

4/Γ
549

ΤΕΙ ΠΕΙΡΑΙΑ
ΣΤΕΦ
ΤΜΗΜΑ ΗΛΕΚΤΡΟΛΟΓΙΑΣ

2011

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ:

Όραση Υπολογιστών

Αξιοποίηση της κάμερας ως βασικού αισθητηρίου ενός
αυτόματου οχήματος

Βασικές αρχές, τεχνικές ελέγχου και εφαρμογή σε αυτόνομο όχημα



ΣΠΟΥΔΑΣΤΕΣ:

ΚΩΝΣΤΑΝΤΙΝΟΣ ΠΑΝΑΓΙΩΤΟΠΟΥΛΟΣ

ΠΑΝΑΓΙΩΤΗΣ ΠΙΕΡΡΑΚΟΣ

ΕΙΣΗΓΗΤΗΣ:

ΠΑΝΑΓΙΩΤΗΣ ΣΙΝΙΟΡΟΣ

ΠΕΡΙΕΧΟΜΕΝΑ

Περίληψη		3
Λέξεις κλειδιά		3
Abstract		3
Key Words		3
Κεφάλαιο 1	ΕΙΣΑΓΩΓΗ	4
1.1	Διαφορετικοί τύποι στόχων και συστημάτων εικόνας	4
1.2	Εξαρτήματα ενός συστήματος εικόνας	7
1.3	Συστήματα απεικόνισης	8
1.4	Αναγνώριση σχεδίων για την όραση υπολογιστών	10
1.5	Αξιολόγηση απόδοσης των αλγορίθμων	11
Κεφάλαιο 2	ΠΡΟΒΟΛΙΚΗ ΓΕΩΜΕΤΡΙΑ	13
2.1	Προβολικός Χώρος	13
2.2	Συντεταγμένες	13
2.3	Προβολική Γεωμετρία και Μετασχηματισμοί στο 2Δ Χώρο	14
2.3.1	Το 2Δ Προβολικό Επίπεδο	14
2.3.2	Προβολικοί Μετασχηματισμοί	18
2.3.3	Ιεραρχία Μετασχηματισμών	19
2.4	Προβολική Γεωμετρία και Μετασχηματισμοί στο 3Δ Χώρο	20
2.4.1	Σημεία και Προβολικοί Μετασχηματισμοί	20
2.4.2	Ιεραρχία Μετασχηματισμών	21
2.5	Μοντέλα Καμερών	22
2.5.1	Πεπερασμένη Κάμερα	23
2.5.2	Πεπερασμένη προβολική κάμερα	25
2.5.3	Η Κάμερα στο Άπειρο	28
Κεφάλαιο 3	ΟΡΑΣΗ ΥΠΟΛΟΓΙΣΤΩΝ ΓΙΑ ΤΙΣ ΑΝΑΓΚΕΣ ΟΧΗΜΑΤΟΣ	30
3.1	Γιατί αντίληψη και δράση;	30
3.2	Γιατί αντίληψη και όχι μόνο όραση ;	31
3.3	Κατάλληλος χώρος ερμηνείας	32
3.4	Διαφορικά πρότυπα για την αντίληψη	36
3.5	Πληροφορίες σε μια εικόνα	37
3.6	Σε ποίο σημείο εισέρχεται η ανθρώπινη γνώση;	38
3.7	Σφαιρικά ολοκληρώματα για την αξιολόγηση της κατάστασης	39
3.8	Επάρκεια συστημάτων εικόνας	40
3.9	Τεχνικά συστήματα εναντίον των βιολογικών	43
3.10	Νοημοσύνη	45
3.11	Βασικές Σχέσεις	49
3.12	Τρισδιάστατος χώρος και χρόνος	52
Κεφάλαιο 4	ΤΟ ΛΟΓΙΣΜΙΚΟ ROBOREALM	55
4.1	Ανίχνευση Ακμών (EDGES)	56
4.2	Φίλτρα	61
4.3	Έλεγχος (CONTROLS)	70

4.4	Display	71
4.5	Ιστογράμμο - histogram	71
4.6	Matching Modules	76
4.7	Μετασηµατισµοί	81
4.8	Προκαθορισµενες δοµικες ενοτητες	85
4.8.1	3d απεικονιση	85
4.8.2	Ανιχνευτής Δαπεδου	88
4.8.3	Γραµμή λέιζερ	90
4.8.4	Σηµείο λέιζερ	90
4.8.5	Μετακίνηση	92
4.8.6	Σχεδιασµος πορειας	93
4.8.7	Σχεδιαση επιφανειας	94
4.8.8	Ανιχνευτής τοίχων	95
4.8.9	Ορίζοντας	96
Κεφάλαιο 5	Η ΕΦΑΡΜΟΓΗ ΜΑΣ	98
5.1	Ανάλυση των βαθµίδων	99
5.2	Βήµατα της εφαρμογης μας	103
	Προτάσεις για µελλοντικη εργασία	106
	Βιβλιογραφία	107

ΠΕΡΙΛΗΨΗ

Στην παρούσα πτυχιακή εργασία παρουσιάζονται οι βασικές αρχές της Όρασης Υπολογιστών και οι τεχνικές που μας επιτρέπουν να σχεδιάσουμε ένα πλήρως αυτονομο όχημα. Η παρουσίαση της λειτουργίας των διαφόρων υποσυστημάτων καθώς και του λογισμικού Roborealms είναι απαραίτητη αφού με την βοήθεια αυτών γίνεται η ανάλυση και ο έλεγχος του οχήματος. Η δημιουργία αυτόνομων οχημάτων που θα κινούνται με την χρησιμοποίηση της όρασης αποτελεί ένα άκρως ενδιαφέρον ζήτημα και δημιουργεί αρκετές μελλοντικές προσδοκίες με την ανάπτυξη των υπάρχουσων τεχνολογιών.

ΛΕΞΕΙΣ ΚΛΕΙΔΙΑ

Όραση Υπολογιστών, ρομποτική όραση, αυτονομο οχημα, roborealms, επεξεργασία εικονας, προσομοίωση, τρισδιαστατος χωρος, δισδιαστατο επιπεδο, αισθητήριο, καμερα.

ABSTRACT

In the present thesis we present the basic principles of Computer Vision and the techniques that will allow us to build an autonomous vehicle. The presentation of the functionality of the various subsystems, as well as the Roborealms software, is necessary, as rely on them for the analysis and the control of the vehicle. Building autonomous vehicles that will use vision as their basic sense is an extremely interesting subject and with the development in current technologies creates great expectations for the future.

KEYWORDS

Computer Vision, robotic vision, autonomous vehicle, Roborealms software, image processing, simulation, 3d space, 2d plane, sensor, camera.

ΚΕΦΑΛΑΙΟ 1

Εισαγωγή

Το πεδίο της όρασης υπολογιστών είναι τόσο πολύπλοκο και υπάρχουν τόσες πολλές διαφορετικές προσεγγίσεις στην εφαρμογή της που φαίνεται λογικό πρώτα να εξετάσουμε και να διευκρινίσουμε την περιοχή στην οποία σκοπεύουμε να εξετάσουμε. Πολλές προσεγγίσεις στη όραση υπολογιστών έχουν αρχίσει με το παράδειγμα ότι τα εύκολα πράγματα πρέπει να αντιμετωπίζονται πρώτα, όπως η ενιαία ερμηνεία εικόνας στιγμιότυπων. Μια επέκταση στις πιο σύνθετες εφαρμογές μπορεί αργότερα να στηριχτεί στην εμπειρία που αποκτιέται. Η προσέγγισή μας αντίθετα ήταν να χωρίσουμε το δυναμικό πεδίο από (το σχεδόν) στατικό αντίστοιχό της ευθύς εξαρχής και να παραγάγουμε τις επαρκείς μεθόδους για αυτήν την συγκεκριμένη περιοχή. Για να προετοιμάσουν το έδαφος για την επιτυχία, οι αρκετά ικανές μέθοδοι και οι αντιπροσωπεύσεις γνώσης πρέπει να εισαχθούν από την αρχή.

1.1 ΔΙΑΦΟΡΕΤΙΚΟΙ ΤΥΠΟΙ ΣΤΟΧΩΝ ΚΑΙ ΣΥΣΤΗΜΑΤΩΝ ΕΙΚΟΝΑΣ

Το σχήμα 1 παρουσιάζει αντιπαραθέσεις διάφορων στόχων όρασης που εμφανίζονται στη καθημερινή ζωή. Για τους ανθρώπους, η ερμηνεία στιγμιότυπων φαίνεται εύκολη, γενικά, όταν είναι καλά γνωστή η περιοχή στην οποία έχει ληφθεί η εικόνα. Τείνουμε να φανταστούμε το χρονικό πλαίσιο και το χρόνο όταν τραβηχτεί η εικόνα. Από την κίνηση της κηλίδας και τις ασυνήθιστες θέσεις, η ενσωμάτωση του στιγμιότυπου σε έναν γνωστό ελιγμό ολοκληρώνεται. Γενικότερα, ακόμη και οι απλές εικόνες απαιτούν τη γνώση υποβάθρου σχετικά με την κίνηση για πιο σε βάθος να κατανοήση αυτό αγνοείται συχνά στην όραση μηχανών ή υπολογιστών. Η προσέγγιση που αναλύεται σε αυτήν την εργασία εξετάζει τις διαδικασίες κινήσεων στο τρισδιάστατο χώρο και το χρόνο ως την βασική γνώση που απαιτείται για την κατανόηση των ακολουθιών εικόνας. Αυτό παράγει ένα φυσικό πλαίσιο για τη γλώσσα και τους όρους που θα χρησιμοποιήσουμε.

Μια άλλη μεγάλη διαφορά στις μεθόδους και τις προσεγγίσεις προκύπτει από το γεγονός ότι η κάμερα που παράγει το τηλεοπτικό σήμα είναι είτε στάσιμη είτε κινούμενη. Εάν κινείται, η γραμμική ή περιστροφική κίνηση της μπορεί επίσης να απαιτήσει ειδική μεταχείριση. Η επιτήρηση γίνεται, συνήθως, από μια στάσιμη θέση ενώ η κάμερα μπορεί να κινείται (περιστροφή γύρω από έναν κάθετο άξονα, συχνά επίσης αποκαλούμενο παρέκκλιση) και κλίση (περιστροφή γύρω από οριζόντιο άξονα) για να αυξήσει το συνολικό οπτικό πεδίο του. Σε αυτήν την περίπτωση, η κίνηση εισάγεται εσκεμμένα και ελέγχεται καλά, έτσι ώστε να μπορεί να ληφθεί υπόψη κατά τη διάρκεια

της αξιολόγησης εικόνας. Εάν η κίνηση πρόκειται να ελεγχθεί βασισμένη στην εικόνα , το σώμα που φέρει τη κάμερα μπορεί να υπόκειται σε ισχυρές διαταραχές, που δεν μπορούν να προβλεφθούν, γενικά.

Εικονογραφική όραση (ενιαία ερμηνεία εικόνας)	Όραση κινήσεων
Επιτήρηση ανίχνευση, επιθεώρηση (θήραμα)	Έλεγχος κινήσεων (θηρευτής)
[υβριδικά συστήματα]	
Μονοφθαλμικό σύστημα κινήσεων Παθητικό	Πολυφθαλμικό σύστημα Ενεργό: ο τύπος άμεσης σταθεροποίησης, προσοχή εστίαστηκε
Δισδιάστατη μορφή Εκτός σύνδεσης	Χωρική ερμηνεία Σύνδεση σε πραγματικό χρόνο
Μονοχρωματικός Ένταση	Όραση χρώματος Εύρος

Σχήμα 1. Τύποι συστημάτων εικόνας και στόχων εικόνας

Σε περιπτώσεις με μεγάλα ποσοστά κίνησης ή περιστροφής, μπορεί να είναι δύσκολη η αξιολόγηση της εικόνας , λόγω του χρόνου καθυστέρησης που εισάγεται με το χειρισμό και την ερμηνεία των μεγάλων ποσοστών πληροφοριών στην εικόνα και ως εκ τούτου ο σταθερός έλεγχος του οχήματος δεν μπορεί πλέον να είναι δυνατός.

Τα βιολογικά συστήματα έχουν αναπτύξει τη συνεργασία μεταξύ αδρανούς και οπτικής αξιολόγησης στοιχείων. Επίσης από βιολογικής άποψης υπάρχει διαφοροποίηση των συστημάτων εικόνας σε δύο κατηγορίες: στα «θήραματα» και τα «αρπακτικά». Τα πρώτα προσπαθούν να καλύψουν ταυτόχρονα ένα μεγάλο οπτικό πεδίο για την ανίχνευση των αρπακτικών ζώων αρκετά πριν αυτά προσεγγίσουν από οποιαδήποτε πιθανή κατεύθυνση. Τα αρπακτικά ζώα κινούνται για να βρουν το θήραμα, και κατά τη διάρκεια της τελικής προσέγγισης καθώς επίσης και στην αναζήτηση πρέπει να υπολογίσουν τη θέση και την ταχύτητά τους σε σχέση με την κίνηση του θηράματος και να είναι αρκετά ακριβές για να πετύχουν το στόχο τους. Η στερεοσκοπική όραση και η υψηλή ανάλυση στη κατεύθυνση της κίνησης παρέχει τα πλεονεκτήματα, και η φύση πέτυχε την ανάπτυξη αυτών των συνδυασμών στο μάτι.

Οι γρήγορα κινούμενοι στόχοι μπορούν να ακολουθηθούν στην ομαλή αναζήτηση, μειώνοντας επίσης τη θαμπάδα κινήσεων για αυτό το συγκεκριμένο αντικείμενο ενδιαφέροντος η επιδείνωση της αναγνώρισης και η καταδίωξη άλλων αντικειμένων λιγότερου ενδιαφέροντος γίνονται αποδεκτές.

Δεδομένου ότι οι εικόνες είναι μόνο σε δύο διαστάσεις, το τρισδιάστατο πλαίσιο φαίνεται το φυσικότερο για την ερμηνεία μιας εικόνας. Αυτό μπορεί να ισχύει για σχεδόν επίπεδα αντιμετωπιζόμενα αντικείμενα, όπως ένα πουλί σε ένα τοπίο. Όταν έχουμε μια επίπεδη επιφάνεια με τον οπτικό άξονα σχεδόν παράλληλο σε αυτή, με μια ανύψωση ελαφρώς επάνω από το έδαφος, η κατάσταση είναι αρκετά διαφορετική. Σε αυτήν την περίπτωση, κάθε γραμμή στην εικόνα αντιστοιχεί σε μια διαφορετική απόσταση από το έδαφος, και το ίδιο τρισδιάστατο αντικείμενο στην επιφάνεια φαίνεται αρκετά διαφορετικό στο μέγεθος σύμφωνα με αυτό που εμφανίζεται στην εικόνα. Αυτός είναι ο λόγος για τον οποίο η ομοιογενώς διανεμημένη επεξεργασία εικόνας από τις διανυσματικές μηχανές έχει έναν σχετικό χρόνο στην παρουσίαση αποδοτικότητάς του, οι τοπικά προσαρμοσμένες μέθοδοι στις περιοχές της εικόνας φαίνονται πιο ελπιδοφόρες σε αυτήν την περίπτωση και έχει αποδείξει την ανωτερότητά τους. Ερμηνεύοντας τις ακολουθίες της εικόνας στο τρισδιάστατο διάστημα με τις αντίστοιχες βάσεις γνώσεων από την αρχή επιτρέπει την εύκολη προσαρμογή στις διαφορές για τα ενιαία αντικείμενα. Φυσικά, η ανάλυση των καταστάσεων που έχουν τα διάφορα αντικείμενα στις διάφορες αποστάσεις πρέπει τώρα να γίνει σε ένα χωριστό επίπεδο, στηριγμένος στα αποτελέσματα όλων των προηγούμενων βημάτων. Αυτό είναι ένας από τους πρωταρχικούς παράγοντες στο σχεδιασμό της αρχιτεκτονικής για τρίτη γενιά σύστημα εικόνας (EMS).

Η κατανόηση των διαδικασιών της κίνησης των τρισδιάστατων αντικειμένων στο τρισδιάστατο χώρο ενώ το σώμα που φέρει την κάμερα κινείται επίσης στο τρισδιάστατο χώρο, φαίνεται να είναι ένα από τα δυσκολότερα έργα στην όραση σε πραγματικό χρόνο. Χωρίς τη βοήθεια της αδρανούς αντίληψης για τον διαχωρισμό από τη σχετική κίνηση, αυτό μπορεί μετά βίας να επιτευχθεί, τουλάχιστον στις δυναμικές καταστάσεις.

Η άμεση μέτρηση σειράς από τους πρόσθετους αισθητήρες όπως οι ανιχνευτές σειράς ραντάρ ή λέιζερ (LRF) θα διευκόλυνε το στόχο της όρασης. Λόγω της σχετικής απλότητας και της χαμηλής ανάγκης τους για δύναμη υπολογισμού, αυτά τα συστήματα έχουν βρει σχετικά διαδεδομένη εφαρμογή στον αυτοκίνητο τομέα. Εντούτοις, όσον αφορά την ανάλυση και την ευελιξία της εκμετάλλευσης στοιχείων καθώς επίσης και του κόστους υλικού και τον όγκο εγκαταστάσεων που

απαιτείται, έχουν την πολύ μικρότερη δυνατότητα από τις παθητικές κάμερες μακροπρόθεσμα με τη δύναμη υπολογισμού άφθονα διαθέσιμη.

1.2 ΕΞΑΡΤΗΜΑΤΑ ΕΝΟΣ ΣΥΣΤΗΜΑΤΟΣ ΕΙΚΟΝΑΣ

Η όραση υπολογιστών είναι ένα σύνθετο θέμα. Έτσι είναι πιο χρήσιμο να το διαιρέσουμε στα διάφορα εξαρτήματα του ή τις ενότητες λειτουργίας του. Σε αυτό το επίπεδο, αυτό είναι πολύ ευκολότερο να συγκρίνουμε ένα τεχνικό σύστημα με ένα βιολογικό σύστημα. Από αυτή την άποψη, η βασική λειτουργία μιας βιολογικής και μηχανικής όρασης περιλαμβάνουν τα ακόλουθα εξαρτήματα:

Πηγή ακτινοβολίας. Εάν καμία ακτινοβολία δεν εκπέμπεται από την εικόνα ή το αντικείμενο ενδιαφέροντος, τίποτα δεν μπορεί να παρατηρηθεί ή να υποβληθεί σε επεξεργασία. Κατά συνέπεια ο κατάλληλος φωτισμός είναι απαραίτητος για τα αντικείμενα που δεν ακτινοβολούν .

Κάμερα. Η «Κάμερα» συλλέγει την ακτινοβολία που παραλαμβάνει από το αντικείμενο κατά τέτοιο τρόπο ώστε η προέλευση της ακτινοβολίας μπορεί να επισημανθεί. Στην απλούστερη περίπτωση αυτό είναι ακριβώς ένας οπτικός φακός. Αλλά θα μπορούσε επίσης να είναι ένα απολύτως διαφορετικό σύστημα, παραδείγματος χάριν, ένα οπτικό φασματόμετρο απεικόνισης, ένα των ακτινών X tomograph, ή ένα πιάτο μικροκυμάτων.

Αισθητήρας. Ο αισθητήρας μετατρέπει τη λαμβανόμενη πυκνότητα ροής της ακτινοβολίας σε ένα κατάλληλο σήμα για την περαιτέρω επεξεργασία. Για ένα σύστημα απεικόνισης κανονικά απαιτείται μια 2 διαστάσεων σειρά αισθητήρων για να συλλάβει τη χωρική διανομή της ακτινοβολίας. Με ένα κατάλληλο σύστημα ανίχνευσης σε μερικές περιπτώσεις ένας ενιαίος αισθητήρας ή μια σειρά των αισθητήρων θα μπορούσε να είναι ικανοποιητικός.

Μονάδα επεξεργασίας. Επεξεργάζεται τα εισερχόμενα, που έχουν γενικά υψηλά –διαστασιακά στοιχεία, που εξάγουν τα κατάλληλα χαρακτηριστικά γνωρίσματα που μπορούν να χρησιμοποιηθούν στο να μετρηθούν οι ιδιότητες των αντικειμένων και να τις ταξινομήσουν σε κατηγορίες. Ένα άλλο σημαντικό συστατικό είναι ένα σύστημα μνήμης για να συλλέξει και να αποθηκεύσει τη γνώση περίπου η σκηνή, συμπεριλαμβανομένων των μηχανισμών για να διαγράψει ασήμαντο πράγματα.

Δράστες. Οι δράστες αντιδρούν στο αποτέλεσμα της οπτικής παρατήρησης. Γίνονται ένα αναπόσπαστο τμήμα του συστήματος όταν το σύστημα αποκρίνεται ενεργά στην παρατήρηση με, παραδείγματος χάριν, την καταδίωξη ενός αντικειμένου ή με τη χρησιμοποίηση μιας Οπτικός - καθοδηγημένης ναυσιπλοΐας (ενεργό όραμα, κύκλος δράσης αντίληψης).

Πίνακας 1: Ενότητες λειτουργίας της ανθρώπινης και της μηχανικής όρασης

Στόχος	Ανθρώπινη όραση	Μηχανική όραση
Απεικόνιση	κυρίως από την αντανάκλαση του φωτός από αδιαφανείς επιφάνειες	Ενεργός (ελεγχόμενος φωτισμός) χρησιμοποιώντας ηλεκτρομαγνητική, μοριακή, και ακουστική ακτινοβολία
Σχηματισμός εικόνας	Διαθλαστικό οπτικό σύστημα	Διάφορα συστήματα
Έλεγχος της ακτινοβολίας	Ελεγχόμενος μυς (ΕΚΠΑΙΔΕΥΣΙΜΟΣ)	Μηχανοποιημένα ανοίγματα, ρόδες φίλτρων, ρυθμιζόμενα φίλτρα
Εστίαση	ελεγχόμενος μυς- αλλαγή του εστιακού μήκους	Συστήματα(Autofocus) αυτόματης εστίασης βασισμένα στις διάφορες αρχές των μετρήσεων απόστασης
Ψήφισμα ακτινοβολίας	Λογαριθμική ευαισθησία	Γραμμική ευαισθησία, κβαντοποίηση μεταξύ 8 - και 16 μπιτ λογαριθμική ευαισθησία
Εντοπισμός	Ιδιαίτερα κινητικός βολβός του ματιού	Σαρωτές και τοποθετημένες σε ρομποτό κάμερες
Επεξεργασία και ανάλυση	Ιεραρχικά οργανωμένη μαζικά παράλληλη επεξεργασία	Τμηματική επεξεργασία κυρίως κυρίαρχη χρήση παράλληλης επεξεργασίας όχι σε γενική χρήση

1.3 ΣΥΣΤΗΜΑΤΑ ΑΠΕΙΚΟΝΙΣΗΣ

Τα συστήματα απεικόνισης καλύπτουν όλες τις διαδικασίες που περιλαμβάνονται στο σχηματισμό μιας εικόνας από τα αντικείμενα και τους αισθητήρες που μετατρέπουν την ακτινοβολία στα ηλεκτρικά σήματα, και περαιτέρω στα ψηφιακά σήματα που μπορούν να υποβληθούν σε επεξεργασία από έναν υπολογιστή. Γενικά ο στόχος είναι να επιτύχουμε ένα σήμα από ένα

αντικείμενο με μια τέτοια μορφή ώστε να ξέρουμε όπου είναι (γεωμετρία), και τι είναι ή ποιες ιδιότητες έχει.

Σχήμα 2: Αλυσίδα των βημάτων που συνδέουν ένα αντικείμενο με το σήμα που μετρείται από ένα σύστημα απεικόνισης.



Είναι σημαντικό να σημειωθεί ότι ο τύπος απάντησης που λαμβάνουμε από αυτές τις δύο υπονοούμενες ερωτήσεις εξαρτώνται από το σκοπό λειτουργίας του συστήματος. Η απάντηση θα μπορούσε να είναι είτε ποιοτικής είτε ποσοτικής φύσης. Για μερικές εφαρμογές θα μπορούσε να είναι ικανοποιητικό να ληφθεί μια ποιοτική απάντηση όπως «εκεί είναι ένα αυτοκίνητο στον αριστερό σας χέρι που έρχεται». Το «Τι» και «όπου» είναι οι ερωτήσεις που μπορούν να καλύψουν ολόκληρη την πρόταση από «εκεί είναι κάτι.» μια συγκεκριμενοποίηση του αντικείμενου υπό μορφή κατηγορίας, σε μια λεπτομερή ποσοτική περιγραφή των διάφορων ιδιοτήτων των αντικειμένων ενδιαφέροντος.

Η σχέση που συνδέει την ιδιότητα του αντικείμενου με το σήμα που μετρείται από ένα σύστημα απεικόνισης είναι μια σύνθετη αλυσίδα διαδικασιών (Σχήμα 2). Η αλληλεπίδραση της ακτινοβολίας με το αντικείμενο (που ενδεχομένως χρησιμοποιεί ένα κατάλληλο σύστημα φωτισμού) αναγκάζει το αντικείμενο να εκπέμψει ακτινοβολία. Μια μερίδα (συνήθως μόνο ένα πολύ μικρό μέρος) της εκπεμπόμενης ενέργειας της ακτινοβολίας συλλέγεται από το οπτικό σύστημα και θεωρείται ως ακτινοβολία (της ακτινοβολίας ενεργή /περιοχή). Ένας αισθητήρας (ή μια σειρά αισθητήρων) μετατρέπει τη λαμβανόμενη ακτινοβολία σε ένα ηλεκτρικό σήμα που επιλέγεται στη συνέχεια και μεταλλάσσεται για να διαμορφώσει μια ψηφιακή εικόνα ως σειρά ψηφιακών αριθμών.

Μόνο τα άμεσα συστήματα απεικόνισης παρέχουν ένα άμεσο σημείο για να δείξουν την αλληλεπίδραση μεταξύ των σημείων των αντικειμένων στον τρισδιάστατο κόσμο και το πλάνο της εικόνας. Τα έμμεσα συστήματα απεικόνισης δίνουν επίσης μια διανεμημένη ακτινοβολία στο χώρο αλλά χωρίς την ένα προς ένα σχέση. Η παραγωγή μιας εικόνας απαιτεί την αναδημιουργία του αντικείμενου από την αντιληπτή ακτινοβολία. Τα παραδείγματα τέτοιων τεχνικών απεικόνισης περιλαμβάνουν την απεικόνιση ραντάρ, διάφορες τεχνικές για τη φασματική απεικόνιση, την ακουστική απεικόνιση, την τομογραφική απεικόνιση, και την απεικόνιση μαγνητικής αντίληξης.

1.4 Αναγνώριση σχεδίων για την όραση υπολογιστών

Ο βασικός στόχος της επεξεργασίας σήματος στην όραση υπολογιστών είναι η εξαγωγή των «κατάλληλων χαρακτηριστικών γνωρισμάτων» για την επόμενη επεξεργασία για να αναγνωρίσει και να ταξινομήσει τα αντικείμενα. Αλλά τι είναι ένα κατάλληλο χαρακτηριστικό γνώρισμα; Αυτό καθορίζεται δυσκολότερα απ'ό, τι σε άλλες εφαρμογές της επεξεργασίας σήματος. Βεβαίως μια από καθορισμένη με σαφήνεια μαθηματική περιγραφή της τοπικής δομής είναι μια σημαντική βάση. Καθώς τα σήματα υποβάλλονται σε επεξεργασία στην όραση υπολογιστών προέρχονται από δυναμικές τρισδιάστατες σκηνές, τα σημαντικά αυτά χαρακτηριστικά γνωρίσματα περιλαμβάνουν επίσης την κίνηση και διάφορες τεχνικές για να προκύψει το βάθος στις εικόνες συμπεριλαμβανομένων των στερεών, το σχήμα της σκίασης και των φωτομετρικών στερεών, και βάθος από την εστίαση.

Υπάρχουν μικρές αμφιβολίες ότι οι μη γραμμικές τεχνικές είναι κρίσιμες για την εξαγωγή χαρακτηριστικών γνωρισμάτων στην όραση υπολογιστών. Εντούτοις, συγκρινόμενες με αυτές του γραμμικού φίλτρου, οι τεχνικές αυτές είναι ακόμα στα σπάργανα. Δεν υπάρχει καμία απλή μη γραμμική τεχνική αλλά υπάρχει ένας πλήθος τέτοιων τεχνικών συχνά προσαρμοσμένες σε έναν ορισμένο σκοπό. Σε αυτό το σημείο, παραθέτουμε μια επισκόπηση των διάφορων κατηγοριών μη γραμμικών τεχνικών φίλτρων και εστιάζουμε σε μια πρώτη εντολής αντιπροσώπευση των μη γραμμικών φίλτρων από το συνδυασμό γραμμικής συνέλιξης και μη γραμμικού σημείου χειρισμούς και μη γραμμικό φιλτράρισμα διάχυσης.

Σε γενικές γραμμές, η ταξινόμηση ρουτινών δεν είναι τίποτα σύνθετο. Πάρτε μερικά κατάλληλα γνωρίσματα και χωρίστε τα σε κατηγορίες. Γιατί τότε είναι τόσο δύσκολο για ένα σύστημα όρασης υπολογιστών να αναγνωρίσει τα αντικείμενα; Το βασικό πρόβλημα συσχετίζεται με το γεγονός ότι η διαστατικότητα του διαστήματος εισαγωγής είναι μεγάλο. Σε γενικές γραμμές, θα ήταν δυνατό να χρησιμοποιηθεί η ίδια η εικόνα ως είσοδος για έναν στόχο ταξινόμησης, αλλά καμία πραγματική τεχνική ταξινόμησης δεν είναι στατιστική, νευρωνικά, ή συγκεχυμένα -θα ήταν σε θέση για να χειριστούν τέτοια υψηλό-διαστατικά διαστήματα γνωρισμάτων. Επομένως, η ανάγκη προκύπτει για να εξαγάγει τα χαρακτηριστικά γνωρίσματα και για να τα χρησιμοποιήσει για την ταξινόμηση.

Δυστυχώς, οι τεχνικές για την επιλογή χαρακτηριστικών γνωρισμάτων πολύ συχνά έχουν παραμεληθεί στην όραση υπολογιστών. Δεν έχουν αναπτυχθεί στον ίδιο βαθμό όπως η ταξινόμηση, όπου γίνεται καλά κατανοητό ότι οι διαφορετικές τεχνικές (ειδικά στατιστικές και νευρικές τεχνικές) μπορούν να εξεταστούν κάτω από μια ενοποιημένη άποψη.

Η αναγνώριση του αντικειμένου μπορεί να εκτελεσθεί μόνο εάν είναι δυνατό να αντιπροσωπευθεί η γνώση με έναν κατάλληλο τρόπο. Σε απλές περιπτώσεις η γνώση μπορεί ακριβώς να στηριχτεί στα απλά πρότυπα. Σε πιο σύνθετες περιπτώσεις αυτό δεν είναι ικανοποιητικό.

1.5 Αξιολόγηση απόδοσης των αλγορίθμων

Η συστηματική αξιολόγηση των αλγορίθμων για την όραση υπολογιστών έχει παραμεληθεί ευρέως. Για έναν νεοφώτιστο στην όραση υπολογιστών με ένα υπόβαθρο εφαρμοσμένης μηχανικής ή μια γενική εκπαίδευση στις φυσικές επιστήμες αυτό είναι μια παράξενη εμπειρία. Του φαίνεται σαν κάποιος να παρουσιάζει τα αποτελέσματα των μετρήσεων χωρίς το δόσιμο των φραγμών λάθους ή ακόμα και σκέψη για πιθανά στατιστικά και συστηματικά λάθη.

Ποια είναι η αιτία αυτής της κατάστασης; Αφ' ενός, είναι αλήθεια ότι μερικά προβλήματα στην όραση υπολογιστών είναι πολύ δύσκολα και ότι είναι ακόμα πιο δύσκολο να εκτελεσθεί μια περίπλοκη ανάλυση λάθους. Αφ' έτερου, η κοινότητα όρασης υπολογιστών έχει αγνοήσει σε μεγάλο βαθμό το γεγονός ότι οποιοσδήποτε αλγόριθμος είναι τόσο καλός όσο η αντικειμενικές και στερεές αξιολόγησης και η επαλήθευσής του.

Ευτυχώς, αυτή η παρερμηνεία έχει αναγνωριστεί στο μεταξύ και υπάρχουν σοβαρές προσπάθειες εν εξελίξει που προσπαθούν να καθιερώσουν γενικά αποδεκτούς κανόνες για την ανάλυση απόδοσης των αλγορίθμων της όρασης υπολογιστών. Τα τρία σημαντικά κριτήρια για την απόδοση των αλγορίθμων όρασης υπολογιστών είναι:

- **Επιτυχής λύση του στόχου.** Οποιοσδήποτε επαγγελματίας δίνει σε αυτό την ύψιστη προτεραιότητα. Αλλά και ο σχεδιαστής ενός αλγορίθμου πρέπει να καθορίσει ακριβώς για ποιο στόχο είναι κατάλληλος και τι όρια έχει.
- **Ακρίβεια.** Αυτό περιλαμβάνει μια ανάλυση των στατιστικών λαθών υπό προσεκτικά καθορισμένους όρους
- **Ταχύτητα.** Αυτό πάλι είναι ένα σημαντικό κριτήριο για τη δυνατότητα εφαρμογής ενός αλγορίθμου.

Υπάρχουν διαφορετικοί τρόποι να αξιολογηθούν οι αλγόριθμοι σύμφωνα με τα κριτήρια που προηγήθηκαν. Ιδανικά αυτό πρέπει να περιλάβει τρεις κατηγορίες μελετών:

- **Αναλυτικές μελέτες.** Αυτό είναι ο από μαθηματική άποψη ο αυστηρότερος τρόπος να ελεγχθούν οι αλγόριθμοι, να ελεγχθεί η διάδοση λάθους, και να προβλεφθούν οι καταστροφικές αποτυχίες.
- **Δοκιμές απόδοσης με παραγμένες τον υπολογιστή εικόνες.** Αυτές οι δοκιμές είναι χρήσιμες δεδομένου ότι μπορούν να πραγματοποιηθούν υπό τις προσεκτικά ελεγχόμενες συνθήκες.
- **Δοκιμές απόδοσης με πραγματικές εικόνες.** Αυτό είναι η τελική δοκιμή για τις πρακτικές εφαρμογές.

ΚΕΦΑΛΑΙΟ 2

Προβολική Γεωμετρία – Βασικές Αρχές

Στο κεφάλαιο αυτό εισάγονται οι κυριότερες ιδέες από το χώρο της προβολικής γεωμετρίας και σκοπός είναι να γίνει η σύνδεση τους με την όραση υπολογιστών. Συγκεκριμένα, δίνεται έμφαση στους μετασχηματισμούς του 2Δ και 3Δ κόσμου καθώς και στην γεωμετρία πολλαπλών όψεων.

Οι προβολικοί μετασχηματισμοί είναι συχνοί στην καθημερινότητα. Σε μια φωτογραφία μπορεί κανείς να δει τετραγωνικά ή κυκλικά σχήματα που δεν ανταποκρίνονται στην πραγματικότητα. Ο μετασχηματισμός που αντιστοιχεί επίπεδα σχήματα του κόσμου στην εικόνα είναι ένα παράδειγμα προβολικού μετασχηματισμού.

Ο προβολικός μετασχηματισμός έχει τις παρακάτω ιδιότητες. Δεν διατηρεί το σχήμα, το μήκος των ευθυγράμμων τμημάτων, γωνίες, αποστάσεις και λόγους μηκών ευθυγράμμων τμημάτων. Δηλαδή, ένας κύκλος ενδέχεται να μετασχηματιστεί σε μία έλλειψη, ένα ορθογώνιο σε ρόμβο. Ωστόσο, η ιδιότητα που διατηρείται είναι αυτή της *ευθύτητας*. Δηλαδή, συνευθειακά σημεία θα μετασχηματιστούν σε συνευθειακά σημεία. Η τελευταία αυτή ιδιότητα της διατήρησης των ευθειών είναι αυτή που ουσιαστικά ορίζει τους προβολικούς μετασχηματισμούς.

2.1 Προβολικός Χώρος

Για τον ορισμό του *προβολικού χώρου* θα γίνει χρήση του Ευκλείδειου χώρου, για λόγους ευκολίας κατανόησης. Η Ευκλείδεια γεωμετρία περιγράφει γωνίες και σχήματα αντικειμένων. Ωστόσο, αποτυγχάνει να περιγράψει το σημείο τομής παράλληλων ευθειών. Συγκεκριμένα, δύο παράλληλες ευθείες τέμνονται στο *άπειρο*. Το *άπειρο*, όμως, δεν υπάρχει και αποτελεί απλά πλάσμα μιας βολικής φαντασίας. Είναι δυνατό να παρακαμφθεί το πρόβλημα αυτό επεκτείνοντας τον Ευκλείδειο χώρο προσθέτοντας τα σημεία στο *άπειρο*, όπου τέμνονται παράλληλες ευθείες. Τα σημεία αυτά καλούνται *ιδανικά σημεία*.

Η προσθήκη των *ιδανικών σημείων* μετασχηματίζει τον Ευκλείδειο χώρο στον *προβολικό χώρο*. Με άλλα λόγια, ο *προβολικός χώρος* είναι μια απλή επέκταση του Ευκλείδειου χώρου.

2.2 Συντεταγμένες

Ένα σημείο στον διδιάστατο Ευκλείδειο χώρο περιγράφεται από το διατεταγμένο ζεύγος πραγματικών αριθμών (x, y) . Η προσθήκη μιας επιπλέον συνιστώσας, δηλαδή $(x, y, 1)$, δηλώνεται να αντιπροσωπεύει το ίδιο σημείο στο χώρο. Μια τέτοια αλλαγή δεν οδηγεί σε σφάλματα, καθώς είναι εύκολο να οδηγηθεί κανείς από τη μία περιγραφή στην άλλη. Στη συνέχεια, η παραπάνω αλλαγή διευρύνεται ορίζοντας ως εναλλακτική περιγραφή του σημείου οποιαδήποτε τριάδα της μορφής (kx, ky, k) , όπου k οποιοσδήποτε μη μηδενικός πραγματικός αριθμός. Και πάλι, είναι δυνατή η μετάβαση από την μία περιγραφή στην άλλη καθώς ο λόγος των δύο πρώτων συνιστωσών προς την Τρίτη συνιστώσα της εναλλακτικής περιγραφής οδηγεί στο αρχικό σημείο (x, y) . Ουσιαστικά λοιπόν, σημεία περιγράφονται από *ισοδύναμες τάξεις* τριάδων, όπου δύο τριάδες είναι *ισοδύναμες* αν διαφέρουν κατά μία πολλαπλασιαστική σταθερά. Οι τριάδες που αποτελούν εναλλακτικούς περιγραφείς των σημείων ονομάζονται *ομογενείς συντεταγμένες*.

