθούν μόνο σε δύο από τα τέσσερα "τεταρτημόρια" που σχηματίζουν η ευθεία με την καμπύλη.

4. Η κατανομή των διανυσμάτων παρουσιάζει λόγω των περιορισμών ελάχιστου και μέγιστου βάθους ένα "στένεμα" στις περιοχές πάνω και κάτω από την ευθεία.

[1] Οι όροι "θετικές" και "αρνητικές" εδώ χρησιμοποιείται καταχρηστικά. Στην πραγματικότητα εννοείται τιμές μεγαλύτερες και μικρότερες από κάποια τιμή u_n.

5. Σε ομάδες διανυσμάτων με την ίδια διεύθυνση, (n_x, n_y), αλλά διαφορετικά μέτρα, $u_{n0}, u_{n1}, u_{n2}, \ldots$, αντιστοιχούν "παράλληλες" υπερβολές αλλά μόνο μια ευθεία, αφού, όπως δείχθηκε, η εξίσωσή της είναι ανεξάρτητη της τιμής του μέτρου. Παρατηρούμε ότι το FOE δεν μπορεί να βρίσκεται εκεί που αλληλοκαλύπτονται οι κατανομές παράλληλων διανυσμάτων κάθετης ροής, αφού τα ζευγάρια υπερβολών που τις ορίζουν δεν τέμνονται πάνω στην ευθεία, στην οποία αυτό ανήκει.

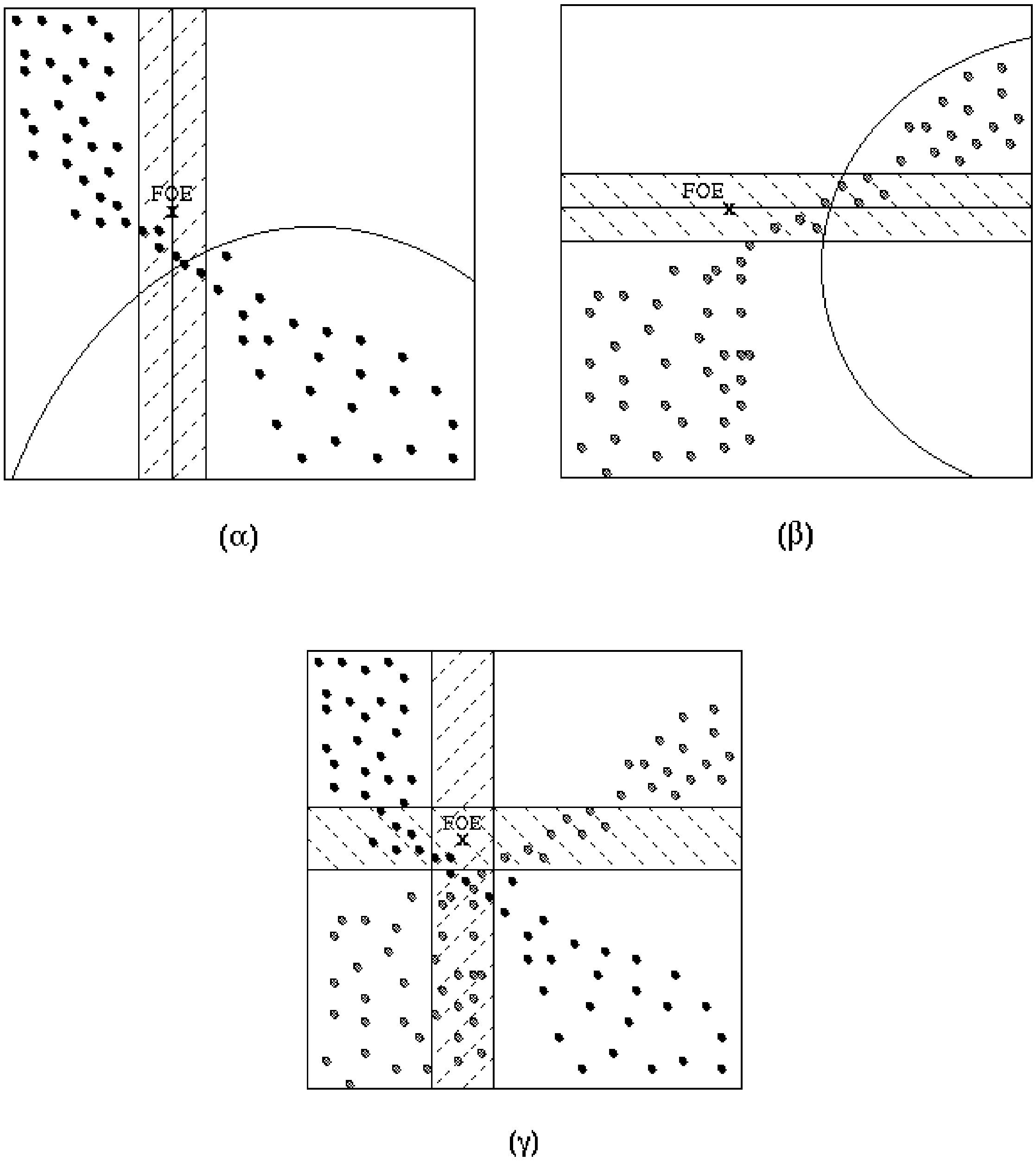
Στόχοι του αλγορίθμου είναι

- σύμφωνα με τα συμπεράσματα 1 και 4, η ανακάλυψη περιοχών όπου το FOE πιθανότατα βρίσκεται (σχήμα (6.1)).
- σύμφωνα με τα συμπεράσματα 1 και 5, ο αποκλεισμός περιοχών όπου το FOE δεν μπορεί να βρίσκεται (σχήμα (6.2)).

Πριν περιγραφεί ο αλγόριθμος, αφιερώνονται μερικές γραμμές στην εξήγηση τριών παραμέτρων του αλγορίθμου οι οποίες κρατούν ιδιαίτερο ρόλο στην υλοποίηση του.

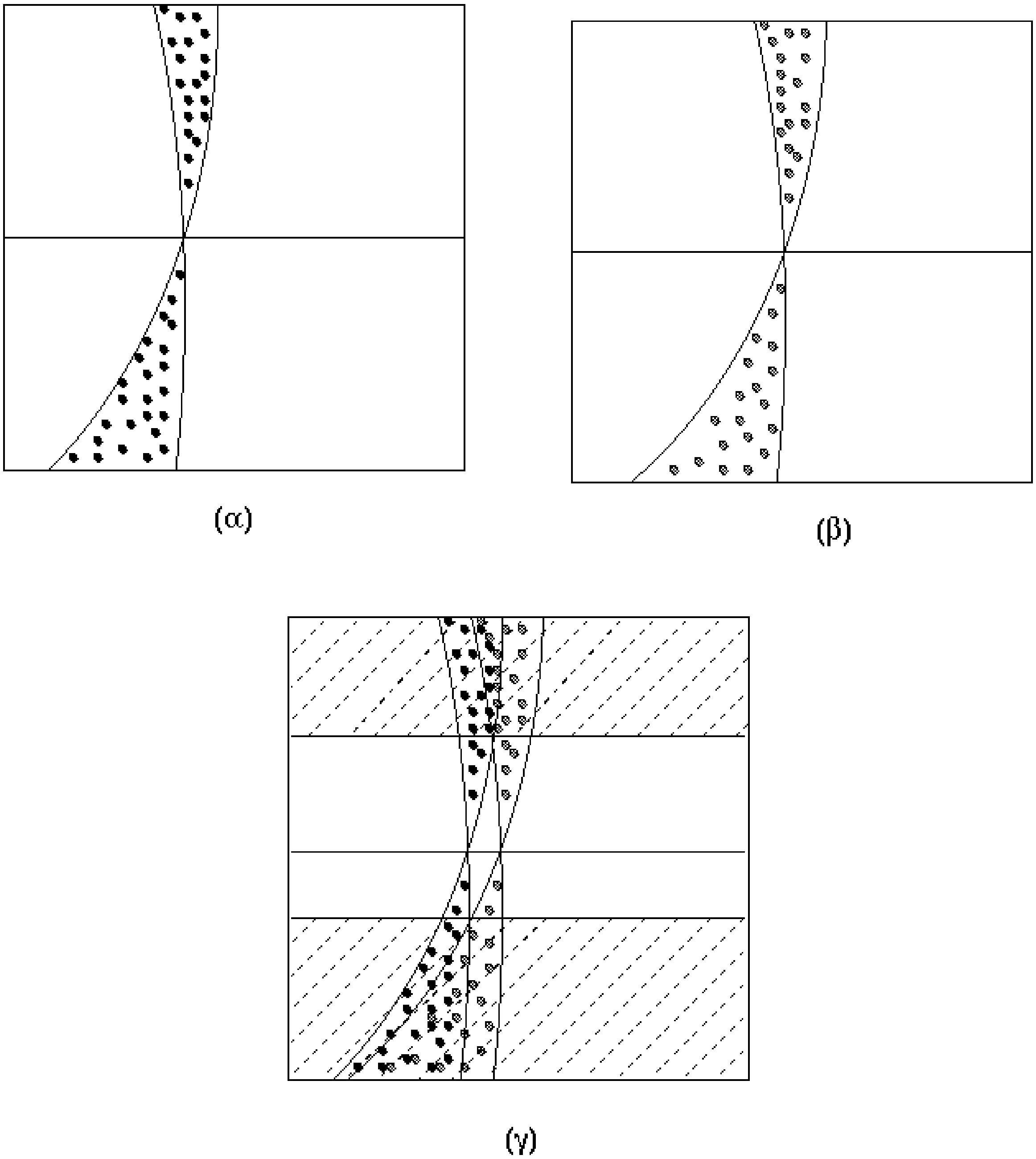
Μια και ασχολούμαστε με πραγματικούς αριθμούς είναι φυσιολογικό να περιμένουμε ότι λόγω των αποκλίσεων στα τελευταία δεκαδικά τους ψηφία ελάχιστα πανομοιότυπα διανύσματα θα ανακαλύπτουμε σε ολόκληρη την εικόνα. Γι' αυτό και συνειδητά αντικαθιστούμε τη συνέχεια με τη διακριτότητα, υιοθετώντας τόσο για τη διεύθυνση όσο και για το μέτρο τη γνωστή από τον απειροστικό λογισμό έννοια της ε - γειτνίασης. Ορίζουμε έτσι δύο μεταβλητές ε_{angle} και ε_{length} , οι οποίες ουσιαστικά δημιουργούν κλάσεις μέτρων και διευθύνσεων. Μέσα σε κάθε κλάση όλοι οι αριθμοί θεωρούνται ίσοι. Π.χ. στην κλάση διευθύνσεων 0 ανήκουν όλες οι διευθύνσεις dir για τις οποίες ισχύει $dir \epsilon [0-\varepsilon_{angle}, 0+\varepsilon_{angle})$. Οι κλάσεις αυτές διαδραματίζουν διπλό ρόλο στον αλγόριθμο, διότι και συνηγορούν στη "διακριτή" ισότητα των αριθμών και καθορίζουν την πολυπλοκότητα του αλγορίθμου, μια και οι δυνατές συγκρίσεις των κατανομών των διανυσμάτων του πεδίου ροής εξαρτώνται από το πλήθος των διαφορετικών τιμών που μπορούν εκείνα να λάβουν. Τυπική τιμή για το ε_{angle} είναι η $0,39 (= \frac{\pi}{8})$ που χωρίζει το κύκλο των διευθύνσεων σε 8 τομείς-κλάσεις (σχήμα 6.3). Μια τυπική τιμή του ε_{length} είναι η 0.07 που για διανύσματα με μέτρο μέχρι 1.5 (το οποίο είναι συνηθισμένη περίπτωση μια και τα διαδοχικά frames καταγράφονται πολύ σύντομα το ένα μετά το άλλο, οπότε και οι ακμές δεν προλαβαίνουν να απομακρυνθούν από την προηγούμενη θέση τους) δημιουργεί 12 κλάσεις μέτρων.

Μια άλλη παράμετρος του συστήματος είναι η $BoxSize$, το μέγεθος δηλαδή της πλευράς των τετραγώνων με τα οποία χωρίζουμε την εικόνα του πεδίου κάθετης ροής. Η κατάτμηση



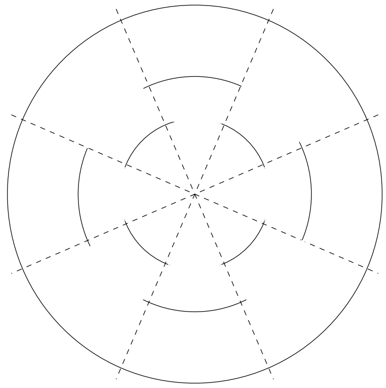
Σχήμα 6.1: Προσδιορισμός FOE από τομή ευθειών

Στο σχήμα (α) παρουσιάζονται τα διανύσματα κάθετης ροής μέτρου u_{n1} και διεύθυνσης 0 (μαύρες κουκίδες). Η ευθεία που περιέχει το FOE και χωρίζει τις θετικές από τις αρνητικές μεταφορικές συνιστώσες έχει κλίση $\frac{\pi}{2}$ και βρίσκεται εκεί που η κατανομή στενεύει. Στο σχήμα (β) παρουσιάζονται τα διανύσματα κάθετης ροής με μέτρο u_{n2} και διεύθυνση $\frac{\pi}{2}$ (γκρι κουκίδες). Η ευθεία που περιέχει το FOE και χωρίζει τις θετικές από τις αρνητικές μεταφορικές συνιστώσες των διανυσμάτων έχει κλίση 0 και βρίσκεται εκεί που η κατανομή των νέων διανυσμάτων στενεύει. Στο (γ) προσδιορίζεται μια περιοχή για το FOE από την τομή των ζωνών που δηλώνουν το "στένεμα" των κατανομών διανυσμάτων.



Σχήμα 6.2: Αποκλεισμός FOE λόγω αλληλοκάλυψης των κατανομών

Στο σχήμα (α) παρουσιάζονται τα διανύσματα κάθετης ροής μέτρου u_{n1} και διεύθυνσης $\frac{\pi}{2}$ (μαύρες κουκίδες). Στο σχήμα (β) παρουσιάζονται τα διανύσματα με ίδια διεύθυνση και μέτρο u_{n2} διάφορο του u_{n1} (γκρι κουκίδες). Στο σχήμα (γ) παρουσιάζονται ταυτόχρονα οι κατανομές στην εικόνα και των δύο ομάδων διανυσμάτων. Επάνω και κάτω στην εικόνα διακρίνονται οι δύο περιοχές όπου υπάρχουν διανύσματα και των δύο κατανομών. Η ζώνες που έχουν διεύθυνση ίδια με της ευθείας που περιέχει το FOE, συγκεκριμένα στο παράδειγμά μας 0 μοίρες, και οι οποίες περνάνε από περιοχές όπου οι δύο κατανομές αλληλοκαλύπτονται θεωρούνται ως υποψήφιες μη κατοχής του FOE.



Σχήμα 6.3: Χωρισμός των διευθύνσεων σε κλάσεις-τομείς.

Η τιμή του ε_{angle} καθορίζει τον αριθμό των κλάσεων διευθύνσεων ($\frac{\pi}{8} \leftrightarrow 8, \frac{\pi}{4} \leftrightarrow 4$). Η τιμή αυτή μαζί με κείνη του ε_{length} επηρεάζει επίσης την πολυπλοκότητα του αλγορίθμου.

