



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΣΗΜΑΤΩΝ, ΕΛΕΓΧΟΥ ΚΑΙ ΡΟΜΠΟΤΙΚΗΣ

**Βαθμονόμηση Ρομποτικού Βραχίονα – Μηχανικής
Όρασης και Σχεδιασμός Συστήματος Ελέγχου
Παρακολούθησης Τροχιάς**

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

Δημήτριος Γ. Αρίστος

Αθήνα, Μάρτιος 2012



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
 ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
 ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
 ΤΟΜΕΑΣ ΣΗΜΑΤΩΝ, ΕΛΕΓΧΟΥ ΚΑΙ ΡΟΜΠΟΤΙΚΗΣ

Βαθμονόμηση Ρομποτικού Βραχίονα – Μηχανικής Όρασης και Σχεδιασμός Συστήματος Ελέγχου Παρακολούθησης Τροχιάς

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

Δημήτριος Γ. Αρίστος

Συμβουλευτική Επιτροπή: Σπυρίδων Τζαφέστας
 Γεώργιος Παπαβασιλόπουλος
 Νικόλαος Μαράτος

Εγκρίθηκε από την επταμελή εξεταστική επιτροπή την 15^η Μαρτίου 2012.

.....
 Σ. Τζαφέστας
 Ομ. Καθηγητής ΕΜΠ

.....
 Γ. Παπαβασιλόπουλος
 Καθηγητής ΕΜΠ

.....
 Ν. Μαράτος
 Καθηγητής ΕΜΠ

.....
 Γ. Στασινόπουλος
 Καθηγητής ΕΜΠ

.....
 Κ. Πεκμεστζή
 Καθηγητής ΕΜΠ

.....
 Ν. Σγούρος
 Καθηγητής Πανεπιστημίου
 Πειραιά

.....
 Γ. Στάμου
 Λέκτορας ΕΜΠ

Αθήνα, Μάρτιος 2012

.....

Δημήτριος Γ. Αρίστος

Διδάκτωρ Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών Ε.Μ.Π.

Copyright Δημήτριος Γ. Αρίστος, 2012.

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα.

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου

ΠΡΟΛΟΓΟΣ

Η παρούσα διατριβή αποτελεί αποτέλεσμα πολύχρονης προετοιμασίας και προσπάθειας τόσο σε θεωρητικό επίπεδο για το χτίσιμο του απαραίτητου θεωρητικού υπόβαθρου πάνω στο οποίο βασίζεται η διατριβή, όσο και σε εργαστηριακό και πειραματικό επίπεδο για την πρακτική εφαρμογή κι επιβεβαίωση των προτεινόμενων τεχνικών και μεθόδων. Η φιλοσοφία καθ' όλη τη διάρκεια της εκπόνησης της διατριβής ήταν η επέκταση υπαρχόντων επιστημονικών μεθόδων στον ευρύτερο κλάδο της επιστήμης της ρομποτικής, αυτομάτου ελέγχου και μηχανικής όρασης, καθώς και η ανακάλυψη νέων αλγορίθμων και τεχνικών οι οποίες παρέχουν μεγαλύτερη ευκολία και καλύτερη ακρίβεια αποτελεσμάτων κατά την εφαρμογή τους σε σύγκριση με τις υπάρχουσες διαθέσιμες μεθοδολογίες. Η πρακτική εφαρμογή των διαφόρων επιστημονικών μεθόδων που παρουσιάζονται σε αυτή τη διατριβή ήταν μια εξαιρετικά δύσκολη, χρονοβόρα κι επίπονη διαδικασία, τόσο σε επίπεδο προετοιμασίας και αντιμετώπισης των διαφόρων τεχνικών δυσκολιών όσο και σε επίπεδο επιβεβαίωσης και τεκμηρίωσης των θεωρητικών αναμενόμενων αποτελεσμάτων. Υπήρχαν πολλές φορές όπου μια φαινομενικά άψογη θεωρητικά επιστημονική μέθοδος είτε αντιμετώπιζε αξεπέραστα προβλήματα στην πρακτική εφαρμογή της με αποτέλεσμα να καθίσταται πρακτικά άχρηστη, είτε δεν έδινε τα αναμενόμενα πειραματικά αποτελέσματα εξ' αιτίας παραγόντων που δεν είχαν προβλεφθεί αρχικά. Η παρούσα διατριβή επικεντρώνεται στις μεθόδους εκείνες που έχουν άμεση πρακτική εφαρμογή και μπορούν να χρησιμοποιηθούν σε αυτοματοποιημένες γραμμές παραγωγής στη βιομηχανία, παρέχοντας αυτοματοποιημένες λύσεις σε συστήματα ρομποτικής και μηχανικής όρασης.

Παρ' όλο που η περάτωση της διατριβής αποτελεί αποτέλεσμα μακροχρόνιας προσωπικής προσπάθειας, αυτή δεν θα ήταν δυνατή χωρίς τη στήριξη κάποιων ανθρώπων, των οποίων η βοήθεια ήταν ανεκτίμητη. Κατ' αρχάς θα ήθελα να ευχαριστήσω τον καθηγητή μου Κ. Τζαφέστα για την στήριξη που μου παρείχε καθ' όλη τη διάρκεια της εκπόνησης της διατριβής μου. Επίσης θα ήθελα να ευχαριστήσω όλους τους συναδέλφους μου στο εργαστήριο ρομποτικής οι οποίοι με βοήθησαν πολλές φορές με την αμέριστη συνεργασία και συμπαράστασή τους. Τέλος θα ήθελα ν' αναφερθώ σε δύο συναδέλφους και φίλους μου Θ. Κρικοχωρίτη και Γ. Πανταζόπουλο, η βοήθεια των οποίων ήταν ανεκτίμητη για την περάτωση της παρούσας διατριβής.

ΠΕΡΙΛΗΨΗ

Η διατριβή πραγματεύεται την μελέτη, υλοποίηση κι αξιολόγηση υπολογιστικών αλγορίθμων, ο συνδυασμός των οποίων χρησιμοποιείται για την ανάπτυξη και κατασκευή ρομποτικών βιομηχανικών εφαρμογών που στοχεύουν στην παρακολούθηση κινούμενων αντικειμένων από το άκρο ενός βιομηχανικού ρομποτικού βραχίονα με βάση την πληροφορία (ανάδραση) που λαμβάνεται από μία ή περισσότερες ψηφιακές κάμερες. Η κατηγορία αυτών των ρομποτικών εφαρμογών αναφέρεται στην διεθνή βιβλιογραφία ως «*visual servoing*» και αποτελεί αντικείμενο έρευνας για πολλά χρόνια. Στις περισσότερες περιπτώσεις μία ή περισσότερες ψηφιακές κάμερες παρατηρούν κάποιο χαρακτηριστικό σημείο ενός κινούμενου αντικειμένου (πχ γωνία ή οπή) και χρησιμοποιώντας κάποιο μοντέλο ελέγχου μεταφράζουν την μεταβολή στις μετρήσεις των καμερών σε μια αντίστοιχη κίνηση του άκρου του ρομποτικού βραχίονα. Αντικείμενο έρευνας της διατριβής είναι η σχεδίαση και δημιουργία ρομποτικών συστημάτων παρακολούθησης τροχιάς, όπου το παρατηρούμενο αντικείμενο δεν έχει εμφανή χαρακτηριστικά υφής τα οποία μπορούν ν' ανιχνευθούν και να μετρηθούν από τους κλασικούς αλγορίθμους επεξεργασίας εικόνας. Οι περιπτώσεις αυτές είναι πολύ συχνές σε βιομηχανικές εφαρμογές όπου απαιτείται η παρακολούθηση κι επεξεργασία κινούμενων αντικειμένων με έντονες καμπύλες και ελάχιστα χαρακτηριστικά σημεία.

Η μελέτη κι ενασχόληση με το παραπάνω αντικείμενο είχε σαν αποτέλεσμα την ανάπτυξη κι εφαρμογή καινοτόμων τεχνικών και μεθοδολογιών, οι οποίες συνέβαλαν στην υλοποίηση του ρομποτικού συστήματος παρακολούθησης τροχιάς. Η πρώτη καινοτομία αφορά την χρησιμοποίηση μετρήσεων ανακατασκευής σημείων για τον υπολογισμό της σχετικής θέσης και προσανατολισμού του άκρου ενός βιομηχανικού ρομποτικού βραχίονα και ενός συστήματος μηχανικής όρασης, το οποίο είναι σταθερά στερεωμένο στο άκρο του βραχίονα. Ο υπολογισμός της σχετικής θέσης και προσανατολισμού ενός συστήματος μηχανικής όρασης και ενός βιομηχανικού ρομποτικού βραχίονα αναφέρεται στην διεθνή βιβλιογραφία ως «*Hand-Eye Calibration*» και «*Robot-World Calibration*» και σε όλες τις περιπτώσεις χρησιμοποιεί τους μετασχηματισμούς από την βαθμονόμηση των καμερών. Η νέα μέθοδος ανακατασκευής σημείων έδωσε συνολικά καλύτερα αποτελέσματα σε σύγκριση με τις υπάρχουσες κλασικές μεθόδους που χρησιμοποιούν μετασχηματισμούς βαθμονόμησης κάμερας, αποδεικνύοντας με αυτό τον τρόπο την ανωτερότητά της.

Η δεύτερη καινοτομία αφορά την χρήση προηγμένων αλγορίθμων υπολογισμού θέσης και προσανατολισμού ενός παρατηρούμενου αντικειμένου χωρίς ιδιαίτερα χαρακτηριστικά υφής (πχ οπές ή γωνίες). Ο υπολογισμός πραγματοποιείται χρησιμοποιώντας τα κατασκευαστικά δεδομένα CAD του αντικειμένου καθώς κι ένα πλήθος δεδομένων ανακατασκευής σημείων από την επιφάνειά του, τα οποία προέρχονται από μετρήσεις μιας μετρητικής κεφαλής αποτελούμενης από μια ψηφιακή κάμερα κι ένα λείζερ πολλαπλών γραμμών. Οι συντεταγμένες των ανακατασκευασμένων σημείων και τα αντίστοιχα CAD δεδομένα της επιφάνειας του αντικειμένου σχετίζονται μεταξύ τους με την εφαρμογή ενός επαναληπτικού αλγορίθμου ICP. Για την βελτιστοποίηση της χρονικής απόκρισης του αλγορίθμου τα CAD δεδομένα έχουν ομαδοποιηθεί με την χρήση δυαδικών δένδρων K-D trees, έτσι ώστε οι υπολογισμοί ελάχιστης απόστασης που απαιτούνται από τον αλγόριθμο ICP να επιταχύνονται

σημαντικά. Η εφαρμογή τέτοιων προηγμένων μεθόδων υπολογισμού θέσης και προσανατολισμού κινούμενων αντικειμένων είναι μια καινοτομία στον χώρο των ρομποτικών συστημάτων παρακολούθησης τροχιάς.

Η διατριβή αυτή αποτελεί σημαντική συνεισφορά στην επιστήμη και τεχνολογία του αυτομάτου ελέγχου, ρομποτικής και μηχανικής όρασης προτείνοντας νέες αρχιτεκτονικές στη σχεδίαση ρομποτικών εφαρμογών παρακολούθησης τροχιάς κινούμενου αντικειμένου και προσφέρει νέες μεθοδολογίες με πολλαπλές χρήσιμες εφαρμογές. Όλες οι καινοτόμες μέθοδοι που αναπτύχθηκαν τεκμηριώνονται θεωρητικά και μαθηματικά, ενώ τα πειραματικά αποτελέσματα αποδεικνύουν την εφικτότητα και την ακρίβεια που προσφέρουν, καθώς και την γρήγορη χρονική απόκρισή τους που τις κάνει ελκυστικές σε πραγματικές ρομποτικές εφαρμογές.

Αντικείμενο μελλοντικής εργασίας αποτελεί η περαιτέρω επιτάχυνση του ρομποτικού συστήματος παρακολούθησης τροχιάς, χρησιμοποιώντας μεθόδους παράλληλης επεξεργασίας σε πολυπύρηνους επεξεργαστές. Στόχος είναι η επίτευξη μεγαλύτερων ταχυτήτων απόκρισης του συστήματος παρακολούθησης, έτσι ώστε ο ρομποτικός βραχίονας να είναι σε θέση να παρακολουθεί αντικείμενα που κινούνται στον χώρο με μεγαλύτερες ταχύτητες. Επίσης αντικείμενο μελλοντικής εργασίας αποτελεί η επίτευξη μιας συνεχούς ομαλής κίνησης παρακολούθησης από τον βραχίονα, όπου ο υπολογισμός της θέσης του κινούμενου αντικειμένου και η διόρθωση τροχιάς του βραχίονα θα πραγματοποιούνται σε συνεχή χρόνο με πολύ μεγάλη ταχύτητα. Τέλος, είναι πιθανή η χρήση κάποιου καταλληλότερου για συστήματα συνεχούς χρόνου νόμου ελέγχου, ο οποίος θα περιλαμβάνει λειτουργίες πρόβλεψης κατάστασης για την αντιστάθμιση του φαινομένου υστέρησης του συστήματος.

Λέξεις Κλειδιά

Βαθμονόμηση κάμερας, Βαθμονόμηση ρομποτικού βραχίονα – μηχανικής όρασης, Παρακολούθηση τροχιάς, Βιομηχανικοί ρομποτικοί βραχίονες, Αλγόριθμος ICP, Δυαδικά δένδρα K-D trees.

SUMMARY

This thesis researches the computational algorithms that are used in developing visual tracking robotic applications, where an industrial robot arm tries to follow the displacements of a moving object based on measurement results acquired from a vision system. This category of robotic applications is usually referred to as “visual servoing applications” and is a subject of research for many years. In most cases a vision system consisted of one or more cameras observes a characteristic feature on the object’s surface (ex. hole or corner) and uses image processing techniques to extract the feature’s location in image coordinates. A proper control law is usually commissioned in order to relate the variance of the sensor measurements with the corresponding movement of the robot arm. This thesis examines the case that no characteristic features can be detected on the surface of the observed object, which can happen if the object has smooth surface without obvious texture features. This case is very common in industrial robotic applications, where some rigid objects that need to be tracked by a robot arm have smooth and featureless surfaces.

Throughout the development of the discussed visual tracking application, various innovative methods and algorithms were developed and tested. The first innovation concerns the calculation of the relative position and orientation of the vision system’s coordinate frame in reference to the robot tool frame. This calculation is usually referred to as “Hand-Eye Calibration” and all currently proposed methods make use of the 3D transformations derived from the calibration of the cameras. The innovation of the proposed method lies in the use of 3D reconstructed points for the calculation of the Hand-Eye transformation rather than the transformation matrices derived from camera calibration. The superiority of the proposed method is theoretically justified and experimentally proven by means of multiple experimental results.

The second innovation concerns the use of 3D surface reconstruction techniques in visual tracking robotic applications, by means of calculating the position and orientation of an observed moving object. Due to the lack of characteristic features on the observed object, its whole visible surface had to be reconstructed by using a vision system consisting of a camera and a multiple laser line device. The cloud of reconstructed points on the object’s surface was correlated with the corresponding CAD data of the object’s model by applying the ICP algorithm. In this way the 3D transformation relating the CAD model’s coordinate frame and the vision system’s measurement frame could be extracted, thus the position and orientation of the observed object could be found. In order to speed up the ICP algorithm the nearest neighbor point extraction function (used by ICP) was accelerated by using 3D space sub-division based on KD-trees. The accuracy and time performance of the proposed method for calculating the position and orientation of an observed object renders the method attractive for practical visual tracking robotic application. The feasibility and performance of the complete visual tracking system is proven by presenting multiple experimental results.

The contribution of this thesis in the scientific fields of automation control and robotics is significant, as many of the presented methods and algorithms constitute an innovation to the corresponding fields. The presented theoretical background and the extensive experimental results provide a complete study for building real-world visual tracking robotic applications.

Throughout the development of the discussed visual tracking application significant effort has been made for optimizing the time performance of the system, especially the algorithms used by the vision system for calculating the position and orientation of the observed object. Although the use of KD-trees significantly increased time performance, the vision system is not fast enough yet for use in time continuous tracking applications, where a robot arm is required to continuously track the displacement of a moving object in real time. Our goal is to use parallel processing techniques in multi-core processors in order to increase the time performance of the vision system, so that the robot arm will be able to track high-speed moving objects by performing an uninterrupted continuous path motion. Furthermore, a more sophisticated control law may be needed for handling a time continuous tracking application, which should include state prediction functionality for compensating for the position hysteresis of the tracking system.

Keywords

Camera calibration, Hand-Eye and Robot-World calibration, Visual tracking, Visual servoing, Industrial robots, ICP algorithm, K-D trees.

ΠΕΡΙΕΧΟΜΕΝΑ

Λίστα σχημάτων	13
1 ΕΙΣΑΓΩΓΗ	15
1.1 Γενικά χαρακτηριστικά αντικειμένου διατριβής	15
1.2 Αντικειμενικός σκοπός και καινοτομία διατριβής	24
1.3 Οργάνωση και δομή της διατριβής	27
2 ΜΗΧΑΝΙΚΗ ΟΡΑΣΗ	29
2.1 Γενικά περί μηχανικής όρασης	29
2.2 Γεωμετρία εικόνας	30
2.3 Βαθμονόμηση κάμερας	35
2.3.1 Αλγόριθμος βαθμονόμησης κάμερας του Z. Zhang	38
2.3.2 Αλγόριθμος βαθμονόμησης κάμερας του R. Tsai.....	45
2.4 Ψηφιακή επεξεργασία εικόνας	53
2.5 Ανακατασκευή τρισδιάστατων σημείων	59
3 ΒΑΘΜΟΝΟΜΗΣΗ ΒΡΑΧΙΟΝΑ – ΜΗΧΑΝΙΚΗΣ ΟΡΑΣΗΣ	65
3.1 Σημασία συσχέτισης συντεταγμένων κάμερας και βραχίονα	65
3.2 Βαθμονόμηση βραχίονα – μηχανικής όρασης για μονή κάμερα.....	69
3.3 Βαθμονόμηση βραχίονα – μηχανικής όρασης για στερεοσκοπική κεφαλή.....	73
3.4 Πειραματικά αποτελέσματα.....	76
4 ΤΑΥΤΟΧΡΟΝΗ ΒΑΘΜΟΝΟΜΗΣΗ ΒΡΑΧΙΟΝΑ, ΜΗΧΑΝΙΚΗΣ ΟΡΑΣΗΣ ΚΑΙ ΧΩΡΟΥ ΕΡΓΑΣΙΑΣ	80
4.1 Ταυτόχρονη βαθμονόμηση βραχίονα, μηχανικής όρασης και χώρου εργασίας για μονή κάμερα.....	80
4.1.1 Γραμμική μέθοδος	82
4.1.2 Αναλυτική μέθοδος.....	84
4.1.3 Μη-γραμμική μέθοδος ελαχιστοποίησης.....	88
4.1.4 Υπολογισμός διανυσμάτων μετατόπισης.....	90
4.2 Ταυτόχρονη βαθμονόμηση βραχίονα, μηχανικής όρασης και χώρου εργασίας για στερεοσκοπική κεφαλή.....	91
4.3 Πειραματικά αποτελέσματα.....	96
5 Έλεγχος τροχιάς ρομποτικού βραχίονα με τη βοήθεια κάμερας και συσκευή λείζερ	105
5.1 Γενική περιγραφή προτεινόμενου συστήματος ελέγχου	105
5.2 Βαθμονόμησης συστοιχίας κάμερας - συσκευής λείζερ	108
5.3 Πραγματοποίηση μετρήσεων με την συστοιχία κάμερας -συσκευής λείζερ	112
5.4 Ο αλγόριθμος ICP	113
5.5 Επιτάχυνση του αλγορίθμου ICP με τη χρήση των κατανομών K-D trees.....	116

5.6	Συσχέτιση συστημάτων συντεταγμένων μέτρησης κάμερας και άκρου ρομποτικού βραχίονα	124
5.7	Συνδυασμός διακριτών μεθόδων και πειραματικά αποτελέσματα.....	125
6	ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΔΗΜΟΣΙΕΥΣΕΙΣ	132
6.1	Συμπεράσματα διατριβής.....	132
6.2	Δημοσιεύσεις.....	133
	ΒΙΒΛΙΟΓΡΑΦΙΑ	135

ΛΕΞΙΚΟ ΚΥΡΙΩΝ ΕΠΙΣΤΗΜΟΝΙΚΩΝ ΟΡΩΝ

3D reconstruction	Τρισδιάστατη ανακατασκευή
Camera Calibration	Βαθμονόμηση κάμερας
Feedback	Ανατροφοδότηση (ανάδραση)
Hand-Eye	Βραχίονας-κάμερα (μηχανικό μάτι)
Hand-Eye Calibration	Βαθμονόμηση βραχίονα - μηχανικής όρασης
Laser Scanner	Σαρωτής λέιζερ
Loop	Βρόχος
Pixel	Εικονοστοιχείο
Robot-World Calibration	Βαθμονόμηση βραχίονα και χώρου εργασίας ρομπότ
Servomotor	Σερβοκινητήρας
Singular Value Decomposition (SVD)	Διάσπαση ιδιόμορφης τιμής
Stereo Head	Στερεοσκοπική κεφαλή
Visual Servoing	Οπτική παρακολούθηση

Λίστα σχημάτων

Σχήμα 1.1	Δομή συστήματος κλειστού βρόχου ελέγχου τροχιάς του άκρου βιομηχανικού βραχίονα, όπου το σφάλμα ελέγχου ανήκει στο χώρο της ψηφιακής εικόνας (2D).....	16
Σχήμα 1.2	Δομή συστήματος ελέγχου τροχιάς ρομπότ κλειστού βρόχου με καρτεσιανό νόμο ελέγχου	20
Σχήμα 1.3	Συσχετισμός συστημάτων συντεταγμένων ρομποτικού βραχίονα και κάμερας σταθερά στερεωμένης στο άκρου του.....	23
Σχήμα 2.1	Η προβολή ενός σημείου στο χώρο (x, y, z) πάνω στο πλάνο εικόνας της κάμερας.....	31
Σχήμα 2.2	Για την αποφυγή αντιστροφής της εικόνας, στο μοντέλο προβολής της κάμερας θεωρούμε ότι το πλάνο εικόνας είναι μπροστά από το κέντρο προβολής.....	32
Σχήμα 2.3	Επίπεδο μοντέλο βαθμονόμησης.....	36
Σχήμα 2.4	Ψηφιακή εικόνα με και χωρίς ακτινική παραμόρφωση.....	44
Σχήμα 2.5	Στέρεο (μη-επίπεδο) μοντέλο βαθμονόμησης κάμερας.....	46
Σχήμα 2.6α	Ψηφιακή εικόνα προς επεξεργασία Σχήμα 2.6β Μάσκα συνέλιξης.....	54
Σχήμα 2.7	Συνέλιξη αρχικής εικόνας και εικόνας-μάσκας.....	55
Σχήμα 2.8	Ανίχνευση περιγράμματος.....	56
Σχήμα 2.9	Βαθμονόμηση στερεοσκοπικής κεφαλής καμερών.....	61
Σχήμα 2.10	Γεωμετρική ερμηνεία της 3D ανακατασκευής σημείου.....	62
Σχήμα 2.11	Οπτικές ευθείες και βέλτιστη λύση ανακατασκευής σημείου.....	63
Σχήμα 3.1	Συστήματα συντεταγμένων ρομποτικού βραχίονα και κάμερας.....	66
Σχήμα 3.2	Παρακολούθηση τροχιάς κινούμενου σημείου.....	68
Σχήμα 3.3	Τοπολογία πειράματος βαθμονόμησης βραχίονα – μηχ. όρασης για μονή κάμερα.....	69
Σχήμα 3.4	Τοπολογία πειράματος βαθμονόμησης βραχίονα – μηχ. όρασης για στερεοσκοπική κεφαλή	73
Σχήμα 3.5	Ρυθμός σύγκλισης του διανύσματος μετατόπισης (αριστερά) και των γωνιών στροφής του εξαγόμενου μετασχηματισμού X.....	77
Σχήμα 3.6	Ρυθμός σύγκλισης του διανύσματος μετατόπισης (αριστερά) και των γωνιών στροφής του εξαγόμενου μετασχηματισμού X.....	78
Σχήμα 4.1	Τοπολογία πειράματος ταυτόχρονης βαθμονόμησης βραχίονα, μηχ. όρασης και χώρου εργασίας για μονή κάμερα.....	80
Σχήμα 4.2	Μοντέλο βαθμονόμησης με τέσσερις χαρακτηριστικούς κύκλους για μετρήσεις τρισδιάστατης ανακατασκευής.....	93
Σχήμα 4.3	Τοπολογία πειράματος ταυτόχρονης βαθμονόμησης βραχίονα, μηχ. όρασης και χώρου εργασίας για την περίπτωση στερεοσκοπικής κεφαλής.....	94
Σχήμα 4.4	Ρυθμός σύγκλισης των διανυσμάτων μετατόπισης των μετασχηματισμών X και Z για την τοπολογία μονής κάμερας.....	99
Σχήμα 4.5	Ρυθμός σύγκλισης των γωνιών στροφής Euler των μετασχηματισμών X και Z για την τοπολογία μονής κάμερας.....	99
Σχήμα 4.6	Ρυθμός σύγκλισης των διανυσμάτων μετατόπισης των μετασχηματισμών X και Z για την τοπολογία στερεοσκοπικής κεφαλής.....	100
Σχήμα 4.7	Ρυθμός σύγκλισης των γωνιών στροφής Euler των μετασχηματισμών X και Z για την τοπολογία στερεοσκοπικής κεφαλής.....	101
Σχήμα 4.8	Ρυθμός σύγκλισης των διανυσμάτων μετατόπισης των μετασχηματισμών X και Z για την τοπολογία στερεοσκοπικής κεφαλής και μικρές μετατοπίσεις του βραχίονα.....	103
Σχήμα 4.9	Ρυθμός σύγκλισης των γωνιών στροφής Euler των μετασχηματισμών X και Z για την τοπολογία στερεοσκοπικής κεφαλής και μικρές μετατοπίσεις του βραχίονα.....	104
Σχήμα 5.1	Τοπολογία πειράματος ελέγχου τροχιάς ρομποτικού βραχίονα με ανάδραση από συστοιχία κάμερας - συσκευής ακτίνων λέιζερ πολλαπλών γραμμών.....	106
Σχήμα 5.2	Μοντέλο CAD στέρεου σώματος.....	107
Σχήμα 5.3	Μη-επίπεδο μοντέλο βαθμονόμησης με την συσκευή γραμμών λέιζερ α) απενεργοποιημένη β) ενεργοποιημένη.....	109
Σχήμα 5.4	Προβολή γραμμής λέιζερ πάνω στην επιφάνεια του μοντέλου βαθμονόμησης.....	111
Σχήμα 5.5	Ψηφιακή εικόνα μιας μεταλλικής επιφάνειας, όταν οι γραμμές λέιζερ είναι ενεργοποιημένες.....	112

Σχήμα 5.6	Κατασκευή κατανομής KD-tree για ένα πλήθος 2D σημείων.	119
Σχήμα 5.7	Κατανομή KD-tree για το πλήθος των σημείων του Σχ. 5.6.....	120
Σχήμα 5.8	Εύρεση του σημείου ελάχιστης απόστασης με τη χρήση της κατανομής KD-tree.....	122
Σχήμα 5.9	Συστήματα και μετασχηματισμοί συντεταγμένων που χρησιμοποιούνται στην παρακολούθηση τροχιάς.....	127
Σχήμα 5.10	CAD μοντέλα δύο μεταλλικών επιφανειών.	129
Σχήμα 5.11	Ψηφιακές εικόνες των μεταλλικών επιφανειών του Σχ. 5.10.	130

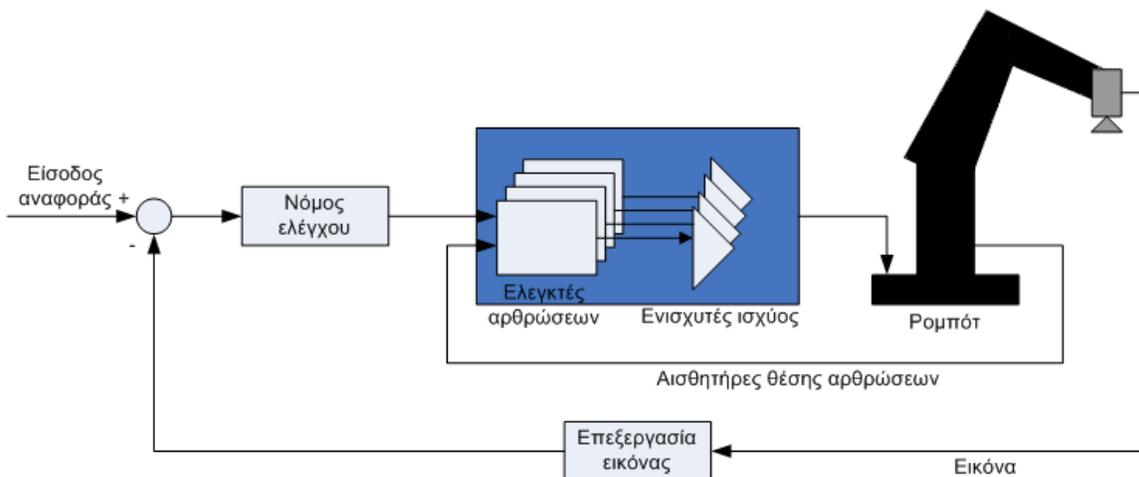
1 ΕΙΣΑΓΩΓΗ

1.1 Γενικά χαρακτηριστικά αντικειμένου διατριβής

Ένα μεγάλο μέρος της επιστήμης της ρομποτικής ασχολείται με τη μελέτη της κινηματικής και δυναμικής ανάλυσης των βιομηχανικών ρομποτικών βραχιόνων. Ένας βιομηχανικός ρομποτικός βραχίονας αποτελείται από έναν αριθμό συμπαγών συνδέσμων οι οποίοι συνδέονται μεταξύ τους με τη βοήθεια αρθρώσεων. Οι αρθρώσεις περιλαμβάνουν ηλεκτρικούς κινητήρες (σερβοκινητήρες) οι οποίοι μπορούν να δημιουργήσουν μηχανική ροπή ανάμεσα στους προσκείμενους συνδέσμους του βραχίονα και να προκαλέσουν μια σχετική κίνησή τους. Συνήθως ο πρώτος σύνδεσμος του ρομποτικού βραχίονα είναι σταθερά στερεωμένος σε κάποια σταθερή επιφάνεια αποτελώντας τη βάση του ρομπότ, ενώ ο τελευταίος σύνδεσμος περιλαμβάνει κάποιο εργαλείο το οποίο αναλαμβάνει να εκτελέσει μια συγκεκριμένη διεργασία. Ο βαθμός ελευθερίας του ρομπότ εκφράζεται με τον αριθμό των ανεξάρτητων μεταβλητών που καθορίζουν πλήρως τη μορφή του βραχίονα, και προφανώς ισούται με τον αριθμό των αρθρώσεων.

Μια άλλη κατηγορία προβλημάτων της ρομποτικής ασχολείται με τον έλεγχο της τροχιάς του άκρου ενός βιομηχανικού ρομποτικού βραχίονα με βάση την ανατροφοδότηση (feedback) που προέρχεται από κάποιον αισθητήρα. Υπάρχει ένας μεγάλος αριθμός τύπων αισθητήρων η χρήση των οποίων αποσκοπεί στην αίσθηση του περιβάλλοντος χώρου και των αλλαγών που συντελούνται σε αυτό. Μια κύρια κατηγορία αισθητήρων περιλαμβάνει οπτικούς αισθητήρες (κάμερες) οι οποίοι έχουν τη δυνατότητα να εγκατασταθούν είτε στο άκρο του ρομποτικού βραχίονα παρατηρώντας κάποια στοιχεία / ιδιότητες (features) του περιβάλλοντος χώρου, είτε σε κάποια σταθερή θέση του χώρου εργασίας του ρομπότ παρατηρώντας τη μεταβολή θέσης και τροχιάς του άκρου του βραχίονα. Άλλες κατηγορίες αισθητήρων που χρησιμοποιούνται σε ανάλογες περιπτώσεις περιλαμβάνουν αισθητήρες υπερήχων (sonar sensors), αισθητήρες ακτίνας φωτός (laser range sensors), αισθητήρες δύναμης, και άλλες λιγότερο χρησιμοποιούμενες κατηγορίες.

Στο πρόβλημα ελέγχου της τροχιάς του άκρου ενός ρομποτικού βραχίονα με βάση την πληροφορία που λαμβάνεται από έναν αισθητήρα μπορούμε να εφαρμόσουμε δύο δυνατές μεθόδους: α) Τη μέθοδο ανοιχτού βρόχου, όπου τα δεδομένα από τον αισθητήρα χρησιμοποιούνται μόνο μια φορά στην αρχή της διαδικασίας, από τα οποία με διάφορους τρόπους και τεχνικές υπολογίζεται η τροχιά που πρέπει ν' ακολουθήσει το άκρο του βραχίονα ώστε να εκτελέσει επιτυχώς την απαιτούμενη διεργασία. Στη συνέχεια οι συντεταγμένες της επιθυμητής τροχιάς στέλνονται στον ελεγκτή του ρομπότ, ο οποίος με τη σειρά του οδηγεί το άκρο του βραχίονα πάνω στην ζητούμενη τροχιά. Είναι απαραίτητο να διευκρινιστεί ότι ο αναφερόμενος ελεγκτής του ρομπότ αποτελεί μέρος του βιομηχανικού ρομπότ και δεν είναι κάποιος ελεγκτής που σχεδιάζεται και υλοποιείται για την εκάστοτε εφαρμογή. β) η μέθοδος κλειστού βρόχου, όπου ο αισθητήρας λαμβάνει συνέχεια μετρήσεις από το περιβάλλον καθ' όλη τη διάρκεια της κίνησης του βραχίονα, οι οποίες χρησιμοποιούνται σε πραγματικό χρόνο (real-time) για τη διόρθωση της τρέχουσας θέσης και τον υπολογισμό της μελλοντικής επιθυμητής τροχιάς. Ουσιαστικά δημιουργείται ένα σύστημα αυτομάτου ελέγχου κλειστού βρόχου, το οποίο συνεχώς ελέγχει και διορθώνει την θέση και τον προσανατολισμό του άκρου του βραχίονα με βάση την ανατροφοδότηση από τις μετρήσεις του αισθητήρα. Στο Σχ. 1.1 φαίνεται η δομή του παραπάνω συστήματος αυτομάτου ελέγχου.



Σχήμα 1.1 Δομή συστήματος κλειστού βρόχου ελέγχου τροχιάς του άκρου βιομηχανικού βραχίονα, όπου το σφάλμα ελέγχου ανήκει στο χώρο της ψηφιακής εικόνας (2D)

Στο παραπάνω σχήμα οι ελεγκτές των αρθρώσεων και οι ενισχυτές ισχύος περιλαμβάνονται στο ρομποτικό ελεγκτή (robot controller) και παρέχονται ως έχουν από τον κατασκευαστή του εκάστοτε ρομποτικού συστήματος. Αυτό σημαίνει ότι ο σχεδιαστής του παραπάνω συστήματος κλειστού βρόχου θα πρέπει να υλοποιήσει τον κατάλληλο για την εκάστοτε εφαρμογή νόμο ελέγχου, καθώς και το κομμάτι της επεξεργασίας εικόνας που αναλύει τις λαμβανόμενες από την κάμερα εικόνες κι αναλαμβάνει να εξαγάγει χρήσιμη πληροφορία απ' αυτές. Η χρήσιμη πληροφορία ποικίλει ανάλογα με την εφαρμογή και μπορεί να συνίσταται στον εντοπισμό των συντεταγμένων εικόνας (image coordinates) ενός παρατηρούμενου χαρακτηριστικού του περιβάλλοντος, την αναγνώριση αντικειμένων ανάλογα με κάποιες γεωμετρικές ιδιότητες ή πρότυπα, κ.α. Ο τεχνολογικός κλάδος που ασχολείται με την επεξεργασία εικόνας είναι εκτενέστατος και χρησιμοποιεί εξειδικευμένους αλγόριθμους και φίλτρα για την εξαγωγή χρήσιμης πληροφορίας από ψηφιακές εικόνες. Ο νόμος ελέγχου αποτελεί το πιο βασικό κομμάτι της εφαρμογής καθώς καθορίζει σε πολύ μεγάλο βαθμό την εφικτότητα και την ακρίβεια του συστήματος. Σύμφωνα με το παραπάνω σχήμα, το σφάλμα του συστήματος ελέγχου (είσοδος αναφοράς – μετρούμενα χαρακτηριστικά λαμβανόμενης εικόνας) ανήκει στον χώρο της ψηφιακής εικόνας, ο οποίος είναι δύο διαστάσεων και χαρακτηρίζεται από συντεταγμένες εικόνας. Ο νόμος ελέγχου θα πρέπει να έχει την ιδιότητα να μετατρέπει το σφάλμα ελέγχου (2D) σε διόρθωση τροχιάς του άκρου του ρομπότ (3D).

Στη γενική περίπτωση όπου έχουμε k κάμερες πάνω στο άκρο του ρομπότ, και από την κάθε λαμβανόμενη εικόνα παίρνουμε πληροφορία δύο διαστάσεων (σε x , y συντεταγμένες εικόνας) τότε η διάσταση του σφάλματος ελέγχου θα είναι $m=2k$. Επίσης αν ο ρομποτικός βραχίονας έχει d βαθμούς ελευθερίας (συνήθως $d=6$ για τους εμπορικούς ρομποτικούς βραχίονες) τότε η έξοδος του νόμου ελέγχου θα πρέπει να δίνει ένα d -διάστατο διάνυσμα ώστε να ελέγξει τις θέσεις όλων των αρθρώσεων (βαθμούς ελευθερίας) του βραχίονα. Επίσης υπάρχουν εφαρμογές όπου ζητείται, εκτός από τον έλεγχο θέσης του άκρου του ρομπότ, και ο έλεγχος της απαιτούμενης στιγμιαίας ταχύτητάς του ανάλογα με τις εκάστοτε απαιτήσεις της εφαρμογής. Σε αυτή την περίπτωση το κομμάτι του ελεγκτή που αποτελεί τον νόμο ελέγχου θα πρέπει να δίνει ένα διάνυσμα εξόδου διάστασης $n=2d$, το οποίο θα καθορίζει εκτός τις μεταβλητές θέσης

των d αρθρώσεων του ρομπότ και τις μεταβλητές των ταχυτήτων τους. Συνοψίζοντας από τα παραπάνω, στην γενική περίπτωση σχεδιασμού ελεγκτή κλειστού βρόχου για τον έλεγχο τροχιάς ενός ρομποτικού βραχίονα d βαθμών ελευθερίας, βασιζόμενοι σε πληροφορία ανάδρασης από k ψηφιακές κάμερες, απαιτείται ο σχεδιασμός ενός νόμου ελέγχου που μεταφράζει το σήμα σφάλματος (διάστασης $m=2k$) στο αντίστοιχο διάνυσμα ελέγχου θέσης-ταχύτητας του άκρου του ρομπότ (διάστασης $n=2d$).

Δεδομένου ότι ο νόμος ελέγχου θα πρέπει να συνδέει μεταβολές εισόδων και εξόδων, δηλαδή τις μεταβολές (σφάλματα) από τον χώρο των ψηφιακών εικόνων, διάστασης m , σε διορθώσεις της τροχιάς (θέσης-ταχύτητας) των αρθρώσεων του ρομπότ διάστασης n , συμπεραίνουμε ότι ο νόμος ελέγχου θα πρέπει να περιλαμβάνει ένα *Ιακωβιανό πίνακα* (*Jacobian Matrix*) διαστάσεων $m \times n$, ο οποίος αποτελείται από μερικές παραγώγους και δίνεται από την παρακάτω εξίσωση.

$$\mathbf{J} = \begin{matrix} m \times n \\ \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix} \end{matrix} \quad (1)$$

Προφανώς \mathbf{f} είναι το m -διάστατο διάνυσμα των συντεταγμένων εικόνας που προέρχονται από την επεξεργασία των ψηφιακών εικόνων, και \mathbf{x} είναι το n -διάστατο διάνυσμα θέσης-ταχύτητας των αρθρώσεων του ρομπότ. Η εξίσωση που συνδέει τα δύο αυτά διανύσματα είναι:

$$\dot{\mathbf{f}} = \mathbf{J}(\mathbf{x}) \dot{\mathbf{x}} \quad (2)$$

$m \times 1 \quad m \times n \quad n \times 1$

όπου $\dot{\mathbf{f}}$ και $\dot{\mathbf{x}}$ είναι τα διανύσματα μεταβολών (πρώτες παράγωγοι) των διανυσμάτων \mathbf{f} και \mathbf{x} αντίστοιχα και $\mathbf{J}(\mathbf{x})$ είναι ο Ιακωβιανός πίνακας ο οποίος είναι συνάρτηση της κατάστασης του ελεγχόμενου συστήματος, δηλαδή του διανύσματος θέσης-ταχύτητας του άκρου του ρομπότ.

Από τα παραπάνω γίνεται προφανές ότι το πρόβλημα σχεδιασμού του κατάλληλου νόμου ελέγχου για την εκάστοτε εφαρμογή συνίσταται στην εξεύρεση του κατάλληλου Ιακωβιανού πίνακα. Μία πρώτη προσέγγιση θα ήταν η προσπάθεια αναλυτικού υπολογισμού του πίνακα \mathbf{J} , ο οποίος θα ίσχυε για κάθε δυνατή κατάσταση του

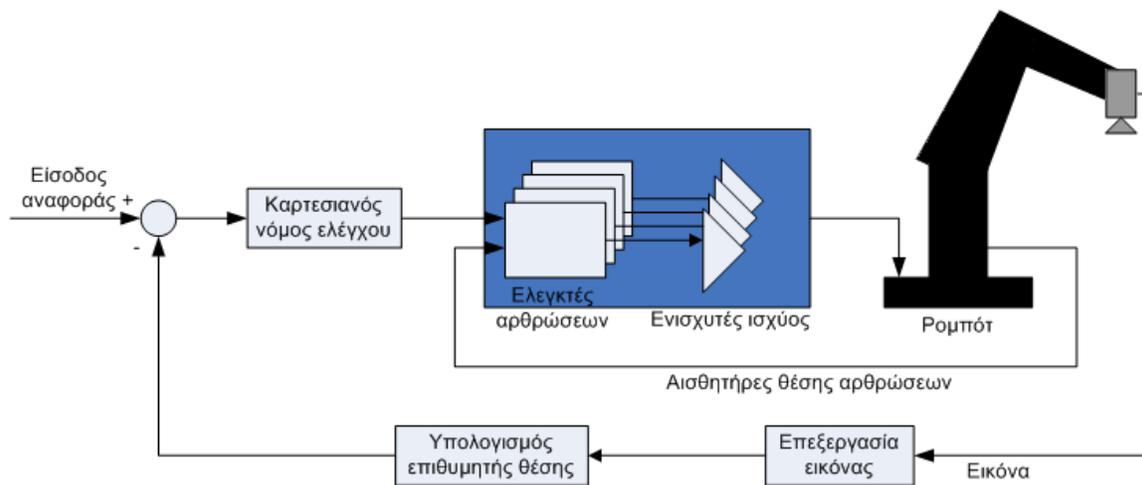
ελεγχόμενου συστήματος, δηλαδή για κάθε δυνατή τιμή του διανύσματος θέσης-ταχύτητας \mathbf{x} του άκρου του ρομπότ. Δυστυχώς στη γενική περίπτωση ο αναλυτικός υπολογισμός είναι πολύ δύσκολος, λόγω της μεγάλης πολυπλοκότητας του συστήματος και της εν γένει μη-γραμμικής φύσης του πίνακα. Στη βιβλιογραφία υπάρχουν κάποιες αναφορές όπου γίνεται αναλυτικός υπολογισμός των εξισώσεων που αποτελούν τον Ιακωβιανό πίνακα [Hutchinson et al., 1996], οι οποίες όμως προκύπτουν από υπεραπλουστευμένα μαθηματικά μοντέλα για κάποια υποσυστήματα της εφαρμογής μας, όπως για παράδειγμα του μοντέλου της κάμερας που συσχετίζει τις 2D συντεταγμένες εικόνας ενός παρατηρούμενου από την κάμερα σημείου με τις αντίστοιχες 3D πραγματικές συντεταγμένες του στο χώρο.

Δεδομένης της δυσκολίας αναλυτικού υπολογισμού των εξισώσεων του Ιακωβιανού πίνακα για κάθε δυνατή κατάσταση του ελεγχόμενου συστήματος, σε πραγματικές εφαρμογές είναι προτιμότερη η χρήση υπολογιστικών αλγορίθμων για τον υπολογισμό μια εκτίμησης του πίνακα \mathbf{J} [Lapreste et al., 2004], η οποία είναι ακριβής σε συγκεκριμένες συνθήκες λειτουργίας του συστήματος (πχ. συγκεκριμένη θέση του βραχίονα και μορφή εικόνων προς επεξεργασία). Η τεχνική αυτή συνίσταται στην γραμμικοποίηση του πίνακα \mathbf{J} γύρω από ένα επιθυμητό σημείο λειτουργίας, η οποία είναι ακριβής μόνο σε μια μικρή περιοχή κοντά στο σημείο γραμμικοποίησης.

Η λύση του κινηματικού και το δυναμικού προβλήματος του ρομπότ [Fu et al., 1987; Τζαφέστας, 2003] συντελείται από τον ρομποτικό ελεγκτή (robot controller) οπότε ο σχεδιαστής του συστήματος ελέγχου δεν χρειάζεται ν' ασχοληθεί με το κομμάτι αυτό. Ειδικά στην περίπτωσή μας όπου το διάνυσμα θέσης του ρομπότ \mathbf{x} περιλαμβάνει συντεταγμένες αρθρώσεων και όχι συντεταγμένες (x, y, z) και προσανατολισμό τελικού στοιχείου δράσης (γωνίες Euler a, b, c) η λύση του ανάστροφου κινηματικού προβλήματος δεν είναι απαραίτητη από τον ρομποτικό ελεγκτή, οπότε απαιτείται μόνο η λύση του ανάστροφου δυναμικού προβλήματος.

Μια εναλλακτική προσέγγιση για την αποφυγή υπολογισμού του Ιακωβιανού πίνακα είναι η εξαγωγή της επιθυμητής θέσης και τροχιάς του άκρου του βραχίονα απ' ευθείας από την ψηφιακή εικόνα (ή τις ψηφιακές εικόνες σε περίπτωση που έχουμε παραπάνω από μία κάμερες). Στην περίπτωση αυτή ο νόμος ελέγχου κλειστού βρόχου γίνεται εξαιρετικά απλός, καθώς το σφάλμα ελέγχου και το σήμα ελέγχου θα έχουν την ίδια

διάσταση, οπότε στις περισσότερες περιπτώσεις απλά χρησιμοποιείται μια απλή ενίσχυση. Η διάσταση αυτή θα είναι ίδια με τη διάσταση του διανύσματος που καθορίζει τη θέση και τον προσανατολισμό του τελικού στοιχείου δράσης του ρομπότ και αποτελείται από 6 μεταβλητές, 3 για τη θέση (x, y, z) και 3 για τον προσανατολισμό (γωνίες Euler a, b, c). Παρατηρούμε ότι το σήμα ελέγχου δεν περιλαμβάνει συντεταγμένες αρθρώσεων όπως στην προηγούμενη προσέγγιση, αλλά ένα 6-διάστατο διάνυσμα που καθορίζει πλήρως την επιθυμητή θέση και προσανατολισμό του τελικού στοιχείου δράσης. Αντίθετα με την προηγούμενη προσέγγιση, εδώ ο ρομποτικός ελεγκτής (robot controller) θα πρέπει να λύσει το ανάστροφο ρομποτικό κινηματικό πρόβλημα για να υπολογίσει τις συντεταγμένες αρθρώσεων (διάστασης n) που θα κατευθύνουν το τελικό στοιχείο δράσης στην επιθυμητή θέση και προσανατολισμό. Η προσέγγιση αυτή φαίνεται στο παρακάτω Σχ. 2.



Σχήμα 1.2 Δομή συστήματος ελέγχου τροχιάς ρομπότ κλειστού βρόχου με καρτεσιανό νόμο ελέγχου

Είναι προφανές ότι η παραπάνω υλοποίηση του συστήματος ελέγχου κλειστού βρόχου μεταθέτει την πολυπλοκότητα του ελεγκτή στον υπολογισμό της επιθυμητής θέσης και προσανατολισμού του τελικού στοιχείου δράσης από την επεξεργασία των ψηφιακών εικόνων. Αυτό σημαίνει ότι θα πρέπει να βρεθεί ένας τρόπος ώστε να περάσουμε από το δισδιάστατο χώρο των ψηφιακών εικόνων στον πραγματικό τρισδιάστατο (ή εξαδιάστατο περιλαμβανομένου και του προσανατολισμού) χώρο εργασίας του ρομποτικού βραχίονα. Το πρώτο βήμα προς αυτή την κατεύθυνση είναι μια διαδικασία επονομαζόμενη *βαθμονόμηση κάμερας (camera calibration)* η οποία μας

βοηθάει να μετασχηματίσουμε τις 2D συντεταγμένες των στοιχείων μια ψηφιακής εικόνας (εικονοστοιχείων) στις αντίστοιχες 3D συντεταγμένες ως προς ένα σταθερό καρτεσιανό σύστημα συντεταγμένων. Με αυτό τον τρόπο μπορούμε να μετρήσουμε πραγματικές Ευκλείδειες αποστάσεις και θέσεις κάποιων παρατηρούμενων από την κάμερα σημείων του χώρου, αναγόμενων πάντα στο προαναφερθέν σταθερό σύστημα συντεταγμένων. Η σχέση που συνδέει τις 2D συντεταγμένες εικόνας (u, v) ενός στοιχείου μια ψηφιακής εικόνας με τις αντίστοιχες πραγματικές 3D συντεταγμένες του στοιχείου αυτού στον τρισδιάστατο χώρο (x, y, z) είναι η παρακάτω:

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \underset{3 \times 3}{\mathbf{C}} \cdot \underset{3 \times 4}{[\mathbf{R} | \mathbf{t}]} \cdot \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (3)$$

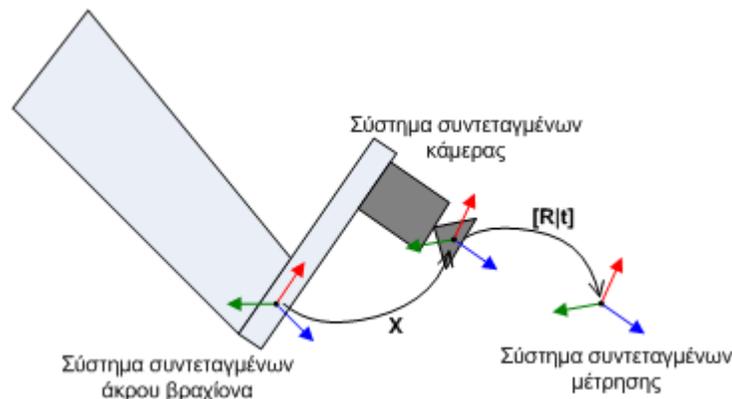
όπου s είναι μια σταθερά αναλογίας, \mathbf{C} είναι ο πίνακας *εσωτερικών παραμέτρων* της κάμερας (internal parameters) και $[\mathbf{R} | \mathbf{t}]$ είναι ο πίνακας *εξωτερικών παραμέτρων* (external parameters). Ο πίνακας εσωτερικών παραμέτρων εξαρτάται από τα κατασκευαστικά στοιχεία της κάμερας (όπως εστιακή απόσταση, κέντρο εικόνας, παραμόρφωση φακού, διαστάσεις εικονοστοιχείου) ενώ ο πίνακας εξωτερικών παραμέτρων είναι ένας ομογενής μετασχηματισμός συντεταγμένων που συνδέει το σύστημα συντεταγμένων στο εστιακό κέντρο της κάμερας με το σταθερό καρτεσιανό σύστημα συντεταγμένων, στο οποίο αναφέρονται οι 3D συντεταγμένες (x, y, z). Υπάρχουν αρκετοί αλγόριθμοι βαθμονόμησης κάμερας [Zhang, 2000; Tsai, 1987; Horn, 2000; Faugeras and Toscani, 1986] και όλοι αποσκοπούν στον υπολογισμό των άγνωστων πινάκων εσωτερικών κι εξωτερικών παραμέτρων. Για την επίτευξη αυτού του στόχου χρησιμοποιούνται κάποια μοντέλα (patterns) τα οποία περιέχουν χαρακτηριστικά (πχ κύκλους ή γωνίες) σε εκ των προτέρων γνωστές προκαθορισμένες θέσεις κι αποστάσεις μεταξύ τους. Η απλούστερη περίπτωση μοντέλου είναι μια εκτύπωση σε εκτυπωτή λείζερ που περιέχει μικρούς κύκλους σε ίσες αποστάσεις μεταξύ τους. Εφόσον οι πραγματικές συντεταγμένες (x, y, z) των χαρακτηριστικών του μοντέλου είναι εκ των προτέρων γνωστές, με βάση ένα οποιοδήποτε προκαθορισμένο σταθερό σύστημα συντεταγμένων, μπορούμε να πάρουμε εικόνες του μοντέλου από μια κάμερα και να υπολογίσουμε τις αντίστοιχες συντεταγμένες εικόνας (u, v) που αντιστοιχούν σε κάθε χαρακτηριστικό του μοντέλου.