Επομένως σύμφωνα με τα παραπάνω, το σημείο $(x, y, 1)$ αντιπροσωπεύει το σημείο (x, y) . Ποιο σημείο όμως αντιπροσωπεύει η τριάδα $(x, y, 0)$; Το σημείο $(x, 0)$ δεν είναι πεπερασμένο. Αυτά τα σημεία αποτελούν τα *ιδανικά σημεία*, δηλαδή τα σημεία στο *άπειρο*. Δηλαδή, τα *ιδανικά σημεία* περιγράφονται από ομογενείς συντεταγμένες με μηδενική την τελευταία συνιστώσα.

Οι ίδιες ιδέες που μετασχηματίζουν το διδιάστατο Ευκλείδειο χώρο σε προβολικό χώρο με σημεία που περιγράφονται από ομογενείς συντεταγμένες, μπορούν να επεκταθούν σε οποιαδήποτε διάσταση. Ο Ευκλείδειος χώρος R^n μπορεί να επεκταθεί στον προβολικό χώρο P με περιγραφείς σημείων μέσω ομογενών διάνυσμάτων. Για $n = 2$ τα σημεία στο άπειρο σχηματίζουν μια ευθεία που καλείται *ευθεία στο άπειρο*, ενώ για $n = 3$ σχηματίζουν επίπεδο που καλείται *επίπεδο στο άπειρο*.

2.3 Προβολική Γεωμετρία και Μετασχηματισμοί στο 2Δ Χώρο

Στην ενότητα αυτή μελετώνται οι ιδιότητες της προβολικής γεωμετρίας στον 2Δ χώρο. Συγκεκριμένα, μελετώνται οι περιγραφείς των σημείων, ευθειών, καμπυλών σε ομογενείς συντεταγμένες, και πώς αυτές μετασχηματίζονται υπό προβολικούς μετασχηματισμούς. Εισάγονται οι έννοιες των *ιδανικών σημείων* και της *ευθείας στο άπειρο*. Στην ενότητα αυτή, έντονοι μικροί χαρακτήρες, όπως x , συμβολίζουν ένα διάνυσμα στήλη, και το ανάστροφό του, x^t , συμβολίζει ένα διάνυσμα γραμμή.

2.3.1 Το 2Δ Προβολικό Επίπεδο

Τα σημεία σε ένα επίπεδο περιγράφονται από το ζεύγος συντεταγμένων (x, y) στο \mathbb{R}^2 . Οι ομογενείς συντεταγμένες του σημείου είναι (kx, ky, k) στο \mathbb{P}^2 .

Ομογενείς περιγραφείς ευθειών. Η ευθεία σε ένα επίπεδο περιγράφεται από την εξίσωση $ax + by + c = 0$. Η επιλογή διαφορετικών τιμών για τις σταθερές a, b και c δίνει διαφορετικές ευθείες στο επίπεδο. Επομένως, μια ευθεία μπορεί να παρασταθεί ως ένα διάνυσμα $(a, b, c)^T$. Ωστόσο, το διάνυσμα αυτό δεν αποτελεί μοναδικό τρόπο περιγραφής της ευθείας αυτής, διότι η εξίσωση $(ka)x + (kb)y + (kc) = 0$ και άρα το διάνυσμα (ka, kb, kc) περιγράφουν την ίδια ευθεία για κάθε μη μηδενικό k . Τα διανύσματα αυτά αν και διαφορετικά είναι ισοδύναμα καθώς παριστάνουν την ίδια ευθεία. Μια *ισοδύναμη τάξη* διανυσμάτων είναι γνωστή ως ομογενές διάνυσμα. Οποιοδήποτε διάνυσμα $(a, b, c)^T$ είναι αντιπροσωπευτικό της ισοδύναμης τάξης.

Ομογενείς περιγραφείς σημείων. Ένα σημείο $\mathbf{x} = (x, y)^T$ ανήκει στην ευθεία $\mathbf{l} = (a, b, c)^T$ αν και μόνο αν $ax + by + c = 0$. Αυτό μπορεί να γραφεί και ως $(x, y, 1)(a, b, c)^T = (x, y, 1)\mathbf{l} = 0$. Να σημειωθεί ότι η για κάθε μη μηδενική σταθερά k και ευθεία \mathbf{l} η εξίσωση $(kx, ky, k)\mathbf{l} = 0$ ισχύει αν και μόνο αν $(x, y, 1)\mathbf{l} = 0$. Επομένως, το σύνολο διανυσμάτων $(kx, ky, k)^T$ αντιστοιχεί στο σημείο (x, y) στο \mathbb{R}^2 . Για το λόγο αυτό, και τα σημεία και οι ευθείς μπορούν να παρασταθούν με ομογενείς συντεταγμένες. Ένα αυθαίρετο ομογενές διάνυσμα της μορφής $\mathbf{x} = (x_1, x_2, x_3)^T$ στο \mathbb{P}^2 αντιπροσωπεύει το σημείο $(,)$ στο \mathbb{R}^2 .

Μια απλή εξίσωση καθορίζει αν ένα σημείο ανήκει στην ευθεία, συγκεκριμένα

Ένα σημείο \mathbf{x} ανήκει στην ευθεία \mathbf{l} εαν και μόνο εαν $\mathbf{x}^T \mathbf{l} = 0$

Τομή ευθειών. Για δύο ευθείες $\mathbf{l} = (a, b, c)^T$ και $\mathbf{l}' = (a', b', c')^T$ ορίζεται το διάνυσμα $\mathbf{x} = \mathbf{l} \chi \mathbf{l}'$, όπου χ αντιπροσωπεύει το εξωτερικό γινόμενο. Από την ιδιότητα του εξωτερικού γινομένου $\mathbf{l}^T (\mathbf{l} \chi \mathbf{l}') = \mathbf{l}'^T (\mathbf{l} \chi \mathbf{l}') = 0$ ισχύει ότι $\mathbf{l}^T \mathbf{x} = \mathbf{l}'^T \mathbf{x} = 0$. Επομένως, αν θεωρηθεί ότι \mathbf{x} αντιπροσωπεύει ένα σημείο, τότε το \mathbf{x} ανήκει και στις δύο ευθείες \mathbf{l} και \mathbf{l}' , και άρα αποτελεί το σημείο τομής των ευθειών.

Το σημείο τομής δύο ευθειών \mathbf{l}, \mathbf{l}' είναι το σημείο $\mathbf{x} = \mathbf{l} \chi \mathbf{l}'$.

Ευθεία ορισμένη από σημεία. Έστω δύο σημεία x και x' . Ορίζοντας το διάνυσμα $l = x \chi x'$ και θεωρώντας ότι παριστάνει ευθεία, εύκολα προκύπτει ότι αυτή διέρχεται και από τα δύο σημεία.

Η ευθεία που διέρχεται από τα σημεία x και x' είναι η $l = x \chi x'$

Ιδανικά Σημεία και η Ευθεία στο Άπειρο

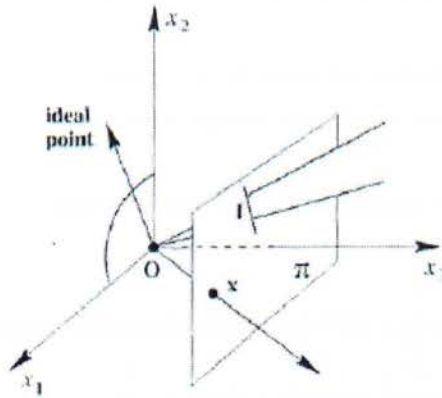
Σημείο τομής παράλληλων ευθειών. Έστω δύο παράλληλες ευθείες $ax + by + c = 0$ και $ax + by + c' = 0$. Αυτές περιγράφονται από τα διανύσματα $l = (a, b, c)^T$ και $l' = (a, b, c')^T$. Σύμφωνα με το Αποτέλεσμα 2.2 το σημείο τομής είναι $l \chi l' = (c' - c)(b, -a, 0)^T$. Αγνοώντας την πολλαπλασιαστική σταθερά, το σημείο τομής των δύο ευθειών είναι το $(b, -a, 0)^T$. Το σημείο αυτό είναι σημείου άπειρο v και δεν αντιστοιχεί σε πεπερασμένο σημείο στο R^2 .

Ευθεία στο άπειρο. Από τα παραπάνω προκύπτει ότι ένα ομογενές διάνυσμα $x = (x_1, x_2, x_3)^T$ με $x_3 = 0$ αντιστοιχεί σε πεπερασμένο σημείο στο R^2 . Τα σημεία με $x_3 = 0$ θεωρούνται σημεία στο άπειρο. Το σύνολο όλων των ιδανικών σημείων (δηλαδή σημείων στο άπειρο) μπορεί να γραφεί ως $(x_1, x_2, 0)^T$, όπου ένα συγκεκριμένο σημείο καθορίζεται από το λόγο $x_1 : x_2$. Είναι φανερό ότι αυτό το σύνολο σημείων ανήκει σε μια ευθεία, την *ευθεία στο άπειρο*, που περιγράφεται από το διάνυσμα $l = (0, 0, 1)^T$. Πράγματι, εύκολα επαληθεύει κανείς ότι $(0, 0, 1)(x_1, x_2, 0)^T = 0$.

Από το Αποτέλεσμα 2.2, το σημείο τομής της ευθείας $l = (a, b, c)^T$ και της ευθείας στο άπειρο είναι το ιδανικό σημείο $(b, -a, 0)^T$. Μια ευθεία $l' = (a, b, c')^T$ παράλληλη στην l , έχει το ίδιο σημείο τομής με την ευθεία στο άπειρο l . Σε μη ομογενείς συντεταγμένες, το διάνυσμα $(b, -a)^T$ είναι ένα διάνυσμα εφαπτόμενο στην ευθεία και ορθογώνιο στο κάθετο διάνυσμα της ευθείας (a, b) και έτσι αντιπροσωπεύει τη διεύθυνση της ευθείας. Για το σκοπό αυτό, η ευθεία στο άπειρο μπορεί να θεωρηθεί ως το σύνολο των διαφορετικών διευθύνσεων των ευθειών στο 2Δ χώρο.

Μοντελοποίηση του προβολικού επιπέδου. Για ευκολότερη κατανόηση του P^2 , μπορεί κανείς να το σκεφτεί ως ένα σύνολο ακτίνων στο R^3 . Το σύνολο όλων των διανυσμάτων της μορφής $k(x_1, x_2, x_3)^T$ με k μεταβαλλόμενο σχηματίζει μια ακτίνα που διέρχεται από την αρχή ($k = 0$). Μια τέτοια ακτίνα αντιπροσωπεύει ένα σημείο στο P^2 . Με το μοντέλο αυτό, μια ευθεία στο P^2 είναι ένα επίπεδο που διέρχεται από την αρχή. Τα σημεία και οι ευθείες στο P^2 μπορούν να προκύψουν από την τομή των ακτίνων και των επιπέδων που διέρχονται από την αρχή με το επίπεδο $x_3 = 1$. Οι ακτίνες που αντιστοιχούν στα ιδανικά σημεία και το επίπεδο που αντιστοιχεί στην είναι παράλληλα στο επίπεδο $x_3 = 1$.

Αναδικότητα. Είναι εμφανές ότι ο ρόλος των ευθειών και των σημείων εναλλάσσεται στις σχέσεις που αφορούν τις ιδιότητες τους. Τα Αποτελέσματα 2.2 και 2.4 είναι τα ίδια, με τους ρόλους των σημείων και ευθειών να εναλλάσσονται.



Σχήμα 2.1: Το μοντέλο του προβολικού επιπέδου. Τα σημεία και οι ευθείες του P^2 αντιστοιχούν σε ακτίνες και επίπεδα, αντίστοιχα, που διέρχονται από την αρχή στο R^3 . Ευθείες που ανήκουν στο x_1x_3 -επίπεδο αντιστοιχούν σε ιδανικά σημεία ενώ το x_1x_3 -επίπεδο αντιπροσωπεύει την T . Το σχήμα προέρχεται από το [22].

Κωνικές Τομές

Μια κωνική τομή είναι μια καμπύλη που περιγράφεται από μια εξίσωση δευτέρου βαθμού στο επίπεδο. Στην Ευκλείδεια γεωμετρία διακρίνονται τρία είδη κωνικών τομών, οι υπερβολοειδείς, οι ελλειψοειδείς και οι παραβολικές.

Η εξίσωση μια κωνικής τομής σε μη ομογενείς συντεταγμένες είναι

$$ax_2 + bx_1y + cy^2 + dx + ey + f = 0 \quad (2.1)$$

Ομογενοποιώντας τις συντεταγμένες με αντικατάσταση της συνιστώσας x με x_1 και της y με x_2 προκύπτει η εξίσωση

$$ax_1^2 + bx_1x_2 + cx_2^2 + dx_1x_3 + ex_2x_3 + fx_3^2 = 0 \quad (2.2)$$

ή σε μορφή πινάκων

$$x^T C x = 0 \quad (2.3)$$

όπου ο πίνακας C δίνεται από

$$C = \begin{bmatrix} a & b/2 & d/2 \\ b/2 & c & e/2 \\ d/2 & e/2 & f \end{bmatrix}$$

Ο πίνακας C είναι συμμετρικός & Ο λόφος των στο χείων του πίνακα είναι σημαντικός καθώς ο πολλαπλασιασμός του πίνακα με μια μη μηδενική σταθερά δεν επηρεάζει την εξίσωση. Επομένως, μια κωνική τομή έχει πέντε βαθμούς ελευθερίας (τα έξι στοιχεία ενός συμμετρικού πίνακα 3×3 πλην μια πολλαπλασιαστική σταθερά).

Ο προσδιορισμός των παραμέτρων του πίνακα C και επομένως της κωνικής τομής απαιτεί τον προσδιορισμό των πέντε στοιχείων και άρα απαιτεί την γνώση πέντε σημείων που ανήκουν στην καμπύλη. Με άλλα λόγια, πέντε σημεία καθορίζουν μια κωνική τομή.

Εφαπτόμενες στις κωνικές τομές. Η ευθεία I εφαπτόμενη σε μια κωνική τομή σε σημείο x δίνεται από το παρακάτω αποτέλεσμα

Η ευθεία I εφαπτόμενη στην C στο σημείο x δίνεται από τη σχέση $I = Cx$.

Πράγματι, η ευθεία αυτή διέρχεται από το σημείο x , διότι $I^T x = x^T C x = 0$. Αν η ευθεία I έχει ένα σημείο επαφής με την κωνική τομή, τότε είναι εφαπτόμενη. Διαφορετικά, αν η ευθεία έχει και δεύτερο σημείο τομής, έστω y , τότε $y^T C y = I^T y = 0$. Τότε όμως ισχύει ότι $(x + ay)^T C (x + ay) = 0$ για κάθε a . Επομένως, όλη η ευθεία $I = Cx$ που ενώνει τα σημεία x, y ανήκει στην κωνική τομή C . Για μη εκφυλισμένες κωνικές τομές (βλ. Παρακάτω), ένα τέτοιο συμπέρασμα δεν ισχύει.

Διαδικές κωνικές τομές. Ο ορισμός της κωνικής τομής παραπάνω βασίζεται στον καθορισμό της καμπύλης από σημεία. Από την ιδιότητα της *δυναδικότητας* μπορεί να οριστεί μια κωνική τομή που προκύπτει από εξισώσεις σε ευθείες. Αυτή η κωνική τομή περιγράφεται επίσης από ένα 3×3 πίνακα $C1$. Μια ευθεία I εφαπτόμενη στην κωνική τομή C ικανοποιεί τη σχέση $I^T C^* I = 0$. Ο συμβολισμός C^* δείχνει ότι ο πίνακας C^* είναι ο συζυγής πίνακας του C . Αν ο πίνακας συμμετρικός και αντιστρέψιμος, τότε ισχύει $C^* = C^{-1}$.

Στην περίπτωση που ο πίνακας C είναι πλήρης, από το Αποτέλεσμα 2.4 η εφαπτόμενη από ένα σημείο x που ανήκει στην κωνική τομή C ικανοποιεί τη σχέση $I = Cx$. Αντιστρέφοντας προκύπτει ότι $x = C^{-1}I$. Το σημείο x ανήκει στην κωνική τομή οπότε ισχύει ότι $x^T C x = 0$. Επομένως, προκύπτει ότι $(C^{-1}I)^T C (C^{-1}I) = I^T C^{-1} I = 0$

Εκφυλισμένες κωνικές τομές. Αν ο πίνακας C δεν είναι πλήρους βαθμού, τότε η κωνική τομή λέγεται εκφυλισμένη.

2.3.2 Προβολικοί Μετασχηματισμοί

Προβολικότητα είναι ένας αντιστρέψιμος μετασχηματισμός h από το P^2 στον εαυτό του τέτοιος ώστε τρία σημεία x_1, x_2 και x_3 ανήκουν στην ίδια ευθεία αν και μόνο αν $h(x_1), h(x_2)$ και $h(x_3)$ είναι συνευθειακά.

Η *προβολικότητα* ονομάζεται αλλιώς και προβολικός μετασχηματισμός ή ομογραφία.

Μια αντιστοίχιση $h : \mathbb{P}^2 \wedge \mathbb{P}^2$ είναι μια προβολικότητα αν και μόνο αν υπάρχει ένας αντιστρέψιμος 3×3 πίνακας H τέτοιος ώστε για κάθε σημείο στο \mathbb{P}^2 που περιγράφεται από ένα διάνυσμα \mathbf{x} να ισχύει ότι $h(\mathbf{x}) = H\mathbf{x}$.

Πράγματι, έστω τρία συνευθειακά σημεία $\mathbf{x}_1, \mathbf{x}_2$ και \mathbf{x}_3 ανήκουν στην ευθεία \mathbf{l} . Τότε $\mathbf{l}^T \mathbf{x}_i = 0$ για $i = 1, \dots, 3$. Έστω ότι ο H είναι ένας μη αντιστρέψιμος 3×3 πίνακας. Ισχύει ότι $\mathbf{l}^T H^{-1} H \mathbf{x}_i = 0$. Επομένως τα σημεία $H \mathbf{x}_i$ ανήκουν στην ευθεία $H^T \mathbf{l}$ και άρα είναι συνευθειακά. Τελικά, διατηρείται η ιδιότητα της συγγραμμικότητας και άρα πρόκειται για *προβολικότητα*.

Από τα παραπάνω, είναι δυνατό να δοθεί ένας εναλλακτικός ορισμός του προβολικού μετασχηματισμού.

Ένας επίπεδος προβολικός μετασχηματισμός είναι ένας γραμμικός μετασχηματισμός σε ομογενείς διανύσματα που περιγράφεται από ένα αντιστρέψιμο 3×3 πίνακα H , δηλαδή $\mathbf{x}' = H\mathbf{x}$,

όπου ο πίνακας H δίνεται από

$$H = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix}$$

Μετασχηματισμός ευθειών. Παραπάνω έγινε φανερό πως όταν ένα σημείο \mathbf{x} ανήκει στην ευθεία \mathbf{l} τότε το νέο σημείο $\mathbf{x}' = H\mathbf{x}$ ανήκει στην ευθεία $\mathbf{l}' = H^T \mathbf{l}$. Συμπερασματικά, υπό το μετασχηματισμό $\mathbf{x}' = H\mathbf{x}$, μια ευθεία μετασχηματίζεται σε

$$\mathbf{l}' = H^T \mathbf{l} \tag{2.4}$$

2.3.3 Ιεραρχία Μετασχηματισμών

Στην ενότητα αυτή περιγράφονται οι ειδικές περιπτώσεις του προβολικού μετασχηματισμού. Στον παρακάτω πίνακα φαίνονται οι περιπτώσεις αυτές.

Για κάθε ένα από τους μετασχηματισμούς δίνεται παρακάτω μια σύντομη περιγραφή των χαρακτηριστικών τους.

- **Προβολικός Μετασχηματισμός** Όπως παρατηρήθηκε και παραπάνω, ο μετασχηματισμός αυτός αφήνει αναλώιωτη την ιδιότητα της συγγραμμικότητας. Διαθέτει οκτώ βαθμούς ελευθερίας, καθώς είναι ομογενής μετασχηματισμός (πίνακας εννέα στοιχείων πλην μιας πολλαπλασιαστικής σταθεράς).
- **Αφινικός Μετασχηματισμός** Ο αφινικός μετασχηματισμός διατηρεί τις παράλληλες ευθείες, τον λόγο εμβαδών, το λόγο μηκών σε παράλληλες ευθείες και την ευθεία στο άπειρο. Διαθέτει έξι βαθμούς ελευθερίας (τέσσερα στοιχεία του πίνακα A και δύο στοιχεία του διανύσματος \mathbf{t}).
- **Ευκλείδειος Μετασχηματισμός** Ο ευκλείδειος μετασχηματισμός διατηρεί τα μήκη και τα εμβαδά. Διαθέτει τρεις βαθμούς ελευθερίας (ένα για την περιστροφή και δύο για την μετακίνηση).

2.4 Προβολική Γεωμετρία και Μετασχηματισμοί στο 3Δ Χώρο

Στην ενότητα αυτή μελετώνται οι ιδιότητες της προβολικής γεωμετρίας στον 3Δ κόσμο. Οι έννοιες που εισάγονται είναι επεκτάσεις των εννοιών που μελετήθηκαν στην προηγούμενη ενότητα. Στην ενότητα αυτή, με κεφαλαία έντονα γράμματα θα συμβολίζουμε ένα διάνυσμα διάστασης

4 χ 1.

2.4.1 Σημεία και Προβολικοί Μετασχηματισμοί

Ένα σημείο \mathbf{X} στο τρισδιάστατο χώρο περιγράφεται σε ομογενείς συντεταγμένες ως ένα διάνυσμα 4 χ 1. Συγκεκριμένα, το ομογενές διάνυσμα $\mathbf{X} = (X_1, X_2, X_3, X_4)^T$ με $X_4 = 0$ αντιστοιχεί στο σημείο (X, Y, Z) του \mathbb{R}^3 όπου

$$X = \frac{X_1}{X_4} \quad Y = \frac{X_2}{X_4} \quad Z = \frac{X_3}{X_4} \quad (2.5)$$

Ομογενή σημεία με $X_4 = 0$ αντιστοιχούν σε σημεία στο άπειρο.

Ένας προβολικός μετασχηματισμός στον \mathbb{P}^3 είναι ένας γραμμικός μετασχηματισμός που δρα σε ομογενούς διανύσματα 4 χ 1. Πρόκειται για έναν αντιστρέψιμο πίνακα H διάστασης 4 χ 4: $\mathbf{X}' = H\mathbf{X}$. Ο πίνακας H που αντιπροσωπεύει τον μετασχηματισμό σε ομογενείς συντεταγμένες έχει 15 βαθμούς ελευθερίας (16 στο \mathbb{P}^3 του πίνακα πλην μιας πολλαπλασιαστικής σταθεράς). Ο προβολικός μετασχηματισμός διατηρεί τις ευθείες, δηλαδή ευθείες μετασχηματίζονται σε ευθείες, καθώς και τα σημεία τομής ευθειών με επίπεδα.

Επίπεδα, Ευθείες και Προβολικοί Μετασχηματισμοί

Στον \mathbb{P}^3 τα σημεία και τα επίπεδα είναι δυαδικά και η περιγραφή τους είναι ανάλογη με την δυαδικότητα σημείου-ευθείας στον \mathbb{P}^2 .

Επίπεδα

Στον τρισδιάστατο ευκλείδειο χώρο ένα επίπεδο περιγράφεται από την εξίσωση

$$\pi_1 X + \pi_2 Y + \pi_3 Z + \pi_4 = 0$$

Ομογενοποιώντας τις συντεταγμένες ενός σημείου που ανήκει στο επίπεδο, δηλαδή $X^{\wedge}, Y^{\wedge}, Z^{\wedge}$ προκύπτει η εξίσωση

$$K_1 X_1 + \pi_2 X_2 + \pi_3 X_3 + K_4 X_4 = 0 \quad (2.6)$$

ή σε διανυσματική μορφή όπου

$$\pi = (\pi_1, \pi_2, \pi_3, \pi_4)^T.$$

$$\hat{X} = 0 \quad (2.7)$$

που εκφράζει ότι το σημείο X ανήκει στο επίπεδο π .

Οι πρώτες τρεις συνιστώσες του π αντιστοιχούν στο κάθετο διάνυσμα του επιπέδου στην ευκλείδεια γεωμετρία.

Τρία σημεία ορίζουν ένα επίπεδο. Έστω τρία σημεία X_i ανήκουν στο επίπεδο π . Τότε για σημείο ισχύει $\pi^T X_i$ κάθε $i = 1, \dots, 3$. Σε μορφή πινάκων οι εξισώσεις γίνονται

$$\begin{bmatrix} X_1^T \\ X_2^T \\ X_3^T \end{bmatrix} \pi = 0 \quad (2.8)$$

Αν τα τρία σημεία είναι μη συγγραμικά, ο πίνακας που σχηματίζεται από τα σημεία θα έχει βαθμό 3. Επομένως, το επίπεδο προκύπτει με ακρίβεια μιας πολλαπλασιαστικής σταθεράς από το δεξί null-space. Αν ο πίνακας έχει βαθμό 2, τότε και το null-space έχει διάσταση 2 και άρα τα σημεία συγγραμικά. Η λύση της εξίσωσης δίνει ένα σύνολο από επίπεδα που διέρχονται από τα σημεία αυτά.

Το επίπεδο στο άπειρο. Το επίπεδο στο άπειρο $\pi^\infty = (0, 0, 0, 1)^T$ περιλαμβάνει όλες τις διευθύνσεις $\mathbf{D} = (X_1, X_2, X_3, 0)^T$. Παρακάτω, περιγράφονται δύο ιδιότητες του επιπέδου αυτού

- Δύο επίπεδα είναι παράλληλα αν και μόνο αν η ευθεία τομής τους ανήκει στο π^∞
- Μια ευθεία είναι παράλληλη σε μία άλλη ευθεία, ή σε ένα επίπεδο, αν το σημείο τομής ανήκει στο π^∞

Το επίπεδο π^∞ είναι μια γεωμετρική αναπαράσταση για τους 3 βαθμούς ελευθερίας που χρειάζονται για να καθοριστούν οι αφινικές ιδιότητες. Το επίπεδο στο άπειρο είναι αναλοίωτο από αφινικούς μετασχηματισμούς.

Ευθείες

Μια ευθεία ορίζεται από την ένωση δύο σημείων ή την τομή δύο επιπέδων. Ο τρόπος περιγραφής τους είναι μέσω αυτών των δύο δυνατοτήτων ορισμού μιας ευθείας. Συγκεκριμένα, μπορούν να παρασταθούν μέσω *null-space and span representation*, *pluecker matrices* και *pluecker line coordinates*.

2.4.2 Ιεραρχία Μετασχηματισμών





Υπάρχει ένα πλήθος περιπτώσεων του προβολικού μετασχηματισμού στον 3D χώρο. Κάθε κατηγορία μετασχηματισμού έχει τις δικές της ιδιότητες. Στον παρακάτω πίνακα φαίνεται η ιεραρχία των μετασχηματισμών, όπου σε κάθε στήλη αναγράφονται μόνο οι επιπρόσθετες ιδιότητες καθώς ανέρχεται κανείς τις στήλες.

2.5 Μοντέλα Καμερών

Η κάμερα είναι μια απεικόνιση μεταξύ του τρισδιάστατου κόσμου (χώρος αντικειμένων) και της δισδιάστατης εικόνας. Στην ενότητα αυτή μελετώνται οι κάμερες που μοντελοποιούνται μέσω πινάκων με διάφορες ιδιότητες. Ενδιαφέρον θα δοθεί κυρίως στις κάμερες *κεντρικής προβολής*. Όλες οι κάμερες κεντρικής προβολής είναι ειδικές περιπτώσεις της *γενικής προβολικής κάμερας*. Για την ανάλυση της γενικότερης αυτής κάμερας χρησιμοποιούνται εργαλεία της προβολικής γεωμετρίας.

Οι ειδικές περιπτώσεις της *γενικής προβολικής κάμερας* κληρονομούν τις ιδιότητες της. Τα ειδικά αυτά μοντέλα κατηγοριοποιούνται σε δύο μεγάλες κλάσεις, αυτές που μοντελοποιούν την κάμερα με πεπερασμένο κέντρο και αυτές που μοντελοποιούν την κάμερα με κέντρο στο άπειρο

Στην ενότητα αυτή, με μικρά έντονα γράμματα θα περιγράφονται σημεία του P^2 ενώ με κεφαλαία έντονα γράμματα σημεία του R^3

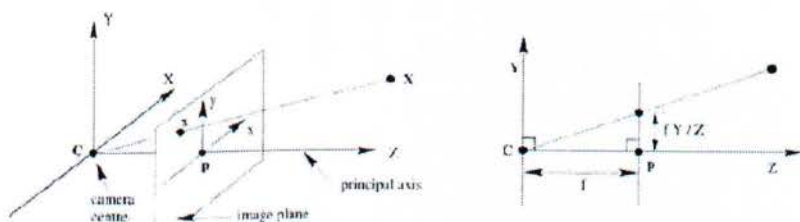
Group	Matrix	Distortion	Invariant properties
Projective 15 dof	$\begin{bmatrix} A & \mathbf{t} \\ \mathbf{v}^T & e \end{bmatrix}$		Intersection and tangency of surfaces in contact. Sign of Gaussian curvature.
Affine 12 dof	$\begin{bmatrix} A & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix}$		Parallelism of planes, volume ratios, centroids. The plane at infinity, π_∞ , (see section 3.5).
Similarity 7 dof	$\begin{bmatrix} sR & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix}$		The absolute conic, ω_∞ , (see section 3.6).
Euclidean 6 dof	$\begin{bmatrix} R & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix}$		Volume.

Γεωμετρικές ιδιότητες των ειδικών κατηγοριών 3Δ μετασχηματισμών. Ο πίνακας $A = [a_{ij}]$ είναι αντιστρέψιμος, ο πίνακας $R = [r_{ij}]$ είναι ένας πίνακας περιστροφής. Η Τρίτη στήλη δείχνει το αποτέλεσμα των μετασχηματισμών σε ένα κύβο. Οι μετασχηματισμοί που βρίσκονται πιο ψηλά στον πίνακα περιλαμβάνουν τους πιο χαμηλούς.

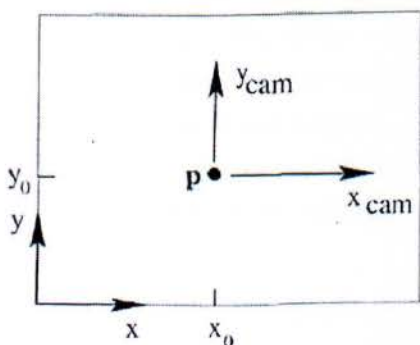
2.5.1 Πεπερασμένη Κάμερα

Στην υποενότητα αυτή μελετώνται τα πιο ειδικά και απλά μοντέλα όπως η pinhole κάμερα και περιγράφονται τα αναγκαία μαθηματικά μοντέλα.

Pinhole Κάμερα. Στο μοντέλο αυτό, τα σημεία του κόσμου προβάλλονται στο επίπεδο της εικόνας. Έστω ότι το κέντρο προβολής είναι η αρχή ενός ευκλείδειου συστήματος συντεταγμένων και έστω ότι το επίπεδο $Z = f$ είναι το επίπεδο της κάμερας. Σύμφωνα με το μοντέλο της pinhole κάμερας, ένα σημείο $\mathbf{X} = (X, Y, Z)^T$ στον 3Δ χώρο απεικονίζεται σε ένα σημείο στο επίπεδο της εικόνας. Το σημείο αυτό είναι το σημείο τομής της ευθείας που ενώνει το κέντρο προβολής με το σημείο του κόσμου \mathbf{X} με το επίπεδο της εικόνας. Τα παραπάνω φαίνονται στο Σχήμα 5



Σχήμα 5: Το μοντέλο της pinhole κάμερας. C είναι το κέντρο της κάμερας και p είναι το κύριο σημείο. Να σημειωθεί ότι το επίπεδο της κάμερας έχει τοποθετηθεί μπροστά από το κέντρο της κάμερας.



Σχήμα 6: Τα συστήματα συντεταγμένων της εικόνας (x, y) και της κάμερας (x_{cam}, y_{cam})

Εύκολα αποδεικνύεται ότι το σημείο $(X, Y, Z)^T$ απεικονίζεται στο σημείο

$$(fX, fY, fZ)^T \text{ στο}$$

επίπεδο της εικόνας. Με άλλα λόγια, η αντιστοιχία δίνεται από την παρακάτω σχέση

$$(X, Y, Z)^T \wedge (fX, fY, fZ)$$

όπου αντιστοιχίζεται σημείο του κόσμου (χώρος R^3) σε σημείο της εικόνας (χώρος R^2).

Η παραπάνω σχέση μπορεί να παρασταθεί με τη μορφή πινάκων. Θεωρώντας τις ομογενείς συντεταγμένες των σημείων προκύπτει ότι

$$\begin{bmatrix} fX \\ fY \\ Z \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

Αν $x = (fX, fY, Z)^T$ και $X = (X, Y, Z, 1)^T$ η παραπάνω σχέση γράφεται

$$x = P X$$

όπου $P = \text{diag}(f, f, 1) / [10]$ ο πίνακας για το μοντέλο της pinhole κάμερας.

Αντιστάθμιση του κύριου σημείου. Γενικεύοντας το μοντέλο της pinhole κάμερας, είναι δυνατό να οριστεί ένα νέο μοντέλο αντισταθμίζοντας τη θέση του κύριου σημείου. Το κύριο σημείο p ορίζεται ως το σημείο τομής του άξονα Z του τοπικού συστήματος συντεταγμένων της κάμερας με το επίπεδο της εικόνας. Στο μοντέλο της pinhole κάμερας, όπως φαίνεται και από το Σχήμα

2.2 το κύριο σημείο είναι και η αρχή του συστήματος z συντεταγμένων της εικόνας για το v P^2 . Είναι δυνατό ωστόσο να γίνει μετατόπιση της αρχής του συστήματος του επιπέδου της εικόνας με αποτέλεσμα να πρέπει να

$$(X, Y, Z)^T \rightarrow (f \frac{X}{Z} + p_x, f \frac{Y}{Z} + p_y)^T \quad (2.12)$$

όπου $(p_x, p_y)^T$ είναι οι συντεταγμένες του κύριου σημείου. Σε μορφή πινάκων η σχέση γίνεται

$$x = K [A0] X_c \quad (2.13)$$

αντισταθμιστεί η απόσταση αυτή στον πίνακα της κάμερας. Τα σημεία του 3D κόσμου

αντιστοιχίζονται στα σημεία της εικόνας σύμφωνα με τον κανόνα όπου ο πίνακας K ονομάζεται *πίνακας βαθμονόμησης* δίνεται από τη σχέση

$$K = \begin{bmatrix} f & 0 & p \\ 0 & f & p \\ 0 & 0 & 1 \end{bmatrix} \quad (2.14)$$

Να σημειωθεί ότι τα σημεία του 3Δ κόσμου με ομογενείς συντεταγμένες $\mathbf{X}_{cam} = (X, Y, Z, 1)^T$ αναφέρονται ως προς σύστημα συντεταγμένων προσαρτημένο στην κάμερα (η αρχή του συστήματος ταυτίζεται με το κέντρο της κάμερας) και με τον κύριο άξονα της κάμερας να έχει την διεύθυνση του άξονα Z. Το σύστημα αυτό καλείται *σύστημα συντεταγμένων της κάμερας*.

Περιστροφή και μετακίνηση της κάμερας. Στην προηγούμενη παράγραφο θεωρήθηκε ότι το *παγκόσμιο σύστημα συντεταγμένων* ταυτίζεται με το *σύστημα συντεταγμένων της κάμερας*. Στη γενική περίπτωση, ωστόσο, τα δύο συστήματα διαφέρουν κατά μια περιστροφή και μία μετατόπιση. Αν \mathbf{X} είναι ένα μη ομογενές διάνυσμα διάστασης 3×1 που περιγράφει ένα σημείο στον 3Δ κόσμο, ενώ \mathbf{X}_{cam} περιγράφει το ίδιο σημείο στο σύστημα συντεταγμένων της κάμερας, τότε μπορεί να γραφεί ότι $\mathbf{X}_{cam} = R(\mathbf{X} - \mathbf{C})$, όπου \mathbf{C} περιγράφει τις συντεταγμένες του κέντρου της κάμερας στο παγκόσμιο σύστημα συντεταγμένων και R είναι ένας 3×3 πίνακας περιστροφής που περιγράφει τον προσανατολισμό του συστήματος συντεταγμένων της κάμερας.

Από τη σχέση 2.13 προκύπτει ότι

$$\mathbf{x} = KR[\mathbf{I} \ \mathbf{C}]\mathbf{X} \quad (2.15)$$

όπου \mathbf{X} είναι σημεία στο παγκόσμιο σύστημα συντεταγμένων.

Είναι φανερό από την τελευταία σχέση ότι μια γενική pinhole κάμερα, $P = KR[\mathbf{I} \ \mathbf{C}]$ έχει 9 βαθμούς ελευθερίας, δηλαδή 3 για τον K , 3 για τον R και 3 για τον \mathbf{C} . Οι παράμετροι του πίνακα βαθμονόμησης K λέγονται *εσωτερικές παράμετροι* της κάμερας, ενώ οι παράμετροι των R και \mathbf{C} λέγονται *εξωτερικές παράμετροι*.

Ο πίνακας της κάμερας πολλές φορές γράφεται στην παρακάτω μορφή

$$P = K [R|\mathbf{t}] \quad (2.16)$$

όπου $\mathbf{t} = -\mathbf{RC}$.

2.5.2 Πεπερασμένη προβολική κάμερα

Στο προηγούμενο μοντέλο θεωρήθηκε ότι οι συντεταγμένες της εικόνας είναι ευκλείδειες με ίση κλίμακα και στις δύο διευθύνσεις των αξόνων. Γενικεύοντας το μοντέλο, υπάρχει η πιθανότητα τα pixels να μην είναι τετράγωνα. Αν, λοιπόν, οι συντεταγμένες μετρώνται σε pixels τότε είναι αναγκαίο να εισαχθούν άνισοι παράγοντες κλίμακας σε κάθε κατεύθυνση. Συγκεκριμένα, αν ο αριθμός των pixels ανά μονάδα απόστασης σε συντεταγμένες εικόνας είναι m_x και m_y στον x- και y-άξονα, τότε η μετατροπή από παγκόσμιες συντεταγμένες σε συντεταγμένες pixel

προκύπτει πολλαπλασιάζοντας τον πίνακα βαθμονόμησης (σχέση) με τον πίνακα $diag(m_x, m_y, 1)$. Επομένως, ο πίνακας βαθμονόμησης γίνεται

$$K = \begin{bmatrix} a & 0 & x_0 \\ 0 & a & y_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (2.17)$$

όπου $a_x = fm_x$ και $alpha_y = fm_y$ αντιπροσωπεύουν το εστιακό μήκος της κάμερας σε διαστάσεις pixel στην x και y-διεύθυνση. Το διάνυσμα $\chi_0 = (x_0, y_0)$ περιγράφει το κύριο σημείο σε pixel συντεταγμένες, δηλαδή $x_0 = m_x p_x$ και $y_0 = m_y p_y$.