αυτή της εικόνας συντελείται για να αποκτήσουμε ένα κριτήριο απόφασης για το αν δυο κατανομές διαφορετικών μέτρων αλληλοκαλύπτονται. Ακριβή γνώση της αλληλοκάλυψης ή μη αλληλοκάλυψης θα αποκτούσαμε μόνο εάν λύναμε γραμμικά συστήματα ανισώσεων που θα ανίχνευαν την παρουσία κάποιου ή κάποιων διανυσμάτων με μέτρο u_{n1} μέσα σε τρίγωνα που θα σχηματίζονταν από τριάδες διανυσμάτων με μέτρο u_{n2}. Αντί για τη χρονοβόρα αυτή διαδικασία προτιμήθηκε η αλληλοκάλυψη των κατανομών να αποφασίζεται απαντώντας το ερώτημα αν σ' ένα στοιχειώδες τετράγωνο βρίσκονται διανύσματα μήκους και u_{n1} και u_{n2}. Προσοχή απαιτείται για μία ορθή επιλογή της $BoxSize$. Αν είναι πιο μεγάλη απ' ότι πρέπει τα τετράγωνα θα περικλείουν συχνά διανύσματα διαφορετικών μέτρων, χωρίς όμως αυτό να ανταποκρίνεται στην πραγματική κατάσταση της αλληλοκάλυψής τους. Από την άλλη, ενδεχόμενη εκχώρηση μιας υπέρ του δέοντος μικρής τιμής στη $BoxSize$, θα επέφερε το αντίθετο αποτέλεσμα. Πολύ δύσκολα πλέον θα ανακαλυπτόταν περιοχή αλληλοκάλυψης και ένας από τους κύριους στόχους του αλγορίθμου θα καθιστόταν *apriori* ανέφικτος. Από μετρήσεις βρέθηκε ότι μια ικανοποιητική τυπική τιμή για τη $BoxSize$ είναι 10 pixels.

Επίσης, πρέπει να σημειωθούν τα εξής σχετικά με το χαρακτήρα του αλγορίθμου. Οπως προαναφέρθηκε, δεύτερο επιδιωκόμενο στόχο του αλγορίθμου αποτελεί ο εντοπισμός περιοχών στην εικόνα όπου το FOE αποκλείεται να βρίσκεται. Η συλλογιστική αποκλεισμών, όμως, συχνά κρύβει παγίδες. Η πιθανότητα κάποιου λάθους στην όλη προεργασία των δεδομένων (στην εύρεση του πεδίου κάθετης ροής ή στην ταξινόμηση των διανυσμάτων σύμφωνα με τις τιμές του μέτρου και της διεύθυνσής τους) ελλοχεύει συνέχεια και άρα θα πρέπει να είμαστε προσεκτικοί με αφορισμούς του στυλ "δεν επανεξετάζονται περιοχές άπαξ και αποδειχτούν μία φορά ακατάλληλες για FOE". Αν και μια τέτοια λογική θα επιτάχυνε την εκτέλεση του αλγορίθμου, θα δημιουργούσε πιθανότατα και δυσάρεστες παρενέργειες όπως ο πλήρης αποπροσανατολισμός του συστήματος, αν για κάποια αιτία το FOE βρισκόταν σε περιοχή που χαρακτηριζόταν ακατάλληλη. Για το λόγο αυτό η υιοθεσία ενός σκληρού, αδιάλλαχτου αλγορίθμου απορρίπτεται και θεσπίζεται ένα πιο "δημοκρατικό" σύστημα ***ψηφίσματος καταλληλότητας*** περιοχής. Σύμφωνα με το σύστημα αυτό, ανάλογα με το πόσο ισχυρή υποψηφιότητα κατοχής του FOE θέτει μια περιοχή (το επίθετο "ισχυρή" αναφέρεται στην εκπλήρωση ή μη κάποιων προϋποθέσεων), προστίθεται στα pixel που την απαρτίζουν μια βαρύνουσα θετική ψήφος. Αντίθετα, όταν κάποια περιοχή θεωρηθεί ισχυρή υποψήφια μη κατοχής του FOE, τότε στα pixel που την απαρτίζουν προστίθεται μία αρνητική ψήφος.

Ο αλγόριθμος που εξετάζει να βρει (και να υπερψηφίσει) τις περιοχές στενέματος των κατανομών διανυσμάτων κάθετης ροής με συγκεκριμένο μέτρο και φορά (σύμφωνα με τα συμπεράσματα 1 και 4) έχει ως εξής:

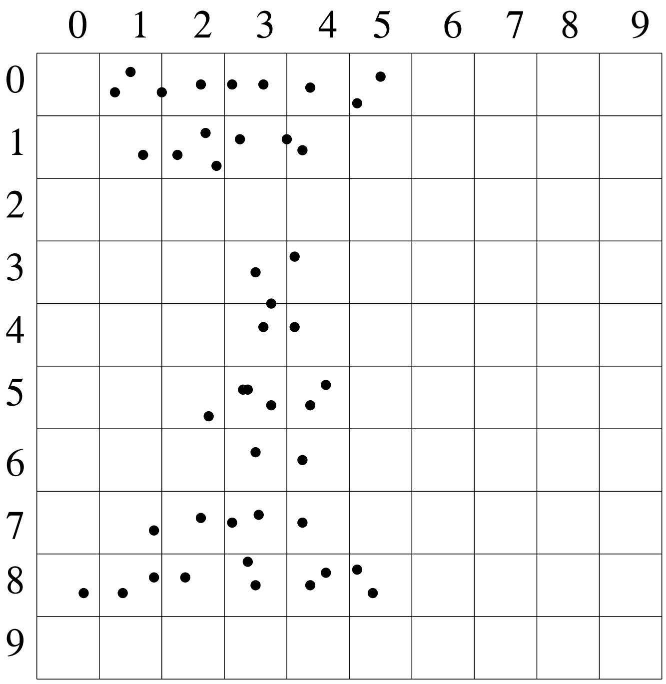
```
# Εύρεση συνόλου κάθετης ροής N
# Καθορισμός διακριτού συνόλου διευθύνσεων D
# Καθορισμός διακριτού συνόλου μέτρων L
# Καθορισμός συνόλου στοιχειωδών τετραγώνων B
# ∀ (d, l, b) ϵ (D × L × B)
    βρες το υποσύνολο κάθετης ροής N_dlb
    ***** διανύσματα στο κουτί b, με μήκος l και διεύθυνση d *****
# ∀ (d_v, l) ϵ (D × L)
    Εάν card(N_{d_v l}) > MinimumPoints **** απαιτείται ελάχιστος αριθμός διανυσμάτων ****
        * Συμπλήρωσε κενά στην κατανομή
        * ∀ ευθεία ε με διεύθυνση d_ε = -1/d_v **** ευθεία κάθετη στα διανύσματα ****
            @ Vote = FitModel1(ε, N_{d_v l}) **** πόσο "καλή" είναι η ευθεία; ****
            @ ∀ b ϵ B και b "δίπλα" στην ε **** ψήφισε σε όλα τα ****
                VoteFOE[b] += Vote        **** στοιχειώδη τετράγωνα της ευθείας ****
```

Ο αλγόριθμος, λοιπόν, ξεκινάει με τη δημιουργία των κλάσεων μέτρου και διεύθυνσης και την επακόλουθη κατάταξη κάθε διανύσματος σε κάποια υπερκλάση (ζεύγος μέτρου και διεύθυνσης), η οποία περιέχει τελικά όλα τα διανύσματα συγκεκριμένου μέτρου και διεύθυνσης. Ταυτόχρονα, για κάθε διάνυσμα βρίσκεται μέσω των καρτεσιανών συντεταγμένων του το στοιχειώδες τετράγωνο στο οποίο ανήκει και η πληροφορία κατακρατείται αθροιστικά για κάθε τετράγωνο, έτσι ώστε στο τέλος της σάρωσης των διανυσμάτων να είναι γνωστό πόσα διανύσματα κάθε υπερκλάσης, δηλαδή συγκεκριμένου μέτρου και διεύθυνσης, ανήκουν σε κάθε τετράγωνο.

Στη συνέχεια ακολουθεί η επαναληπτική, δεύτερη φάση του αλγορίθμου. Η λειτουργία ενός κύκλου της επανάληψης φαίνεται καλύτερα σε ένα εικονικό παράδειγμα. Εστω ότι βρίσκουμε τα διανύσματα κάθετης ροής με μέτρο 1 και διεύθυνση $\frac{\pi}{2}$. Εστω ότι η κατανομή των διανυσμάτων στην εικόνα είναι αυτή του σχήματος 6.4.Η ευθεία που περιέχει το FOE έχει κλίση κάθετη προς τα διανύσματα, δηλαδή 0. Επομένως, για να βρούμε πού στενεύει η κατανομή διανυσμάτων εξετάζουμε όλες τις παράλληλες προς το X άξονα λωρίδες. Ονομάζουμε απόσταση λωρίδας το μέτρο που εκφράζει το πλάτος της κατανομής των διανυσμάτων σε τούτη τη λωρίδα. Ως μονάδα μέτρησης των πλατών κατανομής χρησιμοποιείται το μήκος της πλευράς ενός στοιχειώδους τετραγώνου. Ετσι, στην οριζόντια λωρίδα 0 του παραδείγματός μας η απόσταση είναι 5(=5-1+1), στη λωρίδα 1 4(=4-1+1), στη λωρίδα 3 2(=3-2+1), στη λωρίδα 8 6(=5-0+1) κ.λπ. Παρατηρούμε ότι στις λωρίδες 2 και 9 δε βρέθηκε κάθετη ροή. Επειδή, όμως, τακτική του αλγορίθμου είναι να δίνει ψήφο σε όλες τις περιοχές προβλέπεται η τεχνητή εκχώρηση κάποιας απόστασης σε κάθε τέτοια λωρίδα. Στη λωρίδα 2 π.χ. εκχωρείται η τιμή 3, που προκύπτει από το μέσο όρο των αποστάσεων στις λωρίδες 1 και 3 ($3 = \frac{4+2}{2}$), ενώ στη λωρίδα 9, η οποία βρίσκεται στην άκρη της εικόνας, εκχωρείται η τιμή 6 που αντιστοιχεί απλά στην πλησιέστερη μη κενή λωρίδα.

Στη συνέχεια υπολογίζεται η μέση τιμή των αποστάσεων των λωρίδων. Για το συγκεκριμένο παράδειγμα η μέση τιμή είναι ίση με $3.9 = \frac{5+4+3+3+2+3+3+4+6+6}{10}$. Ενα απλό μέτρο πριμοδότησης των στενών περιοχών της κατανομής αποτελείται από τα τέσσερα παρακάτω κριτήρια:

- αν η απόσταση σε κάποια λωρίδα είναι μεγαλύτερη από τη μέση, τότε η ψήφος είναι ΚΑΚΗ (βλ. λωρίδες 0, 1, 7, 8, 9).
- αν η απόσταση είναι μικρότερη από τη μέση αλλά οι αποστάσεις των πλαϊνών λωρίδων είναι μικρότερες από την εξεταζόμενη, τότε η ψήφος είναι ΜΕΤΡΙΑ (βλ. λωρίδα 5).
- αν η απόσταση είναι μικρότερη από τη μέση και σε μια μόνο από τις διπλανές λωρίδες εμφανίζεται απόσταση μεγαλύτερη από την εξεταζόμενη, τότε η ψήφος είναι ΚΑΛΗ (βλ. λωρίδα 2).
- αν η απόσταση είναι μικρότερη από τη μέση και μικρότερη ή ίση από τις αποστάσεις των διπλανών λωρίδων, τότε η ψήφος είναι ΠΟΛΥ ΚΑΛΗ (βλ. λωρίδες 3, 4, 6).



Πλατος	Ψηφος	Κριτηριο
5	10	Π[ι]>Μεσο
4	10	Π[ι]>Μεσο
3=(4+2)/2	30	Π[ι]<Π[ι−1]&&Π[ι]>Π[ι+1]
2	40	Π[ι]<Π[ι−1]&&Π[ι]=Π[ι+1]
2	40	Π[ι]=Π[ι−1]&&Π[ι]<Π[ι+1]
3	20	Π[ι]>Π[ι−1]&&Π[ι]>Π[ι+1]
2	40	Π[ι]<Π[ι−1]&&Π[ι]<Π[ι+1]
4	10	Π[ι]>Μεσο
6	10	Π[ι]>Μεσο
6	10	Π[ι]>Μεσο

Σχήμα 6.4: Παράδειγμα λειτουργίας αλγορίθμου σταθερής ροής

Στο σχήμα παρουσιάζονται διανύσματα διεύθυνσης $\frac{\pi}{2}$ και μέτρου u_{n1} (μαύρες κουκίδες). Η ευθεία που περιέχει το FOE έχει κλίση 0. Το πλάτος της κατανομής των διανυσμάτων κάθε, οριζόντιας στην περίπτωσή μας, λωρίδας εκφράζεται μέσω της διαφοράς των x συντεταγμένων των ακραίων στοιχειωδών τετραγώνων που περιέχουν διανύσματα του συγκεκριμένου μέτρου και διεύθυνσης. Ανάλογα με τη σχέση του πλάτους κατανομής κάθε λωρίδας με τα αντίστοιχα πλάτη των διπλανών λωρίδων, καθώς και με το μέσο, ελάχιστο και μέγιστο πλάτος, δίδονται ψήφοι σε κάθε στοιχειώδες τετράγωνο.

Η ψήφος προστίθεται σε κάθε στοιχειώδες τετράγωνο που ανήκει στην εξεταζόμενη λωρίδα. Ετσι, στο παράδειγμά μας, αν αντιστοιχήσουμε το ΚΑΚΗ με τον αριθμό 10, το ΜΕΤΡΙΑ με τον αριθμό 20, το ΚΑΛΗ με τον αριθμό 30 και το ΠΟΛΥ ΚΑΛΗ με τον αριθμό 40, οι τελικές ψήφοι κάθε τετραγώνου θα είναι εκείνες που φαίνονται στο σχήμα 6.4. Στην πραγματικότητα το μέτρο ψήφισης είναι πιο περίπλοκο, ώστε να αντανακλά τις αποχρώσεις μεταξύ λωρίδων που εμπίπτουν στο ίδιο κριτήριο. Οι αποχρώσεις αυτές εκφράζουν τη σχέση της απόστασης κάθε λωρίδας με τη μέση, ελάχιστη και μέγιστη απόσταση όλων των λωρίδων. Μ' αυτή τη λογική, στο παράδειγμά μας η λωρίδα 0 θα πάρει χειρότερη ψήφο (στο διάστημα από 1 έως 10) από τη λωρίδα 1, διότι η απόστασή της (5) πλησιάζει περισσότερο τη μέγιστη (6).

Αφού επαναληφθεί η ίδια διαδικασία για όλα τα δυνατά μέτρα και διευθύνσεις, καταμετρούνται οι θετικές ψήφοι κάθε τετραγώνου για να βρεθεί η μέγιστη τιμή. Η τιμή αυτή μπορεί να αντιστοιχεί σε τετράγωνο που περιέχει το FOE, είναι δυνατόν, όμως, να αντιστοιχεί σε κάποιο άλλο, πιθανότατα γειτονικό τού ζητούμενου. Γι' αυτό το λόγο ως τελική λύση προτείνεται το σύνολο των τετραγώνων που συγκεντρώνουν κάποιο προκαθορισμένο ποσοστό ψήφων[2]. Αν, παρ' όλ' αυτά, η προτεινόμενη περιοχή παραμένει σχετικά μεγάλη, το ποσοστό ψήφων μπορεί να αυξάνεται, αποκλείοντας επιπλέον τετράγωνα (και άρα σημεία), μέχρις ότου το εμβαδό της περιοχής μειωθεί κάτω από κάποιο όριο.