Συνδυάζοντας τα ζευγάρια συντεταγμένων εικόνας και πραγματικών συντεταγμένων των στοιχείων μπορούμε με βάση την εξίσωση (3) να υπολογίσουμε τις εσωτερικές κι εξωτερικές συντεταγμένες της κάμερας.

Σ' ένα πραγματικό σύστημα ελέγχου της μορφής του Σχ. 2 μια ή περισσότερες κάμερες παρατηρούν κάποια σημεία ενδιαφέροντος του χώρου εργασίας του ρομπότ και λαμβάνουν ψηφιακές εικόνες που περιλαμβάνουν τα σημεία αυτά. Μέσω των αλγορίθμων επεξεργασίας εικόνας είναι δυνατόν να εντοπιστούν τα σημεία ενδιαφέροντος και να εξαχθούν οι 2D συντεταγμένες εικόνας (u, v) που αντιστοιχούν σε αυτά. Δεδομένου ότι οι συντεταγμένες εικόνας (u, v) ενός στοιχείου της εικόνας (πχ γωνίας, οπής, κύκλου, κτλ) είναι γνωστές, μπορούμε να χρησιμοποιήσουμε την εξίσωση (3) με αντίστροφο τρόπο για να βρούμε τις πραγματικές καρτεσιανές συντεταγμένες που αντιστοιχούν στο στοιχείο αυτό. Δυστυχώς λύνοντας την εξίσωση (3) ως προς τις άγνωστες καρτεσιανές συντεταγμένες (x, y, z) παράγεται ένα σύστημα τριών γραμμικών εξισώσεων με τέσσερις αγνώστους (x, y, z και s). Αυτό σημαίνει ότι το σύστημα έχει μοναδική λύση μόνο εάν μια συντεταγμένη του συγκεκριμένου στοιχείου του χώρου είναι εκ των προτέρων γνωστή (πχ αν το στοιχείο βρίσκεται πάνω στο επίπεδο $z=0$). Για να μπορέσουμε να μετρήσουμε και την τρίτη διάσταση πρέπει να χρησιμοποιήσουμε παραπάνω από μια κάμερες, για τις οποίες έχουμε προηγουμένως εφαρμόσει τον αλγόριθμο βαθμονόμησης κάμερας ξεχωριστά. Με την παραπάνω διαδικασία έχουμε στην διάθεσή μας δύο διαφορετικές εξισώσεις της μορφής (3), ο συνδυασμός των οποίων μπορεί να δημιουργήσει ένα σύστημα έξι γραμμικών εξισώσεων με πέντε αγνώστους (x, y, z, s_1, s_2). Το παραπάνω σύστημα γραμμικών εξισώσεων είναι πλεονάζον και μπορεί να λυθεί εύκολα με τη βοήθεια γραμμικών μεθόδων [Press et al., 1988] (πχ, ψευδοαντιστροφής, SVD, κτλ).

Μια ακόμη διαδικασία που είναι απαραίτητη για το σχεδιασμό του συστήματος ελέγχου του Σχ.2 είναι η συσχέτιση του καρτεσιανού συστήματος συντεταγμένων των 3D μετρήσεων της κάμερας μ' ένα σύστημα συντεταγμένων που είναι γνωστό στον ρομποτικό βραχίονα. Το σύστημα συντεταγμένων των μετρήσεων είναι «πακτωμένο» στην κάμερα και ακολουθεί τις κινήσεις της, και γι' αυτό το λόγο οι 3D μετρήσεις που πραγματοποιεί (x, y, z) είναι τοπικού χαρακτήρα και ανάγονται στο αναφερόμενο κινούμενο σύστημα συντεταγμένων. Εάν θέλουμε να μετασχηματίσουμε τις τοπικές

μετρήσεις σε μια επιθυμητή ή πραγματική θέση του ρομποτικού βραχίονα θα πρέπει να βρούμε το μετασχηματισμό που συνδέει το σύστημα συντεταγμένων των μετρήσεων με κάποιο σύστημα συντεταγμένων που είναι γνωστό στον βραχίονα. Αν η κάμερα είναι σταθερά στερεωμένη στο άκρο του βραχίονα, όπως φαίνεται και στο Σχ. 2, τότε εξυπηρετεί ν' αναγάγουμε τις τοπικές μετρήσεις στο σύστημα συντεταγμένων του άκρου του ρομπότ. Με τον τρόπο αυτό το ρομπότ «γνωρίζει» που βρίσκεται ένα μετρούμενο στοιχείο του περιβάλλοντος χώρου σε σχέση με το δικό του σύστημα συντεταγμένων, οπότε μπορεί να το ακολουθήσει ή να κινηθεί προς αυτό.



Σχήμα 1.3 Συσχετισμός συστημάτων συντεταγμένων ρομποτικού βραχίονα και κάμερας σταθερά στερεωμένης στο άκρο του

Ο υπολογισμός του άγνωστου μετασχηματισμού \mathbf{X} που συνδέει το σύστημα συντεταγμένων της κάμερας και το σύστημα συντεταγμένων του άκρου του βραχίονα ονομάζεται στη βιβλιογραφία ως «Βαθμονόμηση Βραχίονα - Μηχανικής Όρασης» (*Hand-Eye Calibration*) [Horaud and Dornaika, 1995; Andreff et al., 2001; Tsai and Lenz, 1989]. Εφεξής θα ονομάζουμε σύστημα συντεταγμένων κάμερας το καρτεσιανό σύστημα συντεταγμένων που έχει σαν αρχή το εστιακό κέντρο της κάμερας (Σχ. 1.3) και σύστημα συντεταγμένων μέτρησης αυτό που ανάγονται οι μετρήσεις (x, y, z) της κάμερας. Προφανώς τα δύο συστήματα συντεταγμένων συνδέονται με τον ομογενή μετασχηματισμό $[\mathbf{R}|\mathbf{t}]$ των εξωτερικών παραμέτρων της κάμερας, οπότε ο ζητούμενος μετασχηματισμός \mathbf{X} μπορεί ν' αναφέρεται σε οποιοδήποτε από τα δύο συστήματα συντεταγμένων. Υπάρχουν αρκετοί αλγόριθμοι αναφερόμενοι στη βαθμονόμηση βραχίονα – μηχανικής όρασης, οι βασικές αρχές των οποίων θ' αναλυθούν εκτενέστερα σε επόμενα κεφάλαια (Κεφ. 3).

1.2 Αντικειμενικός σκοπός και καινοτομία διατριβής

Το αντικείμενο της διατριβής κινείται στο πλαίσιο που περιγράφεται στην προηγούμενη ενότητα, χρησιμοποιώντας νέες τεχνικές και μεθόδους για την σχεδίαση και υλοποίηση ενός συστήματος αυτομάτου ελέγχου *οπτικής παρακολούθησης τροχιάς (visual servoing)*, όπου το άκρο ενός βιομηχανικού ρομποτικού βραχίονα θα οδηγείται να παρακολουθεί τις μεταβολές θέσης και προσανατολισμού ενός κινούμενου στον χώρο αντικειμένου, με βάση τις μετρήσεις ανάδρασης που πραγματοποιεί μια συστοιχία μέτρησης. Για την επίτευξη του παραπάνω στόχου είναι απαραίτητο να πραγματοποιηθούν συγκεκριμένες διαδικασίες, όπως είναι η βαθμονόμηση της κάμερας ή της συστοιχίας μέτρησης γενικότερα, η συσχέτισή της μ' ένα σύστημα συντεταγμένων που είναι γνωστό στον βραχίονα (βαθμονόμηση βραχίονα – μηχανικής όρασης), η εφαρμογή αλγορίθμων ψηφιακής επεξεργασίας εικόνας για την ανίχνευση συγκεκριμένων παρατηρούμενων χαρακτηριστικών του χώρου και ο συνδυασμός τους για τον υπολογισμό των απαραίτητων μετασχηματισμών συντεταγμένων που απαιτούνται για τον έλεγχο θέσης του άκρου του βραχίονα.

Στη διατριβή αυτή παρουσιάζεται μια νέα τεχνική για τον συσχετισμό των συστημάτων συντεταγμένων της συστοιχίας μέτρησης και του άκρου του βραχίονα (βαθμονόμηση βραχίονα – μηχανικής όρασης), πραγματοποιώντας τρισδιάστατες (3D) μετρήσεις ανακατασκευής συγκεκριμένων χαρακτηριστικών σημείων για την εξαγωγή των απαραίτητων μετασχηματισμών μετατόπισης της κάμερας. Όλες οι υπάρχουσες μέθοδοι βαθμονόμησης κάμερας χρησιμοποιούν τους αλγόριθμους βαθμονόμησης για την εξαγωγή των προαναφερθέντων μετασχηματισμών, με αποτέλεσμα ο υπολογισμός του ζητούμενου μετασχηματισμού Hand-Eye (βραχίονα και κάμερας) να εξαρτάται από την απόλυτη ακρίβεια των πινάκων εξωτερικών παραμέτρων που προκύπτουν από την βαθμονόμηση της κάμερας. Ο υπολογισμός των πινάκων εσωτερικών-εξωτερικών παραμέτρων μιας κάμερας βασίζεται στην διάσπαση ενός πίνακα *προβολής (projection matrix ή homography)* που υπολογίζεται αρχικά από τον εκάστοτε αλγόριθμο βαθμονόμησης. Παρ' όλο που ο πίνακας προβολής μπορεί να υπολογιστεί με πολύ μεγάλη ακρίβεια, η διαδικασία διάσπασής του εισάγει ανακρίβειες στους εξαχθέντες

πίνακες εσωτερικών - εξωτερικών παραμέτρων, το μέγεθος των οποίων εξαρτάται από τον χρησιμοποιούμενο αλγόριθμο βαθμονόμησης κάμερας [González et al., 2005; Sun and Cooperstock, 2006]. Στο Κεφ. 3.3 προτείνεται μια νέα παραλλαγή της μεθόδου βαθμονόμησης βραχίονα – μηχανικής όρασης, όπου αντί για μια μονή κάμερα χρησιμοποιείται μια στερεοσκοπική κεφαλή (stereo head), η οποία είναι σταθερά στερεωμένη στο άκρου του βραχίονα. Αρχικά οι δύο κάμερες βαθμονομούνται, όχι για την χρησιμοποίηση αυτούσιων των εξωτερικών τους παραμέτρων όπως στις κλασσικές μεθόδους αλλά για την απόκτηση της δυνατότητας πραγματοποίησης μετρήσεων 3D ανακατασκευής συγκεκριμένων παρατηρούμενων σημείων. Κατ' αυτόν τον τρόπο, η απόλυτη ακρίβεια των πινάκων εξωτερικών παραμέτρων του χρησιμοποιούμενου αλγορίθμου βαθμονόμησης δεν επηρεάζει την ακρίβεια υπολογισμού του ζητούμενου μετασχηματισμού βραχίονα - κάμερας, καθώς η ακρίβεια των μετρήσεων 3D ανακατασκευής επηρεάζεται μόνο από την ακρίβεια του πίνακα προβολής της κάμερας (γινόμενο εσωτερικών-εξωτερικών παραμέτρων) η οποία είναι κατά κανόνα πολύ μεγάλη. Τα πειραματικά αποτελέσματα που παρουσιάζονται στο Κεφ. 3.4 δικαιολογούν την ανωτερότητα της προτεινόμενης μεθόδου έναντι των υπαρχόντων.

Εκτός από τη μέθοδο βαθμονόμησης βραχίονα – μηχανικής όρασης, υπάρχει μια λιγότερο γνωστή τεχνική ταυτόχρονου υπολογισμού των άγνωστων μετασχηματισμών μεταξύ του άκρου βραχίονα και της χρησιμοποιούμενης κάμερας (Hand-Eye) και μεταξύ της βάσης του ρομπότ και του χρησιμοποιούμενου συστήματος συντεταγμένων μετρήσεων αναφοράς (Robot-World). Η μέθοδος αυτή είναι γνωστή με την ονομασία «*Ταυτόχρονη Βαθμονόμηση Βραχίονα – Μηχανικής Όρασης και Χώρου Εργασίας*» (*Hand-Eye/Robot-World calibration*) [Zhuang et al., 1994; Dornaika and Horaud, 1998], και χρησιμοποιεί παρόμοια διαδικασία με την κλασσική μέθοδο βαθμονόμησης βραχίονα – μηχανικής όρασης. Αυτό σημαίνει ότι όπως και στην προηγούμενη περίπτωση, χρησιμοποιούνται οι πίνακες εξωτερικών παραμέτρων από τον αλγόριθμο βαθμονόμησης της κάμερας για τον υπολογισμό των σχετικών μετατοπίσεων της κάμερας κατά την πραγματοποίηση της μεθόδου, γεγονός που εισάγει ανακρίβειες στον τελικό υπολογισμό των ζητούμενων μετασχηματισμών Hand-Eye και Robot-World. Για την βελτίωση της ακρίβειας των αποτελεσμάτων χρησιμοποιείται εκ νέου η καινούρια τεχνική που περιγράφεται στην προηγούμενη παράγραφο, χρησιμοποιώντας και πάλι μια

στεροσκοπική κεφαλή για την πραγματοποίηση μετρήσεων 3D ανακατασκευής, όπως περιγράφεται αναλυτικά στο Κεφ. 4.2. Η ανωτερότητα της νέας προτεινόμενης μεθόδου έναντι των κλασικών αποδεικνύεται από τα πειραματικά αποτελέσματα που παρουσιάζονται στο Κεφ. 4.3.

Στο τελευταίο κεφάλαιο της διατριβής παρουσιάζεται ο σχεδιασμός και η υλοποίηση ενός συστήματος παρακολούθησης τροχιάς (visual servoing), όπου το άκρο ενός βιομηχανικού ρομποτικού βραχίονα οδηγείται κατά τέτοιο τρόπο ώστε να παρακολουθεί τις μεταβολές θέσης ενός κινούμενου στον χώρο αντικειμένου. Ως αισθητήρας μέτρησης της θέσης του αντικειμένου χρησιμοποιείται μια συστοιχία κάμερας και συσκευής λέιζερ, η οποία είναι σταθερά στερεωμένη στο άκρο του βραχίονα και μπορεί ν' ανακατασκευάσει τις 3D συντεταγμένες ενός μεγάλου πλήθους σημείων στην επιφάνεια του παρατηρούμενου αντικειμένου. Η συστοιχία μέτρησης είναι σχεδιασμένη να ανακατασκευάζει ολόκληρα τμήματα της επιφάνειας του παρατηρούμενου αντικειμένου και όχι συγκεκριμένα χαρακτηριστικά σημεία πάνω στην επιφάνεια, όπως κέντρα οπών, ακμές ή γωνίες. Η εφαρμογή της τεχνικής αυτή είναι πολύ χρήσιμη σε περιπτώσεις όπου τα παρατηρούμενα αντικείμενα δεν έχουν χαρακτηριστικά σημεία υψής στην επιφάνειά τους, αλλά αντίθετα η τοπολογία τους είναι σχετικά απλή και ομαλή. Για τον υπολογισμό των εκάστοτε μετασχηματισμών συντεταγμένων, που αντιστοιχούν στις διορθώσεις θέσης και προσανατολισμού του άκρου του βραχίονα, χρησιμοποιείται μια νέα μέθοδος η οποία χρησιμοποιεί τα CAD δεδομένα της επιφάνειας του παρατηρούμενου αντικειμένου σε συνδυασμό με το πλήθος των μετρούμενων σημείων 3D ανακατασκευής για τον υπολογισμό της μεταβολής θέσης του αντικειμένου στον χώρο. Τα CAD δεδομένα της επιφάνειας ενός αντικειμένου παρέχονται συνήθως από τον κατασκευαστή του κι αποτελούνται από ένα πυκνό πλέγμα 3D τριγώνων τα οποία περιγράφουν την τοπολογία της επιφάνειας του αντικειμένου. Ο σχεδιασμός και η υλοποίηση ενός συστήματος παρακολούθησης τροχιάς χρησιμοποιώντας την παραπάνω τεχνική είναι μια νέα προσέγγιση στην κατηγορία των συστημάτων αυτών, καθώς δεν υπάρχει κάτι ανάλογο στη διεθνή βιβλιογραφία. Όλα τα επιμέρους βήματα που απαιτούνται για την υλοποίηση του δεδομένου συστήματος περιγράφονται στο Κεφ. 5.

1.3 Οργάνωση και δομή της διατριβής

Η διατριβή είναι χωρισμένη σε θεματικές ενότητες, οι οποίες καλύπτονται σε ξεχωριστά κεφάλαια αυτής. Το τελευταίο κεφάλαιο της διατριβής είναι και το πιο σημαντικό, καθώς εκεί περιγράφεται ο συνδυασμός των τεχνικών και αποτελεσμάτων όλων των προηγούμενων κεφαλαίων για το σχεδιασμό και υλοποίηση ενός νέου συστήματος παρακολούθησης τροχιάς ενός κινούμενου αντικειμένου από το άκρο ενός βιομηχανικού βραχίονα, με βάση τις μετρήσεις από μια συστοιχία μέτρησης αποτελούμενη από μια ψηφιακή κάμερα και μια *συσσκευή πολλαπλών γραμμών λέιζερ (triangulation-based laser sensor)*.

Στο Κεφ. 2 αναφέρονται οι βασικές αρχές για το χρησιμοποιούμενο μοντέλο προβολής ενός παρατηρούμενου 3D σημείου στον πραγματικό τρισδιάστατο χώρο πάνω στην ψηφιακή εικόνα μιας τυπικής κάμερας (camera pin-hole model). Στη συνέχεια περιγράφονται αναλυτικά δύο από τους διασημότερους αλγόριθμους βαθμονόμησης κάμερας, ήτοι ο αλγόριθμος του Zhang [Zhang, 2000] και ο αλγόριθμος του Tsai [Tsai, 1987]. Τέλος γίνεται αναφορά σε κάποιες βασικές τεχνικές ψηφιακής επεξεργασίας εικόνας επίσης και στην μέθοδο 3D ανακατασκευής των συντεταγμένων ενός παρατηρούμενου σημείου, χρησιμοποιώντας μια στερεοσκοπική κεφαλή αποτελούμενη από δύο κάμερες.

Στο Κεφ. 3 περιγράφεται ο συνηθέστερα χρησιμοποιούμενος αλγόριθμος για τον συσχετισμό των συστημάτων συντεταγμένων του άκρου του βραχίονα και της χρησιμοποιούμενης κάμερας (βαθμονόμηση βραχίονα – μηχανικής όρασης). Στη συνέχεια προτείνεται μια νέα μέθοδος για τη βελτίωση των αποτελεσμάτων του παραπάνω αλγόριθμου, χρησιμοποιώντας μετρήσεις τρισδιάστατης (3D) ανακατασκευής από μια στερεοσκοπική κεφαλή αντί της μεθόδου βαθμονόμησης μιας μονής κάμερας. Στο τέλος του κεφαλαίου παρουσιάζονται κάποια πειραματικά αποτελέσματα, τα οποία αποδεικνύουν την ανωτερότητα της προτεινόμενης μεθόδου στην ακρίβεια των αποτελεσμάτων.

Στο Κεφ. 4 παρουσιάζονται όλες οι υπάρχουσες μέθοδοι επίλυσης του προβλήματος της ταυτόχρονης βαθμονόμησης βραχίονα – μηχανικής όρασης και χώρου εργασίας ρομπότ, ήτοι η γραμμική μέθοδος (linear solution), η μέθοδος αναλυτικής μορφής (closed-form) και η μέθοδος μη-γραμμικής ελαχιστοποίησης (non-linear minimization).

Στη συνέχεια εφαρμόζεται η τεχνική του προηγούμενου κεφαλαίου για τη βελτίωση της ακρίβειας των αποτελεσμάτων, δηλαδή και πάλι χρησιμοποιούνται μετρήσεις τρισδιάστατης (3D) ανακατασκευής από μια στερεοσκοπική κεφαλή για την εξαγωγή των σχετικών μετασχηματισμών της κάμερας, αντί για τις εξωτερικές παραμέτρους από την βαθμονόμηση μιας μονής κάμερα. Τα πειραματικά αποτελέσματα επιβεβαιώνουν και πάλι τη θετική επίδραση της προτεινόμενης μεθόδου στην ακρίβεια των αποτελεσμάτων.

Τέλος το Κεφ. 5 αποτελεί το πιο σημαντικό μέρος της διατριβής, όπου περιγράφεται ο συνδυασμός των τεχνικών και αποτελεσμάτων όλων των προηγούμενων κεφαλαίων για το σχεδιασμό και την υλοποίηση ενός συστήματος αυτομάτου ελέγχου τροχιάς του άκρου ενός βιομηχανικού βραχίονα ώστε να παρακολουθεί τις μεταβολές θέσης και προσανατολισμού ενός κινούμενου-στο-χώρο αντικειμένου. Η καλή λειτουργία και η ακρίβεια του προτεινόμενου συστήματος αποδεικνύονται από τα πειραματικά αποτελέσματα που παρουσιάζονται στο τέλος του κεφαλαίου.

2 ΜΗΧΑΝΙΚΗ ΟΡΑΣΗ

2.1 Γενικά περί μηχανικής όρασης

Στόχος των συστημάτων μηχανικής όρασης είναι η δημιουργία ενός μοντέλου του πραγματικού κόσμου μέσω της λήψης κι επεξεργασίας ψηφιακών εικόνων από μία ή περισσότερες ψηφιακές κάμερες. Οι ψηφιακές εικόνες (images) είναι στη πραγματικότητα προβολές μιας σκηνής του πραγματικού κόσμου πάνω σε μια επιφάνεια δυο διαστάσεων, η οποία αποτελεί το φωτοευαίσθητο ολοκληρωμένο κύκλωμα (CMOS) της ψηφιακής κάμερας. Βασικό πρόβλημα αυτής της διαδικασίας είναι το γεγονός ότι ενώ ο πραγματικός κόσμος έχει τρεις διαστάσεις, οι αποκτώμενες εικόνες είναι δισδιάστατες, με αποτέλεσμα πολλές από τις πληροφορίες της τρισδιάστατης σκηνής να χάνονται (όπως για παράδειγμα η αίσθηση του βάθους). Υπάρχουν εφαρμογές όπου οι πληροφορίες που αποκτώνται από μια δισδιάστατη εικόνα είναι ικανοποιητικές για τη συγκεκριμένη εφαρμογή. Στην αντίθετη περίπτωση χρησιμοποιούνται συστήματα στερεοσκοπικής όρασης όπου περισσότερες από μία ψηφιακές κάμερες συλλέγουν πληροφορίες μιας σκηνής του πραγματικού κόσμου για την ανάκτηση της τρισδιάστατης δομής του

Οι εφαρμογές των συστημάτων μηχανικής όρασης είναι απεριόριστες και συναντώνται σε πολλά διαφορετικά πεδία. Στη βιομηχανία είναι πολλές φορές απαραίτητος ο αυτόματος έλεγχος της σωστής διεκπεραίωσης μιας συγκεκριμένης διαδικασίας. Έτσι οι πληροφορίες που αποκτώνται από τις ψηφιακές εικόνες που λαμβάνει κι επεξεργάζεται το σύστημα μηχανικής όρασης (πχ ο έλεγχος αν ένα προϊόν έχει την επιθυμητή δομή και ιδιότητες) χρησιμοποιούνται για την λήψη αποφάσεων σχετικά με την εξέλιξη της αυτοματοποιημένης διαδικασίας (πχ αν το προϊόν πρέπει να απορριφθεί ή όχι). Στην ιατρική χρησιμοποιούνται συστήματα μηχανικής όρασης με υπέρηχους ή ακτίνες X για την απόκτηση εικόνων του εσωτερικού του ανθρώπινου σώματος. Περαιτέρω επεξεργασία των εικόνων επιτρέπει τη διάγνωση ασθενειών ή την απόκτηση χρήσιμων πληροφοριών σχετικά με τη δομή του ανθρώπινου οργανισμού.

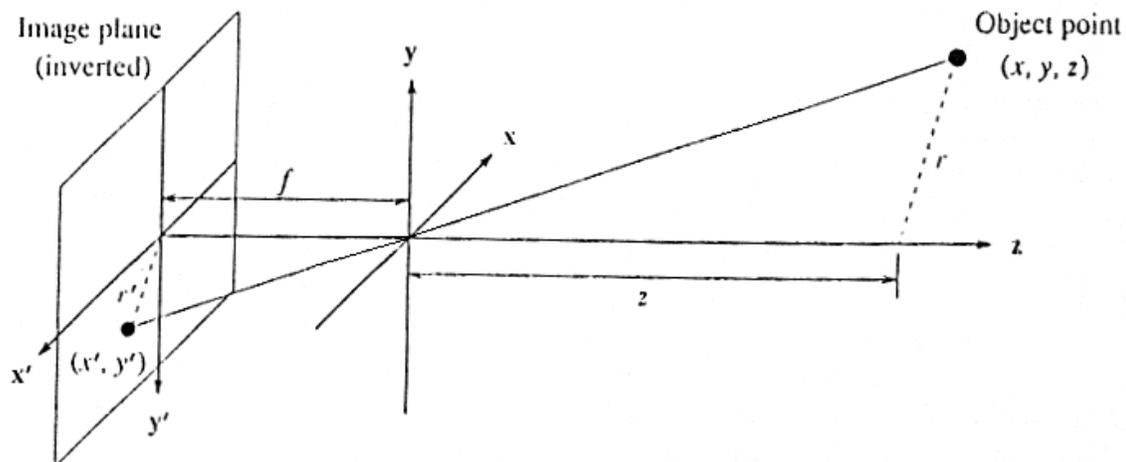
Ίσως οι κυριότερες πρακτικές εφαρμογές της μηχανικής όρασης σχετίζονται με τους κλάδους της ρομποτικής και του αυτομάτου ελέγχου. Ένας ή περισσότεροι αισθητήρες όρασης (ψηφιακές κάμερες) μπορούν να χρησιμοποιηθούν από ένα ρομποτικό σύστημα ή ένα σύστημα αυτομάτου ελέγχου για να αποκτήσουν πληροφορία που σχετίζεται με τη δομή ή τη θέση κάποιων παρατηρούμενων αντικειμένων μέσα στον περιβάλλοντα χώρο. Η πληροφορία αυτή μπορεί να χρησιμοποιηθεί σαν ανάδραση από το σύστημα, με βάση την οποία θα αποφασίσει για την μελλοντική πορεία της διαδικασίας που είναι προγραμματισμένο να εκτελέσει. Τέτοιου είδους συστήματα χρησιμοποιούνται ευρέως στην βιομηχανία, όπου η αυτοματοποίηση των διεργασιών και η μεγιστοποίηση της παραγωγής είναι πάντοτε μέγιστης σημασίας.

Οι παραπάνω εφαρμογές είναι ενδεικτικές και σε καμία περίπτωση δεν περιλαμβάνουν όλο το φάσμα εφαρμογών των συστημάτων μηχανικής όρασης, τα οποία με την πάροδο του χρόνου συναντώνται όλο και περισσότερο σε εμπορικές εφαρμογές.

2.2 Γεωμετρία εικόνας

Στα συστήματα μηχανικής όρασης σχεδόν πάντα χρησιμοποιούνται ψηφιακές κάμερες, οι οποίες έχουν την ιδιότητα να μετατρέπουν ένα πλάνο (σκηνή) του τρισδιάστατου πραγματικού χώρου σε μια ψηφιακή εικόνα δύο διαστάσεων. Μία τυπική κάμερα αποτελείται από μία συστοιχία φακών που συλλαμβάνουν το ορατό φάσμα ακτινοβολίας του φωτός και το προβάλλουν πάνω σε μία φωτοευαίσθητη επιφάνεια. Έτσι, αν υπάρχει επαρκής φωτισμός, η εικόνα ενός παρατηρούμενου αντικειμένου στο χώρο προβάλλεται μέσω των φακών πάνω στη φωτοευαίσθητη επιφάνεια σχηματίζοντας ένα πλάνο εικόνας (*Image Plane*). Στο παρακάτω σχήμα 2.1 φαίνεται μια απλοποιημένη περιγραφή αυτής της διαδικασίας για ένα μόνο σημείο αναφοράς $\mathbf{P}(x,y,z)$ στον τρισδιάστατο χώρο.

Όπως συμβαίνει σε όλους τους τύπους αισθητήρων όρασης, έτσι και στις ψηφιακές κάμερες η εικόνα του παρατηρούμενου πραγματικού χώρου προβάλλεται ανεστραμμένη πάνω στην φωτοευαίσθητη επιφάνεια της κάμερας.



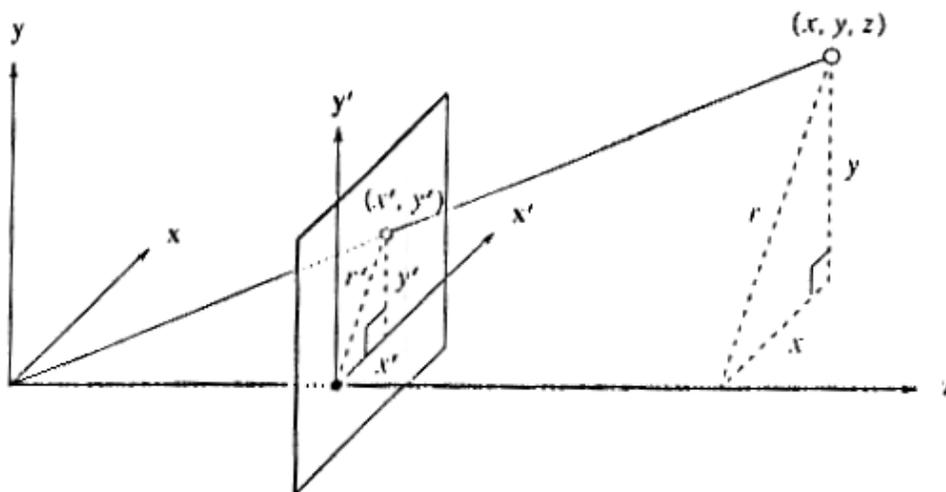
Σχήμα 2.1 Η προβολή ενός σημείου στο χώρο (x, y, z) πάνω στο πλάνο εικόνας της κάμερας

Το σημείο όπου γίνεται η αντιστροφή της εικόνας ονομάζεται “κέντρο προβολής \mathbf{O} ” (*Center of projection*), όπως φαίνεται και στο σχήμα 2.1. Η απόσταση του επιπέδου του πλάνου εικόνας από το κέντρο προβολής \mathbf{O} ονομάζεται εστιακή απόσταση f και η ευθεία που ενώνει το σημείο αναφοράς \mathbf{P} με το κέντρο προβολής ονομάζεται *οπτική ευθεία* (*Line of Sight*). Ορίζουμε ένα τρισσορθογώνιο καρτεσιανό σύστημα συντεταγμένων, που στο εξής θα ονομάζουμε σύστημα συντεταγμένων κάμερας (*Camera coordinate frame*), του οποίου το κέντρο συμπίπτει με το κέντρο προβολής \mathbf{O} της εικόνας, και επίπεδο που ορίζεται από τους άξονες x - y είναι παράλληλο στο επίπεδο του πλάνου εικόνας, όπως φαίνεται στο σχήμα 2.1. Οι συντεταγμένες του σημείου αναφοράς $\mathbf{P}(x,y,z)$ στο χώρο, καθώς και οποιουδήποτε άλλου σημείου του τρισδιάστατου χώρου, στο εξής θα ορίζονται ως προς το σύστημα συντεταγμένων $Oxyz$ της κάμερας.

Για να αποφύγουμε την αντιστροφή της εικόνας κατά την προβολή της, θεωρούμε ότι το επίπεδο του πλάνου εικόνας βρίσκεται μπροστά απ’ το κέντρο προβολής \mathbf{O} , απέχοντας την ίδια απόσταση f απ’ αυτό (σχήμα 2.2). Ορίζουμε ένα δεύτερο σύστημα συντεταγμένων $\mathbf{O}'x'y'$ δύο διαστάσεων πάνω στο επίπεδο του πλάνου εικόνας, το οποίο στο εξής θα ονομάζεται *σύστημα συντεταγμένων επιπέδου εικόνας* (*Image plane coordinates*). Η προβολή του σημείου $\mathbf{P}(x,y,z)$ πάνω στο επίπεδο του πλάνου εικόνας θα έχει συντεταγμένες (x', y') ως προς το σύστημα συντεταγμένων επιπέδου εικόνας. Το κέντρο του προαναφερθέντος συστήματος συντεταγμένων \mathbf{O}' συμπίπτει με την τομή του

άξονα z του συστήματος συντεταγμένων κάμερας με το επίπεδο του πλάνου εικόνας, όπως φαίνεται και στο σχήμα 2.2.

Στη συνέχεια θα γίνει μια προσπάθεια να συσχετισθούν οι συντεταγμένες του σημείου στο χώρο $\mathbf{P}(x,y,z)$ με τις αντίστοιχες συντεταγμένες της προβολής του πάνω στο πλάνο εικόνας (x', y') . Με βάση το σχήμα 2.2 το σημείο (x', y') βρίσκεται στην τομή της οπτικής ευθείας με το επίπεδο του πλάνου εικόνας και οι συντεταγμένες του υπολογίζονται εύκολα με τη χρήση απλών γεωμετρικών σχέσεων.



Σχήμα 2.2 Για την αποφυγή αντιστροφής της εικόνας, στο μοντέλο προβολής της κάμερας θεωρούμε ότι το πλάνο εικόνας είναι μπροστά από το κέντρο προβολής

Σύμφωνα πάντα με το σχήμα 2.2, η απόσταση r του σημείου $\mathbf{P}(x,y,z)$ στο χώρο από τον άξονα z του συστήματος συντεταγμένων κάμερας ισούται με $r = \sqrt{x^2 + y^2}$ ενώ η απόσταση του προβαλλόμενου σημείου (x', y') από το κέντρο του συστήματος συντεταγμένων επιπέδου εικόνας \mathbf{O}' ισούται με $r' = \sqrt{x'^2 + y'^2}$. Ο άξονας z , η οπτική ευθεία από το σημείο $\mathbf{P}(x,y,z)$ και το ευθύγραμμο τμήμα απόστασης r σχηματίζουν ένα τρίγωνο. Αντίστοιχα ο άξονας z , η οπτική ευθεία από το σημείο (x', y') και το ευθύγραμμο τμήμα μήκους r' σχηματίζουν ένα δεύτερο τρίγωνο. Τα δύο τρίγωνα είναι όμοια, με αποτέλεσμα ο λόγος των μηκών των αντίστοιχων πλευρών τους είναι ο ίδιος. Επομένως :

$$\frac{f}{z} = \frac{r'}{r} \quad (2.1)$$

Το τρίγωνο που σχηματίζεται από τις συντεταγμένες x και y και την απόσταση r , και το τρίγωνο που σχηματίζεται από τις συντεταγμένες x' , y' και την αντίστοιχη απόσταση r' είναι επίσης όμοια, επομένως ισχύει :

$$\frac{x'}{x} = \frac{y'}{y} = \frac{r'}{r} \stackrel{(2.1)}{\Rightarrow} \frac{x'}{x} = \frac{y'}{y} = \frac{f}{z} \quad (2.2)$$

Τελικά η προβολή του σημείου $\mathbf{P}(x,y,z)$ πάνω στο πλάνο εικόνας δίνεται απ' τις σχέσεις :

$$x' = \frac{f}{z} x \quad (2.3)$$

και
$$y' = \frac{f}{z} y \quad (2.4)$$

Όπως αναφέρθηκε και προηγουμένως, η ψηφιακή εικόνα του παρατηρούμενου πλάνου δημιουργείται στο επίπεδο του πλάνου εικόνας, όπου στην πραγματικότητα βρίσκεται η φωτοευαίσθητη επιφάνεια της ψηφιακής κάμερας. Αυτή αποτελείται από $M \times N$ φωτοευαίσθητα διακριτά ηλεκτρονικά στοιχεία, τα οποία έχουν την ιδιότητα να μετατρέπουν την ένταση της προσπίπτουσας ακτινοβολίας φωτός σε διακριτές κβαντισμένες τιμές (συνήθως από 0 έως 255). Καθόλου ακτινοβολία (μαύρο χρώμα) θα αντιστοιχεί στην ελάχιστη κβαντισμένη τιμή (0) ενώ μεγάλη ένταση ακτινοβολίας (άσπρο χρώμα) θα αντιστοιχεί στην μέγιστη κβαντισμένη τιμή (255). Οι κβαντισμένες τιμές όλων των $M \times N$ διακριτών στοιχείων της φωτοευαίσθητης επιφάνειας σχηματίζουν ένα διατεταγμένο πίνακα τιμών, ο οποίος στο σύνολό περιέχει την ψηφιακή πληροφορία της εικόνας του περιβάλλοντος που λαμβάνει η κάμερα. Τα $M \times N$ στοιχεία του διατεταγμένου πίνακα ονομάζονται *εικονοστοιχεία* (*pixels*), και το καθένα κρατά πληροφορία για την φωτεινότητα της εικόνας (σκηνής) στο συγκεκριμένο σημείο.

Αναφερόμενοι στο σχήμα 2.2 και τις εξισώσεις (2.3) και (2.4), οι συντεταγμένες (x',y') της προβολής του σημείου \mathbf{P} πάνω στην επιφάνεια του πλάνου εικόνας (φωτοευαίσθητη επιφάνεια) είναι σε πραγματικές μονάδες μήκους (πχ. mm) καθώς προήρθαν από γεωμετρικές σχέσεις όμοιων τριγώνων. Έστω ότι το ορθογώνιο παραλληλόγραμμο του σχήματος 2.2 απεικονίζει την φωτοευαίσθητη επιφάνεια της κάμερας, αποτελούμενη από $M \times N$ εικονοστοιχεία, και ότι το σημείο τομής του άξονα z του συστήματος συντεταγμένων κάμερας με την φωτοευαίσθητη επιφάνεια έχει

συντεταγμένες (u_0, v_0) σε συντεταγμένες εικόνας. Οι συντεταγμένες (u, v) της προβολής του σημείου που αντιστοιχούν στις πραγματικές συντεταγμένες (x', y') δίνονται από τις σχέσεις, σύμφωνα με το σχήμα 2.2:

$$u = u_0 + k_x x' \text{ και } v = v_0 + k_y y' \quad (2.5)$$

Όπου k_x και k_y είναι σταθερές αναλογίας που συνδέουν μεταβολές στις πραγματικές συντεταγμένες (x', y') του σημείου προβολής με τις αντίστοιχες μεταβολές του σε συντεταγμένες εικόνας (u, v) , κατά τον άξονα x και τον άξονα y αντίστοιχα. Προφανώς οι τιμές των u και v στρογγυλοποιούνται σε ακέραιους αριθμούς, για ν' αντιστοιχούν σε συντεταγμένες εικόνας, οπότε οι εξισώσεις (2.3) και (2.4) παίρνουν την παρακάτω μορφή:

$$\begin{aligned} u &= u_0 + \frac{k_x f}{z} x \Rightarrow uz = (k_x f)x + u_0 z \\ v &= v_0 + \frac{k_y f}{z} y \Rightarrow vz = (k_y f)y + v_0 z \end{aligned} \quad (2.6)$$

ή σε μορφή πινάκων:

$$z \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} a & 0 & u_0 \\ 0 & b & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad (2.7)$$

όπου έχουμε θέσει $a=k_x f$ και $b=k_y f$.

Δεδομένου ότι το σύστημα συντεταγμένων της κάμερας είναι συνήθως άγνωστο, στην πράξη οι συντεταγμένες ενός οποιουδήποτε παρατηρούμενου σημείου \mathbf{P} δίνονται συναρτήσει ενός εξωτερικού καθορισμένου συστήματος συντεταγμένων $WXYZ$, ήτοι $\mathbf{P}(X,Y,Z)$. Στην περίπτωση αυτή θα πρέπει να μετασχηματίσουμε τις εξωτερικές συντεταγμένες ως προς το σύστημα συντεταγμένων της κάμερας $Oxyz$, πριν πραγματοποιήσουμε την προβολή του σημείου πάνω στο επίπεδο του πλάνου εικόνας. Χρησιμοποιώντας ομογενείς μετασχηματισμούς συντεταγμένων [Fu et. al., 1987; Τζαφέστας, 2003] η εξίσωση (2.7) παίρνει την παρακάτω μορφή:

$$z \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} a & 0 & u_0 \\ 0 & b & v_0 \\ 0 & 0 & 1 \end{bmatrix}}_{\mathbf{A}} \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \text{---} & \text{---} \\ \text{---} & \text{---} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (2.8)$$

όπου \mathbf{R} είναι ο 3×3 πίνακας στροφής και \mathbf{t} είναι το 3×1 διάνυσμα μετατόπισης, τα οποία σχετίζουν το σύστημα συντεταγμένων της κάμερας $Oxyz$ με το εξωτερικό (καθορισμένο) σύστημα συντεταγμένων $WXYZ$.

Ο 3×3 πίνακας \mathbf{A} της εξίσωσης (2.8) αποτελείται από τιμές που σχετίζονται με κατασκευαστικά στοιχεία της κάμερας, όπως είναι το εστιακό βάθος, τις διαστάσεις ενός εικονοστοιχείου στους άξονες της εικόνας και τη θέση του οπτικού κέντρου σε συντεταγμένες εικόνας. Γι' αυτό το λόγο ο πίνακας αυτός ονομάζεται *πίνακας εσωτερικών παραμέτρων της κάμερας (internal parameters)*. Αντίθετα ο 3×4 μετασχηματισμός συντεταγμένων $[\mathbf{R}|\mathbf{t}]$ που σχετίζει το σύστημα συντεταγμένων της κάμερας με το εξωτερικό σύστημα συντεταγμένων, ονομάζεται *πίνακας εξωτερικών παραμέτρων της κάμερας (external parameters)*. Η τιμή αναλογίας z εξακολουθεί να έχει την ίδια σημασία με πριν, δηλαδή αντιστοιχεί στην συνιστώσα του παρατηρούμενου σημείου $\mathbf{P}(X,Y,Z)$ πάνω στον άξονα z του συστήματος συντεταγμένων της κάμερας $Oxyz$.

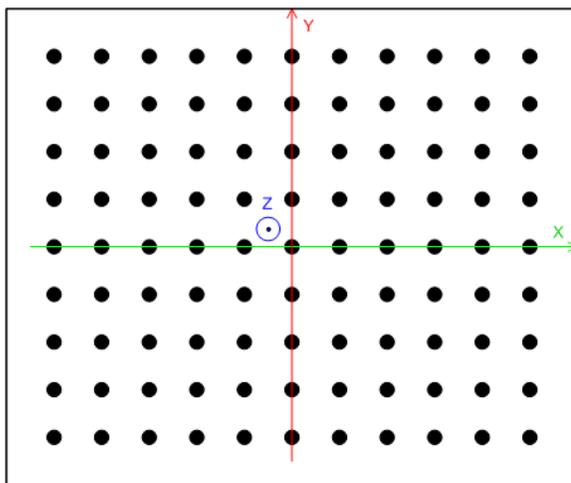
2.3 Βαθμονόμηση κάμερας

Αναφερόμενοι στο παράδειγμα της προηγούμενης ενότητας, έστω ότι έχουμε ένα σημείο $\mathbf{P}(X,Y,Z)$ στο χώρο, με γνωστές συντεταγμένες ως προς ένα αυθαίρετο εξωτερικό σύστημα συντεταγμένων $WXYZ$, το οποίο παρατηρείται από μια ακίνητη ψηφιακή κάμερα. Αν υποθέσουμε ότι ο πίνακας εσωτερικών παραμέτρων \mathbf{A} της κάμερας είναι γνωστός με βάση τα κατασκευαστικά της στοιχεία, και ότι το σύστημα συντεταγμένων της κάμερας $Oxyz$ σχετίζεται με το εξωτερικό σύστημα συντεταγμένων $WXYZ$ με ένα γνωστό μετασχηματισμό συντεταγμένων $[\mathbf{R}|\mathbf{t}]$ (πίνακας εξωτερικών παραμέτρων), τότε η εξίσωση (2.8) μπορεί να λυθεί ως προς τους αγνώστους u , v και z . Αυτό πρακτικά σημαίνει ότι μπορούμε να υπολογίσουμε τις συντεταγμένες εικόνας (u , v) της προβολής του παρατηρούμενου σημείου πάνω στο πλάνο εικόνας της κάμερας, δεδομένου ότι οι πίνακες εσωτερικών κι εξωτερικών παραμέτρων είναι γνωστοί.

Παρ' όλο που οι περισσότεροι κατασκευαστές ψηφιακών καμερών παρέχουν αρκετά κατασκευαστικά στοιχεία των προϊόντων τους, ο υπολογισμός του πίνακα εσωτερικών παραμέτρων της κάμερας μπορεί να γίνει μόνο προσεγγιστικά, καθώς υπεισέρχονται

σφάλματα και ανακρίβειες που δεν μπορούν να προβλεφθούν κατά την κατασκευή της κάμερας. Εκτός αυτού, το εστιακό βάθος f της κάμερας, το οποίο είναι βασικής σημασίας στον υπολογισμό του πίνακα εσωτερικών παραμέτρων, εξαρτάται από την εστίαση (focus) της κάμερας η οποία τις περισσότερες φορές μπορεί να μεταβληθεί από τον χρήστη απλά περιστρέφοντας τον φακό. Ο υπολογισμός του πίνακα εξωτερικών παραμέτρων είναι επίσης μια δύσκολη διαδικασία, καθώς η ακριβής θέση του συστήματος συντεταγμένων $Oxyz$ της κάμερας (η οποία ακολουθεί τις κινήσεις της) μπορεί να υπολογιστεί μόνο προσεγγιστικά.

Στην πράξη, για τον υπολογισμό των πινάκων εσωτερικών-εξωτερικών παραμέτρων μιας ψηφιακής κάμερας χρησιμοποιούνται υπολογιστικοί αλγόριθμοι που ονομάζονται *αλγόριθμοι βαθμονόμησης κάμερας (Camera Calibration)* [Zhang, 2000; Tsai, 1987]. Παρ' όλο που ο καθένας χρησιμοποιεί διαφορετικό τρόπο προσέγγισης κι επίλυσης του προβλήματος, όλοι βασίζονται πάνω σε μια συγκεκριμένη μεθοδολογία, η οποία θα παρουσιαστεί στη συνέχεια.



Σχήμα 2.3 Επίπεδο μοντέλο βαθμονόμησης

Για την βαθμονόμηση της κάμερας είναι απαραίτητη η ύπαρξη ενός αντικειμένου που ονομάζεται *μοντέλο βαθμονόμησης (Calibration Pattern)*. Αυτό ουσιαστικά αποτελείται από ένα πλήθος διακριτών χαρακτηριστικών (πχ. σημείων, ακμών η γωνιών) των οποίων οι συντεταγμένες είναι γνωστές με μεγάλη ακρίβεια με βάση ένα προκαθορισμένο σύστημα συντεταγμένων. Το μοντέλο βαθμονόμησης μπορεί να είναι είτε

επίπεδο, με τα διακριτά χαρακτηριστικά του να βρίσκονται σε συγκεκριμένες θέσεις πάνω στην επιφάνεια του επιπέδου, είτε στέρεο αντικείμενο τριών διαστάσεων (πχ. κύβος ή τραπεζοειδές) του οποίου τα διακριτά χαρακτηριστικά βρίσκονται σε συγκεκριμένες θέσεις πάνω σε δύο ή περισσότερες επιφάνειες του αντικειμένου. Στο σχήμα 2.3 φαίνεται ένα επίπεδο μοντέλο βαθμονόμησης, το οποίο στην πιο απλή του μορφή είναι μια ψηφιακή εκτύπωση από έναν ηλεκτρονικό υπολογιστή. Ορίζουμε

αυθαίρετα το σύστημα συντεταγμένων WXYZ του μοντέλου βαθμονόμησης με αρχή το κεντρικό σημείο του μοντέλου, και του άξονές του όπως φαίνονται στο σχήμα 2.3. Προφανώς όλα τα σημεία \mathbf{P}_i ($i=1, \dots, n$) βρίσκονται πάνω στο ίδιο επίπεδο $z=0$, και απέχουν ίσες αποστάσεις μεταξύ τους με αποτέλεσμα οι συντεταγμένες τους ως προς το σύστημα συντεταγμένων WXYZ να είναι γνωστές. Για παράδειγμα, αν τα σημεία απέχουν μεταξύ τους 20mm οριζόντια και κάθετα, τότε τα σημεία πάνω στον άξονα X θα έχουν συντεταγμένες $(X,Y)=(-100,0), (-80,0) \dots (0,0), \dots (100,0)$. Με την ίδια λογική μπορούν να βρεθούν οι συντεταγμένες όλων των n σημείων \mathbf{P}_i .

Όλοι οι κλασσικοί αλγόριθμοι βαθμονόμησης κάμερας, όπως αυτοί του [Zhang, 2000] και του [Tsai, 1987], οι οποίοι θα παρουσιαστούν σε επόμενα κεφάλαια, απαιτούν την λήψη μίας ή περισσότερων ψηφιακών εικόνων του μοντέλου βαθμονόμησης από την ψηφιακή κάμερα που μας ενδιαφέρει να βαθμονομήσουμε. Στόχος μας είναι από την κάθε ψηφιακή εικόνα στην οποία απεικονίζεται το μοντέλο βαθμονόμησης να εξαγάγουμε τις συντεταγμένες εικόνας (u_i, v_i) της προβολής του κάθε σημείου \mathbf{P}_i του μοντέλου πάνω στην εικόνα. Η διαδικασία αυτή είναι αρκετά πολύπλοκη, και απαιτεί τεχνικές επεξεργασίας εικόνας (image processing), μερικές από τις οποίες αναφέρονται στη βιβλιογραφία [Shortis et al., 1994; Katsuri and Schunck, 1995]. Αποτέλεσμα της διαδικασίας αυτής θα είναι η ανάκτηση n στο πλήθος συντεταγμένων εικόνας (u_i, v_i) $i=1, \dots, n$, με ακρίβεια υπο-εικονοστοιχείου (sub-pixel accuracy), οι οποίες θα αντιστοιχούν μία προς μία στα πραγματικά σημεία \mathbf{P}_i του μοντέλου βαθμονόμησης. Έτσι για κάθε διαφορετική ψηφιακή εικόνα του μοντέλου που λαμβάνουμε με την κάμερα (μετακινώντας κάθε φορά την κάμερα ή το μοντέλο βαθμονόμησης) μπορούμε ν' αποκτήσουμε μια ένα προς ένα αντιστοιχία μεταξύ των πραγματικών συντεταγμένων (X_i, Y_i, Z_i) των σημείων \mathbf{P}_i του παρατηρούμενου μοντέλου και των συντεταγμένων εικόνας των προβολών τους (u_i, v_i) πάνω στην ψηφιακή εικόνα. Η κάθε αντιστοιχία πραγματικών συντεταγμένων $\mathbf{P}_i(X_i, Y_i, Z_i)$ και αντίστοιχων συντεταγμένων εικόνας (u_i, v_i) μπορεί να χρησιμοποιηθεί στην εξίσωση πινάκων (2.8), με τη βοήθεια της οποίας θέλουμε να υπολογίσουμε τους άγνωστους πίνακες εσωτερικών παραμέτρων \mathbf{A} κι εξωτερικών παραμέτρων $[\mathbf{R}|\mathbf{t}]$. Υπάρχουν αρκετοί αλγόριθμοι για τον υπολογισμό των αγνώστων αυτών πινάκων βάσει της παραπάνω λογικής, δύο εκ' των οποίων θα παρουσιαστούν στη συνέχεια.