Γενικεύοντας ακόμα περισσότερο, ο πίνακας βαθμονόμησης μπορεί να έχει τη μορφή

$$K = \begin{bmatrix} \alpha_x & s & x_0 \\ 0 & \alpha_y & y_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (2.18)$$

Η επιπλέον παράμετρος s αναφέρεται ως *παράμετρος λοξότητας*. Για τις περισσότερες κάμερες λαμβάνει μηδενική τιμή. Η κάμερα με

$$P = KR[I \setminus - C] \quad (2.19)$$

όπου ο πίνακας βαθμονόμησης δίνεται από τη σχέση 2.18 ονομάζεται *πεπερασμένη προβολική κάμερα*. Μια τέτοια κάμερα έχει 11 βαθμούς ελευθερίας.

Ιδιότητες Προβολικής Κάμερας

2.3.2.1 Ανατομία Κάμερας

Μια γενικευμένη προβολική κάμερα $P = KR[I \setminus - C]$ μπορεί να αναλυθεί σε μπλοκ σύμφωνα με $P = [M \setminus p_4]$ όπου ο πίνακας M είναι 3×3 .

Κέντρο κάμερας. Ο πίνακας μια προβολικής κάμερας P , διάστασης 3×4 , έχει βαθμό 3. Για το λόγο αυτό ο δεξιός μηδενικός υπόχωρος του έχει διάσταση 1. Έστω ότι ο μηδενικός υπόχωρος προκύπτει από το διάνυσμα C έτσι ώστε $PC = 0$. Είναι εύκολο να δειχθεί (βλ. [22] p.158) ότι το διάνυσμα αυτό περιγράφει το κέντρο της κάμερας σε ομογενείς συντεταγμένες.

Στην περίπτωση των πεπερασμένων καμερών το αποτέλεσμα μπορεί να επαληθευτεί άμεσα, διότι $C = (C^T, 1)^T$ για το οποίο ισχύει $PC = 0$. Το ίδιο ισχύει και για την περίπτωση που ο πίνακας M είναι μη αντιστρέψιμος. Στην περίπτωση αυτή, ένα διάνυσμα που γεννά τον μηδενικό υπόχωρο είναι το $C = (d^T, 0)^T$ με $Md = 0$. Το κέντρο της κάμερας είναι ένα σημείο στο άπειρο (οι κάμερες αυτές ανήκουν στην κατηγορία των *άπειρων καμερών*).

Διανύσματα στήλες. Οι στήλες της προβολικής κάμερας είναι τρισδιάστατα διανύσματα με γεωμετρική ερμηνεία. Συμβολίζοντας με $\mathbf{p}_i, i = 1, \dots, 4$ τις στήλες του πίνακα P εύκολα προκύπτει ότι τα $\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3$ είναι η απεικόνιση των τριών διευθύνσεων του παγκόσμιου συστήματος συντεταγμένων. Για παράδειγμα, ο X -άξονας έχει διεύθυνση $\mathbf{D} = (1, 0, 0, 0)^T$, που απεικονίζεται στο σημείο $\mathbf{p}_1 = P\mathbf{D}$. Τα παραπάνω φαίνονται στο Σχήμα 2.4 Το διάνυσμα \mathbf{p}_4 είναι η απεικόνιση της αρχής των αξόνων.

Διανύσματα γραμμές. Τα διανύσματα γραμμές του πίνακα P είναι διάστασης 4×1 και μπορούν να ερμηνευθούν ως επίπεδα τα οποία θα μελετηθούν παρακάτω. Οι γραμμές του πίνακα συμβολίζονται με $\mathbf{r}_i^T, i = 1, \dots, 3$.

- **Κύριο επίπεδο** Το κύριο επίπεδο είναι το επίπεδο που διέρχεται από το κέντρο της κάμερας και είναι παράλληλο στο επίπεδο της εικόνας. Αποτελείται από το σύνολο σημείων που προβάλλονται στην ευθεία στο άπειρο της εικόνας. Συγκεκριμένα, $P\mathbf{X} = (x, y, 0)^T$. Δηλαδή, ένα σημείο ανήκει στο κύριο επίπεδο της κάμερας αν και μόνο αν $\mathbf{r}_3^T \mathbf{X} = 0$. Με άλλα λόγια, \mathbf{r}_3 είναι το διάνυσμα που περιγράφει το κύριο επίπεδο της κάμερας.
- **Το κύριο σημείο.** Ο κύριος άξονας είναι η ευθεία που διέρχεται από το κέντρο της κάμερας C με διεύθυνση κάθετη στο κύριο επίπεδο \mathbf{r}_3 . Ο άξονας τέμνει το επίπεδο της εικόνας στο κύριο σημείο. Ο προσδιορισμός του σημείου αυτού γίνεται ως εξής: Το κάθετο διάνυσμα ενός επιπέδου $\pi = (\pi_1, \pi_2, \pi_3, \pi_4)^T$ δίνεται από το διάνυσμα $(\pi_1, \pi_2, \pi_3)^T$. Το διάνυσμα αυτό μπορεί να περιγραφεί και από το σημείο $(\pi_1, \pi_2, \pi_3, 0)^T$ του επιπέδου στο άπειρο. Στην περίπτωση του κύριου επιπέδου της κάμερας το σημείο αυτό είναι το $\mathbf{X}_3 = (p_{31}, p_{32}, p_{33}, 0)^T$. Προβάλλοντας το σημείο στην εικόνα προκύπτει το κύριο σημείο, δηλαδή $\mathbf{p} = P\mathbf{X}_3$.
- **Διάνυσμα του κύριου άξονα.** Αν και κάθε σημείο \mathbf{X} που δεν ανήκει στο κύριο επίπεδο μπορεί να αντιστοίχη σε σημείο της εικόνας σύμφωνα με $\mathbf{x} = P\mathbf{X}$, στην πραγματικότητα μόνο τα σημεία στον χώρο που βρίσκονται μπροστά από την κάμερα μπορούν να απεικονιστούν. Έστω ο προβολικός πίνακας $P = [M|\mathbf{p}_4]$. Παραπάνω, έγινε φανερό ότι το διάνυσμα \mathbf{m}_3 (τελευταία γραμμή του M) δείχνει προς την διεύθυνση του κύριου άξονα. Ωστόσο, δεν είναι γνωστό αν το διάνυσμα αυτό δείχνει προς την θετική κατεύθυνση (μπροστά από την κάμερα). Συγκεκριμένα, ο πίνακας P έχει οριστεί με αβεβαιότητα προσήμου. Για το λόγο αυτό δεν μπορεί να καθοριστεί αν το \mathbf{m}_3 ή το $-\mathbf{m}_3$ δείχνει προς τα θετικά.

Αράση της Προβολικής Κάμερας στα Σημεία

Εμπρόσθια προβολή. Όπως έχει αναφερθεί παραπάνω, μια γενική προβολική κάμερα προβάλλει ένα σημείο του χώρου \mathbf{X} σε ένα σημείο στην εικόνα σύμφωνα με τον τύπο $\mathbf{x} = P\mathbf{X}$. Τα σημεία στο άπειρο της μορφής $\mathbf{D} = (\mathbf{d}, 0)^T$ προβάλλονται στα σημεία

$$\mathbf{x} = P\mathbf{D} = [M|\mathbf{p}_4]\mathbf{D} = M\mathbf{d} \quad (2.20)$$

Οπίσθια προβολή σημείων σε ακτίνες. Δεδομένου ενός σημείου \mathbf{x} στην εικόνα, είναι επιθυμητό να προσδιοριστεί ένα σύνολο σημείων στον χώρο που προβάλλονται στο σημείο αυτό. Τα σημεία αυτά αποτελούν μια ακτίνα στο χώρο που διέρχεται από το κέντρο της κάμερας.

Για να οριστεί επαρκώς η ακτίνα χρειάζονται δύο σημεία. Ένα σημείο είναι το κέντρο της κάμερας \mathbf{C} ($\mathbf{PC} = \mathbf{0}$). Ένα δεύτερο σημείο είναι το σημείο $P + \mathbf{x}$ όπου $P +$ είναι ο ψευδο-αντίστροφος του P . Το σημείο αυτό ανήκει στην ακτίνα καθώς η προβολή του στην εικόνα είναι το σημείο \mathbf{x} . Η ακτίνα δίνεται λοιπόν από την εξίσωση

$$\mathbf{X}(\lambda) = P + \mathbf{x} + \lambda \mathbf{C} \quad (2.21)$$

Βάθος Σημείων

Στη συνέχεια, μελετάται η απόσταση των σημείων μπροστά ή πίσω από το κύριο επίπεδο της κάμερας. Έστω ο πίνακας της κάμερας $P = [M | \mathbf{p}_4]$, που προβάλλει το σημείο $\mathbf{X} = (X, Y, Z, 1)^T = (\mathbf{X}, 1)^T$ του κόσμου στο σημείο της εικόνας $\mathbf{x} = w(x, y, 1)^T = P\mathbf{X}$. Όπως αναφέρθηκε παραπάνω, ο οπτικός άξονας της κάμερας έχει διεύθυνση \mathbf{m}_3 (το διάνυσμα της τρίτης γραμμής του πίνακα M). Αν $\mathbf{C} = (\mathbf{C}, 1)^T$ είναι το κέντρο της κάμερας, τότε ισχύει $w = \mathbf{r}^T \mathbf{X} = \mathbf{r}^T (\mathbf{X} - \mathbf{C})$ αφού $\mathbf{PC} = \mathbf{0}$. Ωστόσο, $\mathbf{r}^T (\mathbf{X} - \mathbf{C}) = \mathbf{m}^T (\mathbf{X} - \mathbf{C})$. Επομένως, το $w = \mathbf{m}^T (\mathbf{X} - \mathbf{C})$ μπορεί να ερμηνευθεί ως η προβολή του τμήματος που ενώνει το κέντρο της κάμερας και το σημείο στο 3D χώρο πάνω στον οπτικό άξονα της κάμερας. Αν ο πίνακας M είναι κανονικοποιημένος έτσι ώστε $\det(M) > 0$ και $\|\mathbf{m}_3\| = 1$, τότε το διάνυσμα \mathbf{m}_3 είναι μοναδιαίο και δείχνει προς τα θετικά του οπτικού άξονα. Στην περίπτωση αυτή, η ποσότητα w είναι το βάθος του σημείου \mathbf{X} από το κέντρο της κάμερας \mathbf{C} στη θετική κατεύθυνση του.

Αν ο πίνακας M δεν έχει κανονικοποιηθεί τότε το βάθος προκύπτει από το παρακάτω αποτέλεσμα

Αποτέλεσμα 2.5. Έστω $\mathbf{X} = (X, Y, Z, T)^T$ είναι ένα 3D σημείο και $P = [M | \mathbf{p}_4]$ είναι ο προβολικός πίνακας μια κάμερας. Αν ισχύει $P(X, Y, Z, T)^T = w(x, y, 1)$ τότε

$$depth(\mathbf{X}; P) = \frac{sign(det(M))w}{T\|\mathbf{m}_3\|} \quad (2.22)$$

είναι το βάθος του σημείο \mathbf{X} μπροστά από το κύριο επίπεδο.

2.5.3 Η Κάμερα στο Απειρο

Στην κατηγορία αυτή, το κέντρο των καμερών βρίσκεται στο επίπεδο του απείρου. Αυτό σημαίνει ότι το αριστερό 3×3 μέρος του πίνακα P είναι μη αντιστρέψιμο. Το κέντρο της κάμερας μπορεί να βρεθεί από τη σχέση $\mathbf{PC} = \mathbf{0}$ όπως και στις πεπερασμένες κάμερες.

Οι κάμερες στο άπειρο μπο ρ ύ ν να διαχωριστούν σε δύο κατηγορίες, τις αφινικές και τις μη αφινικές. Συγκεκριμένα, οι αφινικές κάμερες είναι αυτές των οποίων η τελευταία σειρά του πίνακα P είναι της μορφής $(0, 0, 1)T$. Με άλλα λόγια, οι αφινικές κάμερες αντιστοιχίζουν σημεία στο άπειρο σε σημεία στο άπειρο.

ΚΕΦΑΛΑΙΟ 3

Εφαρμογή της όρασης υπολογιστών σε αυτόνομο όχημα

3.1 Γιατί αντίληψη και δράση;

Για τα τεχνικά συστήματα που προορίζονται να βρουν τον δρόμο τους από μόνα τους σε έναν συνεχώς μεταβαλλόμενο κόσμο, είναι αδύνατο να προβλεφθεί κάθε πιθανό γεγονός και να προγραμματιστούν όλες οι απαραίτητες ικανότητες για τις κατάλληλες αντιδράσεις στο λογισμικό του από την αρχή. Για να είναι εύκαμπτες όσον αφορά στις καταστάσεις που αντιμετωπίζονται πραγματικά, το σύστημα πρέπει να έχει τις αντιληπτικές και συμπεριφοριστικές ικανότητες που μπορεί να επεκτείνει από μόνο του σε απάντηση στις νέες απαιτήσεις. Αυτό σημαίνει ότι το σύστημα πρέπει να είναι ικανό να κρίνει την αξία των αποτελεσμάτων ελέγχου σε απάντηση στα μετρούμενα στοιχεία, εντούτοις, δεδομένου ότι τα αποτελέσματα του ελέγχου έχουν επιπτώσεις στις κρατικές μεταβλητές κατά τη διάρκεια ενός συγκεκριμένου ποσού του χρόνου, επόμενος στο ιστορικό του χρόνου πρέπει να παρατηρηθεί και μια χρονικά βαθύτερη κατανόηση πρέπει να αναπτυχθεί. Αυτό είναι ακριβώς τι συλλαμβάνεται στα «δυναμικά πρότυπα» της θεωρίας συστημάτων (και ποια βιολογικά συστήματα μπορούν να αποθηκεύσουν στις νευρώνες γραμμές καθυστέρησης).

Επίσης, μέσω αυτών των χρονικών ιστοριών, το έδαφος προετοιμάζεται για συμπαγέστερο «πεδίου συχνότητας» (ολοκληρωμένες) αναπαραστάσεις. Στο μεγάλο όγκο της βιβλιογραφίας στη γραμμική *θεωρία συστημάτων*, τις χρονικές σταθερές T ως αντίστροφο γεννήτορα των πρωτοκλασάτων τμημάτων συστημάτων, καθώς επίσης και τη συχνότητα, την αναλογία απόσβεσης, και τη σχετική φάση όπως οι χαρακτηριστικές ιδιότητες των δευτεροκλασάτων συστατικών είναι καλά γνωστοί όροι για την περιγραφή των χρονικών χαρακτηριστικών των διαδικασιών, π.χ., [Kailath 1980]. Στη φυσιολογική λογοτεχνία, ο όρος «χρονική τέλεια μορφή» μπορεί ακόμη και να βρεθεί [Ruhnau 1994a, β], δείχνοντας ότι η χρονική μορφή μπορεί να είναι σημαντική και χαρακτηρισμένη ως γνωστή χωρική μορφή.

Συνήθως, ο έλεγχος θεωρείται αποτέλεσμα της ανάλυσης δεδομένων για να επιτευχτεί κάποιος στόχος. Σε ένα σύστημα κλειστών βρόγχων, όπου ένας από τους στόχους του είναι να προσαρμοστεί στις νέες καταστάσεις αυτόνομα, τα αποτελέσματα έλεγχου μπορούν να ερμηνευτούν ως

ερωτήσεις όσον αφορά την πραγματική συμπεριφορά. Οι δυναμικές αντιδράσεις ερμηνεύονται τώρα για να καταλάβουν καλύτερα τη συμπεριφορά ενός σώματος στα διάφορα κράτη και υπό τις διάφορες περιβαλλοντικές συνθήκες. Αυτό ανοίγει μια νέα λεωφόρο για την ερμηνεία σημάτων: εκτός από τη χρήση της για το κρατικό έλεγχο, αυτό τώρα επίσης ερμηνεύεται για τον προσδιορισμό συστημάτων και την εκμάθηση διαμόρφωσης, δηλ., για τα χρονικά συμπεριφοριστικά χαρακτηριστικά της.

Σε ένα ευφύες αυτόνομο σύστημα, αυτή η ικανότητα της προσαρμογής στις νέες καταστάσεις πρέπει να είναι διαθέσιμη για να μειώσει την εξάρτηση στη συντήρηση και την προσαρμογή κοντά ανθρώπινη επέμβαση. Ενώ αυτό δεν είναι ακόμα εφικτό στα παρόντα συστήματα, με τη δύναμη υπολογισμού που διατίθεται στο μέλλον, είναι σαφώς μέσα στο πρόγραμμα. Οι μέθοδοι που απαιτούνται έχουν αναπτυχθεί στους τομείς του προσδιορισμού συστημάτων και του προσαρμοστικού ελέγχου.

Η αίσθηση της όρασης πρέπει να παράγει ικανοποιητικές πληροφορίες για τον κοντινό και μακρύτερο περιβάλλον για να αποφασίσει πότε ο κρατικός έλεγχος δεν είναι τόσο σημαντικός και πότε περισσότερη έμφαση μπορεί να δοθεί στον προσδιορισμό συστημάτων με το να χρησιμοποιηθούν πρόσθετες εισαγωγές ελέγχου για αυτόν το λόγο. Αυτή η προσέγγιση επίσης θα διαδραματίσει έναν ρόλο για τον καθορισμό της έννοιας «αυτόνομο» για το αυτόνομο όχημα.

3.2 Γιατί αντίληψη και όχι μόνο όραση ;

Η λέξη όραση δεν επιτρέπει μια καλά θεμελιωμένη έκφραση σχετικά με την απόλυτη αδρανή κίνηση όταν κινείται ένα άλλο αντικείμενο κοντά στο όχημα και κανένα υπόβαθρο δεν μπορεί να δει στον τομέα αυτής της άποψης (που είναι γνωστό να είναι στάσιμο). Οι αδρανείς αισθητήρες όπως τα επιταχύμετρα και οι γωνιακοί αισθητήρες παράγουν τα αντίστοιχα σήματα για το σώμα που τοποθετούνται το κάνουν αυτό ουσιαστικά χωρίς οποιοδήποτε χρόνο καθυστέρησης και στα υψηλά ποσοστά σημάτων (μέχρι τη κλίμακα kHz).

Στην εικόνα χρειάζεται χρόνος για την ενσωμάτωση της ελαφριάς έντασης στα στοιχεία αισθητήρων (33 1/3, αντίστοιχα, 40ms που αντιστοιχούν στα ευρωπαϊκά πρότυπα ή των Ηνωμένων Πολιτειών), για την εγκλωβισμό πλαισίων και την επικοινωνία των (τεράστιο ποσό) στοιχείων της εικόνας, καθώς επίσης και για την εξαγωγή χαρακτηριστικών γνωρισμάτων, την παραγωγή υπόθεσης, και την κραταιά εκτίμηση. Συνήθως, τρεις έως πέντε τηλεοπτικοί κύκλοι, που είναι 100 έως 200 ms, θα έχουν περάσει μέχρι μια παραγωγή ελέγχου προερχόμενη από την εικόνα θα

επιστρέψει στο πραγματικό κόσμο. Για τον ακριβή έλεγχο των ιδιαίτερα δυναμικών συστημάτων, αυτή τη φορά η καθυστέρηση πρέπει να ληφθεί υπόψη.

Δεδομένου ότι οι διαταραχές πρέπει να εξισορροπηθούν το συντομότερο δυνατόν, και δεδομένου ότι οπτικά τα μετρήσιμα αποτελέσματα των διαταραχών είναι το δεύτερο ολοκλήρωμα των επιταχύνσεων με τους αντίστοιχους χρόνους καθυστέρησης, είναι ενδεδειγμένο να υπάρξουν οι αδρανείς αισθητήρες στο σύστημα για την πρόωρη επανάληψη των διαταραχών. Επειδή η μακροπρόθεσμη σταθεροποίηση μπορεί να είναι πραγματοποιήσιμη χρησιμοποιώντας την εικόνα, δεν είναι απαραίτητο να προσφύγει στους ακριβούς αδρανείς αισθητήρες αντίθετα, όταν χρησιμοποιούνται από κοινού με την εικόνα, οι ανέξοδοι αδρανείς αισθητήρες με τις καλές ιδιότητες για το μέσο στο μέρος υψηλής συχνότητας είναι ικανοποιητικοί όπως καταδεικνύονται από τα προθαλαμιαία συστήματα στα σπονδυλωτά.

Είναι σε θέση τα επιταχύμετρα να μετρήσουν άμεσα τα αποτελέσματα των περισσότερων αποτελεσμάτων ελέγχου; αυτό ανακουφίζει τον προσδιορισμό συστημάτων και εύρεση των αποτελεσμάτων ελέγχου για την αντανάκλαση όπως την αντίδραση των διαταραχών. Συσχέτιση των αδρανών σημάτων με τα οπτικά καθορισμένα σήματα επιτρέπουν χρονικά τη βαθύτερη κατανόηση αυτό που στις φυσικές επιστήμες καλείται «χρονικά ολοκληρώματα» των λειτουργιών εισαγωγής.

Για όλους αυτούς τους λόγους, η κοινή χρήση των οπτικών και αδρανών σημάτων θεωρείται υποχρεωτικός για τις αποδοτικές αυτόνομα κινητές πλατφόρμες επίτευξης. Ομοίως, εάν τα πρόσθετα τμήματα ταχύτητας μπορούν να μετρηθούν ευκολότερα από τις συμβατικές συσκευές, δεν έχει νόημα η προσπάθεια να τα ανακτήσουμε από την εικόνα σε μια «καθαρή» προσέγγιση. Αυτά τα συμβατικά σήματα μπορούν να ανακουφίσουν αρκετά την αντίληψη για το περιβάλλον δεδομένου ότι οι αντίστοιχοι αισθητήρες τοποθετούνται επάνω στο σώμα με έναν σταθερό τρόπο, ενώ στην εικόνα οι μετρημένες τιμές χαρακτηριστικών γνωρισμάτων πρέπει να οριστούν με κάποιο αντικείμενο στο περιβάλλον σύμφωνα ακριβώς με τα οπτικά στοιχεία. Δεν υπάρχει καμία καθιερωμένη σύνδεση για κάθε αξία μέτρησης στην εικόνα όπως συμβαίνει για τους συμβατικούς αισθητήρες.

3.3 Κατάλληλος χώρος ερμηνείας

Οι εικόνες είναι διδιάστατες σειρές στοιχείων το συνηθισμένο μέγεθος σειράς είναι σήμερα από περίπου 64×64 για τα πρόσθετα τσιπ «όρασης» σε περίπου 770 580 για την βιντεοκάμερα (τα πρόσθετα μεγαλύτερα μεγέθη είναι διαθέσιμα αλλά με πολύ υψηλότερο κόστος, π.χ. για τις διαστημικές ή στρατιωτικές εφαρμογές). Ένα ψηφιοποιημένο σήμα τηλεοπτικών στοιχείων είναι

μια γρήγορη ακολουθία αυτών των εικόνων με τα ποσοστά στοιχείων μέχρι ~ 11 MB/s για ασπρόμαυρο και μέχρι τρεις φορές αυτό το ποσό για τα έγχρωμα.

Συχνά, μόνο οι τομείς 320 240pixels (είτε οι μόνο περιέργες είτε ομαλές γραμμές με την αντίστοιχη μείωση του ψηφίσματος μέσα στις γραμμές) αξιολογούνται λόγω έλλειψης δύναμης υπολογισμού. Αυτό οδηγεί σε ένα όγκο στοιχείων ανά κάμερα περίπου 2 MB/s. Ακόμη και σε αυτό το μειωμένο ποσοστό στοιχείων, η δύναμη επεξεργασίας ενός ενιαίου μικροεπεξεργαστή διαθέσιμου σήμερα δεν είναι ικανοποιητική για την ερμηνεία διάφορων τηλεοπτικών σημάτων σε πραγματικό χρόνο. Τα υψηλής ευκρίνειας σήματα TV του μέλλοντος μπορούν να έχουν έως και 1080 γραμμές και 1920 pixels σε κάθε γραμμή με συχνότητες μέχρι 75 Hz αυτό αντιστοιχεί σε ποσοστά στοιχείων μηχανικής όρασης περισσότερων από 155 MB/s. Η μηχανική όραση με τέτοια ανάλυση είναι πολύ μπροστά στο μέλλον.

Ίσως, η ομοιόμορφη επεξεργασία ολόκληρων εικόνων δεν είναι καθόλου επιθυμητή, δεδομένου ότι τα διαφορετικά αντικείμενα θα βρεθούν σε διαφορετικά μέρη των εικόνων, που απαιτούν τη συγκεκριμένη επεξεργασία για την αποδοτικότερη αξιολόγηση. Πολύ συχνά, οι γραμμές ασυνέχειας απατώνται στις εικόνες, οι οποίες πρέπει να αντιμετωπιστούν με τις πρόσθετες μεθόδους που διαφέρουν ουσιαστικά από εκείνους που χρησιμοποιούνται στα ομοιογενή μέρη. Το αντικείμενο και οι ανά περίπτωση εξαρτώμενοι μέθοδοι και οι παράμετροι πρέπει να χρησιμοποιηθούν, ελεγχόμενος από υψηλότερα επίπεδα αξιολόγησης.

Έτσι η ερώτηση είναι, εάν οποιαδήποτε βασική εξαγωγή χαρακτηριστικών γνωρισμάτων πρέπει να εφαρμοστεί ομοιόμορφα πέρα από ολόκληρη την περιοχή της εικόνας. Στα βιολογικά συστήματα, αυτό φαίνεται να συμβαίνει, παραδείγματος χάριν, στο ραβδωτό φλοιό (V1) των σπονδυλωτών όπου τα προσανατολισμένα στοιχεία ακρών ανιχνεύονται με τη βοήθεια των αντίστοιχων δεκτικών πεδίων. Εντούτοις, η όραση των σπονδυλωτών έχει το ανομοιογενές ψήφισμα πέρα από το ολόκληρο οπτικό πεδίο. Η όραση με το υψηλό ψήφισμα στο κέντρο του αμφιβληστροειδή περιβάλλεται από δεκτικά πεδία που έχουν χαμηλότερη πυκνότητα των δεκτών ανά μονάδα της περιοχής στην ακτινωτή κατεύθυνση.

Η όραση που έχουν αναπτύξει τα βιολογικά συστήματα φαίνεται να υποβάλλει τρεις ερωτήσεις κάθε μια από τις οποίες αντιμετωπίζεται από ένα συγκεκριμένο υποσύστημα:

1. Υπάρχει κάτι του ειδικού ενδιαφέροντος στο ευρύτερο οπτικό πεδίο; Τι ακριβώς είναι αυτό, εκείνο που προκάλεσε το ενδιαφέρον στην ερώτηση 1;
2. Μπορεί το μεμονωμένο αντικείμενο να χαρακτηριστεί και να ταξινομηθεί χρησιμοποιώντας τη γνώση υποβάθρου; Ποια είναι η σχετικότητα της δήλωσης «εδώ και τώρα»;
3. Ποια είναι η κατάσταση γύρω από μένα και πώς επηρεάζει στις βέλτιστες αποφάσεις στη συμπεριφορά για την επίτευξη των στόχων μου; Για αυτόν το λόγο, μια σχετική συλλογή

αντικειμένων πρέπει να αναγνωριστεί, και η πιθανή μελλοντική συμπεριφορά πρέπει να προβλεφθεί.

Για να σταθεροποιήσουμε τη διαδικασία όρασης στην αρχή και για να ανιχνεύσει τα νέα αντικείμενα αργότερα, είναι βεβαίως ένα πλεονέκτημα για να έχει ένα από κάτω προς τα επάνω τμήμα ανίχνευσης διαθέσιμο σε όλο το εύρος του οπτικού πεδίου. Ίσως, μερικοί αλγόριθμοι βασισμένοι στην συγκεκριμένη πορεία για την ανίχνευση των ενδιαφερουσών ομάδων χαρακτηριστικών γνωρισμάτων θα είναι επαρκείς για να επιτύχουν αυτόν τον στόχο. Η ερώτηση είναι, πόση προσπάθεια υπολογισμού πρέπει να αφιερωθεί σε αυτό το από κάτω προς τα επάνω συστατικό έναντι των πιο επιμελημένων, πρότυπων βασισμένων για τα αντικείμενα που ανιχνεύθηκαν ήδη και που ακολουθούνται. Συνήθως, τα ενιαία αντικείμενα καλύπτουν μόνο μια μικρή περιοχή σε μια εικόνα.

Για να απαντήσουμε στην ερώτηση 2, τα βιολογικά συστήματα εικόνας κατευθύνουν τη περιοχή με υψηλή ανάλυση με μυς, που είναι πολύ γρήγορες αλλαγές κατεύθυνσης βλέμματος με τα γωνιακά ποσοστά μέχρι αρκετούς βαθμούς ανά δευτερόλεπτο, στην ομάδα χαρακτηριστικών γνωρισμάτων που ξυπνούν το περισσότερο ενδιαφέρον. Οι άνθρωποι είναι σε θέση να εκτελέσουν μέχρι πέντε κινήσεις ανά δευτερόλεπτο με τις ενδιάμεσες φάσεις ομαλής αναζήτησης (καταδίωξη) αυτών των χαρακτηριστικών γνωρισμάτων, που δείχνουν έναν πολύ δυναμικό τρόπο αντίληψης (χρόνος-τεμαχισμένη παράλληλη επεξεργασία). Η καταδίωξη μπορεί να επιτευχθεί αποτελεσματικότερα με τους αλγορίθμους ελεγχόμενους με κάποιο πρότυπο σύμφωνα με την κοντινότερη πρόβλεψη. Οι ικανοποιητικές λύσεις μπορούν να είναι δυνατές μόνο για πρόσθετες περιοχές στόχου για τις οποίες η εμπειρία είναι διαθέσιμη από τις προηγούμενες συγκρούσεις.

Δεδομένου ότι η πρόβλεψη είναι ένα πολύ ισχυρό εργαλείο σε έναν κόσμο με τις συνεχείς διαδικασίες, η ερώτηση προκύπτει: ποιο είναι το καταλληλότερο πλάνο για τη επεξεργασία των συνεχιζόμενων καταστάσεων; Είναι το πλάνο της εικόνας εύκολα - διαθέσιμο ως πλάνο αναφοράς; Εντούτοις, είναι γνωστό ότι η διάσταση βάθους στη χαρτογράφηση προοπτικής έχει χαθεί εντελώς. Όλα τα σημεία σε μια ακτίνα έχουν χαρτογραφηθεί σε ένα ενιαίο σημείο στο πλάνο της εικόνας, ανεξάρτητα από την απόστασή τους, η οποία έχει χαθεί. Θα ήταν καλύτερο να διατυπωθούν όλοι οι όροι συνοχής στο τρισδιάστατο χώρο και χρόνο; Τα αντίστοιχα πρότυπα είναι διαθέσιμα από τις φυσικές επιστήμες δεδομένου ότι ο Newton και Leibniz έχουν διαπιστώσει ότι οι διαφορικές εξισώσεις είναι τα κατάλληλα εργαλεία για να αναπαραστήσουν την συνέχεια αυτών των καταστάσεων με γενική μορφή, κατά τη διάρκεια των τελευταίων δεκαετιών, η τεχνολογία προσομοίωσης παρείχε τις μεθόδους για να επιτύχουμε αυτές τις αντιπροσωπεύσεις στους ψηφιακούς υπολογιστές.

Στην τεχνολογία επικοινωνιών και στον τομέα της αναγνώρισης σχεδίων ή βίντεο η επεξεργασία με το πλάνο εικόνας μπορεί να είναι ο καλύτερος τρόπος δεδομένου ότι καμία κατανόηση του περιεχομένου της σκηνής δεν απαιτείται. Εντούτοις, για τον προσανατολισμό στον πραγματικό κόσμο μέσω της ανάλυσης ακολουθίας εικόνας, η πρόωρη μετάβαση στο φυσικό διάστημα ερμηνείας θεωρείται ιδιαίτερα συμφέρουσα επειδή σε αυτό το διάστημα τα συμπεράσματα γίνονται πιο εύκολα κατανοητά και η συνοχή κινήσεων εμμένει. Επίσης, είναι μέσα σε αυτό το διάστημα που τα αδρανή σήματα πρέπει να ερμηνευθούν και που τα ολοκληρώματα των επιταχύνσεων παράγουν τα τρισδιάστατα τμήματα ταχύτητας; τα ολοκληρώματα αυτών των ταχυτήτων παράγουν τις αντίστοιχες θέσεις και τους γωνιακούς προσανατολισμούς για τους περιστροφικούς βαθμούς ελευθερίας. Επομένως, για την κατανόηση της οπτικής δυναμικής σκηνής, οι εικόνες θεωρούνται ενδιάμεσοι μεταφορείς των στοιχείων που περιέχουν τις πληροφορίες για το χωροχρονικό περιβάλλον. Για να ανακτήσει αυτές τις πληροφορίες αποτελεσματικότερα, όλη η εσωτερική διαμόρφωση στη διαδικασία ερμηνείας γίνεται στο τρισδιάστατο χώρο και το χρόνο, και η μετάβαση σε αυτή την αναπαράσταση πρέπει να πραγματοποιηθεί όσο το δυνατό νωρίτερα. Η γνώση για την επίτευξη αυτού του στόχου είναι συγκεκριμένη για τα ενιαία αντικείμενα και τις γενικές κατηγορίες στα οποία ανήκουν. Επομένως, για να απαντήσουν στην ερώτηση 2 ανωτέρω, οι ειδικές διαδικασίες που συνδέονται με τις κατηγορίες αντικειμένων και τα άτομα αυτών των κατηγοριών που παρατηρούνται στην ακολουθία εικόνας πρέπει να σχεδιαστούν για την άμεση ερμηνεία στο τρισδιάστατο χώρο και το χρόνο.

Μόνο αυτές οι χωροχρονικές αναπαραστάσεις μας επιτρέπουν έπειτα να απαντήσουμε στην ερώτηση 3 κοιτώντας τα δεδομένα όλων των σχετικών αντικειμένων στο κοντινό περιβάλλον για μια πιο εκτεταμένη χρονική περίοδο. Για να είμαστε σε θέση να κατανοήσουμε τις διαδικασίες κινήσεων των αντικειμένων βαθύτερα στο καθημερινό περιβάλλον μας, μια διάκριση πρέπει να γίνει μεταξύ των κατηγοριών αντικειμένων. Εκείνοι που υπακούνε στους απλούς νόμους της κίνησης από τη φυσική είναι οι πιο εύκολα αντιμετωπιζόμενοι (π.χ., από κάποια έκδοση του νόμου του Newton). Ελαφριά αντικείμενα, εύκολος κινούμενα με (ακόμα και ελαφριούς) άνεμους γίνετε δύσκολο να συνθλιφθούν λόγω των μεταβλητών ιδιοτήτων του αέρα.

Μια άλλη μεγάλη κατηγορία αντικειμένων - με πολλές διαφορετικές υποκατηγορίες - διαμορφώνεται από εκείνα που έχουν την ικανότητα να αισθανθούν την κατάσταση του περιβάλλοντός τους και για να αρχίσουν να μετακινούνται από μόνα τους, βασισμένα σε έναν συνδυασμό στοιχείων που αισθάνονται και στην εμπειρία τους. Αυτά τα πρόσθετα αντικείμενα θα ονομαστούν θέματα όλα τα ζώα συμπεριλαμβανομένων των ανθρώπων ανήκουν σε αυτήν την (υπέρ-) κατηγορία καθώς επίσης και τους αυτόνομους πράκτορες που δημιουργούνται από τεχνικά μέσα (όπως τα ρομπότ ή τα αυτόνομα οχήματα). Οι αντίστοιχες υποκατηγορίες διαμορφώνονται

από τους συνδυασμούς αντιληπτικών και συμπεριφοριστικών ικανοτήτων και, φυσικά από τις μορφές τους. Εκτός από τις μορφές τους, τα άτομα των υποκατηγοριών μπορούν να αναγνωριστούν επίσης από τα στερεοτυπικά σχέδια κινήσεων (όπως ένα χοροπηδηκτού καγκουρό ή ένα τυλιγμένο φίδι).

Τα οδικά οχήματα (ανεξάρτητα του ελέγχου από έναν ανθρώπινο οδηγό ή ένα τεχνικό υποσύστημα) παρουσιάζουν χαρακτηριστικές συμπεριφορές ανάλογα με την κατάσταση που αντιμετωπίζουν. Παραδείγματος χάριν, ακολουθούν γραμμές και κάνουν κομβί, κάνουν αλλαγές γραμμών, περνούν άλλα οχήματα ή επιβραδύνουν για να σταθμεύσουν. Όλοι αυτοί οι ελιγμοί είναι γνωστοί στους ανθρώπινους οδηγούς, και αναγνωρίζουν την πρόθεση εκτέλεσης ενός από τη χαρακτηριστική αρχική κίνησή του κατά τη διάρκεια ενός μικρού χρονικού διαστήματος. Παραδείγματος χάριν, ένα αυτοκίνητο που αλλάζει την κατεύθυνση του και κινείται αποφασιστικά προς τη γειτονική λωρίδα κυκλοφορίας υποθέτουμε ότι άρχισε μια αλλαγή λωρίδας. Εάν αυτό συμβεί μέσα στο περιθώριο ασφάλειας, το αυτοκινούμενο όχημα πρέπει να προσαρμοστεί σε αυτήν την (ανάρμοστη) συμπεριφορά άλλων συμμετεχόντων στην κυκλοφορία. Αυτό δείχνει ότι η αναγνώριση της πρόθεσης άλλων θεμάτων είναι σημαντική για ένα αμυντικό ύφος οδήγησης. Αυτό δεν μπορεί να επιτευχθεί χωρίς την γνώση των χρονικά εκτεταμένων ελιγμών και χωρίς την παρατήρηση της συμπεριφοράς των άλλων θεμάτων στο περιβάλλον. Η ερώτηση 3 επομένως δεν μπορεί να απαντηθεί ερμηνεύοντας την εικόνα άμεσα αλλά με την παρατήρηση των συμβολικών αντιπροσωπεύσεων που οδηγούν όπως και στην ερώτηση 2 για διάφορα μεμονωμένα αντικείμενα/θέματα κατά τη διάρκεια μιας εκτεταμένης χρονικής περιόδου.