Ο αλγόριθμος, τώρα, που ελέγχει (σύμφωνα με τα συμπεράσματα 1 και 5 στην αρχή του κεφαλαίου) την ακαταλληλότητα κατοχής του FOE για κάποια περιοχή έχει ως εξής:

\# Εύρεση συνόλου κάθετης ροής N

\# Καθορισμός διακριτού συνόλου διευθύνσεων D

\# Καθορισμός διακριτού συνόλου μέτρων L

\# Καθορισμός συνόλου στοιχειωδών τετραγώνων B

\# $\forall\ (d, l, b) \in (D \times L \times B)$

  βρες το υποσύνολο κάθετης ροής N_{dlb}

  ***** διανύσματα στο κουτί β, με μήκος λ και διεύθυνση δ *****

\# $\forall\ (d_v, l) \in (D \times L)$

  $\forall\ b \in B : N_{dlb} \neq \emptyset$

    $\forall\ l' > l : N_{dl'b} \neq \emptyset$ **** στο b υπάρχουν διανύσματα με μήκη l και l'

      * NegativeVote = WeightIntersect($N_{dlb}, N_{dl'b}$)

        **** πόσο "αξιόπιστη" είναι η αλληλοκάλυψη; ****

[2]Στα δικά μας παραδείγματα χρησιμοποιήθηκε το 80%

* $\forall\, b' \,\epsilon\, B$

και b' ορίζεται από τη διεύθυνση $\frac{-1}{d_v}$ και το b

VoteFOE[b'] += NegativeVote

Περιγραφικά, ο αλγόριθμος, αφού βρει τα διανύσματα κάθετης ροής που ανήκουν σε κάθε τετράγωνο, εξετάζει ποια τετράγωνα χαρακτηρίζονται ως πεδίο σύμπτωσης διανυσμάτων της ίδιας διεύθυνσης αλλά δύο διαφορετικών μέτρων. Οταν βρεθεί ένα τέτοιο τετράγωνο, ακολουθεί η αξιολόγηση της υποτιθέμενης αλληλοκάλυψης περιοχών. Η αξιολόγηση αυτή βασίζεται σε κριτήρια, όπως:

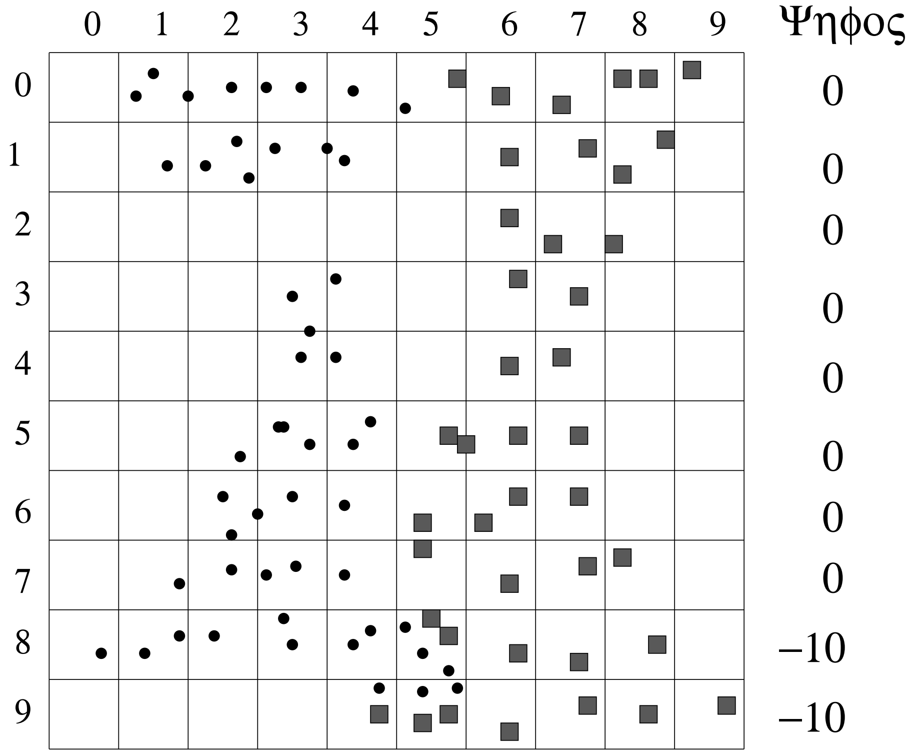
- η διαφορά, $l - l'$, των μέτρων των διανυσμάτων
- το πλήθος, $card(N_{dlb})$ και $card(N_{dl'b})$, των διανυσμάτων κάθετης ροής
- η απόσταση των μέσων τιμών συντεταγμένων των διανυσμάτων κάθε μέτρου

Σύμφωνα με το πρώτο κριτήριο, όσο πιο μεγάλη είναι η διαφορά $l' - l$, τόσο πιο πολύ αυξάνει το βάρος της αρνητικής ψήφου, διότι τόσο μεγαλύτερη γίνεται η πιθανότητα η αλληλοκάλυψη να μην οφείλεται σε απλή γειτνίαση δύο κοντινών μέτρων. Ετσι, πρακτικά, η περίπτωση $l' - l = 1$ δεν εξετάζεται καθόλου, μια και εξ ορισμού οι κατανομές διανυσμάτων με διαδοχικά μέτρα είναι γειτονικές. Αντίθετα, μια διαφορά $l' - l > 1$ αποτελεί ισχυρή ένδειξη αληθούς αλληλοκάλυψης.

Το δεύτερο κριτήριο χρησιμοποιείται για την αποφυγή περιπτώσεων μεμονωμένων εσφαλμένων διανυσμάτων ροής που μπορούν να παραπλανήσουν τον αλγόριθμο. Ετσι, απαιτείται η ύπαρξη ενός ελάχιστου αριθμού από διανύσματα και των δύο μέτρων, ο οποίος να συνηγορεί στο αληθές της ύπαρξης κάθετης ροής λόγω μετακίνησης ακμής στο συγκεκριμένο τετράγωνο.

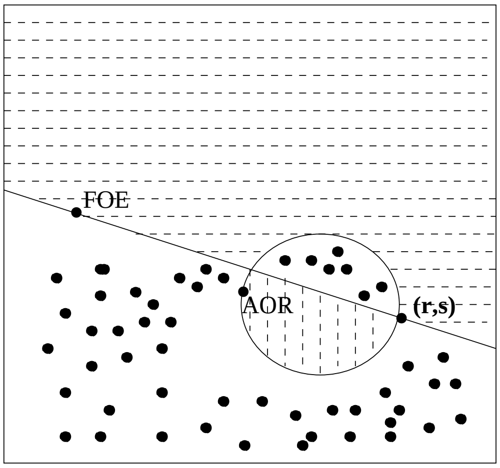
Τέλος, το τρίτο κριτήριο αποτελεί άλλη μια προσπάθεια διασφάλισης εντοπισμού αληθούς και όχι απλά αληθοφανούς αλληλοκάλυψης. Το σημείο που εκφράζει τη μέση τιμή των διανυσμάτων μέτρου l (η πληροφορία για τις συντεταγμένες του παράγεται κατά τη διαδικασία εύρεσης των συνόλων N_{dlb}) θα πρέπει να μην απέχει περισσσότερο από $\frac{BoxSize}{2}$ από το αντίστοιχο σημείο μέσης τιμής των διανυσμάτων μέτρου l'. Αντίθετη περίπτωση θα σήμαινε πιθανότατα ότι όλα τα διανύσματα μέτρου l βρίσκονται π.χ. "αριστερά" στο τετράγωνο και όλα τα διανύσματα μέτρου l' βρίσκονται "δεξιά" στο τετράγωνο, περίπτωση που θα συνεπαγόταν την ανυπαρξία αλληλοκάλυψης των κατανομών.

Με βάση τα παραπάνω κριτήρια εξάγεται από το εξεταζόμενο τετράγωνο b μια αρνητική (ή ακόμα μηδενική) ψήφος. Η ψήφος αυτή δε δίνεται μόνο στο τετράγωνο b, αλλά και σε όλα



Σχήμα 6.5: Παράδειγμα αρνητικής ψήφισης περιοχών

Στο σχήμα διακρίνονται με μαύρες κουκίδες τα διανύσματα με μήκος u_{n1} και διεύθυνση $\frac{\pi}{2}$. Με μαύρα τετραγωνάκια διακρίνονται τα διανύσματα με μήκος u_{n2} και διεύθυνση πάλι $\frac{\pi}{2}$. Αν τα u_{n1} και u_{n2} ήταν διαδοχικά μέτρα οι πιθανές αλληλοκαλύψεις των κατανομών δε θαν επηρέαζαν το τελικό αποτέλεσμα. Εστω ότι $u_{n1} - u_{n2} > 1$. Στην οριζόντια λωρίδα 8 βρίσκεται το τετράγωνο $(5, 8)$ με διανύσματα και των δύο μέτρων. Η διαφορά των δύο "μέσων" σημείων των δύο ομάδων διανυσμάτων δεν απέχει περισσότερο από $\frac{BoxSize}{2}$, οπότε η αλληλοκάλυψη θεωρείται πραγματική και όλα τα τετράγωνα της λωρίδας 8 ψηφίζονται αρνητικά (-10). Το ίδιο ισχύει και για τα τετράγωνα $(4, 9)$ και $(5, 9)$. Αντίθετα, στο τετράγωνο $(5, 0)$ η απόσταση των "μέσων" σημείων είναι μεγαλύτερη από $\frac{BoxSize}{2}$, οπότε η υποτιθέμενη αλληλοκάλυψη δε λαμβάνεται υπόψη και στα τετράγωνα της λωρίδας δεν προστίθεται αρνητική ψήφος.



(γ)

Σχήμα 6.6: Μηδενικά διανύσματα (r, s) coaxis

Οι οριζόντιες γραμμώσεις αντιστοιχούν σε περιοχές με θετικά διανύσματα. Οι κάθετες γραμμώσεις αντιστοιχούν σε περιοχές με αρνητικά διανύσματα. Στις λευκές περιοχές υπάρχουν αναμεμειγμένα θετικά, αρνητικά και μηδενικά διανύσματα. Τα μηδενικά διανύσματα (μαύρες κουκίδες) μπορούν να βρίσκονται μόνο στις λευκές περιοχές.

τα άλλα τετράγωνα που ορίζονται από τη ζώνη που διέρχεται από το b και έχει διεύθυνση $\frac{-1}{d_v}$, εκφράζοντας έτσι το γεγονός μη κατοχής του FOE από κανένα από αυτά. Στο σχήμα (6.5) φαίνεται συνοπτικά η λειτουργικότητα του αλγορίθμου.

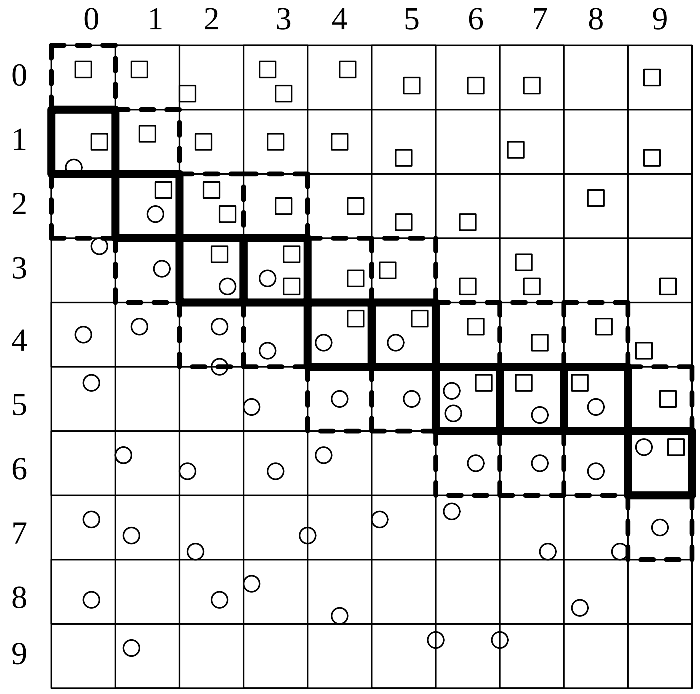
6.2 Εντοπισμός FOE και AOR με διανύσματα (r, s) copoint

Σύμφωνα με το προηγούμενο κεφάλαιο, μια πιθανή θέση της ευθείας και καμπύλης που χωρίζουν τις θετικές από τις αρνητικές μεταφορικές και περιστροφικές συνιστώσες των διανυσμάτων κάθετης ροής που ανήκουν επίσης σε κάποιο σύνολο διανυσμάτων copoint, θα είναι όπως αυτή του σχήματος (6.6). Τα μειονεκτήματα των διανυσμάτων copoint (όπως και των coaxis) είναι ότι:

- μπορούμε να χρησιμοποιήσουμε διανύσματα μηδενικού μόνο μήκους
- το πλήθος γενικά των μηδενικών διανυσμάτων είναι μικρό, με συνέπεια το μοντέλο του σχήματος (6.6) να μην είναι τόσο εμφανές.

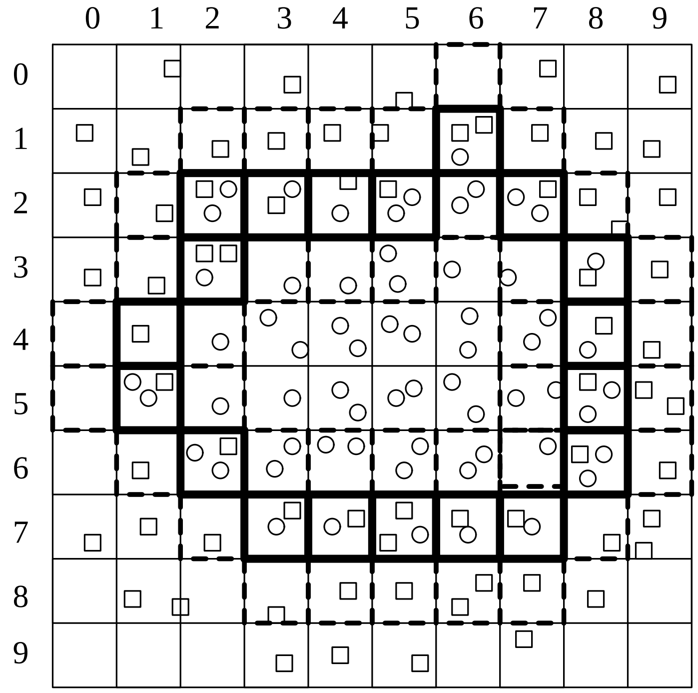
Αντιπαρερχόμαστε το πρώτο μειονέκτημα εφαρμόζοντας τον αλγόριθμο, που θα παρουσιαστεί παρακάτω, σε πολλά σημεία (r,s). Οσον αφορά στη λύση του δεύτερου, παρατηρούμε ότι δε χρειάζεται να περιοριστούμε στη μελέτη των τετραγώνων που περιέχουν μόνο μηδενικά διανύσματα, αλλά κάλλιστα μπορούμε να επεκταθούμε σε μια μελέτη των σαφώς περισσότερων τετραγώνων με αναμεμειγμένες θετικές, αρνητικές και μηδενικές τιμές. Τα σημεία FOE και AOR βρίσκονται σε τέτοιες περιοχές (ή ακριβώς δίπλα σε τέτοιες περιοχές) μια και, όπως φαίνεται και στο σχήμα 6.6, ανήκουν σε καμπύλες που διαχωρίζουν περιοχές της εικόνας με αμιγή σύνολα θετικών ή αρνητικών διανυσμάτων από περιοχές με αναμεμειγμένα σύνολα ετερόσημων διανυσμάτων. Στόχος του αλγορίθμου, λοιπόν, είναι να βρει τις περιοχές στην εικόνα που αποτελούν τα όρια μεταξύ "καθαρότητας" και "μίξης". Η γενική διατύπωσή του είναι η εξής:

```
# Εύρεση συνόλου κάθετης ροής N
# Καθορισμός συνόλου διακριτών διευθύνσεων D
# Καθορισμός συνόλου στοιχειωδών τετραγώνων B
# Για διάφορα (r, s)
    * Εύρεση συνόλου διανυσμάτων copoint CP_rs
    * ∀ b ϵ B
        βρες τα υποσύνολα θετικών CP_rsbp, αρνητικών CP_rsbn
        και μηδενικών CP_rsbz διανυσμάτων
    * Σχημάτισε το Mix = {b : b ϵ B και b όριο μεταξύ
        τετραγώνων με σταθερό πρόσημο και τετραγώνων με
        ποικιλία προσήμων}
    * Εάν PureTranslation(Mix) = TRUE
        ∀ b ϵ Mix VoteFOE[b] += CONST
    * Εάν PureRotation(Mix) = TRUE
        ∀ b ϵ Mix VoteAOR[b] += CONST
    * Εάν NoPure(Mix) = TRUE
        ∀ d ϵ D
            @ ε η ευθεία που διέρχεται από το (r, s) και έχει διεύθυνση d
            @ Vote = FitModel2(ε,CP_rsbz)
            @   ∀ b ϵ B και b "δίπλα" στην ευθεία ε
                    VoteFOE[b] += Vote
```