2.3.1 Αλγόριθμος βαθμονόμησης κάμερας του Z. Zhang

Ο αλγόριθμος βαθμονόμησης κάμερας του Zhang βασίζεται αποκλειστικά στη χρήση επίπεδου μοντέλου βαθμονόμησης, όπως είναι το μοντέλο που απεικονίζεται στο σχήμα 2.3. Στην απλούστερη, και συνηθέστερα χρησιμοποιούμενη περίπτωση, το μοντέλο βαθμονόμησης είναι σχεδιασμένο στον ηλεκτρονικό υπολογιστή με τη βοήθεια μιας εφαρμογής γραφείου (πχ. MS Word ή MS Visio) και είναι τυπωμένο σε έναν εκτυπωτή λέιζερ. Στη συνέχεια το χαρτί τοποθετείται πάνω σε μια επίπεδη επιφάνεια, και συνήθως τοποθετείται πάνω του ένα κομμάτι γυαλί ώστε το μοντέλο (χαρτί) να παραμένει όσο το δυνατόν σε σταθερό επίπεδο. Η ψηφιακή κάμερα, την οποία μας ενδιαφέρει να βαθμονομήσουμε, στερεώνεται σε μια τυχαία θέση πάνω απ' το μοντέλο βαθμονόμησης ώστε η ψηφιακή εικόνα που λαμβάνει από τη θέση αυτή να περιλαμβάνει όλα τα σημεία του μοντέλου, τα οποία θα πρέπει να καταλαμβάνουν όσο το δυνατό μεγαλύτερο εύρος της εικόνας. Οι πραγματικές συντεταγμένες του κάθε σημείου $\mathbf{P}_i(X_i, Y_i, Z_i)$ ως προς το σύστημα συντεταγμένων WXYZ του μοντέλου βαθμονόμησης συνδέονται με τις αντίστοιχες συντεταγμένες εικόνας (u_i, v_i) της προβολής του σημείου πάνω στην ψηφιακή εικόνα με τη γνωστή εξίσωση (2.8), ήτοι:

$$i = 1, 2, \dots, n \quad z_i \begin{bmatrix} u_i \\ v_i \\ 1 \end{bmatrix} = \mathbf{A} [\mathbf{R} | \mathbf{t}] \begin{bmatrix} X_i \\ Y_i \\ Z_i \\ 1 \end{bmatrix} \quad (2.9)$$

$$\text{όπου} \quad \mathbf{A} = \begin{bmatrix} a & c & u_0 \\ 0 & b & v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (2.10)$$

Υπενθυμίζεται ότι η σταθερά αναλογίας z_i αντιστοιχεί στην προβολή του σημείου \mathbf{P}_i πάνω στον άξονα z του συστήματος συντεταγμένων της κάμερας Oxyz. Το καινούριο στοιχείο στον πίνακα εσωτερικών παραμέτρων της εξίσωσης (2.10) είναι η σταθερά c , η οποία περιγράφει την ορθογωνιότητα των αξόνων x' και y' του συστήματος συντεταγμένων του επιπέδου εικόνας (βλ. σχήμα 2.2). Στην περίπτωση που οι δύο άξονες είναι ακριβώς κάθετοι, τότε θα είναι $c=0$ και θα ισχύουν οι εξισώσεις (2.6), (2.7) και

(2.8). Στην αντίθετη περίπτωση, το πρώτο σκέλος της εξίσωσης (2.6) θα παίρνει την παρακάτω μορφή:

$$u = u_0 + \frac{a}{z}x + \frac{c}{z}y \Rightarrow uz = ax + cy + u_0z$$

Αυτό υποδηλώνει ότι η τετμημένη u του συστήματος συντεταγμένων εικόνας εξαρτάται και από τις δύο συντεταγμένες (x, y) της θέσης του σημείου \mathbf{P}_i ως προς το σύστημα συντεταγμένων της κάμερας.

Δεδομένου ότι όλα τα σημεία του μοντέλου βαθμονόμησης βρίσκονται πάνω στο ίδιο επίπεδο $Z=0$, η εξίσωση (2.9) μπορεί να πάρει την παρακάτω μορφή:

$$\begin{aligned} z_i \begin{bmatrix} u_i \\ v_i \\ 1 \end{bmatrix} &= \mathbf{A} \begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{r}_3 & \mathbf{t} \end{bmatrix} \begin{bmatrix} X_i \\ Y_i \\ 0 \\ 1 \end{bmatrix} \\ &= \mathbf{A} \begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{t} \end{bmatrix} \begin{bmatrix} X_i \\ Y_i \\ 1 \end{bmatrix} \end{aligned} \quad (2.11)$$

όπου \mathbf{r}_1 , \mathbf{r}_2 και \mathbf{r}_3 είναι οι στήλες του 3×3 πίνακα στροφής \mathbf{R} . Ορίζουμε τον 3×3 πίνακα \mathbf{H} (projection matrix ή homography) ο οποίος όπως φαίνεται παρακάτω μετασχηματίζει απ' ευθείας τις πραγματικές συντεταγμένες του σημείου στις αντίστοιχες συντεταγμένες εικόνας της προβολής του:

$$\mathbf{H} = [\mathbf{h}_1 \quad \mathbf{h}_2 \quad \mathbf{h}_3] = \lambda \mathbf{A} \begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{t} \end{bmatrix} \quad (2.12)$$

οπότε

$$z_i \begin{bmatrix} u_i \\ v_i \\ 1 \end{bmatrix} = \mathbf{H} \begin{bmatrix} X_i \\ Y_i \\ 1 \end{bmatrix} \quad (2.13)$$

Από την εξίσωση (2.13) φαίνεται ότι ο πίνακας \mathbf{H} πρέπει να οριστεί με βάση μια αυθαίρετη σταθερά αναλογίας λ (scale factor), καθώς κάθε δυνατή αναλογία $\lambda \mathbf{H}$ της (2.13) δίνει τις ίδιες τιμές για τις μεταβλητές (u_i, v_i) . Για μια συγκεκριμένη θέση της ψηφιακής κάμερας, από την οποία έχουμε λάβει μια ψηφιακή εικόνα του μοντέλου βαθμονόμησης, οι αντιστοιχίες μεταξύ των συντεταγμένων εικόνας (u_i, v_i) των n σημείων του μοντέλου και των πραγματικών τους συντεταγμένων (X_i, Y_i) μπορούν να βρεθούν.

Βάσει αυτών των αντιστοιχιών, ο υπολογισμός του πίνακα \mathbf{H} μπορεί να γίνει εφικτός, με τη διαδικασία που περιγράφεται στην πρωτότυπη εργασία [Zhang, 2000].

Οι μεταβλητές της εξίσωσης (2.12) δεν είναι ανεξάρτητες μεταξύ τους, καθώς τα διανύσματα στήλες \mathbf{r}_1 και \mathbf{r}_2 , τα οποία ανήκουν στον αρχικό πίνακα στροφής \mathbf{R} των εξωτερικών παραμέτρων της κάμερας, πρέπει να ικανοποιούν τους περιορισμούς ορθογωνιότητας και μοναδιαίου μέτρου:

$$\mathbf{r}_1 \cdot \mathbf{r}_2 = 0 \Leftrightarrow [\mathbf{r}_1]^T [\mathbf{r}_2] = 0 \quad (2.14)$$

$$\|\mathbf{r}_1\|^2 = \|\mathbf{r}_2\|^2 = 1 \Leftrightarrow \left\| [\mathbf{r}_1]^T [\mathbf{r}_1] \right\| = \left\| [\mathbf{r}_2]^T [\mathbf{r}_2] \right\| = 1 \quad (2.15)$$

Οι παραπάνω εξισώσεις, με τη βοήθεια της εξίσωσης (2.12) μπορεί να δώσουν δύο βασικούς περιορισμούς για τον πίνακα εσωτερικών παραμέτρων \mathbf{A} , δεδομένου ενός πίνακα \mathbf{H} :

$$\mathbf{h}_1^T \mathbf{A}^{-T} \mathbf{A}^{-1} \mathbf{h}_2 = 0 \quad (2.16)$$

$$\mathbf{h}_1^T \mathbf{A}^{-T} \mathbf{A}^{-1} \mathbf{h}_1 = \mathbf{h}_2^T \mathbf{A}^{-T} \mathbf{A}^{-1} \mathbf{h}_2 \quad (2.17)$$

Λαμβάνοντας υπ' όψη την (2.10), μπορούμε να ορίσουμε τον παρακάτω πίνακα:

$$\mathbf{B} = \mathbf{A}^{-T} \mathbf{A}^{-1} = \begin{bmatrix} B_{11} & B_{12} & B_{13} \\ B_{12} & B_{22} & B_{23} \\ B_{13} & B_{23} & B_{33} \end{bmatrix} \stackrel{(2.10)}{=} \begin{bmatrix} 1/a^2 & -c/(a^2b) & (v_0c - u_0b)/(a^2b) \\ -c/(a^2b) & c^2/(a^2b^2) + 1/b^2 & -c(v_0c - u_0b)/(a^2b^2) - v_0/b^2 \\ (v_0c - u_0b)/(a^2b) & -c(v_0c - u_0b)/(a^2b^2) - v_0/b^2 & (v_0c - u_0b)^2/(a^2b^2) + v_0^2/b^2 + 1 \end{bmatrix} \quad (2.18)$$

Ο πίνακας \mathbf{B} είναι συμμετρικός, οπότε ορίζεται πλήρως από ένα 6-διάστατο διάνυσμα:

$$\mathbf{b} = [B_{11}, B_{12}, B_{22}, B_{13}, B_{23}, B_{33}]^T. \quad (2.19)$$

Σύμφωνα με την εξίσωση (2.12) οι στήλες του πίνακα \mathbf{H} ορίζονται ως $\mathbf{h}_i = [h_{i1}, h_{i2}, h_{i3}]^T$, όπου $i=1,2,3$. Τότε θα ισχύει:

$$\mathbf{h}_i^T \mathbf{B} \mathbf{h}_j = \mathbf{v}_{ij}^T \mathbf{b} \quad (2.20)$$

όπου

$$\mathbf{v}_{ij} = [h_{i1}h_{j1}, h_{i1}h_{j2} + h_{i2}h_{j1}, h_{i2}h_{j2}, h_{i3}h_{j1} + h_{i1}h_{j3}, h_{i3}h_{j2} + h_{i2}h_{j3}, h_{i3}h_{j3}]^T.$$

Με βάση την (2.20), οι δύο βασικοί περιορισμοί (2.16) (2.17) του πίνακα εσωτερικών παραμέτρων \mathbf{A} για έναν δεδομένο πίνακα \mathbf{H} , μπορούν να πάρουν την παρακάτω μορφή:

$$\begin{bmatrix} \mathbf{v}_{12}^T \\ (\mathbf{v}_{11} - \mathbf{v}_{22})^T \end{bmatrix} \mathbf{b} = \mathbf{0}. \quad (2.21)$$

Η εξίσωση (2.21) είναι ένα σύστημα δύο γραμμικών εξισώσεων με έξι αγνώστους (διάνυσμα \mathbf{b}), το οποίο προέκυψε από μία μόνο θέση της κάμερας σε σχέση με το μοντέλο βαθμονόμησης, δηλαδή από την επεξεργασία μιας μόνο ψηφιακής εικόνας του μοντέλου. Μετακινώντας την κάμερα σε μια νέα θέση κι επαναλαμβάνοντας τη διαδικασία του περιγράφεται παραπάνω, καταλήγουμε σ' ένα νέα σύστημα γραμμικών εξισώσεων της μορφής της (2.21) με διαφορετικές τιμές για τις μεταβλητές \mathbf{v}_{11} , \mathbf{v}_{12} , \mathbf{v}_{22} αλλά με το ίδιο διάνυσμα αγνώστων \mathbf{b} . Αυτό εξηγείται από το γεγονός ότι στην καινούρια θέση της κάμερας ο πίνακας \mathbf{H} άρα και τα διανύσματα \mathbf{v}_{ij} είναι διαφορετικά, καθώς οι εξωτερικές παράμετροι της κάμερας έχουν αλλάξει (βλ. (2.12)). Αντίθετα ο πίνακας εσωτερικών παραμέτρων \mathbf{A} είναι αμετάβλητος, καθώς εξαρτάται μόνο από τα κατασκευαστικά στοιχεία της κάμερας, άρα και το διάνυσμα \mathbf{b} παραμένει σταθερό (βλ. (2.18), (2.19)).

Στην γενική περίπτωση όπου έχουμε λάβει m ψηφιακές εικόνες του μοντέλου βαθμονόμησης, έχοντας μετακινήσει την κάμερα σε m ($m > 3$) διακριτές θέσεις, κι επαναλαμβάνοντας την διαδικασία που περιγράφεται παραπάνω, μπορούν να προκύψουν m εξισώσεις της μορφής της (2.21) με κοινό άγνωστο το διάνυσμα \mathbf{b} . Στοιβάζοντας τις m εξισώσεις στη μορφή ενός κοινού πίνακα, μπορούμε να δημιουργήσουμε την παρακάτω εξίσωση πινάκων:

$$\mathbf{V}\mathbf{b} = \mathbf{0} \quad (2.22)$$

όπου \mathbf{V} είναι ο $2m \times 6$ πίνακας που προκύπτει από του 2×6 πίνακες της (2.21) για τις m διακριτές θέσεις της κάμερας, στοιβαγμένους τον έναν κάτω απ' τον άλλον. Η βέλτιστη λύση της (2.22), δηλαδή το διάνυσμα \mathbf{b} που ελαχιστοποιεί τη συνάρτηση $\|\mathbf{V}\mathbf{b}\|^2$, είναι το ιδιοδιάνυσμα του πίνακα $\mathbf{V}^T\mathbf{V}$ που αντιστοιχεί στην μικρότερη ιδιοτιμή του [Press et. al., 1988; Golub and Van Loan, 1983].

Από τον υπολογισμό του διανύσματος \mathbf{b} είναι πλέον δυνατή η εύρεση του πίνακα εσωτερικών παραμέτρων \mathbf{A} . Από την (2.22) φαίνεται ότι το \mathbf{b} ορίζεται με βάση μια αυθαίρετη σταθερά αναλογίας λ , καθώς όλες οι δυνατές αναλογίες της μορφής $\lambda\mathbf{b}$ αποτελούν λύσεις της εξίσωσης. Οπότε με βάση την (2.18) έχουμε:

$$\mathbf{B} = \lambda \mathbf{A}^{-T} \mathbf{A} \quad (2.23)$$

Με τη βοήθεια των (2.23) και (2.18) είναι δυνατός ο αναλυτικός προσδιορισμός των εσωτερικών παραμέτρων της κάμερας, καθώς και της σταθεράς αναλογίας λ της εξίσωσης (2.23):

$$\begin{aligned} v_0 &= (B_{12}B_{13} - B_{11}B_{23}) / (B_{11}B_{22} - B_{12}^2) \\ \lambda &= (B_{33} - [B_{13}^2 + v_0(B_{12}B_{13} - B_{11}B_{23})]) / B_{11} \\ a &= \sqrt{\lambda / B_{11}} \\ b &= \sqrt{\lambda B_{11} / (B_{11}B_{22} - B_{12}^2)} \\ c &= -B_{12}a^2b / \lambda \\ u_0 &= cv_0 / b - B_{13}a^2 / \lambda \end{aligned} \quad (2.24)$$

Από την (2.12) μπορεί επίσης να υπολογιστεί ο πίνακας εξωτερικών παραμέτρων για κάθε διακριτή θέση της κάμερας, δηλαδή για κάθε διαφορετικό πίνακα \mathbf{H} :

$$\begin{aligned} \mathbf{r}_1 &= \lambda \mathbf{A}^{-1} \mathbf{h}_1 \\ \mathbf{r}_2 &= \lambda \mathbf{A}^{-1} \mathbf{h}_2 \\ \mathbf{r}_3 &= \mathbf{r}_1 \times \mathbf{r}_2 \\ \mathbf{t} &= \lambda \mathbf{A}^{-1} \mathbf{h}_3 \end{aligned} \quad (2.25)$$

όπου $\lambda = 1 / \|\mathbf{A}^{-1} \mathbf{h}_1\| = 1 / \|\mathbf{A}^{-1} \mathbf{h}_2\|$

Ο πίνακας στροφής \mathbf{R} που υπολογίζεται από τις τρεις πρώτες εξισώσεις (2.18) δεν είναι συνήθως ορθοκανονικός εξαιτίας του θορύβου και της ανακρίβειας που υπεισέρχονται στις μετρήσεις των σημείων του μοντέλου βαθμονόμησης. Αυτό σημαίνει ότι ο πίνακας \mathbf{R} θα πρέπει να βελτιστοποιηθεί, ώστε να ικανοποιούνται οι συνθήκες ορθοκανονικότητας $\mathbf{R}\mathbf{R}^T = \mathbf{R}^T\mathbf{R} = \mathbf{I}_{3 \times 3}$.

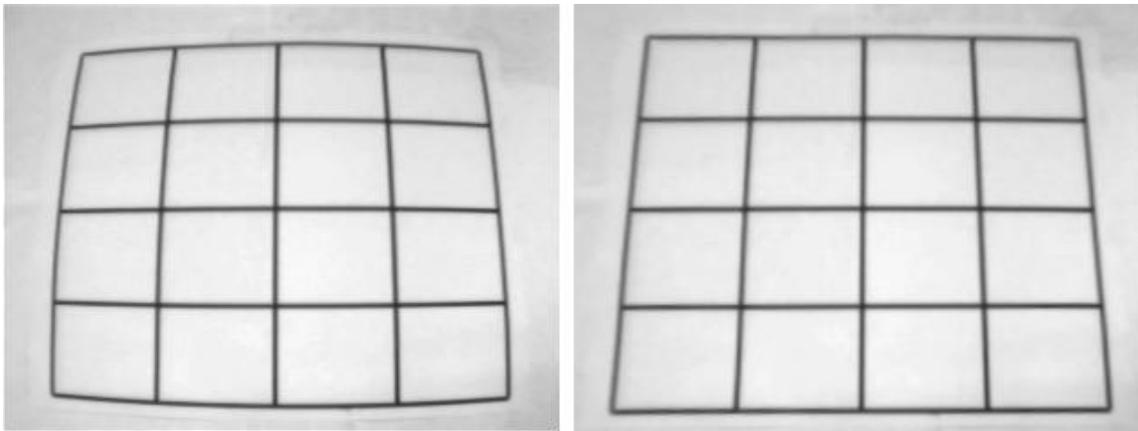
Ο αλγόριθμος υπολογισμού των εσωτερικών-εξωτερικών παραμέτρων που περιγράφεται παραπάνω αποτελεί μια αναλυτική λύση (closed-form) του προβλήματος βαθμονόμησης της κάμερας. Δυστυχώς η ανακρίβεια μέτρησης των συντεταγμένων εικόνας των παρατηρούμενων σημείων του μοντέλου βαθμονόμησης, καθώς και κατασκευαστικοί παράγοντες της κάμερας που δεν έχουν ληφθεί υπ' όψη απ' το μοντέλο προβολής (pin-hole model), μειώνουν την ακρίβεια των υπολογισμών και των εξαχθέντων αποτελεσμάτων. Υπάρχει η δυνατότητα βελτιστοποίησης των

υπολογισθέντων πινάκων εσωτερικών-εξωτερικών παραμέτρων, ορίζοντας μια θετικά ορισμένη μη-γραμμική συνάρτηση κόστους τετραγωνικής μορφής, η οποία θα περιγράφει το σφάλμα υπολογισμού, και στη συνέχεια επιχειρώντας την ελαχιστοποίησή της χρησιμοποιώντας μια μη-γραμμική μέθοδο ελαχιστοποίησης. Η πιο γνωστή και ευρύτερα χρησιμοποιούμενη μέθοδος μη-γραμμικής ελαχιστοποίησης είναι μια παραλλαγή της μεθόδου Levenberg-Marquardt [Fletcher, 1990; Gill et al., 1989], όπως αυτή υλοποιείται από το πακέτο λογισμικού MINPACK [Moré et al., 1980, 1984], όπου ο υπολογισμός μιας εκτίμησης του διανύσματος παραγώγων (gradient vector) υλοποιείται εσωτερικά από τον αλγόριθμο. Η συνάρτηση κόστους θα έχει τη μορφή:

$$f(\mathbf{A}, \mathbf{R}_i, \mathbf{t}_i) = \sum_{i=1}^m \sum_{j=1}^n \left\| m_{ij} - \hat{m}(\mathbf{A}, \mathbf{R}_i, \mathbf{t}_i, \mathbf{M}_j) \right\|^2 \quad (2.26)$$

όπου n είναι ο συνολικός αριθμός των παρατηρούμενων σημείων πάνω στο μοντέλο βαθμονόμησης, m είναι ο αριθμός των διακριτών θέσεων της κάμερας από τις οποίες έχουμε λάβει ψηφιακές εικόνες, m_{ij} είναι οι συντεταγμένες εικόνας του j παρατηρούμενου σημείου του μοντέλου βαθμονόμησης ($j \in [1, m]$) από την i ψηφιακή εικόνα ($i \in [1, n]$), και $\hat{m}(\mathbf{A}, \mathbf{R}_i, \mathbf{t}_i, \mathbf{M}_j)$ είναι η εκτίμηση του σημείου m_{ij} βάσει της εξίσωσης (2.11). Ο αλγόριθμος μη-γραμμικής βελτιστοποίησης απαιτεί την ύπαρξη μιας αρχικής εκτίμησης για τις τιμές των αγνώστων παραμέτρων \mathbf{A} , \mathbf{R}_i και \mathbf{t}_i , η οποία μπορεί να υπολογιστεί με την αναλυτική μέθοδο που περιγράφεται παραπάνω.

Μέχρι στιγμής, δεν έχει ληφθεί υπ' όψη στους υπολογισμούς βαθμονόμησης κάμερας, η παραμόρφωση της λαμβανόμενης ψηφιακής εικόνας εξ' αιτίας της κυρτότητας του φακού της κάμερας. Ο παράγοντας που έχει τη μεγαλύτερη επίδραση στην παραμόρφωση της ψηφιακής εικόνας είναι η ακτινική παραμόρφωση (radial distortion) που προκαλεί ο φακός της κάμερας. Στο Σχ. 2.4 φαίνεται η πραγματική ψηφιακή εικόνα όπως ακριβώς λαμβάνεται από την κάμερα, και δίπλα η ίδια ιδανική εικόνα αν δεν υπήρχε καθόλου ακτινική παραμόρφωση.



Σχήμα 2.4 Ψηφιακή εικόνα με και χωρίς ακτινική παραμόρφωση

Έστω (u, v) οι ιδανικές συντεταγμένες εικόνας ενός παρατηρούμενου σημείου και (u', v') οι πραγματικές παραμορφωμένες συντεταγμένες του ίδιου σημείου, εξ' αιτίας της ακτινικής παραμόρφωσης. Οι ιδανικές συντεταγμένες αποτελούν στην πραγματικότητα την προβολή του παρατηρούμενου σημείου πάνω στο πλάνο εικόνας, με βάση το χρησιμοποιούμενο μοντέλο προβολής (pin-hole model). Ομοίως, (x, y) και (x', y') είναι οι ιδανικές και πραγματικές συντεταγμένες εικόνας, κανονικοποιημένες ως προς το κέντρο προβολής. Αυτές σχετίζονται με τις παρακάτω εξισώσεις:

$$\begin{aligned} x' &= x + x[k_1(x^2 + y^2) + k_2(x^2 + y^2)^2] \\ y' &= y + y[k_1(x^2 + y^2) + k_2(x^2 + y^2)^2] \end{aligned} \quad (2.27)$$

όπου k_1 και k_2 είναι οι δύο βασικοί συντελεστές ακτινικής παραμόρφωσης. Δεδομένων των σχέσεων $u' = u_0 + ax' + cy'$ και $v' = v_0 + by'$, έχουμε:

$$\begin{aligned} u' &= u + (u - u_0)[k_1(x^2 + y^2) + k_2(x^2 + y^2)^2] \\ v' &= v + (v - v_0)[k_1(x^2 + y^2) + k_2(x^2 + y^2)^2] \end{aligned} \quad (2.28)$$

Η ακτινική παραμόρφωση είναι αρκετά μικρή για να δημιουργήσει μεγάλες ανακρίβειες στον υπολογισμό των εσωτερικών-εξωτερικών παραμέτρων της κάμερας, δίχως αυτή να έχει ληφθεί υπ' όψη στους υπολογισμούς. Μία καλή προσέγγιση για την εξεύρεση των αγνώστων συντελεστών ακτινικής παραμόρφωσης k_1 και k_2 είναι ο υπολογισμός τους να γίνεται εκ των υστέρων του υπολογισμού των εσωτερικών-εξωτερικών παραμέτρων της κάμερας. Το μοντέλο προβολής χρησιμοποιείται για την εύρεση των ιδανικών (χωρίς ακτινική παραμόρφωση) συντεταγμένων εικόνας (u, v) των παρατηρούμενων σημείων,

χρησιμοποιώντας την βασική εξίσωση (2.11). Για ένα παρατηρούμενο σημείο, η (2.28) μπορεί να αναδιατυπωθεί ως εξής:

$$\begin{bmatrix} (u-u_0)(x^2+y^2) & (u-u_0)(x^2+y^2)^2 \\ (v-v_0)(x^2+y^2) & (v-v_0)(x^2+y^2)^2 \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} = \begin{bmatrix} u'-u \\ v'-v \end{bmatrix}. \quad (2.29)$$

Για όλα τα n παρατηρούμενα σημεία του μοντέλου βαθμονόμησης και για όλες τις m ψηφιακές εικόνες που έχουν ληφθεί από τις m διακριτές θέσεις της κάμερας, οι εξισώσεις της μορφής (2.29) μπορούν να συνδυαστούν για να σχηματίσουν ένα σύστημα $2mn$ γραμμικών εξισώσεων με δύο αγνώστους, της μορφής:

$$\mathbf{D}\mathbf{k} = \mathbf{d} \quad \text{όπου} \quad \mathbf{k} = [k_1, k_2]^T. \quad (2.30)$$

Η βέλτιστη λύση ελαχίστων τετραγώνων της (2.29) δίνεται με τη βοήθεια της ψευδο-αντίστροφης του πίνακα \mathbf{D} , όπως φαίνεται παρακάτω:

$$\mathbf{k} = (\mathbf{D}^T \mathbf{D})^{-1} \mathbf{D}^T \mathbf{d}. \quad (2.31)$$

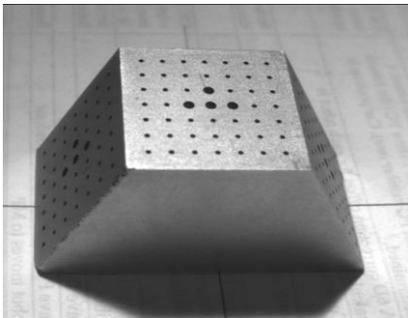
Εφόσον οι συντελεστές ακτινικής παραμόρφωσης k_1 και k_2 έχουν υπολογιστεί, μπορούμε να βελτιστοποιήσουμε όλες τις υπόλοιπες παραμέτρους της κάμερας (πίνακες εσωτερικών-εξωτερικών παραμέτρων), ελαχιστοποιώντας την παρακάτω μη-γραμμική τετραγωνική συνάρτηση κόστους:

$$f(\mathbf{A}, \mathbf{R}_i, \mathbf{t}_i) = \sum_{i=1}^m \sum_{j=1}^n \left\| m_{ij} - \hat{m}(\mathbf{A}, k_1, k_2, \mathbf{R}_i, \mathbf{t}_i, \mathbf{M}_j) \right\|^2 \quad (2.32)$$

όπου $\hat{m}(\mathbf{A}, k_1, k_2, \mathbf{R}_i, \mathbf{t}_i, \mathbf{M}_j)$ είναι η προβολή του σημείου \mathbf{M}_j του μοντέλου βαθμονόμησης πάνω στην ψηφιακή εικόνα i , χρησιμοποιώντας την βασική εξίσωση (2.11), υποβαλλόμενη σε τεχνητή ακτινική παραμόρφωση βάσει των εξισώσεων (2.28). Η χρησιμοποιούμενη μέθοδος μη-γραμμικής ελαχιστοποίησης είναι μια παραλλαγή της μεθόδου Levenberg-Marquardt, όπως και στην περίπτωση της ελαχιστοποίησης της (2.26).

2.3.2 Αλγόριθμος βαθμονόμησης κάμερας του R. Tsai

Ο αλγόριθμος βαθμονόμησης κάμερας του Tsai βασίζεται στη χρήση ενός μοντέλου βαθμονόμησης, το οποίο μπορεί να είναι είτε ένα επίπεδο αντικείμενο, όπως αυτό που χρησιμοποιείται στη μέθοδο βαθμονόμησης του Zhang, είτε ένα στέρεο αντικείμενο με



Σχήμα 2.5 Στέρεο (μη-επίπεδο) μοντέλο βαθμονόμησης κάμερας

διακριτά χαρακτηριστικά σημεία σε συγκεκριμένες θέσεις πάνω σε κάποιες επιφάνειές του. Στο Σχ. 2.5 φαίνονται δύο αντιπροσωπευτικά στέρεα μοντέλα βαθμονόμησης, των οποίων οι επιφάνειες περιέχουν χαρακτηριστικά σημεία (κουκίδες) με γνωστές 3D συντεταγμένες $\mathbf{P}_i(X_i, Y_i, Z_i)$ ως προς το σύστημα συντεταγμένων WXYZ του μοντέλου. Αυτό συνήθως ορίζεται αυθαίρετα στο κέντρο του στέρεου αντικειμένου, καθώς έτσι γίνεται ευκολότερος ο γεωμετρικός προσδιορισμός των 3D συντεταγμένων των χαρακτηριστικών σημείων. Όπως θα γίνει κατανοητό στη συνέχεια, η μεθοδολογία που ακολουθείται απ' τον αλγόριθμο του Tsai για την βαθμονόμηση της κάμερας είναι διαφορετική ανάλογα με το αν χρησιμοποιείται επίπεδο ή στέρεο μοντέλο βαθμονόμησης.

Ξεκινώντας από το γνωστό μοντέλο προβολής (pin-hole model) για μια ψηφιακή κάμερα, η εξίσωση που σχετίζει ένα παρατηρούμενο σημείο i , με συντεταγμένες $\mathbf{P}_i(X_i, Y_i, Z_i)$ ως προς ένα εξωτερικό σύστημα συντεταγμένων, με την προβολή του (u_i, v_i) πάνω στην ψηφιακή εικόνα, είναι:

$$i = 1, 2, \dots, n \quad z_i \begin{bmatrix} u_i \\ v_i \\ 1 \end{bmatrix} = \mathbf{A} [\mathbf{R} | \mathbf{t}] \begin{bmatrix} X_i \\ Y_i \\ Z_i \\ 1 \end{bmatrix} \quad (2.33)$$

$$\text{όπου} \quad \mathbf{A} = \begin{bmatrix} a & 0 & u_0 \\ 0 & b & v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (2.34)$$

$$\text{και} \quad \mathbf{R} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \quad \mathbf{t} = \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} \quad (2.35)$$

Από την (2.34) παρατηρούμε ότι ο αλγόριθμος του Tsai δεν λαμβάνει υπ' όψη τη σταθερά ορθογωνιότητας των αξόνων x' και y' του συστήματος συντεταγμένων του επιπέδου εικόνας, τους οποίους θεωρεί πάντα κάθετους μεταξύ τους ($c=0$). Απαλείφοντας από την (2.33) τη μεταβλητή z_i , και αντικαθιστώντας τις (2.34) και (2.35) μπορούμε να καταλήξουμε στις παρακάτω εξισώσεις:

$$u_i - u_0 = a \frac{r_{11}X_i + r_{12}Y_i + r_{13}Z_i + t_x}{r_{31}X_i + r_{32}Y_i + r_{33}Z_i + t_z} \quad (2.36)$$

$$\text{και} \quad v_i - v_0 = b \frac{r_{21}X_i + r_{22}Y_i + r_{23}Z_i + t_y}{r_{31}X_i + r_{32}Y_i + r_{33}Z_i + t_z} \quad (2.37)$$

Εφόσον έχουμε $a=k_x f$ και $b=k_y f$ (βλ. εξ. (2.6) και (2.7)), όπου f είναι η εστιακή απόσταση της κάμερας (focal length) και k_x, k_y είναι οι σταθερές αναλογίας που σχετίζονται με τις διαστάσεις του κάθε εικονοστοιχείου της ψηφιακής εικόνας, οι παραπάνω εξισώσεις μπορούν να αναδιατυπωθούν ως εξής:

$$\frac{u_i - u_0}{f} = k_x \frac{r_{11}X_i + r_{12}Y_i + r_{13}Z_i + t_x}{r_{31}X_i + r_{32}Y_i + r_{33}Z_i + t_z} \quad (2.38)$$

$$\text{και} \quad \frac{v_i - v_0}{f} = k_y \frac{r_{21}X_i + r_{22}Y_i + r_{23}Z_i + t_y}{r_{31}X_i + r_{32}Y_i + r_{33}Z_i + t_z} \quad (2.39)$$

α) Μη-επίπεδο μοντέλο βαθμονόμησης

Εξετάζουμε πρώτα την περίπτωση που το χρησιμοποιούμενο μοντέλο βαθμονόμησης δεν είναι επίπεδο ($Z_i \neq 0$). Έστω ότι λαμβάνουμε μια ψηφιακή εικόνα του μοντέλου, με την επεξεργασία της οποίας βρίσκουμε τις συντεταγμένες εικόνας (u_i, v_i) που αντιστοιχούν στις πραγματικές 3D συντεταγμένες $\mathbf{P}_i(X_i, Y_i, Z_i)$ των χαρακτηριστικών σημείων του μοντέλου. Αρχικά, υποθέτουμε ότι το κέντρο της εικόνας (u_0, v_0) είναι γνωστό από τα κατασκευαστικά στοιχεία της κάμερας. Σε περίπτωση που αυτά δεν είναι διαθέσιμα, μπορούμε να θέσουμε σαν αρχική εκτίμηση το κεντρικό σημείο της εικόνας ($u_0=x_{\max}/2, v_0=y_{\max}/2$). Διαιρώντας τις (2.38) και (2.39) κατά μέλη, και θέτοντας $s=k_x/k_y$, $u_i' = u_i - u_0, v_i' = v_i - v_0$, έχουμε:

$$\frac{u'_i}{v'_i} = s \frac{r_{11}X_i + r_{12}Y_i + r_{13}Z_i + t_x}{r_{21}X_i + r_{22}Y_i + r_{23}Z_i + t_y}. \quad (2.40)$$

Αναλύοντας την (2.40) στην μορφή γραμμικής εξίσωσης, έχουμε:

$$\begin{aligned} (X_i v'_i) s r_{11} + (Y_i v'_i) s r_{12} + (Z_i v'_i) s r_{13} + (v'_i) s t_x \\ - (X_i u'_i) r_{21} - (Y_i u'_i) r_{22} - (Z_i u'_i) r_{23} - (u'_i) t_y = 0 \end{aligned} \quad (2.41)$$

Η (2.41) είναι μια ομογενής γραμμική εξίσωση, με αγνώστους τις μεταβλητές $s \cdot r_{11}$, $s \cdot r_{12}$, $s \cdot r_{13}$, $s \cdot t_x$, r_{21} , r_{22} , r_{23} , t_x , και προέρχεται από την αντιστοιχία ενός γνωστού σημείου $\mathbf{P}_i(X_i, Y_i, Z_i)$ του μοντέλου βαθμονόμησης με την προβολή του (u_i, v_i) πάνω στην ψηφιακή εικόνα. Επαναλαμβάνοντας την διαδικασία για όλα τα $i=1,2,\dots,n$ σημεία του μοντέλου βαθμονόμησης λαμβάνουμε n εξισώσεις της μορφής της (2.41) με διαφορετικούς συντελεστές, αλλά κοινούς αγνώστους, σχηματίζοντας έτσι ένα ομογενές σύστημα n γραμμικών εξισώσεων της μορφής $\mathbf{Ax}=\mathbf{0}$, όπου \mathbf{x} το 8×1 διάνυσμα των αγνώστων και \mathbf{A} είναι ο $n \times 8$ πίνακας των συντελεστών. Δεδομένου ότι το σύστημα είναι ομογενές, το διάνυσμα των αγνώστων που αποτελεί την λύση του πρέπει να οριστεί με βάση μια σταθερά αναλογίας (scale factor). Αυτό δικαιολογείται από το γεγονός, ότι αν ένα διάνυσμα \mathbf{x} αποτελεί λύση του ομογενούς συστήματος γραμμικών εξισώσεων $\mathbf{Ax}=\mathbf{0}$, τότε το κάθε διάνυσμα $\mathbf{x}'=\lambda\mathbf{x}$ θα αποτελεί επίσης λύση του συστήματος, όπου λ μια αυθαίρετη σταθερά. Η βέλτιστη λύση του συστήματος, υπό την έννοια των ελαχίστων τετραγώνων, είναι το ιδιοδιάνυσμα του πίνακα $\mathbf{A}^T\mathbf{A}$ που αντιστοιχεί στην μικρότερη ιδιοτιμή του. Έστω ότι η βέλτιστη λύση ελαχίστων τετραγώνων είναι η παρακάτω:

$$\mathbf{x} = \left[s r'_{11}, s r'_{12}, s r'_{13}, s t'_x, r'_{21}, r'_{22}, r'_{23}, t'_y \right]^T, \quad (2.42)$$

τότε λαμβάνοντας υπ' όψη και την σταθερά αναλογίας λ του ομογενούς συστήματος, οποιοδήποτε διάνυσμα της μορφής:

$$\mathbf{x}' = \left[\lambda s r'_{11}, \lambda s r'_{12}, \lambda s r'_{13}, \lambda s t'_x, \lambda r'_{21}, \lambda r'_{22}, \lambda r'_{23}, \lambda t'_y \right]^T \quad (2.43)$$

αποτελεί επίσης λύση του ομογενούς συστήματος $\mathbf{Ax}=\mathbf{0}$. Η σταθερά αναλογίας λ μπορεί να υπολογιστεί από τη συνθήκη μοναδιαίου μέτρου για τις γραμμές και τις στήλες του ορθοκανονικού πίνακα στροφής \mathbf{R} της σχέσης (2.35). Για την πρώτη και τη δεύτερη γραμμή του πίνακα στροφής έχουμε:

$$r_{11}^2 + r_{12}^2 + r_{13}^2 = 1 \quad \text{και} \quad r_{21}^2 + r_{22}^2 + r_{23}^2 = 1, \quad (2.44)$$

$$\text{οπότε} \quad (\lambda r'_{21})^2 + (\lambda r'_{22})^2 + (\lambda r'_{23})^2 = 1 \Rightarrow \lambda = 1/\sqrt{r_{21}^{\prime 2} + r_{22}^{\prime 2} + r_{23}^{\prime 2}} \quad (2.45)$$

$$\text{και} \quad (\lambda r'_{11})^2 + (\lambda r'_{12})^2 + (\lambda r'_{13})^2 = 1 \Rightarrow \lambda/s = 1/\sqrt{(sr'_{11})^2 + (sr'_{12})^2 + (sr'_{13})^2}. \quad (2.46)$$

Από τις (2.45) και (2.46) υπολογίζονται η σταθερά αναλογίας λ και ο συντελεστής αναλογίας διαστάσεων εικονοστοιχείου $s=k_x/k_y$ συναρτήσει του γνωστού διανύσματος \mathbf{x} της σχέσης (2.42). Επίσης υπολογίζονται και οι δύο πρώτες γραμμές του πίνακα στροφής \mathbf{R} ως $r_{ij} = \lambda r'_{ij}$, οι οποίες στη συνέχεια μπορούν να βελτιστοποιηθούν ώστε τα διανύσματά τους να είναι κάθετα μεταξύ τους (συνθήκη ορθοκανονικότητας). Η τρίτη γραμμή του \mathbf{R} γίνεται γνωστή από το εξωτερικό γινόμενο των διανυσμάτων των δύο πρώτων γραμμών.

β) Επίπεδο μοντέλο βαθμονόμησης

Στην περίπτωση που το χρησιμοποιούμενο μοντέλο βαθμονόμησης είναι επίπεδο, είναι βολικό να ορίσουμε την αρχή του συστήματος συντεταγμένων του στο κεντρικό του σημείο, με αποτέλεσμα όλα τα χαρακτηριστικά του σημεία $i=1,2,\dots,n$ να έχουν 3D συντεταγμένες $\mathbf{P}_i(X_i, Y_i, Z_i)$ όπου $Z_i=0$. Δυστυχώς, με τη χρήση επίπεδου μοντέλου βαθμονόμησης είναι αδύνατη η εύρεση της αναλογίας διαστάσεων εικονοστοιχείου $s=k_x/k_y$, οπότε πρέπει να υποθέσουμε ότι οι διαστάσεις όλων των εικονοστοιχείων των ψηφιακών εικόνων είναι ακριβώς τετραγωνικές ($k_x=k_y$, $s=1$). Με βάση τα παραπάνω, η εξίσωση (2.40) που σχετίζει τις 3D συντεταγμένες $\mathbf{P}_i(X_i, Y_i, Z_i)$ των χαρακτηριστικών σημείων του μοντέλου βαθμονόμησης με τις αντίστοιχες συντεταγμένες εικόνας (u_i, v_i) των προβολών τους, παίρνει την παρακάτω μορφή:

$$\frac{u_i'}{v_i'} = \frac{r_{11}X_i + r_{12}Y_i + t_x}{r_{21}X_i + r_{22}Y_i + t_y}. \quad (2.47)$$

Αναλύοντας την (2.47) μπορούμε να καταλήξουμε στην γραμμική εξίσωση:

$$(X_i v_i')r_{11} + (Y_i v_i')r_{12} + (v_i')t_x - (X_i u_i')r_{21} - (Y_i u_i')r_{22} - (u_i')t_y = 0 \quad (2.48)$$

η οποία έχει αγνώστους τις μεταβλητές r_{11} , r_{12} , r_{21} , r_{22} , t_x , t_y . Επαναλαμβάνοντας την διαδικασία για όλα τα $i=1,2,\dots,n$ χαρακτηριστικά σημεία του μοντέλου βαθμονόμησης και τις αντίστοιχες προβολές τους (u_i, v_i) πάνω στην ψηφιακή εικόνα, μπορούμε από τις εξισώσεις (2.48) να σχηματίσουμε ένα ομογενές σύστημα n γραμμικών εξισώσεων με 6

αγνώστους της μορφής $\mathbf{Ax}=\mathbf{0}$, όπου \mathbf{x} το 6×1 διάνυσμα των αγνώστων και \mathbf{A} είναι ο $n \times 6$ πίνακας των συντελεστών. Όπως και στην περίπτωση του μη-επίπεδου μοντέλου βαθμονόμησης, η λύση του ομογενούς γραμμικού συστήματος πρέπει να δηλωθεί με βάση μια αυθαίρετη σταθερά αναλογίας λ , καθώς αν το διάνυσμα \mathbf{x} αποτελεί λύση του συστήματος τότε οποιοδήποτε διάνυσμα $\mathbf{x}'=\lambda\mathbf{x}$ αποτελεί επίσης λύση. Η βέλτιστη λύση του συστήματος, υπό την έννοια των ελαχίστων τετραγώνων, είναι το ιδιοδιάνυσμα του πίνακα $\mathbf{A}^T\mathbf{A}$ που αντιστοιχεί στην μικρότερη ιδιοτιμή του. Έστω η βέλτιστη λύση του συστήματος είναι η παρακάτω:

$$\mathbf{x} = \left[r'_{11}, r'_{12}, r'_{21}, r'_{22}, t'_x, t'_y \right]^T, \quad (2.49)$$

τότε λαμβάνοντας υπ' όψη και την σταθερά αναλογίας λ του ομογενούς συστήματος, οποιοδήποτε διάνυσμα της μορφής:

$$\mathbf{x}' = \left[\lambda r'_{11}, \lambda r'_{12}, \lambda r'_{21}, \lambda r'_{22}, \lambda t'_x, \lambda t'_y \right]^T \quad (2.50)$$

αποτελεί επίσης λύση του ομογενούς συστήματος $\mathbf{Ax}=\mathbf{0}$. Η σταθερά αναλογίας λ , καθώς και οι συντελεστές r_{13} και r_{23} του πίνακα στροφής \mathbf{R} που λείπουν από τις εξισώσεις για το επίπεδο μοντέλο βαθμονόμησης, μπορούν να υπολογιστούν από τις παρακάτω εξισώσεις ορθοκανονικότητας του πίνακα στροφής \mathbf{R} :

$$r_{11}^2 + r_{12}^2 + r_{13}^2 = 1 \Rightarrow r_{11}'^2 + r_{12}'^2 + r_{13}'^2 = 1/\lambda^2 \quad (2.51)$$

$$r_{21}^2 + r_{22}^2 + r_{23}^2 = 1 \Rightarrow r_{21}'^2 + r_{22}'^2 + r_{23}'^2 = 1/\lambda^2 \quad (2.52)$$

$$\vec{r}_1 \cdot \vec{r}_2 = 0 \Rightarrow r'_{11}r'_{21} + r'_{12}r'_{22} + r'_{13}r'_{23} = 0 \quad (2.53)$$

Συνδυάζοντας όλες τις παραπάνω εξισώσεις, κι απαλείφοντας τις μεταβλητές r_{13} και r_{23} καταλήγουμε στην παρακάτω εξίσωση:

$$1/\lambda^4 - 1/\lambda^2 (r_{11}'^2 + r_{12}'^2 + r_{21}'^2 + r_{22}'^2) + (r'_{11}r'_{22} - r'_{12}r'_{21})^2 = 0 \quad (2.54)$$

η οποία μπορεί να λυθεί ως προς $1/\lambda^2$. Στη συνέχεια, συντελεστές r_{13} και r_{23} μπορούν να βρεθούν με τη χρήση των (2.51) και (2.52):

$$r_{13}'^2 = 1/\lambda^2 - (r_{11}'^2 + r_{12}'^2) \quad (2.55)$$

$$r_{23}'^2 = 1/\lambda^2 - (r_{21}'^2 + r_{22}'^2) \quad (2.56)$$

Μόνο η μεγαλύτερη τιμή των λύσεων (περισσότερο θετική) της (2.54) ως προς $1/\lambda^2$ μπορεί να δώσει θετικά πρόσημα για τα δεξιά μέρη των εξισώσεων (2.55) και (2.56). Προφανώς, η εξ' αυτών κρατάμε μόνο την θετική ρίζα για την ζητούμενη σταθερά

αναλογίας λ , από την οποία προκύπτουν οι πραγματικές τιμές των r_{11} , r_{12} , r_{21} , r_{22} , t_x , t_y σύμφωνα με τη σχέση (2.50). Τα πρόσημα των r_{13} και r_{23} μπορούν να βρεθούν με τη βοήθεια της (2.53), σύμφωνα με την οποία το πρόσημο του γινομένου των μεταβλητών πρέπει να ισούται με:

$$r_{13}'r_{23}' = -(r_{11}'r_{21}' + r_{12}'r_{22}') \quad (2.57)$$

Τέλος, η τρίτη γραμμή του πίνακα στροφής \mathbf{R} μπορεί να υπολογιστεί ως το εξωτερικό γινόμενο των διανυσμάτων των δύο πρώτων γραμμών:

$$\vec{r}_3 = \vec{r}_1 \times \vec{r}_2 \quad (2.58)$$

Μέχρι στιγμής, είτε για την περίπτωση του επίπεδου είτε του μη-επίπεδου μοντέλου βαθμονόμησης, έχουμε υπολογίσει τον πίνακα στροφής \mathbf{R} καθώς και τις μετατοπίσεις t_x και t_y των εξωτερικών παραμέτρων της κάμερας, έχοντας χρησιμοποιήσει ως αρχική εκτίμηση ότι το οπτικό κέντρο της κάμερας συμπίπτει με το κέντρο της ψηφιακής εικόνας ($u_0=x_{\max}/2$, $v_0=y_{\max}/2$). Η εστιακή απόσταση f και η μετατόπιση t_z μπορούν να υπολογιστούν με βάση τις (2.38) και (2.39), φέρνοντας τις στην μορφή γραμμικών εξισώσεων ως εξής:

$$(r_{11}X_i + r_{12}Y_i + r_{13}Z_i + t_x)sk_y f - u_i t_z = (r_{31}X_i + r_{32}Y_i + r_{33}Z_i)u_i \quad (2.59)$$

$$(r_{21}X_i + r_{22}Y_i + r_{23}Z_i + t_y)k_x f - v_i t_z = (r_{31}X_i + r_{32}Y_i + r_{33}Z_i)v_i \quad (2.60)$$

Προφανώς στην περίπτωση επίπεδου μοντέλου βαθμονόμησης ισχύει $k_x=k_y=k$ ($s=1$) και $Z_i=0$ για όλα τα χαρακτηριστικά σημεία του, οπότε οι εξισώσεις (2.59) και (2.60) διαμορφώνονται ανάλογα. Λαμβάνοντας υπ' όψη όλα τα n χαρακτηριστικά σημεία του μοντέλου βαθμονόμησης και εξάγοντας τις εξισώσεις της μορφής των (2.59) και (2.60) για καθένα απ' αυτά, μπορούμε να σχηματίσουμε ένα σύστημα $2n$ γραμμικών εξισώσεων με αγνώστους τις μεταβλητές $k_y f$ και t_z , το οποίο μπορούμε εύκολα να λύσουμε με την ψευδο-αντίστροφη του πίνακα των συντελεστών. Στην περίπτωση της χρήσης επίπεδου μοντέλου βαθμονόμησης ίσως είναι προτιμότερο το σύστημα των γραμμικών εξισώσεων να προκύψει μόνο από εξισώσεις της μορφής της (2.60), καθώς σε κάποιες περιπτώσεις η προσέγγιση $s=1$ (εικονοστοιχείο τετραγωνικών διαστάσεων) δεν είναι ακριβής. Η μεταβλητές $a=k_x f$ και $b=k_y f$ είναι στην πραγματικότητα η εστιακή απόσταση της κάμερας μετασχηματισμένη σε συντεταγμένες εικόνας, κατά τον x και κατά τον y άξονα αντίστοιχα του συστήματος συντεταγμένων του πλάνου εικόνας. Σύμφωνα με τις

εξισώσεις (2.36-2.39) ο υπολογισμός των μεταβλητών a και b είναι επαρκής για τον μετασχηματισμό των 3D συντεταγμένων $\mathbf{P}_i(X_i, Y_i, Z_i)$ ενός παρατηρούμενου σημείου στις αντίστοιχες συντεταγμένες εικόνας της προβολής του πάνω στην ψηφιακή εικόνα. Αντιθέτως, ο υπολογισμός των απόλυτων τιμών των μεταβλητών k_x , k_y και f δεν είναι απαραίτητος.