Η ταυτόχρονη διερμηνεία των ακολουθιών εικόνας στις πολλαπλές κλίμακες στο τρισδιάστατο χώρο και χρόνο είναι ο τρόπος να ικανοποιήσουμε όλες οι απαιτήσεις για την ασφαλή και προσανατολισμένη συμπεριφορά.

3.4 Διαφορικά πρότυπα για την αντίληψη

Η εμπειρία έχει δείξει ότι η ταυτόχρονη χρήση των διαφορικών και ακέραιων προτύπων στις διαφορετικές παραγωγές κλιμάκων είναι ο καλύτερος τρόπος τήξης και ερμηνείας των στοιχείων. Το σχήμα 1.2 παρουσιάζει σε μια συστηματική μέθοδο το σχέδιο ερμηνείας που αναπτύσσεται. Κάθε ένας από τους άξονες υποδιαιρείται σε τέσσερις σειρές κλίμακας. Στην πάνω αριστερή γωνία το σημείο «εδώ και τώρα» παρουσιάζεται ως σημείο όπου πραγματοποιείται όλη η αλληλεπίδραση με τον πραγματικό κόσμο. Η δεύτερη σειρά κλίμακας καλύπτει τον τοπικό (όπως αντιτάσσεται στο παγκόσμιο) περιβάλλον που επιτρέπει να ανακαλύψουμε νέες διαφορικές έννοιες έναντι στην

υποδεικνυόμενη κατάσταση. Τοπική ενσωμάτωση, με τις χαρακτηριστικές ιδιότητες όπως τα χωρικά ή χρονικά ποσοστά αλλαγής, οι χωρικές κλίσεις, ή οι κατευθύνσεις των ακραίων τιμών όπως οι κλίσεις έντασης είναι χαρακτηριστικά παραδείγματα.

ΠΙΝΑΚΑΣ

Αυτά τα διαφορικά έχουν παρουσιάσει σαν ισχυρές έννοιες για την αντιπροσώπευση της γνώσης για τις σωματικές ιδιότητες των κατηγοριών αντικειμένων. Οι διαφορικές εξισώσεις αντιπροσωπεύουν το φυσικό μαθηματικό στοιχείο για τη γνώση κωδικοποίησης για τις διαδικασίες κινήσεων στο πραγματικό κόσμο. Με την εμφάνιση του φίλτρου Kalman [Kalman 1960], έχουν γίνει το κλειδί για τη λήψη της καλύτερης εκτίμησης των μεταβλητών που περιγράφουν το σύστημα, βασισμένης στις επαναλαμβανόμενες μεθόδους εφαρμόζοντας μια πρότυπη τακτοποίηση σε ένα όσο το δυνατόν μικρότερο πίνακα με τετράγωνα. Η σε πραγματικό χρόνο οπτική αντίληψη για την κίνηση των αντικειμένων είναι μετά βίας δυνατή χωρίς αυτήν την πολύ αποδοτική προσέγγιση.

3.5 Πληροφορίες σε μια εικόνα

Είναι γνωστό ότι οι πληροφορίες σε μια εικόνα περιλαμβάνονται στις τοπικές εναλλαγές της έντασης: Μια ομοιόμορφα γκριζα εικόνα έχει μόνο μερικά bits πληροφορία, δηλαδή, (1) την τιμή του γκρι και (2) ομοιόμορφη διανομή αυτής της τιμής σε ολόκληρη την εικόνα. Η εικόνα μπορεί να περιγραφεί εντελώς από τρία bytes, ακόμα κι αν το ποσό στοιχείων μπορεί να είναι περίπου 400 000 bytes σε ένα πλαίσιο TV ή ακόμα και 4 MB (2k 2k pixels). Εάν υπάρχουν ορισμένοι τομείς των ομοιόμορφων γκριζών τιμών, οι γραμμές ορίου αυτών των περιοχών συν τις εσωτερικές γκριζες τιμές περιέχουν όλες τις πληροφορίες στην εικόνα. Αυτό το αντικείμενο στο πλάνο εικόνας μπορεί να περιγραφεί με πολύ λιγότερα στοιχεία από τις τιμές pixels που καλύπτει.

Σε μια γενικότερη μορφή, οι περιοχές εικόνας που καθορίζονται από ένα σύνολο ιδιοτήτων (μορφή, σύσταση, χρώμα, κοινή κίνηση, κ.λπ.) μπορούν να θεωρηθούν αντικείμενα εικόνας, τα οποία προήλθαν από τα τρισδιάστατα αντικείμενα από τη χαρτογράφηση προοπτικής. Λόγω των πολυάριθμων πτυχών, που ένα τέτοιο αντικείμενο μπορεί να υιοθετήσει σχετικά με τη κάμερα, οι πιθανές εμφανίσεις της στο πλάνο εικόνας είναι πολύ διαφορετικές. Η αντιπροσώπευσή τους θα απαιτήσει περισσότερα στοιχεία για μια εξαντλητική περιγραφή από την αντιπροσώπευσή της στο τρισδιάστατο διάστημα συν τους νόμους της χαρτογράφησης προοπτικής, οι οποίοι είναι οι ίδιοι για όλα τα αντικείμενα. Επομένως, ένα αντικείμενο καθορίζεται από την τρισδιάστατη μορφή του, η οποία μπορεί να θεωρηθεί τοπικό χωρικό ολοκλήρωμα της διαφορικής περιγραφής γεωμετρίας της στους όρους κυρτότητας. Ανάλογα με το στόχο και πόσο εφικτός είναι, και η διαφορική και ακέραια αντιπροσώπευση, ή ένας συνδυασμός και των δύο μπορεί να χρησιμοποιηθεί για την οπτική

αναγνώριση. Όπως θα παρουσιαστεί για το παράδειγμα της καθοδήγησης οδικών οχημάτων, η παράλληλη χρήση αυτών των προτύπων στα διαφορετικά μέρη της γενικότερης διαδικασίας και του συστήματος ελέγχου αναγνώρισης μπορεί να είναι η αποδοτικότερη.

3.6 Σε ποίο σημείο εισέρχεται η ανθρώπινη γνώση;

Τα αντικείμενα και οι κατηγορίες αντικειμένου διαδραματίζουν έναν σημαντικό ρόλο στην ανθρώπινη γλώσσα και στο να μάθουμε να κατανοούμε «τον κόσμο». Αυτό ισχύει για την εμφάνισή τους σε πρώτο χρόνο, και επίσης για τη συμπεριφορά κινήσεών τους με την πάροδο του χρόνου. Στο χρονικό άξονα, η συνδυασμένη χρήση των διαφορετικών και ακέραιων προτύπων μπορεί να μας επιτρέψει σε να αποσχιστούμε από τον υπολογισμό των οπτικών διανυσματικών τομέων ροής ή μετατοπίσεων, όποιοι είναι πολύ έντονοι υπολογισμοί και ευαίσθητοι στο θόρυβο. Λόγω του τεράστιου ποσού στοιχείων σε μια ενιαία εικόνα, αυτό δεν θεωρείται καλύτερος τρόπος να πάει, δεδομένου ότι μια πρόωρη μετάβαση στην έννοια των φυσικών αντικειμένων ή των θεμάτων με τους όρους συνοχής στο τρισδιάστατο χώρο και το χρόνο έχει διάφορα πλεονεκτήματα: (1) βοηθά να κόψει το ποσό το στοιχείο που απαιτείται για την επαρκή περιγραφή, και (2) αυτό παράγει το κατάλληλο πλαίσιο για τη γνώση που προέρχεται από τις προηγούμενες συγκρούσεις (δυναμικά πρότυπα, stereotypical ελιγμοί ελέγχου, κ.λπ.). Αυτό το βήμα αντικαθίσταται από τις γνωστές τεχνικές παρατηρητών στη δυναμική συστημάτων (φίλτρο Kalman και παράγωγα, παρατηρητές Luenberger). Αυτές οι επαναλαμβανόμενες μέθοδοι αναδημιουργούν τα χρονικά παράγωγα του κράτους οι μεταβλητές από το λάθος πρόβλεψης ανατροφοδοτούν και γνώση για τη δυναμική συμπεριφορά από το αντικείμενο και (για το φίλτρο Kalman) των στατιστικών ιδιοτήτων του συστήματος (που μεταγλωττίζεται «τις εγκαταστάσεις» στη δυναμική συστημάτων) και των διαδικασιών μέτρησης. οι στερεοτυπικές συμπεριφοριστικές ικανότητες των θεμάτων στις διαφορετικές καταστάσεις διαμορφώνουν ένα σημαντικό μέρος της βάσης γνώσεων.

Δύο ευδιάκριτα διαφορετικοί τύποι «τοπικών χρονικών ολοκληρωμάτων» χρησιμοποιούνται ευρέως: Ολοκληρώματα ενιαίων βημάτων για την τηλεοπτική δειγματοληψία και πολλαπλάσια τοπικά ολοκληρώματα βημάτων (για την κατανόηση ελιγμού. Μέσω της διαδικασίας απεικόνισης, η αναλογική διαδικασία κινήσεων μέσα ο πραγματικός κόσμος γίνεται ιδιαίτερος κατά μήκος του χρονικού άξονα. Με τη διαμόρφωση (προσεγγίστε, από καμένος γραμμικό) τα ολοκληρώματα, χρονική έκταση του αναλογικού τηλεοπτικού κύκλος ζωών (33 1/3 κα στις Ηνωμένες Πολιτείες και

την κα 40 στην Ευρώπη, αντίστοιχα, μισές από αυτές τις τιμές για τους τομείς) γεφυρώνονται από τις ιδιαίτερες μήτρες μετάβασης από το $kT (K + 1) T$, $K =$ τρέχοντας δείκτης.

Ακόμα κι αν οι τιμές έντασης κάθε εικονοκτυτάρου είναι ολοκληρώματα πέρα από τη μεγάλη έκταση ή μέρος αυτής της περιόδου, ερμηνεύονται ως πραγματικά επιλεγείσα αξία έντασης κατά την διάρκεια της ανάγνωσης φωτογραφικών μηχανών. Δεδομένου ότι όλες οι βασικές ερμηνείες της κατάστασης στηρίζονται σε αυτά τα στοιχεία, η παραγωγή ελέγχου υπολογίζεται πρόσφατα μόνο μετά από αυτήν την περίοδο κατά συνέπεια, είναι σταθερό κατά τη διάρκεια του βασικού κύκλου ζωών. Αυτό επιτρέπει τον αναλυτικό υπολογισμό των αντίστοιχων κρατικών μεταβάσεων, οι οποίες αξιολογούνται αριθμητικά για κάθε κύκλο επαναλαμβανόμενη διαδικασία εκτίμησης (κεφάλαιο 6) αυτοί χρησιμοποιούνται για την κρατική πρόβλεψη και τον ευφυή έλεγχο της εξαγωγής χαρακτηριστικών γνωρισμάτων εικόνας.

3.7 Σφαιρικά ολοκληρώματα για την αξιολόγηση της κατάστασης

Οι πιο σύνθετες καταστάσεις που καλύπτουν πολλά αντικείμενα ή αποστολές που αποτελούνται από τις ακολουθίες στοιχείων αποστολής αντιπροσωπεύονται στη χαμηλότερη δεξιά γωνία του σχεδίου 1.2. Πάλι, πώς να επιλέξετε τις καλύτερες υποδιαιρέσεις και απόλυτες κλίμακες στο χρονικό άξονα ή στο διάστημα εξαρτώνται πάρα πολύ από την προβληματική περιοχή που μελετάμε. Αυτό θα είναι τελείως διαφορετική προσέγγιση για την κατασκευή μικροσυστημάτων έναντι ενός συστήματος στη διαστημική πτήση. Η βασική αρχή στην υποδιείρεση του γενικού στόχο, εντούτοις, μπορεί να είναι σύμφωνα με το ίδιο σχέδιο που δίνεται στο σχήμα 1.2, ακόμα κι αν τα τεχνικά στοιχεία που χρησιμοποιούνται μπορούν να είναι τελείως διαφορετικά.

Σε ένα πολύ μεγαλύτερο χρονοδιάγραμμα, η επίδραση όλων των feed-forward χρονικών ιστοριών ελέγχου μπορούν να προβλεφτούν ώστε να έχουν στόχο μερικές πρόσθετες κρατικές αλλαγές ή μεταβάσεις. Παραδείγματος χάριν, αλλαγή παρόδων ενός οδικού οχήματος σε έναν αυτοκινητόδρομο, ο οποίος μπορεί να διαρκέσει 2 έως 10 δευτερόλεπτα στο σύνολο, μπορεί να περιγραφεί ως μια καλά δομημένη ακολουθία ελέγξτε τα αποτελέσματα με συνέπεια σε μια ορισμένη τροχιά του οχήματος. Στο τέλος του ελιγμού, το όχημα πρέπει να είναι στη γειτονική πάροδο με τις ίδιες κρατικές μεταβλητές διαφορετικά (ταχύτητα, πλευρική θέση στην πάροδο, κατεύθυνση). Το σύμβολο «αλλαγή κατεύθυνσης», ωστόσο, ισχύει για ένα σχετικά σύνθετο στοιχείο ελιγμού που μπορεί να προκληθεί από τα πιο υψηλά επίπεδα κατόπιν παραγγελίας

χρησιμοποιώντας αυτό το σύμβολο (ίσως μαζί με μερικές παραμέτρους που διευκρινίζουν το χρόνο ελιγμού και, με αυτόν τον τρόπο, μέγιστη πλευρική επιτάχυνση που μπορεί να αντιμετωπιστεί).

Αυτά τα «ελισσόμενα στοιχεία», καθορισμένα κατάλληλα, επιτρέπουν σε μας να αποσυνθέσουμε σύνθετους ελιγμούς σε στερεοτυπικά στοιχεία που μπορούν να συναρμολογηθούν σύμφωνα με τις πραγματικές ανάγκες. Μεγάλα μέρη αυτών των αποστολών μπορούν να εκτελεστούν με την εκμετάλλευση του ανατροφοδοτούμενου έλεγχου, όπως η πορεία σε γραμμή και η διατήρηση απόστασης που κρατούν για τα οδικά οχήματα. Με αυτόν τον τρόπο, οι κλίμακες των αποστάσεων για τις ολόκληρες αποστολές εξαρτώνται από τη διαδικασία που ελέγχονται, αυτοί θα είναι απολύτως διαφορετικοί για τα «αυτόνομα καθοδηγημένα οχήματα» (AGVs) στο πάτωμα εργοστασίων (εκατοντάδες μετρά) συγκρινόμενα με τα οδικά οχήματα (δεκάδες χλμ) ή ακόμα και τα αεροσκάφη (εκατοντάδες ή χιλιάδες χλμ).

Το σχέδιο του συστήματος εικόνας πρέπει να επιλεχτεί ανάλογα με το στόχο.

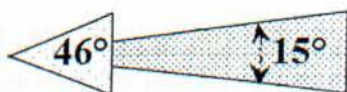
3.8 Επάρκεια συστημάτων εικόνας

Για τον έλεγχο κινήσεων, λόγω της αδράνειας ενός σώματος, το πραγματικό διάνυσμα ταχύτητας αποφασίζει πού πρέπει να κοιτάξει για να αποφύγει τις συγκρούσεις με άλλα αντικείμενα. Δεδομένου ότι ο πλευρικός έλεγχος μπορεί να εφαρμοστεί ως ένα ορισμένο βαθμό και δεδομένου ότι άλλα αντικείμενα και θέματα μπορούν να έχουν ένα δικό τους διάνυσμα ταχύτητας, η εξεταζόμενη απόσταση πρέπει να είναι αρκετά μεγάλη για την ανίχνευση όλων των πιθανών πορειών σύγκρουσης με άλλα αντικείμενα. Επομένως, το ταυτόχρονο οπτικό πεδίο είναι κρισιμότερο σε κοντινές αποστάσεις.

Αφ' ενός εάν απαιτείται οδήγηση σε υψηλή ταχύτητα, η όραση μπροστά πρέπει να έχει αρκετά μεγάλο βεληνεκές για να ανιχνεύσει αξιόπιστα τα αντικείμενα στις αποστάσεις που επιτρέπουν το ασφαλές φρενάρισμα. Με μια ταχύτητα 30 m/s (108 km/h ή περίπου 65 mph), η απόσταση για το φρενάρισμα [με ένα επίπεδο επιβράδυνσης 0.4 γήινης βαρύτητας $g(9.81 \text{ m/s}^2)$, το οποίο είναι $ax = -4 \text{ m/s}^2$] και με το χρόνο αντίδρασης 0.5 δευτερολέπτων] είναι $15 + 113 = 128 \text{ μ}$. Για το μισό μέγεθος στην επιβράδυνση (-2 m/s^2 , π.χ., υπό τους δυσμενείς οδικούς όρους) η απόσταση φρεναρίσματος θα ήταν 240 μ.

Η αξιόπιστη εκτίμηση απόστασης για τα οδικά οχήματα εμφανίζεται υπό τους όρους χαρτογράφησης με τουλάχιστον περίπου 20 pixels στο πλάτος του οχήματος (χαρακτηριστικά περίπου 2 μ στη διάσταση). Το συνολικό οπτικό πεδίο μιας κάμερας σε μια απόσταση 130 μ, όπου

αυτός ο όρος ικανοποιεί, θα είναι περίπου 76 μ (για ~ 760 pixels ανά γραμμή). Αυτό αντιστοιχεί σε μια γωνία ανοιγμάτων ~ 34°. Αυτό δεν είναι βεβαίως αρκετό να καλύψει ένα επαρκές οπτικό πεδίο στην κοντινή απόσταση. Επομένως, τουλάχιστον μια διεστιακή ρύθμιση καμερών απαιτείται με δύο διαφορετικά εστιακά μήκη,

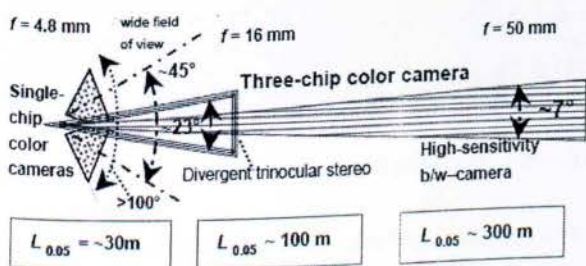


$$f = 7.5 \text{ mm} \quad f = 24 \text{ mm}$$

Στην απόσταση $L_5 \sim 20 \text{ m}$ ($\sim 60 \text{ m}$), η ανάλυση είναι 5 cm/pixel

Για μια μάλλον ευέλικτη υψηλών επιδόσεων «τεχνικό μάτι» μια trifocal ρύθμιση καμερών όπως φαίνεται στο σχήμα 1.4 συστήνεται. Η διπλής ευρείας γωνίας CCD κάμερα με το εστιακό μήκος 4 έως 6 χιλ. και με τους διάφορους οπτικούς άξονες έχει το κεντρικό βεληγεκές των επικαλυπτομένων περιοχών εικόνας, η οποία επιτρέπει την κοντινή στερεοφωνική-ερμηνεία. Στο σύνολο, ένα οπτικό πεδίο περίπου 100 έως 130 μοιρών μπορεί να καλυφθεί αυτό επιτρέπει να επιβλεφτεί το ένα τρίτο του ολόκληρου πανοράματος.

Μια ήπια ασύρματη κάμερα με τρεις έως τέσσερις φορές το εστιακό μήκος της ευρείας γωνίας πρέπει να είναι μια έγχρωμη κάμερα με τρία -τσιπ για την ακριβέστερη αναγνώριση του αντικειμένου. Το οπτικό πεδίο του περιλαμβάνεται στο στερεοφωνικό οπτικό πεδίο των ευρείας γωνίας καμερών έτσι ώστε η τριεστιακή στερεοσκοπική ερμηνεία να είναι δυνατή [Rieder 1996].

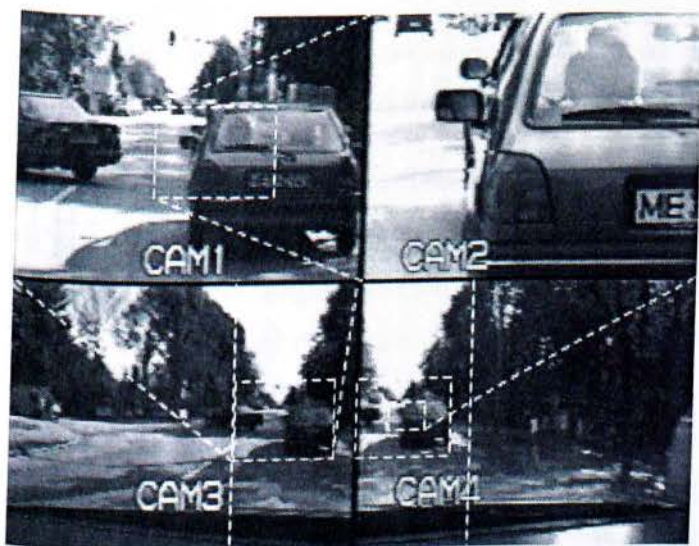


Σχήμα 1.4. Ρύθμιση τρισκοπικής (Trifocal) κάμερας με ευρύ οπτικό πεδίο

Για να ανιχνευτούν τα αντικείμενα σε συγκεκριμένους τομείς ενδιαφέροντος μακριά, μια κάμερα με ένα 3x εστιακό μήκος (πάλι με έναν παράγοντα 3 έως 4 σχετικά με τα ήπιους τηλεφακούς), και το

οπτικό πεδίο μέσα στο οποίο η ήπια ασύρματη κάμερα πρέπει να προστεθεί (δείτε το σχήμα 1.4). Αυτή η κάμερα μπορεί να επιλεχτεί για να είναι ιδιαίτερα φωτοευαίσθητη, οι ασπρόμαυρες εικόνες μπορεί να είναι επαρκής για να περιορίσουν το ποσό των δεδομένων. Η εστιακή αναλογία μήκους 4 έχει το πλεονέκτημα ότι η πιο χονδροειδής εικόνα αντιπροσωπεύει την ίδια σκηνή σε μια ανάλυση που αντιστοιχεί στο δεύτερο πυραμιδικό στάδιο του καλύτερου.

Αυτός ο τύπος συνδυασμού αισθητήρων είναι ιδανικά ταιριαγμένος για τον ενεργό έλεγχο κατεύθυνσης : η χονδροειδής ανάλυση ,το μεγάλο εξομοιούμενο οπτικό πεδίο επιτρέπει την ανίχνευση αντικείμενων πιθανού ενδιαφέροντος για μια ευρεία περιοχή, και μια εμφανής αλλαγή κατεύθυνσης θα φέρει αυτό το αντικείμενο στο κέντρο των εικόνων με την υψηλότερη ανάλυση. Συγκρίνοντας με μια κάμερα με τη μέγιστη ανάλυση στο ίδιο ολόκληρο οπτικό πεδίο, η λύση που παρουσιάζεται δείχνει να έχει μόνο 2με 4% του ποσοστού στοιχείων. Επιτυγχάνει αυτό σε αντάλλαγμα της ανάγκης του γρήγορου ελέγχου της κατεύθυνσης και εις βάρος των χρόνων καθυστέρησης που απαιτούνται για να εκτελεστούν αυτές τις αλλαγές <<βλέμματος>>. Το σχήμα 1.5 δίνει μια εντύπωση των οπτικών πεδίων αυτής της τρισκοπικής κάμερας.



Σχήμα 1.5. Οπτικά πεδία της trifocal ρύθμισης φωτογραφικών μηχανών.

Οι δύο χαμηλότερες ευρείας γωνίας εικόνες έχουν μια κεντρική περιοχή της επικάλυψης σημειούμενες με κάθετες άσπρες γραμμές. Στην αριστερή, φαίνετε όλη η διασταύρωση και εμφανίζετε ένα αυτοκίνητο που βγαίνει από το σταυροδρόμι και ένα άλλο ένα ακριβώς την ώρα που μπαίνει στο σταυροδρόμι, το οπίσθιο τμήμα αυτού του οχήματος και το όχημα ακριβώς μπροστά μπορούμε να το δούμε στην ανώτερη αριστερή εικόνα της ήπιας ασύρματης κάμερας. Αυτό

επιτρέπει ακόμη και τη τρισκοπική στερεοφωνική ερμηνεία. Η περιοχή που χαρακτηρίζεται στο λευκό σε αυτό την ήπια τηλε-εικόνα παρουσιάζεται επάνω δεξιά ως πλήρης εικόνα από μια ισχυρή ασύρματη κάμερα. Εδώ, τα γράμματα στη πινακίδα αριθμού κυκλοφορίας μπορούν να διαβαστούν, και μπορούμε να δούμε από το δεύτερο οπισθοσκόπο καθρέφτη στην αριστερή πλευρά ότι υπάρχει ένα δεύτερο αυτοκίνητο αμέσως μπροστά από το προπορευόμενο αυτοκίνητο. Ο αριθμός των pixels ανά περιοχή στο ίδιο αντικείμενο σε αυτήν την εικόνα είναι εκατό φορές αυτός των ευρείας γωνίας εικόνων.

Για την αδρανή σταθεροποίηση της αντιλαμβανόμενης κατεύθυνσης κατά οδήγηση πέρα από μια ανώμαλη επιφάνεια ή για τα αεροσκάφη που πετούν σε έναν ταραχώδη αέρα, μια ενεργή ανάρτηση καμερών απαιτείται οπωσδήποτε. Η ταυτόχρονη χρήση σχεδόν των χωρίς καθυστέρηση αδρανών μετρήσεων (χρονικά παράγωγα όπως τα γωνιακά ποσοστά και οι γραμμικές επιταχύνσεις) και των εικόνων, των οποίων η ερμηνεία εισάγει διάφορα δέκατα ενός δεύτερου χρόνου καθυστέρησης, απαιτεί τις εκτεταμένες αντιπροσωπεύσεις κατά μήκος του χρονικού άξονα. Δεν υπάρχει κανένας χρόνος για τον οποίο είναι πιθανός να καταστήσει τη συνεπή αίσθηση όλων των στοιχείων διαθέσιμη. Μόνο η έννοια μιας «εκτεταμένης παρουσίας» επιτρέπει σε μια αποδοτική αμετάβλητη ερμηνεία (στο 4D). Για αυτόν τον λόγο, το πολυεστιακό, saccadic σύστημα εικόνας θεωρείται η προτιμητέα λύση για τα αυτόνομα οχήματα γενικά.

3.9 Τεχνικά συστήματα εναντίον των βιολογικών

Τα βιολογικά συστήματα εικόνας έχουν εξελιχθεί πέρα από τα εκατομμύρια των γενεών με την επιλογή του καταλληλότερου οικολογικού περιβάλλοντος που αντιμετωπίζουν. Το βασικό ο νευρικό υποστρώμα που αναπτύχθηκε (βασισμένο σε άνθρακα) μπορεί να χαρακτηριστεί από μερικούς αριθμούς. Οι ηλεκτροχημικές μονάδες έχουν τους χρόνους διακοπής της τάξεως των χιλιοστών του δευτερολέπτου (ms) η ταχύτητα διακίνησης των σημάτων είναι της τάξεως των 10 έως 100 m/s. Οι συνδέσεις μεταξύ των μονάδων υπάρχουν στην αφθονία (1000 έως 10 000 ανά νευρώνα). Ένας εγκέφαλος αποτελείται από μέχρι και 10¹¹ αυτών των μονάδων. Το κύριο βήμα επεξεργασίας είναι άθροισμα των σταθμισμένων σημάτων εισαγωγής που περιέχουν μέχρι τώρα άγνωστους (πολλαπλούς;) ανατροφοδοτούμενους βρόχους [εγχειρίδιο της φυσιολογίας 1984, 1987].

Αυτά τα συστήματα χρειάζονται τους μακροχρόνιους χρόνους εκμάθησης και προσαρμόζονται στις νέες καταστάσεις αργά. Αντίθετα, τα τεχνικά υποστρώματα για τους αισθητήρες και οι μικροεπεξεργαστές (βασισμένοι σε πυρίτιο) έχουν χρόνους διακοπής της τάξεως νανοδευτερολέπτου (έναν παράγοντα 10⁶ έναντι των βιολογικών συστημάτων). Είναι εύκολα προγραμματιζόμενοι και έχουν διάφορους υπολογιστικούς τρόπους μεταξύ των οποίων κάποιοι μπορούν να κλείσουν σχεδόν

Ενδιαφέρουσες μελέτες έχουν γίνει σε διάφορα ερευνητικά κέντρα που προσπάθησαν να εξερευνήσουν την αναλογική επεξεργασία στοιχείων - στα τσιπ σιλικόνης [Koch 1995] οι μελλοντικές συγκρίσεις των αποτελεσμάτων θα πρέπει να παρουσιάσουν εάν το διάστημα που απαιτείται στο τσιπ μπορεί να δικαιολογηθεί από τα πλεονεκτήματα που απαιτούνται.

Η επικρατούσα ανάπτυξη σήμερα οδηγείται από την εμπορική TV για τους αισθητήρες και από τα προσωπικούς Η/Υ και παιχνίδια για τους επεξεργαστές. Με μια αναμενόμενη αύξηση στη δύναμη υπολογισμού ενός μεγέθους κάθε 4 έως 5 έτη πέρα από επόμενη δεκαετία, η σε πραγματικό χρόνο μηχανική όραση θα είναι έτοιμη για ένα ευρύ φάσμα εφαρμογών που χρησιμοποιούν τις συμβατικές μεθόδους εφαρμοσμένης μηχανικής όπως αντιπροσωπεύεται από την 4-D προσέγγιση.

Μερικοί (ίσως μια δωδεκάδα) από αυτούς τους επεξεργαστές θα είναι ικανοποιητικοί για να λύσουν ακόμη και σύνθετους στόχους όπως και την καθοδήγηση οχημάτων εδάφους ή και αέρος, οι διπλοί επεξεργαστές σε ένα ενιαίο τσιπ μόλις άρχισαν να μπαίνουν στην αγορά.

3.10 Νοημοσύνη

Οι αισθητήρες των σύνθετων αυτόνομων βιολογικών ή τεχνικών συστημάτων παράγουν ένα τεράστιο ποσοστό δεδομένων που περιέχει τις πληροφορίες και για την κατάσταση του σώματος των οχημάτων σχετικά με το περιβάλλον και για άλλα αντικείμενα ή θέματα στο περιβάλλον. Είναι ο στόχος ενός ευφυούς συστήματος εξαγωγής πληροφοριών (ερμηνεία στοιχείων) να ξεφορτωθούν γρήγορα όσο το δυνατόν περισσότερα στοιχεία, παρ' όλα αυτά ταυτόχρονα, να διατηρηθούν όλες οι ουσιαστικές πληροφορίες για το στόχο που πρέπει να επιτευχτεί. Οι ουσιαστικές πληροφορίες συνδέονται με τις περιοχές στόχου εντούτοις, τα πολυσύνθετα συστήματα όπως τα ζώα και τα αυτόνομα οχήματα δεν έχουν μόνο έναν ενιαίο στόχο να εκτελέσουν. Ανάλογα με τις περιστάσεις, αρκετά διαφορετικοί στόχοι μπορούν να υπερισχύσουν.

Τα συστήματα θα ονομαστούν ευφυή εάν μπορούν να:

αναγνωρίζουν τις καταστάσεις που απαιτούν ορισμένες συμπεριφοριστικές ικανότητες και να προκαλέσουν αυτήν την συμπεριφορά νωρίς και σωστά, έτσι ώστε η γενική προσπάθεια να εξεταστεί η κατάσταση να είναι χαμηλότερη από για την άμεση αντίδραση σε κάποιο συνδυασμό τιμών που μετριοούνται αλλά που εμφανίζονται αργότερα (τακτική - στρατηγική διαφοροποίηση).

Αυτή η «διορατικότητα» στις διαδικασίες στο πραγματικό κόσμο είναι ενδεικτική ενός εσωτερικού χρονικού προτύπου για αυτήν την διαδικασία στο σύστημα ερμηνείας. Επομένως, η κατανόηση της

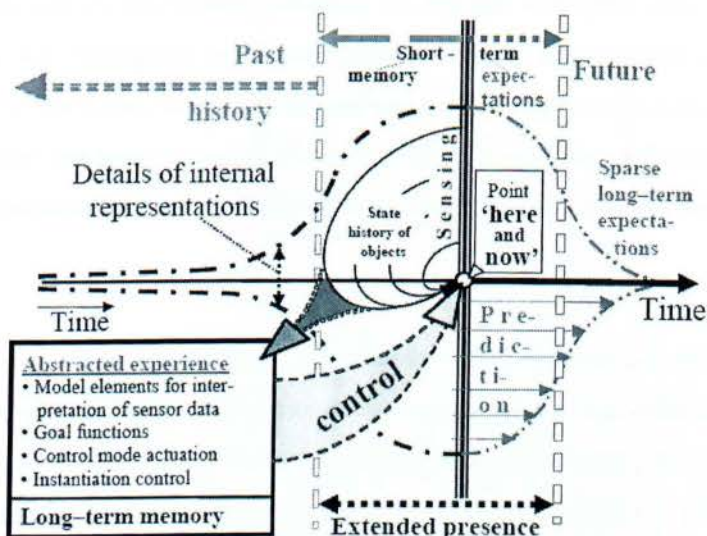
νοημοσύνης από αυτή την άποψη απαιτεί τη γνώση υποβάθρου για τις διαδικασίες για να γίνει αντιληπτή και η ικανότητα να αναγνωριστούν οι παρόμοιες ή ελαφρώς διαφορετικές καταστάσεις προκειμένου να είναι σε θέση να επεκτείνει τη βάση γνώσεων για τη σωστή χρήση.

Δεδομένου ότι το ίδιο ευφυές σύστημα θα πρέπει να εξετάσει πολλές διαφορετικές καταστάσεις, εκείνα τα άτομα θα είναι ανώτερα και θα μπορούν να εξαγάγουν τις πληροφορίες από πραγματική εμπειρία όχι μόνο για την περίπτωση που αντιμετωπίζει αλλά και για την κατάλληλη χρήση σε άλλες καταστάσεις. Αυτός ο τύπος «μεταφοράς γνώσης» είναι χαρακτηριστικός των αληθινά ευφυών συστημάτων. Από αυτήν την άποψη, η νοημοσύνη δεν είναι η ικανότητα να χειριστεί μερικά αφηρημένα σύμβολα στην απομόνωση αλλά για να έχει τις συμβολικές αντιπροσωπεύσεις διαθέσιμες που επιτρέπουν ευνοϊκές αποφάσεις για τη δράση στις διαφορετικές καταστάσεις που πρέπει να αναγνωριστούν νωρίς και σοβαρά.

Αυτές οι ενέργειες μπορούν να είναι ανατροφοδοτούμενοι νόμοι ελέγχου με πολύ γρήγορες εφαρμογές στην παραγωγή ελέγχου άμεσα από τις μετρούμενες ποσότητες (σαν αντανακλαστική συμπεριφορά), ή στερεοτυπικές feed-forward ιστορίες ελέγχου που επικαλούνται μετά από κάποιο γεγονός, γνώση για να επιτύχει το επιθυμητό αποτέλεσμα (βασισμένο στους κανόνες εκκίνησης). Για να ασχοληθούν με τις κοινές διαταραχές στο πραγματικό κόσμο, οι προσδοκίες των κρατικών μεταβλητών χρόνου ή ανταποκρινόμενα σε κάποια feed-forward παραγωγή ελέγχου μπορούν να καθοριστούν. Διαφορές μεταξύ των αναμενόμενων και παρατηρηθέντων καταστάσεων χρησιμοποιούνται σε ανατροφοδοτούμενους βρόχους για να τροποποιήσουν τη συνολική παραγωγή ελέγχου έτσι ώστε οι αναμενόμενες καταστάσεις να επιτυγχάνονται τουλάχιστον στο περίπου παρά τις απρόβλεπτες διαταραχές.

Ελέγχοντας αυτά τα αντικείμενα ελέγχου και τις προκύπτουσες ιστορίες κρατικού μεταβλητές χρόνου, η πρόκληση «ισόπεδης γνώση» έχει όλη τη διαθέσιμη πληροφορία για τον έλεγχο των εσωτερικών προτύπων στα οποία βάσισε τις προβλέψεις του και τις αποφάσεις του. Σε ένα διανεμημένο σύστημα επεξεργασίας, αυτό το επίπεδο γνώσης δεν χρειάζεται να περιληφθεί σε οποιοδήποτε από τους γρήγορους βρόχους εφαρμογής ελέγχου και κρατικής εκτίμησης. Εάν υπάρχουν συστηματικά λάθη πρόβλεψης, αυτά μπορούν να χρησιμοποιηθούν για να τροποποιήσουν τα πρότυπα. Επομένως, η ελαχιστοποίηση λάθους πρόβλεψης μπορεί να χρησιμοποιηθεί όχι μόνο για την τωρινή εκτίμηση σύμφωνα με κάποιο πρότυπο αλλά και για την προσαρμογή του ίδιου του προτύπου, με αυτόν τον τρόπο μαθαίνοντας να καταλαβαίνει καλύτερα τα συμπεριφοριστικά χαρακτηριστικά ενός σώματος ή του περιβάλλοντος διαταραχής στην πραγματική κατάσταση. Και οι δύο μπορούν να χρησιμοποιηθούν στο μέλλον σαν πλεονέκτημα. Η γνώση που αποθηκεύεται έτσι

είναι συμπυκνωμένες πληροφορίες για το (υλικό) κόσμο συμπεριλαμβανομένου του σώματος του οχήματος που φέρνει τους αισθητήρες και τον εξοπλισμό επεξεργασίας εάν μπορεί να επικαλεσθεί στις αντίστοιχες καταστάσεις στο μέλλον, θα βοηθήσει να ελέγξει καλύτερα τη συμπεριφορά κάποιου σε παρόμοιες περιπτώσεις.