Σχήμα 6.7: Παράδειγμα λειτουργίας αλγόριθμου χρήσης διανυσμάτων copoint και coaxis (1)

A) Περίπτωση διανυσμάτων copoint. Εστω ότι οι κυκλίσκοι δείχνουν τις θέσεις, όπου βρέθηκαν αρνητικά διανύσματα κάθετης ροής που επιπλέον ανήκουν σε κάποιο σύνολο διανυσμάτων copoint (r, s). Τα τετραγωνάκια δείχνουν τις θέσεις των ανάλογων θετικών διανυσμάτων. Τα τετράγωνα που περιέχουν αμφότερες θετικές και αρνητικές τιμές (ορίζονται από έντονες μη διακεκομμένες γραμμές) βρίσκονται πάνω σε κάποιο σχηματισμό ευθείας, άρα η μεταφορική συνιστώσα της γενικής κίνησης κυριαρχεί έναντι της περιστροφικής. Μια και εξετάζουμε διανύσματα copoint, τα τετράγωνα αυτά, αλλά και τα γειτονικά τους (ορίζονται από έντονες διακεκομμένες γραμμές) για λόγους χαλαρότητας, θεωρούνται υποψήφια κατοχής του FOE και λαμβάνουν την ανάλογη θετική ψήφο.
B) Αν ανάλογο σχήμα εξαγόταν από την εξέταση διανυσμάτων coaxis, τα τετράγωνα κατοχής δύο προσήμων θα ψηφίζονταν ως υποψήφια κατοχής του AOR.



Σχήμα 6.8: Παράδειγμα λειτουργίας αλγορίθμου χρήσης διανυσμάτων copoint και coaxis (2)

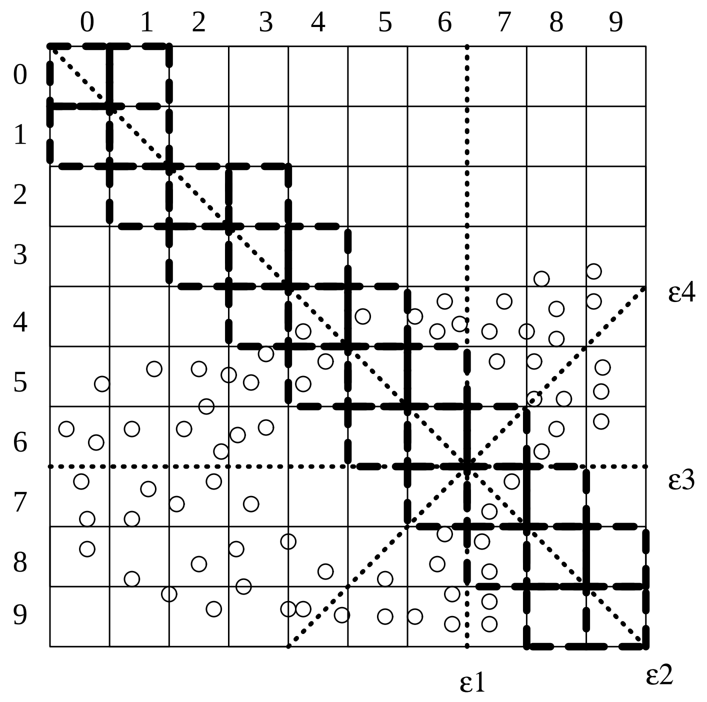
A) Περίπτωση διανυσμάτων copoint. Οι κυκλίσκοι δείχνουν τα αρνητικά διανύσματα κάθετης ροής που ανήκουν σε κάποιο σύνολο διανυσμάτων copoint. Τα μικρά τετράγωνα δείχνουν τις αντίστοιχες θέσεις των θετικών διανυσμάτων. Τα τετράγωνα που περιέχουν αμφότερες θετικές και αρνητικές τιμές (ορίζονται από έντονες μη διακεκομμένες γραμμές) σχηματίζουν μια κλειστή καμπύλη, άρα η περιστροφική συνιστώσα της γενικής κίνησης υπερισχύει της μεταφορικής. Τα τετράγωνα αυτά, καθώς και τα γειτονικά τους (ορίζονται από έντονες διακεκομμένες γραμμές) για λόγους χαλαρότητας, ψηφίζονται ως υποψήφια κατοχής του AOR. B) Αν ανάλογο σχήμα εξαγόταν από την εξέταση διανυσμάτων coaxis, τα τετράγωνα κατοχής δύο προσήμων θα ψηφίζονταν ως υποψήφια κατοχής του FOE.

Ο σχηματισμός του Mix πραγματοποιείται με μια απλή αναδρομική διαδικασία depth-first-search. Η διαδικασία ξεκινάει από κάποιο τετράγωνο που περιέχει διανύσματα μηδενικά ή διανύσματα δύο προσήμων. Στη συνέχεια, εξετάζεται η υπεροχή ή όχι κάποιας από τις δύο συνιστώσες κίνησης (περιστροφική, μεταφορική). Αν η κίνηση είναι μεταφορική (η γνώση αυτή προέρχεται από το γενικότερο σύστημα ελέγχου των κινήσεων του mobile), τότε σύμφωνα με το σχήμα 5.7 το σύνολο των τετραγώνων του Mix θα αποτελούν μία ευθεία και όλοι οι ψήφοι θα αφορούν στο FOE (σχ. 6.7). Στην αντίθετη περίπτωση, τα τετράγωνα του Mix θα σχηματίζουν (σχεδόν) ένα κύκλο και όλοι οι ψήφοι θα καταλογίζονται για το AOR (σχ. 6.8). Στη γενική περίπτωση κίνησης, τα τετράγωνα b του συνόλου Mix δε θα σχηματίζουν κάποιο κανονικό γεωμετρικό σχήμα. Σ'αυτήν την περίπτωση ακολουθείται μια ανάλογη με την περίπτωση σταθερής κάθετης ροής διαδικασία: εξετάζουμε ποια απ' όλες τις ευθείες που περνάνε από το (r, s) ταιριάζει καλύτερα με το μοντέλο του σχήματος 6.6. Κριτήριο ταιριάσματος αποτελεί ο χωρισμός στα δύο της κατανομής των μηδενικών διανυσμάτων, με τρόπο ώστε η κατανομή να "στενεύει" στην περιοχή κοντά στην ευθεία (σχ. 6.9) Η μέτρηση του "στενέματος" στους αλγορίθμους γίνεται με μεθόδους παρόμοιες με εκείνες του αλγορίθμου συγκεκριμένης κάθετης ροής. Τα τετράγωνα που βρίσκονται δίπλα στην επικρατέστερη ευθεία ψηφίζονται ως κατάλληλα κατοχής του FOE.

6.3 Εντοπισμός FOE και AOR με διανύσματα (A, B, C) coaxis

Ο αλγόριθμος υπολογισμού FOE και AOR με τη χρησιμοποίηση των διανυσμάτων α, β, γ (ειδικών περιπτώσεων των coaxis (A, B, C)) παρουσιάζει πολλές ομοιότητες με εκείνον που χρησιμοποιεί τα διανύσματα copoint. Μια και πάλι τα διανύσματα με μηδενικό μέτρο είναι λίγα, επεκτεινόμαστε σε όλα τα διανύσματα που περιέχουν αναμεμειγμένα πρόσημα. Στόχος μας είναι να βρούμε τις περιοχές όπου παρατηρείται αλλαγή από τετράγωνα αμιγών προσήμων σε τετράγωνα δύο προσήμων. Ο αλγόριθμος έχει ως εξής:

Εύρεση συνόλου κάθετης ροής N
Καθορισμός συνόλου στοιχειωδών τετραγώνων B
Καθορισμός συνόλου διακριτών διευθύνσεων D
$\forall\ w \in \{\alpha, \beta, \gamma\}$ **** τρεις περιπτώσεις διανυσμάτων coaxis ****
    * Εύρεση διανυσμάτων coaxis CA_w
    * $\forall\ b \in B$



Σχήμα 6.9: Παράδειγμα λειτουργίας αλγόριθμου χρήσης διανυσμάτων copoint και coaxis (3)

Με τους κυκλίσκους αναπαριστώνται τα διανύσματα κάθετης ροής μηδενικού μήκους που ανήκουν σε κάποιο σύνολο διανυσμάτων (r, s) copoint. Το (r, s) βρίσκεται στην τομή των ευθειών ε1, ε2, ε3, ε4. Η ευθεία που θα θεωρηθεί ότι ανταποκρίνεται καλύτερα στο μοντέλο του σχήματος 6.6 θα θεωρηθεί υποψήφια κάτοχος του FOE και σε όλα τα στοιχειώδη τετράγωνα που γειτονεύουν με αυτή θα δοθεί μία θετική ψήφος. Στο παράδειγμά μας η ευθεία ε2 αποδεικνύεται η καλύτερη πρόταση, αφού η κατανομή των μηδενικών διανυσμάτων "στενεύει" και στις δύο περιοχές που συναντάται με την ευθεία. Τα τετράγωνα που σχηματίζονται με διακεκομμένες γραμμές ψηφίζονται ως κάτοχοι του FOE.

βρες τα υποσύνολα θετικών CA_{wbp}, αρνητικών CA_{wbn}
και μηδενικών CA_{wbz} διανυσμάτων

* Σχημάτισε το $Mix = \{b\ :\ b \epsilon B$ και b όριο μεταξύ
τετραγώνων με σταθερό πρόσημο και τετραγώνων με
ποικιλία προσήμων$\}$

* Εάν PureTranslation(Mix_w) = TRUE
$\forall\ b\ \epsilon\ Mix_w$ VoteFOE[b] += CONST

* Εάν PureRotation(Mix_w) = TRUE
$\forall\ b\ \epsilon\ Mix_w$ VoteAOR[b] += CONST

* Εάν NoPure(Mix_w) = TRUE
Εάν $w = \alpha \Rightarrow d = 0$ **** α διανύσματα ****
Εάν $w = \beta \Rightarrow d = \frac{\pi}{2}$ **** β διανύσματα ****
$\forall$ ευθεία με διεύθυνση d
Vote = FitModel3(ε,CA_{wbz})
$\forall\ b\ \epsilon\ B$ και b "δίπλα"στην ε
VoteAOR[b] += Vote
Εάν $w = \gamma$ **** γ διανύσματα ****
$\forall\ d\ \epsilon\ D$
ε η ευθεία που διέρχεται από το $(0,0)$ και έχει διεύθυνση d
Vote = FitModel4(ε,CP_{wbz})
$\forall\ b\ \epsilon\ B$ και b "δίπλα" στην ευθεία ε
VoteAOR[b] += Vote

Η εύρεση της περιοχής Mix γίνεται, όπως και στην περίπτωση των διανυσμάτων copoint, με μια διαδικασία depth-first-search, η οποία ξεκινάει από κάποιο στοιχειώδες τετράγωνο στο οποίο περιέχονται και θετικά και αρνητικά (ή απλά και μόνο μηδενικά) διανύσματα. Αν η κίνηση είναι μεταφορική (και πάλι, όπως και στην περίπτωση των διανυσμάτων copoint, η γνώση αυτή προέρχεται από το γενικότερο σύστημα ελέγχου των κινήσεων του mobile), τότε τα τετράγωνα που αποτελούν το Mix θα σχηματίζουν μια καμπύλη (υπερβολή αν εξετάζουμε τα α ή β διανύσματα και κύκλο αν εξετάζουμε τα γ διανύσματα) και όλα τα τετράγωνα θα ψηφιστούν ως υποψήφια κατοχής του FOE. Αν η κίνηση είναι περιστροφική, τότε τα τετράγωνα του Mix, τα οποία θα σχηματίζουν (περίπου) μία ευθεία, θα ψηφιστούν ως υποψήφια κατοχής του AOR. Αν καμιά από τις δύο περιπτώσεις δεν ισχύει, τότε ο αλγόριθμος προσπαθεί να εντοπίσει την ευθεία διαχωρισμού των περιστροφικών συνιστωσών με τη γνωστή, από τους προηγούμενους αλγορίθμους, μέθοδο δοκιμής ευθειών διαφόρων τοποθετήσεων στην εικόνα. Από τις εξισώσεις των καμπυλών και ευθειών που παρουσιάστηκαν στο προηγούμενο κεφάλαιο και χαρακτηρίζουν τα α, β και γ διανύσματα, συνεπάγεται ότι για τα α και β διανύσματα, εξετάζονται μόνο οι παράλληλες προς το X και Y αντίστοιχα άξονα ευθείες (η FitModel3, δηλαδή του αλγορίθμου είναι ειδική περίπτωση της FitModel1 του αλγορίθμου σταθερής κάθετης ροής), ενώ για τα γ διανύσματα, εξετάζονται μόνο οι ευθείες διαφόρων διευθύνσεων που διέρχονται από το (0,0) (η FitModel4, δηλαδή, είναι ειδική περίπτωση της FitModel2 του αλγορίθμου διανυσμάτων copoint).

6.4 Αποτελέσματα

Η απόδοση των αλγορίθμων που περιγράφηκαν προηγουμένως δοκιμάστηκε σε μια σειρά από εικόνες που παρουσιάζουν το πεδίο κάθετης ροής που προκύπτει από τη μετακίνηση της σκηνής. Στις δοκιμές χρησιμοποιήθηκαν κυρίως συνθετικές, αλλά και οριομένα ζεύγη πραγματικών εικόνων.

6.4.1 Συνθετικά δεδομένα

Τα συνθετικά δεδομένα δημιουργούνται γνωρίζοντας εκ των προτέρων τις τιμές των διανυσμάτων (U, V, W) και (α, β, γ) που καθορίζουν το μέγεθος της περιστροφής και της μεταφοράς. Με βάση αυτή τη γνώση και τις εξιοώσεις (4.7) και (4.8), βρίσκεται αρχικά η οπτική ροή για κάθε σημείο (x, y) της εικόνας και στη συνέχεια υπολογίζεται η κάθετη ροή προβάλλοντας κάθε διάνυσμα σε διευθύνσεις που αντιστοιχούν σε υποθετικές ακμές. Για να προσομοιωθούν, όμως, όσο το δυνατόν περισσότερο οι ρεαλιστικές συνθήκες εύρεσης κάθετης ροής χρησιμοποιούνται τα εξής τεχνάσματα:

- Για κάθε σημείο (x, y) ορίζεται μια τυχαία τιμή βάθους Z, έτσι ώστε να μην αποκλείονται φαινόμενα όπου σημεία γειτονικά στην εικόνα, αλλά απομακρυσμένα στη σκηνή παρουσιάζουν εντελώς διαφορετική μετατόπιση ακριβώς εξαιτίας της διαφοράς του βάθους τους. Η τιμή αυτή επιλέγεται ανάμεσα σε εκείνες που ορίζονται από τα προκαθοριομένα όρια (Z_{min} και Z_{max}) της σκηνής.