Ας εξετάσουμε την ειδική περίπτωση όπου όλα τα χαρακτηριστικά σημεία $\mathbf{P}_i(X_i, Y_i, Z_i)$ ανήκουν σ' ένα κοινό επίπεδο (επίπεδο μοντέλο βαθμονόμησης), το οποίο είναι κάθετο με το z -άξονα του συστήματος συντεταγμένων της κάμερας $Oxyz$ (βλ. σχήμα 2.1). Η προβολή των σημείων πάνω στον z -άξονα της κάμερας θα είναι:

$$z_i = r_{31}X_i + r_{32}Y_i + r_{33}Z_i + t_z \Rightarrow z_i = t_z \quad (2.61)$$

καθώς θα ισχύει
$$r_{31}X_i + r_{32}Y_i + r_{33}Z_i = 0. \quad (2.62)$$

Στην περίπτωση αυτή το σύστημα των γραμμικών εξισώσεων που σχηματίζεται από τις (2.59) και (2.60) θα γίνει ομογενές της μορφής $\mathbf{Ax}=\mathbf{0}$, οπότε μπορεί να προκύψει μοναδική λύση μόνο για την αναλογία $k_y f/t_z$. Στην πράξη αυτό σημαίνει ότι για να προκύψουν μοναδικές λύσεις για τις μεταβλητές $k_y f$ και t_z πρέπει είτε το μοντέλο βαθμονόμησης να είναι μη-επίπεδο, είτε στην περίπτωση που είναι επίπεδο η θέση του να μην είναι ακριβώς κάθετη με τον οπτικό άξονα της κάμερας.

Η παραπάνω γραμμική μέθοδος βαθμονόμησης κάμερας του R. Tsai δεν μπορεί να δώσει ακριβή αποτελέσματα, παρά μόνο μια αρχική εκτίμηση των ζητούμενων εσωτερικών-εξωτερικών παραμέτρων της κάμερας, για τους παρακάτω λόγους:

α) Όλες οι εξισώσεις βασίζονται στην υπόθεση ότι η θέση του οπτικού κέντρου της κάμερας (u_0, v_0) συμπίπτει με το κεντρικό σημείο της εικόνας ($u_0=x_{\max}/2, v_0=y_{\max}/2$), η οποία αρκετές φορές απέχει πολύ από την πραγματικότητα.

β) Τα συστήματα των γραμμικών εξισώσεων, στα οποία καταλήγει η γραμμική μέθοδος βαθμονόμησης, σχηματίζονται από τον συνδυασμό διαφόρων μεταβλητών, καθεμιά από τις οποίες εκφράζει διαφορετικά φυσικά μεγέθη. Αυτό έχει σαν αποτέλεσμα τα διανύσματα των λύσεων τους να περιέχουν συνδυασμούς μεταβλητών διαφόρων φυσικών μεγεθών, τα οποία στην πράξη δεν έχουν ξεκάθαρη φυσική σημασία.

γ) Στην περίπτωση χρήσης επίπεδου μοντέλου βαθμονόμησης, δεν είναι δυνατός ο υπολογισμός της αναλογίας διαστάσεων των εικονοστοιχείων $s=k_x/k_y$, οπότε όλοι οι

υπολογισμοί βασίζονται στην υπόθεση ότι τα εικονοστοιχεία των ψηφιακών εικόνων είναι τετραγωνικών διαστάσεων ($s=1$).

δ) Ακόμη δεν έχει ληφθεί υπ' όψη η ακτινική παραμόρφωση των εικόνων ούτε από το μοντέλο προβολής της κάμερας (pin-hole model) ούτε από τους υπολογισμούς που σχετίζονται με την γραμμική μέθοδο βαθμονόμησής της.

Οι παραπάνω ανακρίβειες της γραμμικής μεθόδου μπορούν εν μέρει να εξαλειφθούν, χρησιμοποιώντας την μη-γραμμική μέθοδο βελτιστοποίησης, όπως αυτή περιγράφεται στο κεφ. 2.3.2 του αλγόριθμου βαθμονόμησης του Zhang. Ορίζουμε την παρακάτω θετική τετραγωνική συνάρτηση σφάλματος:

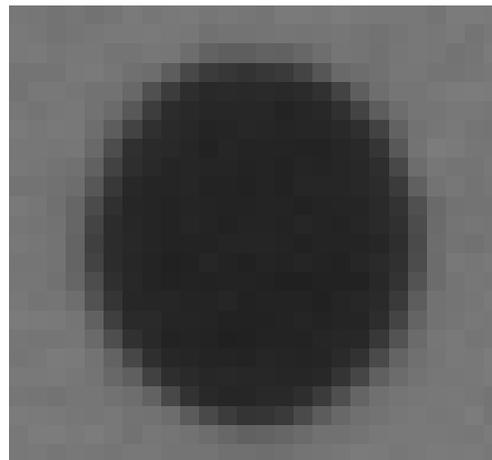
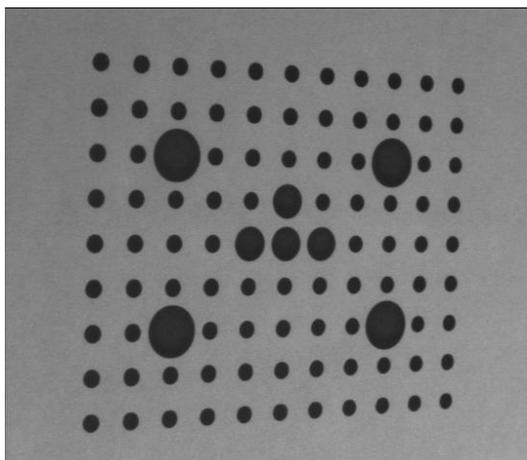
$$h(f, s, \mathbf{R}, \mathbf{t}, k_1, k_2) = \sum_{i=1}^n (u_i - u_i')^2 + \sum_{i=1}^n (v_i - v_i')^2 \quad (2.62)$$

όπου k_1 και k_2 είναι οι συντελεστές της ακτινικής παραμόρφωσης, (u_i, v_i) είναι οι πραγματικές συντεταγμένες εικόνας του i -χαρακτηριστικού σημείου του μοντέλου βαθμονόμησης, όπως αυτές προέρχονται από την ψηφιακή επεξεργασία της εικόνας, και (u_i', v_i') είναι οι συντεταγμένες εικόνας του ίδιου σημείου σύμφωνα με το χρησιμοποιούμενο μοντέλο προβολής της κάμερας (εξ. (2.33)), οι οποίες είναι τεχνητά ακτινικά παραμορφωμένες σύμφωνα με τις εξισώσεις (2.28). Εφόσον η γραμμική μέθοδος του αλγόριθμου του Tsai δεν υπολογίζει τους συντελεστές k_1 και k_2 , χρησιμοποιούμε μηδενικές τιμές σαν αρχικές εκτιμήσεις για τη μη-γραμμική μέθοδο βελτιστοποίησης. Όπως και στην περίπτωση του αλγορίθμου του Zhang, η χρησιμοποιούμενη μέθοδος μη-γραμμικής βελτιστοποίησης είναι μια παραλλαγή της μεθόδου Levenberg-Marquardt [Fletcher, 1990; Gill et al., 1989], όπως αυτή υλοποιείται από το πακέτο λογισμικού MINPACK [Moré et al., 1980, 1984], όπου ο υπολογισμός του διανύσματος παραγώγων (gradient vector) υλοποιείται εσωτερικά από τον αλγόριθμο.

2.4 Ψηφιακή επεξεργασία εικόνας

Οι αλγόριθμοι βαθμονόμησης κάμερας που παρουσιάστηκαν στα προηγούμενα κεφάλαια προϋποθέτουν τη χρήση εξειδικευμένων τεχνικών για την εξαγωγή χρήσιμης πληροφορίας από τις ψηφιακές εικόνες που λαμβάνουν οι κάμερες. Η χρήσιμη

πληροφορία, που στην προκειμένη περίπτωση συνίσταται στις συντεταγμένες εικόνας των χαρακτηριστικών σημείων του μοντέλου βαθμονόμησης πάνω στην ψηφιακή εικόνα, μπορεί να γίνει διαθέσιμη με την χρήση κατάλληλων ειδικών αλγορίθμων επεξεργασίας εικόνας. Υπάρχει ένας πολύ μεγάλος αριθμός αλγορίθμων επεξεργασίας εικόνας, καθένας από τους οποίους είναι κατάλληλος για την ανίχνευση συγκεκριμένων χαρακτηριστικών πάνω σε μια ψηφιακή εικόνα, όπως για παράδειγμα ακμών, γωνιών, κύκλων, κτλ. Στην παρούσα φάση θα περιοριστούμε στην ανίχνευση κουκίδων και κύκλων, διαδικασία απαραίτητη για την βαθμονόμηση των χρησιμοποιούμενων καμερών και για την μετέπειτα χρήση τους για την μέτρηση αυθαίρετων χαρακτηριστικών σημείων στον περιβάλλοντα χώρο εργασίας.



Σχήμα 2.6α Ψηφιακή εικόνα προς επεξεργασία

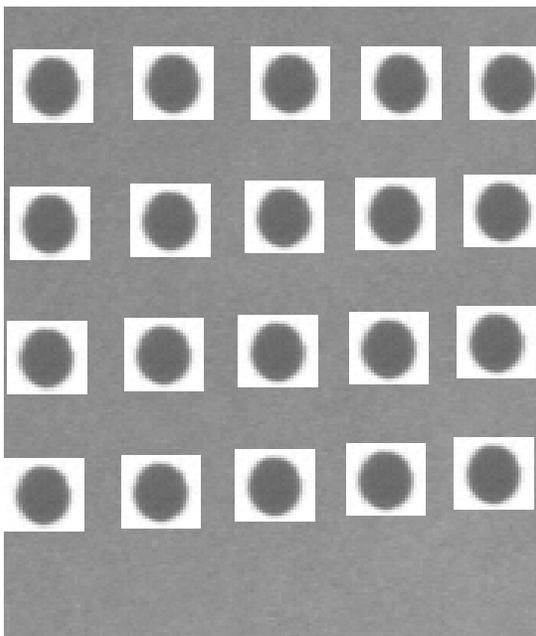
Σχήμα 2.6β Μάσκα συνέλιξης

Ας υποθέσουμε ότι έχουμε μια ψηφιακή εικόνα ενός επίπεδου μοντέλου βαθμονόμησης, όπως αυτήν του Σχ. 2.6α, από την οποία απαιτείται να εξαχθούν οι συντεταγμένες εικόνας όλων των χαρακτηριστικών σημείων (κουκίδων). Μία προσέγγιση για την λύση του προβλήματος είναι ένας χονδρικός προσδιορισμός της θέσης των κουκίδων πάνω στην ψηφιακή εικόνα, με τη χρήση ενός αλγορίθμου συνέλιξης [Katsuri and Schunck, 1995] (pattern matching), και ο μετέπειτα ακριβής υπολογισμός των συντεταγμένων των κέντρων τους με τη χρήση αλγορίθμων υπολογισμού κέντρου βάρους [Shortis et al., 1994]. Ο αλγόριθμος συνέλιξης χρησιμοποιεί μια εικόνα-μάσκα, όπως αυτή που φαίνεται στο Σχ. 2.6β, την οποία

«σαρώνει» πάνω στην αρχική εικόνα προσπαθώντας να βρει σημεία βέλτιστης ταύτισης. Ουσιαστικά η μάσκα εκφράζει ένα μικρό παράθυρο το οποίο τοποθετείται σε κάθε δυνατή θέση πάνω στην αρχική εικόνα, ψάχνοντας την θέση (ή τις θέσεις) όπου υπάρχει βέλτιστη ταύτιση (ταίριασμα) μεταξύ της μάσκας και του εξεταζόμενου παραθύρου της αρχικής εικόνας. Εάν η αρχική εικόνα έχει διαστάσεις $M \times N$ εικονοστοιχεία και η χρησιμοποιούμενη μάσκα $K \times L$ εικονοστοιχεία (προφανώς $K \ll M$ και $L \ll N$) τότε η τιμή της συνέλιξης σε μια τυχαία θέση της αρχικής εικόνας (u, v) είναι:

$$(f * g)(u, v) = \sum_{i=1}^k \sum_{j=1}^l f[u+i, u+j]g[i, j] \quad (2.63)$$

όπου $f[u, v]$ και $g[u, v]$ είναι οι διακριτές συναρτήσεις που χαρακτηρίζουν τη φωτεινότητα των εικονοστοιχείων (pixel brightness) της αρχικής ψηφιακής εικόνας και της εικόνας-μάσκας αντίστοιχα, στις θέσεις με συντεταγμένες (u, v) . Χρησιμοποιώντας την ίδια τεχνική για όλες τις δυνατές θέσεις πάνω στην αρχική εικόνα ($u=1, \dots, M$ και $v=1, \dots, N$) μπορούμε να βρούμε κάποια σημεία (u_0, v_0) όπου η συνάρτηση συνέλιξης της μορφής (2.63) παρουσιάζει τοπικά μέγιστα. Τα σημεία αυτά αντιστοιχούν στις θέσεις της αρχικής εικόνας όπου η εικόνα-μάσκα «ταιριάζει» καλύτερα πάνω στη εξεταζόμενη περιοχή της αρχικής εικόνας. Αν υποθέσουμε ότι διακριτή συνάρτηση $f[u, v]$ χαρακτηρίζει την



Σχήμα 2.7 Συνέλιξη αρχικής εικόνας και εικόνας-μάσκας

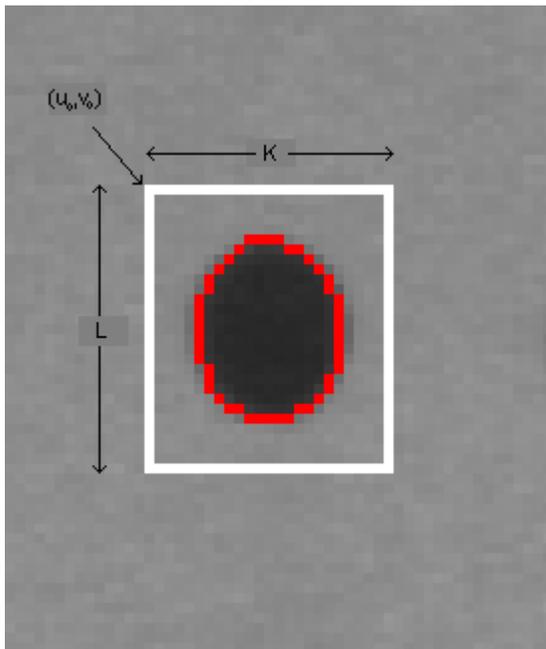
ψηφιακή εικόνα του Σχ. 2.6α και η διακριτή συνάρτηση $g[u, v]$ χαρακτηρίζει την εικόνα-μάσκα του Σχ. 2.6β, τότε η συνάρτηση συνέλιξης θα παρουσιάζει τοπικά μέγιστα σε $9 \times 11 = 99$ σημεία, όσα και τα χαρακτηριστικά σημεία (κουκίδες) της εικόνας του μοντέλου. Η διαδικασία γίνεται καλύτερα κατανοητή με την βοήθεια του Σχ. 2.7, όπου φαίνονται κάποια σημεία ταιριάσματος της μάσκας πάνω στην αρχική εικόνα.

Αποτέλεσμα του αλγορίθμου συνέλιξης είναι η εύρεση 99 σημείων πάνω στην αρχική εικόνα με συντεταγμένες (u_0, v_0) , τα οποία αντιστοιχούν στα σημεία βέλτιστης ταύτισης

της εξεταζόμενης περιοχής της εικόνας με την εικόνα-μάσκα (βλ. Σχ. 2.7). Στην πραγματικότητα οι συντεταγμένες (u_0, v_0) περιγράφουν την πάνω-αριστερή γωνία του παραθύρου ταύτισης (ταιριάσματος) καθώς η αρχή συντεταγμένων των ψηφιακών εικόνων βρίσκεται στην πάνω-αριστερή γωνία του πλαισίου τους. Οι συντεταγμένες αυτές προφανώς δεν αντιστοιχούν στα ζητούμενα κέντρα των κύκλων (κουκίδων) αλλά ορίζουν τις περιοχές (παράθυρα) πάνω στην εξεταζόμενη εικόνα μέσα στις οποίες βρίσκονται οι ζητούμενες κουκίδες.

Ο απλούστερος υπολογισμός των συντεταγμένων εικόνας των κέντρων των κουκίδων μπορεί να γίνει αν υποθέσουμε ότι αυτές συμπίπτουν με το κέντρο του παραθύρου «ταιριάσματος» της εικόνας-μάσκας πάνω στην αρχική εικόνα. Έτσι αν η συνάρτηση συνέλιξης (2.63) δώσει ένα τοπικό μέγιστο στο σημείο (u_0, v_0) και η εικόνα μάσκα έχει διαστάσεις $K \times L$ εικονοστοιχεία τότε οι συντεταγμένες του κέντρου δίνονται από τη σχέση:

$$\mathbf{k} = (k_x, k_y) = (u_0 + K/2, v_0 + L/2) \quad (2.64)$$



Σχήμα 2.8 Ανίχνευση περιγράμματος

Η παραπάνω μέθοδος συνήθως δεν χρησιμοποιείται σε πρακτικές εφαρμογές εξαιτίας της περιορισμένης ακρίβειάς της και της μεγάλης ευαισθησίας της στο θόρυβο της ψηφιακής εικόνας. Εξάλλου ο αλγόριθμος συνέλιξης είναι κατάλληλος μόνο για μια χονδρική εκτίμηση της θέσης ενός ζητούμενου χαρακτηριστικού πάνω σε μια ψηφιακή εικόνα. Μία διαφορετική προσέγγιση του προβλήματος είναι η εύρεση του περιγράμματος (*contour*) της κουκίδας που ανήκει στην περιοχή του εξεταζόμενου παραθύρου «ταιριάσματος», βάσει του οποίου θα γίνει ο υπολογισμός του

ζητούμενου κέντρου. Το περίγραμμα της κουκίδας μπορεί να βρεθεί με την εφαρμογή ενός αλγορίθμου υπολογισμού ακμών (*edge detection*) μέσα στο εξεταζόμενο παράθυρο της αρχικής εικόνας, ο οποίος στην πραγματικότητα βρίσκει τα εικονοστοιχεία που

παρουσιάζουν μέγιστη μεταβολή φωτεινότητας σε σχέση με τα γειτονικά τους εικονοστοιχεία. Αποτέλεσμα του αλγορίθμου είναι ένα σύνολο εικονοστοιχείων τα οποία αντιπροσωπεύουν το περίγραμμα της συγκεκριμένης κουκίδας, όπως φαίνεται και στο Σχ. 2.8. Εάν $\mathbf{I}_i = (I_{i(x)}, I_{i(y)})$, $i = 1, 2, \dots, n$, είναι οι συντεταγμένες εικόνας του περιγράμματος, τότε οι συντεταγμένες του κέντρου της κουκίδας μπορούν να βρεθούν με έναν από τους παρακάτω τρόπους [Shortis et al., 1994]:

α) Μέσος όρος περιμέτρου: $\mathbf{k} = 1/n \cdot \sum_{i=1}^n \mathbf{I}_i$.

β) Δυαδικό κέντρο (Binary centroid): $\mathbf{k} = \begin{bmatrix} k_x \\ k_y \end{bmatrix} = \frac{\sum_{j=1}^L \sum_{i=1}^K \begin{bmatrix} i \\ j \end{bmatrix} \cdot I_{i,j}}{\sum_{j=1}^L \sum_{i=1}^K I_{i,j}}$, όπου $I_{i,j}$

είναι 1 εάν το εικονοστοιχείο (i, j) βρίσκεται στην εσωτερική επιφάνεια του περιγράμματος, αλλιώς $I_{i,j} = 0$.

γ) Διαβαθμίσμο κέντρο (Grey scale centroid): το ίδιο με την περίπτωση β) αλλά το $I_{i,j}$ είναι η τιμή φωτεινότητας του εικονοστοιχείου (i, j) εάν αυτό βρίσκεται στην εσωτερική επιφάνεια του περιγράμματος, αλλιώς $I_{i,j} = 0$.

δ) Τετραγωνικό διαβαθμίσμο κέντρο (Squared grey scale centroid): το ίδιο με την περίπτωση γ) αλλά οι τιμές φωτεινότητας $I_{i,j}$ είναι τετραγωνισμένες.

ε) Προσαρμογή επίπεδης έλλειψης (Ellipse fitting): η μέθοδος αυτή επιχειρεί να ταιριάξει μια αυθαίρετη 2D έλλειψη πάνω στα εικονοστοιχεία του περιγράμματος της κουκίδας, υπό την έννοια των ελαχίστων τετραγώνων. Μία 2D έλλειψη χαρακτηρίζεται από το κέντρο της, τις διαστάσεις του μικρού και του μεγάλου άξονά της, καθώς και από την γωνία στροφής του μεγάλου άξονα ως προς τον άξονα x του σταθερού συστήματος συντεταγμένων. Ο αλγόριθμος προσαρμογής της έλλειψης πάνω στο περίγραμμα είναι επαναληπτικός (iterative) και απαιτεί την ύπαρξη κάποιων αρχικών εκτιμήσεων για τις μεταβλητές της έλλειψης. Μία αρχική εκτίμηση του κέντρου της έλλειψης μπορεί να βρεθεί με έναν από τους παραπάνω τρόπους (α)-(δ) ενώ οι διαστάσεις των αξόνων μπορούν να προσεγγιστούν με τον υπολογισμό των μέγιστων αποστάσεων του περιγράμματος από το κέντρο. Η γωνία στροφής μπορεί αρχικά να θεωρηθεί μηδενική. Αφού ο αλγόριθμος προσαρμογής συγκλίνει, το κέντρο της τελικής έλλειψης εκλαμβάνεται ως το ζητούμενο κέντρο \mathbf{k} της εξεταζόμενης κουκίδας.

στ) Προσαρμογή Gaussian συνάρτησης (Gaussian fitting): κατ' αντιστοιχία με την προσαρμογή επίπεδης έλλειψης, η μέθοδος αυτή επιχειρεί να ταιριάξει μια 2D Gaussian κατανομή με την εσωτερική επιφάνεια του περιγράμματος της εξεταζόμενης κουκίδας. Η κατανομή δίνεται από την παρακάτω διακριτή συνάρτηση:

$$G(x, y) = \frac{K}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \exp\left\{\frac{-1}{2(1-\rho^2)}\left((x/\sigma_x)^2 - 2\rho(x/\sigma_x)(y/\sigma_y) + (y/\sigma_y)^2\right)\right\}$$

όπου:

σ_x και σ_y είναι οι τυπικές αποκλίσεις της κατανομής στους άξονες x και y αντίστοιχα, ρ είναι ο συντελεστής συσχέτισης (correlation coefficient),

K είναι ο συντελεστής αναλογίας,

(x, y) είναι οι τιμές απόκλισης από τη μέση τιμή της κατανομής.

Ο αλγόριθμος προσαρμογής είναι επαναληπτικός (iterative) και απαιτεί την ύπαρξη κάποιων αρχικών τιμών για τους συντελεστές της κατανομής. Κατ' αναλογία με την μέθοδο προσαρμογής έλλειψης, η μέση τιμή της κατανομής (κεντρικό σημείο) μπορεί να βρεθεί τη χρήση μίας από τις μεθόδους (α)-(δ), οι τυπικές αποκλίσεις σ_x και σ_y μπορούν να εξαχθούν από τις μέγιστες αποστάσεις του κεντρικού σημείου από το περίγραμμα της κουκίδας, ενώ ο συντελεστής συσχέτισης ρ μπορεί αρχικά να θεωρηθεί μηδενικός. Η τιμή του συντελεστή αναλογίας K εξαρτάται από την μέγιστη τιμή φωτεινότητας που μπορούν να πάρουν τα εικονοστοιχεία της εικόνας (πχ. $K=255$ για εικόνες με 256 διαβαθμίσεις χρωμάτων). Μετά τη σύγκλιση του αλγορίθμου προσαρμογής, η μέση τιμή της τελικής Gaussian κατανομής εκλαμβάνεται ως το ζητούμενο κέντρο \mathbf{k} της εξεταζόμενης κουκίδας.

Όλες οι παραπάνω μέθοδοι υπολογισμού των κέντρων των κουκίδων λειτουργούν με ακρίβεια μεγαλύτερη από τις διαστάσεις ενός εικονοστοιχείου της ψηφιακής εικόνας (sub-pixel accuracy). Τα πλεονεκτήματα κάποιων μεθόδων έναντι κάποιων άλλων εξαρτώνται από τις διαστάσεις της εξεταζόμενης κουκίδας πάνω στην ψηφιακή εικόνα, καθώς και από τα ποσοστά θορύβου στην εικόνα που προκαλούνται από εξωγενείς παράγοντες (ανομοιόμορφος φωτισμός, χαμηλή ποιότητα εικόνας, τυχαίες σκιές, κτλ.). Οι μέθοδοι προσαρμογής (ε) και (στ) είναι συνήθως προτιμότερες εάν οι διαστάσεις των κουκίδων είναι αρκετά μεγάλες ώστε το περίγραμμά τους να έχει αρκετά μεγάλη ανάλυση. Επίσης είναι πιο ανεκτικές στον θόρυβο της εικόνας από τις μεθόδους (α)-(δ),

οι οποίες συνήθως προτιμώνται είτε για την ανίχνευση κουκίδων μικρών διαστάσεων (μόλις μερικών εικονοστοιχείων) είτε σε περιπτώσεις που οι ψηφιακές εικόνες είναι πολύ καλής ποιότητας και ανάλυσης. Σε αντίθετη περίπτωση χρησιμοποιούνται μόνο για να δώσουν μια αρχική εκτίμηση των κέντρων, οι οποίες χρησιμοποιούνται σαν αρχικές τιμές στους επαναληπτικούς αλγορίθμους προσαρμογής (ϵ) και (σ).

2.5 Ανακατασκευή τρισδιάστατων σημείων

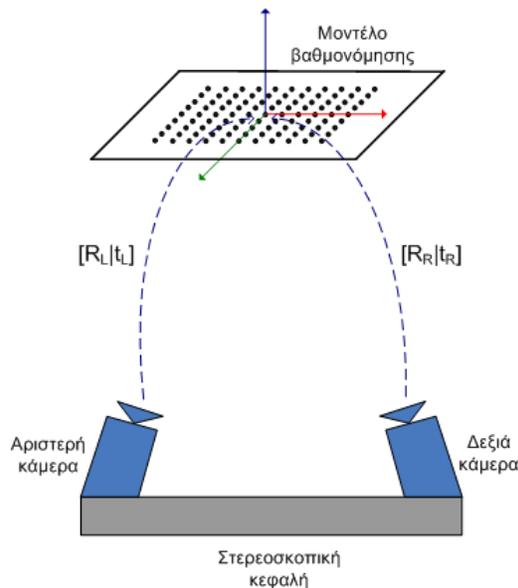
Στο Κεφ. 2.2 γίνεται μια προσπάθεια ανάλυσης του συνηθέστερα χρησιμοποιούμενου μοντέλου προβολής ενός παρατηρούμενου σημείου που ανήκει στον πραγματικό τρισδιάστατο χώρο πάνω στην φωτοευαίσθητη επιφάνεια μιας ψηφιακής κάμερας (camera pin-hole model). Η βασική εξίσωση που συνδέει τις 3D συντεταγμένες (X, Y, Z) του παρατηρούμενου σημείου με τις αντίστοιχες συντεταγμένες εικόνας (u, v) της προβολής του πάνω στο πλάνο εικόνας της κάμερας (φωτοευαίσθητο ολοκληρωμένο CMOS) είναι η εξής:

$$z \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \mathbf{A} [\mathbf{R} | \mathbf{t}] \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (2.65)$$

όπου \mathbf{A} και $[\mathbf{R} | \mathbf{t}]$ είναι οι πίνακες εσωτερικών κι εξωτερικών παραμέτρων αντίστοιχα, και z είναι μια (άγνωστη) σταθερά που περιγράφει την προβολή του παρατηρούμενου σημείου πάνω στον άξονα z του συστήματος συντεταγμένων της κάμερας (βλ. Κεφ. 2.2). Όπως αναλύθηκε στο Κεφ. 2.3, υπάρχουν διάφορες τεχνικές και αλγόριθμοι για τον υπολογισμό των αγνώστων πινάκων εσωτερικών-εξωτερικών παραμέτρων, οι οποίοι στο εξής θα θεωρούνται γνωστοί.

Στην περίπτωση που οι συντεταγμένες (X, Y, Z) του παρατηρούμενου σημείου είναι γνωστές τότε βάσει της (2.65) μπορούμε μονοσήμαντα να υπολογίσουμε τις συντεταγμένες εικόνας της προβολής του (u, v), καθώς και τη σταθερά z . Σε πρακτικές εφαρμογές συνήθως μας ενδιαφέρει η αντίστροφη διαδικασία, δηλαδή ο υπολογισμός των 3D συντεταγμένων ενός παρατηρούμενου σημείου βάσει των γνωστών συντεταγμένων της προβολής του πάνω στην ψηφιακή εικόνα. Η διαδικασία αυτή

Ο συνδυασμός των (2.67) και (2.68), εκτός από εξισώσεις 3D ευθείας, αποτελεί επίσης ένα σύστημα δύο γραμμικών εξισώσεων με αγνώστους τις πραγματικές 3D συντεταγμένες (X, Y, Z) του παρατηρούμενου σημείου. Προφανώς οι εξισώσεις αυτές δεν είναι αρκετές για την πλήρη ανακατασκευή του ζητούμενου σημείου, δηλαδή τον υπολογισμό των 3D συντεταγμένων του βάσει των γνωστών συντεταγμένων εικόνας (u, v) της προβολής του. Το γεγονός αυτό μας οδηγεί στο συμπέρασμα ότι μια ψηφιακή εικόνα δεν είναι αρκετή για την πλήρη ανακατασκευή ενός παρατηρούμενου σημείου, παρά μόνο εάν η μία συντεταγμένη του είναι εκ των προτέρων γνωστή. Έτσι εάν γνωρίζουμε ότι το παρατηρούμενο σημείο ανήκει σε κάποιο γνωστό επίπεδο στον τρισδιάστατο χώρο (πχ. $Z=0$) τότε είναι δυνατός ο μονοσήμαντος προσδιορισμός των υπόλοιπων συντεταγμένων του βάσει των (2.67) και (2.68).



Σχήμα 2.9 Βαθμονόμηση στερεοσκοπικής κεφαλής καμερών

Μία μέθοδος επίλυσης του παραπάνω προβλήματος είναι η χρήση μιας δεύτερης κάμερας, η οποία σε συνδυασμό με την υπάρχουσα υλοποιούν μία *στερεοσκοπική κεφαλή (stereo head of cameras)*. Η γεωμετρία της στερεοσκοπικής κεφαλής είναι σταθερή, πράγμα που σημαίνει ότι η σχετική θέση των δύο καμερών δεν μπορεί να μεταβληθεί. Οι δύο κάμερες βαθμονομούνται ταυτόχρονα χρησιμοποιώντας εικόνες από το ίδιο μοντέλο

βαθμονόμησης, με αποτέλεσμα το σύστημα συντεταγμένων μέτρησης-βαθμονόμησης να είναι κοινό και για τις δύο κάμερες και να

ορίζεται από τους πίνακες εξωτερικών παραμέτρων τους, όπως φαίνεται και στο Σχ. 2.9.

Ας υποθέσουμε ότι \mathbf{H}_L και \mathbf{H}_R είναι οι πίνακες \mathbf{H} (projection matrices ή homographies) της αριστερής και της δεξιάς κάμερας αντίστοιχα, και οι συντεταγμένες εικόνας των προβολών ενός κοινά παρατηρούμενου σημείου με 3D συντεταγμένες (X, Y, Z) πάνω στις ψηφιακές εικόνες της αντίστοιχης κάμερας είναι (u_L, v_L) και (u_R, v_R). Χρησιμοποιώντας τις (2.65) και (2.66) για κάθε κάμερα ξεχωριστά, μπορούμε να καταλήξουμε σε δύο ζευγάρια εξισώσεων της μορφής των (2.67) και (2.68):

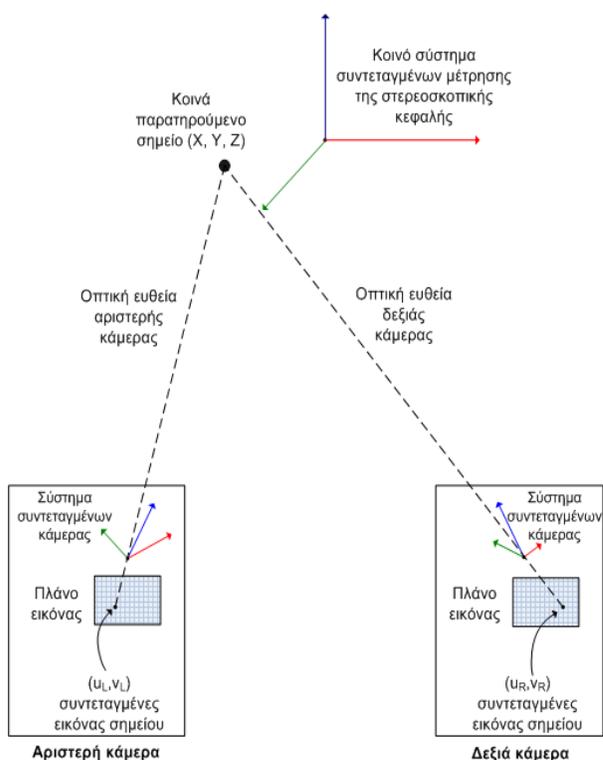
$$(h_{L(11)} - u_L h_{L(31)})X + (h_{L(12)} - u_L h_{L(32)})Y + (h_{L(13)} - u_L h_{L(33)})Z = u_L h_{L(34)} - h_{L(14)} \quad (2.69)$$

$$(h_{L(21)} - v_L h_{L(31)})X + (h_{L(22)} - v_L h_{L(32)})Y + (h_{L(23)} - v_L h_{L(33)})Z = v_L h_{L(34)} - h_{L(24)} \quad (2.70)$$

$$(h_{R(11)} - u_R h_{R(31)})X + (h_{R(12)} - u_R h_{R(32)})Y + (h_{R(13)} - u_R h_{R(33)})Z = u_R h_{R(34)} - h_{R(14)} \quad (2.71)$$

$$(h_{R(21)} - v_R h_{R(31)})X + (h_{R(22)} - v_R h_{R(32)})Y + (h_{R(23)} - v_R h_{R(33)})Z = v_R h_{R(34)} - h_{R(24)} \quad (2.72)$$

Οι ζητούμενες συντεταγμένες (X, Y, Z) του παρατηρούμενου σημείου είναι κοινές και για τις δύο κάμερες, καθώς αυτές αναφέρονται σε ένα κοινό σύστημα συντεταγμένων μέτρησης, ήτοι το σύστημα συντεταγμένων μέτρησης της στερεοσκοπικής κεφαλής. Οι παραπάνω εξισώσεις αποτελούν ένα σύστημα τεσσάρων γραμμικών εξισώσεων με τρεις αγνώστους, του οποίου η βέλτιστη λύση, υπό την έννοια ελαχίστων τετραγώνων, μπορεί να βρεθεί με τη χρήση της ψευδο-αντίστροφης ή του αλγορίθμου διάσπασης ιδιόμορφης τιμής (*Singular Value Decomposition, SVD*) [Press et al., 1988; Golub and Van Loan, 1983].

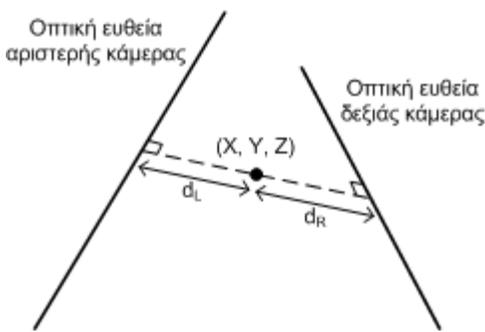


Σχήμα 2.10 Γεωμετρική ερμηνεία της 3D ανακατασκευής σημείου

Μία γεωμετρική ερμηνεία της παραπάνω προσέγγισης μπορεί να δοθεί με τη βοήθεια του Σχ. 2.10, όπου απεικονίζεται μία στερεοσκοπική κεφαλή καμερών η οποία παρατηρεί ένα αυθαίρετο χαρακτηριστικό σημείο στον τρισδιάστατο χώρο (πχ. κουκίδα, κύκλο, ακμή, κτλ.). Με τη βοήθεια κατάλληλων αλγορίθμων επεξεργασίας εικόνας, μερικοί από τους οποίους παρουσιάστηκαν στο Κεφ. 2.4, μπορούμε να ανιχνεύσουμε με ακρίβεια τις συντεταγμένες εικόνας των προβολών του σημείου (u_L, v_L) και (u_R, v_R) πάνω στις ψηφιακές εικόνες της αριστερής και της δεξιάς κάμερας αντίστοιχα. Οι οπτικές ευθείες που

σχηματίζονται από το σημείο προβολής στο πλάνο της εκάστοτε εικόνας και από το κέντρο του συστήματος συντεταγμένων της εκάστοτε κάμερας, περιγράφονται από τις

εξισώσεις (2.69 & 2.70) και (2.71 & 2.72) για την αριστερή και την δεξιά κάμερα αντίστοιχα, και αναφέρονται ως προς το κοινό σύστημα συντεταγμένων μέτρησης της στερεοσκοπικής κεφαλής. Λύνοντας το γραμμικό σύστημα των τεσσάρων εξισώσεων ως προς τους τρεις αγνώστους (X, Y, Z) ουσιαστικά βρίσκουμε το σημείο τομής των δύο ευθειών στον τρισδιάστατο χώρο.



Σχήμα 2.11 Οπτικές ευθείες και βέλτιστη ανακατασκευής σημείου

Σε πρακτικές εφαρμογές, η πεπερασμένη ακρίβεια στον υπολογισμό των συντεταγμένων εικόνας (u_L, v_L) και (u_R, v_R) από τους αλγορίθμους ψηφιακής επεξεργασίας εικόνας, καθώς και των πινάκων εσωτερικών-εξωτερικών παραμέτρων των καμερών κατά τη διαδικασία της βαθμονόμησής τους, έχουν σαν αποτέλεσμα οι ευθείες των εξισώσεων (2.69-2.71) να μην τέμνονται ιδανικά στο ζητούμενο σημείο $(X, Y,$

$Z)$. Μαθηματικά αυτό μεταφράζεται στο ότι το σύστημα των γραμμικών εξισώσεων δεν έχει πραγματική λύση, γεγονός που είναι αναμενόμενο καθώς στην γενική περίπτωση ένα γραμμικό σύστημα με παραπάνω εξισώσεις από αγνώστους είναι αδύνατο να λυθεί. Στην πράξη, μας ενδιαφέρει να βρούμε ένα σημείο στον τρισδιάστατο χώρο του οποίου η απόσταση και από τις δύο ευθείες να είναι όσο το δυνατόν μικρότερη, υπό την έννοια των ελαχίστων τετραγώνων. Αυτή ακριβώς είναι και η λειτουργία του αλγορίθμου της διάσπασης ιδιόμορφης τιμής, ο οποίος χρησιμοποιείται για την εύρεση της βέλτιστης λύσης για τους αγνώστους (X, Y, Z) . Εναλλακτικά μπορεί να χρησιμοποιηθεί η ψευδο-αντίστροφη του πίνακα των συντελεστών του γραμμικού συστήματος, με παρόμοια αποτελέσματα. Η μέση τετραγωνική απόσταση της βέλτιστης λύσης για τις συντεταγμένες (X, Y, Z) από τις οπτικές ευθείες των δύο καμερών συνήθως χρησιμοποιείται σαν μέτρο ποιότητας της 3D ανακατασκευής. Όσο η τομή των δύο ευθειών πλησιάζει την τέλεια τομή, δηλαδή όσο η μέση τετραγωνική απόσταση της βέλτιστης λύσης για το ζητούμενο σημείο από τις δύο οπτικές ευθείες τείνει προς το μηδέν, τόσο η ακρίβεια της 3D ανακατασκευής αυξάνεται. Προφανώς η ακρίβεια της ανακατασκευής εξαρτάται από την ακρίβεια της ανίχνευσης των συντεταγμένων εικόνας (u_L, v_L) και (u_R, v_R) των προβολών του παρατηρούμενου σημείου, καθώς και από την

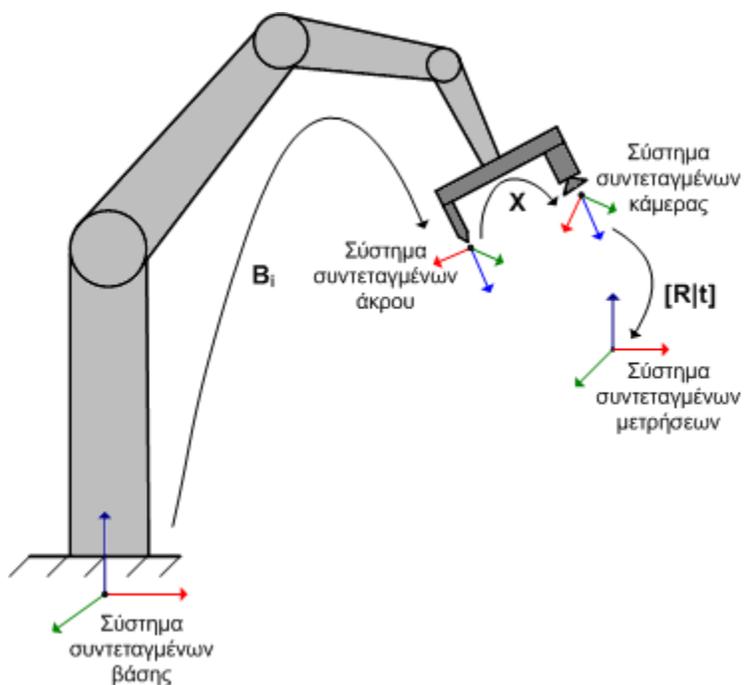
ακρίβεια των πινάκων \mathbf{H}_L και \mathbf{H}_R που υπολογίζονται κατά την διαδικασία βαθμονόμησης της στερεοσκοπικής κεφαλής. Αν για παράδειγμα η προβολή του σημείου για μία τουλάχιστο κάμερα έχει υπολογιστεί σε λάθος θέση, τότε οι οπτικές ευθείες του Σχ. 2.10 θα απέχουν πολύ απ' το να τέμνονται σ' ένα κοινό σημείο, και η μέση τετραγωνική απόσταση του βέλτιστου υπολογισθέντος σημείου από τις δύο ευθείες θα είναι πολύ μεγάλη. Ορίζοντας μία αποδεκτή τιμή κατωφλίου για την μέση τετραγωνική απόσταση μπορούμε είτε ν' αποδεχθούμε την 3D ανακατασκευή ενός παρατηρούμενου σημείου είτε να την απορρίψουμε ως ανακριβή ή λανθασμένη.

3 ΒΑΘΜΟΝΟΜΗΣΗ ΒΡΑΧΙΟΝΑ – ΜΗΧΑΝΙΚΗΣ ΟΡΑΣΗΣ

3.1 Σημασία συσχέτισης συντεταγμένων κάμερας και βραχίονα

Υπάρχουν πολλές πρακτικές εφαρμογές όπου απαιτείται το άκρο ενός βιομηχανικού ρομποτικού βραχίονα να ακολουθεί τις κινήσεις ενός κινούμενου σώματος ή σημείου που βρίσκεται στον περιβάλλοντα χώρο εργασίας, με βάση την πληροφορία ανάδρασης που προέρχεται από τις μετρήσεις μίας κάμερας ή στερεοσκοπικής κεφαλής (stereo head) η οποία είναι σταθερά στερεωμένη στο άκρο του βραχίονα. Οι εφαρμογές αυτές στην διεθνή βιβλιογραφία κατηγοριοποιούνται κάτω από μία κοινή ονομασία, ήτοι *εφαρμογές οπτικής παρακολούθησης (Visual Servoing Applications)*. Στο Σχ. 3.1 απεικονίζεται ένας βιομηχανικός ρομποτικός βραχίονας με μια κάμερα σταθερά στερεωμένη στο άκρο του, όλα τα συστήματα συντεταγμένων που εμπλέκονται στην τοπολογία, καθώς και οι μετασχηματισμοί συντεταγμένων που τα συνδέουν μεταξύ τους. Στην εισαγωγή της παρούσας διατριβής έγινε αναφορά στην σημασία και αναγκαιότητα υπολογισμού του μετασχηματισμού συντεταγμένων X που σχετίζει το σύστημα συντεταγμένων του άκρου του βραχίονα με το σύστημα συντεταγμένων του σένσορα (κάμερα ή stereo head).

Αναφερόμενοι στο Σχ. 3.1, το σύστημα συντεταγμένων βάσης του ρομπότ βρίσκεται στο σημείο στήριξης του βραχίονα πάνω σε μια σταθερή επιφάνεια (πχ. πάτωμα), ενώ η θέση και ο προσανατολισμός του είναι σταθερά και δεν μεταβάλλονται με την κίνηση του ρομπότ. Το σύστημα συντεταγμένων του άκρου ταυτίζεται με το άκρο του τελευταίου συνδέσμου του βραχίονα, ενώ η θέση και ο προσανατολισμός του ακολουθούν την εκάστοτε κίνησή του. Ο μετασχηματισμός συντεταγμένων B_i που σχετίζει το κινούμενο σύστημα συντεταγμένων του άκρου με το σταθερό σύστημα συντεταγμένων της βάσης είναι αποτέλεσμα της λύσης του *ευθέως κινηματικού προβλήματος (direct kinematics problem)* του βραχίονα [Fu et al., 1987; Τζαφέστας, 2003] και δίνεται απ' ευθείας από τον ρομποτικό ελεγκτή (robot controller).



Σχήμα 3.1 Συστήματα συντεταγμένων ρομποτικού βραχίονα και κάμερας

Δυστυχώς, οι ανακρίβειες στις μεταβλητές που σχετίζονται με την ρομποτική αλυσίδα, όπως για παράδειγμα οι διαστάσεις των συνδέσμων και οι γωνίες θέσης των αρθρώσεων, καθώς και ένα πλήθος εξωγενών παραγόντων όπως οι μεταβολές στη θερμοκρασία και η κάμψη των συνδέσμων λόγω βάρους στήριξης, εμποδίζουν τον ακριβή υπολογισμό του μετασχηματισμού B_i , ιδιαίτερα όσον αφορά τις μεταβλητές μετατόπισης. Το σύστημα συντεταγμένων του άκρου του βραχίονα συνδέεται με το σύστημα συντεταγμένων του κέντρου της κάμερας (pin-hole center) μ' έναν άγνωστο προς το παρόν μετασχηματισμό X , ο οποίος είναι ανεξάρτητος της θέσης του βραχίονα καθώς η κάμερα είναι σταθερά στερεωμένη πάνω στο άκρο του. Τέλος το σύστημα συντεταγμένων των μετρήσεων της κάμερας σχετίζεται με το σύστημα συντεταγμένων του κέντρου της με τον μετασχηματισμό εξωτερικών παραμέτρων της κάμερας $[R|t]$, ο οποίος έχει υπολογιστεί κατά τη διαδικασία της βαθμονόμησής της (βλ. Κεφ. 2.3). Όπως αναφέρθηκε και στο προηγούμενο κεφάλαιο, όλες οι μετρήσεις που πραγματοποιεί η κάμερα (τρισδιάστατη ανακατασκευή παρατηρούμενου σημείου) αναφέρονται στο σύστημα συντεταγμένων μέτρησης, το οποίο ορίζεται κατά την διαδικασία της βαθμονόμησής της. Το σύστημα συντεταγμένων μέτρησης είναι σταθερά προσαρτημένο πάνω στην κάμερα σύμφωνα με

τον πίνακα (μετασχηματισμό) εξωτερικών παραμέτρων $[\mathbf{R}|\mathbf{t}]$ και συνεπώς ακολουθεί πάντα τις κινήσεις της.