Σχήμα 1.6. Συμβολική αντιπροσώπευση των αλληλεπιδράσεων μεταξύ του 'διανοητικού «και του πραγματικού κόσμου ' (σημείο 'εδώ και τώρα') με μορφή κλειστών βρόγχων

Η νοημοσύνη, επομένως, ορίζεται επιτρέποντας να καταλάβει βαθιά τις διαδικασίες και ο τρόπος που το «ίδιο σώμα» μπορεί να το εκμεταλλευτεί αυτό. Δεδομένου ότι οι κατάλληλες αντιδράσεις εξαρτώνται από την κατάσταση που αντιμετωπίζεται, αναγνωρίζοντας τις καταστάσεις νωρίς και σωστά και ξέροντας τι να κάνει σε αυτές τις περιπτώσεις (λήψη αποφάσεων) είναι στον πυρήνα της νοημοσύνης. η αίσθηση της σταθερής εκμάθησης, όλες οι προκύπτουσες ενέργειες ελέγχονται και χρησιμοποιούνται για να βελτιωθούν οι εσωτερικές αντιπροσωπεύσεις για την καλύτερη χρήση στο μέλλον. Το σχήμα 1.6 παρουσιάζει μια συμβολική αντιπροσώπευση της γενικής αλληλεπίδρασης μεταξύ του (άτομου) και «του διανοητικού κόσμου» ως δραστηριότητα χειρισμού στοιχείων σε μια λάθος πρόβλεψη- ανατροφοδοτεί το βρόχο. Εκτείνεται μέρος του χρονικού άξονα (οριζόντια γραμμή) και του «πραγματικού κόσμου» που αντιπροσωπεύεται από το χωρικό σημείο «εδώ» (όπου είναι οι αισθητήρες). Το χωρικό σημείο «εδώ», με το τοπικό περιβάλλον του, και το χρονικό σημείο «τώρα», όπου η αλληλεπίδραση του θέματος με ο πραγματικό κόσμο πραγματοποιείται, είναι το μόνο 4-D σημείο για το αυτόνομο σύστημα που πρέπει να την κάνει πραγματική εμπειρία. Όλες οι αλληλεπιδράσεις με τον κόσμο πραγματοποιούνται «εδώ και τώρα» (δείτε το κεντρικό κιβώτιο). Ο

υπόλοιπος κόσμος, οι επεκτάσεις του στο χώρο και το χρόνο, είναι μεμονωμένα κατασκευάσματα στο «διανοητικό κόσμο» για «να κατανοήσει» τα στοιχεία των αισθητήρων και τις ιδιότητες σταθερότητάς του που παρατηρούνται χωριστά, και ως κοινωνική προσπάθεια μεταξύ πρακτόρων ικανών για κατάλληλη ανταλλαγή πληροφοριών.

Οι ευρέως ποικίλες ερμηνείες των παρόμοιων γεγονότων στους διαφορετικούς ανθρώπινους πολιτισμούς είναι μια ένδειξη της ευρείας ποικιλίας των σχετικά σταθερών πιθανών συστημάτων ερμηνείας. Τα βιολογικά συστήματα έπρεπε να αρχίσουν από την αρχή οι κοινωνικές ομάδες ήταν ικανοποιημένες με τις ερμηνείες, οι οποίες τους επέτρεψαν να ρυθμίσουν τις ζωές τους αντίστοιχα. Οι ασυνέπειες γίνονταν αποδεκτές, γενικά, εάν οι εξηγήσεις μπορούσαν να βρεθούν ικανοποιητικές. Η πρόοδος προς τα συνεπέστερα γενικά πρότυπα "του κόσμου" ήταν αργή και πήρε χιλιετίες για την ανθρωπότητα.

Οι φυσικές επιστήμες ως συγκεκριμένη προσπάθεια ατόμων από διαφορετικές πολιτιστικές κοινότητες που ψάχνουν μια συνεπή περιγραφή «του κόσμου» και που προσπαθούν να αποφύγουν τις προκαταλήψεις που επιβάλλονται από τους συγκεκριμένους πολιτισμούς τους έχουν βρει ένα σύνολο «παγκόσμιων προτύπων», τα οποία παράγουν πολύ καλές προβλέψεις. Ειδικά κατά τη διάρκεια των τελευταίων τριών αιώνων μετά από την ανακάλυψη του διαφορικού λογισμού από Leibniz και Newton και ο πιο συγκεκριμένα κατά τη διάρκεια των τελευταίων πέντε δεκαετιών μετά από τους ηλεκτρονικούς υπολογιστές που έγιναν ικανοί να λύσουν τα προκύπτοντα σύνολα εξισώσεων με την γενικότερη μορφή τους, αυτές οι ικανότητες πρόβλεψης πέταξαν στα ύψη.

Μπροστά από αυτό το υπόβαθρο, φαίνεται λογικό να εξοπλιστούν τα σύνθετα τεχνικά συστήματα με μια ομοίως προηγμένη ακολουθία αισθητήρων όπως έχουν οι άνθρωποι, με μια ερμηνεία υπόβαθρου στο τελευταίο επίπεδο της ανάπτυξης στις φυσικές επιστήμες και στην εφαρμοσμένη μηχανική. Πρέπει να καλύψει μια (για όλους τους πρακτικούς λόγους) σωστή περιγραφή των φαινομένων που είναι άμεσα αισθητά με τα συστήματα αισθητήρων του. Αυτό περιλαμβάνει τους όρους φωτισμού μέσω του ήλιου και του φεγγαριού, οι καιρικές συνθήκες όπως αντιμετωπίζονται ανάλογα με τον χρόνο και πέρα από τις διαφορετικές θέσεις στην υδρόγειο, και βασικά φυσικά αποτελέσματα που εξουσιάζουν τη μετακίνηση όπως η γήινη βαρύτητα, την ξηρή και υγρή τριβή, καθώς επίσης και τις πηγές δύναμης και πληροφοριών. Όσον αφορά τους τελευταίους, τα τεχνικά συστήματα έχουν το πλεονέκτημα να μετρήσουν άμεσα τη θέση τους στη σφαίρα μέσω του «συστήματος παγκόσμιας πλοήγησης» (GPS). Αυτό είναι ένα πρόσφατο επίτευγμα της ανθρώπινης τεχνολογίας μικρότερο από δύο δεκαετίες ζωής, η οποία είναι βασισμένη σε μια συλλογή από τεχνητούς Γήινους δορυφόρους που περιστρέφονται σε κατάλληλα επιλεγμένες τροχιές.

Με αυτές τις πληροφορίες και με τους ψηφιακούς χάρτες των ηπείρων, τεχνικώς τα αυτόνομα συστήματα θα έχουν την ικανότητα ναυσιπλοΐας μακρύτερα από τα βιολογικά συστήματα. Η προσθήκη των παντός καιρού αισθητήρων απεικόνισης στη σειρά κυμάτων χιλιοστόμετρου θα καταστήσει αυτά τα συστήματα αληθινά παγκόσμια όσον αφορά το χώρο και το χρόνο στο μέλλον.

3.11 Βασικές Σχέσεις

Η όραση είναι μια διαδικασία στην οποία οι χρονικά μεταβαλλόμενες τιμές έντασης και χρώματος στο πλάνο εικόνας πρέπει να ερμηνευθούν ως διαδικασίες στο πραγματικό κόσμο που συμβαίνουν σε 3-D διάστημα με την πάροδο του χρόνου. Κάθε εικόνα των σημερινών τηλεοπτικών καμερών περιέχει περίπου μισό εκατομμύριο pixels. Είκοσι πέντε (ή τριάντα) από αυτές τις εικόνες λαμβάνονται ανά δευτερόλεπτο. Αυτό υψηλό το ποσοστό πλαισίων εικόνας έχει επιλεγεί για να προκαλέσει την εντύπωση της σταθερής και συνεχούς κίνησης στους ανθρώπινους παρατηρητές. Εάν κάθε εικόνα ήταν απολύτως διαφορετική από την άλλη, όπως σε μια επίδειξη φωτογραφικών διαφανειών με τα στιγμιότυπα από τις σκηνές που λήφθηκαν σε διαφορετικό χρόνο και χώρο, και επιδείχθηκαν σαν τηλεοπτικό σποτ σαν ταινία, κανένας δεν θα καταλάβαινε τι παρουσιάζεται. Η συνεχής ανάπτυξη της δράσης που καθιστά τις ταινίες κατανοητές λείπει.

Αυτό που πρέπει να καταστεί σαφές είναι ότι δεν είναι το περιεχόμενο κάθε ενιαίας εικόνας, η οποία αποτελεί τις πληροφορίες που μεταβιβάζονται στον παρατηρητή, αλλά η σχετικά αργή ανάπτυξη της κίνησης και της δράσης με την πάροδο του χρόνου. Η κοινή μονάδα του 1ος δευτέρου καθορίζεται ως το χρονικό ψήφισμα που είναι επαρκέστερο για την ανθρώπινη κατανόηση. Κατά συνέπεια, τα σχετικά αργά κινούμενα αντικείμενα και τα αργά στο να ενεργήσουν θέματα είναι οι ουσιαστικοί μεταφορείς των πληροφοριών σε αυτό το πλαίσιο. Μια σφαίρα που πετά μέσω της σκηνής μπορεί να γίνει αντιληπτή μόνο από την επίδραση που έχει σε άλλα αντικείμενα ή θέματα. Επομένως, η ικανότητα της οπτικής αντίληψης είναι βασισμένη στη δυνατότητα να παραχθούν οι εσωτερικές αντιπροσωπεύσεις των χρονικών διαδικασιών στα τρισδιάστατα διαστήματα και το χρόνο με τα αντικείμενα και τα θέματα (σύνθεση), τα οποία υποστηρίζονται από τις ροές χαρακτηριστικών γνωρισμάτων από τις ακολουθίες εικόνας (ανάλυση). Αυτό είναι μια ζωτική διαδικασία με γενικώς γνωστά στοιχεία, και οι δύο παράμετροι που καθορίζουν πραγματικά την τρισδιάστατη μορφή και τη χρονική ιστορία των κρατικών μεταβλητών των αντικειμένων που παρατηρούνται πρέπει να καθοριστούν από την εικόνα.

Σε αυτήν την διαδικασία «ανάλυσης από τη σύνθεση» που επιλέγεται στην 4-D προσέγγιση στη δυναμική όραση, οι εσωτερικές αντιπροσωπεύσεις στη διαδικασία ερμηνείας έχουν τέσσερις ανεξάρτητες μεταβλητές: τρία ορθογώνια διαστημικά συστατικά (τρισεδιάστατο διάστημα) και χρόνος. Για τους κοινούς μας στόχους στο φυσικό μας (μεσοδιάστημα, αυτό δεν είναι πάρα πολύ μικρό και όχι και πάρα πολύ μεγάλο) περιβάλλον, αυτές οι μεταβλητές είναι γνωστό ότι είναι αρκετά αντιπροσωπευτικές υπό την κλασσική nonrelativistic έννοια.

Όπως αναφέρεται στην εισαγωγή, οι γρήγορες ακολουθίες εικόνας περιέχουν ένα κομμάτι πλεονασμού, δεδομένου ότι μόνο οι μικρές αλλαγές εμφανίζονται από το ένα πλαίσιο στο επόμενο, γενικά οι ογκώδεις οργανισμοί παρουσιάζουν συνοχή στην κίνησή τους. Οι χαρακτηριστικές συχνότητες της ανθρώπινης και πιο ζωικής κίνησης είναι λιγότερο από μερικές ταλαντώσεις ανά δευτερόλεπτο (Hz), έτσι στο τηλεοπτικό ποσοστό, τουλάχιστον μια δωδεκάδα πλαισίων εικόνας λαμβάνονται ανά περίοδο ταλάντωσης. Σύμφωνα με την επιλεγείσα θεωρία στοιχείων, αυτό επιτρέπει την καλή αναγνώριση των δυναμικών παραμέτρων στο διάστημα συχνότητας (χρονικές σταθερές, παραγόμενες συχνότητες, και απόσβεση). Έτσι, ο στόχος της οπτικής δυναμικής κατανόησης σκηνής μπορεί να περιγραφεί ως εξής:

Εξετάζοντας τις 2-D σειρές στοιχείων που παράγονται από αρκετές χιλιάδες αισθητήρες, τα στοιχεία βρίσκουν μια διανομή των αντικειμένων στο πραγματικό κόσμο και της σχετικής κίνησής τους. Τα στοιχεία αισθητήρων τακτοποιούνται σε μια ομοιόμορφη σειρά στο τσιπ, συνήθως. Επί των οχημάτων, δεν μπορεί να υποτίθεται ότι ο προσανατολισμός αισθητήρων είναι γνωστός εκ των προτέρων ή ακόμα και στάσιμος. Εντούτοις, οι αδρανείς αισθητήρες για τα γραμμικά τμήματα επιτάχυνσης και τα περιστροφικά ποσοστά είναι διαθέσιμοι για την αντίληψη της αυτόνομης - κίνησης.

Είναι αμέσως σαφές ότι γνώση για τις κατηγορίες αντικειμένου και ο τρόπος τους τα ορατά χαρακτηριστικά γνωρίσματα χαρτογραφούνται στο πλάνο εικόνας είναι μεγάλου ενδιαφέροντος για την κατανόηση ακολουθίας της εικόνας. Αυτά τα αντικείμενα μπορούν να ομαδοποιηθούν στις κατηγορίες με παρόμοια λειτουργία ή/και εμφάνιση. Το σώμα του οχήματος που φέρει τους αισθητήρες και η παροχή των μέσων για τη μετακίνηση είναι εξαιρετικά σημαντική, φυσικά. η μεγάλη περιγραφή της προηγούμενης πρότασης θα βραχυθθεί από τον όρο: «σώμα». Για να καταλάβει την κίνησή του άμεσα και ανεξάρτητα από την εικόνα, σήματα από άλλους αισθητήρες όπως τα οδόμετρα, τους αδρανείς γωνιακούς αισθητήρες ποσοστού και τα γραμμικά επιταχύμετρα καθώς επίσης και το GPS (από την παροχή «συστημάτων παγκόσμιας πλοήγησης» γεωγραφικός συντονισμός) που χρησιμοποιούνται ευρέως.

Τα σημεία στοιχείων εικόνας δεν φέρνουν καμία άμεση πληροφορία για την απόσταση των πηγών φωτός τους, που έχουν υποκινηθεί ότι το σήμα αισθητήρων είναι στο πραγματικό κόσμο η τρίτη διάσταση (σειρά) χάνεται εντελώς σε μια ενιαία εικόνα (εκτός ίσως για τη μείωση έντασης πέρα από τις μεγαλύτερες αποστάσεις). Επιπλέον, δεδομένου ότι οι διαταραχές μπορούν να ακυρώσουν το περιεχόμενο πληροφοριών ενός ενιαίου pixel σχεδόν τελείως, χρήσιμα χαρακτηριστικά γνωρίσματα εικόνας αποτελούν τα σήματα από τις ομάδες στοιχείων αισθητήρων όπου οι τοπικές διαταραχές τείνουν να είναι ισοπεδωμένες. Στα βιολογικά συστήματα, αυτά είναι τα δεκτικά πεδία στα τεχνικά συστήματα, αυτές είναι μάσκες αξιολόγησης των διάφορων μεγεθών. Αυτό επιτρέπει τώρα μια ακριβέστερη δήλωση του στόχου της όρασης:

Για τα απλά ογκώδη αντικείμενα (π.χ., μια πέτρα, ο ήλιος και το φεγγάρι μας) και τα κατασκευασμένα από τον άνθρωπο οχήματα, τα καλά «δυναμικά πρότυπα» που περιγράφουν τους περιορισμούς κινήσεων είναι γνωστά πολύ συχνά. Για να περιγράψουν τη σχετική ή απόλυτη κίνηση των αντικειμένων ακριβώς, τα κατάλληλα ισότιμα συστήματα αναφοράς πρέπει να εισαχθούν. Σύμφωνα με την ευρεία κλίμακα του διαστήματος προσιτή από την εικόνα, ορισμένες κλίμακες της αντιπροσώπευσης είναι συμφέρουσες:

- τα στοιχεία αισθητήρων έχουν τις διαστάσεις στη σειρά μικρόμετρου (μ.μ.).
- Οι άνθρωποι αναπτύσσουν δραστηριότητες της τάξεως του μέτρου(m): επίτευξη του χώρου, μοναδιαίου βήματος (μέγεθος σωμάτων).
- Για τα βλήματα και τα γρήγορα οχήματα, η σειρά των άμεσων αντιδράσεων επεκτείνεται σε αρκετά μέτρα ή χιλιόμετρα (χλμ).
- Οι αποστολές μπορούν να εκταθούν σε αρκετές χιλιάδες χιλιόμετρα, ακόμη και ένα τρίτο ως και το μισό ολόκληρης της υδρόγειου με άμεση πτήση.
- Η διαστημική πτήση και ο φωτισμός από τον ήλιο και το φεγγάρι μας επεκτείνονται μέχρι και 150 εκατομμύρια χλμ ως χαρακτηριστικό βεληγεκές (ακτίνα της γήινης τροχιάς).
- Τα ορατά αστέρια είναι αρκετά πέρα από αυτές τις αποστάσεις (όχι ενδιαφέροντος εδώ).

Είναι πιθανό να βρεθεί ένας ενιαίος τύπος αντιπροσώπευσης που να καλύπτει όλες τις μονάδες μέτρησης; Αυτό δεν είναι βεβαίως επιτεύξιμο με τις μεθόδους που χρησιμοποιούν τα πλέγματα των διαφορετικών κλιμάκων όπως συχνά γίνεται στη «τεχνητή νοημοσύνη» - προσεγγίσεις. Μάλλον, η προσέγγιση που αναπτύσσεται στην ηλεκτρονική γραφιστική με τις ομαλοποιημένες περιγραφές μορφής και τους γενικούς παράγοντες ξελεπιάσματος είναι ο πρωταρχικός υποψήφιος. Οι

ομοιογενείς συντεταγμένες όπως εισάγονται από [Roberts 1965, Blinn 1977] επίσης επιτρέπουν, εκτός από το ξελέπιασμα, ενσωματώνοντας τη διαδικασία χαρτογράφησης προοπτικής στο ίδιο πλαίσιο. Αυτό παράγει μια ενοποιημένη προσέγγιση για την όραση υπολογιστών και την ηλεκτρονική γραφιστική εντούτοις, στην όραση υπολογιστών, πολλές από τις μεταβλητές που εισάγουν τις ομοιογενείς μήτρες μετασχηματισμού είναι τα άγνωστα του προβλήματος. Μια άμεση εφαρμογή των μεθόδων από την ηλεκτρονική γραφιστική είναι έτσι αδύνατη, δεδομένου ότι η αντιστροφή της προβολής προοπτικής είναι ένα έντονα μη γραμμικό πρόβλημα με την ανάγκη να ανακτηθεί ένα διαστημικό συστατικό που χάνεται εντελώς στη χαρτογράφηση.

Η εισαγωγή των ισχυρών περιορισμών στη χρονική εξέλιξη (των τρισδιάστατων) χωρικών τροχιών, εντούτοις, επιτρέπει την ανάκτηση μέρους των πληροφοριών που χάνονται με την εκμετάλλευση των πρώτης - διαταγής παραγώγων. Αυτό είναι το μεγάλο πλεονέκτημα των χωροχρονικών προτύπων και της επαναλαμβανόμενης εκτίμησης πέρα από την άμεση αντιστροφή προοπτικής (υπολογιστική όραση). Η ιακωβιανή αυτής της προσέγγισης που συζητείται σε όλο το κείμενο διαδραματίζει έναν ζωτικής σημασίας ρόλο στην 4-D προσέγγιση και στην κατανόηση ακολουθίας εικόνας.

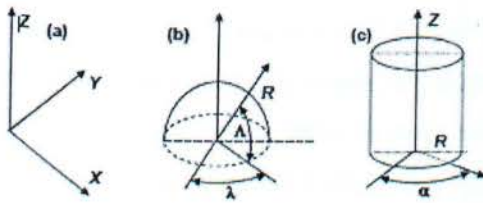
Προτού να μπορέσει να εκτιμηθεί πλήρως αυτό, η αλυσίδα των ισότιμων μετασχηματισμών από μια αντικείμενο-κεντροθετημένη διανομή χαρακτηριστικών γνωρισμάτων για κάθε αντικείμενο στο τρισδιάστατο διάστημα στην αποθήκευση της 2-D εικόνας στη μνήμη υπολογιστών πρέπει να γίνουν κατανοητές.

3.12 Τρισδιάστατος χώρος και χρόνος

Κάθε σημείο στο χώρο μπορεί να διευκρινιστεί πλήρως με το δόσιμο τριών συντεταγμένων σε ένα καθορισμένο με σαφήνεια πλαίσιο αναφοράς. Αυτό το πλαίσιο αναφοράς μπορεί να είναι ένα «καρτεσιανό» σύστημα με τρεις άξονες κατεύθυνσης (σχήμα 2.1a), ένα σφαιρικό (πολικό) σύστημα με μια (ακτινωτή) απόσταση και δύο γωνίες (αριθμός 2.1b), ή ένα κυλινδρικό σύστημα ή ως μίγμα και οι δύο, με δύο άξονες και μια γωνία (σχήμα 2.1c).

Το βασικό πλάνο αναφοράς επιλέγεται συνήθως για να παραγάγει την απλούστερη περιγραφή του προβλήματος: Για να περιγράψουν τη μορφή των αντικειμένων, τα συμμετρικά πλάνα προτιμώνται παραδείγματος χάριν, το σχήμα 2.2 παρουσιάζει ορθογώνιο κιβώτιο με μήκος L , το πλάτος B και το ύψος H . Το συνολικό κέντρο βάρους St βρίσκεται λαμβάνοντας υπόψη τη διατομή δύο διαγώνιων. Μπορεί να εξεταστεί το κιβώτιο σαν να περιβάλλει έναν όχημα δρόμου ; τυπικά, το L είναι ο

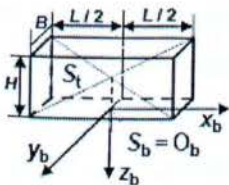
μεγαλύτερο και η κατεύθυνσή του καθορίζει την κατεύθυνση του οχήματος (αντικειμένου). Επομένως, η κεντρική γραμμή από τη χαμηλότερη επιφάνεια επιλέγεται



Σχήμα 2.1. Βασικά ισότιμα συστήματα (καίσιο): (α) καρτεσιανό , (β) σφαιρικό , (γ) κυλινδρικό

σαν κατεύθυνση ενός σώματος σε ένα καρτεσιανό σύστημα (x_b, y_b, z_b) με την αρχή του στο σημείο O_b στην προβολή S_b του S_t επάνω στο επίπεδο πλάνο.

Για να περιγράψει την κίνηση σε όλον τον εξουσιάζοντα τομέα της βαρύτητας, το πλάνο της αναφοράς μπορεί να περιέχει και τη βαρύτητα και το διάνυσμα ταχύτητας με την προέλευση από το κέντρο βαρύτητας του κινούμενου αντικειμένου. Το «οριζόντιο» επίπεδο κανονικό στο διάνυσμα βαρύτητας έχει επίσης μερικά πλεονεκτήματα, ειδικά για τη δυναμική οχημάτων δεδομένου ότι κανένα τμήμα βαρύτητας δεν έχει επιπτώσεις στην κίνηση του.



Σχήμα 2.2. Αντικειμενοστρεφές ισότιμο σύστημα για ένα ορθογώνιο κουτί

Εάν ένα άκαμπτο αντικείμενο κινηθεί στο τρισδιάστατο χώρο , αυτό είναι προς συμφέρον μας για να περιγράψουμε τη μορφή του αντικειμένου σε ένα αντικειμενοστρεφές πλαίσιο αναφοράς με την προέλευσή της στο κέντρο (ενδεχομένως ακόμη και το κέντρο βαρύτητας) ή κάποιο άλλο κατάλληλο, εύκολα ευπροσδιόριστο σημείο (πιθανώς στην επιφάνεια του). Στο σχήμα 2.2, η μορφή του ορθογώνιου κιβωτίου καθορίζεται από τα μήκη των πλευρών του $L, B,$ και $H.$ Η προέλευση επιλέγεται στο κέντρο του της κάτω πλευράς του (S_b). Εάν η θέση και ο προσανατολισμός αυτού του κιβωτίου πρέπει να περιγραφεί σχετικά με ένα άλλο αντικείμενο, το πλαίσιο της αναφοράς που δίνεται στο σχέδιο πρέπει να αφορά ένα άλλο ανεξάρτητα καθορισμένο αντικείμενο από τρεις μεταφράσεις και τρεις περιστροφές, γενικότερα.

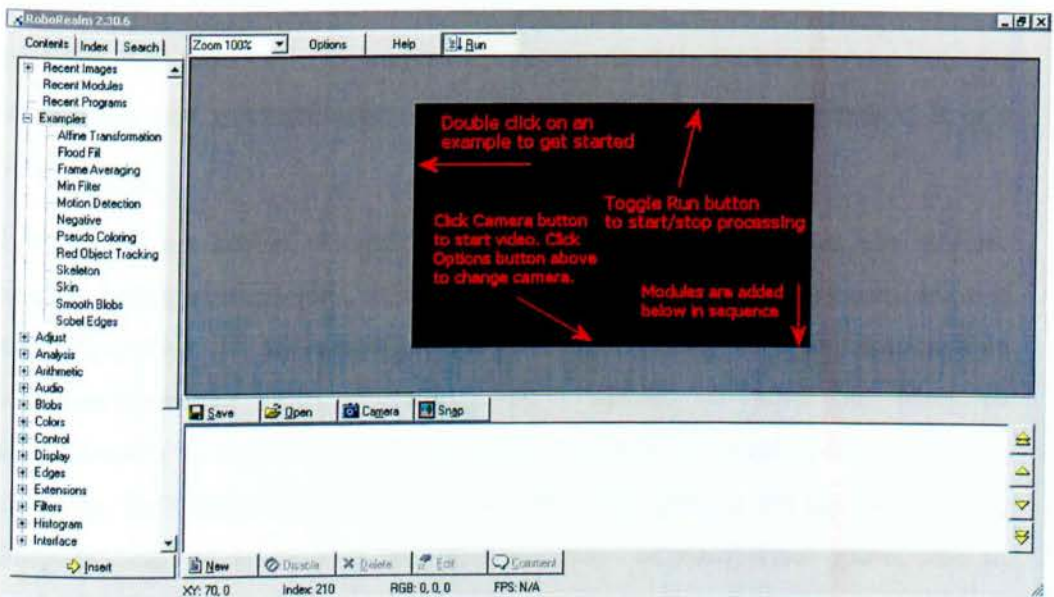
Για να περιγράψει τη μορφή αντικειμένου (κιβώτιο) στο νέο πλαίσιο της αναφοράς, ένας μετασχηματισμός των συντεταγμένων για όλα τα σημαντικά σημεία που καθορίζουν τη μορφή πρέπει να εκτελεσθεί. Για το ορθογώνιο κιβώτιο, αυτά είναι τα οκτώ σημεία γωνιών του που βρίσκονται στο $\pm L/2$ και το $\pm B/2$ για $Z_b = 0$ και $-H$. Οι ευθείες άκρες του κιβωτίου παραμένουν γραμμικές συνδέσεις μεταξύ αυτών των σημείων. [Η επιλογή των ισότιμων αξόνων έχει εκτελεσθεί σύμφωνα με το διεθνές πρότυπο για τα αεροδιαστημικά οχήματα. Χ είναι στην τυποποιημένη κατεύθυνση της κίνησης, το Χ και το Ζ είναι στο πλάνο της συμμετρίας οχημάτων, και το Υ ολοκληρώνει ένα δεξιόστροφο το σύνολο συντεταγμένων. Το διάνυσμα στο χαμηλότερο άκρο έξω από το σώμα απλουστεύει τις μετρήσεις και είναι ιδιαίτερα εύχρηστο για τα επίγεια οχήματα, όπου το περιβάλλοντα κιβώτιο αγγίζει το έδαφος λόγω της βαρύτητας, στην κανονική περίπτωση. Η μέτρηση του ύψους (ανύψωση) θετικά προς τα πάνω απαιτεί μια αλλαγή σημαδιών από την θετική κατεύθυνση Ζ (κατεύθυνση του διανύσματος βαρύτητας στην κανονική πτήση επιπέδων). Για αυτόν τον λόγο, κάποια διεθνή πρότυπα για τα επίγεια οχήματα περιστρέφουν το σύστημα συντεταγμένων 180° γύρω από τον άξονα των Χ (το Ζ προς τα πάνω και το Υ αριστερά).

Γενικά, οι μετασχηματισμοί συντεταγμένων μεταξύ δύο συστημάτων στο τρισδιάστατο χώρο έχει τρία μεταφραστικά και τρία περιστροφικά συστατικά. Στη δεκαετία του '70, όταν αυτοί οι τύποι διαδικασιών έγιναν κοινοί στην ηλεκτρονική γραφιστική, μαζί με τη χαρτογράφηση προοπτικής ως τελική φάση της απεικόνισης για τους ανθρώπινους παρατηρητές, αποκαλούμενη οι «ομοιογενείς συντεταγμένες» εισήχθησαν [Roberts 1965, Blinn 1977]. Επιτρέπουν την αντιπροσώπευση όλων των μετασχηματισμών που απαιτούνται από τις εξισώσεις μετασχηματισμού του μεγέθους 4×4 με διαφορετικές καταχωρήσεις. Ειδικοί μικροεπεξεργαστές έχουν αναπτυχθεί στη δεκαετία του '70 που μας επιτρέπουν να χειριστούμε αυτές τις διαδικασίες αποτελεσματικά. Οι εκτεταμένες αλληλουχίες διάφορων διαδοχικών μετασχηματισμών αποδεικνύονται προϊόντα αυτών των εξισώσεων, για να επιτύχουμε τη σε πραγματικό χρόνο απόδοση για τις ρεαλιστικές προσομοιώσεις με οπτική ανατροφοδότηση και οι άνθρωποι χειριστές στο βρόχο, αυτές οι διαδικασίες έχουν διαμορφώσει το σχέδιο υλικού ηλεκτρονικής γραφιστικής (ο υπολογιστής παράγαγε τις εικόνες, CGI [Foley και λοιποί. 1990]

Κεφάλαιο 4

Το λογισμικό Roborealms

Το Roborealms είναι μια εφαρμογή που χρησιμοποιείται στην όραση υπολογιστών, την ανάλυση και επεξεργασία εικόνας και στα συστήματα ρομποτικής όρασης.



Βασίζεται σε ένα σύστημα server, του οποίου οι λειτουργίες είναι διαθέσιμες μέσω κλήσεων απομακρυσμένων διαδικασιών (Remote procedure calls). Υπάρχει διαθέσιμη διεπαφή προγραμματισμού εφαρμογών (Application Programming Interface – API) για διάφορες γλώσσες προγραμματισμού. Το Roborealms υποστηρίζει μια πληθώρα ρομποτικών και προγραμματιστικών πλατφορμών, όπως το Microsoft Robotics Developer Studio και Lego Mindstorms NXT. Επίσης είναι ενσωματωμένοι αρκετοί αλγόριθμοι επεξεργασίας εικόνας καθώς και άλλες λειτουργίες που χρησιμοποιούνται στο πεδίο της όρασης υπολογιστών.

Μπορεί να χρησιμοποιηθεί σαν εργαλείο της όρασης υπολογιστών για εφαρμογές στο πεδίο της ρομποτικής, καθώς περιλαμβάνει μεθόδους αναγνώρισης θέσης ενός συστήματος στον χώρο (robot localization), απλές ικανότητες πλοήγησης και επικοινωνίας με τους πιο κοινούς αισθητήρες που χρησιμοποιούνται στην ρομποτική.

Στο κεφάλαιο αυτό παρουσιάζουμε ενδεικτικά κάποια module της εφαρμογής καθώς και αυτά που θα χρησιμοποιηθούν στο συστημά μας.

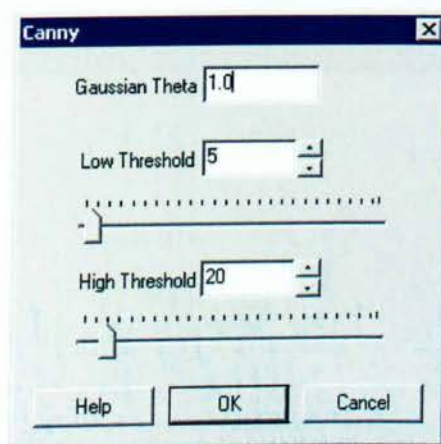
4.1 Ανίχνευση Ακμών (EDGES)

Canny : Το module αυτό χρησιμοποιεί τον Ανιχνευτή ακμών Canny, που παρουσιάστηκε από τον John Canny στην εργασία του “ A Computational Approach to Edge Detection “ . Ο ανιχνευτής ακμών Canny θεωρείται ένας από τους καλύτερους που χρησιμοποιείται αυτή την στιγμή. Ο βασικός αλγόριθμος είναι ο εξής:

1. Θολώστε την εικόνα ελαφρώς. Αυτό εξαλείφει τα αποτελέσματα των τυχαίων μαύρων ή άσπρων πιξελ (δηλ. πιξελ θορύβου) που μπορούν να έχουν επιπτώσεις στα επόμενα στάδια. Το θόλωμα ολοκληρώνεται χρησιμοποιώντας μια “γκαουσιανή θάμβωση”(Gaussian blur). Εντούτοις, θα πρέπει να είστε σε θέση να χρησιμοποιήσετε ένα Μέσο φίλτρο (Mean Filter) για να ολοκληρώσετε την ίδια επίδραση. Το Μέσο φίλτρο μπορεί να εκτελεσθεί γρηγορότερα από μια γκαουσιανή θάμβωση, έτσι το προτιμούμε για τις εφαρμογές σε πραγματικό χρόνο, που το υπολογιστικό πλεονέκτημα μπορεί να είναι σημαντικό. Τα φίλτρα θα αναλυθούν σε επόμενη ενότητα.
2. Υπολογίζουμε τις παραγώγους των X και Y της εικόνας.. Παρόμοια με τον ανιχνευτή ακμών Sobel οι παράγωγοι των X και Y υπολογίζονται απλά με την αφαίρεση των τιμών των δύο περιβαλλόντων γειτονικών πίξελ ανάλογα με τα οποία ο άξονας υπολογίζεται (αριστερά και δεξιά για την παράγωγο X και την κορυφή και το κατώτατο σημείο για το παράγωγο Y).
3. Από τις παραγώγους των X και Y μπορούμε να υπολογίσουμε το πλάτος των ακμών που δηλώνει βασικά τη δύναμη της ακμής που ανιχνεύεται στο τρέχον σημείο.

4. Εάν επρόκειτο να εξετάσουμε την εικόνα ακμών σε αυτό το σημείο θα μας φαινόταν παρόμοια με το φίλτρο ακμών Sobel. Δηλαδή, οι άκρες θα ήταν παχιές σε πολλούς τομείς της προκύπτουσας εικόνας ακμών. Για να λεπτύνουμε αυτές τις ακμές πρέπει να εφαρμόσουμε ένα μη-μεγιστο φίλτρο που θα κρατήσει μόνο τις ακμές που είναι η κορυφή ή η κορυφογραμμή μιας ανιχνευμένης ακμής. Το μη-μεγιστο φίλτρο εφαρμόζεται πρώτα αναλύοντας ποιά είναι η κλίση της ακμής και να την χρησιμοποιήσει ώστε να αντιληφθεί ποιά είναι η διεύθυνση της συγκεκριμένης ακμής. Μπορούμε έπειτα να υποδείξουμε την κάθετο δύο ακμών στην τρέχουσα ακμή για να καθορίσετε εάν το τρέχον πιξελ ανήκει σε μια κορυφογραμμή ακμών ή όχι. Εάν το τρέχον πιξελ είναι σε μια κορυφογραμμή το μέγεθός του θα είναι μεγαλύτερο από των δύο μεγεθών των κάθετων πιξελ (σκεφτείτε το περπάτημα σε μια κορυφογραμμή βουνών με το έδαφος που μειώνεται και στις δύο πλευρές του πλαγιας). Η χρησιμοποίηση αυτής της ανίχνευσης κορυφογραμμών και της επακόλουθης αποβολής θα οδηγήσει στις πολύ λεπτές ευθυγραμμισμένες ακμές.

5. Μόλις λεπτύνουν οι ακμές το τελευταίο βήμα είναι να θέσουμε το κατώτατο όριο (threshold) των ανιχνευμένων ακμων. Τα μεγέθη ακμών πάνω από το ανώτερο κατώτατο όριο συντηρούνται. Οι ακμές κάτω από το ανώτερο κατώτατο όριο αλλά επάνω από το χαμηλότερο κατώτατο όριο συντηρούνται ΜΟΝΟ εάν συνδέονται με τις ακμές που είναι επάνω από το ανώτερο κατώτατο όριο. Αυτή η διαδικασία είναι γνωστή ως υστέρηση και επιτρέπει στις ακμές για να γίνουν μεγαλύτερες μόνο με τη χρησιμοποίηση ενός ενιαίου κατώτατου ορίου χωρίς εισαγωγή περισσότερου θορύβου στην προκύπτουσα εικόνα ακμών.

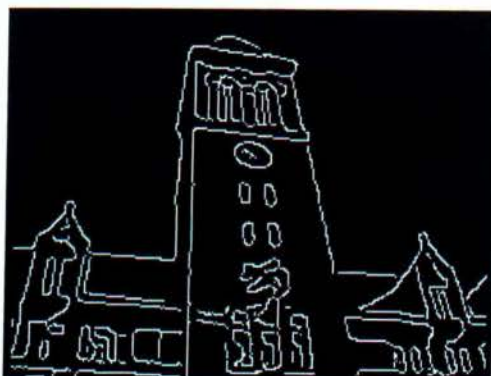


Παράδειγμα

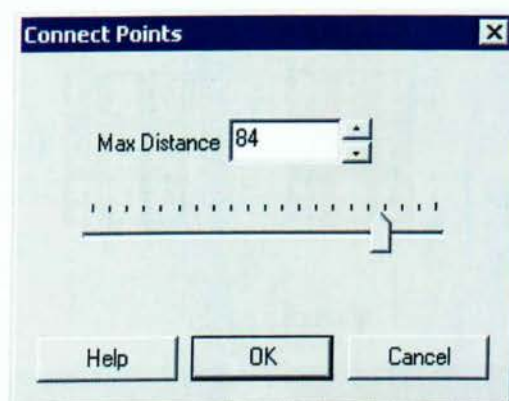
Πηγη



Ανιχνευτής Ακμών Canny

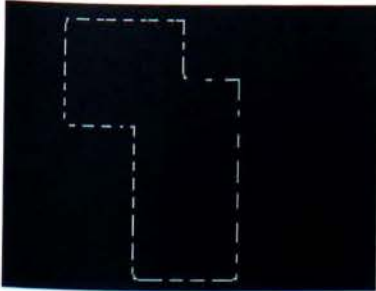


Connect Points (Σύνδεση σημείων) : Θα συνδέσει κοντινά σημεία για να δημιουργήσει συνδεδεμένες ακμές. Αυτή η λειτουργία είναι χρήσιμη μετά από thresholding όπου σημεία ακμών μπορούν να μείνουν ασύνδετα. Αντίθετα από την τεχνική υστέρησης που χρησιμοποιείται στον προηγούμενο ανιχνευτή ακμών, τα σημεία συνδέονται βασιζόμενα στην απόσταση μεταξύ τους.