- Επιλέγεται επίσης τυχαία σε κάθε σημείο η διεύθυνση προβολής της οπτικής ροής, ώστε η κάθετη ροή να αντιστοιχεί σε ακμές ποικίλων διευθύνσεων.

- Ορίζονται περιοχές στην εικόνα όπου το οπτικό πεδίο μηδενίζεται απόλυτα αντιστοιχώντας στη ρεαλιστική κατάσταση της απουσίας οπτικής-κάθετης ροής σε περιοχές χωρίς καθόλου ακμές, όπως π.χ. συμβαίνει στους τοίχους. Στις υπόλοιπες περιοχές όπου υπάρχουν ακμές - επομένως και διανύσματα κάθετης ροής, με μια συνάρτηση πυκνότητας πιθανότητας εξετάζεται κάθε σημείο (x, y) αν ανήκει όντως σε ακμή οπότε και δημιουργείται το ανάλογο διάνυσμα κάθετη ροής.

Ακολουθώντας τις παραπάνω αρχές τα πεδία κάθετης ροής που εξάγονται είναι μεν συνθετικά, προσομοιάζουν, όμως, σε αρκετό βαθμό τη συμπεριφορά των πεδίων κάθετης ροής πραγματικών εικόνων.

Τα πειράματα που εξετάστηκαν διέφεραν μεταξύ τους σε βασικές παραμέτρους του συστήματος, όπως

- το μέγεθος της μεταφοράς (τάξη μεγέθους μείον τέσσερα (π.χ. 0.0006) ως μείον ένα (π.χ.0.2))
- το μέγεθος της περιστροφής (τάξη μεγέθους μείον τέσσερα (π.χ. 0.0006) ως μείον ένα (π.χ. 0.2))
- η θέση των σημείων FOE και AOR στην εικόνα (κέντρο ή περιφέρεια)
- το ποσοστό της εικόνας όπου υπάρχουν περιοχές με πληροφορία κίνησης
- η τιμή της εστιακής απόστασης f (150mm-599mm)

Οι τιμές των υπόλοιπων παραμέτρων διατηρήθηκαν σταθερές και συγκεκριμένα, το ελάχιστο βάθος Z_{min} στα 150mm, το μέγιστο βάθος Z_{max} στα 5000mm, οι διαστάσεις των εικόνων στο 300×300, το κέντρο $(0, 0)$ του επιπέδου προβολής στο $(150, 150)$ συντεταγμένων εικόνας.

Για καθένα από τα παρακάτω πειράματα εξετάστηκε το ποσοστό επιτυχίας και η συνεισφορά κάθε αλγορίθμου στην τελική λύση. Η ανασκόπηση των αποτελεσμάτων έδειξε ότι στις περιπτώσεις που η περιστροφική κίνηση είναι σε απόλυτα μεγέθη αρκετά μεγαλύτερη από τη μεταφορική, ο εντοπισμός του AOR βασίζεται στη χρησιμοποίηση τόσο των copoint όσο και των coaxis διανυσμάτων. Στις περιπτώσεις, όμως, που η μεταφορική κίνηση γίνεται ίση ή και μεγαλύτερη από την περιστροφική, η γενική κατεύθυνση για την εύρεση του AOR προέρχεται από τα διανύσματα coaxis. Οσον αφορά στο FOE, όπως αναμενόταν οι βασικές κατευθύνσεις δίνονται από τις μεθόδους εύρεσης του στενέματος της κατανομής των διανυσμάτων κάθετης ροής και αναγνώρισης των προτύπων των διανυσμάτων copoint (συνήθως χρησιμοποιούνται τέσσερα ή οκτώ σημεία αναφοράς (r,s)). Αντίθετα, ο αλγόριθμος αποκοπής περιοχών όπου διαπιστώνονται αλληλοκαλύψεις σε κατανομές διανυσμάτων διαφορετικού μήκους και ίδιας διεύθυνσης αποδεικνύεται και πρακτικά ότι χρησιμεύει για να αποκλείει περιοχές στα άκρα της εικόνας, παρά στο να καθορίζει με ακρίβεια θέσεις όπου το FOE μπορεί να βρίσκεται.

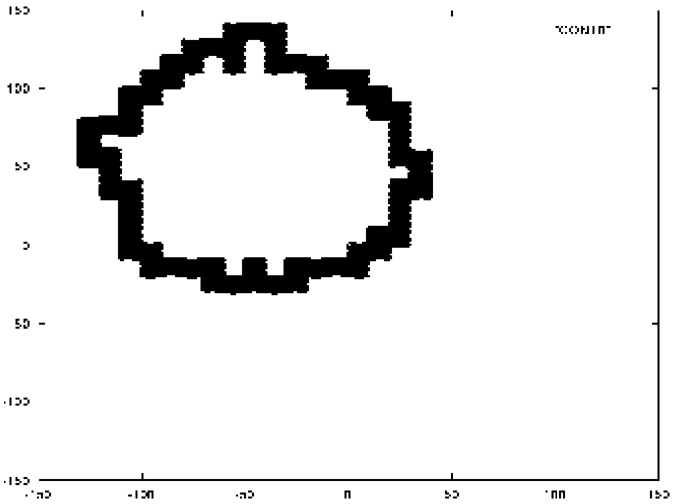
Στις εικόνες (6.10-6.18) φαίνεται ένα παράδειγμα αποτελέσματος αναζήτησης του AOR. Η κίνηση της κάμερας ήταν $(U,V,W) = (0.03, 0.03, 0.06)$ και $(\alpha,\beta,\gamma) = (0.005, 0.0055, 0.008)$, που συνεπάγεται ότι το AOR βρίσκεται στο σημείο (93,103). Στις εικόνες (6.10),(6.11),(6.12) σχηματίζονται τα ενδιάμεσα αποτελέσματα μελέτης των προτύπων των διανυσμάτων copoint με σημεία αναφοράς τα (0,0), (-97,107) και (128,109), και στο (6.13) η τελική πρόταση αυτής της μεθόδου για τις περιοχές που περικλείουν το AOR. Οι εικόνες (6.14), (6.15), (6.16), (6.17) παρουσιάζουν τα α,β,γ διανύσματα για την ίδια είσοδο και την συνεπαγόμενη πρόταση για το AOR. Στην εικόνα (6.18) παρουσιάζεται το αποτέλεσμα από το συνδυασμό των δύο προτάσεων.

Στην περίπτωση που όλες οι περιοχές της εικόνας προσφέρουν πληροφορίες για το πεδίο κάθετης ροής η επιτυχία της μεθόδου ήταν απόλυτη. Οι υπολογιζόμενες περιοχές για τα FOE και AOR είχαν εμβαδό της τάξης του 1%. Η σταδιακή μείωση του ποσοστού, 80%, 60%, 40%, της εικόνας που παρέχει διανύσματα κάθετης ροής επέφερε μια πτώση της αποδοτικότητας της μεθόδου καθώς οι μέσες τιμές των εμβαδών των περιοχών για τα FOE και AOR ανέβηκαν στο 3%, 8% και 12% αντίστοιχα. Και πάλι, όμως, η σύμπτωση των FOE και AOR με τις προτεινόμενες περιοχές ήταν ακριβής. Δείγμα, επίσης, της σταθερότητας του αλγορίθμου αποτελεί η παρατήρηση ότι τα αποτελέσματα δεν επηρεάζονται από τη σχέση των διάφορων τάξεων μεγεθών μεταφοράς και περιστροφής που δοκιμάστηκαν.

Οι χρόνοι εκτέλεσης των προγραμμάτων σε όλες τις περιπτώσεις ήταν αρκετά χαμηλοί, αποτέλεσμα της αναζήτησης απλών και όχι σύνθετων μέτρων αναγνώρισης των προτύπων. Στους πρώτους δύο αλγορίθμους οι χρόνοι εξαρτώνται αποκλειστικά από το σύνολο των διαφορετικών μέτρων και διευθύνσεων που χαρακτηρίζουν τα διανύσματα. Στον τρίτο αλγόριθμο οι επαναλήψεις καθορίζονται από τον αριθμό των σημείων αναφοράς για τα οποία βρίσκουμε τα διανύσματα copoint, ενώ στον τελευταίο ο χρόνος εκτέλεσης είναι σταθερός μια και κάθε φορά εξετάζουμε τις τρεις απλές περιπτώσεις των διανυσμάτων coaxis, τα α, βκαιγ.

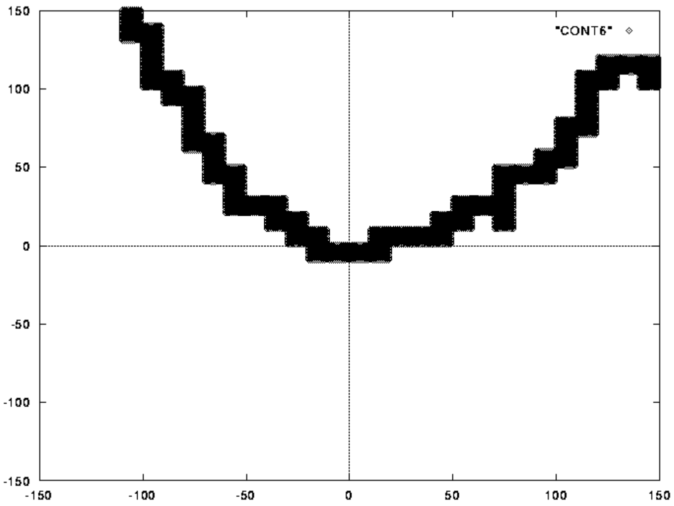
6.4.2 Πραγματικές εικόνες

Η εφαρμογή των αλγορίθμων σε πραγματικές εικόνες απέφερε αποτελέσματα χαμηλότερης ακρίβειας απ' ό,τι στην περίπτωση των συνθετικών εικόνων. Οι λόγοι που συντελούν σ' αυτή την πτώση της αποδοτικότητας είναι η μείωση του πλήθους των διαθέσιμων διανυσμάτων κάθετης ροής (στις εικόνες που χρησιμοποιήθηκαν το ποσοστό των pixel με κάθετη ροή δεν ξεπερνούσε το 10%), επομένως και της ωφέλιμης πληροφορίας και η παρείσφρηση μεταξύ των διανυσμάτων του πεδίου κάθετης ροής διανυσμάτων που δεν οφείλονται στην κίνηση της κάμερας αλλά στην απρόσμενη αλλαγή της φωτεινότητας κάποιας περιοχής. Τα διανύσματα αυτά φυσικά δεν ακολουθούν τα μοντέλα κατανομών που περιγράφηκαν στο κεφάλαιο 5, οπότε οδηγούν τους αλγορίθμους σε λαθεμένες εκτιμήσεις. Η κατάσταση αυτή συνεπάγεται μια χαλαρότητα στην εφαρμογή των κριτηρίων ανίχνευσης των προτύπων που αναζητούνται, η οποία αποφέρει αυξήσεις στο εμβαδό των περιοχών των FOE και AOR. Στα διάφορα πειράματα που έγιναν το εμβαδό αυτό κυμαινόταν γύρω στο 15%. Στα σχήματα (6.19)(6.20)(6.21) παρουσιάζονται τα αποτελέσματα του αλγορίθμου για το πεδίο κάθετης ροής του σχήματος (4.4). Στο σχήμα (6.19) παρουσιάζεται η πρόταση του αλγορίθμου εύρεσης FOE μέσω διανυσμάτων σταθερής ροής. Στο σχήμα (6.20) παρουσιάζεται η πρόταση του αλγορίθμου που χρησιμοποιεί διανύσματα copoint (τα $(0,0), (-26,34), (38,30), (-59,-129)$ στην περίπτωσή μας). Στο σχήμα (6.21) παρουσιάζεται η τελική πρόταση που προκύπτει από το συνδυασμό των δύο προτάσεων. Τα τετράγωνα που ψηφίστηκαν ως πιθανοί κάτοχοι του FOE συγκέντρωσαν πάνω από το 70% της μέγιστης ψήφου. Το FOE στην πραγματικότητα βρίσκεται στο σημείο $(6,24)$.



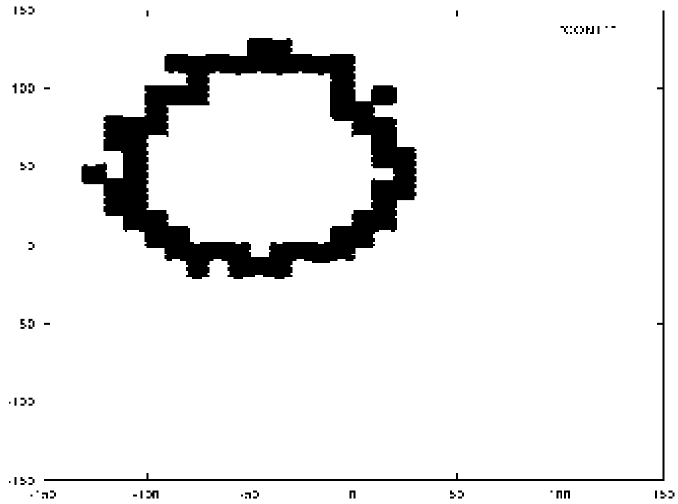
Σχήμα 6.10: Πρότυπο διανυσμάτων (0,0) copoint

Η διαδικασία εύρεσης και εντοπισμού των προτύπων για τα διανύσματα (0,0) copoint κατέληξε στην ψήφιση των περιοχών που διακρίνονται στο σχήμα ως περιοχών υποψήφιων για το AOR.



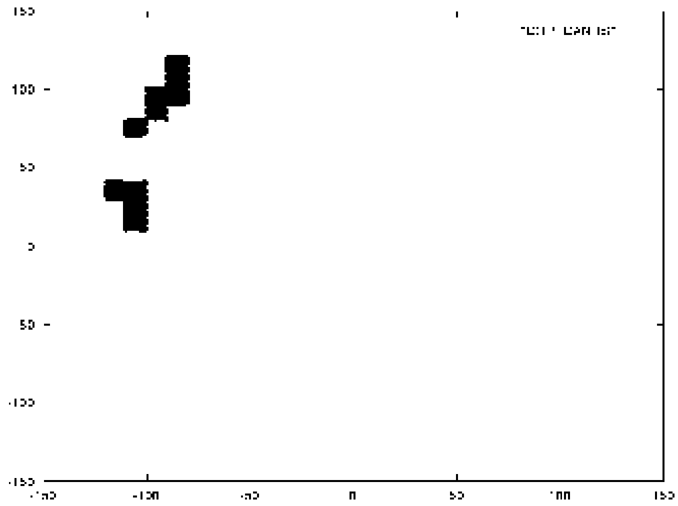
Σχήμα 6.11: Πρότυπο διανυσμάτων (128,109) copoint

Η διαδικασία εύρεσης και εντοπισμού των προτύπων για τα διανύσματα (128,109) copoint κατέληξε στην ψήφιση των περιοχών που διακρίνονται στο σχήμα ως περιοχών υποψήφιων για το AOR.