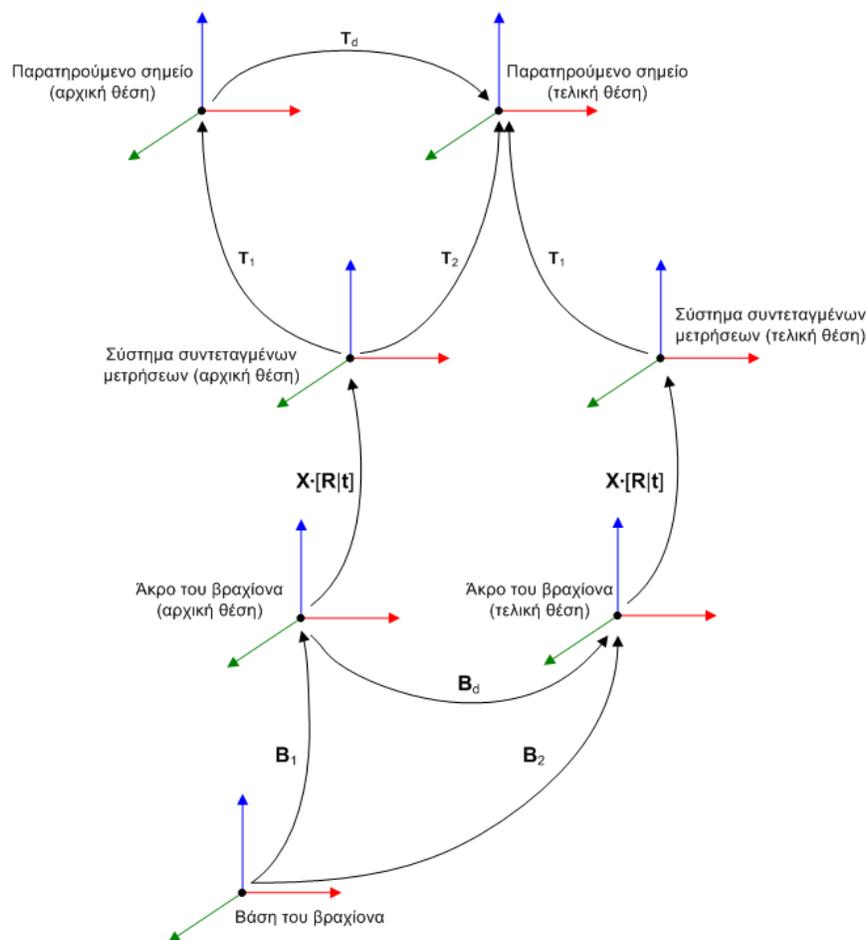
Η σημασία υπολογισμού του άγνωστου μετασχηματισμού \mathbf{X} είναι πολύ μεγάλη για τις περισσότερες πρακτικές ρομποτικές εφαρμογές, καθώς μας επιτρέπει να αναγάγουμε τις τρισδιάστατες μετρήσεις μίας κάμερας ή μιας στερεοσκοπικής κεφαλής σ' ένα σύστημα συντεταγμένων που «γνωρίζει» ο ρομποτικός βραχίονας. Ας υποθέσουμε για παράδειγμα ότι η κάμερα παρατηρεί ένα κινούμενο σημείο, του οποίου η θέση σε μια αρχική και μια τυχαία χρονική στιγμή είναι $\mathbf{t}_1=[X_1, Y_1, Z_1]^T$ και $\mathbf{t}_2=[X_2, Y_2, Z_2]^T$ αντίστοιχα. Ο υπολογισμός των συντεταγμένων θέσης του κινούμενου σημείου είναι εφικτός με την τρισδιάστατη ανακατασκευή του στις δύο χρονικές στιγμές, αναφερόμενοι πάντα στο σύστημα συντεταγμένων μέτρησης της κάμερας. Η αρχική και η τελική θέση του κινούμενου σημείου, καθώς και η μεταβολή της, μπορούν να εκφραστούν ως μετασχηματισμοί συντεταγμένων ως προς το σύστημα συντεταγμένων μέτρησης, των οποίων οι πίνακες στροφής θα συμπίπτουν με τον μοναδιαίο 3×3 πίνακα, ήτοι:

$$\mathbf{T}_1 = \begin{bmatrix} \mathbf{I}_{3 \times 3} & \mathbf{t}_1 \\ \mathbf{0}_{1 \times 3} & 1 \end{bmatrix} \quad \mathbf{T}_2 = \begin{bmatrix} \mathbf{I}_{3 \times 3} & \mathbf{t}_2 \\ \mathbf{0}_{1 \times 3} & 1 \end{bmatrix} \quad \mathbf{T}_d = \begin{bmatrix} \mathbf{I}_{3 \times 3} & d\mathbf{t} \\ \mathbf{0}_{1 \times 3} & 1 \end{bmatrix} \quad (3.1)$$

Σκοπός μας είναι να κινήσουμε το άκρο του βραχίονα κατά τέτοιο τρόπο ώστε στην τελική του θέση η κάμερα να «βλέπει» το κινούμενο σημείο με τον ίδιο ακριβώς τρόπο όπως στην αρχική θέση του βραχίονα. Η διαδικασία αυτή γίνεται καλύτερα κατανοητή με τη βοήθεια του Σχ. 3.2, όπου απεικονίζονται όλα συστήματα συντεταγμένων που εμπλέκονται στο πείραμα, καθώς και οι σχετικοί μετασχηματισμοί που τα συνδέουν μεταξύ τους. Το άκρο του βραχίονα βρίσκεται σε μια αρχική θέση \mathbf{B}_1 σε σχέση με το σταθερό σύστημα συντεταγμένων της βάσης, ενώ το σύστημα συντεταγμένων μέτρησης της κάμερας συνδέεται με το άκρο του βραχίονα με τον μετασχηματισμό $\mathbf{X} \cdot [\mathbf{R}|\mathbf{t}]$ (βλ. επίσης Σχ. 3.1). Ας υποθέσουμε ότι η κάμερα μετράει τις συντεταγμένες ενός παρατηρούμενου σημείου ως \mathbf{T}_1 , αναφερόμενες στο σύστημα συντεταγμένων μέτρησης. Εάν το σημείο μετακινηθεί στο χώρο τότε την επόμενη χρονική στιγμή η κάμερα θα μετρήσει τις νέες του συντεταγμένες ως \mathbf{T}_2 . Η μεταβολή αυτή των μετρήσεων της κάμερας θα πρέπει να αναχθεί σε ένα σύστημα συντεταγμένων που είναι γνωστό στον ρομποτικό βραχίονα, όπως είναι το σύστημα συντεταγμένων του άκρου του. Αν η κίνηση

του βραχίονα εκτελεστεί σωστά, τότε η κάμερα στην νέα της θέση θα πρέπει να παρατηρεί το σημείο με τον ίδιο ακριβώς τρόπο όπως και στην αρχική της θέση. Το γεγονός αυτό μεταφράζεται μαθηματικά ότι στην νέα της θέση η κάμερα θα πρέπει να μετράει τις συντεταγμένες του παρατηρούμενου σημείου ως T_1 , όπως ακριβώς και στην αρχική της θέση. Σύμφωνα με το Σχ. 3.2, ο μετασχηματισμός συντεταγμένων B_d που αντιστοιχεί στην κίνηση που πρέπει να εκτελέσει το άκρο του βραχίονα για να ακολουθήσει τη μεταβολή θέσης του κινούμενου σημείου είναι η εξής:

$$B_d = (X[R|t])T_1T_dT_1^{-1}(X[R|t])^{-1} = (X[R|t])T_2T_1^{-1}(X[R|t])^{-1} \quad (3.2)$$

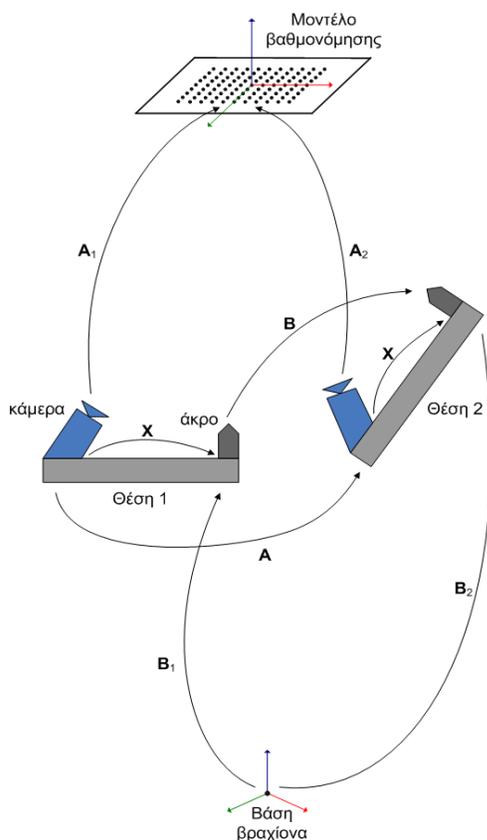


Σχήμα 3.2 Παρακολούθηση τροχιάς κινούμενου σημείου

Η μέθοδος παρακολούθησης που περιγράφεται παραπάνω είναι αδύνατη δίχως τον εκ των προτέρων υπολογισμό του άγνωστου μετασχηματισμού X , ο οποίος στην γενική

περίπτωση συνδέει τη θέση του χρησιμοποιούμενου οργάνου μέτρησης με το άκρο του βραχίονα, πάνω στο οποίο είναι σταθερά στερεωμένο. Στα επόμενα κεφάλαια είναι συγκεντρωμένες οι συνηθέστερα χρησιμοποιούμενες μέθοδοι υπολογισμού του άγνωστου μετασχηματισμού X , στην περίπτωση που το χρησιμοποιούμενο όργανο μέτρησης είναι μια μονή κάμερα, η οποία είναι σταθερά στερεωμένη στο άκρο του βραχίονα. Στη συνέχεια θα παρουσιαστεί μια καινούρια μέθοδος που μας επιτρέπει να επεκτείνουμε τις υπάρχουσες μεθοδολογίες για την περίπτωση χρήσης στερεοσκοπικής κεφαλής, και μάλιστα επιτυγχάνοντας πολύ μεγαλύτερη ακρίβεια στον υπολογισμό του μετασχηματισμού X σε σχέση με την περίπτωση της μονής κάμερας.

3.2 Βαθμονόμηση βραχίονα – μηχανικής όρασης για μονή κάμερα



Σχήμα 3.3 Τοπολογία πειράματος βαθμονόμησης βραχίονα – μηχανικής όρασης για μονή κάμερα

Στο παρόν κεφάλαιο γίνεται μια προσπάθεια ανάλυσης της διαδικασίας που ακολουθείται για τον συσχετισμό των συστημάτων συντεταγμένων του άκρου ενός βιομηχανικού ρομποτικού βραχίονα και μιας ψηφιακής κάμερας η οποία είναι σταθερά στερεωμένη πάνω του. Επίσης γίνεται μια ανάλυση των συνηθέστερα χρησιμοποιούμενων μαθηματικών αλγορίθμων που επιστρατεύονται για την επίλυση του προβλήματος, καθώς και των πλεονεκτημάτων και μειονεκτημάτων που τους χαρακτηρίζουν.

Στο Σχ. 3.3 φαίνεται η τοπολογία του πειράματος που χρησιμοποιείται για τον υπολογισμό του άγνωστου μετασχηματισμού X , που συνδέει το σύστημα συντεταγμένων του κέντρου της κάμερας (pin-hole center) με το σύστημα συντεταγμένων του άκρου του βραχίονα. Ο συνδυασμός της κάμερας και του

άκρου του βραχίονα αποτελούν ένα στέρεο σώμα, καθώς η κάμερα είναι σταθερά στερεωμένη πάνω του, το οποίο συνήθως ονομάζεται *συστοιχία κάμερας-άκρου* (*Hand-Eye device*). Η κάμερα παρακολουθεί ένα επίπεδο μοντέλο βαθμονόμησης με τη βοήθεια του οποίου μπορεί να βαθμονομηθεί σε οποιαδήποτε αυθαίρετη θέση. Ας υποθέσουμε ότι για δύο διαφορετικές θέσεις της συστοιχίας κάμερας-άκρου, ήτοι τις θέσεις 1 και 2 του Σχ. 3.3, η θέση και ο προσανατολισμός του άκρου του βραχίονα, σε σχέση με το σταθερό σύστημα συντεταγμένων της βάσης του ρομπότ, δίνονται από τους μετασχηματισμούς \mathbf{B}_1 και \mathbf{B}_2 αντίστοιχα, ενώ η θέση και ο προσανατολισμός του σταθερού συστήματος συντεταγμένων του μοντέλου βαθμονόμησης σε σχέση με το σύστημα συντεταγμένων του κέντρου της κάμερας είναι \mathbf{A}_1 και \mathbf{A}_2 αντίστοιχα. Οι μετασχηματισμοί \mathbf{B}_1 και \mathbf{B}_2 δίνονται απ' ευθείας απ' τον ρομποτικό ελεγκτή (robot controller) ως λύση του ευθέως κινηματικού προβλήματος του βραχίονα, ενώ οι μετασχηματισμοί \mathbf{A}_1 και \mathbf{A}_2 αποτελούν τις εξωτερικές παραμέτρους της κάμερας, προερχόμενες από την εφαρμογή του αλγορίθμου βαθμονόμησης στις δύο διαφορετικές θέσεις 1 και 2, και συνδέουν το κινούμενο σύστημα συντεταγμένων της κάμερας (pin-hole center) με το σταθερό σύστημα συντεταγμένων του μοντέλου βαθμονόμησης. Απ' ευθείας από το Σχ. 3.3 παράγεται η βασική ομογενής εξίσωση μετασχηματισμών για τη βαθμονόμηση βραχίονα – μηχανικής όρασης:

$$\mathbf{AX} = \mathbf{XB} \quad (3.3)$$

όπου $\mathbf{A} = \mathbf{A}_1\mathbf{A}_2^{-1}$ και $\mathbf{B} = \mathbf{B}_1^{-1}\mathbf{B}_2$ είναι οι μετασχηματισμοί που εκφράζουν τις μετατοπίσεις του κέντρου της κάμερας και του άκρου του βραχίονα αντίστοιχα, από τη θέση 1 στην θέση 2. Προφανώς οι μετασχηματισμοί \mathbf{A} και \mathbf{B} εξαρτώνται από την κίνηση της συστοιχίας κάμερας-άκρου, ενώ ο μετασχηματισμός \mathbf{X} είναι σταθερός κι εξαρτάται μόνο από τη σχετική θέση της κάμερας και του άκρου του βραχίονα.

Η ομογενής εξίσωση (3.3) μπορεί να διασπαστεί σε δύο επιμέρους εξισώσεις, μία για τους 3×3 πίνακες στροφής \mathbf{R}_A , \mathbf{R}_B και \mathbf{R}_X , και μία για τα 3×1 διανύσματα μετατόπισης \mathbf{t}_A , \mathbf{t}_B και \mathbf{t}_X , ως εξής:

$$\begin{bmatrix} \mathbf{R}_A & \mathbf{t}_A \\ \mathbf{0} & 1 \end{bmatrix} \begin{bmatrix} \mathbf{R}_X & \mathbf{t}_X \\ \mathbf{0} & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{R}_X & \mathbf{t}_X \\ \mathbf{0} & 1 \end{bmatrix} \begin{bmatrix} \mathbf{R}_B & \mathbf{t}_B \\ \mathbf{0} & 1 \end{bmatrix} \Leftrightarrow$$

$$\mathbf{R}_A\mathbf{R}_X = \mathbf{R}_X\mathbf{R}_B \quad (3.4)$$

$$(\mathbf{R}_A - \mathbf{I})\mathbf{t}_X = \mathbf{R}_X \mathbf{t}_B - \mathbf{t}_A \quad (3.5)$$

Είναι προφανές ότι αν ο πίνακας στροφής \mathbf{R}_X γίνει γνωστός με την επίλυση της εξίσωσης (3.4), τότε ο υπολογισμός του διανύσματος μετατόπισης \mathbf{t}_X ανάγεται στην επίλυση ενός συστήματος γραμμικών εξισώσεων, σύμφωνα με την εξίσωση (3.5).

Υπάρχουν αρκετές μέθοδοι για την επίλυση της (3.4) [Horaud and Dornaika, 1995; Andreff et al., 2001; Tsai and Lenz, 1989; Park and Martin, 1994], ο πιο δημοφιλής από τους οποίους θα παρουσιαστεί στην συνέχεια [Horaud and Dornaika, 1995]. Είναι γνωστό ότι όλοι οι πίνακες στροφής έχουν τουλάχιστο μία μοναδιαία ιδιοτιμή. Ας υποθέσουμε ότι \mathbf{n}_a , \mathbf{n}_b και \mathbf{n}_x είναι τα 3×1 ιδιοδιανύσματα των 3×3 πινάκων στροφής \mathbf{R}_A , \mathbf{R}_B και \mathbf{R}_X αντίστοιχα, τα οποία αντιστοιχούν στις μοναδιαίες ιδιοτιμές τους. Πολλαπλασιάζοντας την (3.4) από τα δεξιά με το ιδιοδιάνυσμα \mathbf{n}_b έχουμε:

$$\mathbf{R}_A \mathbf{R}_X \mathbf{n}_b = \mathbf{R}_X \mathbf{R}_B \mathbf{n}_b \stackrel{\lambda=1}{\Leftrightarrow} \mathbf{R}_A \mathbf{R}_X \mathbf{n}_b = \mathbf{R}_X \mathbf{n}_b \quad (3.6)$$

Από την (3.6) φαίνεται ότι το 3×1 διάνυσμα $(\mathbf{R}_X \cdot \mathbf{n}_b)$ αποτελεί το ιδιοδιάνυσμα του πίνακα στροφής \mathbf{R}_A , που αντιστοιχεί στην μοναδιαία ιδιοτιμή του, οπότε ισχύει:

$$\mathbf{n}_a = \mathbf{R}_X \mathbf{n}_b \quad (3.7)$$

Η εξίσωση (3.7) αντιστοιχεί σε μία μετατόπιση της συστοιχίας κάμερας-άκρου, ήτοι από την θέση 1 στην θέση 2 (Σχ. 3.3). Στην γενική περίπτωση που έχουμε n στον αριθμό μετατοπίσεις της συστοιχίας ($n+1$ διακριτές θέσεις) τότε θα λάβουμε n εξισώσεις της μορφής της (3.7), με διαφορετικές τιμές για τα ιδιοδιανύσματα \mathbf{n}_a και \mathbf{n}_b αλλά κοινή τιμή για τον άγνωστο πίνακα στροφής \mathbf{R}_X .

$$(i=1,2,\dots,n) \quad \mathbf{n}_a(i) = \mathbf{R}_X \mathbf{n}_b(i) \quad (3.8)$$

Στην παραπάνω εξίσωση τα n ιδιοδιανύσματα $\mathbf{n}_a(i)$ και $\mathbf{n}_b(i)$, τα οποία αντιστοιχούν στις μοναδιαίες ιδιοτιμές, μπορούν να βρεθούν από τους γνωστούς πίνακες στροφής $\mathbf{R}_A(i)$ και $\mathbf{R}_B(i)$, οπότε η μόνη άγνωστη μεταβλητή παραμένει ο πίνακας στροφής \mathbf{R}_X . Η επίλυση της εξίσωσης (3.8) ανάγεται στην ελαχιστοποίηση της παρακάτω μη-γραμμικής συνάρτησης κόστους:

$$f_1(\mathbf{R}_X) = \sum_{i=1}^n \|\mathbf{n}_a(i) - \mathbf{R}_X \mathbf{n}_b(i)\|^2 \quad (3.9)$$

Για την ελαχιστοποίηση της συνάρτησης (3.9) ως προς \mathbf{R}_X υπάρχει αναλυτική μέθοδος κλειστής μορφής (closed-form), η οποία μας αποτρέπει απ' το να χρησιμοποιήσουμε μη-γραμμικές μεθόδους ελαχιστοποίησης, οι οποίες είναι χρονοβόρες κι απαιτούν καλή αρχική εκτίμηση. Η μεθοδολογία υπολογισμού του πίνακα \mathbf{R}_X που ελαχιστοποιεί την συνάρτηση (3.9) περιγράφεται στη βιβλιογραφία [Horn, 1987; Horaud and Dornaika, 1995] και χρησιμοποιεί τις μεταβλητές *Unit Quaternions* [Kravchenko, 2003] για να περιγράψει τους πίνακες στροφής.

Εφόσον ο πίνακας στροφής \mathbf{R}_X έχει γίνει γνωστός, ο υπολογισμός του άγνωστου διανύσματος μετατόπισης \mathbf{t}_X γίνεται απλή υπόθεση. Επεκτείνοντας την εξίσωση (3.5) για n κινήσεις της συστοιχίας μπορούμε να δημιουργήσουμε την παρακάτω γραμμική (ως προς \mathbf{t}_X) συνάρτηση κόστους:

$$(i = 1, 2, \dots, n) \quad f_2(\mathbf{t}_X) = \sum_{i=1}^n \|\mathbf{R}_X \mathbf{t}_B(i) - (\mathbf{R}_A(i) - \mathbf{I})\mathbf{t}_X - \mathbf{t}_A(i)\|^2, \quad (3.10)$$

την οποία θέλουμε να ελαχιστοποιήσουμε. Εφόσον η παραπάνω συνάρτηση κόστους είναι γραμμική, μπορούμε να χρησιμοποιήσουμε κλασσικές μεθόδους ελαχίστων τετραγώνων [Press et al., 1988] για τον υπολογισμό του βέλτιστου διανύσματος \mathbf{t}_X . Εναλλακτικά μπορούμε να δημιουργήσουμε ένα σύστημα $3 \times n$ γραμμικών εξισώσεων με τρεις αγνώστους (τις παραμέτρους του διανύσματος \mathbf{t}_X), το οποίο μπορούμε να επιλύσουμε υπό την έννοια ελαχίστων τετραγώνων χρησιμοποιώντας τη μέθοδο της ψευδο-αντίστροφης ή τον αλγόριθμο SVD. Ο συνδυασμός των υπολογισθέντων \mathbf{R}_X και \mathbf{t}_X μας δίνει τελικά τον άγνωστο ομογενή μετασχηματισμό \mathbf{X} .

Μία προαιρετική μέθοδος για την περαιτέρω βελτιστοποίηση των αποτελεσμάτων είναι η προσπάθεια ταυτόχρονης ελαχιστοποίησης του αθροίσματος των δύο συναρτήσεων κόστους, ήτοι:

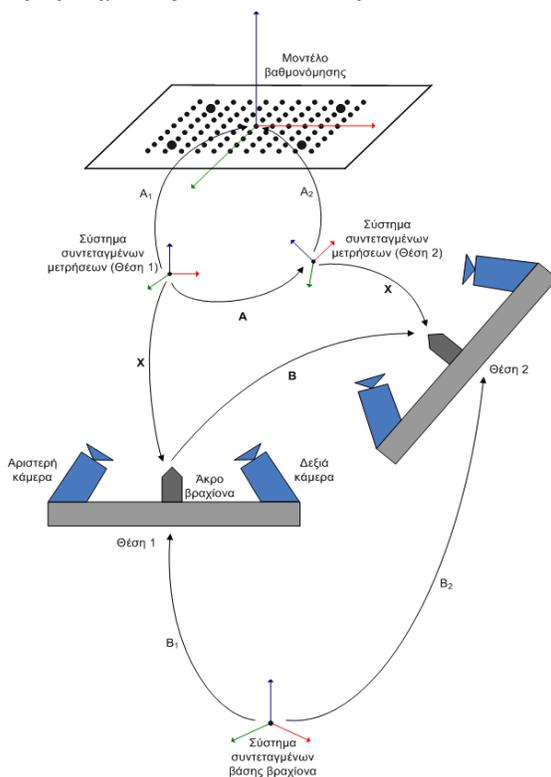
$$f(\mathbf{R}_X, \mathbf{t}_X) = f_1 + f_2 \quad (3.11)$$

Η παραπάνω συνάρτηση κόστους είναι μη-γραμμική ως προς τις παραμέτρους της, οπότε για την εύρεση των παραμέτρων \mathbf{R}_X και \mathbf{t}_X που την ελαχιστοποιούν είναι απαραίτητη η χρήση μιας μη-γραμμικής μεθόδου ελαχιστοποίησης, όπως είναι ο αλγόριθμος των Levenberg-Marquardt [Fletcher, 1990; Gill et al., 1989; Moré et al., 1980, 1984]. Ο αλγόριθμος απαιτεί την ύπαρξη μιας καλής αρχικής εκτίμησης για τις παραμέτρους (μεταβλητές) της συνάρτησης κόστους, οι οποίες είναι διαθέσιμες από την μέθοδο

επίλυσης κλειστής μορφής (closed form) που έχει προηγηθεί. Το κύριο μειονέκτημα των μεθόδων μη-γραμμικής βελτιστοποίησης είναι το γεγονός ότι απαιτούν αρκετό χρόνο και υπολογιστική ισχύ για να εφαρμοσθούν, καθώς η φύση τους είναι επαναληπτική (iterative) και όχι κλειστής μορφής (closed form). Επιπλέον, η εγκυρότητα των αποτελεσμάτων τους εξαρτάται σε μεγάλο βαθμό από την ποιότητα της αρχικής εκτίμησης για τις παραμέτρους της συνάρτησης κόστους, με αποτέλεσμα κάποιες φορές η ακρίβεια των αποτελεσμάτων μετά την βελτιστοποίηση να είναι πολύ χειρότερη από την αρχική τους εκτίμηση.

3.3 Βαθμονόμηση βραχίονα – μηχανικής όρασης για στερεοσκοπική κεφαλή

Στην παρούσα παράγραφο θα παρουσιαστεί μια καινούρια προσέγγιση του προβλήματος του Hand-Eye Calibration, για την περίπτωση χρήσης στερεοσκοπικής



Σχήμα 3.4 Τοπολογία πειράματος βαθμονόμησης βραχίονα – μηχανικής όρασης για στερεοσκοπική κεφαλή

κεφαλής καμερών (Stereo Head) αντί μιας μονής κάμερας. Μια πρώτη σκέψη θα ήταν να εφαρμόσουμε την μεθοδολογία της παραγράφου 3.2 για την κάθε κάμερα ξεχωριστά, υπολογίζοντας δηλαδή δύο μετασχηματισμούς συντεταγμένων X_1 και X_2 , οι οποίοι θα σχετίζονται με το κέντρο της κάθε κάμερας (pin-hole center) με το σύστημα συντεταγμένων του άκρου του βραχίονα. Το γεγονός ότι η χρήση στερεοσκοπικής κεφαλής μας επιτρέπει να πραγματοποιήσουμε μετρήσεις τρισδιάστατης ανακατασκευής (3D reconstruction) μας ωθεί να ακολουθήσουμε μια διαφορετική προσέγγιση για την επίλυση του προβλήματος, όπως θα παρουσιαστεί παρακάτω.

Στο Σχ. 3.4 απεικονίζεται η τοπολογία του πειράματος για την επίλυση του προβλήματος βαθμονόμησης βραχίονα – μηχανικής όρασης, για την περίπτωση χρήσης στερεοσκοπικής κεφαλής. Το επίπεδο μοντέλο βαθμονόμησης περιλαμβάνει, εκτός από τις κουκίδες στις προκαθορισμένες θέσεις, επιπλέον τέσσερις κύκλους στις γωνίες του, οι οποίοι θα χρησιμοποιηθούν για τις τρισδιάστατες μετρήσεις ανακατασκευής. Κάποιοι μετασχηματισμοί συντεταγμένων έχουν διαφορετική σημασία σε σχέση με την περίπτωση χρήσης μιας μονής κάμερας, όπως αναλύεται παρακάτω. Ο μετασχηματισμός \mathbf{A} αντιπροσωπεύει την μετατόπιση του συστήματος συντεταγμένων μέτρησης της στερεοσκοπικής κεφαλής, σε αντίθεση με την περίπτωση χρήσης μονής κάμερας όπου αντιπροσώπευε την μετατόπιση του κέντρου (pin-hole center) της κάμερας. Επίσης, ο μετασχηματισμός \mathbf{X} συνδέει το σύστημα συντεταγμένων μέτρησης της κεφαλής με το σύστημα συντεταγμένων του άκρου του βραχίονα, σε αντίθεση με την προηγούμενη περίπτωση που συνέδεε το κέντρο της κάμερας με το άκρο του βραχίονα. Ο υπολογισμός του μετασχηματισμού \mathbf{X} είναι μεγάλης σημασίας, καθώς επιτρέπει σε οποιαδήποτε εφαρμογή την απ' ευθείας αναγωγή των τρισδιάστατων μετρήσεων ανακατασκευής στο σύστημα συντεταγμένων του άκρου του βραχίονα. Ο μετασχηματισμός \mathbf{B} διατηρεί την παλιά του σημασία, και εκφράζει την μετατόπιση του συστήματος συντεταγμένων του άκρου του βραχίονα. Αναφερόμενοι στο Σχ. 3.4, η βασική ομογενής εξίσωση (3.3) εξακολουθεί να ισχύει και για την περίπτωση χρήσης στερεοσκοπικής κεφαλής, με διαφορετική όμως σημασιολογία για τους μετασχηματισμούς \mathbf{A} και \mathbf{X} , όπως αναφέρθηκε προηγουμένως. Ο μετασχηματισμός $\mathbf{B} = \mathbf{B}_1^{-1}\mathbf{B}_2$ δίνεται απ' ευθείας από τον ρομποτικό ελεγκτή (robot controller) μέσω της λύσης του ευθέως κινηματικού προβλήματος του βραχίονα. Ο μετασχηματισμός $\mathbf{A} = \mathbf{A}_1\mathbf{A}_2^{-1}$ μπορεί να υπολογιστεί με τη βοήθεια μετρήσεων τρισδιάστατης ανακατασκευής των κέντρων των τεσσάρων κύκλων που βρίσκονται πάνω στο μοντέλο βαθμονόμησης, όπως θα αναλυθεί στην συνέχεια.

Οι τέσσερις κύκλοι βρίσκονται σε συγκεκριμένα σημεία πάνω στο επίπεδο μοντέλο βαθμονόμησης, οπότε οι συντεταγμένες των κέντρων τους \mathbf{C}_{1N} , \mathbf{C}_{2N} , \mathbf{C}_{3N} και \mathbf{C}_{4N} είναι γνωστές ως προς το σύστημα συντεταγμένων του μοντέλου, έστω $\mathbf{C}_{1N}=(-40, -30, 0)$, $\mathbf{C}_{2N}=(40, -30, 0)$, $\mathbf{C}_{3N}=(-40, 30, 0)$ και $\mathbf{C}_{4N}=(40, 30, 0)$ (όλα σε mm). Προφανώς για όλες τις συντεταγμένες ισχύει $Z=0$, καθώς το επίπεδο που ανήκουν οι τέσσερις κύκλοι ταυτίζεται με το επίπεδο του μοντέλου βαθμονόμησης. Όπως έχει ήδη αναλυθεί σε

προηγούμενη ενότητα, στη θέση που πραγματοποιείται η βαθμονόμηση της στερεοσκοπικής κεφαλής το σύστημα συντεταγμένων μέτρησης ταυτίζεται με το σύστημα συντεταγμένων του μοντέλου βαθμονόμησης. Εάν από τη θέση αυτή της κεφαλής πραγματοποιήσουμε μέτρηση τρισδιάστατης ανακατασκευής των κέντρων των τεσσάρων κύκλων, τότε οι υπολογισθέντες συντεταγμένες τους θα ταυτίζονται κατά μεγάλη προσέγγιση με τις ονομαστικές τους τιμές \mathbf{C}_{1N} , \mathbf{C}_{2N} , \mathbf{C}_{3N} και \mathbf{C}_{4N} . Εάν στη συνέχεια μετακινήσουμε την στερεοσκοπική κεφαλή σε μια τυχαία θέση, έστω την θέση 1 του Σχ. 3.4, το σύστημα συντεταγμένων μέτρησης θ' ακολουθήσει την κίνηση της κεφαλής κατά ανάλογο τρόπο. Εφόσον το σύστημα συντεταγμένων μέτρησης δεν θα συμπίπτει πλέον με το σύστημα συντεταγμένων του μοντέλου βαθμονόμησης, μια νέα μέτρηση τρισδιάστατης ανακατασκευής των κέντρων των κύκλων θα δώσει διαφορετικές συντεταγμένες $\mathbf{C}_{1(1)}$, $\mathbf{C}_{2(1)}$, $\mathbf{C}_{3(1)}$ και $\mathbf{C}_{4(1)}$ σε σχέση με τις αρχικές ονομαστικές τους τιμές. Υπενθυμίζεται ότι οι μετρήσεις τρισδιάστατης ανακατασκευής αναφέρονται πάντα στο σύστημα συντεταγμένων μέτρησης της κεφαλής, οπότε η μεταβολή θέσης του συστήματος συντεταγμένων μέτρησης θα επηρεάσει τις τιμές των μετρήσεων, παρόλο που στην πραγματικότητα τα κέντρα των μετρούμενων κύκλων παραμένουν ακίνητα.

Αρχικός μας στόχος είναι να βρούμε το μετασχηματισμό συντεταγμένων \mathbf{A}_1 (βλ. Σχ. 3.4), αναφερόμενοι στην θέση 1 της συστοιχίας κεφαλής-άκρου βραχίονα, με την βοήθεια των μετρήσεων τρισδιάστατης ανακατασκευής των κέντρων των κύκλων. Στην πραγματικότητα έχουμε στην διάθεσή μας τις συντεταγμένες των κέντρων των τεσσάρων κύκλων ως προς δύο ξεχωριστά συστήματα συντεταγμένων, ήτοι το σταθερό σύστημα συντεταγμένων του μοντέλου βαθμονόμησης και το κινούμενο μαζί με την στερεοσκοπική κεφαλή σύστημα συντεταγμένων μέτρησης. Για να βρούμε τον μετασχηματισμό που συνδέει τα δύο αυτά συστήματα συντεταγμένων αρκεί να υπολογίσουμε τον βέλτιστο μετασχηματισμό που σχετίζει τις ονομαστικές συντεταγμένες των κέντρων των κύκλων \mathbf{C}_{1N} , \mathbf{C}_{2N} , \mathbf{C}_{3N} και \mathbf{C}_{4N} και τις μετρούμενες τιμές τους $\mathbf{C}_{1(1)}$, $\mathbf{C}_{2(1)}$, $\mathbf{C}_{3(1)}$ και $\mathbf{C}_{4(1)}$ από την θέση 1 της στερεοσκοπικής κεφαλής. Το πρόβλημα ανάγεται στην ελαχιστοποίηση της παρακάτω μη-γραμμικής συνάρτησης κόστους:

$$f(\mathbf{R}_{A_1}, \mathbf{t}_{A_1}) = \sum_{i=1}^4 \left\| \mathbf{C}_{i(1)} - (\mathbf{R}_{A_1} \mathbf{C}_{i(N)} + \mathbf{t}_{A_1}) \right\|^2, \quad (3.12)$$

όπου \mathbf{R}_{A_1} είναι ο 3×3 πίνακας στροφής και \mathbf{t}_{A_1} το 3×1 διάνυσμα μετατόπισης που αντιστοιχούν στον ζητούμενο μετασχηματισμό \mathbf{A}_1 . Για την εύρεση του βέλτιστου μετασχηματισμού συντεταγμένων \mathbf{A}_1 που ελαχιστοποιεί την (3.12) υπάρχει αναλυτική λύση (closed-form) [Horn, 1987] η οποία μας αποτρέπει από την χρήση των επαναληπτικών (iterative) και υπολογιστικά χρονοβόρων μεθόδων μη-γραμμικής βελτιστοποίησης, όπως είναι οι αλγόριθμοι Levenberg-Marquardt και Steepest Descent [Fletcher, 1990; Gill et al., 1989; Press et al., 1988]. Επαναλαμβάνοντας την ίδια διαδικασία για την θέση 2 της συστοιχίας κεφαλής-άκρου, μπορούμε να υπολογίσουμε επίσης τον μετασχηματισμό συντεταγμένων \mathbf{A}_2 (βλ. Σχ. 3.4), ο οποίος συνδέει το σταθερό σύστημα συντεταγμένων του μοντέλου βαθμονόμησης με την καινούρια θέση του συστήματος συντεταγμένων μέτρησης της κεφαλής. Με την βοήθεια των παραπάνω μπορούμε τελικά να υπολογίσουμε τον ζητούμενο μετασχηματισμό μετατόπισης $\mathbf{A} = \mathbf{A}_1 \mathbf{A}_2^{-1}$ του συστήματος συντεταγμένων μέτρησης της κεφαλής, από την θέση 1 στην θέση 2.

Όπως περιγράφεται παραπάνω, η βασική ομογενής εξίσωση (3.3) για την περίπτωση χρήσης στερεοσκοπικής κεφαλής αναφέρεται σε μία κίνηση της συστοιχίας κάμερας-άκρου βραχίονα (από την θέση 1 στην θέση 2). Όπως και στην περίπτωση χρήσης μιας μονής κάμερας, έτσι και τώρα μπορούμε να επεκτείνουμε την (3.3) για n κινήσεις της συστοιχίας ($n+1$ διακριτές θέσεις), οπότε καταλήγουμε στην παρακάτω σχέση:

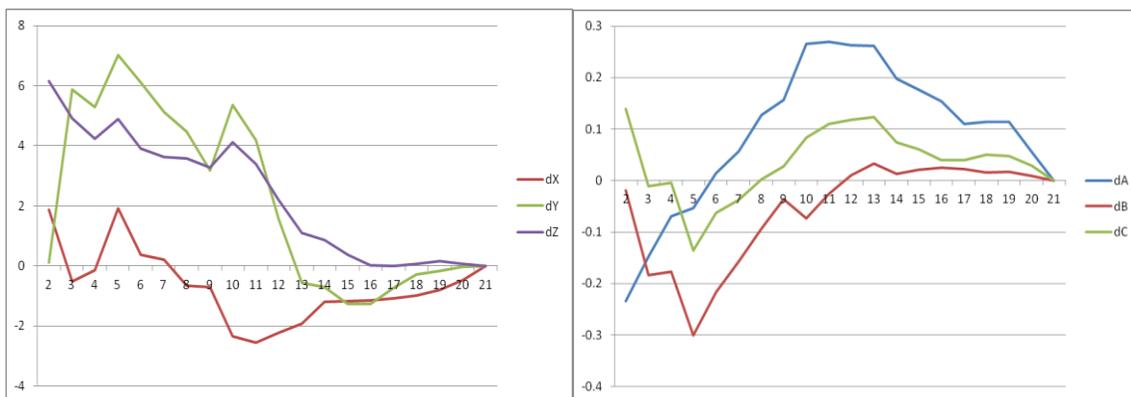
$$i=1,2,\dots,n \quad \mathbf{A}_i \mathbf{X} = \mathbf{X} \mathbf{B}_i, \quad (3.13)$$

Για την επίλυση της (3.13) ως προς \mathbf{X} ακολουθείται ακριβώς η ίδια μεθοδολογία με την περίπτωση χρήσης μονής κάμερας, ακολουθώντας τις εξισώσεις (3.4-3.11).

3.4 Πειραματικά αποτελέσματα

Στη συνέχεια θα παρουσιαστούν ορισμένα πειραματικά αποτελέσματα, τα οποία αποδεικνύουν την υπεροχή της προτεινόμενης μεθόδου του Κεφ. 3.3 έναντι της κλασσικής μεθόδου του Κεφ. 3.2. Για την εφαρμογή των εργαστηριακών πειραμάτων, η τοπολογία των οποίων απεικονίζονται στα Σχ. 3.3 και Σχ. 3.4, έχει χρησιμοποιηθεί ένας βιομηχανικός ρομποτικός βραχίονας τύπου KR15/1 της KUKA, ο οποίος ελέγχεται από έναν ρομποτικό ελεγκτή τύπου KRC1. Όλοι οι αλγόριθμοι που περιγράφονται στα Κεφ.

3.1-3.3, καθώς και οι αλγόριθμοι επεξεργασίας εικόνας, έχουν υλοποιηθεί σε γλώσσα προγραμματισμού C++ σε περιβάλλον Microsoft Visual Studio 6.0, το οποίο είναι εγκατεστημένο σ' έναν τυπικό ηλεκτρονικό υπολογιστή (συμβατό IBM PC). Για την εφαρμογή του αλγορίθμου βαθμονόμησης κάμερας του Tsai έχει χρησιμοποιηθεί η βιβλιοθήκη [<http://www.cs.cmu.edu/~rgw/TsaiCode.html>], ενώ για τον αντίστοιχο του Zhang έχει χρησιμοποιηθεί το εκτελέσιμο αρχείο Easycalib [<http://research.microsoft.com/en-us/um/people/zhang/calib>]. Στο PC είναι εγκατεστημένη μια PCI κάρτα γραφικών Matrox Meteor II MC (Frame Grabber), η οποία μπορεί να λαμβάνει εικόνες από δύο ψηφιακές κάμερες Sony HR50 (640×495 pixel resolution) που είναι συνδεδεμένες πάνω του.



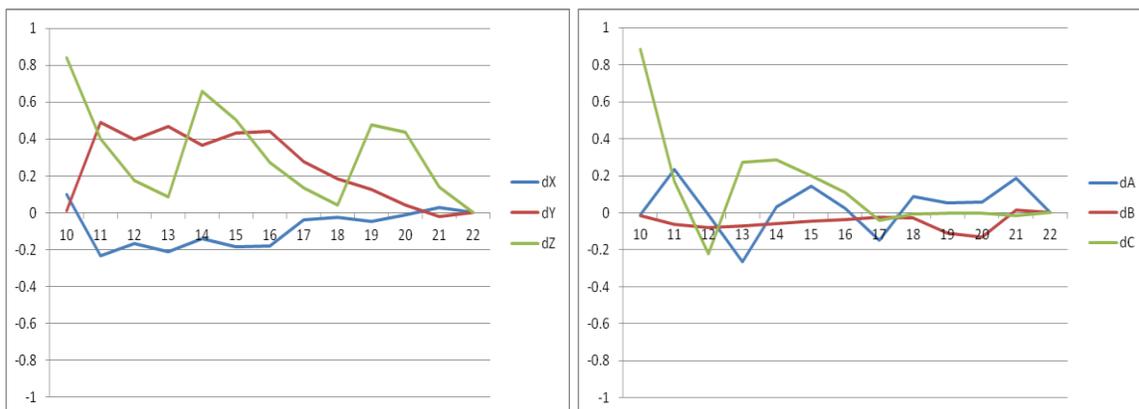
Σχήμα 3.5 Ρυθμός σύγκλισης του διανύσματος μετατόπισης (αριστερά) και των γωνιών στροφής του εξαγόμενου μετασχηματισμού X

Αρχικά παρουσιάζονται τ' αποτελέσματα της κλασσικής μεθόδου που περιγράφεται στο Κεφ. 3.2, όπου οι μετασχηματισμοί συντεταγμένων A_i της κάμερας για κάθε μετατόπιση του ρομπότ προέρχονται από τις εξωτερικές παραμέτρους της βαθμονόμησης της κάμερας. Για το πείραμα αυτό προτιμήθηκε ο αλγόριθμος βαθμονόμησης κάμερας του Z.Zhang έναντι του αντίστοιχου του R.Tsai καθότι παρατηρήθηκε ότι ο πρώτος υπολόγιζε τις εξωτερικές παραμέτρους της κάμερας με μεγαλύτερη απόλυτη ακρίβεια σε σύγκριση με τον δεύτερο. Στο Σχ. 3.5 απεικονίζεται ο ρυθμός σύγκλισης των αποτελεσμάτων του ζητούμενου μετασχηματισμού κάμερας-ρομποτικού βραχίονα X , όσον αφορά το διάνυσμα μετατόπισης (x, y, z) και τις γωνίες στροφής roll-pitch-yaw που προκύπτουν από τον πίνακα στροφής. Τα αποτελέσματα δίνονται σε σχέση με τις τελικές τιμές του υπολογισθέντος μετασχηματισμού X , ο οποίος προκύπτει χρησιμοποιώντας όλα

τα διαθέσιμα δεδομένα απ' όλες τις μετατοπίσεις του βραχίονα. Από το Σχ. 3.5 είναι προφανές ότι ο αλγόριθμος αρχίζει να συγκλίνει ικανοποιητικά μετά από 10 μετατοπίσεις του βραχίονα ενώ παρέχει ακρίβεια της τάξης του 1mm για το διάνυσμα μετατόπισης και 0.1° για τις γωνίες στροφής χρησιμοποιώντας περισσότερες από 20 μετατοπίσεις. Ο περιοριστικός παράγοντας στην ακρίβεια και τον ρυθμό σύγκλισης των αποτελεσμάτων δεν είναι ο ίδιος ο αλγόριθμος βαθμονόμησης κάμερας-ρομποτικού βραχίονα αλλά η ποιότητα των δεδομένων εισόδου, ήτοι οι μετασχηματισμοί της κάμερας \mathbf{A}_i και του άκρου του βραχίονα \mathbf{B}_i . Το μέσο τετραγωνικό σφάλμα για το διάνυσμα μετατόπισης και τις γωνίες στροφής του μετασχηματισμού \mathbf{X} , χρησιμοποιώντας όλα τα διαθέσιμα δεδομένα απ' όλες τις μετατοπίσεις του βραχίονα είναι $E_t=3.017231$ και $E_r=0.299305$ αντίστοιχα, όπου:

$$\mathbf{E} = \sum_n \|\mathbf{A}_i \mathbf{X} - \mathbf{X} \mathbf{B}_i\|^2$$

Στη συνέχεια παρουσιάζονται τ' αποτελέσματα της προτεινόμενης μεθόδου που περιγράφεται στο Κεφ. 3.3, όπου οι μετασχηματισμοί συντεταγμένων \mathbf{A}_i της κάμερας για κάθε μετατόπιση i του ρομπότ προέρχονται από τις μετρήσεις τρισδιάστατης ανακατασκευής από την στερεοσκοπική κεφαλή. Για την βαθμονόμηση της στερεοσκοπικής κεφαλής χρησιμοποιήθηκε ο αλγόριθμος του R.Tsai, καθότι στην περίπτωση αυτή δεν μας ενδιαφέρει η απόλυτη ακρίβεια των εξωτερικών παραμέτρων των καμερών, παρά μόνο η πραγματοποίηση ακριβών μετρήσεων ανακατασκευής των συντεταγμένων των τεσσάρων κύκλων.

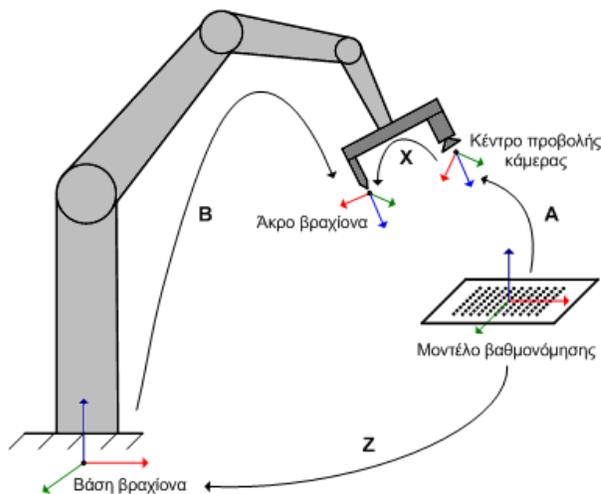


Σχήμα 3.6 Ρυθμός σύγκλισης του διανύσματος μετατόπισης (αριστερά) και των γωνιών στροφής του εξαγόμενου μετασχηματισμού \mathbf{X}

Από το Σχ. 3.6 είναι προφανές είναι ότι το διάνυσμα μετατόπισης του υπολογισθέντος μετασχηματισμού X συγκλίνει με μεγαλύτερη ακρίβεια σε σχέση με τ' αποτελέσματα της κλασικής μεθόδου που παρουσιάζονται στο Σχ. 3.5, ενώ οι γωνίες στροφής υπολογίζονται με περίπου την ίδια ακρίβεια. Για την σύγκλιση των αποτελεσμάτων απαιτούνται δεδομένα από περισσότερες από 10 μετατοπίσεις του βραχίονα, ενώ με περισσότερες από 20 μετατοπίσεις οι ανακρίβειες του διανύσματος μετατόπισης και των γωνιών στροφής είναι περίπου 0.2mm και 0.1° αντίστοιχα. Το μέσο τετραγωνικό σφάλμα χρησιμοποιώντας όλα τα διαθέσιμα δεδομένα είναι $E_t = 0.964983$ για το διάνυσμα μετατόπισης και 0.171481 για τις γωνίες στροφής, τα οποία είναι εμφανώς μικρότερα σε σύγκριση με τα αντίστοιχα της κλασικής μεθόδου. Συμπέρασμα της σύγκρισης των δύο μεθόδων είναι ότι οι μετρήσεις τρισδιάστατης ανακατασκευής παρέχουν καλύτερη απόλυτη ακρίβεια από τις εξωτερικές παραμέτρους βαθμονόμησης κάμερας, γεγονός που βελτιώνει την ακρίβεια των αποτελεσμάτων της μεθόδου βαθμονόμησης κάμερας-ρομποτικού βραχίονα.

4 ΤΑΥΤΟΧΡΟΝΗ ΒΑΘΜΟΝΟΜΗΣΗ ΒΡΑΧΙΟΝΑ, ΜΗΧΑΝΙΚΗΣ ΟΡΑΣΗΣ ΚΑΙ ΧΩΡΟΥ ΕΡΓΑΣΙΑΣ

4.1 Ταυτόχρονη βαθμονόμηση βραχίονα, μηχανικής όρασης και χώρου εργασίας για μονή κάμερα



Σχήμα 4.1 Τοπολογία πειράματος ταυτόχρονης βαθμονόμησης βραχίονα, μηχ. όρασης και χώρου εργασίας για μονή κάμερα

Στην προηγούμενη ενότητα παρουσιάστηκε κι αναλύθηκε η μεθοδολογία επίλυσης του προβλήματος βαθμονόμησης βραχίονα – μηχανικής όρασης, η οποία μας επιτρέπει να υπολογίζουμε με μεγάλη ακρίβεια την σχετική θέση και προσανατολισμό των συστημάτων συντεταγμένων του άκρου ενός ρομποτικού βραχίονα και μιας συσκευής μηχανικής όρασης (μονής κάμερας ή στερεοσκοπικής κεφαλής) η οποία είναι σταθερά στερεωμένη πάνω του. Υπάρχουν εφαρμογές όπου η

πληροφορία αυτή δεν είναι αρκετή, και απαιτείται παράλληλα ο υπολογισμός και κάποιου άλλου μετασχηματισμού συντεταγμένων. Το Σχ. 4.1 απεικονίζει την τοπολογία του πειράματος της προηγούμενης ενότητας για την περίπτωση χρήσης μιας μονής κάμερας, σκοπός του οποίου ήταν ο υπολογισμός του άγνωστου μετασχηματισμού X που συνδέει το κέντρο προβολής της κάμερας με το άκρο του βραχίονα. Παρατηρούμε ότι εκτός από τον X , ο μετασχηματισμός Z , που συνδέει τα συστήματα συντεταγμένων του χρησιμοποιούμενου μοντέλου βαθμονόμησης και της βάσης του ρομπότ, είναι σταθερός κι ανεξάρτητος της θέσης και προσανατολισμού του άκρου του βραχίονα και της κάμερας. Αν B είναι ο μετασχηματισμός που συνδέει τα συστήματα συντεταγμένων της βάσης του βραχίονα και του άκρου του, και A ο μετασχηματισμός που συνδέει τα συστήματα συντεταγμένων του κέντρου προβολής της κάμερας και του μοντέλου

βαθμονόμησης, τότε απ' ευθείας από το Σχ. 4.1 μπορεί να εξαχθεί η παρακάτω εξίσωση ομογενών μετασχηματισμών:

$$\mathbf{AX} = \mathbf{ZB} \quad (4.1)$$

Οι μεθοδολογίες που επιτρέπουν τον ταυτόχρονο υπολογισμό των αγνώστων μετασχηματισμών \mathbf{X} και \mathbf{Z} ονομάζονται «*Ταυτόχρονη βαθμονόμηση βραχίονα, μηχανικής όρασης και χώρου εργασίας*» (*Hand-Eye and Robot-World calibration*) [Zhuang et al., 1994; Dornaika and Horaud, 1998]. Υπενθυμίζεται ότι οι μετασχηματισμοί \mathbf{X} και \mathbf{Z} είναι σταθεροί, ενώ οι \mathbf{A} και \mathbf{B} εξαρτώνται από την θέση και τον προσανατολισμό του άκρου του βραχίονα (άρα και της κάμερας που είναι σταθερά στερεωμένη πάνω του).

Μια διαφορετική τοπολογία πειράματος, η οποία επίσης καταλήγει στην εξίσωση (4.1), είναι αυτή όπου η κάμερα είναι σταθερά στερεωμένη μέσα στο χώρο εργασίας του ρομπότ, και παρατηρεί κάποιο μοντέλο βαθμονόμησης που είναι σταθερά τοποθετημένο πάνω στο κινούμενο άκρο του βραχίονα [Dornaika and Horaud, 1998]. Στην περίπτωση αυτή ο μετασχηματισμός \mathbf{X} αναφέρεται στον άγνωστο συσχετισμό των συστημάτων συντεταγμένων του παρατηρούμενου μοντέλου και του άκρου του βραχίονα, ενώ ο μετασχηματισμός \mathbf{Z} αναφέρεται στον συσχετισμό των συστημάτων συντεταγμένων του κέντρου προβολής της κάμερας και της βάσης του βραχίονα. Όπως και στην αρχική τοπολογία του Σχ. 4.1, οι μετασχηματισμοί \mathbf{X} και \mathbf{Z} είναι σταθεροί, ενώ οι \mathbf{A} και \mathbf{B} εξαρτώνται από την θέση και τον προσανατολισμό του άκρου του βραχίονα (άρα και του μοντέλου βαθμονόμησης που είναι σταθερά στερεωμένο πάνω του).

Η εξίσωση ομογενών μετασχηματισμών (4.1) μπορεί να αποσυντεθεί σε μια εξίσωση 3×3 πινάκων στροφής και σε μια διανυσματική εξίσωση, ήτοι:

$$\mathbf{R}_A \mathbf{R}_X = \mathbf{R}_Z \mathbf{R}_B \quad (4.2)$$

$$\mathbf{R}_A \mathbf{t}_X + \mathbf{t}_A = \mathbf{R}_Z \mathbf{t}_B + \mathbf{t}_Z \quad (4.3)$$

όπου \mathbf{R}_A , \mathbf{t}_A , \mathbf{R}_B , \mathbf{t}_B , \mathbf{R}_X , \mathbf{t}_X και \mathbf{R}_Z , \mathbf{t}_Z είναι οι 3×3 πίνακες στροφής και τα 3-διάστατα διανύσματα θέσης που αντιστοιχούν στους μετασχηματισμούς \mathbf{A} , \mathbf{B} , \mathbf{X} και \mathbf{Z} αντίστοιχα. Με την επίλυση της (4.2) είναι δυνατός ο υπολογισμός των αγνώστων πινάκων στροφής \mathbf{R}_X και \mathbf{R}_Z , δεδομένου ότι οι πίνακες στροφής \mathbf{R}_A και \mathbf{R}_B είναι ήδη γνωστοί από τις εξωτερικές παραμέτρους βαθμονόμησης της κάμερας και από την επίλυση του ευθέως κινηματικού προβλήματος από τον ρομποτικό ελεγκτή αντίστοιχα. Ο μετέπειτα υπολογισμός των αγνώστων διανυσμάτων μετατόπισης \mathbf{t}_X και \mathbf{t}_Z γίνεται απλή υπόθεση,

καθώς η (4.3) ανάγεται σε ένα σύστημα γραμμικών εξισώσεων με έξι αγνώστους. Υπάρχουν τρεις διαφορετικές προσεγγίσεις για την επίλυση της (4.2), ήτοι α) η γραμμική μέθοδος, β) η αναλυτική μέθοδος κλειστής μορφής και γ) η μη-γραμμική μέθοδος, οι οποίες θ' αναλυθούν εκτενώς στην συνέχεια.