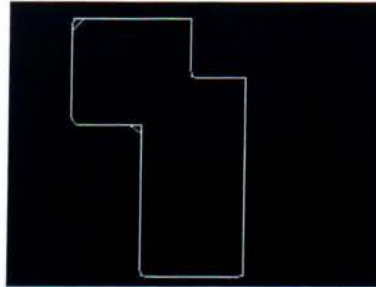


Παραδειγμα

Πηγη



Συνδεδεμένα σημεία



Outline Filter : Το φίλτρο αυτό χρησιμοποιείται για να εμφανίσει τις ακμές που βρίσκονται μέσα στην εικόνα. Παρόμοιο με ένα φίλτρο συνελίξης (convolution - βλεπε φίλτρα) αυτό το φίλτρο αντιδρά στις αλλαγές της έντασης ή του χρώματος στις ακμές. Οι αιχμηρότερες ακμές θα οδηγήσουν σε τιμές υψηλότερης έντασης. Τα αποτελέσματα είναι ομαλοποιημένα και επιδεικνύονται ως grayscale τιμές.

Prewitt Edge : Το φίλτρο ακμών Prewitt χρησιμοποιείται για να ανιχνεύσει τις ακμές εφαρμόζοντας ένα οριζόντιο και ένα κάθετο φίλτρο στη σειρά. Και τα δύο φίλτρα εφαρμόζονται στην εικόνα και αθροίζονται για να διαμορφώσουν το τελικό αποτέλεσμα. Τα δύο φίλτρα είναι βασικά φίλτρα συνελίξεων της μορφής:

Οριζόντιο Φίλτρο Κάθετο Φίλτρο

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

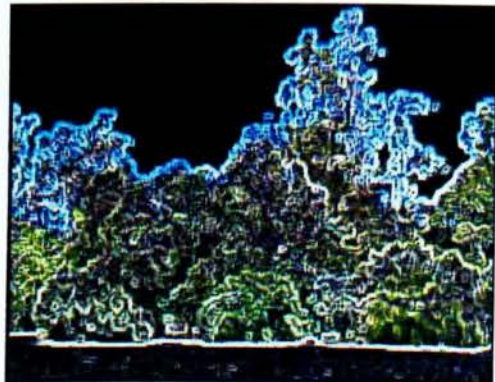
-1	0	1
-1	0	1
-1	0	1

Παράδειγμα

Πηγή



Prewitt



Roberts Edge : Πρατει ομοίως με το προηγούμενο φίλτρο. Τα δύο φίλτρα είναι βασικά φίλτρα συνελίξεων της μορφής:

Οριζόντιο Φίλτρο

1	0
0	-1

Κάθετο Φίλτρο

0	1
-1	0

Παράδειγμα

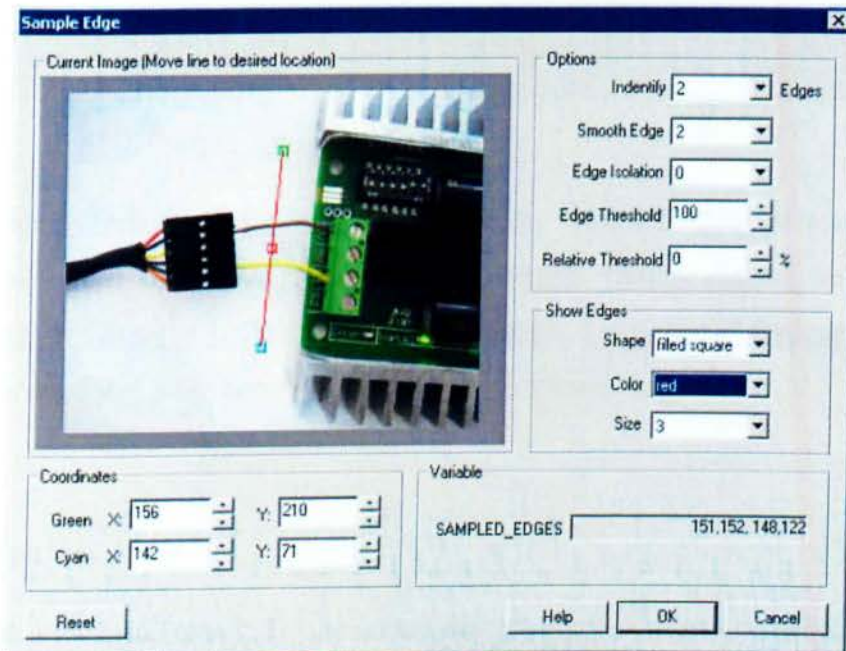
Πηγή



Roberts



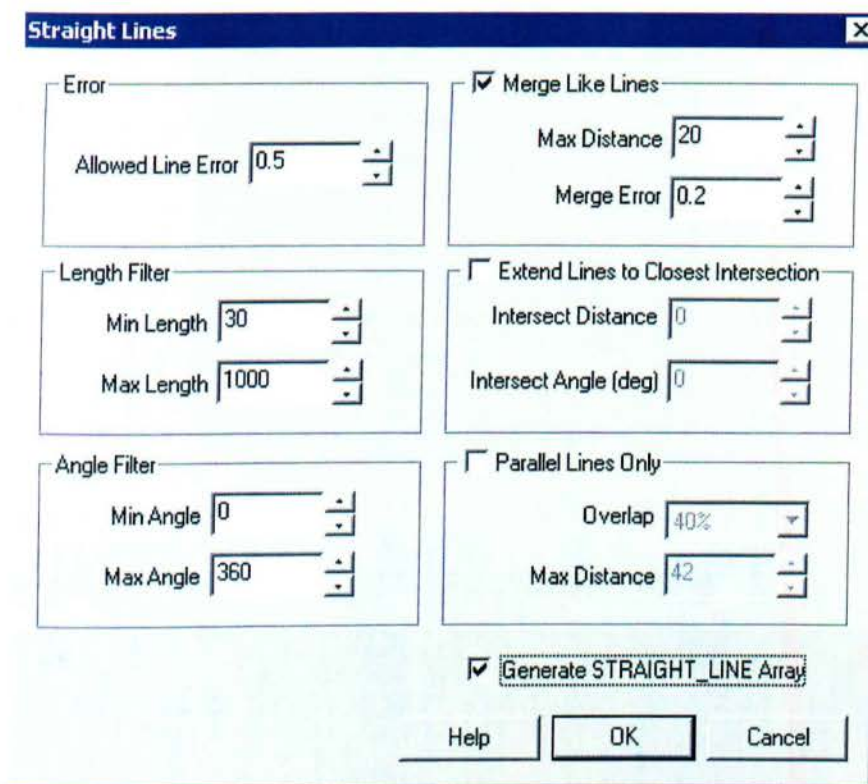
Sample edge : Η επιλογή αυτή μας παρέχει έναν τρόπο να ανιχνευθούν οι άκρες κατά μήκος μιας διευκρινισμένης πορείας. Η πορεία είναι μια ευθεία γραμμή όπου μπορούμε με το χέρι να την ορίσουμε στην εικόνα ή να ελέγξετε τη θέση της χρησιμοποιώντας μεταβλητές. Οι ακμές που ανιχνεύονται σώζονται σε μια μεταβλητή για μετέπειτα χρήση. Αυτή η επιλογή είναι χρήσιμη για την τοποθέτηση ενός ιδιαίτερου αντικειμένου μέσα σε μια εικόνα ή για να ανιχνεύσει ότι ορισμένες ακμές πρέπει να είναι παρούσες υπό μορφή εξασφάλισης συστατικής ποιότητας.



Μεταβλητές

SAMPLED_EDGES – μια x,y ακολουθία με τις συντεταγμένες των ανιχνευθέντων ακμών.

Straight Lines (Ευθείες Γραμμές) : Το module αυτό παρέχει έναν τρόπο να φιλτραριστούν οι ευθείες γραμμές από μια εικόνα. Αυτό είναι λειτουργικά παρόμοιο με τον Μετασχηματισμό Hough αλλά θα λειτουργήσει με μεγαλύτερη ταχύτητα.

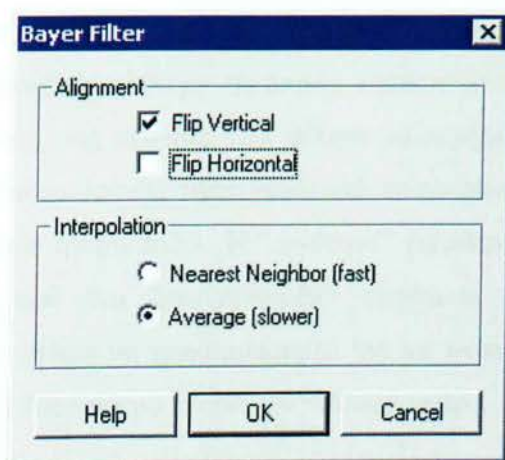


Τα module [Difference of Boxes \(DOB\)](#), έχουν τις ίδιες ιδιότητες και λειτουργούν βάσει των ίδιων αρχών με τον Ανιχνευτή Ακμών Canny. Επίσης τα module [Difference of Gaussian \(DOG\)](#), [Frei & Chen](#), [Kirsch](#), [Laplacian of Gaussian \(LOG\)](#), [Robinson](#) και [Sobel Edge](#) λειτουργούν όμοια με το [Prewitt Edge](#).

4.2 Φίλτρα

Bayer Filter (Φίλτρο Bayer) : Η χρησιμοποίηση του φίλτρου Bayer θα επαναδιατάξει τα τρέχοντα πιξελ της εικόνας από φορμα Bayer σε RGB. Το φορμα Bayer

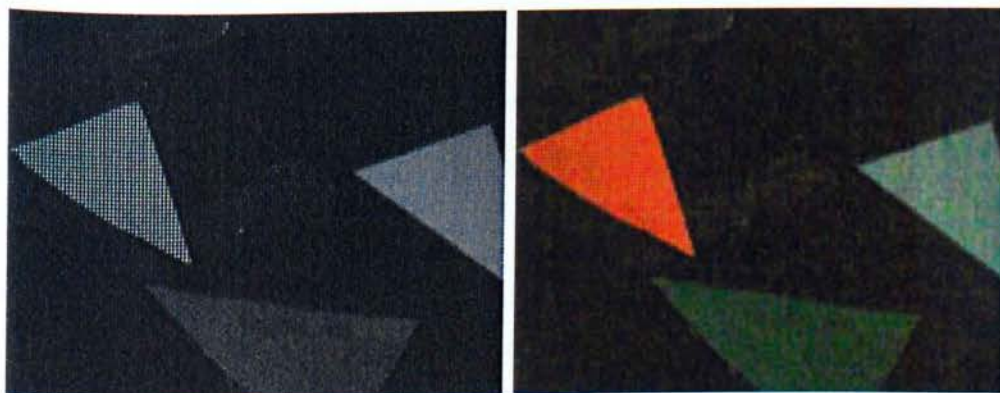
χρησιμοποιείται από τις περισσότερες CCD κάμερες ώστε να δειγματοποιήσει το κόκκινο, πράσινο και μπλε χρώμα χρησιμοποιώντας διαφορετικούς αισθητήρες χρώματος. Αυτοί οι αισθητήρες χρώματος παράγουν έπειτα μια εικόνα που συμπλέκει τα κόκκινα, πράσινα και μπλε χρώματα το ένα γύρω από το άλλο. Προκειμένου να δημιουργηθεί μια λογική αντιπροσώπευση χρώματος αυτά τα μονοκάναλα πιξελ χρώματος πρέπει να συνδυαστούν με τέτοιο τρόπο ώστε να παραχθεί μια εικόνα χρώματος.



Παράδειγμα

Πηγή Bayer

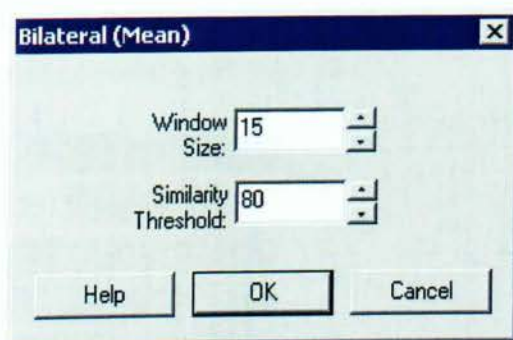
Χρωματισμένη εικόνα



Background removal (Αφαίρεση Παρασκηνίου) : Το φίλτρο αφαίρεσης παρασκηνίου προσπαθεί να απομακρύνει τις μεγάλες περιοχές της ίδιας σύστασης. Το μέγεθος παραθύρου (window size) χρησιμοποιείται για να δημιουργήσει μια μέση αξία πιξελ (RGB) με το άθροισμα όλων των τιμών των πιξελ στο παράθυρο και τη διαίρεση με το μέγεθος του παραθυρου. Αυτή η μέση αξία πιξελ συγκρίνεται έπειτα με την

τρέχουσα αξία πιξελ της εικόνας στο κέντρο της περιοχής παραθύρων/δειγμάτων. Εάν η αξία πιξελ εικόνας είναι μέσα σε ένα προκαθορισμένο όριο από το μέσο όρο, το πιξελ θεωρείται ότι είναι ένα πιξελ παρασκηνίου και γίνεται μαυρο. Τα πιξελ που δεν είναι μέσα στη μέση απόσταση στο μέσο πιξελ θεωρούνται ένα πιξελ πρώτου πλάνου και η αξία του παραμένει άθικτη. Το παράθυρο μετατοπίζεται ένα εικονοκύτταρο και η διαδικασία επαναλαμβάνεται για να επεξεργαστεί ολόκληρη η εικόνα.

Bilateral : Το διμερές (Bilateral) φίλτρο είναι πολύ παρόμοιο με το μέσο φίλτρο (Mean filter) αλλά διατηρεί καλύτερα τις ακμές υπολογίζοντας κατά μέσο όρο και θολώνοντας άλλα μέρη της εικόνας. Το φίλτρο ολοκληρώνει αυτόν τον στόχο υπολογίζοντας κατά μέσο όρο τις τιμές γύρω από το τρέχον πίξελ που είναι κοντά στην αξία χρώματος στο τρέχον πιξελ. Η “εγγύτητα” γειτονικών πιξελ στα τρέχοντα πιξελ καθορίζονται από ένα διευκρινισμένο κατώτατο όριο (threshold). Το συγκεκριμένο φίλτρο μπορεί να χρησιμοποιηθεί για να αφαιρέσει θόρυβο από μια εικόνα χωρίς να χαθούν τα όρια των περιοχών ενδιαφέροντος.



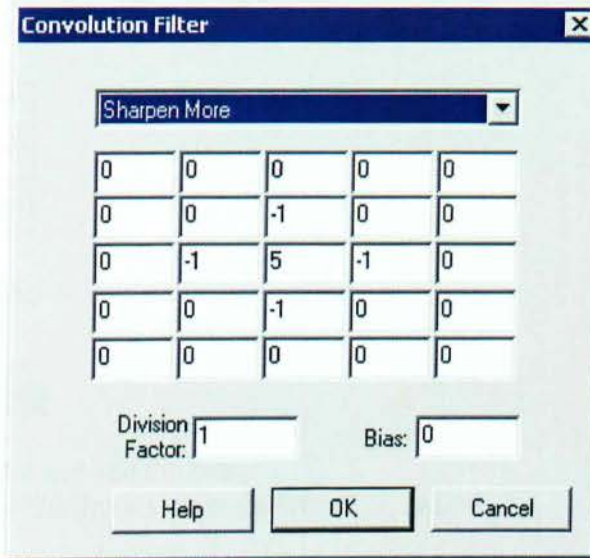
Παράδειγμα

Πηγή

Διμερές Φίλτρο



Convolution Filter (Φίλτρο Συνέλιξης) : Τα φίλτρα συνέλιξης είναι ένας πολύ καλός τρόπος να υποβληθούν σε επεξεργασία οι εικόνες για ορισμένα χαρακτηριστικά γνωρίσματα. Τα χαρακτηριστικά γνωρίσματα καθορίζονται από ένα n από τη μήτρα μ που εφαρμόζεται στην εικόνα με τον ακόλουθο τρόπο: (μόνο για λόγους του παραδείγματος χρησιμοποιούμε γκρι εικόνα)



Παράδειγμα ασπρόμαυρης εικόνας μεγέθους (10x10):

34	22	77	48	237	205	29	212	107	41
50	150	77	158	233	251	112	165	47	229
93	0	77	219	43	56	42	113	140	94
32	19	44	30	36	94	151	101	28	84
10	90	48	73	63	148	159	183	99	22
192	70	27	88	20	230	53	34	38	106
239	202	196	205	50	123	192	88	41	37
230	174	14	22	127	100	189	186	214	187
227	86	195	6	53	168	46	166	36	249
215	165	237	110	125	191	191	94	123	8

Ένα παράδειγμα συνελκτικού φίλτρου για ανίχνευση γραμμών:

-1	-1	-1
-1	8	-1
-1	-1	-1

Το σειρά=2, στήλη=2 pixel και τα γειτονικά απο την παραπάνω εικόνα:

34	22	77
50	150	77
93	0	77

Για να εφαρμόσουμε το φίλτρο συνελίξεως πολλαπλασιάζουμε τις αξίες του φίλτρου και τις εικόνας. Εργαζόμαστε για κάθε πιξελ και τα 3x3 γειτονικά του:

-1*34	-1*22	-1*77
-1*50	8*150	-1*77
-1*93	-1*0	-1*77

Αθροίζουμε τις τιμές:
 $(-34)+(-22)+(-77)+$
 $(-50)+(1200)+(-77)+$
 $(-93)+(0)+(-77) = 770$

Divide by the divisor and add the bias.
 $(770/\text{divisor})+\text{bias}=770$ (in this example divisor=1, bias=0)

Αν η νέα τιμή του πιξελ είναι > 255 set it to 255
Αν η νέα τιμή του πιξελ είναι < 0 set it to 0

Η νέα τιμή είναι 255. Αποθηκεύουμε σε μια νέα εικόνα:

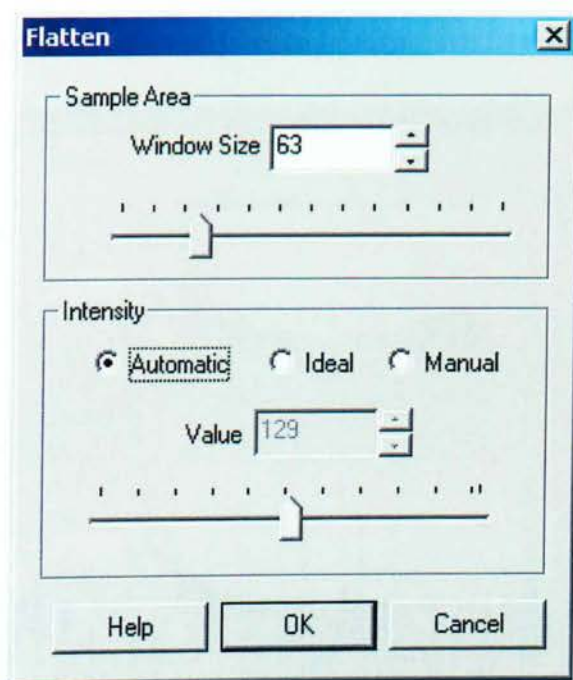
34	22	77
50	255	77
93	0	77

Συνεχίζουμε με τις υπόλοιπες 3x3 περιοχές της εικόνας χρησιμοποιώντας τις αρχικές τιμές. Για παράδειγμα η συνέχεια θα μπορούσε να είναι:

22	77	48
150	77	158
0	77	219

Σημειώνουμε οτι το 3x3 "παράθυρο" μετατοπίζεται προς τα δεξια κατα ένα πίξελ και οτι οι νέες τιμές των πιξελ δεν χρησιμοποιούνται αλλα απο θηκεύονται σε μια νέα εικόνα.

Flatten : προσπαθεί να ομαλοποιήσει τον μη ισορροπημένο φωτισμό μέσα σε μια εικόνα. Οι εικόνες λαμβάνονται σε καταστάσεις χαμηλού φωτισμού όπως με τη χρήση φλας ή με φωτογραφικές μηχανές με ανώμαλες διανομές φωτισμού. Αυτό μας δημιουργεί σημαντικές δυσκολίες μόλις προσπαθήσουμε να χρησιμοποιήσουμε διάφορες επεξεργαστικές τεχνικές όπως φίλτρα, modules κ.λπ. δεδομένου ότι η κατάτμηση εμφανίζεται ανώμαλη καθώς ο φωτισμός αλλάζει.



Παράδειγμα

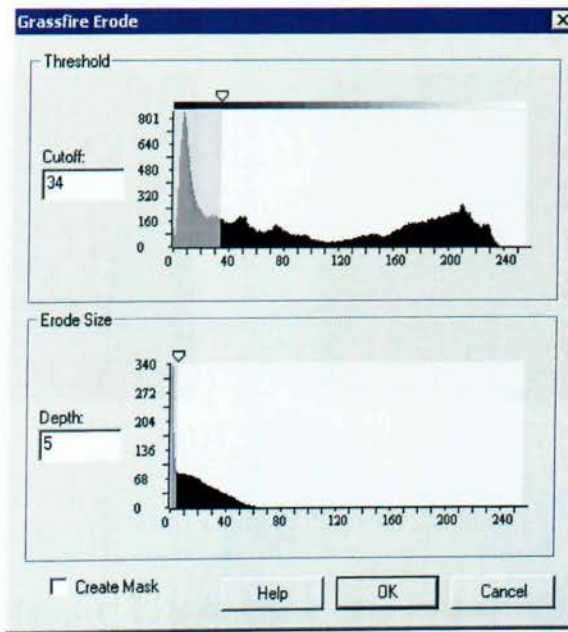
Πηγη



Ισορροπημένος Φωτισμός



Grassfire : Το εφε Grassfire χρησιμοποιείται για να “διαβρώσει” (erode) τις εικόνες κατα ένα καθορισμένο ποσό. Κατά την επεξεργασία των εικόνων πολλά μικρά αντικείμενα (λανθασμένα πιξελ) μπορούν να είναι παρόντα και να δημιουργούν θόρυβο στα τελικά αποτελέσματα. Η διάβρωση της εικόνας μπορεί να αφαιρέσει τα ενιαία πιξελ και να λειανει τις ακμές των μεγαλύτερων αντικειμένων.



Παράδειγμα

Πηγή

Grassfire



Kuwahara : Το φίλτρο Kuwahara είναι ένα φίλτρο συντήρησης ακμών που “μαλακώνει” την τρέχουσα εικόνα αλλά προσπαθεί να συντηρήσει τις άκρες. Παρόμοιο με το μέσο φίλτρο (mean filter), το φίλτρο Kuwahara αντικαθιστά το τρέχον πιξελ με το μέσο όρο ενός γειτονικού 3x3 παραθύρου που έχει τη μικρότερη διαφορά.

Παραδειγμα

Πηγή

Kuwahara

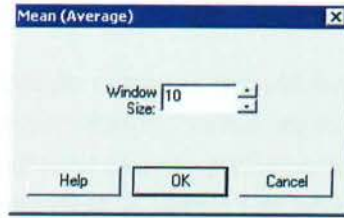


Mean Filter (Μεσο Φίλτρο) : Το μέσο φίλτρο χρησιμοποιείται για να μαλακώσει μια εικόνα μεσω του υπολογισμό του μέσου όρου των περιβαλλουσών τιμών πιξελ. Παραδείγματος χάριν, λαμβάνοντας υπόψη το παράθυρο πιξελ μιας ασπρομαυρης εικόνας 3x3 :

22	77	48
150	77	158
0	77	219

$$\text{Κεντρικό πιξελ} = (22+77+48+150+77+158+0+77+219)/9$$

Η τιμη του κεντρικου πιξελ θα αλλαξει απο 77 σε 92 καθως αυτη είναι η μέση τιμή των πιξελ που το περιβάλλουν.



Παραδειγμα

Πηγη

Μεσο φιλτρο



Τα φίλτρα [Median](#), [Midpoint](#), [Min](#) και [Pixelate](#) λειτουργουν με την λογική του μέσου φίλτρου και δεν χρίζουν περαιτέρω ανάλυσης. Το φίλτρο Salt and Pepper αφαιρεί τα πιξελ τα οποία διαφέρουν σημαντικά σε ένταση απο την πλειονότητα, δηλαδή αυτα που έχουν ασυνήθιστα υψηλές ή χαμηλές τιμές. Χρησιμοποιείται κυρίως για την επεξεργασία εικόνων που μας έρχονται απο διαστημοσυσκευές.

4.3 Έλεγχος (CONTROLS)

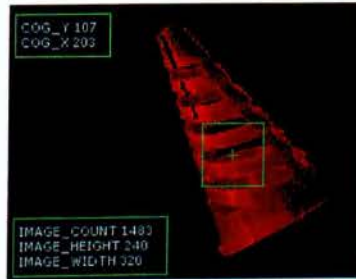
Η ενότητα Controls μας προσφέρει έτοιμο λογισμικό ώστε να χρησιμοποιήσουμε συγκεκριμένα περιφεριακά, όπως κάμερες, μοτέρ, σέρβο ή και ακόμα έτοιμα ρομπότ του εμπορίου. Έτσι λύνουμε αρκετά από τα προβλήματα επικοινωνίας μεταξύ των συσκευών που θα χρησιμοποιήσουμε. Η χρήση τους είναι αρκετά απλή και με λίγες δοκιμές και πειραματισμούς πετυχαίνουμε ένα πολύ καλό επίπεδο ρύθμισης. Οι υποενότητες είναι :

- **Cameras**
 - [Creative Live Motion](#)
 - [DLink Internet Camera](#)
 - [Firewire Camera](#)
 - [Logitech Orbit](#)
 - [TRENDnet Internet Camera](#)
- **Motors**
 - [Dimension Engineering Sabertooth](#)
 - [DT300](#)
- **Robots**
 - [IRobot Create](#)
 - [IRobot Roomba](#)
 - [Lego Mindstorm](#)
 - [Lego NXT](#)
 - [Parallax Boe-Bot](#)
 - [Surveyor SRV-1](#)
 - [Surveyor SRV-1b \(WiFi\)](#)
 - [RoboticsConnection Traxster](#)
 - [Innovation First Vex Controller](#)
- **Servos**
 - [Endurance 25 Servo](#)
 - [Lynxmotion SSC-32](#)
 - [Parallax](#)
 - [Phidgets](#)
 - [Pololu SSC](#)
 - [SSC](#)
 - [SV200](#)
 - [YEI ServoCenter 3.0](#)
- **Other**
 - [Dream Cheeky USB Car](#)
 - [Dream Cheeky USB Missile Launcher](#)
 - [Endurance R/C PCTx](#)
 - [IDVT LED-Wiz](#)
 - [Lynxmotion Seq](#)
 - [Parallel Port](#)
 - [Serial](#)
 - [Striker USB Missile Launcher](#)
 - [USB HID](#)

4.4 DISPLAY

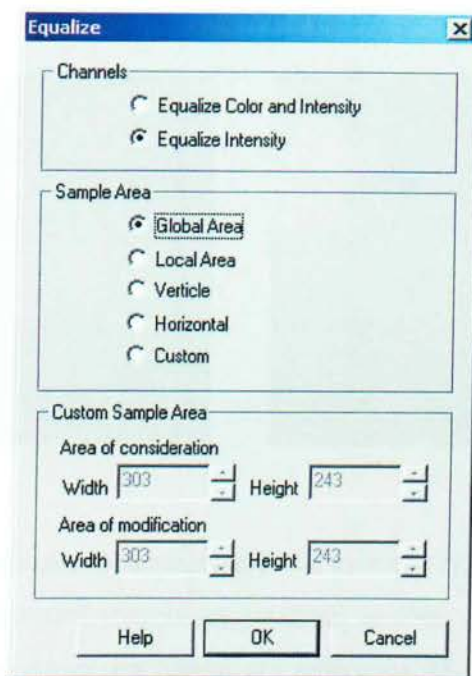
Οι επιλογές στην ενότητα Display μας δίνουν την δυνατότητα να εμφανίστουν στην οθόνη διάφορα στοιχεία που μπορεί να είναι χρήσιμα για τον χρήστη όπως η ημερομηνία, οι χρησιμοποιούμενες μεταβλητές ή ακόμα και κάποια σχήματα που θα διαφοροποιήσουν την τελική εικόνα επεξεργασίας. Οι επιλογές είναι :

- [Button Interface](#)
- [Display Circle](#)
- [Display Date/Time](#)
- [Display Image](#)
- [Display Line](#)
- [Display Point](#)
- [Display Rectangle](#)
- [Display Text](#)
- [Display Triangle](#)
- [Display Variables](#)



4.5 ΙΣΤΟΓΡΑΜΜΑ - HISTOGRAM

Equalize (Εξισορροπηση) : Η ενότητα equalize εξισορροπει την αντιθεση μιας εικονας (contrast) βασιζόμενη στα διάφορα μεγέθη δειγμάτων. Προσπαθεί να εξισώσει τον αριθμό εικονοκυττάρων σε ένα δεδομένο χρώμα και τείνει να “ισιώσει” και να αυξήσει το ιστόγραμμα της εικονας. Η επιλογή global θα εξισορροπησει την εικόνα βασισμένη σε όλα τα εικονοκύτταρα στην εικόνα. Η κάθετη εξισορροπηση χρησιμοποιεί μόνο μια ενιαία στήλη των εικονοκυττάρων στη διαδικασία εξίσωσης. Η οριζόντια εξισορροπηση χρησιμοποιεί μια ενιαία σειρά των εικονοκυττάρων. Η ενότητα εξισορροπηση μπορεί να λειτουργήσει χρησιμοποιώντας μονο τις εντάσεις εικονοκυττάρου ή σε όλα τα RGB κανάλια. Αυτό επιτρέπει την τροποποιηση των χρωμάτων της εικόνας για να τα διανειμει καλύτερα στο φάσμα.



Πηγή



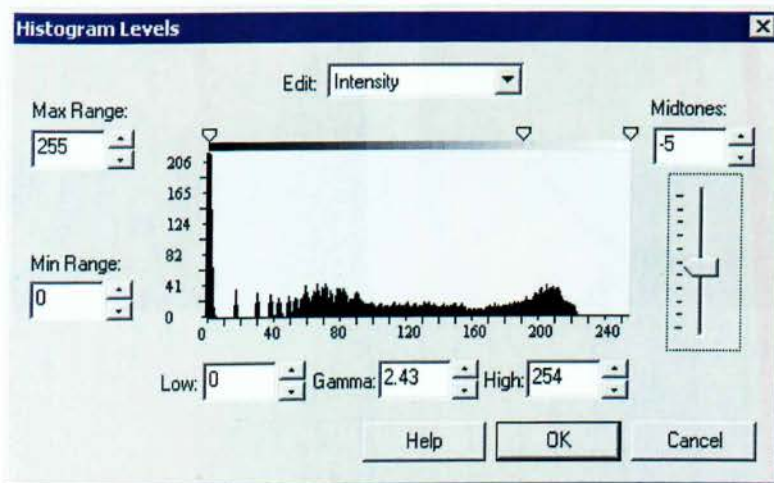
Εξισοροπημένη εικόνα



Exponential Histogram(Εκθετικό ιστόγραμμα) :Το εκθετικό ιστόγραμμα επιτρέπει τον μετασχηματισμο – επαναφορά μιας εικονας, πίσω στο γραμμικό διάστημα φωτισμού. Η εκτέλεση μιας εκθετικής ρύθμισης του ιστογράμματος θα τείνει να σκουρύνει τα φωτεινότερα εικονοκύτταρα κρατώντας σκοτεινά τα ήδη σκουρα εικονοκυτταρα.



Histogram Levels (Επίπεδα ιστογράμμων) : Η ενότητα Histogram Levels παρέχει έναν τρόπο να ελεγχθεί το ιστόγραμμα μιας εικόνας με έναν προσαρμοσμένο απο τον χρηστη τρόπο για να επιτύχει μια επιθυμητή επίδραση. Η τροποποίηση μπορεί να γίνει σε πολλαπλά κανάλια ή ένα ενιαίο κανάλι.



Πηγη



Histogram Levels



Logarithmic Histogram (Λογαριθμικό ιστόγραμμα): Η εκτέλεση μιας λογαριθμικής ρύθμισης του ιστογράμματος θα τείνει να λαμπρύνει τα σκοτεινότερα εικονοκύτταρα κρατώντας τα φωτεινά εικονοκύτταρα φωτεινά. Η λογαριθμική ενότητα βοηθά να φέρει έξω τα αντικείμενα απο τις σκιές.

Πηγη



Λογαριθμικό ιστόγραμμα



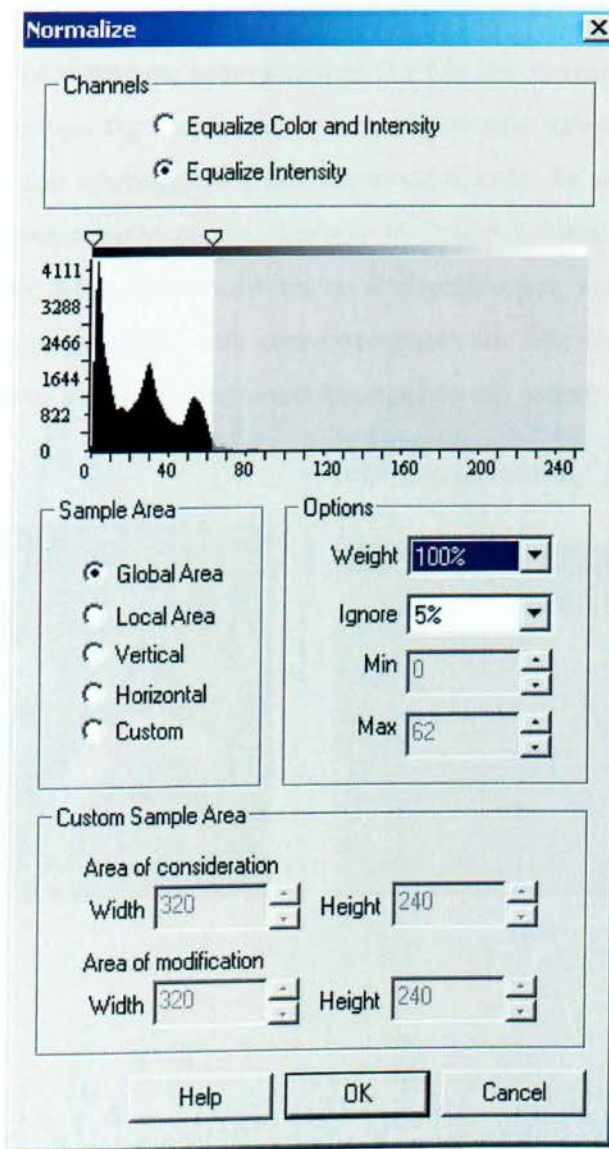
Normalize (Κανονικοποίηση) : Η ενότητα αυτη μεταβαλει τις τιμες των αξιων των εικονοκυταρων μιας εικονας με σκοπο να καλυψουν το πληρες ευρος τιμων (0-255). Η συναρτηση επεξεργάζεται καθε ζώνη χρώματος (RGB) και καθορίζει την ελάχιστη και μέγιστη αξία σε κάθε μια από τις τρεις ζώνες χρώματος. Μόλις υπολογιστούν αυτές οι τιμές η εικόνα επανεπεξεργάζεται με την αφαίρεση της ελάχιστης αξίας κάθε ζώνης από κάθε εικονοκύτταρο και τη διαίρεση της με την ανώτατη-ελάχιστη σειρά της (3 φορές για κάθε RGB εικονοκύτταρο). Αυτό έχει ως αποτελεσμα να μεταβληθει η τιμη καθε εικονοκυττάρου στο πληρες ευρος τιμων εικονοκυττάρου 0-255.

$$R = ((R - R_{min}) / (R_{max} - R_{min})) * 255.0$$

$$G = ((G - G_{min}) / (G_{max} - G_{min})) * 255.0$$

$$B = ((B - B_{min}) / (B_{max} - B_{min})) * 255.0$$

Η κανονικοποίηση είναι ένα καλό εργαλείο για να καταπολεμήσει τις αλλαγές φωτισμού καθώς η κάμερα κινείται. Πολλά φίλτρα στηρίζονται στις απόλυτες τιμές εικονοκυττάρου (όπως το κατώτατο όριο-threshold) που αλλάζουν εύκολα στις διαφορετικές συνθήκες φωτισμού. Η κανονικοποίηση είναι ένα από τα εργαλεία που χρησιμοποιούνται για να μειώσουν αυτήν την ευαισθησία στο φωτισμό των σκηνών.



Πηγη



Κανονικοποιημενη εικονα

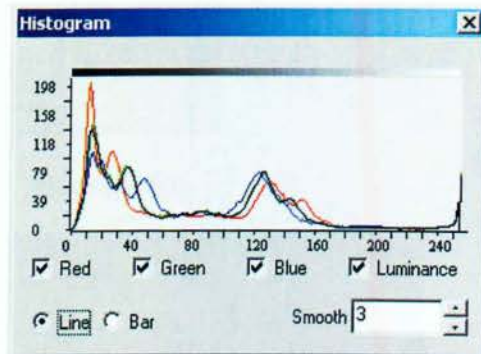


View Histogram (Απεικονιση ιστογραμματος) : Με την ενοτητα αυτη μπορουμε να δουμε το ιστογραμμα της τρεχουσας εικονας, εξαρτομενο παντα απο την θεση του module στην ροη του προγραμματος που κατασκευαζουμε. Το ιστογραμμα εικονας ειναι ειναι η γραφικη απεικονιση του αριθμου των πιξελ (αξονας Y) και των τιμων τους (0-255, αξονας X). Χρησιμοποιωντας το ιστογραμμα μιας εικονας μπορουμε να δουμε την κατανομη χρωματος των εικονοστοιχειων και πως τα διαφορα στοιχεια επεξεργασιας εικονας που χρησιμοποιουμε επηρεαζουν την τελικη εικονα.