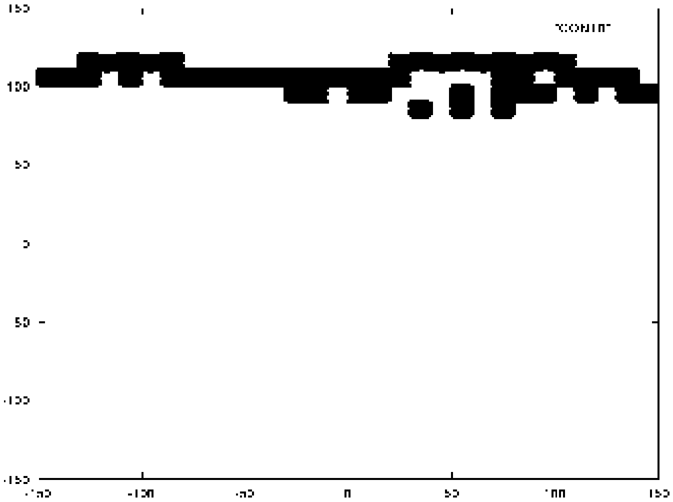
Σχήμα 6.12: Πρότυπο διανυσμάτων (-97,107) copoint

Η διαδικασία εύρεσης και εντοπισμού των προτύπων για τα διανύσματα (-97,107) copoint κατέληξε στην ψήφιση των περιοχών που διακρίνονται στο σχήμα ως περιοχών υποψήφιων για το AOR.



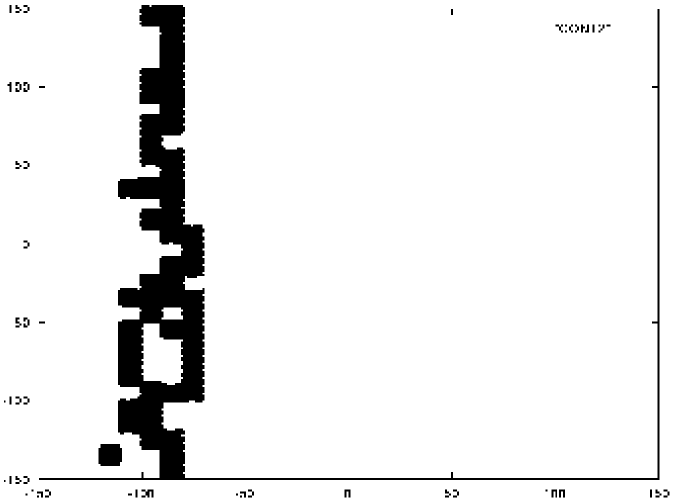
Σχήμα 6.13: Μια πρώτη πρόταση για το AOR

Τα μαυρισμένα τετραγωνάκια συγκέντρωσαν πάνω από το 80% του μέγιστου των θετικών ψήφων ύστερα από το συνδυασμό των παραπάνω τριών προτάσεων των διανυσμάτων copoint.



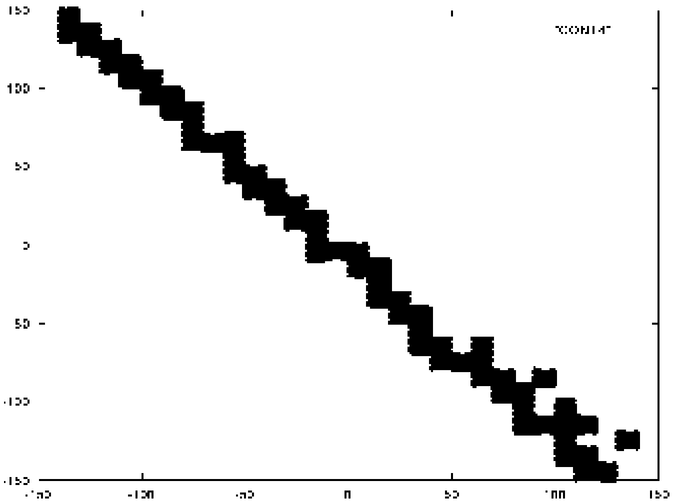
Σχήμα 6.14: Πρότυπο διανυσμάτων α

Το μαυρισμένο περίγραμμα βρέθηκε ότι χωρίζει θετικά από αρνητικά διανύσματα α.



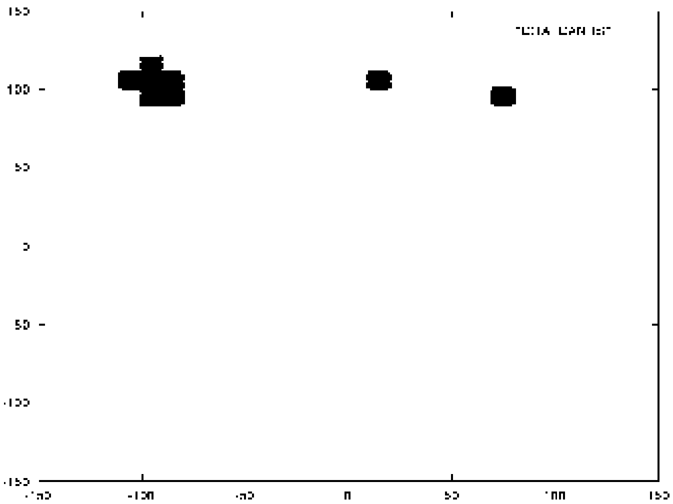
Σχήμα 6.15: Πρότυπο διανυσμάτων β

Το μαυρισμένο περίγραμμα βρέθηκε ότι χωρίζει θετικά από αρνητικά διανύσματα β.



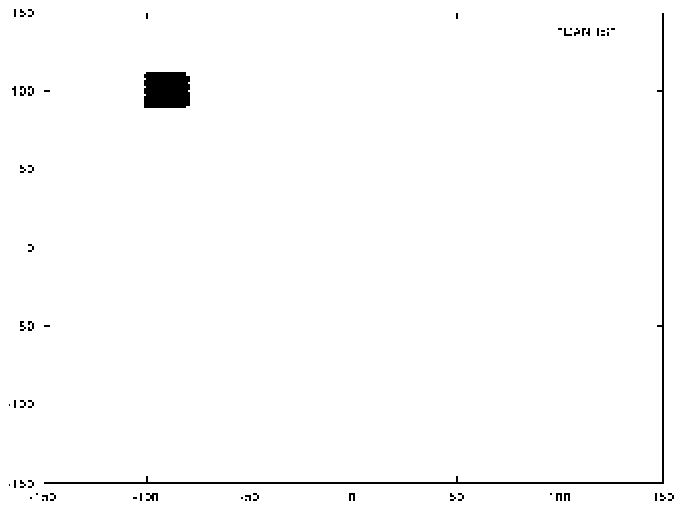
Σχήμα 6.16: Πρότυπο διανυσμάτων γ

Το μαυρισμένο περίγραμμα βρέθηκε ότι χωρίζει θετικά από αρνητικά διανύσματα γ.



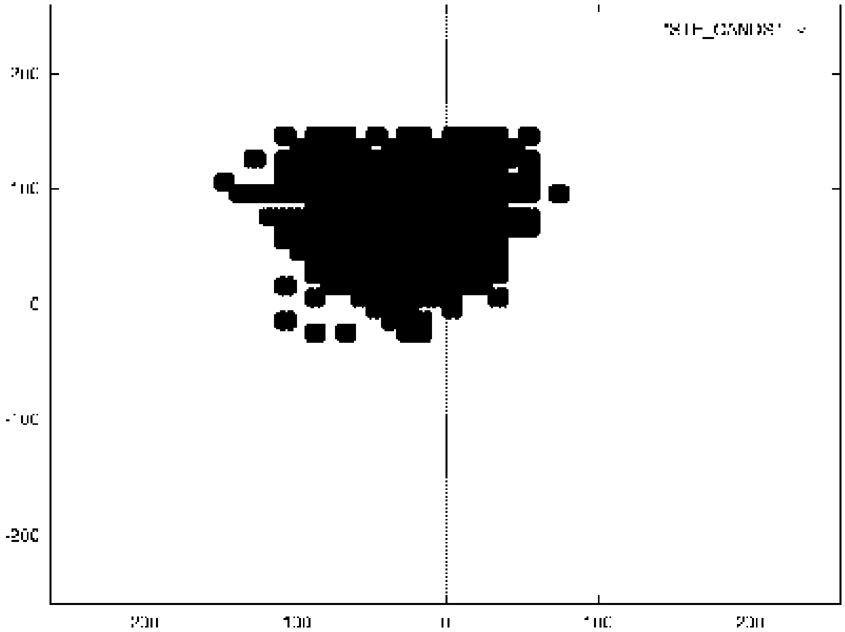
Σχήμα 6.17: Δεύτερη πρόταση για το AOR

Οι μαυρισμένες περιοχές συγκέντρωσαν τους περισσότερους ψήφους για το AOR ύστερα από το συνδυασμό των παραπάνω τριών προτάσεων των διανυσμάτων α, β και γ.



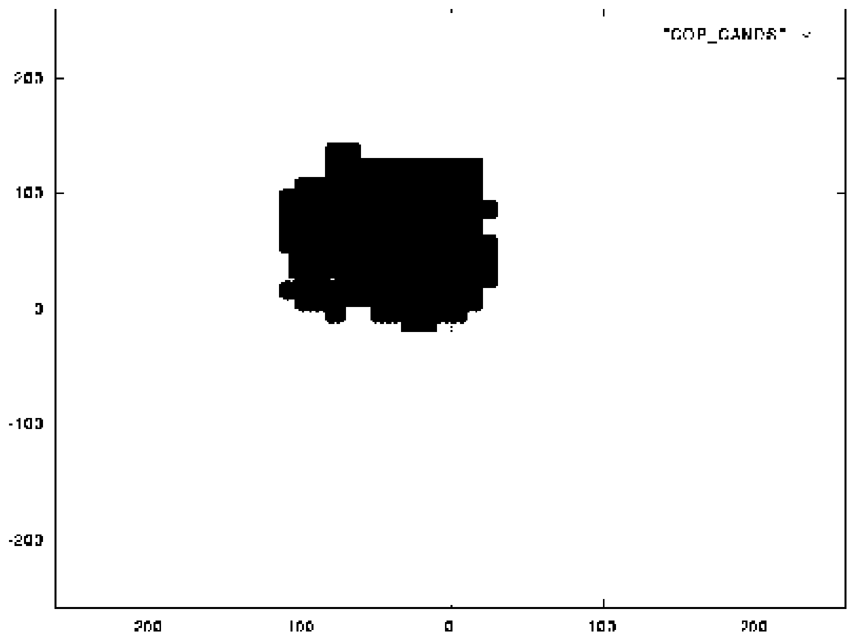
Σχήμα 6.18: Τελική τοποθέτηση του AOR

Συνδυάζοντας τις προτάσεις που φαίνονται στα σχήματα 6.13 και 6.17, τις περισσότερες ψήφους απέσπασαν τα τέσσερα στοιχειώδη τετράγωνα που εικονίζονται στο επάνω αριστερό κομμάτι της εικόνας. Οι πραγματικές συντεταγμένες του AOR είναι (93,103). Οι θετικές τιμές του X άξονα στην πραγματικότητα είναι αντίθετα από ό,τι δείχνει το σχήμα, μια και στο δεξιόστροφο σύστημα συντεταγμένων ο θετικός ημιάξονας OX δείχνει προς τα αριστερά. Η σύμπτωση AOR στην προτεινόμενη περιοχή είναι ακριβής. Το ποσοστό της περιοχής είναι 4 μέρη στα 900, δηλαδή 0.44%



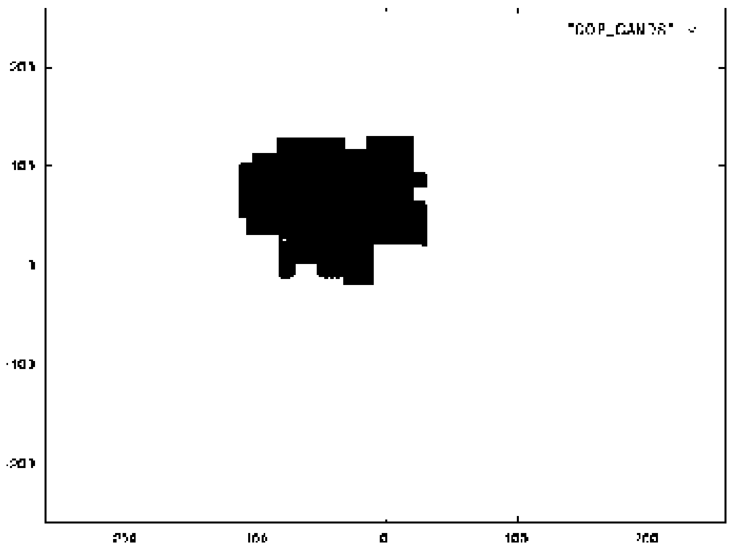
Σχήμα 6.19: Πρώτη πρόταση για το FOE

Στο σχήμα παρουσιάζονται τα τετράγωνα που συνέλεξαν τουλάχιστο το 70% της μέγιστης ψήφου, σύμφωνα με τον αλγόριθμο σταθερής κάθετης ροής



Σχήμα 6.20: Δεύτερη πρόταση για το FOE

Στο σχήμα παρουσιάζεται το αποτέλεσμα της ψηφοφορίας για την κατοχή του FOE με τη χρήση τεσσάρων συνόλων διανυσμάτων copoint



Σχήμα 6.21: Τελική πρόταση για το FOE

Ο συνδυασμός της πρότασης (1) και της πρότασης (2) έχει ως αποτέλεσμα την πρόκριση των μαυριομένων τετραγώνων ως πιθανά κατοχής του FOE. Το FOE στην πραγματικότητα βρίσκεται στο $(6, 24)$

Κεφάλαιο 7

Επίλογος

7.1 Συμπεράσματα

Στα πλαίσια αυτής της εργασίας μελετήθηκε το πρόβλημα του προσδιορισμού της ιδίας κίνησης ενός παρατηρητή ο οποίος διαθέτει μία κάμερα λήψης ψηφιακών εικόνων και κινείται σε ένα στατικό περιβάλλον. Η ικανότητα αυτοπροσδιορισμού των κινήσεων ενός φυσικού σώματος είναι θεμελιώδης όχι μόνο για τα έμβια όντα, αλλά και για τα αυτόνομα μηχανικά συστήματα που προσομοιώνουν "έννοες" συμπεριφορές. Κατέχοντας αυτήν την ικανότητα υποβοηθείται η κατόρθωση δυσκολότερων έργων, όπως η αποφυγή εμποδίων και η παρακολούθηση κινούμενων αντικειμένων.

Στην παρούσα εργασία δόθηκε έμφαση στην εκμετάλλευση του πεδίου κάθετης ροής και όχι σε κείνο της οπτικής, διότι ο υπολογισμός του πρώτου απαιτεί λιγότερες πράξεις και παραδοχές για το σκηνικό. Οπως έχει αποδειχθεί, στο πεδίο κάθετης ροής, αν και εξ ορισμού περιέχονται λιγότερες πληροφορίες από το πεδίο οπτικής ροής, είναι δυνατόν να ανιχνευτούν στοιχεία που να παρέχουν μέτρα υπόθεσης και προσδιορισμού των ακριβών παραμέτρων κίνησης. Τα στοιχεία αυτά παραμένουν μεταμφιεσμένα μέσα σε εξισώσεις απλών καμπυλών, ευθειών, υπερβολών και κύκλων, οι οποίες εκφράζουν τις ποσότητες της περιστροφικής και της μεταφορικής συνιστώσας της κίνησης. Η ανακάλυψη των στοιχείων αυτών αποδεικνύεται ότι δεν απαιτεί ακριβείς μετρήσεις μεγεθών, αλλά αντίθετα μπορεί να στηριχτεί σε ποιοτικές αναλύσεις των εικόνων, όπως η εξερεύνησή τους για τον εντοπισμό και "χαρτογράφηση" προτύπων που προκαλεί η ιδιόμορφη κατανομή των θετικών και αρνητικών διανυσμάτων του πεδίου κάθετης ροής.