4.1.1 Γραμμική μέθοδος

Για ν' απλοποιηθεί η εξίσωση (4.2) μπορούμε να μετατρέψουμε τους 3×3 πίνακες στροφής στις αντίστοιχες αναπαραστάσεις τους υπό την μορφή μεταβλητών *Unit Quaternions* [Kranchenko, 2003]. Όπως αναλύεται στην βιβλιογραφία, η αναπαράσταση ενός πίνακα στροφής \mathbf{R} υπό την μορφή των μεταβλητών Unit Quaternions συντελείται με τη μορφή ενός 4-διάστατου διανύσματος $\mathbf{q}_R = (q_0, q_1, q_2, q_3) = (q_0, \mathbf{q})$, το οποίο συνδέεται μονοσήμαντα με τον αντίστοιχο 3×3 πίνακα στροφής με την παρακάτω εξίσωση:

$$\mathbf{R}(\mathbf{q}_R) = \begin{bmatrix} q_0^2 + q_1^2 - q_2^2 - q_3^2 & 2(q_1q_2 - q_0q_3) & 2(q_1q_3 + q_0q_2) \\ 2(q_1q_2 + q_0q_3) & q_0^2 + q_2^2 - q_1^2 - q_3^2 & 2(q_2q_3 - q_0q_1) \\ 2(q_1q_3 - q_0q_2) & 2(q_2q_3 + q_0q_1) & q_0^2 + q_3^2 - q_1^2 - q_2^2 \end{bmatrix} \quad (4.4)$$

Οι μεταβλητές Unit Quaternions που αντιστοιχούν στους πίνακες στροφής \mathbf{R}_A , \mathbf{R}_B , \mathbf{R}_X και \mathbf{R}_Z θα είναι αντίστοιχα $\mathbf{q}_A = (a_0, \mathbf{a})$, $\mathbf{q}_B = (b_0, \mathbf{b})$, $\mathbf{q}_X = (x_0, \mathbf{x})$ και $\mathbf{q}_Z = (z_0, \mathbf{z})$. Επαναδιατυπώνοντας την εξίσωση (4.2), χρησιμοποιώντας τις μεταβλητές Unit Quaternions για την αναπαράσταση των πινάκων στροφής, καταλήγουμε στην παρακάτω εξίσωση:

$$\mathbf{q}_A * \mathbf{q}_X = \mathbf{q}_Z * \mathbf{q}_B \quad (4.5)$$

όπου ο τελεστής '*' αντιστοιχεί στην πράξη πολλαπλασιασμού μεταξύ δύο μεταβλητών τύπου Unit Quaternion, και δεν σχετίζεται σε καμία περίπτωση με την πράξη πολλαπλασιασμού πινάκων ή εσωτερικού γινομένου διανυσμάτων. Αναλύοντας την (4.5) καταλήγουμε σε μία βαθμωτή και μία διανυσματική εξίσωση, ως εξής:

$$a_0x_0 - \mathbf{a} \cdot \mathbf{x} = z_0b_0 - \mathbf{b} \cdot \mathbf{z} \quad (4.6)$$

$$a_0\mathbf{x} + x_0\mathbf{a} + \mathbf{a} \times \mathbf{x} = z_0\mathbf{b} + b_0\mathbf{z} - \mathbf{b} \times \mathbf{z} \quad (4.7)$$

όπου οι τελεστές ' \cdot ' και ' \times ' αντιπροσωπεύουν το εσωτερικό και το εξωτερικό γινόμενο για τα αντίστοιχα τρισδιάστατα διανύσματα. Εάν $a_0 \neq 0$, τότε η (4.6) μπορεί να γραφτεί ως εξής:

$$x_0 = (\mathbf{a}/a_0) \cdot \mathbf{x} + (b_0/a_0)z_0 - (\mathbf{b}/a_0) \cdot \mathbf{z} \quad (4.8)$$

Αντικαθιστώντας την (4.8) στην (4.6) και εκφράζοντας τα 3D διανύσματα υπό την μορφή 3×1 πινάκων στήλης, μπορούμε να καταλήξουμε στην παρακάτω εξίσωση πινάκων:

$$\begin{aligned} & (a_0 \mathbf{I} + \mathbf{a}\mathbf{a}^T / a_0 + \Omega(\mathbf{a}))\mathbf{x} + \\ & (-b_0 \mathbf{I} - \mathbf{a}\mathbf{b}^T / a_0 + \Omega(\mathbf{b}))\mathbf{z} = z_0 \mathbf{b} - z_0 (b_0 / a_0) \mathbf{a} \end{aligned} \quad (4.9)$$

όπου $\Omega(\mathbf{u})$ είναι ένας 3×3 πίνακας που σχετίζεται με ένα οποιοδήποτε 3D διάνυσμα $\mathbf{u} = [u_x, u_y, u_z]^T$ σύμφωνα με την παρακάτω σχέση:

$$\Omega(\mathbf{u}) = \begin{bmatrix} 0 & -u_z & u_y \\ u_z & 0 & -u_x \\ -u_y & u_x & 0 \end{bmatrix} \quad (4.10)$$

Στην πραγματικότητα ο πίνακας της (4.10) εκφράζει τον τελεστή του εξωτερικού γινομένου 3-διάστατων διανυσμάτων υπό την μορφή πολλαπλασιασμού πινάκων, ήτοι:

$$\mathbf{u} \times \mathbf{v} = \Omega(\mathbf{u})\mathbf{v} = \begin{bmatrix} 0 & -u_z & u_y \\ u_z & 0 & -u_x \\ -u_y & u_x & 0 \end{bmatrix} \begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix} \quad (4.11)$$

Για την εξαγωγή της (4.9) έχει επίσης χρησιμοποιηθεί η ταυτότητα $(\mathbf{u} \cdot \mathbf{v}) \cdot \mathbf{w} = (\mathbf{w}\mathbf{v}^T)\mathbf{u}$ για τρία τυχαία 3D διανύσματα \mathbf{u} , \mathbf{v} , \mathbf{w} εκφρασμένα υπό την μορφή 3×1 πινάκων στήλης. Η (4.9) μπορεί να αναδιατυπωθεί υπό την μορφή ενός γραμμικού συστήματος τριών εξισώσεων με έξι αγνώστους ($z_0 \neq 0$):

$$\underset{3 \times 6}{\mathbf{J}} \cdot \underset{6 \times 1}{\mathbf{u}} = \mathbf{b} - (b_0 / a_0) \mathbf{a} \quad (4.12)$$

όπου $\mathbf{u} = [u_1, u_2, u_3, u_4, u_5, u_6]^T = [\mathbf{x}/z_0, \mathbf{z}/z_0]^T$ είναι το 6-διάστατο διάνυσμα των αγνώστων και \mathbf{J} είναι ένας γνωστός 3×6 πίνακας, ο οποίος μπορεί να υπολογιστεί από τις μεταβλητές \mathbf{q}_A και \mathbf{q}_B που προκύπτουν από του γνωστούς πίνακες στροφής \mathbf{R}_A and \mathbf{R}_B αντίστοιχα. Η εξίσωση (4.12) αποτελεί ένα σύστημα τριών γραμμικών εξισώσεων με έξι αγνώστους, το οποίο προφανώς δεν έχει μοναδική λύση.

Για την επίλυση του συστήματος (4.12) είναι απαραίτητη η επανάληψη της παραπάνω διαδικασίας για τουλάχιστο έναν ακόμα διαφορετικό προσανατολισμό του άκρου του βραχίονα, δηλαδή για διαφορετικές τιμές για τους πίνακες στροφής \mathbf{R}_A και \mathbf{R}_B . Κατ' αυτόν τον τρόπο θα καταλήξουμε σ' ένα ακόμα γραμμικό σύστημα της μορφής της (4.12) με διαφορετικούς συντελεστές αλλά το ίδιο διάνυσμα των αγνώστων \mathbf{u} , καθώς

οι άγνωστες μεταβλητές Unit Quaternions $\mathbf{q}_X = (x_0, \mathbf{x})$ $\mathbf{q}_Z = (z_0, \mathbf{z})$ είναι σταθερές και ανεξάρτητες από την θέση και προσανατολισμό του άκρου του βραχίονα και της κάμερας. Συνδυάζοντας τα δύο γραμμικά συστήματα της μορφής της (4.12) μπορούμε να καταλήξουμε σ' ένα 6×6 γραμμικό σύστημα, το οποίο μπορεί να δώσει μοναδική λύση για το διάνυσμα των αγνώστων \mathbf{u} . Στην γενική περίπτωση που χρησιμοποιούμε δεδομένα από n διαφορετικούς προσανατολισμούς του άκρου του βραχίονα, τότε καταλήγουμε σ' ένα σύστημα $3n$ γραμμικών εξισώσεων με έξι αγνώστους, το οποίο μπορούμε να επιλύσουμε με τη βοήθεια κλασσικών γραμμικών μεθόδων [Press et al., 1988; Golub and Van Loan, 1983], όπως του ψευδοαντίστροφου πίνακα ή της SVD. Προφανώς, περιμένουμε η ακρίβεια και η αξιοπιστία των αποτελεσμάτων να αυξάνεται συναρτήσει του αριθμού n των διαφορετικών προσανατολισμών του άκρου του βραχίονα, καθώς συνυπολογίζονται όλο και περισσότερα δεδομένα για την εξαγωγή της βέλτιστης λύσης για το διάνυσμα των αγνώστων \mathbf{u} .

Εφόσον το διάνυσμα \mathbf{u} γίνει γνωστό, τα παράγωγα των \mathbf{q}_X και \mathbf{q}_Z μπορούν να βρεθούν από τις παρακάτω εξισώσεις:

$$\begin{aligned} z_0 &= \pm(1 + u_4 + u_5 + u_6)^{-1/2} \\ \mathbf{z} &= z_0[u_4, u_5, u_6] \\ \mathbf{x} &= z_0[u_1, u_2, u_3] \\ x_0 &= \pm(1 - \|\mathbf{x}\|^2)^{1/2} \end{aligned} \quad (4.13)$$

Τέλος οι άγνωστοι πίνακες στροφής \mathbf{R}_X και \mathbf{R}_Z που αντιστοιχούν στις μεταβλητές \mathbf{q}_X και \mathbf{q}_Z μπορούν να βρεθούν με τη βοήθεια της εξίσωσης (4.4).

4.1.2 Αναλυτική μέθοδος

Η αναλυτική μέθοδος κλειστής μορφής [Dornaika and Horaud, 1998] δημιουργεί μια θετικά ορισμένη τετραγωνική συνάρτηση σφάλματος, την οποία στην συνέχεια προσπαθεί να ελαχιστοποιήσει για να εξαγάγει τις βέλτιστες λύσεις για τους άγνωστους πίνακες στροφής \mathbf{R}_X και \mathbf{R}_Z . Χρησιμοποιώντας τις μεταβλητές *Unit Quaternion* \mathbf{q}_X και \mathbf{q}_Z για την αναπαράσταση των \mathbf{R}_X και \mathbf{R}_Z αντίστοιχα, και εκφράζοντας τον τελεστή πολλαπλασιασμού '*' των μεταβλητών υπό την μορφή πολλαπλασιασμού πινάκων, έχουμε:

$$\mathbf{q}_A * \mathbf{q}_X = Q(\mathbf{q}_A) \mathbf{q}_X \quad (4.14)$$

$$\mathbf{q}_Z * \mathbf{q}_B = W(\mathbf{q}_B) \mathbf{q}_Z \quad (4.15)$$

Αντικαθιστώντας τις παραπάνω εξισώσεις στην βασική εξίσωση (4.2), μπορούμε να καταλήξουμε στην παρακάτω 4-διάστατη συνάρτηση σφάλματος:

$$Q(\mathbf{q}_A) \mathbf{q}_X - W(\mathbf{q}_B) \mathbf{q}_Z = \mathbf{0} \quad (4.16)$$

Σύμφωνα με την άλγεβρα των μεταβλητών Unit Quaternion, οι πίνακες $Q(\mathbf{q})$ and $W(\mathbf{q})$ για μια αυθαίρετη μεταβλητή \mathbf{q} είναι αντίστοιχα της παρακάτω μορφής:

$$Q(\mathbf{q}) = \begin{bmatrix} q_0 & -q_x & -q_y & -q_z \\ q_x & q_0 & -q_z & q_y \\ q_y & q_z & q_0 & -q_x \\ q_z & -q_y & q_x & q_0 \end{bmatrix} \quad (4.17)$$

$$W(\mathbf{q}) = \begin{bmatrix} q_0 & -q_x & -q_y & -q_z \\ q_x & q_0 & q_z & -q_y \\ q_y & -q_z & q_0 & q_x \\ q_z & q_y & -q_x & q_0 \end{bmatrix} \quad (4.18)$$

Οι πίνακες αυτοί είναι ορθογώνιοι, οπότε:

$$Q(\mathbf{q})^T Q(\mathbf{q}) = \mathbf{q}^T \mathbf{q} \mathbf{I} = \mathbf{I} \quad (4.19)$$

$$W(\mathbf{q})^T W(\mathbf{q}) = \mathbf{q}^T \mathbf{q} \mathbf{I} = \mathbf{I} \quad (4.20)$$

Η τετραγωνική νόρμα της παραπάνω συνάρτησης σφάλματος (4.16) μπορεί να εκφραστεί με την παρακάτω μορφή:

$$\begin{aligned} & \|Q(\mathbf{q}_A) \mathbf{q}_X - W(\mathbf{q}_B) \mathbf{q}_Z\|^2 = \\ & [Q(\mathbf{q}_A) \mathbf{q}_X - W(\mathbf{q}_B) \mathbf{q}_Z]^T [Q(\mathbf{q}_A) \mathbf{q}_X - W(\mathbf{q}_B) \mathbf{q}_Z] = \\ & \mathbf{q}_X^T Q(\mathbf{q}_A)^T Q(\mathbf{q}_A) \mathbf{q}_X + \mathbf{q}_Z^T W(\mathbf{q}_B)^T W(\mathbf{q}_B) \mathbf{q}_Z - \\ & \mathbf{q}_Z^T W(\mathbf{q}_B)^T Q(\mathbf{q}_A) \mathbf{q}_X - \mathbf{q}_X^T Q(\mathbf{q}_A)^T W(\mathbf{q}_B) \mathbf{q}_Z \end{aligned} \quad (4.21)$$

Ορίζουμε ως \mathbf{u} το παρακάτω 8-διάστατο διάνυσμα των αγνώστων:

$$\mathbf{u}^T = [\mathbf{q}_X^T, \mathbf{q}_Z^T] \quad (4.22)$$

Η τετραγωνική συνάρτηση σφάλματος (quadratic form) μπορεί να πάρει την μορφή:

$$\|Q(\mathbf{q}_A) \mathbf{q}_X - W(\mathbf{q}_B) \mathbf{q}_Z\|^2 = \mathbf{u}^T \mathbf{S}_i \mathbf{u} \quad (4.23)$$

Ο 8×8 θετικά ορισμένος συμμετρικός πίνακας \mathbf{S}_i είναι της μορφής:

$$\mathbf{S}_i = \begin{bmatrix} \mathbf{I} & \mathbf{C}_i \\ \mathbf{C}_i^T & \mathbf{I} \end{bmatrix} \quad (4.24)$$

όπου $\mathbf{C}_i = -\mathbf{Q}(\mathbf{q}_A)^T \mathbf{W}(\mathbf{q}_B)$ είναι ένας ορθογώνιος πίνακας με τάξη (rank) 4.

Επεκτείνοντας την παραπάνω διαδικασία για την περίπτωση n διακριτών προσανατολισμών του άκρου του βραχίονα, δηλαδή για διαφορετικές τιμές των μεταβλητών \mathbf{q}_A και \mathbf{q}_B , η παραπάνω τετραγωνική συνάρτηση σφάλματος μπορεί να γενικευθεί ώστε να περιλαμβάνει όλα τα διαθέσιμα δεδομένα.

$$f(\mathbf{q}_X, \mathbf{q}_Z) = \sum_{i=1}^n \mathbf{u}^T \mathbf{S}_i \mathbf{u} = \mathbf{u}^T \left(\sum_{i=1}^n \mathbf{S}_i \right) \mathbf{u} = \mathbf{u}^T \mathbf{S} \mathbf{u} \quad (4.25)$$

όπου:

$$\mathbf{S} = \begin{bmatrix} n\mathbf{I} & \mathbf{C} \\ \mathbf{C}^T & n\mathbf{I} \end{bmatrix} \quad (4.26)$$

Ο πίνακας $\mathbf{C} = \sum_{i=1}^n \mathbf{C}_i$ αποτελεί το άθροισμα n ορθογώνιων πινάκων, και στην γενική περίπτωση είναι πλήρους τάξης (rank=4). Πρακτικά είναι πολύ δύσκολο να βρεθεί περίπτωση όπου ο πίνακας \mathbf{C} είναι ιδιόμορφος (rank<4), οπότε στο εξής ο \mathbf{C} θα θεωρείται πάντα πλήρους τάξης.

Ένας τρόπος για την ελαχιστοποίηση της παραπάνω τετραγωνικής συνάρτησης σφάλματος (4.25), είναι χρησιμοποιώντας τους πολλαπλασιαστές *Lagrange* (*Lagrange Multipliers*), ως εξής:

$$\min_{\mathbf{q}_X, \mathbf{q}_Z} F(\mathbf{q}_X, \mathbf{q}_Z) = \min_{\mathbf{q}_X, \mathbf{q}_Z} (f(\mathbf{q}_X, \mathbf{q}_Z) + \lambda_1(1 - \mathbf{q}_X^T \mathbf{q}_X) + \lambda_2(1 - \mathbf{q}_Z^T \mathbf{q}_Z)) \quad (4.27)$$

Οι δύο τελευταίοι όροι της παραπάνω εξίσωσης επιβάλλουν μοναδιαίο μέτρο για τα 4D διανύσματα των αγνώστων \mathbf{q}_X and \mathbf{q}_Z , καθώς αυτά αντιστοιχούν σε μεταβλητές unit quaternion. Χρησιμοποιώντας τις εξισώσεις (4.25), (4.26) και (4.27), καταλήγουμε στην παρακάτω θετικά ορισμένη τετραγωνική συνάρτηση σφάλματος:

$$F(\mathbf{q}_X, \mathbf{q}_Z) = (n - \lambda_1) \mathbf{q}_X^T \mathbf{q}_X + (n - \lambda_2) \mathbf{q}_Z^T \mathbf{q}_Z + \mathbf{q}_X^T \mathbf{C} \mathbf{q}_Z + \mathbf{q}_Z^T \mathbf{C}^T \mathbf{q}_X + \lambda_1 + \lambda_2 \quad (4.28)$$

Η συνάρτηση (4.28) ελαχιστοποιείται για τιμές των αγνώστων \mathbf{q}_X και \mathbf{q}_Z οι οποίες μηδενίζουν τις πρώτες μερικές παραγώγους της, ως εξής:

$$\frac{\partial F}{\partial \mathbf{q}_X} = 2(n - \lambda_1) \mathbf{q}_X + 2\mathbf{C} \mathbf{q}_Z = 0 \quad (4.29)$$

$$\frac{\partial F}{\partial \mathbf{q}_Z} = 2(n - \lambda_2)\mathbf{q}_Z + 2\mathbf{C}^T \mathbf{q}_X = 0 \quad (4.30)$$

Από την (4.29) έχουμε:

$$\mathbf{q}_X = \frac{1}{\lambda_1 - n} \mathbf{C} \mathbf{q}_Z \quad (4.31)$$

και αντικαθιστώντας την \mathbf{q}_X στην (4.30) έχουμε:

$$\mathbf{C}^T \mathbf{C} \mathbf{q}_Z = (\lambda_1 - n)(\lambda_2 - n)\mathbf{q}_Z \quad (4.32)$$

Αντικαθιστώντας τις (4.29) και (4.30) στην αρχική συνάρτηση σφάλματος (4.28) μπορούμε να υπολογίσουμε την ελάχιστη τιμή της συνάρτησης σφάλματος, στα σημεία μηδενισμού των πρώτων μερικών παραγώγων ως προς \mathbf{q}_X και \mathbf{q}_Z . Η ελάχιστη τιμή θα είναι:

$$F(\mathbf{q}_X, \mathbf{q}_Z) = \lambda_1 + \lambda_2 \quad (4.33)$$

Από την (4.32) συμπεραίνουμε ότι το διάνυσμα \mathbf{q}_Z είναι ιδιοδιάνυσμα του συμμετρικού και θετικά ορισμένου πίνακα $\mathbf{C}^T \mathbf{C}$, με αντίστοιχη ιδιοτιμή την $(\lambda_1 - n)(\lambda_2 - n)$. Στην γενική περίπτωση ο πίνακας $\mathbf{C}^T \mathbf{C}$ έχει τέσσερις πραγματικές και θετικές ιδιοτιμές α_i , ($i=1, \dots, 4$) και τέσσερα αντίστοιχα ιδιοδιανύσματα \mathbf{e}_i , οπότε ισχύει:

$$\mathbf{C}^T \mathbf{C} \mathbf{e}_i = \alpha_i \mathbf{e}_i \quad (4.34)$$

Από τις τέσσερις πραγματικές ιδιοτιμές α_i θα πρέπει να επιλέξουμε εκείνη που δίνει την ελάχιστη θετική τιμή στην τετραγωνική συνάρτηση σφάλματος στα σημεία μηδενισμού των μερικών παραγώγων, η οποία δίνεται από την (4.33). Στην πρωτότυπη εργασία [Dornaika and Horaud, 1998] προτείνεται μία μέθοδος, η οποία είναι μαθηματικά σωστή αλλά στην πράξη δημιουργεί προβλήματα σε κάποιες περιπτώσεις. Πιο συγκεκριμένα, προτείνεται η επιλογή της ιδιοτιμής α_i η οποία δίνει την μικρότερη θετική τιμή στην παρακάτω συνάρτηση:

$$\lambda(\alpha_i) = n \pm \sqrt{\alpha_i} \quad (4.35)$$

όπου η μεταβλητή n αντιπροσωπεύει τον αριθμό των διακριτών θέσεων (προσανατολισμών) του άκρου του βραχίονα, από τις οποίες έχουμε λάβει δεδομένα για τις μεταβλητές \mathbf{q}_A και \mathbf{q}_B . Το ιδιοδιάνυσμα που αντιστοιχεί στην συγκεκριμένη ιδιοτιμή λαμβάνεται ως λύση για την άγνωστη μεταβλητή unit quaternion \mathbf{q}_Z . Στην συνέχεια, η άλλη άγνωστη μεταβλητή quaternion \mathbf{q}_X μπορεί εύκολα να γίνει γνωστή από την εξίσωση

(4.31). Στην περίπτωση όμως που έχει επιλεγεί λανθασμένη ιδιοτιμή α_i του πίνακα $\mathbf{C}^T\mathbf{C}$, τότε τα αποτελέσματα για τους αγνώστους \mathbf{q}_X και \mathbf{q}_Z θα είναι τελείως λανθασμένα.

Μία πιο αξιόπιστη μέθοδος για την επιλογή της κατάλληλης ιδιοτιμής α_i και του αντίστοιχου ιδιοδιανύσματος είναι η ελαχιστοποίηση της βασικής μορφής της τετραγωνικής συνάρτησης σφάλματος:

$$f(\mathbf{q}_X, \mathbf{q}_Z) = \sum_{i=1}^n \left\| Q(\mathbf{q}_{A(i)})\mathbf{q}_X - W(\mathbf{q}_{B(i)})\mathbf{q}_Z \right\|^2 \quad (4.36)$$

η οποία προκύπτει από την (4.23) γενικευμένη ώστε να περιλαμβάνει τα δεδομένα για τις μεταβλητές \mathbf{q}_A και \mathbf{q}_B απ' όλες τις n διακριτές θέσεις του άκρου του βραχίονα. Το σωστό ζευγάρι για τις άγνωστες μεταβλητές \mathbf{q}_X και \mathbf{q}_Z είναι αυτό που ελαχιστοποιεί την (4.36).

4.1.3 Μη-γραμμική μέθοδος ελαχιστοποίησης

Στόχος της μη-γραμμικής μεθόδου ελαχιστοποίησης είναι ο ταυτόχρονος υπολογισμός των βέλτιστων τιμών για τους αγνώστους πίνακες στροφής \mathbf{R}_X και \mathbf{R}_Z , με την ελαχιστοποίηση της παρακάτω μη-γραμμικής τετραγωνικής συνάρτησης κόστους:

$$F(\mathbf{R}_X, \mathbf{R}_Z) = \mu_1 \sum_{i=1}^n \left\| \mathbf{R}_{A(i)}\mathbf{R}_X - \mathbf{R}_Z\mathbf{R}_{B(i)} \right\|^2 + \mu_2 \sum_{i=1}^n \left\| \mathbf{R}_X\mathbf{R}_X^T - \mathbf{I} \right\|^2 + \mu_3 \sum_{i=1}^n \left\| \mathbf{R}_Z\mathbf{R}_Z^T - \mathbf{I} \right\|^2 \quad (4.37)$$

όπου $\mathbf{R}_{A(i)}$ και $\mathbf{R}_{B(i)}$ είναι οι γνωστοί 3×3 πίνακες στροφής που προκύπτουν από τον κάθε διακριτό προσανατολισμό του άκρου του βραχίονα ($i=1,2,\dots,n$), ενώ \mathbf{R}_X και \mathbf{R}_Z είναι οι άγνωστοι πίνακες στροφής. Είναι προφανές ότι η (4.37) προκύπτει από την βασική εξίσωση για τους πίνακες στροφής (4.2), γενικευμένη για να περιλαμβάνει τα δεδομένα από όλους τους n διακριτούς προσανατολισμούς του άκρου του βραχίονα. Ο δεύτερος και ο τρίτος όρος της (4.37) επιβάλλουν την ορθοκανονικότητα των 3×3 πινάκων \mathbf{R}_X και \mathbf{R}_Z , έτσι ώστε αυτοί να εξακολουθούν ν' αντιπροσωπεύουν πίνακες στροφής και μετά το πέρας της ελαχιστοποίησης. Οι σταθερές μ_1 , μ_2 και μ_3 εκφράζουν την βαρύτητα του κάθε όρου της (4.37) κατά την διαδικασία της μη-γραμμικής ελαχιστοποίησης. Τυπικές τιμές για τις σταθερές αυτές είναι οι εξής: $\mu_1=1$, $\mu_2=\mu_3=10^{-6}$.

Μία πιο κομψή κι αποτελεσματική μέθοδος για την επιβολή της ορθοκανονικότητας των 3×3 πινάκων \mathbf{R}_X και \mathbf{R}_Z , έτσι ώστε αυτοί ν' αντιπροσωπεύουν πάντα πίνακες

στροφής, είναι η αναπαράστασή τους υπό την μορφή των τριών γωνιών Euler [Fu et al., 1987; Τζαφέστας, 2003]. Οι γωνίες αυτές είναι γραμμικά ανεξάρτητες μεταξύ τους και αναπαριστούν μονοσήμαντα οποιονδήποτε πίνακα στροφής. Η χρήση των γωνιών Euler μας επιτρέπει να μειώσουμε τον αριθμό των ανεξάρτητων μεταβλητών της (4.37) από 18 (9+9 για τους \mathbf{R}_X και \mathbf{R}_Z) σε 6 (3+3 για τις γωνίες Euler), ενώ ταυτόχρονα να επιβάλλουμε την ορθοκανονικότητα των πινάκων \mathbf{R}_X και \mathbf{R}_Z . Έστω $\mathbf{r}_X=(a_X, b_X, c_X)$ οι γωνίες Euler που αντιστοιχούν στον πίνακα στροφής \mathbf{R}_X και $\mathbf{r}_Z=(a_Z, b_Z, c_Z)$ οι γωνίες Euler που αντιστοιχούν στον πίνακα στροφής \mathbf{R}_Z , τότε η μη-γραμμική τετραγωνική συνάρτηση κόστους (4.37) μπορεί να πάρει την παρακάτω μορφή:

$$F(\mathbf{r}_X, \mathbf{r}_Z) = \mu_1 \sum_{i=1}^n \left\| \mathbf{R}_{A(i)} \mathbf{R}_X(\mathbf{r}_X) - \mathbf{R}_Z(\mathbf{r}_Z) \mathbf{R}_{B(i)} \right\|^2 \quad (4.38)$$

όπου

$$\mathbf{R}(\mathbf{r}) = \begin{bmatrix} \cos(a)\cos(b) & \cos(a)\sin(b)\sin(c) - \sin(a)\cos(c) & \cos(a)\sin(b)\cos(c) + \sin(a)\sin(c) \\ \sin(a)\cos(b) & \sin(a)\sin(b)\sin(c) + \cos(a)\cos(c) & \sin(a)\sin(b)\cos(c) - \cos(a)\sin(c) \\ -\sin(b) & \cos(b)\sin(c) & \cos(b)\cos(c) \end{bmatrix} \quad (4.39)$$

είναι η σχέση μετατροπής των τριών γωνιών Euler στον αντίστοιχο πίνακα 3×3 στροφής. Όπως είναι γνωστό, ο πίνακας στροφής που προκύπτει από διαδοχικές περιστροφές γύρω από τους κύριους άξονες του εξεταζόμενου συστήματος συντεταγμένων (γωνίες Euler) εξαρτάται από την αλληλουχία των περιστροφών. Η σχέση (4.39) ισχύει για περιστροφή κατά γωνία a γύρω από τον z -άξονα, ακολουθούμενη από περιστροφή κατά γωνία b γύρω από τον y -άξονα, ακολουθούμενη τέλος από περιστροφή κατά γωνία c γύρω από τον x -άξονα. Σε περίπτωση διαφορετικής σειράς περιστροφών, η σχέση που μετατρέπει τις τρεις γωνίες Euler στον αντίστοιχο πίνακα στροφής είναι διαφορετική.

Η ελαχιστοποίηση της παραπάνω τετραγωνικής μη-γραμμικής συνάρτησης κόστους (4.38) μπορεί να πραγματοποιηθεί με τη χρήση τυπικών μεθόδων μη-γραμμικής ελαχιστοποίησης-βελτιστοποίησης, όπως είναι ο αλγόριθμος των Levenberg-Marquardt [Fletcher, 1990; Gill et al., 1989]. Θεωρητικά, η εφαρμογή του παραπάνω αλγορίθμου απαιτεί την ύπαρξη μιας καλής αρχικής εκτίμησης για τις μεταβλητές προς βελτιστοποίηση (γωνίες Euler \mathbf{r}_X και \mathbf{r}_Z που αντιστοιχούν στους πίνακες στροφής \mathbf{R}_X και \mathbf{R}_Z), έτσι ώστε ο αλγόριθμος να συγκλίνει στο ολικό ελάχιστο της συνάρτησης κόστους (global minimum), που αντιστοιχεί στις βέλτιστες τιμές για τις προς βελτιστοποίηση μεταβλητές. Μια καλή αρχική εκτίμηση για τις μεταβλητές \mathbf{r}_X και \mathbf{r}_Z μπορεί να γίνει

Επεκτείνοντας την εξίσωση (4.40) ώστε να περιλαμβάνει όλα τα διαθέσιμα δεδομένα από τις n διαφορετικές θέσεις του άκρου του βραχίονα μπορούμε να καταλήξουμε σ' ένα σύστημα $3 \times n$ γραμμικών εξισώσεων με έξι αγνώστους, της παρακάτω μορφής:

$$\begin{bmatrix} \mathbf{F}_1 \\ \mathbf{F}_2 \\ \vdots \\ \mathbf{F}_n \end{bmatrix}_{3n \times 6} \cdot \begin{bmatrix} \mathbf{t}_X \\ \mathbf{t}_Z \end{bmatrix}_{6 \times 1} = \begin{bmatrix} \mathbf{d}_1 \\ \mathbf{d}_2 \\ \vdots \\ \mathbf{d}_n \end{bmatrix}_{3n \times 1} \quad (4.42)$$

Η βέλτιστη λύση για το διάνυσμα των αγνώντων (υπό την έννοια των ελάχιστων τετραγώνων) μπορεί πολύ εύκολα να γίνει γνωστή με την χρήση του ψευδοαντίστροφου πίνακα ή του αλγορίθμου SVD [Press et al., 1988; Golub and Van Loan, 1983]. Έχοντας υπολογίσει τους πίνακες στροφής και τα διανύσματα μετατόπισης μπορούμε εύκολα να εξάγουμε τους ζητούμενους μετασχηματισμούς \mathbf{X} και \mathbf{Z} .

4.2 Ταυτόχρονη βαθμονόμηση βραχίονα, μηχανικής όρασης και χώρου εργασίας για στερεοσκοπική κεφαλή

Για την επίλυση του προβλήματος της ταυτόχρονης βαθμονόμησης βραχίονα, μηχανικής όρασης και χώρου εργασίας για την περίπτωση χρήσης μιας μονής κάμερας απαιτείται η εκ των προτέρων γνώση των μετασχηματισμών \mathbf{B}_i και \mathbf{A}_i για κάθε i διακριτή θέση του άκρου του βραχίονα, όπου $i=1,2,3,\dots,n$. Όπως έχει ήδη αναφερθεί, ο μετασχηματισμός \mathbf{B}_i αποτελεί λύση του ευθέως κινηματικού προβλήματος του βραχίονα, και παρέχεται αυτούσιος από τον ρομποτικό ελεγκτή. Ο μετασχηματισμός \mathbf{A}_i μπορεί να βρεθεί αν χρησιμοποιήσουμε ψηφιακές εικόνες από την κάθε i θέση του άκρου του βραχίονα για εφαρμόσουμε τον αλγόριθμο βαθμονόμησης κάμερας (camera calibration) και κρατήσουμε τον πίνακα εξωτερικών παραμέτρων για την κάθε θέση. Οι περισσότεροι αλγόριθμοι βαθμονόμησης κάμερας υπολογίζουν αρχικά τους πίνακες προβολής (projection matrices ή homographies, βλ. Κεφ. 2.3.1) για όλες τις n διαθέσιμες θέσεις της κάμερας, και στην συνέχεια επιχειρούν να τους διασπάσουν σ' έναν μοναδικό πίνακα εσωτερικών παραμέτρων και n πίνακες εξωτερικών παραμέτρων. Η διαδικασία αυτή εισάγει μεγάλες ανακρίβειες [Gonzalez et al., 2005; Sun and Cooperstock, 2006], με αποτέλεσμα η φυσική σημασία των τιμών των πινάκων (εστιακή απόσταση, κέντρο προβολής, μετασχηματισμοί συντεταγμένων, κτλ.) να μην είναι πάντοτε σωστή. Αυτός

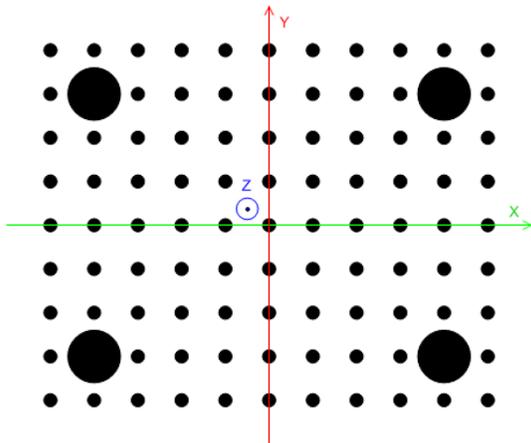
είναι και ο κύριος λόγος για τον οποίο οι πίνακες εξωτερικών παραμέτρων, οι οποίοι χρησιμοποιούνται αυτούσιοι ως μετασχηματισμοί συντεταγμένων \mathbf{A}_i για την επίλυση του προβλήματος της ταυτόχρονης βαθμονόμησης βραχίονα, μηχανικής όρασης και χώρου εργασίας, έχουν περιορισμένη ακρίβεια.

Μια μέθοδος για την αντιμετώπιση του παραπάνω προβλήματος είναι η χρησιμοποίηση των πινάκων προβολής στους υπολογισμούς αντί των ανακριβών πινάκων εξωτερικών παραμέτρων. Η προσέγγιση αυτή έχει χρησιμοποιηθεί επιτυχώς για την επίλυση του προβλήματος της βαθμονόμησης βραχίονα – μηχανικής όρασης (Κεφ. 3), όπου χρησιμοποιούνται οι πίνακες προβολής από διάφορες θέσεις μιας μονής κάμερας για τον υπολογισμό του άγνωστου μετασχηματισμού \mathbf{X} (Hand-Eye). Δυστυχώς δεν υπάρχει αντίστοιχη μεθοδολογία για την επίλυση του προβλήματος της ταυτόχρονης βαθμονόμησης βραχίονα, μηχανικής όρασης και χώρου εργασίας, όπου απαιτείται ο ταυτόχρονος υπολογισμός των αγνώστων μετασχηματισμών \mathbf{X} (Hand-Eye) και \mathbf{Z} (Robot-World).

Στην παρούσα εργασία περιγράφεται μια καινούρια προσέγγιση για την επίλυση του προβλήματος της ταυτόχρονης βαθμονόμησης βραχίονα, μηχανικής όρασης και χώρου εργασίας, όπου γίνεται χρήση μιας στερεοσκοπικής κεφαλής (Stereo Head) αντί μιας μονής κάμερας. Η νέα μέθοδος στοχεύει στην επίλυση του προβλήματος ανακρίβειας των μετασχηματισμών \mathbf{A}_i , η οποία προφανώς έχει σημαντική επίπτωση στην ποιότητα των αποτελεσμάτων για τους άγνωστους μετασχηματισμούς \mathbf{X} και \mathbf{Z} . Είναι προφανές ότι, ανεξαρτήτως της χρησιμοποιούμενης μεθόδου (γραμμικής, κλειστής μορφής ή μη-γραμμικής), η ακρίβεια των αποτελεσμάτων βασίζεται στην ποιότητα των δεδομένων εισόδου \mathbf{A}_i και \mathbf{B}_i . Στόχος μας είναι η χρήση μετρήσεων τρισδιάστατης ανακατασκευής (βλ. Κεφ. 2.5) για την εξαγωγή του ζητούμενου μετασχηματισμού \mathbf{A}_i , για κάθε $i=1,2,\dots,n$ διακριτή θέση του άκρου του βραχίονα. Όπως αναλύθηκε στο Κεφ. 2.5, δεδομένων των 2D συντεταγμένων εικόνας ενός παρατηρούμενου σημείου (προερχόμενων από αλγορίθμους επεξεργασίας εικόνας) και ενός γνωστού πίνακα προβολής \mathbf{H} για μια δεδομένη θέση της κάμερας, είναι δυνατός ο υπολογισμός μιας 3D εξίσωσης ευθείας (οπτική ευθεία) πάνω στην οποία θα βρίσκεται το παρατηρούμενο σημείο. Για την πλήρη 3D ανακατασκευή των συντεταγμένων $\mathbf{P}(x, y, z)$ του σημείου είναι απαραίτητη η χρήση και μιας δεύτερης κάμερας, οπότε το ζητούμενο σημείο θα βρίσκεται στο σημείο τομής

των οπτικών ευθειών των δύο καμερών. Η ακρίβεια υπολογισμού των εξισώσεων των οπτικών ευθειών για τις δύο κάμερες, άρα και η ακρίβεια ανακατασκευής των 3D συντεταγμένων του παρατηρούμενου σημείου \mathbf{P} , εξαρτώνται από την ακρίβεια του πίνακα προβολής της κάθε κάμερας στην παρούσα θέση, καθώς και από την ακρίβεια ανίχνευσης των 2D συντεταγμένων pixel του παρατηρούμενου σημείου πάνω στην κάθε ψηφιακή εικόνα. Αντίθετα, η απόλυτη ακρίβεια των πινάκων εσωτερικών-εξωτερικών παραμέτρων δεν έχει καμία επίπτωση στην ακρίβεια της 3D ανακατασκευής.

Στο πείραμα που ακολουθεί χρησιμοποιείται ένα διαφορετικό επίπεδο μοντέλο βαθμονόμησης (Σχ. 4.2), το οποίο εκτός από τις χαρακτηριστικές κουκίδες που χρησιμοποιούνται για την βαθμονόμηση των καμερών, περιλαμβάνει και τέσσερις

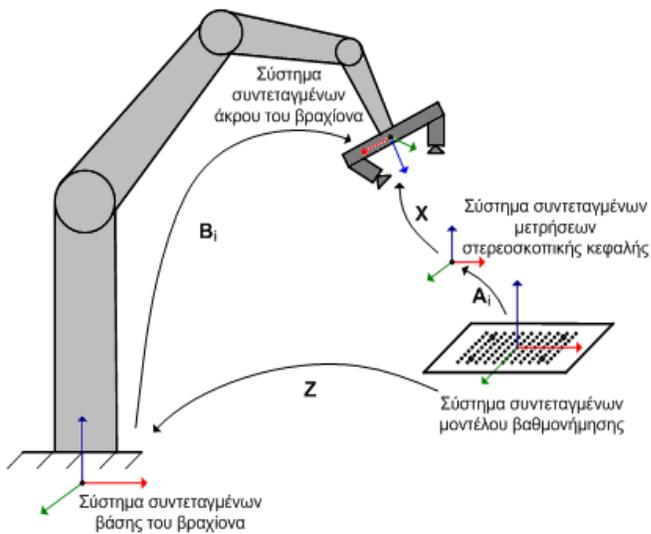


Σχήμα 4.2 Μοντέλο βαθμονόμησης με τέσσερις χαρακτηριστικούς κύκλους για μετρήσεις τρισδιάστατης ανακατασκευής

μεγάλους κύκλους στις γωνίες του. Σκοπός μας είναι να χρησιμοποιήσουμε τις μετρήσεις τρισδιάστατης ανακατασκευής των τεσσάρων κύκλων για την έμμεση εξαγωγή των μετασχηματισμών συντεταγμένων A_i για την κάθε i διακριτή θέση του άκρου του βραχίονα. Υπενθυμίζεται ότι το επίπεδο x - y του συστήματος συντεταγμένων του μοντέλου συμπίπτει με το επίπεδο που βρίσκονται όλα τα χαρακτηριστικά του σημεία (κουκίδες και κύκλοι), ενώ το κέντρο του συμπίπτει με το κέντρο της μεσαίας κουκίδας. Αν υποθέσουμε

ότι οι κουκίδες απέχουν μεταξύ τους οριζόντια και κάθετα απόσταση a mm, τότε οι πραγματικές 3D συντεταγμένες των κέντρων των τεσσάρων κύκλων ως προς το σύστημα συντεταγμένων του μοντέλου θα είναι: $(-4a, -3a, 0)$, $(+4a, -3a, 0)$, $(-4a, +3a, 0)$ και $(+4a, +3a, 0)$ αντίστοιχα.

Η τοπολογία του παρόντος πειράματος είναι ίδια μ' αυτή του Σχ. 4.1, με τη διαφορά ότι χρησιμοποιούμε το μοντέλο βαθμονόμησης του Σχ. 4.2 και μια στερεοσκοπική κεφαλή σταθερά στερεωμένη στο άκρο του βραχίονα αντί για μια μονή κάμερα. Σε μια αρχική θέση του άκρου του βραχίονα ($i=0$), όπου και οι δυο κάμερες είναι σε θέση να βλέπουν το χρησιμοποιούμενο μοντέλο, λαμβάνουμε ψηφιακές εικόνες και εφαρμόζουμε



Σχήμα 4.3 Τοπολογία πειράματος ταυτόχρονης βαθμονόμησης βραχίονα, μηχ. όρασης και χώρου εργασίας για την περίπτωση στερεοσκοπικής κεφαλής

τον αλγόριθμο βαθμονόμησης του R.Tsai (βλ. Κεφ. 2.3.2) για κάθε μία κάμερα ξεχωριστά. Ο λόγος που στο συγκεκριμένο πείραμα προτιμάται ο αλγόριθμος βαθμονόμησης του R.Tsai έναντι του αλγορίθμου του Z.Zhang (βλ. Κεφ. 2.3.1) είναι το γεγονός ότι ο πρώτος μπορεί να λειτουργήσει ακόμα και με μία μόνο ψηφιακή εικόνα από κάθε κάμερα, σε αντίθεση με τον δεύτερο που απαιτεί εικόνες από τουλάχιστο τρεις διαφορετικές θέσεις. Με την εφαρμογή του αλγορίθμου

βαθμονόμησης των καμερών ορίζεται το σύστημα συντεταγμένων μετρήσεων της στερεοσκοπικής κεφαλής, το οποίο στην αρχική αυτή θέση συμπίπτει με το σύστημα συντεταγμένων του μοντέλου βαθμονόμησης. Όπως έχει αναλυθεί και σε προηγούμενες ενότητες, το σύστημα συντεταγμένων μετρήσεων είναι «πακτωμένο» σε σταθερή θέση και προσανατολισμό σε σχέση με την στερεοσκοπική κεφαλή, και ακολουθεί τις εκάστοτε κινήσεις της. Αυτό σημαίνει ότι εάν το άκρο του βραχίονα, άρα και η στερεοσκοπική κεφαλή που είναι σταθερά στερεωμένη πάνω του, απομακρυνθεί από την αρχική του θέση τότε το σύστημα συντεταγμένων μετρήσεων θα πάψει να ταυτίζεται με το σύστημα συντεταγμένων του μοντέλου βαθμονόμησης και θ' ακολουθήσει την κίνηση της στερεοσκοπικής κεφαλής (βλ. Σχ. 4.3).

Δεδομένων των γνωστών πινάκων προβολής για κάθε κάμερα (γινόμενο πινάκων εσωτερικών-εξωτερικών παραμέτρων) μπορούμε να χρησιμοποιήσουμε τις τεχνικές που περιγράφονται στο Κεφ. 2.5 για την πλήρη τρισδιάστατη ανακατασκευή των συντεταγμένων των κέντρων των τεσσάρων κύκλων. Εάν η στερεοσκοπική κεφαλή βρίσκεται στην αρχική θέση βαθμονόμησης ($i=0$) τότε οι μετρήσεις τρισδιάστατης ανακατασκευής θα δώσουν τιμές πολύ κοντά στα ονομαστικά κέντρα των κύκλων $(-4a, -3a, 0)$, $(+4a, -3a, 0)$, $(-4a, +3a, 0)$ και $(+4a, +3a, 0)$, καθώς τα συστήματα συντεταγμένων

μετρήσεων και μοντέλου βαθμονόμησης συμπίπτουν. Εάν μετακινήσουμε το άκρο του βραχίονα σε μια τυχαία θέση κι επαναλάβουμε την διαδικασία, τότε τα αποτελέσματα της τρισδιάστατης ανακατασκευής θα είναι διαφορετικά, καθώς το σύστημα συντεταγμένων μετρήσεων θα βρίσκεται πλέον σε διαφορετική θέση. Η τοπολογία του πειράματος στην τυχαία αυτή θέση του άκρου του βραχίονα, μαζί με τους σχετικούς μετασχηματισμούς συντεταγμένων, απεικονίζεται στο Σχ. 4.3. Συγκρίνοντας τους μετασχηματισμούς του Σχ. 4.3 με τους αντίστοιχους του κλασσικού πειράματος του Σχ. 4.1 μπορούμε να διαπιστώσουμε μερικές διαφορές. Ο μετασχηματισμός \mathbf{X} δεν συνδέει πλέον το σύστημα συντεταγμένων του άκρου του βραχίονα με το κέντρο προβολής κάποιας κάμερας, αλλά με το σύστημα συντεταγμένων μετρήσεων της στερεοσκοπικής κεφαλής. Ο μετασχηματισμός \mathbf{X} είναι σταθερός και ανεξάρτητος της θέσης του άκρου του βραχίονα, καθώς το σύστημα συντεταγμένων μετρήσεων είναι «πακτωμένο» στην στερεοσκοπική κεφαλή, σύμφωνα με τους μετασχηματισμούς των εξωτερικών παραμέτρων των καμερών, συνεπώς και στο άκρο του βραχίονα. Επίσης ο ζητούμενος μετασχηματισμός \mathbf{A}_i δεν συνδέει πλέον το σύστημα συντεταγμένων του ακίνητου μοντέλου βαθμονόμησης μ' εκείνο του κέντρου προβολής της κάμερας, όπως στην τοπολογία του Σχ. 4.1, αλλά με το κινούμενο σύστημα συντεταγμένων μετρήσεων. Από το Σχ. 4.3 είναι προφανές ότι η βασική εξίσωση $\mathbf{A}_i\mathbf{X}=\mathbf{ZB}_i$ του προβλήματος της ταυτόχρονης βαθμονόμησης βραχίονα, μηχανικής όρασης και χώρου εργασίας εξακολουθεί να ισχύει, με αποτέλεσμα να ισχύουν και όλες οι μέθοδοι υπολογισμού των αγνώστων μετασχηματισμών \mathbf{X} και \mathbf{Z} που περιγράφονται στα Κεφ. 4.1.1-4.1.4. Η ιδιαιτερότητα της νέας μεθόδου έγκειται στην χρήση της στερεοσκοπικής κεφαλής για την πραγματοποίηση μετρήσεων τρισδιάστατης ανακατασκευής, με την βοήθεια των οποίων γίνεται ο υπολογισμός των μετασχηματισμών \mathbf{A}_i για κάθε i θέση του άκρου του βραχίονα, όπως περιγράφεται παρακάτω.

Έστω \mathbf{c}_0 , \mathbf{c}_1 , \mathbf{c}_2 και \mathbf{c}_3 οι τρισδιάστατες συντεταγμένες των κέντρων των τεσσάρων κύκλων ως προς το σύστημα συντεταγμένων μετρήσεων της στερεοσκοπικής κεφαλής, όπου \mathbf{c}_0 αντιστοιχεί στον κάτω-αριστερά κύκλο του μοντέλου (βλ. Σχ. 4.2), \mathbf{c}_1 στον κάτω-δεξιά, \mathbf{c}_2 στον πάνω-αριστερά και \mathbf{c}_3 στον πάνω-δεξιά. Ο μετασχηματισμός \mathbf{A}_i^{-1} που ορίζει το σύστημα συντεταγμένων του μοντέλου βαθμονόμησης ως προς εκείνο των μετρήσεων της στερεοσκοπικής κεφαλής δίνεται από τις παρακάτω σχέσεις:

$$\mathbf{A}^{-1} = \begin{bmatrix} \mathbf{r}_x & \mathbf{r}_y & \mathbf{r}_z & \mathbf{t} \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4.43)$$

$$\mathbf{r}_x = \frac{(\mathbf{c}_1 - \mathbf{c}_0)}{\|\mathbf{c}_1 - \mathbf{c}_0\|}, \quad \mathbf{r}_y = \frac{(\mathbf{c}_2 - \mathbf{c}_0)}{\|\mathbf{c}_2 - \mathbf{c}_0\|}, \quad \mathbf{r}_z = \mathbf{r}_x \times \mathbf{r}_y, \quad \mathbf{t} = \frac{(\mathbf{c}_0 + \mathbf{c}_1 + \mathbf{c}_2 + \mathbf{c}_3)}{4}$$

όπου \mathbf{r}_x , \mathbf{r}_y και \mathbf{r}_z είναι οι κύριοι άξονες του συστήματος συντεταγμένων του μοντέλου βαθμονόμησης ως προς το σύστημα συντεταγμένων μετρήσεων. Ο ζητούμενος μετασχηματισμός \mathbf{A}_i του Σχ. 4.3 είναι ο αντίστροφος απ' αυτόν που υπολογίζεται βάσει της (4.43) και μπορεί να υπολογισθεί από μια απλή αντιστροφή πινάκων.

Μια πιο ακριβής και κομψή μέθοδος υπολογισμού του μετασχηματισμού \mathbf{A}_i από τις 3D μετρήσεις ανακατασκευής των κέντρων των κύκλων συνίσταται στην ελαχιστοποίηση της παρακάτω συνάρτησης κόστους:

$$f(\mathbf{q}) = \frac{1}{N} \sum_{i=1}^N \|\mathbf{x}_i - \mathbf{R}(\mathbf{q}_R)\mathbf{c}_i - \mathbf{q}_T\|^2 \quad (4.44)$$

όπου $N=4$ (για τους τέσσερις κύκλους), \mathbf{c}_i είναι οι μετρούμενες τρισδιάστατες συντεταγμένες των κέντρων και \mathbf{x}_i είναι τα ονομαστικά κέντρα ως προς το σύστημα συντεταγμένων του μοντέλου. Ελαχιστοποιώντας την παραπάνω συνάρτηση με τη βοήθεια της αναλυτικής λύσης του [Horn, 1987] βρίσκουμε τις μεταβλητές *Unit Quaternions* \mathbf{q}_R που αντιστοιχούν στο πίνακα στροφής του ζητούμενου μετασχηματισμού \mathbf{A}_i , καθώς και το αντίστοιχο διάνυσμα μετατόπισης \mathbf{q}_T .