Παραδειγμα



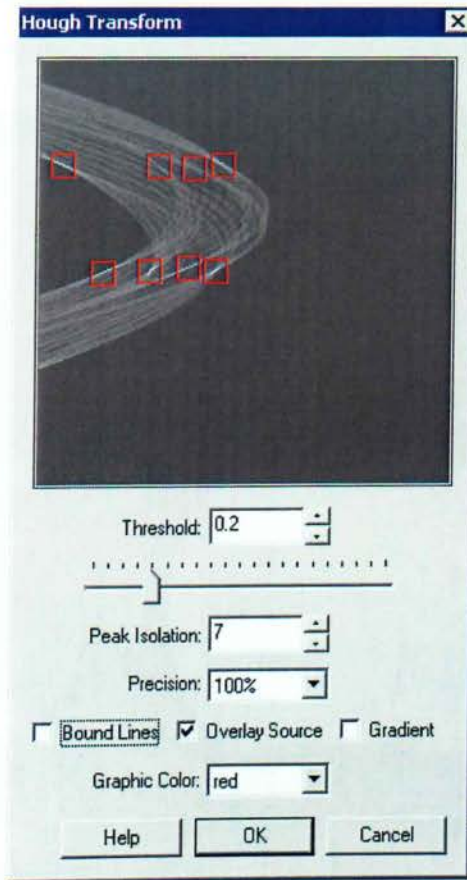
Ιστογραμμα εικονας



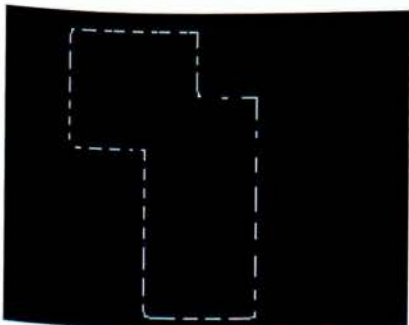
4.6 Matching Modules

Hugh Transform (Μετασχηματισμος Hugh) : Ο μετασχηματισμος Hugh ειναι μια τεχνικη για την δημιουργια γραμμων βασιζομενων σε σημεια. Τα αποτελεσματα μιας

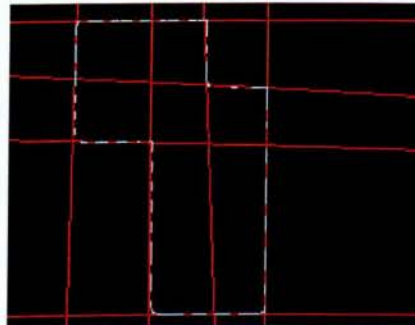
χαρακτηριστικής ρουτίνας ανίχνευσης ακμών είναι πολλά αποσυνδεδεμένα σημεία. Σε μας είναι προφανές ότι αυτά τα σημεία αντιπροσωπεύουν σχήματα αλλά επειδή τα σημεία δεν συνδέονται είναι δύσκολο για μια μηχανή να γίνει κατανοητή η ελλοχεύουσα μορφή. Ο μετασηματισμος Hugh παίρνει ως είσοδο πολλά σημεία και θα παραγάγει εικασίες για το ποιες γραμμές εκείνα τα σημεία αντιπροσωπεύουν.



Πηγη

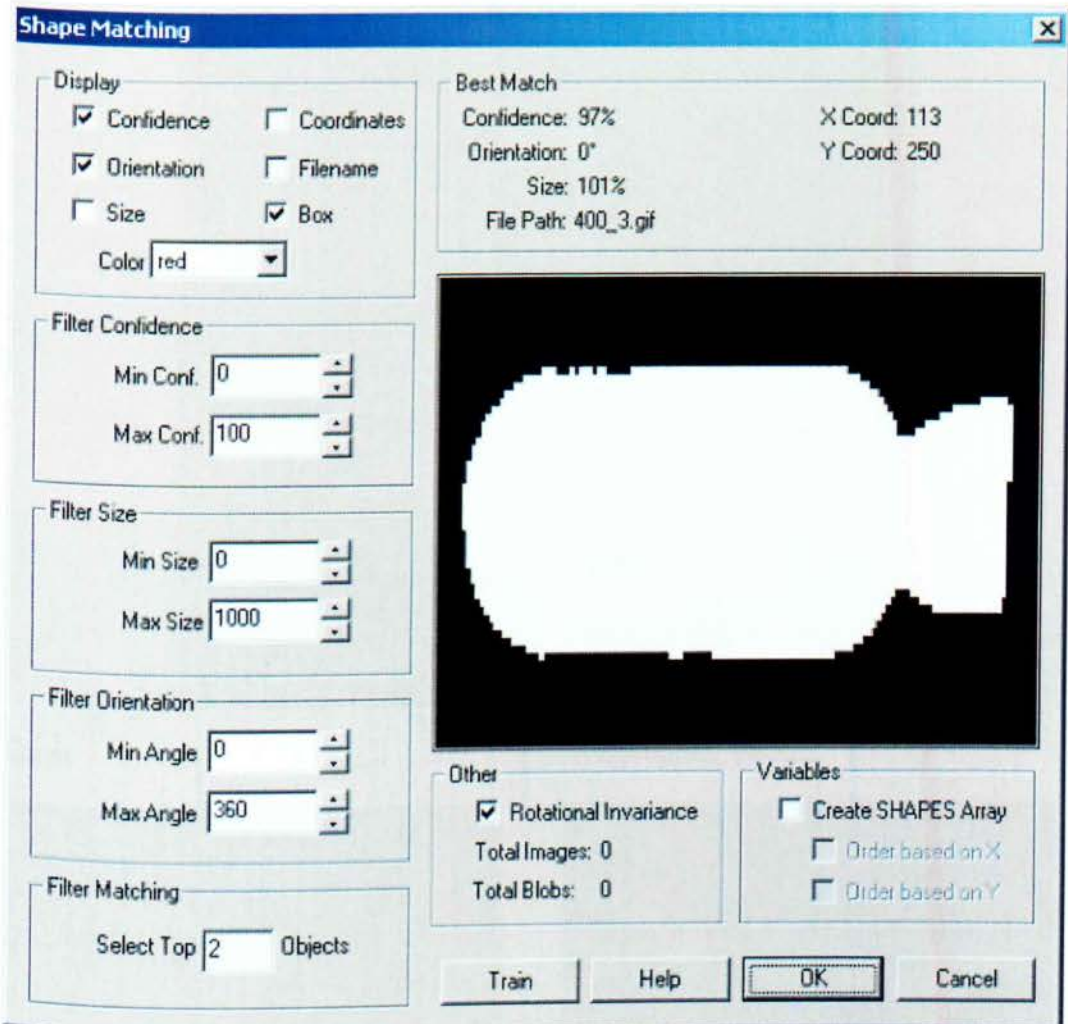


Γραμμες Hugh

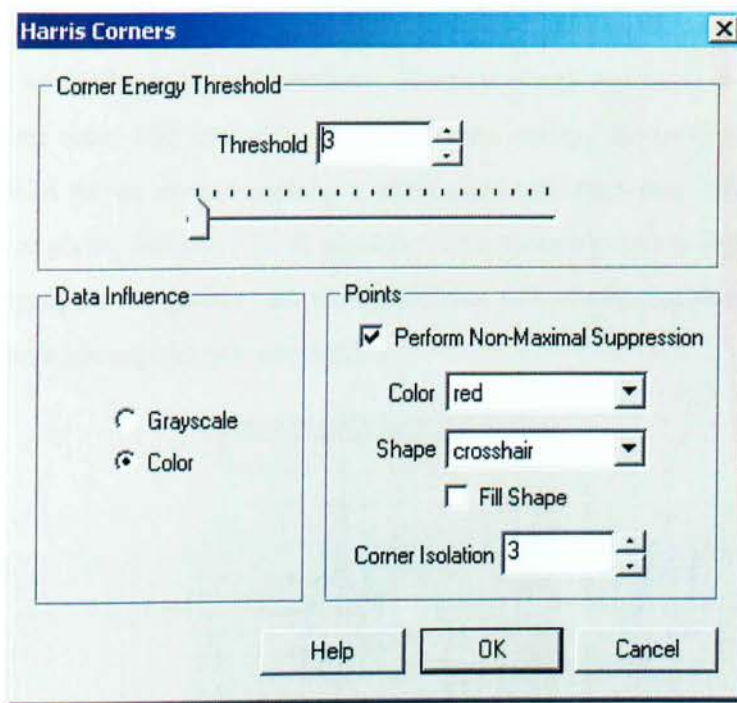


Shape matching (Αναγνώριση σχηματος) : Μας παρέχει έναν τρόπο να αναγνωριστεί μια δυαδική εικόνα χρησιμοποιώντας μια γνωστή βάση δεδομένων

εικόνων. Η ενότητα χρησιμοποιείται για να αναγνωρίσει μορφές και να μας παρέχει τις στατιστικές σχέσεις μεταξύ του αναγνωριζομένου αντικειμένου και αυτού που βρίσκεται αποθηκευμένο στη βάση δεδομένων μας. Σημειώστε ότι η αναγνωριση σχηματος βασίζεται στην πραγματική μορφή της εικόνας και ΟΧΙ στην κατα εικονοκυτταρο συγκριση εικονων όπως γίνεται στο ταίριασμα προτύπων εικόνας (Image template matching). Κατά αυτόν τον τρόπο αναγνωριση γίνεται χωρίς να παίζουν ρόλο το μεγεθος και ο προσανατολισμος των αντικειμενων.

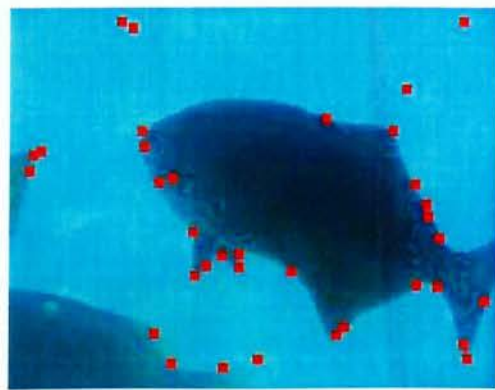


Harris Corners (Γωνίες Harris): Προσδιορίζει τις παρούσες στην εικόνα γωνίες χρησιμοποιώντας την τεχνική ανίχνευσης γωνιών Harris. Η τεχνική προσδιορίζει αρχικά τις κάθετες και οριζόντιες άκρες χρησιμοποιώντας έναν ανιχνευτή ακμών τύπων Sobel. Εκείνες οι ακμές είναι έπειτα θολωμένες για να μειώσουν την επίδραση οποιουδήποτε θορύβου εικόνας. Οι προκύπτουσες ακμές συνδυάζονται έπειτα για να διαμορφώσουν έναν ενεργειακό χάρτη που περιέχει τις “αιχμές” και τις “κοιλιάδες” (peaks and valleys). Οι αιχμές δείχνουν την παρουσία μιας γωνίας.



Πηγη

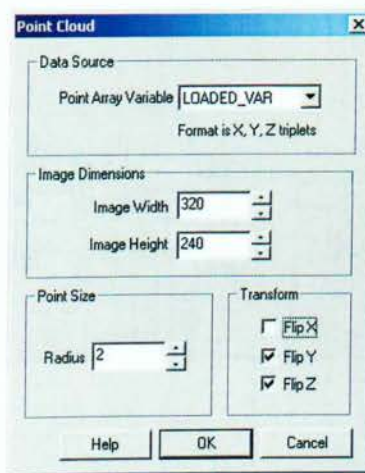
Ανιχνευμενες γωνίες



Νεφη σημείων

Η ενότητα νέφη σημείων χρησιμοποιείται για να δημιουργήσει μια στερεή δομή βασισμένη σε ένα νεφος σημείων. Αυτό απαιτείται συνήθως μετά από ληψη εικονας απο ορισμενες συσκευες των οποίων τελικό αποτέλεσμα είναι ένα σύνολο επιλεγεντων σημείων σε μη διακριτα διαστήματα. Τα νεφη σημείων είναι ένας τροπος ένωσης αυτών των αραιών σημείων σε ένα ομοιόμορφο στερεό προκειμένου να επιτευθει η πρόσθετη επεξεργασία.

Μόλις προσεγγίζονται αυτά τα σημεία είναι ομαλοποιημένα ωστε να εφαρμοσουν σε μια εικόνα με προκαθορισμενες διαστάσεις. Κατά συνέπεια ακόμα κι αν το ευρος του σημειου x είναι από -100 έως -50 η προκύπτουσα εικόνα θα αντιπροσωπεύσει το ευρος αυτο αλλά με τις συντεταγμένες x μετέφρασμενες σε 0 έως 320 υποθέτοντας ότι ορισαμε ενα ευρος εικόνας 320. Η προκύπτουσα γραφική εικόνα δημιουργείται με την ομαλοποίηση των σημείων και το σχεδιασμό ενός υπολογισμένου, κατά μέσο όρο, κύκλου σε κάθε σημείο στη νέα εικόνα.

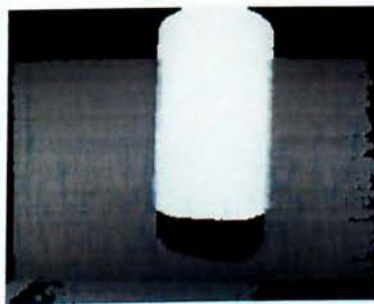


Παράδειγμα

Σημεία με την ακτίνα 0



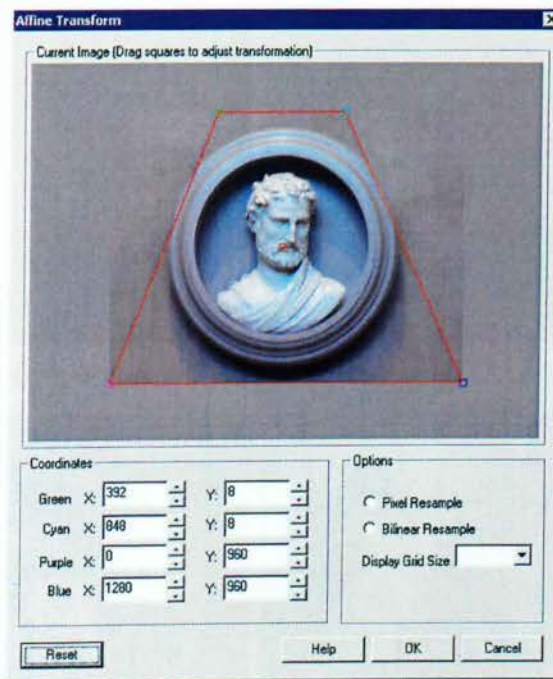
Σημεία με την ακτίνα 2



4.7 Μετασχηματισμοί

Συσχετισμένος μετασχηματισμός

Η ενότητα συσχετισμένος μετασχηματισμός (affine transformation) παρέχει έναν τρόπο ώστε να στρεβλώσει μια εικόνα σε μια διαφορετική εικόνα. Θα χρησιμοποιήσει τα τέσσερα επιλεγμένα σημεία για να δημιουργήσει μια νέα εικόνα που αλλάζει μέγεθος, περιστρέφεται και μεταφράζεται βασισμένος στα τέσσερα σημεία. Αυτή η μετατροπή είναι χρήσιμη για τη διαστρέβλωση προοπτικής ή για την αποκατάσταση των αντικειμένων μέσα στο οπτικό πεδίο.



Παράδειγμα

Πηγή



Καθαρή μετατροπή



Διόρθωση προοπτικής



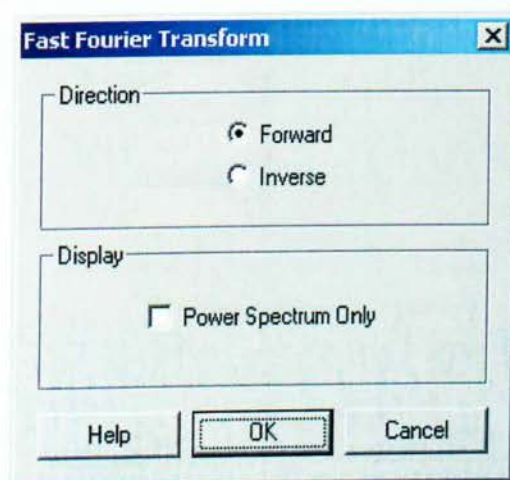
Ευθυγράμμιση αντικειμένου



FFT - Γρήγορος μετασχηματισμός Φουριέ

Η ενότητα FFT παρέχει έναν τρόπο να μετασχηματιστεί η τρέχουσα εικόνα από το χωρικό διάστημα στο διάστημα συχνότητας. Το FFT είναι μια βασικές ρουτίνες στην επεξεργασία σήματος δεδομένου ότι μπορεί να χρησιμοποιηθεί ώστε να εξαλειφει επαναλαμβανόμενα σήματα από το αρχικό σήμα. Η ενότητα FFT θα αποσυνθέσει μια εικόνα στις θεμελιώδεις συχνότητες έντασής της που μπορούν να φιλτραριστούν και να επανασυνδυαστούν για να δημιουργήσουν μια νέα εικόνα.

Τα αποτελέσματα ενός μετασχηματισμού κατά Φουριέ είναι δύο κανάλια δεδομένων. Το πρώτο κανάλι που περιέχει την ένταση ή τους πραγματικούς αριθμούς αναφέρεται συχνά ως δυναμικό φάσμα. Αυτό είναι το κανάλι που μπορεί να υποβληθεί σε επεξεργασία για να αφαιρέσει ορισμένα ανεπιθύμητα δεδομένα από μια εικόνα. Το δεύτερο κανάλι καλείται κανάλι φάσης. Αυτό το κανάλι περιέχει τις φανταστικές ή πληροφορίες φάσης ως αποτέλεσμα του μετασχηματισμού κατά Φουριέ και χρειάζεται μόνο εάν επιδιώκεται ένας αντιστροφος μετασχηματισμός. Αυτό το κανάλι δεν υποβάλλεται σε επεξεργασία και περνά αμετάβλητο στον αντιστροφο μετασχηματισμό. Η κύρια χρήση του FFT στην επεξεργασία εικόνας είναι για την αφαίρεση του επαναλαμβανόμενου θορύβου από μια εικόνα. Αυτό ολοκληρώνεται με το μετασχηματισμό της εικόνας στο διάστημα συχνότητας χρησιμοποιώντας την ενότητα FFT και αφηνοντας έπειτα έξω τα σημεία από το φάσμα ισχύος που απεικονίζουν την παρουσία ενός ιδιαίτερου σήματος. Μολις το σήμα μετασχηματισθεί στο διάστημα συχνότητας είναι εύκολο να δούμε τα σημεία ή τις γραμμές στο δυναμικό φάσμα που δημιουργούνται λόγω του ημιτονοειδούς θορύβου στην αρχική εικόνα.



Παράδειγμα

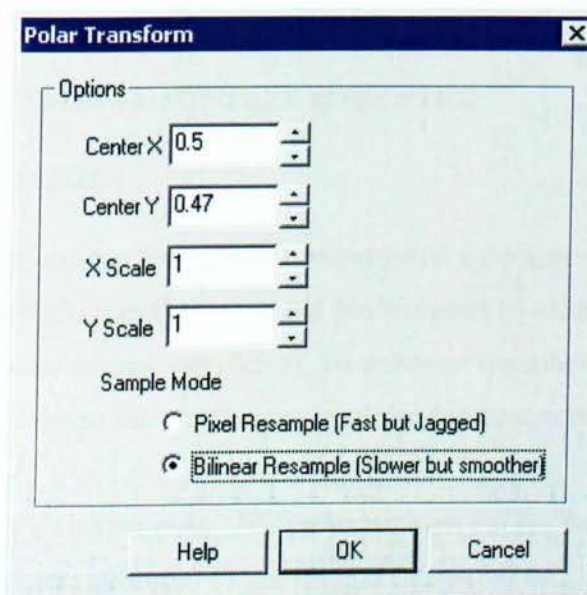
Πηγή

Επεξεργασμένο FFT



Πολική μετατροπή

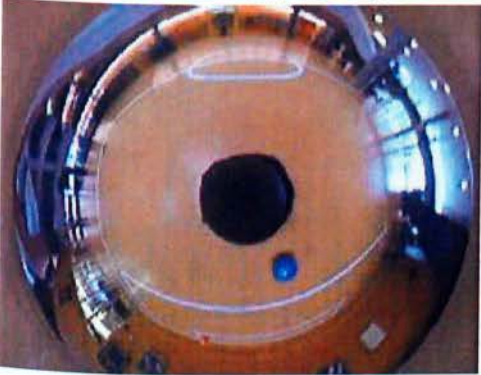
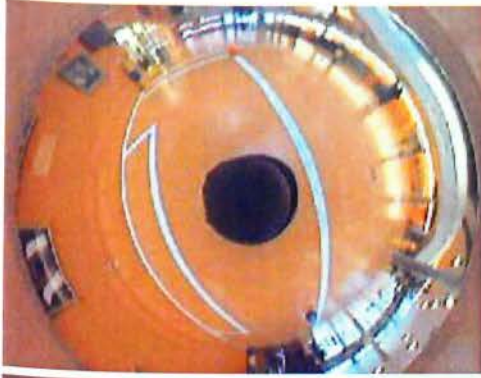
Η ενότητα πολικής μετατροπής “ξετυλιγει” μια πανκατευθυντική παραβολική εικόνα έτσι ώστε η προκύπτουσα εικόνα είναι μια κάθετη ερμηνεία της αρχικής κυκλικής εικόνας. Αυτό είναι χρήσιμο ώστε να μετασχηματίσει την εικόνα σε μια πιό ανθρωπινως κατανοητή μορφή.



Παράδειγμα

Πηγή

Πολική μετατροπή

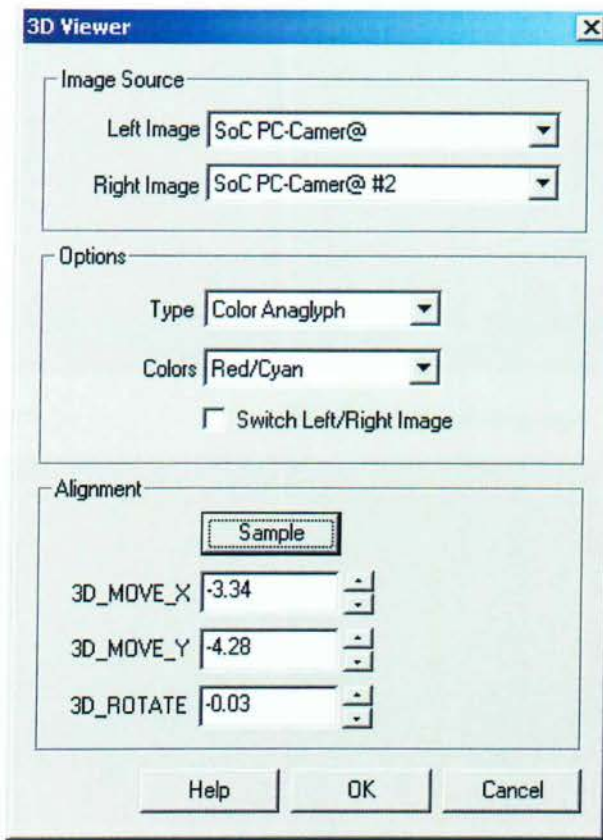


4.8 ΠΡΟΚΑΘΟΡΙΣΜΕΝΕΣ ΔΟΜΙΚΕΣ ΕΝΟΤΗΤΕΣ

4.8.1 3D ΑΠΕΙΚΟΝΙΣΗ

Η ενότητα αυτή μας παρέχει ένα τρόπο να απεικονιστεί καλύτερα η τρισδιάστατη εικόνα που λαμβανουμε. Η ενότητα απαιτεί δύο webcams (ή άλλες πηγές εικόνας) τοποθετημένα περίπου σε απόσταση 6.5cm. Τα webcams τροφοδοτούν έπειτα με δύο εικόνες αυτήν την ενότητα που παράγει μια ενιαία εικόνα βασισμένη στις προδιαγραφές μας.

Σημειώστε ότι αυτή η ενότητα είναι μόνο για απεικόνιση και δεν δημιουργεί στερεοσκοπικούς χάρτες βάθους.

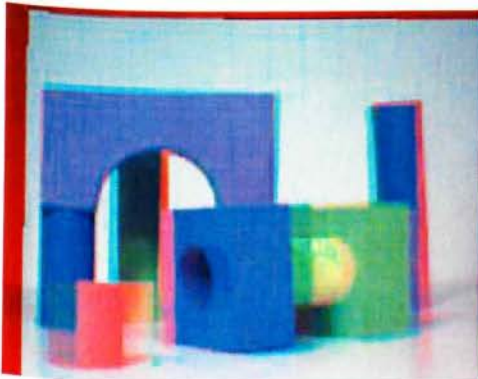


Παράδειγμα

Οι ακόλουθες επιδείξεις το ρομπότ Traxster που επισημαίνει ένα κοκ μπορούν.

Ανάγλυφο χρώματος

Βελτιστοποιημένο χρώμα



Ανάγλυφο χρώματος

Μισό χρώμα



Ανάγλυφο χρώματος



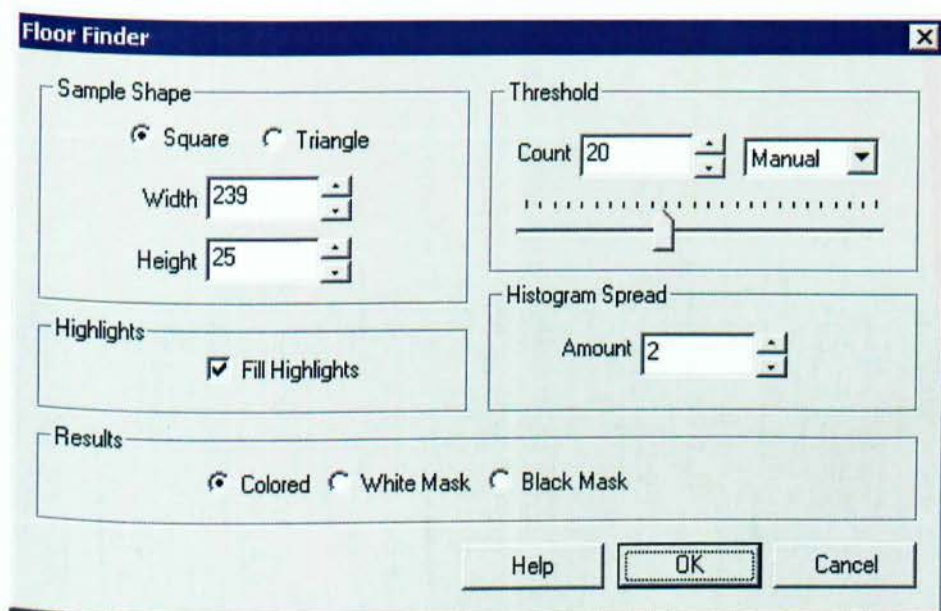
Βελτιστοποιημένο χρώμα



4.8.2 Ανιχνευτής Δαπεδου

Η ενότητα ανιχνευτης δαπεδου χρησιμοποιείται για να προσδιορίσει το δαπεδομέσα σε μια εικόνα. Οι υποθέσεις που καθιστούν αυτό εφικτο είναι ότι το ρομπότ ή η φωτογραφική μηχανή είναι σε ένα επίπεδο πάτωμα που εκτείνεται από το κατώτατο σημείο της καμερας και μακριά από την καμερα. Αυτή η υπόθεση απαιτείται δεδομένου ότι η ενότητα θα επιλέξει τα εικονοκύτταρα στο κατώτατο σημείο της περιοχής εικόνας (το δειγμα) και θα χρησιμοποιήσει εκείνα τα εικονοκύτταρα για να προσδιορίσει τα παρόμοια εικονοκύτταρα στο υπόλοιπο της εικόνας. Κατά συνέπεια υποτίθεται ότι το δικαίωμα διαστήματος πατωμάτων μπροστά από το ρομπότ είναι σχετικά απαλλαγμένο από τα εμπόδια και αντιπροσωπεύει μια μερίδα του πατώματος.

Διεπαφή

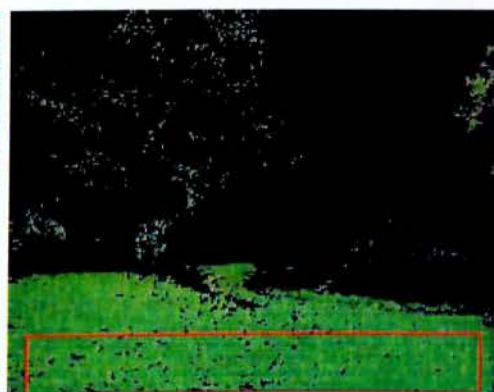


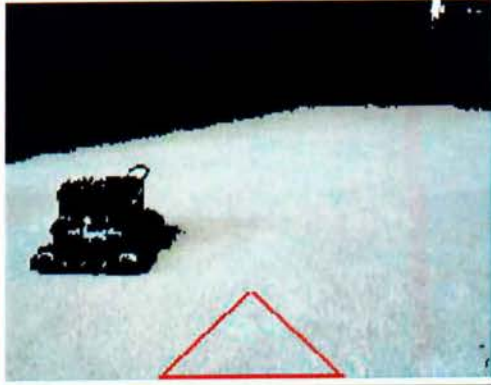
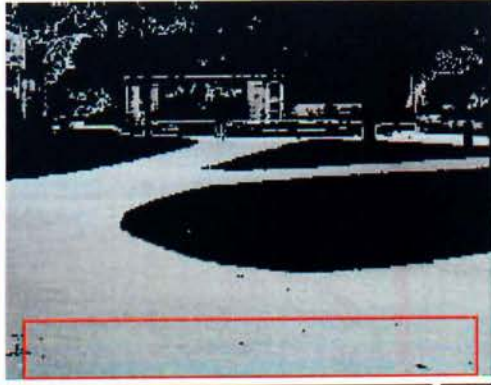
Παραδείγματα

Πηγή



Πάτωμα που βρίσκεται





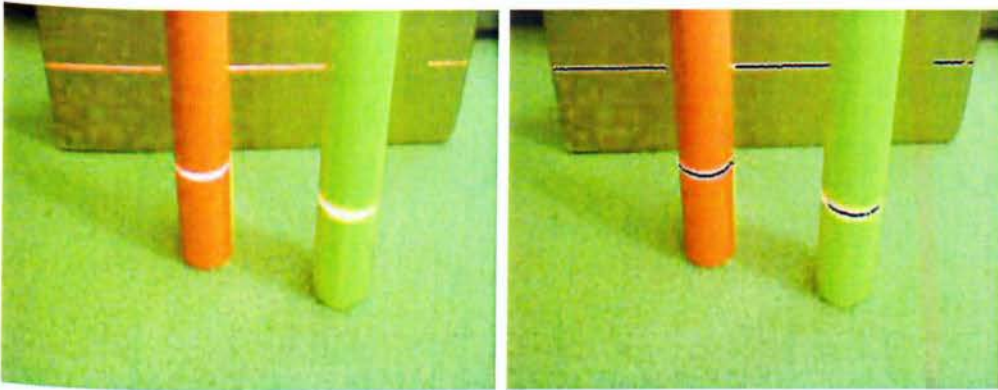
4.8.3 Γραμμή λέιζερ

Η ενότητα γραμμών λέιζερ ανιχνεύει μια ενιαία κόκκινη γραμμή λέιζερ μέσα στην τρέχουσα εικόνα που υποτίθεται ότι διαγραφεί ην εικόνα σε μια οριζόντια κατεύθυνση από τη δεξιά στην αριστερή πλευρά της εικόνας. Ο αρχικός σκοπός της ενότητας γραμμών λέιζερ είναι να ανιχνευθεί μια οριζόντια γραμμή λέιζερ έτσι ώστε να χρησιμοποιήσει τις ανιχνευμένες τιμές για τη αποφυγή σύγκρουσης ή εμποδίων. Αυτή η ενότητα έχει εξεταστεί με την δίοδο γραμμών λέιζερ [Instapark](#).

Παράδειγμα

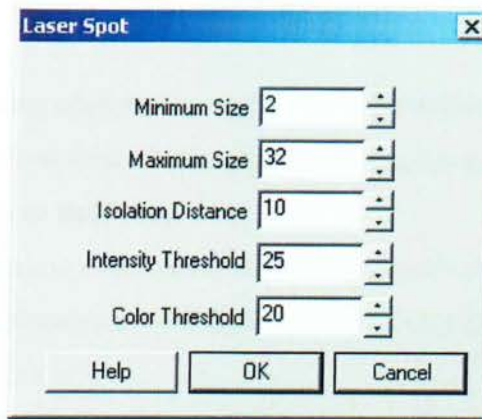
Πηγή

Laser_Line



4.8.4 Σημείο λέιζερ

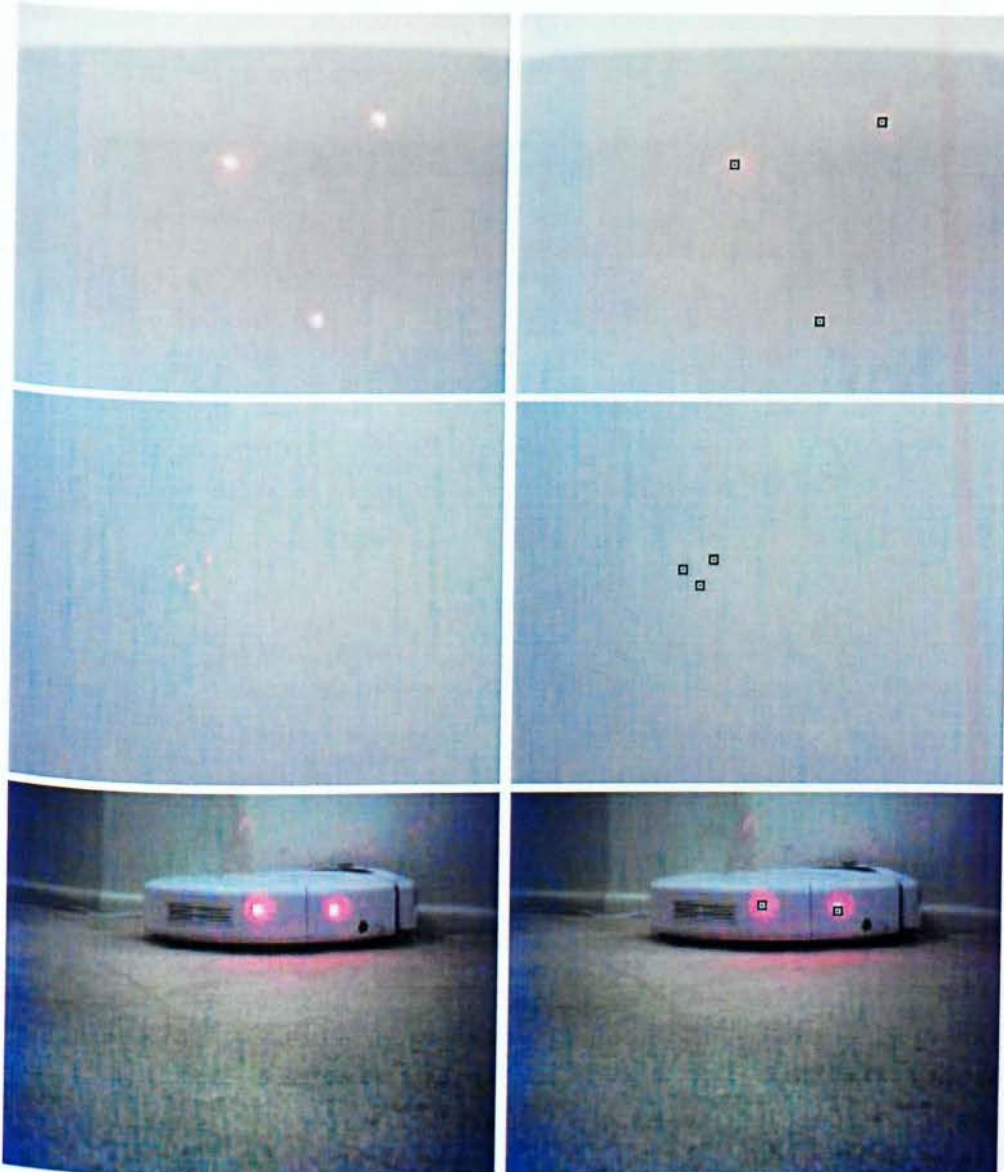
Η ενότητα σημείων λέιζερ παρέχει έναν εύκολο τρόπο να προσδιοριστεί ένα κόκκινο σημείο λέιζερ μέσα στην τρέχουσα εικόνα. Η ρουτίνα ανίχνευσης ψάχνει τα χαρακτηριστικά που δείχνουν την παρουσία ενός κόκκινου σημείου που περιλαμβάνει την ένταση, το χρώμα και τη μορφή. Η ενότητα θα δείξει γραφικά τη θέση των σημείων λέιζερ και θα δημιουργήσει επίσης μια μεταβλητή σειρά `LASER_POINTS` που κρατά τις θέσεις `X` και `Y` των ανιχνευμένων σημείων.



Παράδειγμα

Πηγή

Σημεία λέιζερ

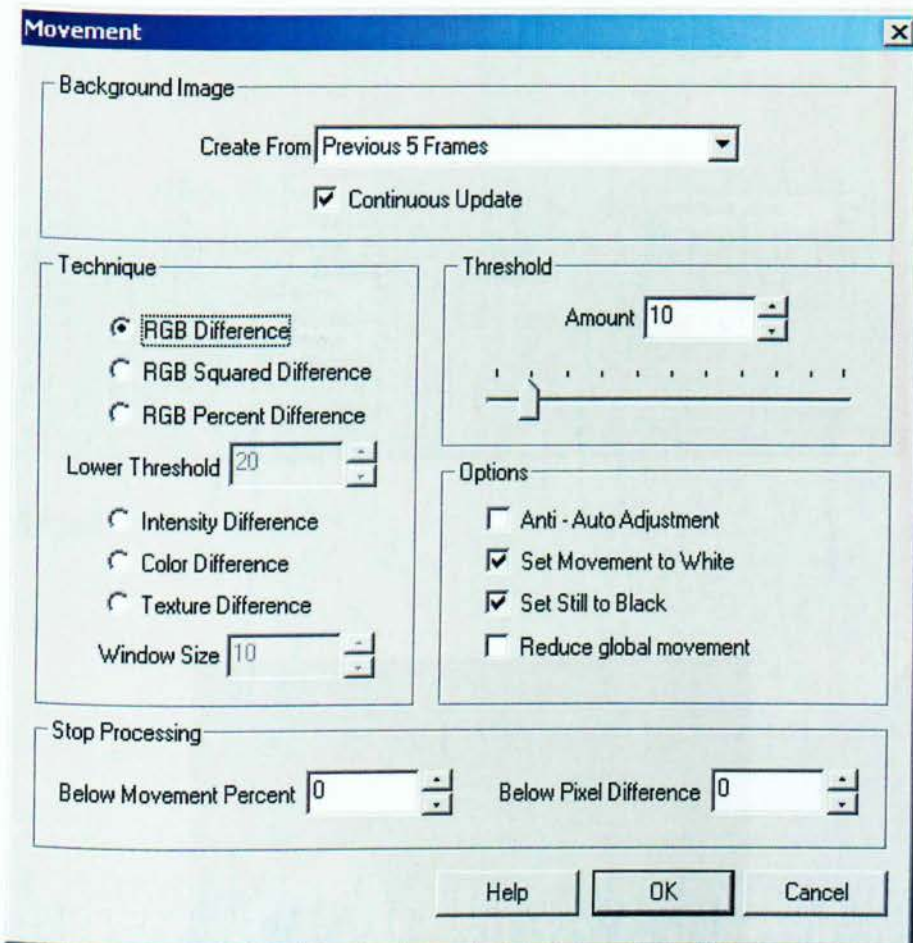


4.8.5 Μετακίνηση

Η λειτουργία μετακίνησης παρέχει έναν τρόπο να ανιχνευθούν οι αλλαγές εικόνας. Η μετακίνηση μπορεί να είναι ένας εύκολος τρόπος να ξεχωρίσουμε ένα αντικείμενο που μας ενδιαφέρει από το υπόβαθρο.