Η μέθοδος, λοιπόν, που χρησιμοποιείται για τον προσδιορισμό της ιδίας κίνησης βασίζεται στην εύρεση αυτών των γενικών προτύπων, τα οποία παρέχουν πληροφορίες για τη θέση στην εικόνα δύο ενδιαφέροντων σημείων του επιπέδου, του FOE και του AOR, των σημείων δηλαδή στα οποία τα διανύσματα μεταφοράς και περιστροφής συναντούν το επίπεδο προβολής. Τελικός στόχος είναι η οριοθέτηση δύο όσο το δυνατόν μικρότερων περιοχών, για τις οποίες με ασφάλεια θα αποφαινόμαστε ότι περιέχουν τα FOE και AOR.

Κάθε ένας από τους τέσσερις αλγόριθμους που υλοποιήθηκαν ακολουθεί τη δική του συλλογιστική βασιζόμενος στο διαφορετικό πρότυπο που επιχειρεί να ανακαλύψει. Μια και στις εφαρμογές της ρομποτικής δίνεται μεγάλη έμφαση στα θέματα χρόνου, κύριο μέλημα κάθε αλγορίθμου είναι η θέσπιση απλών και εύκολα υπολογίσιμων κριτηρίων για την ανίχνευση των προτύπων, έτσι ώστε ο χρόνος εκτέλεσης του προγράμματος να διατηρείται όσο το δυνατόν μικρότερος. Στο τέλος κάθε εκτέλεσης προτείνονται ομάδες σημείων που διακρίνονται για τη μεγάλη πιθανότητα να συγκαταλέγονται ανάμεσά τους τα FOE και AOR. Μια συμπερασματική διαδικασία αξιολογεί και συγχωνεύει τα αποτελέσματα των τεσσάρων αλγορίθμων δημιουργώντας την τελική εκτίμηση για τις πιθανές θέσεις των AOR και FOE.

Τα αποτελέσματα που παρήγαγαν οι αλγόριθμοι κατά τις δοκιμές μπορούν να χωριστούν με κριτήριο την αληθοφάνειά τους σε τρεις κατηγορίες: άριστα, ικανοποιητικά και όχι αποτρεπτικά. Χρησιμοποιώντας συνθετικές εικόνες με αρκετή πυκνότητα πεδίου κάθετης ροής ο συνολικός αλγόριθμος επέτυχε υψηλότατες αποδόσεις εντοπίζοντας πάντα με ακρίβεια και σε στενά γεωμετρικά πλαίσια τα σημεία FOE και AOR. Οταν η ένταση του πεδίου κάθετης ροής μειώνεται αισθητά στην εικόνα, τότε τα πλαίσια αυτά διευρύνονται ανάλογα, παρέχοντας όμως πάντα αξιόπιστα αποτελέσματα. Στην περίπτωση, τέλος, πραγματικών εικόνων η απόδοση του αλγορίθμου μειώνεται αισθητά -όχι όμως απογοητευτικά-, μείωση που οφείλεται κυρίως στις συνθήκες δημιουργίας κάθετης ροής από πραγματικές εικόνες.

Το μέγεθος των περιοχών θα μπορούσε να ελαττωθεί αν χρησιμοποιούνταν περισσότερο sophisticated τεχνικές εύρεσης της κάθετης ροής (ουσιαστικά, εύρεση ακμών σε διαδοχικές εικόνες), ώστε οι αλγόριθμοι να δέχονται ως είσοδο περισσότερα και πιο αξιόπιστα διανύσματα κάθετης ροής. Τέτοιες τεχνικές απαιτούν μεγαλύτερους χρόνους εκτέλεσης, όμως η αύξηση αυτή ισοσκελίζεται εν μέρει από τη μείωση του χρόνου του ακριβούς προσδιορισμού των παραμέτρων κίνησης από τα fuzzy-FOE και fuzzy-AOR, μια και οι αλγόριθμοι με ποιοτικότερη είσοδο παράγουν μικρότερες περιοχές. Από την άλλη, εξαντλητικότερες διαδικασίες αναζήτησης των διάφορων προτύπων (π.χ. με τη χρησιμοποίηση μεθόδων ελάχιστου τετραγωνικού σφάλματος (least square error) για την επιτυχημένη παρεμβολή των συναρτήσεων των καμπυλών που δημιουργούν τα πρότυπα) πιθανότατα θα απέφεραν μικρότερες περιοχές για τα fuzzy-FOE και fuzzy-AOR, το τίμημα, όμως, θα ήταν η επιβάρυνση των χρόνων εκτέλεσης των αλγορίθμων. Γι' αυτό στην παρούσα εργασία προτιμήθηκε να διατηρηθεί η απλότητα των κριτηρίων ανακάλυψης των προτύπων, αφού οι αλγόριθμοι όντως οριοθετούν σωστά τις περιοχές των FOE και AOR.

7.2 Επεκτάσεις

Δεδομένης της μείωσης της ικανότητας του αλγορίθμου σε πραγματικές εικόνες βασική και πρώτη έγνοια επέκτασής του αποτελεί η προσπάθεια σταθεροποίησής του σε υψηλά επίπεδα απόδοσης. Η προσπάθεια αυτή θα πρέπει να αποτελείται από δύο υποενότητες, η μία θα αφορά στην ποιοτικότερη εξαγωγή του πεδίου κάθετης ροής και η άλλη στην εξαντλητικότερη διερεύνηση και ανάλυση των προτύπων χαρακτηρισμού της περιστροφικής και μεταφορικής συνιστώσας της κίνησης. Οι εξαντλητικότερες και αναλυτικότερες αυτές αναζητήσεις συνεπάγονται επιβάρυνση των προγραμμάτων με επιπλέον χρόνο, ο οποίος για εφαρμογές ρομποτικής είναι ιδιαίτερα σημαντικός και περιορισμένος. Μια λύση στο πρόβλημα αυτό θα αποτελούσε η χρησιμοποίηση αρχιτεκτονικών παράλληλων που θα επιτρέπουν την εκτέλεση έξυπνων και γρήγορων αλγορίθμων για την αποκομιδή ολοένα και καλύτερων αποτελεσμάτων. Η έννοια μάλιστα της παράλληλης εκτέλεσης τμημάτων ενός προγράμματος είναι ιδιαίτερα βολική για τους κατασκευασθέντες αλγορίθμους καθώς αυτοί εμφανίζουν διπλή παραλληλία: εσωτερική, καθώς σε κάθε αλγόριθμο αναζητούνται επαναληπτικά τα πρότυπα που δημιουργούνται από διανύσματα κάθετης ροής που έχουν χωριστεί σε διάφορες ανεξάρτητες κατηγορίες, αλλά και εξωτερική, καθώς στο τέλος της όλης διαδικασίας οι προτάσεις κάθε αλγορίθμου συλλέγονται και συγκρίνονται για την εξαγωγή των τελικών συμπερασμάτων.

Παράρτημα A

Υπολογισμός κάθετης ροής

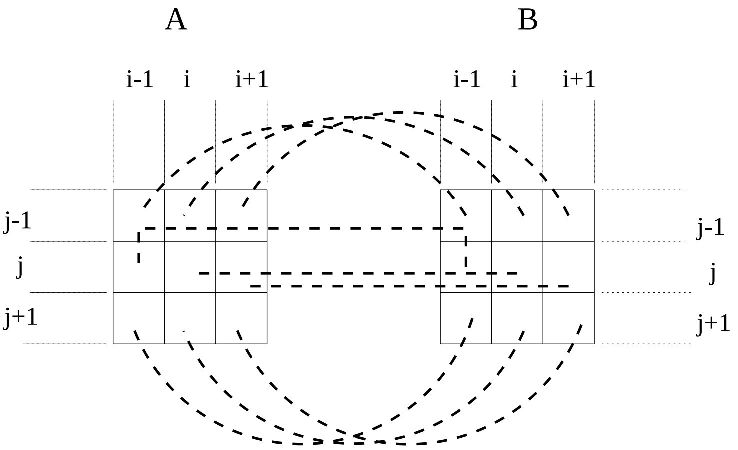
Εστω ότι η κάμερα έχει καταγράψει σε ένα σύντομο χρονικό διάστημα δύο εικόνες του πραγματικού κόσμου, οι οποίες έχουν αποθηκευθεί στους πίνακες A και B. Για να βρεθεί η κάθετη ροή σε κάποιο σημείο (i,j) της εικόνας χρειάζεται να παραγωγίσουμε ως προς χρόνο και χώρο. Η παραγώγιση ως προς χρόνο γίνεται βρίσκοντας τη μέση τιμή της διαφοράς της έντασης των σημείων της 3×3 γειτονιάς του (i,j) (σχ. A.1)

$$\begin{aligned} I_t \;=\; & (B(i+1,j+1) - A(i+1,j+1)+ \\ & B(i+1,j) - A(i+1,j)+ \\ & B(i+1,j-1) - A(i+1,j-1)+ \\ & B(i,j+1) - A(i,j+1)+ \\ & B(i,j) - A(i,j)+ \\ & B(i,j-1) - A(i,j-1)+ \\ & B(i-1,j+1) - A(i-1,j+1)+ \\ & B(i-1,j) - A(i-1,j)+ \\ & B(i-1,j-1) - A(i-1,j-1))\tfrac{1}{9} \end{aligned} .$$

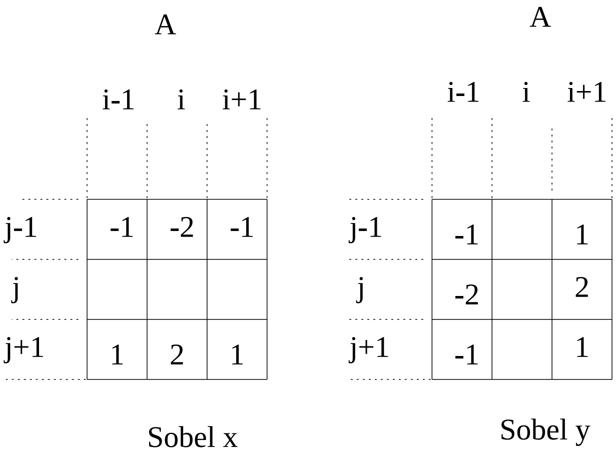
Η παραγώγιση ως προς τις δύο διαστάσεις του χώρου πραγματοποιείται με την εφαρμογή των φίλτρων Sobel στη 3×3 γειτονιά του (i,j) (σχ. A.2)

$$\begin{aligned} I_y \;=\; & (B(i+1,j-1) - B(i-1,j-1)+ \\ & + 2 * (B(i+1,j) - B(i-1,j))+ \\ & + B(i+1,j+1) - B(i-1,j+1)+ \\ & + A(i+1,j-1) - B(i-1,j-1)+ \\ & + 2 * (A(i+1,j) - A(i-1,j))+ \\ & + A(i+1,j+1) - A(i-1,j+1))\tfrac{1}{16} \end{aligned} .$$

$$
\begin{aligned}
I_x \;=\; & (B(i-1,j+1) - B(i-1,j-1) + \\
& + 2 * (B(i,j+1) - B(i,j-1)) + \\
& + B(i+1,j+1) - B(i+1,j-1) + \\
& + A(i-1,j+1) - A(i-1,j-1) + \\
& + 2 * (A(i,j+1) - A(i,j-1)) + \\
& + A(i+1,j+1) - A(i+1,j-1)) \tfrac{1}{16}
\end{aligned}
$$



Σχήμα Α.1: Παραγώγιση ως προς χρόνο



Σχήμα Α.2: Φίλτρα Sobel

Παράρτημα B

Προβολή οπτικής ροής σε διανύσματα copoint και coaxis

B.1 Προβολή σε διανύσματα coaxis

Τα διανύσματα coaxis ως προς μια τριάδα πραγματικών αριθμών (A, B, C) για την οποία ισχύει $A^2 + B^2 + C^2 = 1$, ορίζονται ως εκείνα τα διδιάστατα διανύσματα (x, y) που διεύθυνσή τους είναι η:

$$(m_x, m_y) = \left\{ \left[-A\left(y^2 + f^2\right) + Bxy + Cxf\right], \left[Axy - B\left(x^2 + f^2\right) + Cyf\right] \right\}$$

Η προβολή των διανυσμάτων οπτικής ροής στα διανύσματα (A, B, C) coaxis ισούται με

$$(u_x, \upsilon_y) = ((u, \upsilon) \cdot (M_x, M_y))(M_x, M_y)$$

όπου $M = (M_x, M_y) = \frac{1}{||\vec{M}||}(m_x, m_y)$ το μοναδιαίο coaxis (A, B, C) στο σημείο (x, y) και (u, υ) η οπτική ροή στο σημείο (x, y) όπως δίνεται από τους τύπους

$$\begin{aligned}
(u, \upsilon) &= (u_{trans} + u_{rot}, \upsilon_{trans} + \upsilon_{rot}) \\
&= \left\{ \left(\frac{-Uf + xW}{Z}\right) + \left[\alpha\frac{xy}{f} - \beta\left(\frac{x^2}{f} + f\right) + \gamma y\right], \left(\frac{-Vf + yW}{Z}\right) + \left[\alpha\left(\frac{y^2}{f} + f\right) - \beta\frac{xy}{f} - \gamma x\right] \right\} \\
&= \frac{W}{Z}(x - x_0, y - y_0) + \left\{ \left[\alpha\frac{xy}{f} - \beta\left(\frac{x^2}{f} + f\right) + \gamma y\right], \left[\alpha\left(\frac{y^2}{f} + f\right) - \beta\frac{xy}{f} - \gamma x\right] \right\}
\end{aligned}$$

όπου (U, V, W) και (α, β, γ) τα διανύσματα που εκφράζουν αντίστοιχα τη μεταφορική και περιστροφική συνιστώσα της κίνησης και $(x_0, y_0) = (\frac{Uf}{W}, \frac{Vf}{W})$ το σημείο FOE στο οποίο το διάνυσμα μεταφοράς συναντάει το επίπεδο προβολής. Το μέτρο της προβολής της οπτικής ροής στο (M_x, M_y) ισούται με το άθροισμα των μέτρων των προβολών της μεταφορικής t_n

και της περιστροφικής r_n συνιστώσας της οπτικής ροής στο (M_x, M_y). Μας ενδιαφέρει να μάθουμε πότε το άθροισμα αυτό γίνεται 0.