4.3 Πειραματικά αποτελέσματα

Στη συνέχεια θα παρουσιαστούν ορισμένα πειραματικά αποτελέσματα, τα οποία αποδεικνύουν την υπεροχή της προτεινόμενης μεθόδου του Κεφ. 4.2 έναντι της κλασσικής μεθόδου του Κεφ. 4.1. Για την εφαρμογή των εργαστηριακών πειραμάτων, η τοπολογία των οποίων απεικονίζονται στα Σχ. 4.1 και Σχ. 4.2, έχει χρησιμοποιηθεί ένας βιομηχανικός ρομποτικός βραχίονας τύπου KR15/1 της KUKA, ο οποίος ελέγχεται από έναν ρομποτικό ελεγκτή τύπου KRC1. Όλοι οι αλγόριθμοι που περιγράφονται στα Κεφ. 4.1.1-4.1.4, καθώς και οι αλγόριθμοι επεξεργασίας εικόνας, έχουν υλοποιηθεί σε γλώσσα προγραμματισμού C++ σε περιβάλλον Microsoft Visual Studio 6.0, το οποίο είναι εγκατεστημένο σ' έναν τυπικό ηλεκτρονικό υπολογιστή (συμβατό IBM PC). Για την

εφαρμογή του αλγορίθμου βαθμονόμησης κάμερας του Tsai έχει χρησιμοποιηθεί η βιβλιοθήκη [<http://www.cs.cmu.edu/~rgw/TsaiCode.html>], ενώ για τον αντίστοιχο του Zhang έχει χρησιμοποιηθεί το εκτελέσιμο αρχείο Easycalib [<http://research.microsoft.com/en-us/um/people/zhang/calib>]. Στο PC είναι εγκατεστημένη μια PCI κάρτα γραφικών Matrox Meteor II MC (Frame Grabber), η οποία μπορεί να λαμβάνει εικόνες από δύο ψηφιακές κάμερες Sony HR50 (ανάλυση εικονοστοιχείων 640×495) που είναι συνδεδεμένες πάνω του.

Στους παρακάτω πίνακες 4.1 και 4.2 φαίνονται τα πειραματικά αποτελέσματα για τους ζητούμενους μετασχηματισμούς \mathbf{X} και \mathbf{Z} αντίστοιχα, συναρτήσεως του αριθμού n των χρησιμοποιούμενων διακριτών θέσεων του άκρου του βραχίονα. Για την εξαγωγή των αποτελεσμάτων έχει εφαρμοσθεί η κλασική μέθοδος μονής κάμερας του Κεφ. 4.1, ενώ οι μετασχηματισμοί \mathbf{A}_i έχουν υπολογισθεί με την εφαρμογή του αλγορίθμου βαθμονόμησης του Zhang, ο οποίος στην πράξη δίνει πολύ πιο ακριβή και αξιόπιστα αποτελέσματα συγκρινόμενος με τον αντίστοιχο του Tsai, όσον αφορά την απόλυτη ακρίβεια των πινάκων εξωτερικών παραμέτρων [Gonzalez, et al., 2005]. Από τις διαθέσιμες μεθόδους επίλυσης έχει προτιμηθεί ο αλγόριθμος κλειστής μορφής του Κεφ. 4.1.2, ακολουθούμενος από τον αλγόριθμο μη-γραμμικής βελτιστοποίησης μόνο για τους πίνακες στροφής του Κεφ. 4.1.3. Η γραμμική μέθοδος του Κεφ. 4.1.1 δίνει ελαφρώς χειρότερα αποτελέσματα σε σχέση με την μέθοδο κλειστής μορφής, όσον αφορά το μέσο τετραγωνικό σφάλμα και την ακρίβεια σύγκλισης των αποτελεσμάτων, γι' αυτό και δεν προτιμάται για την πρακτική εφαρμογή. Τέλος, τα διανύσματα μετατόπισης έχουν υπολογιστεί με την μέθοδο που περιγράφεται στο Κεφ. 4.1.4.

Πίνακας 4.1 Πειραματικά αποτελέσματα του μετασχηματισμού \mathbf{X} για την τοπολογία μονής κάμερας

n	$\mathbf{X}(x)$	$\mathbf{X}(y)$	$\mathbf{X}(z)$	$\mathbf{X}(a)$	$\mathbf{X}(b)$	$\mathbf{X}(c)$	Err_Rot
3	116,1689	215,5331	-316,8666	88,6592	0,2100	-0,8667	0,0040
5	116,7216	218,6735	-317,2372	88,7412	0,1619	-0,7845	0,0089
10	117,1908	218,9937	-316,3215	88,7650	0,1821	-0,8558	0,0200
15	117,2966	219,1032	-316,6850	88,7547	0,1651	-0,8452	0,0170
20	117,1863	218,9156	-316,9491	88,7526	0,1640	-0,8305	0,0154
21	117,0937	219,0170	-316,7966	88,7541	0,1645	-0,8318	0,0150
22	117,2424	218,7995	-316,7663	88,7541	0,1650	-0,8372	0,0148

Πίνακας 4.2 Πειραματικά αποτελέσματα του μετασχηματισμού \mathbf{Z} για την τοπολογία μονής κάμερας

n	$\mathbf{Z}(x)$	$\mathbf{Z}(y)$	$\mathbf{Z}(z)$	$\mathbf{Z}(a)$	$\mathbf{Z}(b)$	$\mathbf{Z}(c)$	Err_Trans
3	-137,2049	-1047,8201	-788,7375	104,5308	0,3487	-0,4352	0,4678
5	-135,9269	-1049,3188	-790,1377	104,4503	0,3144	-0,5220	0,8623
10	-136,3225	-1049,5900	-790,8044	104,4514	0,3205	-0,4976	6,6954
15	-135,9408	-1049,5045	-790,5594	104,4616	0,3119	-0,5051	5,6869
20	-135,8758	-1049,4615	-790,2945	104,4590	0,3159	-0,5131	4,9778
21	-135,9698	-1049,5155	-790,4837	104,4577	0,3155	-0,5115	4,9313
22	-135,9795	-1049,3242	-790,3655	104,4580	0,3170	-0,5066	4,7308

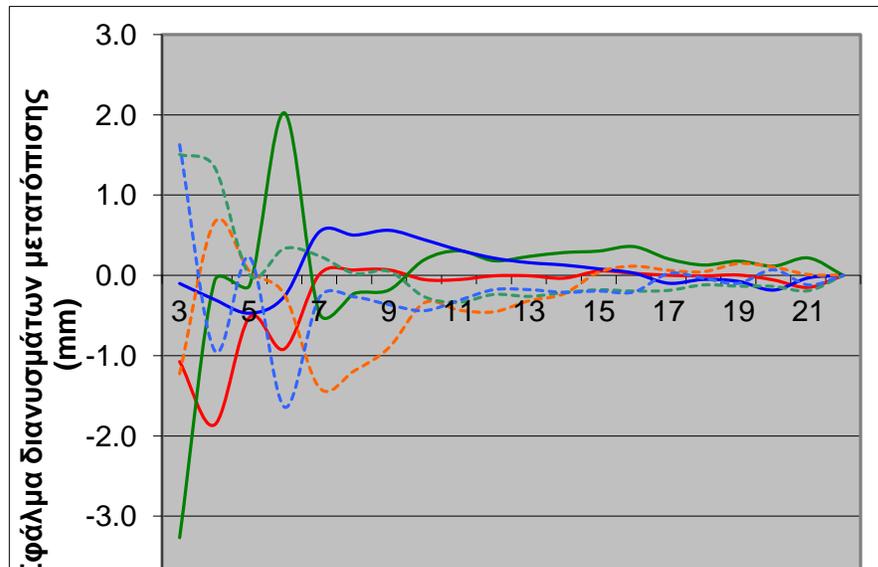
Στους παραπάνω πίνακες 4.1 και 4.2 οι τιμές των διανυσμάτων μετατόπισης δίνονται σε mm, ενώ οι τιμές των γωνιών στροφής Euler δίνονται σε μοίρες. Οι τιμές των μέσων τετραγωνικών σφαλμάτων για τις γωνίες στροφής Err_Rot και για τα διανύσματα μετατόπισης Err_Trans είναι κοινές και για τους δυο άγνωστους μετασχηματισμούς \mathbf{X} και \mathbf{Z} , και δίνονται από τις παρακάτω σχέσεις:

$$\text{Err_Rot} = \sqrt{\frac{\sum_{i=1}^n \|\mathbf{R}_{A(i)} \mathbf{R}_X - \mathbf{R}_Z \mathbf{R}_{B(i)}\|^2}{n}} \quad (4.45)$$

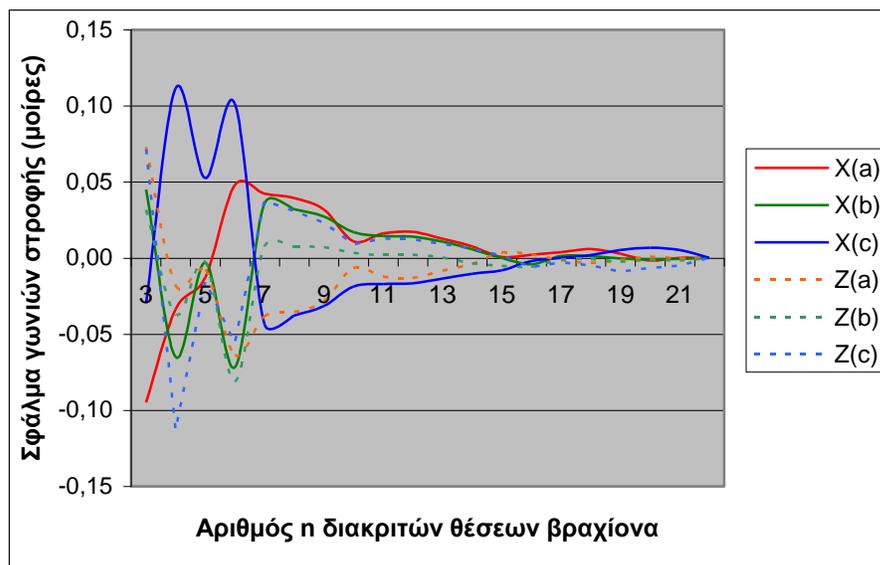
και

$$\text{Err_Trans} = \sqrt{\frac{\sum_{i=1}^n \|\mathbf{R}_{A(i)} \mathbf{t}_X + \mathbf{t}_{A(i)} - \mathbf{R}_Z \mathbf{t}_{B(i)} - \mathbf{t}_Z\|^2}{n}} \quad (4.46)$$

Οι τιμές των μέσων τετραγωνικών σφαλμάτων προέρχονται από τις βασικές εξισώσεις (4.2) και (4.3), γενικευμένες ώστε να περιλαμβάνουν όλα τα διαθέσιμα δεδομένα από τις n διακριτές θέσεις του άκρου του βραχίονα. Αυτές δίνουν μια ενδεικτική τιμή για την ακρίβεια και την αξιοπιστία των αποτελεσμάτων βάσει των δεδομένων εισόδου \mathbf{A}_i και \mathbf{B}_i , για όλα τα $i=1,2,\dots,n$. Τα διαγράμματα των Σχ. 4.4 και 4.5 δείχνουν τον ρυθμό σύγκλισης των αποτελεσμάτων για τους ζητούμενους μετασχηματισμούς \mathbf{X} και \mathbf{Z} αντίστοιχα, συναρτήσει του αριθμού n των διακριτών θέσεων. Η ιδανική τιμή των ζητούμενων μετασχηματισμών, βάσει της οποίας υπολογίζονται τα σφάλματα, βρίσκεται χρησιμοποιώντας έναν μεγάλο αριθμό διακριτών θέσεων του άκρου ($n=100$), όπου τ' αποτελέσματα έχουν συγκλίνει με πολύ μεγάλη ακρίβεια. Παρατηρούμε ότι τ' αποτελέσματα αρχίζουν να συγκλίνουν για αριθμό θέσεων $n>10$, ενώ για $n>20$ επιτυγχάνονται ακρίβειες της τάξης των 0.1mm για τα διανύσματα μετατόπισης και 0.02° για τις γωνίες στροφής.



Σχήμα 4.4 Ρυθμός σύγκλισης των διανυσμάτων μετατόπισης των μετασχηματισμών X και Z για την τοπολογία μονής κάμερας



Σχήμα 4.5 Ρυθμός σύγκλισης των γωνιών στροφής Euler των μετασχηματισμών X και Z για την τοπολογία μονής κάμερας

Στην συνέχεια το πείραμα επαναλαμβάνεται ακολουθώντας την προτεινόμενη μέθοδο του Κεφ. 4.2, χρησιμοποιώντας δεδομένα από μετρήσεις τρισδιάστατης ανακατασκευής που έχουν ληφθεί από τις ίδιες θέσεις του βραχίονα σε σχέση με το προηγούμενο

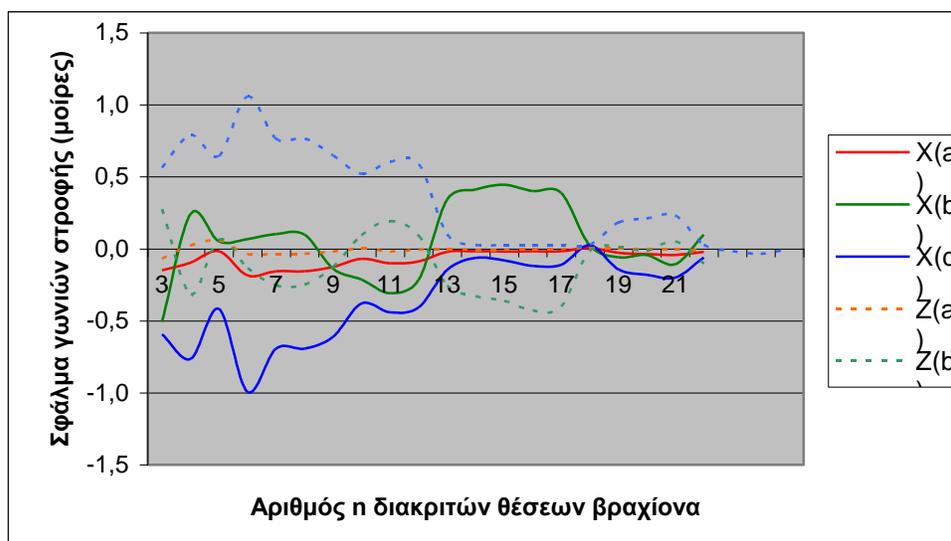
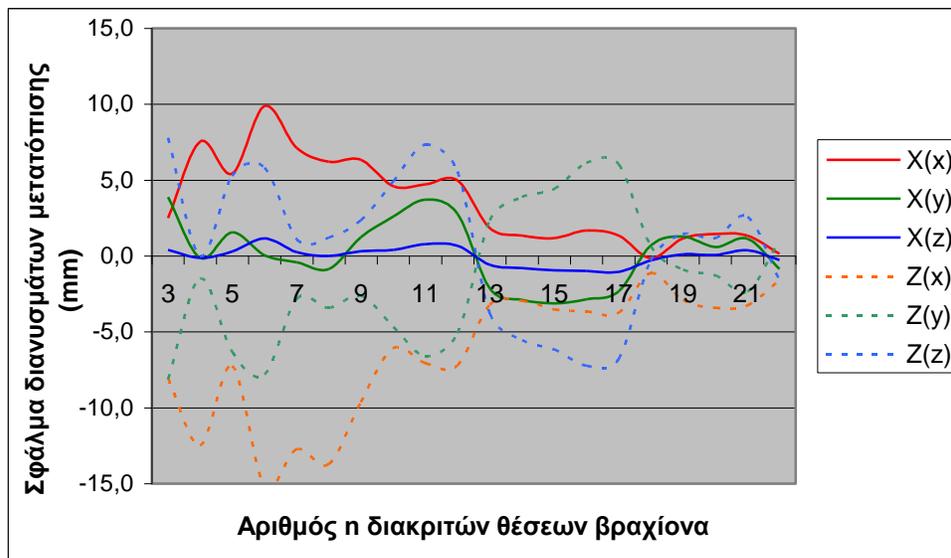
πείραμα. Και στην περίπτωση αυτή, η προτεινόμενη μεθοδολογία επίλυσης είναι η μέθοδος κλειστής μορφής ακολουθούμενη από τη μέθοδο μη-γραμμικής βελτιστοποίησης μόνο για τους πίνακες στροφής. Στους πίνακες 4.3 και 4.4 φαίνονται τα αποτελέσματα για τους ζητούμενους μετασχηματισμούς **X** και **Z** αντίστοιχα, συναρτήσει του αριθμού n των χρησιμοποιούμενων διακριτών θέσεων του άκρου του βραχίονα. Επίσης στα Σχ. 4.6 και 4.7 απεικονίζεται ο ρυθμός σύγκλισης των αποτελεσμάτων συναρτήσει του αριθμού των διακριτών θέσεων n , για τα διανύσματα μετατόπισης και της γωνίες στροφής αντίστοιχα.

Πίνακας 4.3 Πειραματικά αποτελέσματα του μετασχηματισμού **X** για την τοπολογία στερεοσκοπικής κεφαλής

n	X (x)	X (y)	X (z)	X (a)	X (b)	X (c)	Err_Rot
3	-12,5071	-116,4223	516,5104	-94,7324	6,4832	-173,1906	1,1584
5	-9,5816	-118,7446	516,4016	-94,5999	7,0411	-173,0142	1,1093
10	-10,4268	-117,6914	516,5093	-94,6530	6,7760	-172,9763	1,4253
15	-13,8343	-123,4172	515,1684	-94,6011	7,4352	-172,6790	1,2821
20	-13,5661	-119,7048	516,1732	-94,6201	6,9467	-172,7790	1,2911
21	-13,6749	-119,1645	516,4953	-94,6260	6,8810	-172,7990	1,2779
22	-14,8681	-121,1511	515,8331	-94,6035	7,0869	-172,6578	1,2530

Πίνακας 4.4 Πειραματικά αποτελέσματα του μετασχηματισμού **Z** για την τοπολογία στερεοσκοπικής κεφαλής

n	Z (x)	Z (y)	Z (z)	Z (a)	Z (b)	Z (c)	Err_Trans
3	55,3726	-1048,4822	-847,5882	117,0424	1,8017	0,5347	0,0067
5	56,1768	-1046,6939	-850,0883	117,1700	1,5913	0,6151	1,3965
10	57,3550	-1045,0366	-850,4549	117,1133	1,6198	0,4916	2,0856
15	59,9231	-1035,9589	-861,5222	117,1054	1,1638	0,0000	1,6626
20	60,0328	-1041,6297	-854,1891	117,1114	1,5119	0,1834	2,7435
21	60,1618	-1042,7042	-852,7747	117,1070	1,5776	0,2017	2,7771
22	61,8519	-1039,3351	-856,9795	117,1028	1,4241	0,0000	2,4730



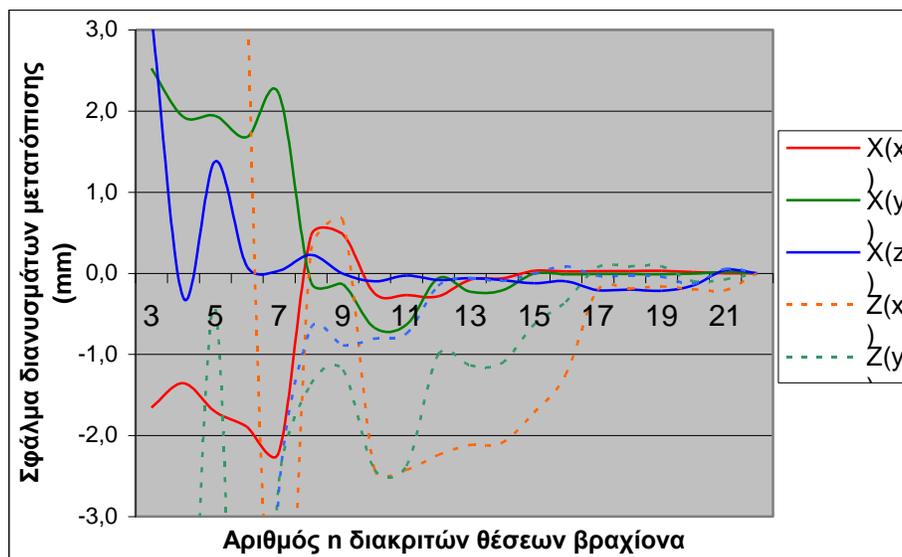
Σχήμα 4.7 Ρυθμός σύγκλισης των γωνιών στροφής Euler των μετασχηματισμών X και Z για την τοπολογία στερεοσκοπικής κεφαλής

Αντίθετα με τις προσδοκίες μας, τα Σχ. 4.4-4.7 μαρτυρούν ότι η κλασική μέθοδος μονής κάμερας συγκλίνει πολύ πιο γρήγορα και με μεγαλύτερη ακρίβεια, συγκρινόμενη με την προτεινόμενη μέθοδο της στερεοσκοπικής κεφαλής. Επίσης, από τους πίνακες 4.1-4.4 είναι φανερό ότι παρόλο που η προτεινόμενη μέθοδος παρουσιάζει μικρότερο μέσο τετραγωνικό σφάλμα για τα διανύσματα μετατόπισης σε σχέση με το αντίστοιχο της κλασικής μεθόδου, αντιθέτως το μέσο τετραγωνικό σφάλμα για τις γωνίες στροφής

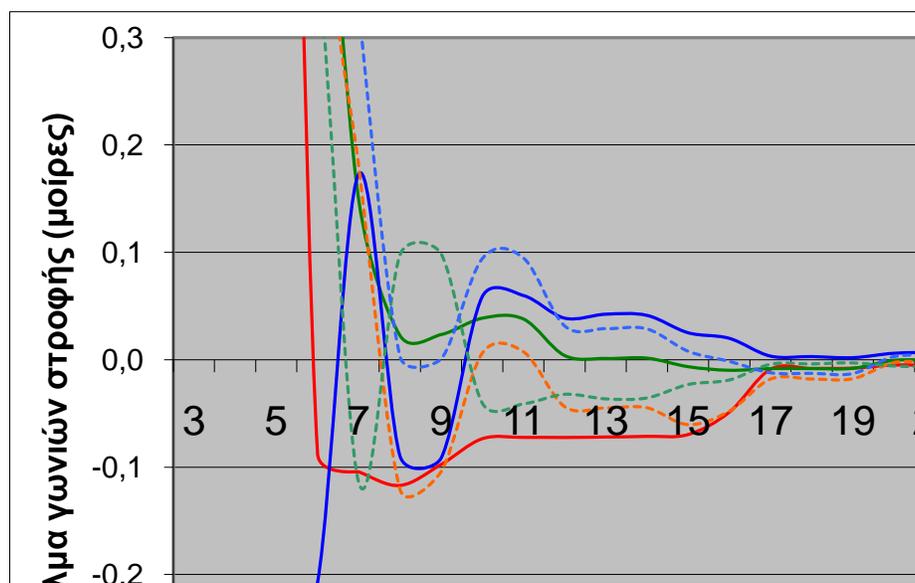
είναι πολύ μεγαλύτερο. Η μειωμένη ακρίβεια των αποτελεσμάτων της προτεινόμενης μεθόδου δικαιολογείται από την μεγάλη ανακρίβεια των μετρήσεων τρισδιάστατης ανακατασκευής σε θέσεις της στερεοσκοπικής κεφαλής μακριά από την αρχική θέση όπου πραγματοποιήθηκε η βαθμονόμηση των καμερών. Πράγματι, οι πίνακες προβολής των καμερών (γινόμενο εσωτερικών-εξωτερικών παραμέτρων) που υπολογίζονται από τον αλγόριθμο βαθμονόμησης εγγυώνται την πραγματοποίηση ακριβών μετρήσεων τρισδιάστατης ανακατασκευής, μόνο όταν σχετική θέση της στερεοσκοπικής κεφαλής και του παρατηρούμενου αντικειμένου βρίσκεται κοντά στην αρχική σχετική θέση της κεφαλής και του μοντέλου βαθμονόμησης, όπου πραγματοποιήθηκε η βαθμονόμηση των καμερών. Στην χρησιμοποιούμενη τοπολογία του πειράματος (Σχ. 4.3) που το μοντέλο βαθμονόμησης είναι ακίνητο, η ακρίβεια των μετρήσεων τρισδιάστατης ανακατασκευής μειώνεται ανάλογα με το μέγεθος της μετακίνησης της στερεοσκοπικής κεφαλής από την αρχική της θέση βαθμονόμησης.

Ο παραπάνω περιορισμός μας ωθεί να επαναλάβουμε το τελευταίο πείραμα, πραγματοποιώντας μικρότερες κινήσεις του άκρου του βραχίονα, έτσι ώστε ν' αναγκάζουμε την στερεοσκοπική κεφαλή να μην απομακρύνεται πολύ από την αρχική της θέση βαθμονόμησης. Πάντως οι κινήσεις του βραχίονα θα πρέπει να είναι επαρκώς μεγάλες, ώστε το εύρος των τιμών για τους μετασχηματισμούς εισόδου \mathbf{A}_i και \mathbf{B}_i να είναι αρκετά μεγάλο για να οδηγήσει στην σύγκλιση των αποτελεσμάτων. Πράγματι τα αποτελέσματα του τελευταίου πειράματος επιβεβαιώνουν τους ισχυρισμούς μας, και οι τιμές των μέσων τετραγωνικών σφαλμάτων για τα διανύσματα μετατόπισης (0.322459) και για τις γωνίες στροφής (0.014) των μετασχηματισμών \mathbf{X} και \mathbf{Z} είναι μικρότερες από κάθε προηγούμενο πείραμα. Οι τιμές των αποτελεσμάτων στα τελευταία βήματα σύγκλισης είναι πολύ κοντά στις τιμές σύγκλισης του προηγούμενου πειράματος (Πίνακες 4.3 και 4.4), καθώς οι θέσεις της στερεοσκοπικής κεφαλής και του μοντέλου βαθμονόμησης δεν έχουν αλλάξει. Στα Σχ. 4.8 και 4.9 απεικονίζεται ο ρυθμός σύγκλισης για τα διανύσματα μετατόπισης και τις γωνίες στροφής αντίστοιχα, συναρτήσει του αριθμού n των χρησιμοποιούμενων διακριτών θέσεων του άκρου του βραχίονα. Μοναδικό μειονέκτημα αποτελεί το γεγονός ότι η σύγκλιση των αποτελεσμάτων απαιτεί μεγαλύτερο αριθμό δεδομένων ($n > 15$) ώστε να επιτευχθεί ακρίβεια της τάξης των 0.1mm για τα διανύσματα μετατόπισης και 0.01° για τις γωνίες στροφής. Αυτό δικαιολογείται

από το μικρότερο εύρος τιμών των δεδομένων εισόδου A_i και B_i σε σχέση με τις αντίστοιχες του αρχικού πειράματος τοπολογίας μονής κάμερας, όπου οι μετακινήσεις του άκρου του βραχίονα ήταν μεγαλύτερες.



Σχήμα 4.8 Ρυθμός σύγκλισης των διανυσμάτων μετατόπισης των μετασχηματισμών X και Z για την τοπολογία στερεοσκοπικής κεφαλής και μικρές μετατοπίσεις του βραχίονα



Σχήμα 4.9 Ρυθμός σύγκλισης των γωνιών στροφής Euler των μετασχηματισμών X και Z για την τοπολογία στερεοσκοπικής κεφαλής και μικρές μετατοπίσεις του βραχίονα

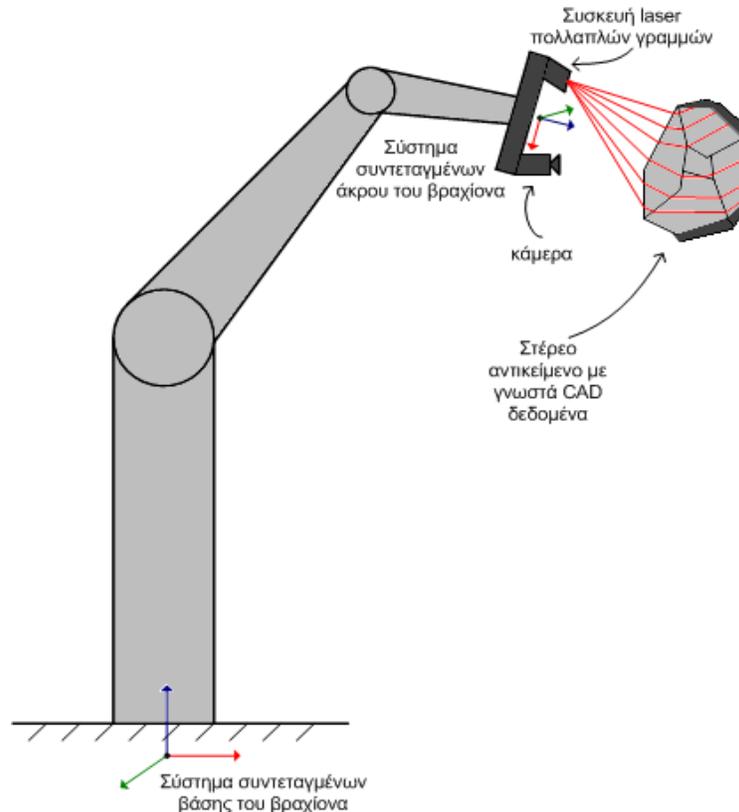
5 Έλεγχος τροχιάς ρομποτικού βραχίονα με τη βοήθεια κάμερας και συσκευή λέιζερ

5.1 Γενική περιγραφή προτεινόμενου συστήματος ελέγχου

Στο κεφάλαιο αυτό θα παρουσιαστεί η μεθοδολογία σχεδίασης και υλοποίησης ενός συστήματος ελέγχου τροχιάς του άκρου ενός βιομηχανικού ρομποτικού βραχίονα, χρησιμοποιώντας πληροφορία ανάδρασης από μια συστοιχία ψηφιακής κάμερας και συσκευής λέιζερ (laser) σταθερά στερεωμένης στο άκρο του βραχίονα. Η συσκευή ακτινών φωτός εκπέμπει πολλαπλές γραμμές λέιζερ οι οποίες προσπίπτουν πάνω στην επιφάνεια ενός παρατηρούμενου αντικειμένου κατά τέτοιο τρόπο ώστε να γίνονται ορατές από την ψηφιακή κάμερα. Η συστοιχία ψηφιακής κάμερας και συσκευής ακτινών λέιζερ αναφέρεται στην βιβλιογραφία [Blais, 2004] ως *Laser-triangulation range sensor* εξαιτίας της τριγωνικής γεωμετρίας της συστοιχίας. Η περιγραφόμενη συστοιχία μέτρησης και γενικότερα η τοπολογία του πειράματος γίνονται καλύτερα αντιληπτές με τη βοήθεια του Σχ. 5.1.

Σκοπός του προτεινόμενου συστήματος ελέγχου είναι το άκρο του βραχίονα να παρακολουθεί τις μεταβολές θέσης του παρατηρούμενου στέρεου αντικειμένου, έτσι ώστε η σχετική θέση και προσανατολισμός του άκρου και του αντικειμένου να είναι πάντα η ίδια. Στην πράξη μια τέτοια διάταξη είναι πολύ χρήσιμη, καθώς ο βραχίονας μπορεί να εκτελεί κάποιες διεργασίες πάνω σε συγκεκριμένα σημεία της επιφάνειας του αντικειμένου (πχ, διάνοιξη οπών, ηλεκτροσυγκόλληση, κτλ.) ανεξαρτήτως της θέσης που αυτό τυγχάνει να βρίσκεται. Εφόσον ο ελεγκτής λειτουργεί σωστά και με ακρίβεια, το ρομπότ θα έχει τη δυνατότητα να παρακολουθεί και ν' αντισταθμίζει τις εκάστοτε κινήσεις του αντικειμένου στο χώρο, με αποτέλεσμα να εκτελεί τις προγραμματισμένες διεργασίες πάνω στην επιφάνειά του με τον ίδιο ακριβώς τρόπο, ανεξαρτήτως της (μεταβαλλόμενης) θέσης του αντικειμένου. Χωρίς τον προτεινόμενο ελεγκτή κλειστού βρόχου (βλ. Σχ. 1.1 & 1.2), η παραπάνω διεργασία θα ήταν εφικτή μόνο αν το

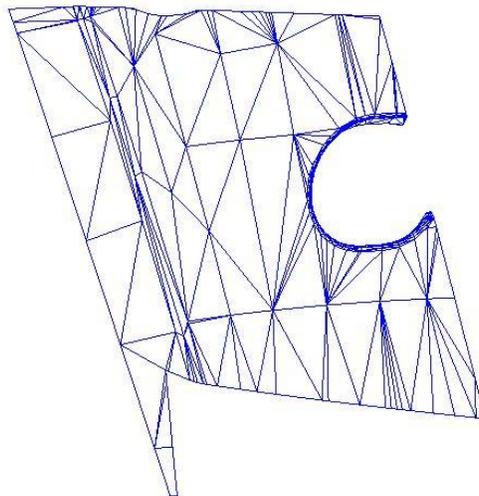
αντικείμενο ήταν ακίνητο στο χώρο, οπότε το ρομπότ θα εκτελούσε τις επιθυμητές διεργασίες πάντα στα ίδια ακριβώς σημεία της επιφάνειας του αντικειμένου.



Σχήμα 5.1 Τοπολογία πειράματος ελέγχου τροχιάς ρομποτικού βραχίονα με ανάδραση από συστοιχία κάμερας - συσκευής ακτινών λέιζερ πολλαπλών γραμμών.

Όπως αναφέρθηκε και στο 1^ο κεφάλαιο, οι εφαρμογές που στοχεύουν στην παρακολούθηση τροχιάς ενός κινούμενου στον χώρο παρατηρούμενου αντικειμένου από το άκρο ενός ρομποτικού βραχίονα ομαδοποιούνται κάτω από την κοινή ονομασία *Visual Servoing Systems* [Hutchinson et al., 1996]. Υπάρχουν δύο κύριες κατηγορίες των αναφερόμενων συστημάτων ανάλογα με τον χώρο που ανήκει το διάλυσμα σφάλματος του συστήματος ελέγχου, ήτοι τα συστήματα όπου το σφάλμα ανήκει στον δισδιάστατο χώρο των ψηφιακών εικόνων (*image-based visual servoing systems*) και τα συστήματα όπου το σφάλμα ανήκει στον πραγματικό Ευκλείδειο τρισδιάστατο χώρο (*position-based visual servoing systems*). Τα τελευταία χρόνια πάντως έχουν προταθεί υβριδικά συστήματα ελέγχου που συνδυάζουν χαρακτηριστικά και από τις δύο παραπάνω κατηγορίες (*hybrid visual servoing systems*) όπως η εφαρμογή αυτόματης

ηλεκτροσυγκόλλησης που προτείνεται από τους [Xu et al., 2005]. Στην βιβλιογραφία δεν υπάρχουν πολλές εφαρμογές παρακολούθησης τροχιάς κινούμενων αντικειμένων που χρησιμοποιούν ανάδραση ελέγχου είτε από συστοιχίες ψηφιακών καμερών και συσκευών λέιζερ (laser-triangulation sensors) είτε από συσκευές προβολής δομημένου φωτός (structured light systems). Η εφαρμογή που προτείνεται από τους [Khadraoui et al., 1996] χρησιμοποιεί ως οπτική ανάδραση τις ψηφιακές εικόνες μιας κάμερας, η οποία παρακολουθεί τις προβολές δύο κάθετων μεταξύ τους γραμμών λέιζερ που προσπίπτουν πάνω σε μια παρατηρούμενη σφαιρική επιφάνεια. Στην μέθοδο που προτείνεται από τους [Pagés et al., 2005] περιγράφεται ένα σύστημα ελέγχου τριών βαθμών ελευθερίας ενός ρομποτικού βραχίονα χρησιμοποιώντας τέσσερις παράλληλες και συμμετρικά τοποθετημένες κουκίδες λέιζερ (laser pointers). Μια άλλη εφαρμογή [Pagés et al., 2006] χρησιμοποιεί έναν προβολέα δομημένου φωτός με συγκεκριμένους κωδικούς χρωμάτων (*coded structured light*) για ν' ανιχνεύσει την μορφή και την θέση αντικειμένων χωρίς ιδιαίτερα χαρακτηριστικά υφής, όπως γωνίες, οπές, ακμές, κτλ.



Σχήμα 5.2 Μοντέλο CAD στέρεου σώματος.

Η μέθοδος ελέγχου τροχιάς ρομποτικού βραχίονα που προτείνεται στο κεφάλαιο αυτό ακολουθεί διαφορετική προσέγγιση σε σχέση με τις προαναφερθείσες εφαρμογές. Για την μέτρηση της θέσης και προσανατολισμού του παρατηρούμενου αντικειμένου χρησιμοποιούνται τα CAD δεδομένα του (*Computer-aided design*), τα οποία συνήθως παρέχονται από τον κατασκευαστή και αποτελούνται από ένα πυκνό πλέγμα

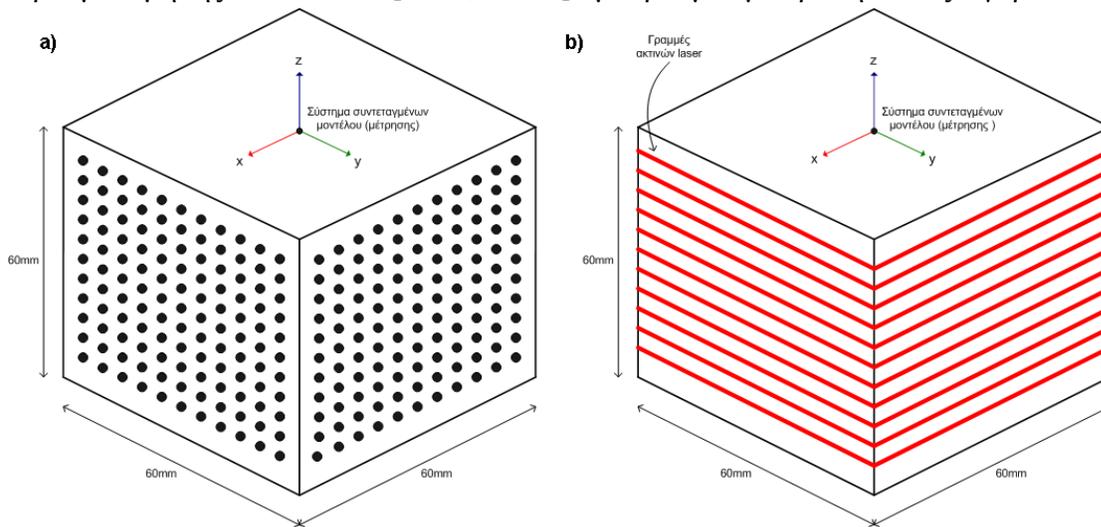
τρισδιάστατων τριγώνων που σχηματίζουν την τοπολογία της επιφάνειας του αντικειμένου στον χώρο (βλ. Σχ. 5.2). Στόχος της προτεινόμενης μεθόδου ελέγχου τροχιάς είναι η μέτρηση της θέσης και προσανατολισμού του παρατηρούμενου αντικειμένου χρησιμοποιώντας τα CAD δεδομένα του, και η αναγωγή της μέτρησης σε ένα σύστημα συντεταγμένων που είναι γνωστό στο ρομποτικό βραχίονα. Είναι προφανές ότι το προτεινόμενο σύστημα ελέγχου θ' ανήκει στην κατηγορία των *συστημάτων παρακολούθησης τροχιάς βάσει συντεταγμένων (position-based visual servoing)*, εξαιτίας της εν γένει τρισδιάστατης υπόστασης του χρησιμοποιούμενου CAD μοντέλου.

5.2 Βαθμονόμησης συστοιχίας κάμερας - συσκευής λέιζερ

Η λειτουργία οποιουδήποτε συστήματος παρακολούθησης τροχιάς βάσει συντεταγμένων απαιτεί την εκ των προτέρων βαθμονόμηση της χρησιμοποιούμενης κάμερας (camera calibration) έτσι ώστε να γίνει δυνατή η συσχέτιση των συντεταγμένων των ψηφιακών εικόνων (pixel coordinates) με τις αντίστοιχες πραγματικές συντεταγμένες στο χώρο (x, y, z) ενός παρατηρούμενου σημείου. Η χρησιμοποίηση της συσκευής γραμμών λέιζερ σε συνδυασμό με την ψηφιακή κάμερα για την πραγματοποίηση μετρήσεων απαιτεί επιπλέον την βαθμονόμηση της συσκευής σε σχέση με την κάμερα, διαδικασία που μεταφράζεται στην εύρεση των εξισώσεων επιπέδων που σχηματίζουν οι γραμμές φωτός ως προς το σύστημα συντεταγμένων μέτρησης (βαθμονόμησης) της κάμερας. Στην ενότητα αυτή θα δοθεί μια αναλυτική περιγραφή της διαδικασίας βαθμονόμησης της συστοιχίας κάμερας-συσκευής λέιζερ (triangulation-based laser range sensor).

Όπως αναφέρθηκε στο 2ο κεφάλαιο, η βαθμονόμησης μιας ψηφιακής κάμερας απαιτεί την ύπαρξη ενός μοντέλου βαθμονόμησης (calibration pattern) με χαρακτηριστικά σημεία που βρίσκονται σε γνωστές θέσεις ως προς ένα σταθερό σύστημα συντεταγμένων στον τρισδιάστατο χώρο. Για την παρούσα εφαρμογή επιλέξαμε ένα μη-επίπεδο μοντέλο βαθμονόμησης κυβικού σχήματος, όπως αυτό που απεικονίζεται στο Σχ. 5.3α, το οποίο περιλαμβάνει πλέγματα 11×11 κουκίδων σε δύο επιφάνειές του. Η χρησιμοποίηση ενός μη-επίπεδου μοντέλου βαθμονόμησης είναι απαραίτητη για να γίνει δυνατή η ταυτόχρονη βαθμονόμηση των γραμμών λέιζερ, διαδικασία που θα αναλυθεί

εκτενώς στην συνέχεια. Η διάσταση της κάθε πλευράς του κύβου είναι 60mm και οι κουκίδες απέχουν μεταξύ τους οριζόντια και κάθετα κατά 5mm. Δεδομένου ότι το σύστημα συντεταγμένων του μοντέλου επιλέγεται στο κέντρο του κύβου και με άξονες παράλληλους στις πλευρές του, οι τρισδιάστατες συντεταγμένες θέσης όλων των κουκίδων ως προς το αναφερθέν σύστημα συντεταγμένων μπορούν να γίνουν γνωστές. Για παράδειγμα η πάνω σειρά κουκίδων στην αριστερή επιφάνεια του κύβου (βλ. Σχ. 5.3α) θα έχει συντεταγμένες (x, y, z) : $(30, -25, 25)$, $(30, -20, 25)$, $(30, -15, 25)$, ... , $(30, 25, 25)$. Ουσιαστικά μεταβάλλεται μόνο η τεταγμένη y , καθώς το επίπεδο της αριστερής επιφάνειας του κύβου είναι $x=30\text{mm}$ και το z είναι το ίδιο για όλη την πάνω σειρά των κουκίδων. Στη συνέχεια παίρνουμε μια ψηφιακή εικόνα του μοντέλου βαθμονόμησης, αφού πρώτα το τοποθετήσουμε μπροστά στην κάμερα. Χρησιμοποιώντας τεχνικές επεξεργασίας εικόνας [Shortis et al., 1994] μπορούμε να εντοπίσουμε τα κέντρα των κουκίδων πάνω στην ψηφιακή εικόνα, και να εξαγάγουμε τις 2D συντεταγμένες εικόνας που αντιστοιχούν σε αυτά. Χρησιμοποιώντας τον βελτιστοποιημένο αλγόριθμο βαθμονόμησης του Tsai [Horn, 2000] μπορούμε με βάση τα ζευγάρια 2D και



Σχήμα 5.3 Μη-επίπεδο μοντέλο βαθμονόμησης με την συσκευή γραμμών λέιζερ α) απενεργοποιημένη β) ενεργοποιημένη

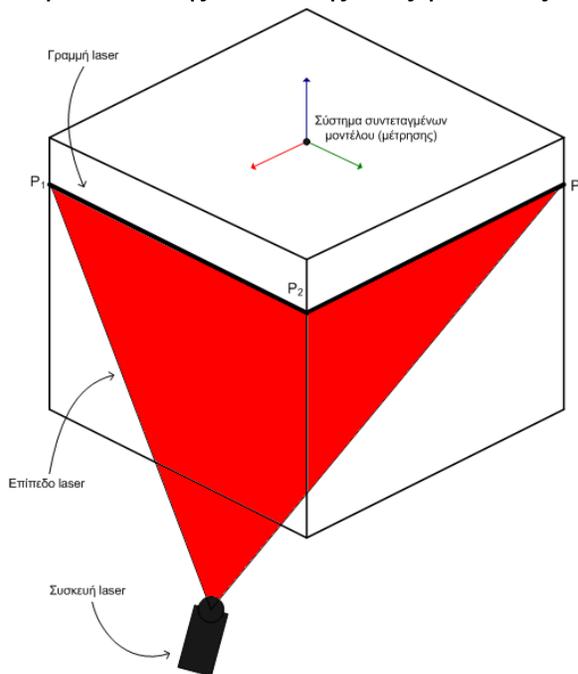
πραγματικών 3D συντεταγμένων των κουκίδων να υπολογίσουμε τους πίνακες εσωτερικών κι εξωτερικών παραμέτρων της κάμερας (βλ. Κεφ. 2.3), οι οποίοι ως γνωστόν σχετίζουν τις πραγματικές 3D συντεταγμένες (X_i, Y_i, Z_i) ενός παρατηρούμενου σημείου με τις αντίστοιχες 2D συντεταγμένες (u_i, v_i) της προβολή του πάνω σε μια ψηφιακή εικόνα σύμφωνα με την σχέση:

$$z_i \begin{bmatrix} u_i \\ v_i \\ 1 \end{bmatrix} = \mathbf{A} [\mathbf{R} | \mathbf{t}] \begin{bmatrix} X_i \\ Y_i \\ Z_i \\ 1 \end{bmatrix} \quad (5.1)$$

Η βαθμονόμηση της συσκευής λέιζερ ως προς την ψηφιακή κάμερα ισοδυναμεί με τον υπολογισμό των 3D εξισώσεων επιπέδων που χαρακτηρίζουν τις γραμμές λέιζερ ως προς το σύστημα συντεταγμένων μέτρησης (βαθμονόμησης) της κάμερας. Η χρησιμότητα της διαδικασίας αυτής θα γίνει περισσότερο κατανοητή στην επόμενη ενότητα, όπου περιγράφεται η διαδικασία πραγματοποίησης μετρήσεων με την συστοιχία κάμερας - συσκευής λέιζερ. Έστω ότι η κάμερα και το μοντέλο βαθμονόμησης (κύβος) παραμένουν στην αρχική τους θέση, στην οποία πραγματοποιήθηκε η βαθμονόμηση της κάμερας. Η ενεργοποίηση της συσκευής λέιζερ στην θέση αυτή θα έχει σαν αποτέλεσμα την προβολή των γραμμών φωτός πάνω στις ορατές πλευρές του κύβου, όπως φαίνεται στο Σχ. 5.3β. Υπενθυμίζεται ότι στην θέση αυτή το σύστημα συντεταγμένων του μοντέλου (κύβου) συμπίπτει με το σύστημα συντεταγμένων μέτρησης (βαθμονόμησης) της κάμερας. Αν χρησιμοποιήσουμε ειδικά φίλτρα στο φακό της κάμερας, τα οποία επιτρέπουν τη διέλευση μόνο της συχνότητας φωτός των ακτινών λέιζερ, τότε στην ψηφιακή εικόνα που θα λάβει η κάμερα από την παρούσα θέση θα είναι ορατές μόνο οι φωτεινές γραμμές πάνω στις επιφάνειες του κύβου. Αρχικός μας στόχος είναι να υπολογίσουμε τις 3D εξισώσεις ευθειών, οι οποίες θα χαρακτηρίζουν την κάθε φωτεινή γραμμή που προβάλλεται πάνω στις δύο επιφάνειες του κύβου. Είναι προφανές από το Σχ. 5.3β ότι με βάση το σύστημα συντεταγμένων του μοντέλου βαθμονόμησης, μία συντεταγμένη των ζητούμενων 3D εξισώσεων είναι γνωστή, καθώς τα επίπεδα των επιφανειών του κύβου είναι γνωστά εκ των προτέρων. Έτσι η αριστερή επιφάνεια του κύβου βρίσκεται στο επίπεδο $x=30\text{mm}$, και η δεξιά βρίσκεται στο επίπεδο $y=30\text{mm}$, οπότε και οι αντίστοιχες 3D εξισώσεις των γραμμών θ' ανήκουν στα παραπάνω επίπεδα. Για τον υπολογισμό μιας ζητούμενης εξίσωσης ευθείας αρκεί να βρούμε τις 3D συντεταγμένες δύο τουλάχιστο σημείων που ανήκουν στην αντίστοιχη φωτεινή γραμμή. Για το σκοπό αυτό εντοπίζουμε πάνω στην ψηφιακή εικόνα δύο 2D σημεία (u_1, v_1) , (u_2, v_2) που ανήκουν στην ίδια γραμμή και χρησιμοποιούμε την εξίσωση (5.1) κατά την αντίστροφη φορά για την ανακατασκευή των δύο άγνωστων 3D συντεταγμένων των

σημείων (η μία συντεταγμένη είναι γνωστή από την εξίσωση επιπέδου της πλευράς του κύβου). Για παράδειγμα, αναφερόμενοι σε ένα σημείο (u, v) που αντιστοιχεί στην αριστερή πλευρά του κύβου, επιλύουμε την εξίσωση (5.1) ως προς αγνώστους Y, Z και z , δεδομένου ότι $X=30\text{mm}$.

Οι 3D εξισώσεις ευθειών που χαρακτηρίζουν τις φωτεινές γραμμές πάνω στις επιφάνειες του κύβου δεν μπορούν να χαρακτηρίσουν απόλυτα την τοπολογία των γραμμών λέιζερ στο χώρο, καθώς αυτές εξαρτώνται άμεσα από την σχετική θέση του κύβου και της συσκευής λέιζερ, καθώς και τις διαστάσεις του κύβου. Αντίθετα το



Σχήμα 5.4 Προβολή γραμμής λέιζερ πάνω στην επιφάνεια του μοντέλου βαθμονόμησης.