Οποιοδήποτε εικονοκύτταρο που είναι αρκετά διαφορετικό στο χρώμα ή την ένταση διατηρείται. Τα εικονοκύτταρα που δεν έχουν αλλάξει θετονται ως Μαύρο (προεπιλογή).

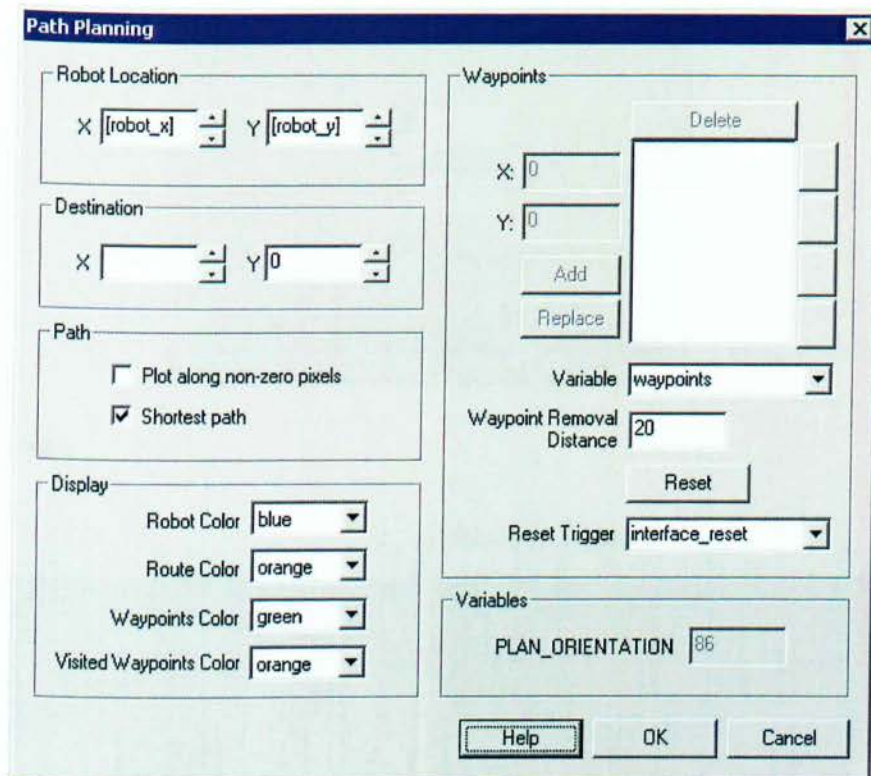
Σημειώστε ότι η χρησιμοποίηση ενός μέσου φίλτρου πριν από τη ανιχνευση μετακίνησης συστήνεται συχνά ως ένας τρόπος να μειωθούν οποιεσδήποτε απότομες αλλαγές εικονοκυττάρου. Οι πλαστές τιμές εικονοκυττάρου θα είναι πίο κοινές στις φωτογραφικές μηχανές χαμηλότερης ποιότητας έτσι οι τιμές που χρησιμοποιείτε μπορεί να πρέπει να ενημερωθούν κατά τον αλλαγή των φωτογραφικών μηχανών.



4.8.6 Σχεδιασμος πορειας

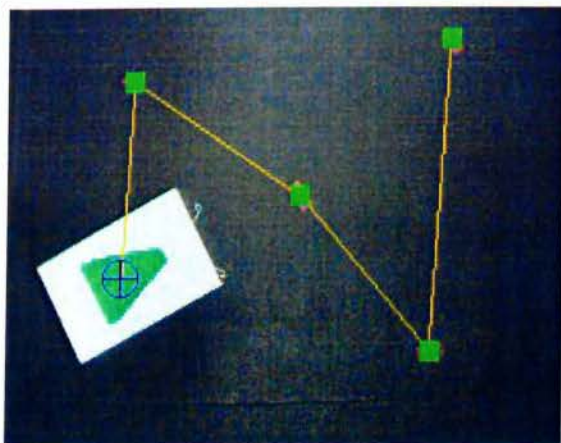
Η ενότητα σχεδιασμος πορειας χρησιμοποιείται για να καθορίσει μια διαδρομή από μια θέση συντεταγμενων σε άλλη κατά μήκος ενός συνόλου waypoints.

Παραδείγματος χάριν, εάν είχατε μια εικόνα ενός λαβυρίνθου και έπρεπε να καθορίσουμε την καλύτερη πορεία από όπου το ρομπότ βρίσκεται εως το σημειο που θελουμε να βρεθει θα χρησιμοποιησουμε την ενότητα σχεδιασμου πορειας για να καθορίσουμε την συντομοτερη ή καλύτερη πορεία στην επιθυμητή θέση.



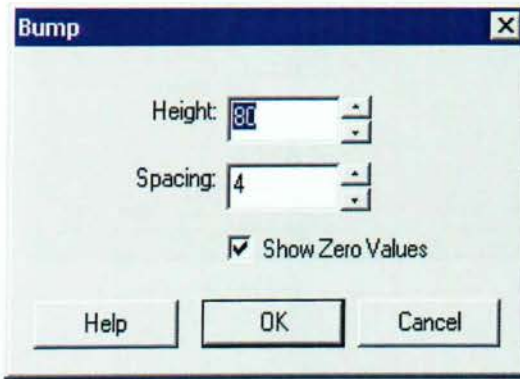
Παράδειγμα

Πηγή



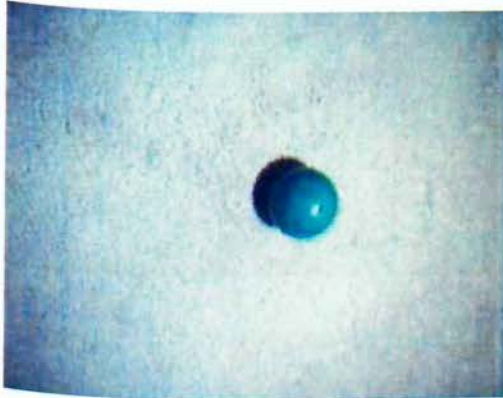
4.8.7 Σχεδιαση επιφανειας

Για να απεικονισουμε καλύτερα τις τιμες των εικονοκυτόταρων στην εικόνα χρησιμοποιούμε την λειτουργια σχεδιαση επιφανειας. Η ενοτητα αυτη χρησιμοποιεί την ένταση εικονοκυττάρων ως άξονα Y και σχεδιάζει τη θέση χρησιμοποιώντας ένα άσπρο εικονοκυτόταρο. Το τελιοκ αποτελεσμα είναι μια τρισδιάστατη εικόνα που μπορεί να χρησιμοποιηθεί για να δουμε λεπτομερέστερα πώς η επεξεργασία επιδρα στις σειρές εικονοκυττάρου.

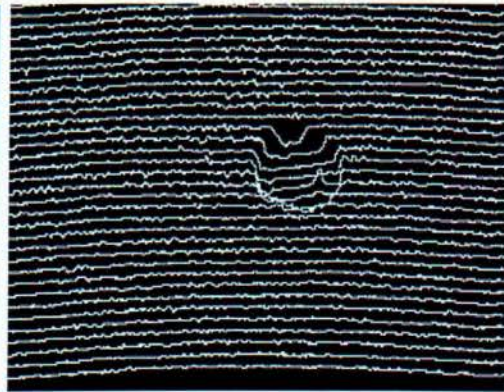


Παράδειγμα

Πηγή

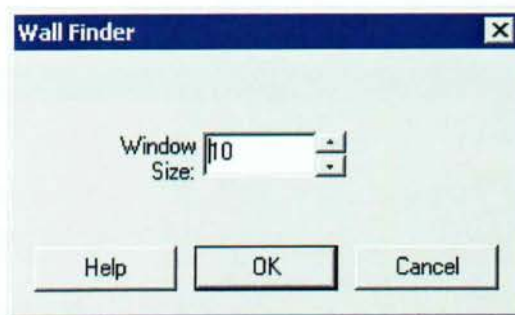


Επιφάνεια Πλοκή



4.8.8 Ανιχνευτής τοίχων

Ο ανιχνευτής τοίχων προσπαθεί να καθορίσει πού το πάτωμα και ο τοίχος συναντιούνται. Υποθέτει ότι η κάμερα εξετάζει το πάτωμα έτσι ώστε το πάτωμα να φαίνεται στο κάτω σημείο της εικόνας. Η λειτουργία ψάχνει τη μετάβαση μεταξύ του δαπέδου και του τοίχου και δείχνει εκείνη την γωνία χρησιμοποιώντας ένα κόκκινο σημείο.



Παράδειγμα

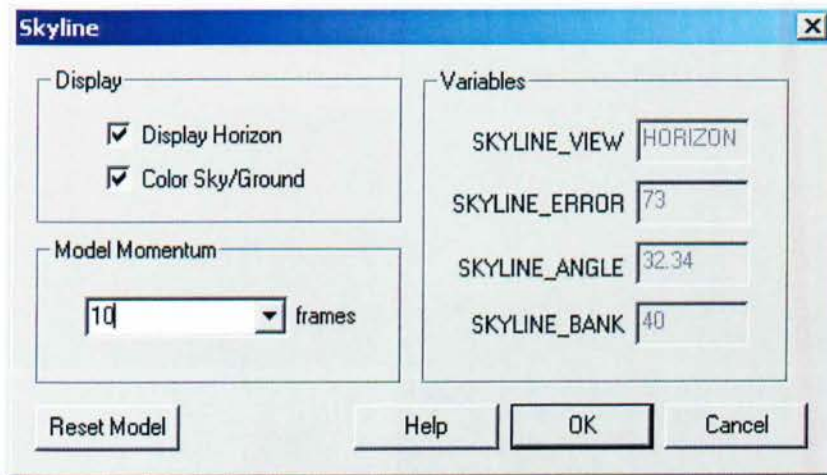
Πηγή

Ανιχνευτής τοίχων



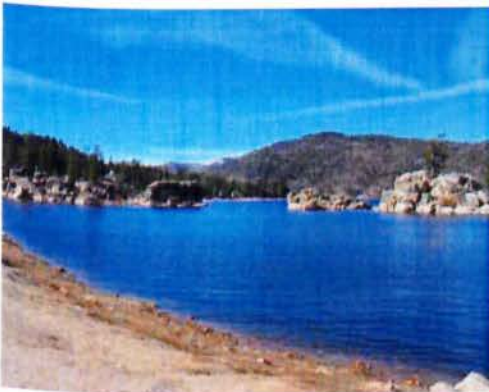
4.8.9 Ορίζοντας

Η ενότητα οριζόντων αναλύει την τρέχουσα εικόνα και προτείνει που μπορεί να είναι ο ορίζοντας. Πραγματοποιεί αυτή την εργασία με την ελαχιστοποίηση της διαφοράς μεταξύ δύο μισών της εικόνας στο σημείο όπου μια ομαλή κατάσταση επιτυγχάνεται. Αυτό σπάει ουσιαστικά την εικόνα σε δύο μέρη. Δεδομένου ότι οι περισσότεροι ουρανοί είναι πολύ διαφορετικής σύστασης από το έδαφος αυτό είναι μια αποτελεσματική τεχνική.

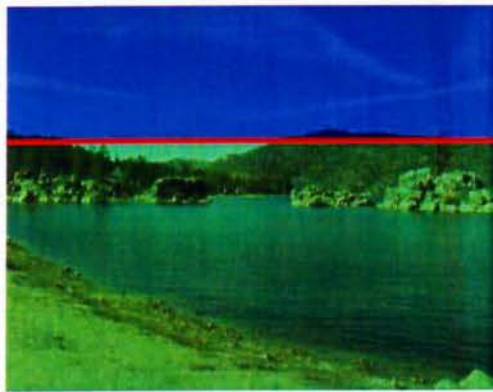


Παράδειγμα

Πηγή



Ορίζοντας



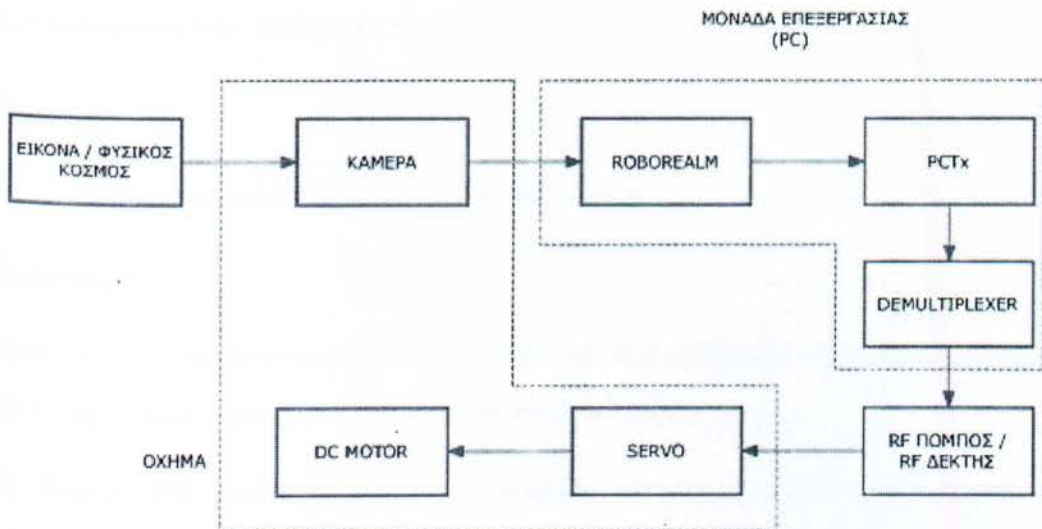


Κεφάλαιο 5

Εφαρμογή

Σε αυτό το κεφάλαιο περιγράφεται αναλυτικά η διαδικασία με την οποία θα υλοποιήσουμε ένα σύστημα υπολογιστικής όρασης, χρησιμοποιώντας το λογισμικό Roboreal. Στόχος μας είναι να καταφέρουμε να κατασκευάσουμε ένα σύστημα το οποίο θα μπορεί να κινηθεί αυτόνομα ανάμεσα σε δύο ευδιακριτές γραμμές στο έδαφος.

Εφαρμόζοντας την αρχές που είδαμε στα προηγούμενα κεφάλαια καταλήγουμε στο παρακάτω block διάγραμμα:



Απο το διαγραμμα της εφαρμογης είναι κατανοητο οτι η επιθυμητη αναδραση επιτυγχανεται μεσω της επεξεργασιας της εικονας και της “αντίληψης” του οχηματος για την θεση του στον χωρο. Δεν υπαρχει καποια αλλη ηλεκτρικη σύνδεση που μας βοηθα να πετυχουμε αυτο τον στοχο. Αν η πορεια του χρειάζεται διόρθωση αυτο γινεται στο σταδιο επεξεργασιας της εικόνας μεσω του Roboreal. Η κάμερα μας βρίσκεται τοποθετημένη πάνω στο οχημά μας και το σήμα της μεταφέρεται ασύρματα ενσύρματα στην μονάδα επεξεργασίας.

Στην περίπτωση μας δεν χρησιμοποιήσαμε τον RF πομποδέκτη του οχήματος για αυτό το σήμα μας θα μεταφερθεί ενσύρματα στον οχήμα.

Το σύστημα μας χωρίζεται στις εξής βαθμίδες:

1. Κάμερα
2. Λογισμικό RoborealM
3. PCTx
4. Demultiplexer (Αποπολυπλέκτης)
5. Όχημα

5.1 Ανάλυση των βαθμίδων

Κάμερα

Μπορεί να χρησιμοποιηθεί μία οποιαδήποτε κάμερα

RoborealM

Πυρήνας του συστήματος μας είναι το λογισμικό RoborealM, το οποίο παρουσιάσαμε σε προηγούμενο κεφάλαιο. Το επιλέξαμε για δύο λόγους:

- A. Παρουσιάζει μεγάλη ποικιλία στα modules και μας επιτρέπει συνδυαστικά να έχουμε το επιθυμητό αποτέλεσμα.
- B. Αν και σχετικά περιορισμένων δυνατοτήτων κρίνεται ακρως ικανοποιητικό για τους σκοπούς αυτής της πτυχιακής εργασίας.

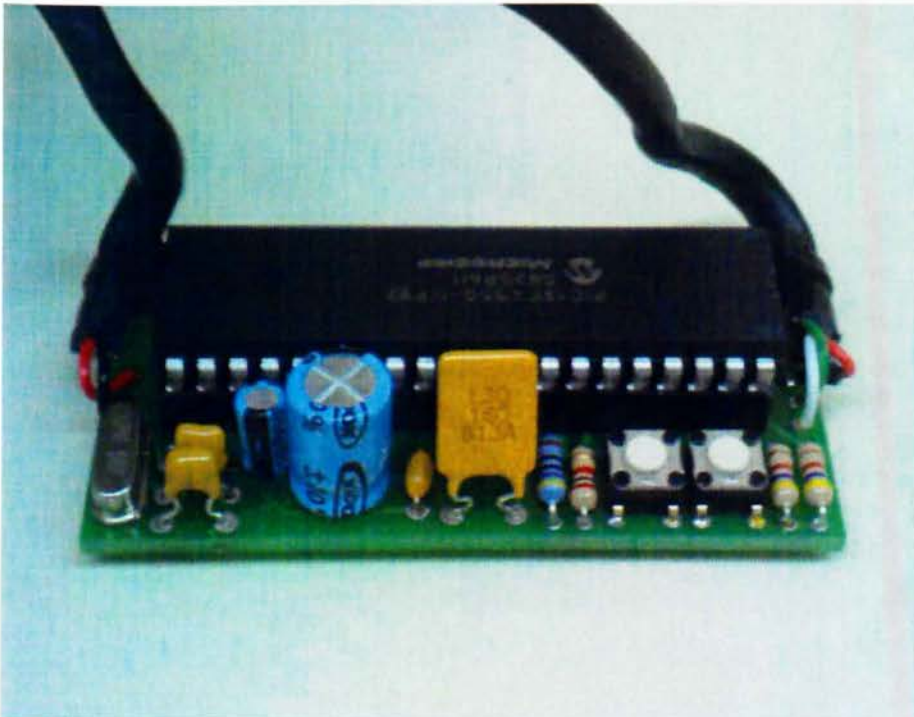
PCTx

Το εξάρτημα PCTx της εταιρίας Endurance-RC μας επιτρέπει να συνδέσουμε μια τηλεκατεύθυνση με τον ηλεκτρονικό μας υπολογιστή. Με την πραγματοποίηση αυτής της

σύνδεσης έχουμε την δυνατότητα να ελέγξουμε ασύρματα το οχημα μας μέσω προγραμμάτων που εκτελούνται στον υπολογιστή μας.

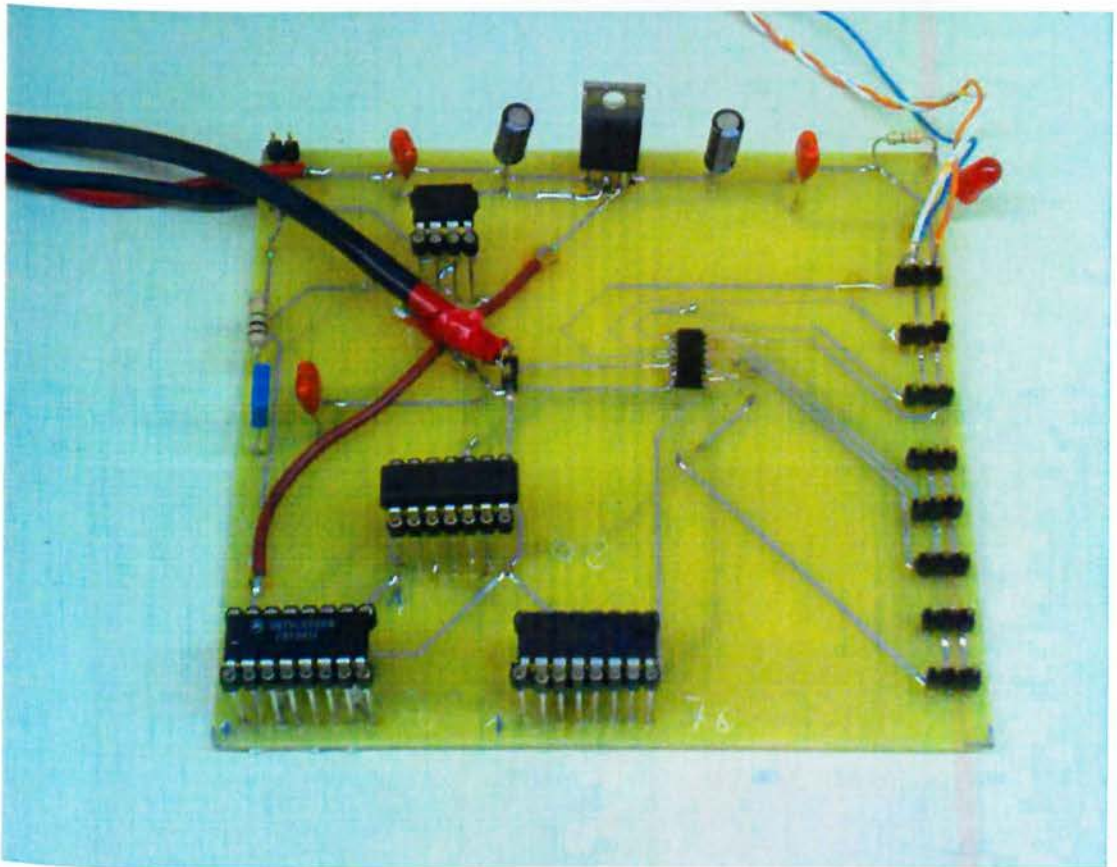


Το PCTX χρησιμοποιεί τον μικροελεγκτή PIC 18F4550 για να πραγματοποιήσει τον απαιτούμενο στόχο του.



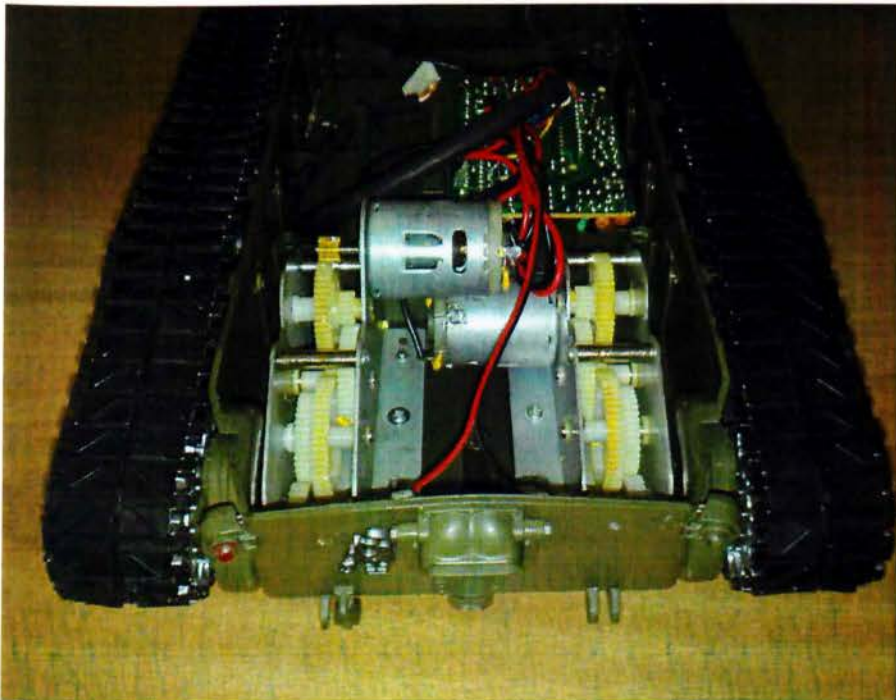
Demultiplexer – Αποπολυπλέκτης

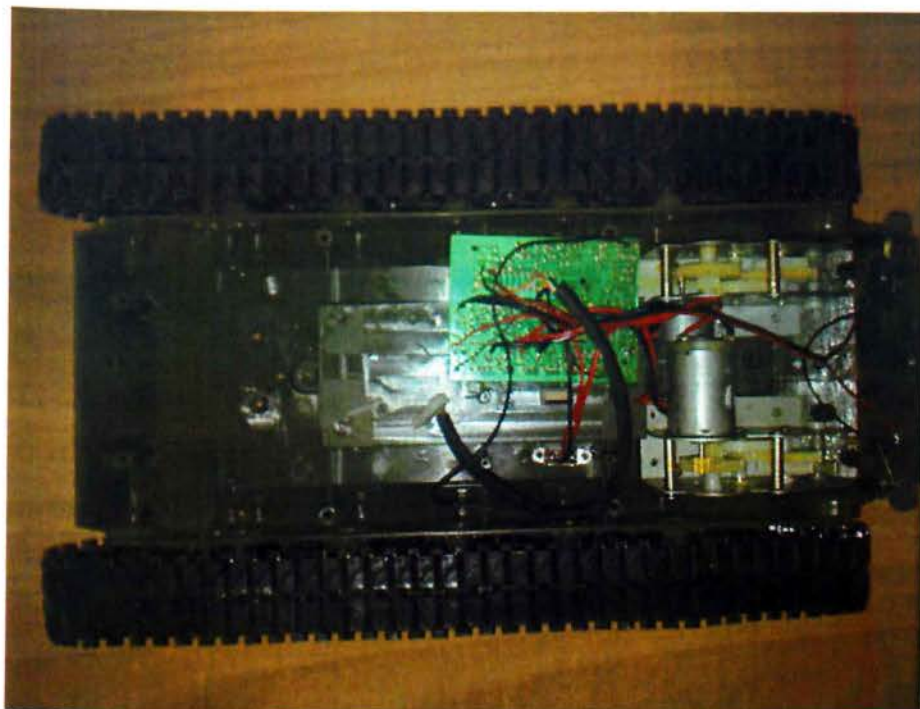
Ο πάλμος στην έξοδο του PCTx είναι διαμορφωμένος με την τεχνική Διαμόρφωσης Θέσης Παλμών (PPM – Pulse Position Modulation). Χρησιμοποιώντας αυτή την τεχνική έχουμε μια παλμοσειρά όπου οι παλμοί (8 στην περιπτωσή μας), είναι πολυπλεγμένοι σε ένα σήμα εξόδου. Για να μπορέσουμε να χρησιμοποιήσουμε τα απαραίτητα κανάλια για την εφαρμογή μας πρέπει να αποπολυπλέξουμε το σήμα εξόδου του PCTx, και να εφαρμόσουμε τον κατάλληλο παλμο στα servo που χρησιμοποιούμε για τον έλεγχο των dc μοτερ.



Όχημα

Για την παρούσα εργασία χρησιμοποιήσαμε ένα τηλεκατευθυνόμενο μοντέλο το οποίο προσαρμόσαμε κατάλληλα στις ανάγκες μας. Η χρησιμοποίηση ενός έτοιμου οχήματος έγινε για να αποφύγουμε να το σχεδιάσουμε από την αρχή καθώς είναι κάτι που είναι έξω από τους σκοπούς αυτής της εργασίας.





5.2 Βήματα της εφαρμογής μας

Σαν στόχο μας θέτουμε το οχημα μας να ακολουθήσει μια προδιαγεγραμμένη πορεία που θα την ορίσουμε με μια μαυρη γραμμή στο έδαφος. Η μοναδική οδηγία προς το όχημα θα προέρχεται από την κάμερα και μετά από επεξεργασία μέσω του λογισμικού RoborealM.

Η γραμμή στο έδαφος που θα προδιαγράψει την πορεία, θα είναι όπως στην παρακάτω εικόνα.

Καμπύλη



Ευθεία



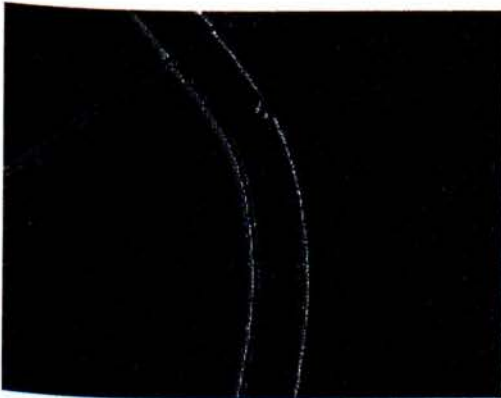
Αφού ισορροπήσουμε τον φωτισμό της εικόνας σε καλά επίπεδα, θα πρέπει να κάνουμε μια αναγνώριση ακμών ώστε το λογισμικό μας να αναγνωρίσει την γραμμή στο έδαφος. Δεν μας ενδιαφέρει το χρώμα της γραμμής αφού ο τρόπος που διαλέγουμε να επεξεργαστούμε την εικόνα βασίζεται στο σχήμα και όχι στο χρώμα. Θα εφαρμόσουμε ένα φίλτρο συνεξέλιξης για να το πετύχουμε αυτό.

Το φίλτρο μας στις αρχικές τιμές από το πρόγραμμα είναι ένας πίνακας αυτής της μορφής:

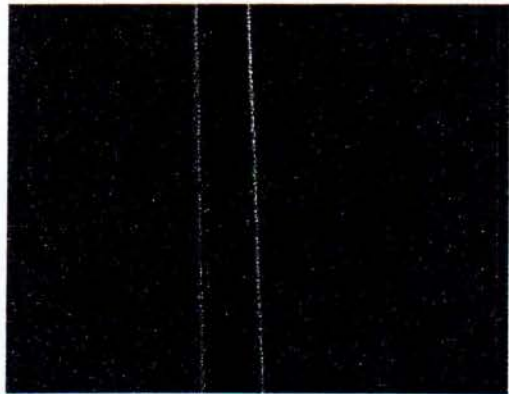
-1	-1	-1
-1	8	-1
-1	-1	-1

Το αποτέλεσμα αυτού φαίνεται στην παρακάτω εικόνα

Καμπύλη



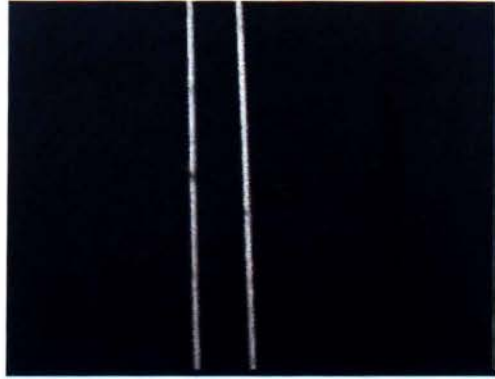
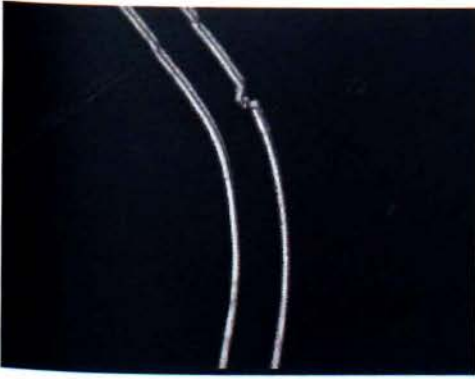
Ευθεία



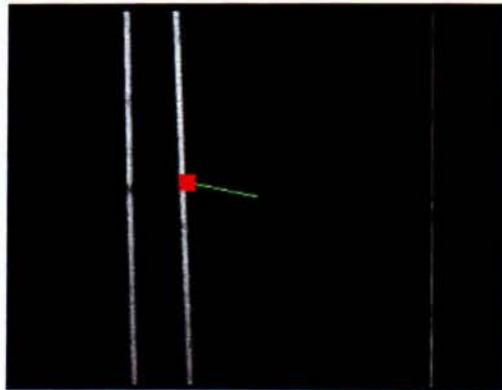
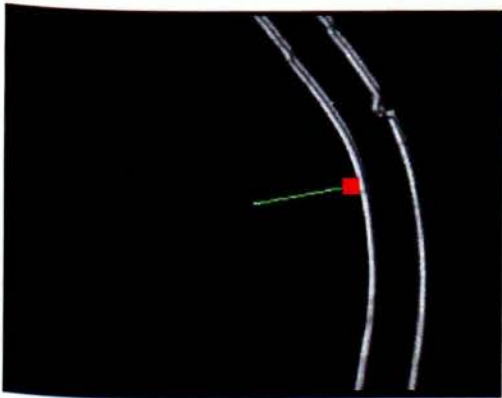
Το αποτέλεσμα κρίνεται μη ικανοποιητικό και γι αυτό θα χρησιμοποιήσουμε ένα φίλτρο συνεξέλιξης 5x5 της μορφής

-1	-1	-1	-1	-1
-1	0	0	0	-1
-1	0	16	0	-1
-1	0	0	0	-1
-1	-1	-1	-1	-1

που θα έχει σαν αποτέλεσμα να μας δώσει ακμές μεγαλύτερου πάχους. Πράγματι όπως βλέπουμε και στις παρακάτω εικόνες τα αποτελέσματα είναι σαφώς καλύτερα.



Στο επόμενο βήμα μας θα χρησιμοποιήσουμε την λειτουργία του Roboreal, που ονομάζεται Κέντρο βάρους (Center of Gravity – COG). Το Κέντρο βάρους πολύ απλά βλέπει τα λευκά pixel στην εικόνα και προσπαθεί να τα μοιράσει ισόποσα δεξιά και αριστερά του κεντρου της εικόνας. Αυτό δημιουργεί την ανάγκη στο όχημα να κρατήσει τις ακμές της γραμμής που έχουμε ήδη ορίσει όσο περισσότερο στο κέντρο της εικόνας μπορεί. Με απλά λογια προσπαθεί να ισορροπήσει πάνω στην γραμμή καθώς μετακινείται εμπρός. Τα αποτελέσματα είναι τα εξής.



Βασιζόμενοι τώρα στην θέση του Κέντρου βάρους μπορούμε να δώσουμε τις εξής οδηγίες στο οχημά μας:

- Όταν το κέντρο βάρους βρίσκεται στην αριστερή πλευρά της οθόνης στρίψε αριστερά
- Όταν το κέντρο βάρους βρίσκεται στην δεξιά πλευρά της οθόνης στρίψε δεξιά
- Όταν το κέντρο βάρους βρίσκεται στο κέντρο της οθόνης διατήρησε την πορεία σου προς τα εμπρός

Οι παραπάνω απλές οδηγίες θα μεταφραστούν μέσω του προγράμματος σε κατάλληλους παλμούς που θα οδηγήσουν αποτελεσματικά τα servo και αυτά με την σειρά τους το όχημα μας πάνω στην επιθυμητή πορεία.

ΠΡΟΤΑΣΕΙΣ ΓΙΑ ΜΕΛΛΟΝΤΙΚΗ ΕΡΓΑΣΙΑ

Κατα την διεξαγωγή της πτυχιακής εργασίας μελετήσαμε κατα κύριο λόγο τις μεθόδους και τις τεχνικές που απαιτούνται ώστε ένα όχημα να κινηθεί αυτόνομα μέσα στον χώρο. Η χρήση του λογισμικού RoborealM απαιτεί καλή γνώση υπολογιστών καθώς και του θεωρητικού μέρους της επεξεργασίας εικόνας ώστε να έχουμε τα επιθυμητά αποτελέσματα.

Σαν μελλοντική εργασία θα μπορούσαν να τεθούν προς διερεύνηση πολύπλοκότερα προβλήματα όπως η κίνηση του οχήματος μέσα σε ένα περιβάλλον με περισσότερους κανόνες. Αυτό θα έχει σαν αποτέλεσμα την ανάγκη σχεδίασης προγραμμάτων αρκετά εκτενέστερων και αρκετά πιο πολύπλοκων που θα μπορούν όμως να ανταποκρίνονται σε πιο ρεαλιστικές ανάγκες και καταστάσεις.

Συμβαδίζοντας με την εξέλιξη της τεχνολογίας θα μπορούσαν να χρησιμοποιηθούν νέοι αλγόριθμοι επεξεργασίας εικόνας καθώς και νέου τύπου υποσυστήματα, που αν και θα ανέβαζαν το κόστος της κατασκευής θα βελτίωναν σημαντικά τις δυνατότητες του οχήματος.

Τέλος θα μπορούσαν οι τεχνικές που αναφέραμε να χρησιμοποιηθούν απο ένα στατικό σύστημα και να πράξουν εντελώς διαφορετικές εργασίες όπως η αναγνώριση αντικειμένων του πραγματικού κόσμου, η διαλογή και η ταξινόμησή τους βάσει προκαθορισμένου σετ κανόνων, όπως για παράδειγμα σε ένα εργοστάσιο ανακύκλωσης.

Σε κάθε περίπτωση, πρωταρχικό ρόλο έχει το ενδιαφέρον που μπορεί κανείς να διαθέτει καθώς επίσης και το υπόβαθρο των γνώσεων, τόσο των θεωρητικών όσο και των πρακτικών ανάλογα με την πολυπλοκότητα της εφαρμογής.

ΒΙΒΛΙΟΓΡΑΦΙΑ

1. Dynamic Vision For Perception And Control Of Motion – Ernst D. Dickmanns
2. Visual Perception and robotic manipulation – Taylor & Kleeman
3. Computer Vision A Modern Approach - Forsyth , Ponce
4. Robot vision – Berthold Klaus & Paul Horn
5. Robot Vision - Stefan Florezyk
6. Τρισδιάστατη Ανακατασκευή Αντικειμένων και Κτηρίων με Γεωμετρία Πολλαπλών Όψεων – Γεωργία Γκοξιάρη (διπλωματική εργασία Ε.Μ.Π)
7. Πηγές Internet
http://en.wikipedia.org/wiki/Computer_vision
<http://www.roborealm.com/>
http://en.wikipedia.org/wiki/Driverless_car
<http://www.endurance-rc.com/pctx.php>
<http://www.gizmag.com/>
<http://ocw.mit.edu/courses/electrical-engineering-and-computer-science/>
<http://ocw.mit.edu/courses/electrical-engineering-and-computer-science/6-801-machine-vision-fall-2004/>