Η προβολή της μεταφορικής συνιστώσας στο (M_x, M_y) ισούται με:

$$\begin{aligned}
t_n &= \tfrac{W}{Z}(x-x_0, y-y_0)\cdot(M_x,M_y) = \tfrac{W}{Z}(x-x_0,y-y_0)\cdot\left[\tfrac{1}{||\vec{M}||}(m_x,m_y)\right] \\
&= \tfrac{W}{Z}(x-x_0,y-y_0)\cdot\left\{\tfrac{1}{||\vec{M}||}\left[(-A(y^2+f^2)+Bxy+Cxf),(Axy-B(x^2+f^2)+Cyf)\right]\right\} \\
&= \tfrac{W}{Z}(-Axy^2-Axf^2+Ax_0y^2+Af^2x_0+Bx^2y+Cx^2f-Bxx_0y-Cxx_0f)\ \tfrac{1}{||\vec{M}||} \\
&\quad + \tfrac{W}{Z}(Axy^2-Axyy_0-Bx^2y-Byf^2+Bx^2y_0+By_0f^2+Cy^2f-Cyy_0f)\ \tfrac{1}{||\vec{M}||} \\
&= \tfrac{W}{Z}\left[x^2(Cf+By_0)+y^2(Cf+Ax_0)-xy(Ay_0+Bx_0)\right]\ \tfrac{1}{||\vec{M}||} \\
&\quad + \tfrac{W}{Z}\left[-xf(Af+Cx_0)-yf(Bf+Cy_0)+f^2(Ax_0+By_0)\right]\ \tfrac{1}{||\vec{M}||} \\
&= \tfrac{W}{Z}\ k(A,B,C,x_0,y_0;x,y)\ \tfrac{1}{||\vec{M}||}
\end{aligned}$$

όπου

$$\begin{aligned}
k(A,B,C,x_0,y_0;x,y) &= x^2(Cf+By_0)+y^2(Cf+Ax_0)-xy(Ay_0+Bx_0) \\
&\quad -xf(Af+Cx_0)-yf(Bf+Cy_0)+f^2(Ax_0+By_0)
\end{aligned}$$

Η εξίσωση $k = 0$ περιγράφει μία καμπύλη η οποία διέρχεται από το FOE και χωρίζει τις θετικές από τις αρνητικές μεταφορικές συνιστώσες των διανυσμάτων (A, B, C) coaxis μια και το πρόσημο της είναι ίδιο με κείνο της t_n.

Η προβολή της περιστροφικής συνιστώσας στο (M_x, M_y) ισούται με:

$$\begin{aligned}
r_n &= \left\{\left[\alpha\tfrac{xy}{f}-\beta(\tfrac{x^2}{f}+f)+\gamma y\right],\left[\alpha(\tfrac{y^2}{f}+f)-\beta\tfrac{xy}{f}-\gamma x\right]\right\}\cdot(M_x,M_y) \\
&= \left\{\left[\alpha\tfrac{xy}{f}-\beta(\tfrac{x^2}{f}+f)+\gamma y\right],\left[\alpha(\tfrac{y^2}{f}+f)-\beta\tfrac{xy}{f}-\gamma x\right]\right\}\cdot\left[\tfrac{1}{||\vec{M}||}(m_x,m_y)\right]
\end{aligned}$$

$$
\begin{aligned}
&= (-\alpha A\tfrac{xy^3}{f} - \alpha Axyf + \beta A\tfrac{x^2y^2}{f} + \beta Af^3 + \beta Ax^2 f + \beta Ay^2 f \\
&\quad + \alpha B\tfrac{x^2y^2}{f} - \beta B\tfrac{x^3y}{f} - \beta Bxyf + \gamma Bxy^2 \\
&\quad + \alpha Cx^2y - \beta Cx^3 - \beta Cxf^2 + \gamma Cxyf \\
&\quad + \alpha A\tfrac{xy^3}{f} + \alpha Axyf - \beta A\tfrac{x^2y^2}{f} - \gamma Ax^2y \\
&\quad - \alpha B\tfrac{x^2y^2}{f} - \alpha By^2f - \alpha Bf^3 - \alpha Bx^2f - \alpha Bf^3 + \beta B\tfrac{x^3y}{f} + \beta Bxyf + \gamma Bx^3 + \gamma Bxf^2 \\
&\quad + \alpha Cy^3 + \alpha Cyf^2 - \beta Cxy^2 - \gamma Cxyf)\,\tfrac{1}{||\vec{M}||} \\
&= (x^2+y^2+f^2)\,(\alpha Cy - \gamma Ay - \alpha Bf + \beta Af + \gamma Bx - \beta Cx)\,\tfrac{1}{||\vec{M}||} \\
&= (x^2+y^2+f^2)\,l(A,B,C,\alpha,\beta,\gamma;x,y)\,\tfrac{1}{||\vec{M}||}
\end{aligned}
$$

όπου $l(A,B,C,\alpha,\beta,\gamma;x,y) = \alpha Cy - \gamma Ay - \alpha Bf + \beta Af + \gamma Bx - \beta Cx$

Η εξίσωση $l = 0$ είναι μία ευθεία που διέρχεται από το AOR, ακολουθεί το πρόσημο της r_n και άρα χωρίζει τις θετικές από τις αρνητικές περιστροφικές συνιστώσες.

Τελικά έχουμε ότι το μέτρο της προβολής της οπτικής ροής στη διεύθυνση (M_x, M_y) είναι ίσο με

$$t_n + r_n = \frac{1}{||\vec{M}||}\left[\frac{W}{Z}\,k(A,B,C,x_0,y_0;x,y) + l(A,B,C,\alpha,\beta,\gamma;x,y)\right]$$

Οταν $k > 0$ και $l > 0$, τότε $t_n + r > 0$. Αντίθετα, όταν $k < 0$ και $l < 0$, τότε $t_n + r < 0$. Αρα πιθανά διανύσματα (A, B, C) coaxis με μέτρο 0 πρέπει να αναζητηθούν στα δύο τεταρτημόρια που σχηματίζουν οι $l = 0$ και $k = 0$ και για τα οποία οι k και l λαμβάνουν ετερόσημες τιμές.

B.2 Προβολή σε διανύσματα copoint

Τα διανύσματα copoint ως προς μια δυάδα πραγματικών αριθμών (r, s) ορίζονται ως εκείνα τα διδιάστατα διανύσματα (x, y) που διεύθυνσή τους είναι η:

$$(m_x, m_y) = (-y+s, x-r)$$

Το μέτρο της προβολής της μεταφορικής συνιστώσας στο μοναδιαίο (M_x, M_y) της διεύθυνσης $(-y+s, x-r)$ ισούται με:

$$\begin{aligned}
t_n &= \tfrac{W}{Z}(x-x_0, y-y_0)\cdot(M_x, M_y) = \tfrac{W}{Z}(x-x_0, y-y_0)\cdot\left[\tfrac{1}{||\vec{M}||}(m_x, m_y)\right] \\
&= \tfrac{W}{Z}(x-x_0, y-y_0)\cdot\left[\tfrac{1}{||\vec{M}||}(-y+s, x-r)\right] \\
&= \tfrac{W}{Z}(-xy+xs+x_0y-x_0s+xy-xy_0-yr+y_0r)\tfrac{1}{||\vec{M}||} \\
&= \tfrac{W}{Z}\left[y(x_0-r)-x(y_0-s)-x_0s+y_0r\right]\tfrac{1}{||\vec{M}||} \\
&= \tfrac{W}{Z}\, m(r,s,x_0,y_0;x,y)\,\tfrac{1}{||\vec{M}||}
\end{aligned}$$

όπου $m(r,s,x_0,y_0;x,y) = y(x_0-r)-x(y_0-s)-x_0s+y_0r$. Η εξίσωση $m=0$ παριστάνεται από μια ευθεία, η οποία διέρχεται από το FOE και το (r,s) και χωρίζει τις θετικές από τις αρνητικές μεταφορικές συνιστώσες.

Το μέτρο της προβολής της περιστροφικής συνιστώσας στο (M_x, M_y) ισούται με:

$$\begin{aligned}
r_n &= \left\{\left[\alpha\tfrac{xy}{f}-\beta(\tfrac{x^2}{f}+f)+\gamma y\right], \left[\alpha(\tfrac{y^2}{f}+f)-\beta\tfrac{xy}{f}-\gamma x\right]\right\}\cdot(M_x, M_y) \\
&= \left\{\left[\alpha\tfrac{xy}{f}-\beta(\tfrac{x^2}{f}+f)+\gamma y\right], \left[\alpha(\tfrac{y^2}{f}+f)-\beta\tfrac{xy}{f}-\gamma x\right]\right\}\cdot\left[\tfrac{1}{||\vec{M}||}(m_x, m_y)\right] \\
&= \left\{\left[\alpha\tfrac{xy}{f}-\beta(\tfrac{x^2}{f}+f)+\gamma y\right], \left[\alpha(\tfrac{y^2}{f}+f)-\beta\tfrac{xy}{f}-\gamma x\right]\right\}\cdot\left[\tfrac{1}{||\vec{M}||}(-y+s, x-r)\right] \\
&= (-\alpha\tfrac{xy^2}{f}+\alpha s\tfrac{xy}{f}+\beta\tfrac{x^2y}{f}+\beta yf-\beta s\tfrac{x^2}{f}-\beta sf-\gamma y^2+\gamma sy \\
&\quad +\alpha\tfrac{xy^2}{f}+\alpha xf-\alpha r\tfrac{y^2}{f}-\alpha rf-\beta\tfrac{x^2y}{f}+\beta r\tfrac{xy}{f}-\gamma x^2+\gamma rx)\tfrac{1}{||\vec{M}||}
\end{aligned}$$

$$= \tfrac{1}{f}(x^2(\beta s + \gamma f) - y^2(\alpha r + \gamma f) + xy(\alpha s + \beta r)$$

$$+xf(\alpha f + \gamma r) + yf(\beta f + \gamma s) - f^2(\alpha r + \beta s))\tfrac{1}{\|\vec{M}\|}$$

$$= \tfrac{1}{f}\; n(r, s, \alpha, \beta, \gamma; x, y)\; frac1\|\vec{M}\|$$

όπου $n(r, s, \alpha, \beta, \gamma; x, y) = x^2(\beta s+\gamma f)-y^2(\alpha r+\gamma f)+xy(\alpha s+\beta r)+xf(\alpha f+\gamma r)+yf(\beta f+\gamma s) - f^2(\alpha r + \beta s)$. Η εξίσωση $n = 0$ παριστάνεται από μία καμπύλη δευτέρου βαθμού που διέρχεται από το AOR και χωρίζει τις αρνητικές από τις θετικές μεταφορικές συνιστώσες.

Τελικά έχουμε ότι το μέτρο της προβολής ενός διανύσματος οπτικής ροής στη διεύθυνση (M_x, M_y) ισούται με $t_n + r_n = \left[\frac{W}{Z}\, m(r, s, x_0, y_0; x, y) + \frac{1}{f}\, n(r, s, \alpha, \beta, \gamma; x, y)\right] \frac{1}{\|\vec{M}\|}$. Τα διανύσματα μηδενικού μήκους βρίσκονται μόνο στις περιοχές όπου $m > 0$ και $n < 0$ ή $m < 0$ και $n > 0$.

Βιβλιογραφία

[1] Dynamic scene understanding for autonomous mobile robots. *Proceedings IEEE Conference on Computer Vision and Pattern Recognition*, pages 736–741, June 1988.

[2] On computing a 'fuzzy' focus of expansion for autonomous navigation. *Proceedings IEEE Conference on Computer Vision and Pattern Recognition*, pages 563–568, June 1989.

[3] A. M. Weng T. S. Huang N. Ahuja. Motion and Structure from Line Correspndences: Closed-Form Solution, Uniqueness and Optimazation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 318–337, 1992.

[4] R. Bajcsy. Active Perception. *Proceedings IEEE*, pages 76:996–1005, 1988.

[5] D. Ballard. Animate Vision. *Artificial Intelligence*, pages 48:57–86, 1991.

[6] J. Alloimonos I. Weiss A. Bandopadhay. Active Vision. *International Journal of Computer Vision*, 2:333–356, 1988.

[7] A. M. Waxman J. Wu F. Bergholm. Convected Activation Profiles and Receptive Fields for Real Time Measurement of Short Range Visual Motion. *Proceedings Conference on Computer Vision and Pattern Recognition*, pages 771–723, 1988.

[8] B.K.P Horn B.Schunck. Determining Optical Flow. *Artificial Intelligence*, pages 185–203, 1981.

[9] Z. Duric A. Rosenfeld L. Davis. Egomotion Analysis Based on the Frenet-Serret Motion Model. *International Journal of Computer Vision*, 15(1/2):105–122, June 1995.

[10] Y. Aloimonos Z. Duric. Estimating the Heading Direction Using Normal Flow. *International Journal of Computer Vision*, 13(13):33–56, September 1994.

[11] C. Fermuller. Basic Visual Capabilities. *PhD thesis, Computer Vision Laboratory, Center for Automation Research, University of Maryland, College Park*, May 1993.

[12] D. Heeger. Optical flow using spatiotemporal filters. *International Journal of Computer Vision*, 1:279–302, 1988.

[13] B.K.P Horn, editor. *Robot Vision.* McGraw Hill, New York, 1986.

[14] R.Y Tsai T.S Huang. Uniqueness and estimation of three-dimensional motion parameters of rigid objects with curved surfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 6:13–27, 1984.

[15] S. S. Beauchemin J. L. Barron, D. J. Fleet. Performance of Optical Flow Techniques. *International Journal of Computer Vision*, pages 12:43–79, 1994.

[16] D. Fleet A. Jepson. Computation of Component Image Velocity from Local Phase Information. *International Journal of Computer Vision*, pages 77–104, 1990.

[17] C. Tomasi T. Kanade. Shape and motion from Iimage stremas under Orthography: a Factorization Method. *International Journal of Computer Vision*, pages 137–154, 1992.

[18] D. McReynolds D. Lowe. Rigidy Checking of 3D Point Correspondences Under Perspective Projection. *Technical report, Laboratory for Computational Intelligence Department of Computer Science, University of British Columbia, Vancouver Canada.*

[19] D. Marr, editor. *Vision.* W.H. Freeman, San Francisco, 1982.

[20] D. Sinclair A. Blake D. Murray. Robust Estimation of Egomotion from Normal Flow. *International Journal of Computer Vision*, 13(13):57–69, September 1994.

[21] H.H Nagel. Displacement vectors derived from second order intensity variations in image sequences. *Computer Vision, Graphics and Image Processing*, pages 21:85–117, 1983.

[22] H.H Nagel. On the estimation of optical flow : Relations between different approaches and some new results. *Artificial Intelligence*, pages 33:459–483, 1987.

[23] V. Nalwa, editor. *A Guided Tour of Computer Vision.* Winston:New York, 1993.

[24] P.Anandan. A computational Framework and an Algorithm for the Measurement of Visual Motion. *International Journal of Computer Vision*, 2:283–310, 1989.

[25] K. Prazdny. Egomotion and relative depth map from optical flow. *Biological Cybernetics*, pages 36:87–102, 1980.

[26] A. Shashua. Projective Structure from two Uncalibrated Images: Structure from Motion and Recognition. *Tecnical report, Massachusetts institute of tecnology, Artificial intelligence laboratory*, September 1992.

[27] A. Singh. Optic Flow Computation: A Unified Perspective. *IEEE Computer Society Press*, 1992.

[28] O.D Faugeras F. Lustman G. Toscani. Motion and structure from motion from point and line matches. *Proc. International Conference on Computer Vision*, pages 25–34, 1987.

[29] S. Ullman. The Interpretation of Structure from Motion. *Proceedings of the Royal Society, London*, pages 405–426, 1979.

[30] E. Micheli V. Torre S. Uras. The Accuracy of the Computation of Optical Flow and of the Recovery of Motion Parameters. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(5):434–447, May 1993.

[31] B. Bhanu W. Burger. Estimating 3-D Egomotion from Perspective Image Sequences. *IEEE Trabsactions on Pattern Analysis and MAchine Intelligence*, 12:1040–1058, 1990.

[32] B.K.P Horn E.J Weldon. Computationally efficient methods for recovering translational motion. *Proc. International Conference on Computer Vision*, pages 2–11, 1987.

[33] S. Zeki. Visual Image in Mind and Brain. *Scientific American*, pages 69–76, September 1992.