επίπεδο που δημιουργεί η κάθε γραμμή λέιζερ στο χώρο (βλ. Σχ. 5.4) χαρακτηρίζει απόλυτα την τοπολογία της, ανεξάρτητα από την εκάστοτε επιφάνεια πάνω στην οποία μπορεί να προβάλλεται. Η εξίσωση επιπέδου της κάθε γραμμής εξαρτάται αποκλειστικά από τα κατασκευαστικά χαρακτηριστικά της συσκευής (πχ. γωνία ανοίγματος ακτινών) και μπορεί να υπολογιστεί ως προς το σύστημα συντεταγμένων μέτρησης της κάμερας, καθώς η σχετική θέση της κάμερας και της συσκευής λέιζερ είναι αμετάβλητη. Για να βρούμε την εξίσωση επιπέδου της κάθε γραμμής λέιζερ αρκεί να βρούμε τις 3D

συντεταγμένες τριών μη-γραμμικών σημείων που ανήκουν στο ίδιο επίπεδο, όπως των $P_1(x_1, y_1, z_1)$, $P_2(x_2, y_2, z_2)$ και $P_3(x_3, y_3, z_3)$ (βλ. Σχ. 5.4). Η εξίσωση του επιπέδου που προκύπτει από τα τρία σημεία αυτά είναι:

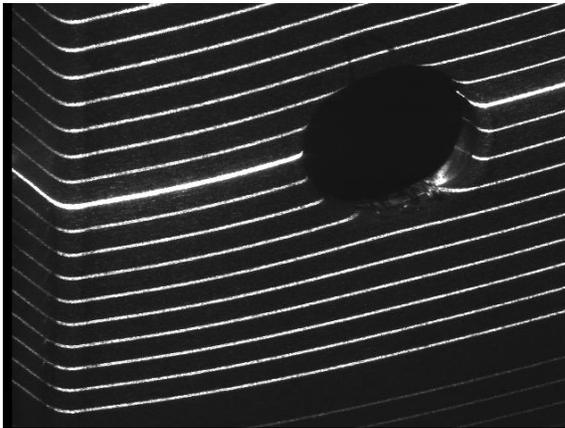
$$\begin{vmatrix} x - x_1 & y - y_1 & z - z_1 \\ x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \\ x_3 - x_1 & y_3 - y_1 & z_3 - z_1 \end{vmatrix} = 0 \quad (5.2)$$

και είναι της μορφής $Ax + By + Cz + D = 0$. Οι 3D συντεταγμένες των τριών σημείων υπολογίζονται και πάλι με την επίλυση της εξίσωσης (5.1) ως προς X, Y, Z , δεδομένου

ότι η μία συντεταγμένη (X ή Y) είναι γνωστή. Πραγματοποιώντας την ίδια διαδικασία για κάθε γραμμή λέιζερ μπορούμε να εξαγάγουμε την εξίσωση επιπέδου της κάθε γραμμής ξεχωριστά, αναφερόμενοι πάντα στο σύστημα συντεταγμένων μέτρησης της κάμερας. Η πληροφορία αυτή μας είναι απαραίτητη για να πραγματοποιήσουμε μετρήσεις, όπως θα γίνει κατανοητό στην επόμενη ενότητα.

5.3 Πραγματοποίηση μετρήσεων με την συστοιχία κάμερας - συσκευής λέιζερ

Έστω ότι μετακινούμε το άκρο του ρομποτικού βραχίονα μπροστά σε μια προς μέτρηση μεταλλική επιφάνεια. Αν στη θέση αυτή ενεργοποιήσουμε τις γραμμές λέιζερ τότε η ψηφιακή εικόνα που θα λάβει η κάμερα θα μοιάζει με αυτήν του Σχ. 5.5. Είναι προφανές ότι οι γραμμές λέιζερ παίρνουν τη μορφή της επιφάνειας πάνω στην οποία προβάλλονται, δίνοντας μια ξεκάθαρη εικόνα του προς μέτρηση αντικειμένου. Είναι αρκετά εύκολο, χρησιμοποιώντας τεχνικές επεξεργασίας εικόνας [Fisher and Naidu,



Σχήμα 5.5 Ψηφιακή εικόνα μιας μεταλλικής επιφάνειας, όταν οι γραμμές λέιζερ είναι ενεργοποιημένες.

1996], να βρούμε τις 2D συντεταγμένες εικόνας (pixel coordinates) του συνόλου των σημείων που απαρτίζουν τις φωτεινές γραμμές πάνω στην ψηφιακή εικόνα. Στη συνέχεια θα πρέπει να μετατρέψουμε τις 2D συντεταγμένες (εικονοστοιχεία) των φωτεινών σημείων της εικόνας στις αντίστοιχες πραγματικές τους 3D συντεταγμένες. Όπως αναλύθηκε κεφάλαιο (2.5), η χρήση μιας μονής κάμερας στην γενική περίπτωση δεν επιτρέπει την πλήρη

ανακατασκευή και των τριών διαστάσεων (συντεταγμένων) ενός παρατηρούμενου σημείου, καθώς χρησιμοποιώντας την εξίσωση (5.1) κατά αντίστροφο τρόπο, δεδομένου ότι οι 2D συντεταγμένες εικόνας (u , v) της προβολής του σημείου είναι γνωστές, παράγεται ένα σύστημα τριών γραμμικών εξισώσεων με τέσσερις αγνώστους (X , Y , Z και z). Στην περίπτωση όμως που η ψηφιακή κάμερα συνδυάζεται με μια συσκευή λέιζερ

(laser-triangulation sensor), η τέταρτη εξίσωση μπορεί να προέρθει από την γνωστή εξίσωση επιπέδου της γραμμής λέιζερ που αντιστοιχεί στο δεδομένο παρατηρούμενο σημείο. Σημειώνεται ότι τόσο η εξίσωση επιπέδου της γραμμής ($Ax+By+Cz+D=0$) όσο και οι άγνωστες 3D συντεταγμένες (X, Y, Z) αναφέρονται στο ίδιο σύστημα συντεταγμένων, ήτοι το σύστημα συντεταγμένων μέτρησης της κάμερας, γι' αυτό και μπορούν να συνδυαστούν με τη χρήση της (5.1) σχηματίζοντας ένα σύστημα τεσσάρων γραμμικών εξισώσεων με τέσσερις αγνώστους. Ακολουθώντας την παραπάνω διαδικασία για έναν μεγάλο αριθμό φωτεινών σημείων πάνω στην ψηφιακή εικόνα μπορούμε να ανακατασκευάσουμε τις αντίστοιχες 3D συντεταγμένες, οι οποίες αντιστοιχούν στα σημεία των γραμμών λέιζερ που προβάλλονται πάνω στην επιφάνεια του προς μέτρηση αντικειμένου (βλ. Σχ. 5.5). Αν το πλήθος των μετρούμενων (ανακατασκευασμένων) 3D σημείων είναι αρκετά μεγάλο τότε ο συνδυασμός τους παρέχει την πλήρη μορφή της επιφάνειας του αντικειμένου στον τρισδιάστατο χώρο.

5.4 Ο αλγόριθμος ICP

Η τρισδιάστατη ανακατασκευή των συντεταγμένων ενός μεγάλου πλήθους σημείων πάνω στην επιφάνεια του παρατηρούμενου αντικειμένου χρησιμεύει ως ενδιάμεσο βήμα για τον προσδιορισμό της θέσης και του προσανατολισμού του στο χώρο. Εκτός από τα σημεία τρισδιάστατης ανακατασκευής, στην διάθεσή μας έχουμε επίσης τα CAD δεδομένα της επιφάνειας του αντικειμένου, τα οποία όπως ήδη αναφέρθηκε αποτελούνται από ένα πυκνό πλέγμα τριγωνικών επιφανειών συνδεδεμένων μεταξύ τους. Τα CAD δεδομένα συνήθως παρέχονται από τον κατασκευαστή του αντικειμένου και η μορφή τους για το εξεταζόμενο αντικείμενο απεικονίζεται στο Σχ. 5.2. Οι συντεταγμένες των κορυφών των τριγώνων αναφέρονται σ' ένα ξεχωριστό σύστημα συντεταγμένων, το οποίο ονομάζεται σύστημα συντεταγμένων μοντέλου CAD. Προφανώς τα CAD δεδομένα του Σχ. 5.2 και το πλήθος των σημείων 3D ανακατασκευής, τα οποία υπολογίστηκαν με τη βοήθεια της ψηφιακής εικόνας του Σχ. 5.5, έχουν την ίδια μορφή καθώς περιγράφουν την επιφάνεια του ίδιου αντικειμένου. Το πρόβλημα είναι ότι το καθένα αναφέρεται σε ξεχωριστό σύστημα συντεταγμένων, ήτοι στο σύστημα συντεταγμένων μέτρησης της κάμερας και στο σύστημα συντεταγμένων του CAD

μοντέλου αντίστοιχα. Επόμενος στόχος μας είναι να συσχετίσουμε τα δύο αυτά συστήματα συντεταγμένων, υπολογίζοντας τον ομογενή μετασχηματισμό που τα συνδέει μεταξύ τους. Κατ' αυτό τον τρόπο θα έχουμε τη δυνατότητα να αναγάγουμε τα CAD δεδομένα (συντεταγμένες των τριγώνων) στο σύστημα συντεταγμένων μέτρησης της κάμερας.

Ο αλγόριθμος για τον υπολογισμό του παραπάνω μετασχηματισμού αναφέρεται στην διεθνή βιβλιογραφία ως *Iterative Closest Point (ICP)* [Besl and McKay, 1992]. Έστω ότι έχουμε ένα σύνολο N_p μετρούμενων 3D σημείων στο χώρο $P = \{\bar{p}_i\}$, τα οποία ανακτώνται από την 3D ανακατασκευή της ψηφιακής εικόνας του Σχ. 5.5, και ένα σύνολο N_x 3D τριγώνων, τα οποία παρέχονται από τα CAD δεδομένα του αντικειμένου $X = \{\bar{x}_i\}$. Ζητούμενο είναι να βρεθεί ο ομογενής μετασχηματισμός που συνδέει το σύστημα συντεταγμένων του CAD μοντέλου με το σύστημα συντεταγμένων μέτρησης (βαθμονόμησης) της κάμερας. Για τη διευκόλυνση της επίλυσης του παραπάνω προβλήματος χρησιμοποιούνται οι μεταβλητές *Unit Quaternions* [Kravchenko, 2003] $\bar{q}_R = [q_0 q_1 q_2 q_3]^T$, οι οποίες μπορούν να αναπαραστήσουν οποιονδήποτε 3×3 πίνακα στροφής. Για τις συγκεκριμένες μεταβλητές ισχύει $q_0 \geq 0$ και $q_0^2 + q_1^2 + q_2^2 + q_3^2 = 1$, ενώ ο αντίστοιχος πίνακας στροφής παράγεται από την παρακάτω εξίσωση:

$$\mathbf{R}(\mathbf{q}) = \begin{bmatrix} q_0^2 + q_1^2 - q_2^2 - q_3^2 & 2(q_1 q_2 - q_0 q_3) & 2(q_1 q_3 + q_0 q_2) \\ 2(q_1 q_2 + q_0 q_3) & q_0^2 + q_2^2 - q_1^2 - q_3^2 & 2(q_2 q_3 - q_0 q_1) \\ 2(q_1 q_3 - q_0 q_2) & 2(q_2 q_3 + q_0 q_1) & q_0^2 + q_3^2 - q_1^2 - q_2^2 \end{bmatrix} \quad (5.3)$$

Στην περίπτωσή μας, εκτός από την εύρεση του πίνακα στροφής που συνδέει τα δύο συστήματα συντεταγμένων, απαιτείται κι ο υπολογισμός της σχετικής μετατόπισης των κέντρων τους. Για το σκοπό αυτό, εκτός από τον υπολογισμό του 4D διανύσματος των unit quaternions απαιτείται κι ο υπολογισμός ενός 3D διανύσματος μετατόπισης $\bar{q}_T = [q_4 q_5 q_6]^T$ που συνδέει τα κέντρα των δύο συστημάτων συντεταγμένων. Συνολικά απαιτείται ο υπολογισμός ενός 7D διανύσματος $\bar{q} = [\bar{q}_R | \bar{q}_T]^T$, το οποίο αποτελεί τη συγχώνευση των δύο προαναφερθέντων διανυσμάτων. Ο αλγόριθμος ICP περιλαμβάνει τα παρακάτω βήματα:

- I. Αρχικοποίηση του αλγορίθμου, θέτοντας $P_k = P_0$, $\vec{q}_0 = [1 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0]^T$ και $k=0$, όπου P_0 είναι το (αρχικό) πλήθος των μετρούμενων σημείων, \vec{q}_0 είναι ο μηδενικός μετασχηματισμός και k είναι το βήμα επανάληψης (iteration step).
- II. Για όλα τα σημεία του συνόλου P_k βρίσκουμε τ' αντίστοιχα σημεία ελάχιστων αποστάσεων Y_k από το σύνολο X των CAD τριγώνων. Τα σημεία ελάχιστων αποστάσεων είναι 3D σημεία πάνω στην επιφάνεια του πλέγματος των CAD τριγώνων.
- III. Υπολογισμός του 7D μετασχηματισμού \vec{q}_k που στοιχίζει κατά τον καλύτερο τρόπο το αρχικό πλήθος μετρήσεων P_0 με τ' αντίστοιχα σημεία ελάχιστων αποστάσεων Y_k , τα οποία υπολογίστηκαν στο προηγούμενο βήμα. Η εύρεση του ζητούμενου μετασχηματισμού επιτυγχάνεται με την ελαχιστοποίηση της παρακάτω συνάρτησης κόστους:

$$f(\vec{q}_k) = \frac{1}{N_p} \sum_{i=1}^{N_p} \left\| \vec{y}_i(k) - \mathbf{R}(\vec{q}_R) \vec{p}_i(0) - \vec{q}_T \right\|^2 \quad (5.4)$$

Μια αναλυτική λύση υπολογισμού του ζητούμενου μετασχηματισμού \vec{q}_k , ο οποίος ελαχιστοποιεί την παραπάνω συνάρτηση κόστους, μπορεί να βρεθεί στο [Horn, 1987]. Εκτός από τον μετασχηματισμό, η προτεινόμενη αναλυτική λύση παρέχει μια τιμή μέσου τετραγωνικού σφάλματος d_k , η οποία χαρακτηρίζει την ποιότητα αντιστοίχισης του συνόλου των σημείων P_0 με τ' αντίστοιχα Y_k .

- IV. Εφαρμογή του μετασχηματισμού \vec{q}_k στο αρχικό πλήθος των μετρούμενων σημείων P_0 , ήτοι $P_{k+1} = \vec{q}_k(P_0)$.
- V. Τερματισμός της ακολουθίας εάν το μέσο τετραγωνικό σφάλμα d_k του βήματος III πέσει κάτω από μια προκαθορισμένη τιμή κατωφλίου (threshold), ή εάν δεν παρατηρηθεί σημαντική μεταβολή του σε δύο διαδοχικές επαναλήψεις του αλγορίθμου. Στην αντίθετη περίπτωση η ακολουθία επαναλαμβάνεται ξεκινώντας από το βήμα II, και θέτοντας $k=k+1$.

Πειραματικά αποτελέσματα έχουν δείξει πρέπει να υπάρχει μια αρχική χονδρική στοίχιση των CAD δεδομένων (τριγώνων) και του πλήθους των σημείων μέτρησης, ώστε ο αλγόριθμος ICP να συγκλίνει στο σωστό σημείο, παρέχοντας τον ζητούμενο μετασχηματισμό συντεταγμένων $\vec{q} = [\vec{q}_R | \vec{q}_T]^T$. Αν δεν υπάρχει αρχική στοίχιση τότε υπάρχει περίπτωση ο αλγόριθμος να συγκλίνει σ' ένα τοπικό ελάχιστο, δίνοντας μια εντελώς λανθασμένη εκτίμηση για τον ζητούμενο μετασχηματισμό. Η αρχική στοίχιση μπορεί να γίνει στρέφοντας τα CAD τρίγωνα γύρω από τους κύριους άξονες του συστήματος συντεταγμένων του μοντέλου (με τη βοήθεια κάποιου προγράμματος επεξεργασίας 3D), ώστε η μορφή τους να μοιάζει αρκετά με τη μορφή του πλήθους των μετρούμενων σημείων. Σύμφωνα με τα πειραματικά αποτελέσματα της παρούσας υλοποίησης ο αλγόριθμος ICP συγκλίνει στο σωστό βέλτιστο σημείο αν η αρχική στοίχιση έχει μια απόκλιση ± 10 μοίρες από την ιδανική στοίχιση, ενώ αποκλίσεις στις σχετικές θέσεις των κέντρων των δύο συστημάτων συντεταγμένων δεν φαίνεται να επηρεάζουν το σημείο σύγκλισης του αλγόριθμου.

Αποτέλεσμα της χρήσης του αλγορίθμου ICP είναι η συσχέτιση του συστήματος συντεταγμένων του CAD μοντέλου με το σύστημα συντεταγμένων μέτρησης της κάμερας, χρησιμοποιώντας ένα μεγάλο πλήθος σημείων 3D ανακατασκευής που αντιστοιχούν στην επιφάνεια του παρατηρούμενου αντικειμένου. Το σύστημα συντεταγμένων του CAD μοντέλου είναι ιδεατά «πακτωμένο» πάνω στο μετρούμενο αντικείμενο και ακολουθεί πάντοτε τις κινήσεις του, οπότε λαμβάνοντας ψηφιακές εικόνες του παρατηρούμενου αντικειμένου και χρησιμοποιώντας τον αλγόριθμο ICP μπορούμε ν' ανιχνεύσουμε τις μεταβολές θέσης και προσανατολισμού του κινούμενου αντικειμένου (συστήματος συντεταγμένων CAD) και να τις αναγάγουμε στο σύστημα συντεταγμένων μέτρησης της κάμερας. Η παραπάνω διαδικασία θα αναλυθεί εκτενέστερα στο Κεφ. 5.7, όπου γίνεται συνδυασμός όλων των διακριτών βημάτων για την επίτευξη του ζητούμενου στόχου, ο οποίος είναι η παρακολούθηση τροχιάς ενός κινούμενου αντικειμένου από το άκρο ενός βιομηχανικού ρομποτικού βραχίονα.

5.5 Επιτάχυνση του αλγορίθμου ICP με τη χρήση των κατανομών K-D trees

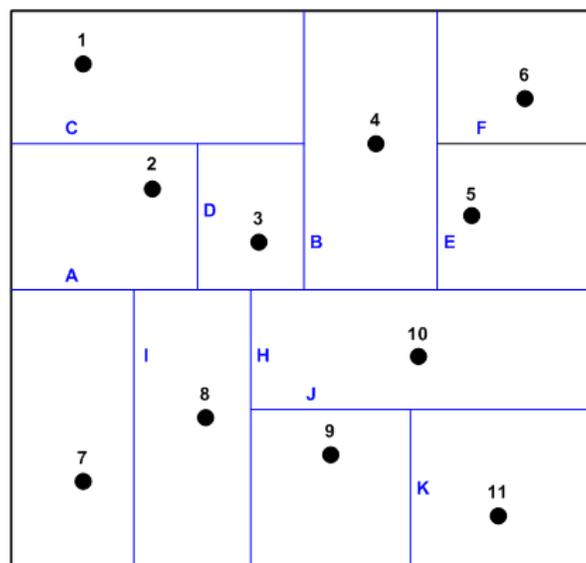
Ένας πολύ σημαντικός παράγοντας για τους υπολογιστικούς αλγόριθμους που προορίζονται για πρακτικές εφαρμογές είναι η χρονική απόκρισή τους (time performance). Ο αλγόριθμος ICP είναι επαναληπτικός (iterative) δηλαδή βασίζεται στην επανάληψη της ίδιας ακολουθίας μαθηματικών πράξεων μέχρι να υπάρξει σύγκλιση των αποτελεσμάτων. Ανάλογα με την αρχική σχετική θέση του πλήθους των μετρούμενων σημείων και του CAD μοντέλου, την πολυπλοκότητα της επιφάνειας του παρατηρούμενου αντικειμένου, καθώς και την απαιτούμενη ακρίβεια των αποτελεσμάτων, ο αριθμός των επαναλήψεων που απαιτείται από τον αλγόριθμο ICP για τη σύγκλιση των αποτελεσμάτων μπορεί να είναι από 5 μέχρι 100. Στην πρωτότυπη εργασία [Besl and McKay, 1992] προτείνεται μια παραλλαγή του αλγορίθμου, η οποία μπορεί να μειώσει σημαντικά τον απαιτούμενο αριθμό επαναλήψεων. Σε κάθε περίπτωση πάντως, η πιο υπολογιστικά χρονοβόρα διαδικασία που σχετίζεται με τον αλγόριθμο ICP είναι ο υπολογισμός των ζευγών ελαχίστων αποστάσεων μεταξύ του πλήθους των μετρούμενων σημείων και του πλέγματος των CAD τριγώνων (βήμα II του Κεφ. 5.4). Εάν το πλήθος των μετρούμενων σημείων είναι N_p και το πλήθος των τριγώνων που απαρτίζουν το CAD μοντέλο είναι N_x , τότε είναι προφανές ότι κάθε μετρούμενο σημείο θα πρέπει να ψάξει όλο το πλήθος των N_x τριγώνων για να βρει εκείνο που αντιστοιχεί στην ελάχιστη απόσταση. Εξετάζοντας το θέμα από υπολογιστικής άποψης, σε κάθε επανάληψη του αλγορίθμου ICP μια συνάρτηση υπολογισμού απόστασης μεταξύ ενός 3D σημείου και της επιφάνειας ενός 3D τριγώνου πρέπει να κληθεί $N_p \times N_x$ φορές, έτσι ώστε όλα τα N_p μετρούμενα σημεία ν' αντιστοιχιστούν με τα εκάστοτε τρίγωνα ελαχίστων αποστάσεων. Εκτός από την τιμή της απόστασης, η συνάρτηση υπολογισμού θα πρέπει επίσης να επιστρέφει το σημείο πάνω στην επιφάνεια του τριγώνου που αντιστοιχεί στην ελάχιστη απόσταση. Υπενθυμίζεται ότι το σύνολο των σημείων ελαχίστων αποστάσεων (Y_k) πάνω στις επιφάνειες των αντίστοιχων τριγώνων είναι απαραίτητο για το επόμενο βήμα του αλγορίθμου ICP (βήμα III). Τα CAD μοντέλα των περισσότερων βιομηχανικών αντικειμένων απαρτίζονται από αρκετές εκατοντάδες ή και χιλιάδες τριγώνων, έτσι ώστε η περιγραφή της τοπολογίας των επιφανειών τους να είναι αρκετά ακριβής. Το γεγονός αυτό, σε συνδυασμό με την ανάγκη τρισδιάστατης ανακατασκευής ενός αρκετά μεγάλου πλήθους σημείων για την κάλυψη μιας ικανοποιητικής περιοχής της επιφάνειας του μετρούμενου αντικειμένου, μας οδηγεί στο

συμπέρασμα ότι η συνάρτηση υπολογισμού απόστασης μεταξύ ενός 3D σημείου και της επιφάνειας ενός 3D τριγώνου μπορεί να καλείται εκατομμύρια φορές σε κάθε επανάληψη του αλγορίθμου ICP. Η χρονική καθυστέρηση που συνεπάγεται τον μεγάλο αυτό υπολογιστικό φόρτο μπορεί ν' αποτρέψει την χρήση του αλγορίθμου ICP σε πρακτικές εφαρμογές, στις οποίες ο χρονική απόκριση ενδεχομένως παίζει σημαντικό ρόλο.

Η πιο κομψή και αποτελεσματική μέθοδος για την επιτάχυνση των υπολογισμών ελαχίστων αποστάσεων είναι η προσέγγιση των κατανομών *KD-Tree*, η οποία προτάθηκε για πρώτη φορά από τους [Friedman et al., 1977] και εφαρμόστηκε επιτυχώς για την επιτάχυνση του αλγορίθμου ICP από τους [Cheverikov et al., 2002; Rusinkiewicz and Levo, 2001; Greenspan and Yurick, 2003]. Η βασική ιδέα συνίσταται στην κατηγοριοποίηση των δομικών στοιχείων του CAD μοντέλου (τριγώνων στην περίπτωσή μας) με βάση κάποιους συγκεκριμένους κανόνες, με σκοπό τη δημιουργία μιας δενδροειδούς κατανομής. Η κατηγοριοποίηση των CAD τριγώνων κατ' αυτόν τον τρόπο ωφελεί σε πολύ μεγάλο βαθμό την ταχύτητα εύρεσης του τριγώνου ελάχιστης απόστασης μεταξύ ενός 3D σημείου στο χώρο και του δεδομένου CAD πλέγματος. Ανάλογα με τη σχετική θέση του 3D σημείου ως προς το CAD πλέγμα των τριγώνων, η αναζήτηση του τριγώνου ελάχιστης απόστασης περιορίζεται σε συγκεκριμένες μόνο περιοχές της δενδροειδούς κατανομής, αποκλείοντας άλλες που δεν είναι δυνατό να περιέχουν το ζητούμενο τρίγωνο. Ο αποκλεισμός ολόκληρων περιοχών του πλέγματος των τριγώνων επιταχύνει σε πολύ μεγάλο βαθμό την διαδικασία εύρεσης του τριγώνου ελάχιστης απόστασης, καθώς δεν είναι πλέον απαραίτητο να ελέγχεται όλο το πλήθος των N_x τριγώνων ως υποψήφια λύση. Η δημιουργία της δενδροειδούς κατανομής (KD-tree) για το πλέγμα των τριγώνων πραγματοποιείται μόνο μια φορά (πχ. στην αρχή του προγράμματος), καθώς αυτή παραμένει αμετάβλητη για ένα δεδομένο CAD μοντέλο.

Στην συνέχεια θα περιγραφεί η κατασκευή μια κατανομής KD-tree για ένα σύνολο 2D σημείων, τα οποία απεικονίζονται στο Σχ. 5.6. Η κατασκευή της κατανομής KD-tree που αντιστοιχεί στο σύνολο των 3D τριγώνων του CAD μοντέλου είναι λίγο πιο σύνθετη, αλλά βασίζεται στις ίδιες αρχές με αυτές που θα χρησιμοποιηθούν για το σύνολο των 2D σημείων. Το πρώτο βήμα είναι η επιλογή μιας κατάλληλης *διαχωριστικής γραμμής* κατά τη διεύθυνση του x ή y άξονα, η οποία θα βρίσκεται όσο το δυνατόν πιο κοντά στο μέσο της περιοχής που περικλείει τα σημεία, και θα τα χωρίζει σε δύο σύνολα του ίδιου (ή

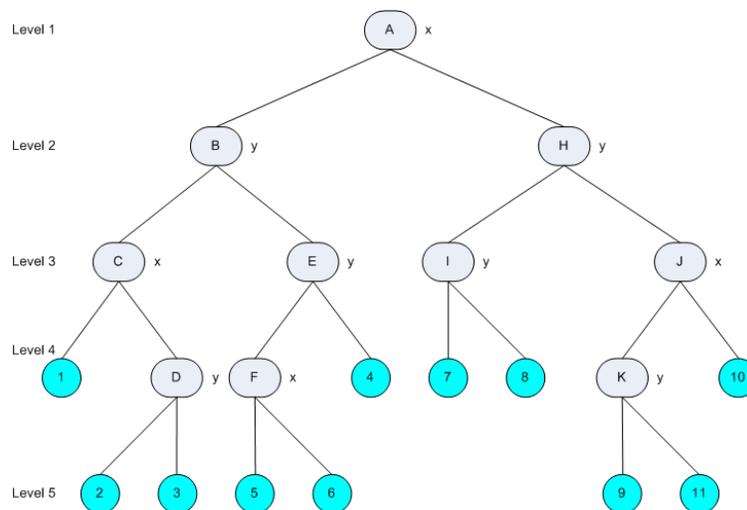
σχεδόν του ίδιου) μεγέθους. Στο παράδειγμα του Σχ. 5.6, η πρώτη επιλεγμένη διαχωριστική γραμμή είναι κατά τον άξονα x και ονομάζεται A . Στην συνέχεια, από τις δύο υπο-περιοχές που προκύπτουν, η πάνω περιοχή διαιρείται από την διαχωριστική γραμμή B κατά τον άξονα y , ενώ η κάτω περιοχή διαιρείται από την διαχωριστική γραμμή H κατά τον ίδιο άξονα. Η ίδια διαδικασία εκτελείται διαδοχικά, μέχρι όλες οι διαιρεμένες υπο-περιοχές να περιλαμβάνουν μόνο ένα σημείο. Υπάρχουν αρκετές μέθοδοι για την επιλογή της βέλτιστης θέσης και διεύθυνσης της εκάστοτε διαχωριστικής γραμμής κατά τη διαδικασία κατασκευής μιας κατανομής KD-tree, μερικές από τις οποίες αναφέρονται από τους [Sproull, 1991; Zinber et al., 2003]. Η απλούστερη μέθοδος συνίσταται στην επιλογή εναλλασσόμενων διαχωριστικών διευθύνσεων (x - y - x - y - x -...), επιλέγοντας κάθε φορά το μέσο των υπό διαίρεση περιοχών. Άλλες πιο εξελιγμένες μέθοδοι χρησιμοποιούν μια κατάλληλη συνάρτηση κόστους, από την ελαχιστοποίηση της οποίας επιχειρούν να προσδιορίσουν την βέλτιστη θέση και διεύθυνση της εκάστοτε διαχωριστικής γραμμής. Η κατανομή KD-tree που παράγεται για το απλό παράδειγμα των 2D σημείων του Σχ. 5.6 φαίνεται στο Σχ. 5.7.



Σχήμα 5.6 Κατασκευή κατανομής KD-tree για ένα πλήθος 2D σημείων.

Τα στρογγυλεμένα κουτιά του Σχ. 5.7 που περιέχουν τα κεφαλαία γράμματα A-K ονομάζονται *κόμβοι* (*nodes*) και αντιπροσωπεύουν διαχωριστικές γραμμές κατά τους

άξονες x ή y . Αναφερόμενοι σ' ένα πρόγραμμα υπολογιστή, η πληροφορία που θα πρέπει να εμπεριέχει ένας κόμβος είναι η διεύθυνση της διαχωριστικής γραμμής (x ή y), μια αριθμητική τιμή που αντιπροσωπεύει το σημείο διαχωρισμού (πάνω στον άξονα y ή x αντίστοιχα) και δύο δείκτες (pointers) που δείχνουν τις δύο παραχθείσες υπο-περιοχές στο αμέσως χαμηλότερο επίπεδο του δένδρου. Οι κύκλοι του Σχ. 5.7 που περιέχουν τους αριθμούς 1-11 ονομάζονται *φύλλα* (leaves) και αντιστοιχούν στα πραγματικά 2D σημεία που είναι κατηγοριοποιημένα σύμφωνα με την κατανομή του δένδρου. Η παραχθείσα κατανομή KD-tree χαρακτηρίζεται από ένα συγκεκριμένο βάθος, το οποίο συνήθως ονομάζεται *επίπεδο* (level). Κάθε διαίρεση μιας επίπεδης περιοχής από μια διαχωριστική γραμμή παράγει δύο υπο-περιοχές, οι οποίες χαρακτηρίζονται από ένα αύξοντα αριθμό επιπέδου σε σχέση με την αρχική περιοχή. Στο Σχ. 5.7 φαίνεται καθαρά ότι τα φύλλα 1-11 ανήκουν στα επίπεδα 4 και 5 της κατανομής.



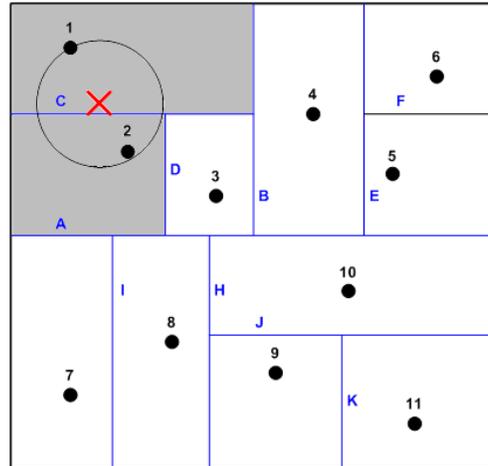
Σχήμα 5.7 Κατανομή KD-tree για το πλήθος των σημείων του Σχ. 5.6.

Στην παρούσα εφαρμογή, τα δομικά στοιχεία που πρέπει να κατηγοριοποιηθούν κάτω από μια κατανομή KD-tree είναι 3D τριγωνικές επιφάνειες. Στην περίπτωση αυτή θα πρέπει να διαιρέσουμε τον τρισδιάστατο χώρο σε περιοχές, χρησιμοποιώντας αντί για διαχωριστικές γραμμές *διαχωριστικά επίπεδα*, τα οποία θα χαρακτηρίζονται από μια σταθερή συντεταγμένη (x , y , ή z). Σε κάθε βήμα κατασκευής της κατανομής το διαχωριστικό επίπεδο μπορεί να τέμνει τις επιφάνειες κάποιων τριγώνων του CAD

μοντέλου, τα οποία θα πρέπει να συμπεριληφθούν και στις δύο παραχθείσες υπο-περιοχές. Το γεγονός αυτό καθιστά απίθανη την περίπτωση όπου η διαδοχική υποδιαίρεση του τρισδιάστατου χώρου θα παράγει υπο-περιοχές που θα περιλαμβάνουν ακριβώς ένα τρίγωνο. Είναι υπολογιστικά αποδεκτό να θεωρήσουμε μια υπο-περιοχή ως φύλλο της κατανομής εάν ο αριθμός των περικλειόμενων τριγώνων πέσει κάτω από κάποια προκαθορισμένη τιμή.

Η επιλογή της διεύθυνσης και της θέσης του εκάστοτε διαχωριστικού επιπέδου σε κάθε βήμα υποδιαίρεσης του τρισδιάστατου χώρου είναι ζωτικής σημασίας για την κατασκευή ενός καλά κατανεμημένου KD-tree. Στην παρούσα εφαρμογή δεν χρησιμοποιήσαμε εναλλασσόμενες διευθύνσεις διαχωριστικών επιπέδων (x-y-z-x-y-z...) σε κάθε επίπεδο (level) κατασκευής της κατανομής KD-tree, αλλά επιλέγαμε την εκάστοτε διεύθυνση που ήταν κάθετη στο μεγαλύτερο εύρος συντεταγμένων των τριγώνων που περικλείονταν στην τρέχουσα υπό διαίρεση περιοχή. Για την επιλογή του βέλτιστου σημείου διαχωρισμού, δημιουργήσαμε ένα πυκνό πλέγμα διαχωριστικών επιπέδων της επιλεγμένης διεύθυνσης καθ' όλο το εύρος συντεταγμένων των τριγώνων, και πραγματοποιώντας πειραματικές διαιρέσεις της συγκεκριμένης περιοχής επιλέγαμε το επίπεδο εκείνο που διαιρούσε το πλήθος των τριγώνων σε δύο σύνολα όσο το δυνατόν του ίδιου μεγέθους. Κατ' αυτόν τον τρόπο εξασφαλίζαμε ότι κάθε υποδιαίρεση του χώρου δημιουργούσε δύο υπο-περιοχές που περιλάμβαναν περίπου τον ίδιο αριθμό τριγώνων.

Όπως αναφέρθηκε παραπάνω, ο κύριος λόγος κατασκευής μιας κατανομής KD-tree είναι για να επιταχύνει τους υπολογισμούς ελάχιστης απόστασης. Η περιγραφή της διαδικασίας επίτευξης του παραπάνω στόχου θα βασιστεί στο απλό παράδειγμα του συνόλου των 2D σημείων του Σχ. 5.6, του οποίου η κατανομή KD-tree έχει ήδη υλοποιηθεί και απεικονίζεται στο Σχ. 5.7. Ας υποθέσουμε ότι έχουμε στη διάθεσή μας ένα σταθερό 2D σημείο αναφοράς μέσα στην εξεταζόμενη περιοχή και θέλουμε από το σύνολο των σημείων 1-11 να βρούμε εκείνο που αντιστοιχεί στην ελάχιστη απόσταση, χωρίς όμως να υπολογίσουμε όλες τις δυνατές αποστάσεις.



Σχήμα 5.8 Εύρεση του σημείου ελάχιστης απόστασης με τη χρήση της κατανομής KD-tree.

Αναφερόμενοι στο Σχ. 5.8, το σημείο αναφοράς (x_0, y_0) αντιστοιχεί στον κόκκινο σταυρό, ενώ οι μαύρες κουκίδες αριθμημένες από 1 ως 11 είναι το σύνολο των εξεταζόμενων σημείων που θέλουμε να ελέγξουμε ως υποψήφιες λύσεις. Ο αλγόριθμος εύρεσης του σημείου ελάχιστης απόστασης αποτελείται από τα παρακάτω βήματα:

- α) Εύρεση του φύλλου της κατανομής KD-tree που περιλαμβάνει το σημείο αναφοράς. Ξεκινώντας από την ρίζα της κατανομής (Σχ. 5.7) εξετάζουμε σε ποια πλευρά της διαχωριστικής γραμμής A βρίσκεται το σημείο αναφοράς. Στο παράδειγμά μας ισχύει $y_0 > y_A$, οπότε ακολουθούμε την αριστερή διακλάδωση του κόμβου A καταλήγοντας στον κόμβο B. Επαναλαμβάνοντας της ίδια διαδικασία για τους κόμβους B και C καταλήγουμε στο φύλλο (υπο-περιοχή) που αντιστοιχεί στο σημείο 1.
- β) Υπολογισμός της απόστασης μεταξύ του σημείου αναφοράς και του σημείου που αντιστοιχεί στο φύλλο του βήματος (α). Στο παράδειγμά μας αναφερόμαστε στο σημείο 1, οπότε η ζητούμενη απόσταση είναι $D_{\min} = D_1$.
- γ) Πραγματοποίηση οπισθοδρόμησης (backtracking) στους γονικούς κόμβους της κατανομής KD-tree, δηλαδή εκείνους που ανήκουν στα υψηλότερα επίπεδα της κατανομής. Πρώτα εξετάζεται ο κόμβος C, ο οποίος αντιπροσωπεύει μια διαχωριστική γραμμή κατά τον άξονα x. Εάν ο κύκλος που σχηματίζεται με κέντρο

το σημείο αναφοράς (x_0, y_0) και ακτίνα D_{\min} τέμνει την διαχωριστική γραμμή C (βλ. Σχ. 5.8) τότε ακολουθούμε και την άλλη πλευρά του κόμβου C καταλήγοντας στον κόμβο D. Ελέγχοντας και πάλι σε ποια μεριά της διαχωριστικής γραμμής D ανήκει το σημείο αναφοράς ($x_0 < x_D$) καταλήγουμε στο φύλλο της κατανομής που αντιστοιχεί στο σημείο 2. Εάν η απόσταση μεταξύ του σημείου (x_0, y_0) και του σημείου 2 είναι μικρότερη από την τρέχουσα τιμή D_{\min} τότε ανανεώνουμε την τιμή αυτή ως $D_{\min} = D_2$.

- δ) Επανάληψη του βήματος (γ) μέχρι να εξεταστούν όλοι οι απαραίτητοι κόμβοι της κατανομής. Στο παράδειγμά μας η διαδρομή που ακολουθεί η διαδικασία backtracking είναι περνώντας διαδοχικά από εξής σημεία της κατανομής (βλ. Σχ. 5.7): A-B-C-1-D-2-B-A. Ο κύκλος με ακτίνα την ελάχιστη απόσταση D_{\min} δεν τέμνει ούτε την διαχωριστική γραμμή B (κόμβος B) ούτε την A (κόμβος A), οπότε δεν χρειάζεται να εξετάσουμε τα τμήματα της κατανομής (κλαδιά) που έχουν σαν ρίζες τους κόμβους E και H, αποκλείοντας κατ' αυτόν τον τρόπο όλα τα σημεία 3-11 ως υποψήφιες λύσεις ελάχιστης απόστασης. Ο αλγόριθμος backtracking τερματίζεται στον ρίζα της κατανομής (κόμβος A) έχοντας βρει ως σημείο ελάχιστης απόστασης το σημείο 2.

Το πλεονέκτημα της εύρεσης του σημείου ελάχιστης απόστασης με τη βοήθεια της κατανομής KD-tree είναι προφανές. Στο παράδειγμά μας ο υπολογισμός της απόστασης μεταξύ δύο σημείων χρειάστηκε να πραγματοποιηθεί μόνο δύο φορές, ήτοι μεταξύ του σημείου αναφοράς και των σημείων 1 και 2. Χρησιμοποιώντας τη συμβατική μέθοδο θα έπρεπε να υπολογιστούν όλες οι δυνατές αποστάσεις μεταξύ του σημείου αναφοράς και του συνόλου των 11 σημείων.

Η μεθοδολογία υπολογισμού ελάχιστης απόστασης που περιγράφεται για την περίπτωση των 2D σημείων μπορεί πολύ εύκολα να γενικευθεί για την περίπτωση του πλέγματος 3D τριγώνων ενός CAD μοντέλου. Οι κύριες διαφορές είναι ότι θα πρέπει επιπλέον να χρησιμοποιήσουμε την z-διάσταση όπου αυτό είναι απαραίτητο (πχ. στους κόμβους που αντιστοιχούν σε διαχωριστικά επίπεδα παράλληλα στον άξονα z) και να χρησιμοποιήσουμε μια συνάρτηση υπολογισμού απόστασης μεταξύ ενός 3D σημείου και της επιφάνειας ενός 3D τριγώνου, αντί της απόστασης μεταξύ δύο σημείων. Επιπλέον σε

κάθε βήμα του αλγορίθμου backtracking θα εξετάζεται εάν η τρισδιάστατη σφαίρα με κέντρο το σημείο αναφοράς και ακτίνα την τρέχουσα ελάχιστη απόσταση D_{\min} θα τέμνει το επίπεδο που αντιστοιχεί στον εξεταζόμενο κόμβο της κατανομής.

Εάν η κατανομή KD-tree είναι ιδανικά κατανεμημένη σε όλη την έκταση του χώρου που περιλαμβάνει τα δομικά στοιχεία (N_x τρίγωνα) τότε ο χρόνος υπολογισμού του τριγώνου ελάχιστης απόστασης ως προς ένα 3D σημείο αναφοράς μπορεί να πέσει από $O(N_x)$ σε $O(\log N_x)$, όπου O είναι μια αυθαίρετη σταθερά αναλογίας. Όπως θα φανεί και στα πειραματικά αποτελέσματα του Κεφ. 5.7, η επιτάχυνση που επιφέρει η χρήση των κατανομών KD-tree στον αλγόριθμο ICP είναι πολύ σημαντική, ειδικά στην περίπτωση χρήσης CAD μοντέλων αποτελούμενων από πολύ μεγάλο πλήθος τριγώνων.

5.6 Συσχέτιση συστημάτων συντεταγμένων μέτρησης κάμερας και άκρου ρομποτικού βραχίονα

Το αμέσως επόμενο βήμα είναι η συσχέτιση του συστήματος συντεταγμένων μέτρησης της κάμερας με το αντίστοιχο του άκρου του ρομποτικού βραχίονα. Κατ' αυτό τον τρόπο οι μετρήσεις που πραγματοποιεί η συστοιχία κάμερας και συσκευής λέιζερ (laser triangulation sensor), οι οποίες χαρακτηρίζονται ως τοπικού χαρακτήρα καθώς εξαρτώνται από την θέση της κάμερας στο χώρο, μπορούν με αναχθούν σ' ένα σύστημα συντεταγμένων που είναι γνωστό στον ρομποτικό βραχίονα. Εφόσον τα συστήματα συντεταγμένων μέτρησης της κάμερας και του κέντρου προβολής της συσχετίζονται με τον σταθερό μετασχηματισμό $[R|t]$ (πίνακας εξωτερικών παραμέτρων από την βαθμονόμηση της κάμερας) αρκεί να υπολογίσουμε τον άγνωστο μετασχηματισμό X μεταξύ οποιουδήποτε εκ των δύο συστημάτων συντεταγμένων και του αντίστοιχου του άκρου του βραχίονα. Όπως αναφέρθηκε και σε προηγούμενα κεφάλαια, η διαδικασία υπολογισμού του άγνωστου μετασχηματισμού X είναι γνωστή ως *βαθμονόμηση βραχίονα – μηχανικής όρασης (Hand-Eye Calibration)*. Η συσκευή λέιζερ δεν χρησιμοποιείται καθ' όλη τη διάρκεια της διαδικασίας, οπότε η συστοιχία μέτρησης κάμερας - συσκευής λέιζερ ουσιαστικά αντιμετωπίζεται σαν μια μονή κάμερα. Ως μοντέλο βαθμονόμησης μπορεί είτε να χρησιμοποιηθεί το επίπεδο μοντέλο του Κεφ. 3.2, είτε το μοντέλο κυβικού σχήματος του Κεφ. 5.2. Στα πειράματά μας χρησιμοποιήσαμε ένα επίπεδο μοντέλο

βαθμονόμησης, σε συνδυασμό με τον αλγόριθμο βαθμονόμησης κάμερας του Zhang για την εξαγωγή των μετασχηματισμών της κάμερας $A_1-A_2-\dots-A_n$.

5.7 Συνδυασμός διακριτών μεθόδων και πειραματικά αποτελέσματα

Τελικός στόχος της εφαρμογής μας είναι ο έλεγχος της τροχιάς του άκρου ενός βιομηχανικού ρομποτικού βραχίονα ώστε να παρακολουθεί τις μεταβολές θέσης και προσανατολισμού ενός κινούμενου στο χώρο αντικειμένου, χρησιμοποιώντας ως ανάδραση τις μετρήσεις από μια συστοιχία κάμερας - συσκευής λέιζερ (laser triangulation sensor) που είναι σταθερά στερεωμένη στο άκρο του βραχίονα. Η θέση και ο προσανατολισμός του κινούμενου αντικειμένου ουσιαστικά χαρακτηρίζονται από το σύστημα συντεταγμένων του CAD μοντέλου, το οποίο είναι πακτωμένο πάνω στο αντικείμενο κι ακολουθεί όλες τις κινήσεις του. Για να γίνει δυνατή η παρακολούθηση του αντικειμένου από το άκρο του βραχίονα θα πρέπει το κινούμενο σύστημα συντεταγμένων του CAD μοντέλου να συσχετιστεί με κάποιο σύστημα συντεταγμένων που είναι γνωστό στο βραχίονα, όπως είναι εκείνο της βάσης του (*world or base frame*) ή του τελικού στοιχείου δράσης (*end-effector or TCP frame*).

Μια πρώτη προφανής προσέγγιση είναι η απ' ευθείας αναγωγή του συστήματος συντεταγμένων του CAD μοντέλου στο σταθερό σύστημα συντεταγμένων βάσης του βραχίονα, χρησιμοποιώντας όλους τους ενδιάμεσους μετασχηματισμούς. Ας υποθέσουμε ότι το άκρο του βραχίονα βρίσκεται σε μια αρχική θέση όπου η συστοιχία κάμερας-συσκευής λέιζερ μπορεί να μετρήσει (ανακατασκευάσει) τις 3D συντεταγμένες ενός μεγάλου πλήθους σημείων πάνω στην επιφάνεια του παρατηρούμενου αντικειμένου. Εφαρμόζοντας τον αλγόριθμο ICP για το πλήθος των μετρούμενων σημείων μπορεί να εξαχθεί ο μετασχηματισμός \mathbf{M}_i που συνδέει το σύστημα συντεταγμένων του CAD μοντέλου με το σύστημα συντεταγμένων μέτρησης της συστοιχίας. Έχοντας επίσης στην διάθεσή μας τον μετασχηματισμό \mathbf{X} που συνδέει το σύστημα συντεταγμένων μέτρησης με το αντίστοιχο του άκρου του βραχίονα, καθώς και τον μετασχηματισμό \mathbf{B}_i του άκρου ως προς τη βάση του βραχίονα μέσω της επίλυσης του ευθέως κινηματικού προβλήματος,

μπορούμε απ' ευθείας να συσχετίσουμε το σύστημα συντεταγμένων του CAD μοντέλου με το αντίστοιχο της βάσης του βραχίονα σύμφωνα με τον παρακάτω μετασχηματισμό:

$$\mathbf{T}_i = \mathbf{B}_i \mathbf{X} \mathbf{M}_i \quad (5.5)$$

Κατ' αυτό τον τρόπο ο βραχίονας ξέρει ανά πάσα στιγμή την θέση και τον προσανατολισμό του παρατηρούμενου αντικειμένου ως προς το δικό του σύστημα συντεταγμένων αναφοράς (βάσης). Επαναλαμβάνοντας την διαδικασία σε διαδοχικές χρονικές στιγμές ($i=1,2,\dots,n$) οι απόλυτες συντεταγμένες της τροχιάς του αντικειμένου μπορούν να υπολογίζονται συνεχώς και να στέλνονται στον ρομποτικό ελεγκτή ως εντολές κίνησης του βραχίονα.

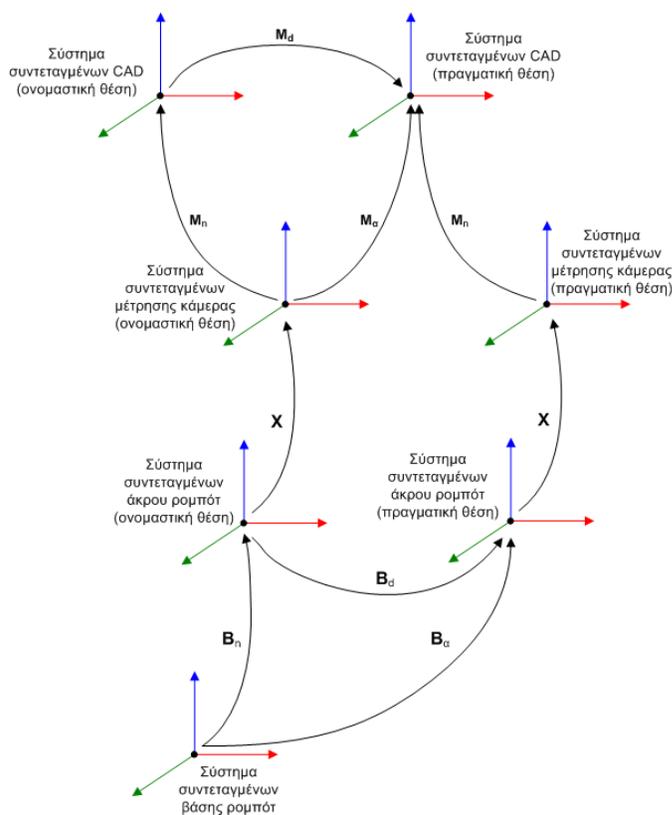
Δυστυχώς τα πειραματικά αποτελέσματα αποδεικνύουν ότι η παραπάνω προσέγγιση έχει πολύ περιορισμένη ακρίβεια, γεγονός που οφείλεται κυρίως στην συσσώρευση των ανακρίβειών υπολογισμού των ενδιάμεσων μετασχηματισμών \mathbf{B}_i , \mathbf{X} και \mathbf{M}_i στον τελικό μετασχηματισμό \mathbf{T}_i . Ειδικά ο μετασχηματισμός \mathbf{B}_i είναι ιδιαίτερα επιρρεπής σε ανακρίβειες, καθώς η κινηματική αλυσίδα του βραχίονα επηρεάζεται από παράγοντες όπως οι διαστολές-συστολές των συνδέσμων λόγω μεταβολών θερμοκρασίας, η μηχανική κάμψη τους λόγω βάρους, τα σφάλματα θέσης των κινητήρων, καθώς και από ένα πλήθος άλλων παραγόντων [Fu et al., 1987; Τζαφέστας, 2003]. Παρ' όλο που οι περισσότεροι κατασκευαστές βιομηχανικών ρομπότ δίνουν στις προδιαγραφές τους ακρίβειες της τάξης των 0.1-0.3mm, αυτές αναφέρονται στην επαναληψιμότητα του βραχίονα και όχι στην απόλυτη ακρίβεια του παρεχόμενου από τον ρομποτικό ελεγκτή μετασχηματισμού \mathbf{B}_i , η οποία είναι προφανώς πολύ χειρότερη.

Η περιορισμένη ακρίβεια της μεθόδου απ' ευθείας συσχέτισης των συστημάτων συντεταγμένων του CAD μοντέλου και της βάσης του βραχίονα επιβάλλει την χρήση μιας διαφορετικής προσέγγισης για την επίτευξη του στόχου. Αναφερόμενοι στο Σχ. 5.9, έστω ότι το κινούμενο αντικείμενο βρίσκεται σε μια αρχική θέση μέσα στο χώρο εργασίας του βραχίονα, την οποία εφεξής θα ονομάζουμε *ονομαστική θέση*, η οποία βρίσκεται μέσα στο πεδίο όρασης της συστοιχίας μέτρησης. Λαμβάνοντας μια ψηφιακή εικόνα με την κάμερα και ανιχνεύοντας τις φωτεινές γραμμές λείζερ (βλ. Σχ. 5.5) μπορούμε να ανακατασκευάσουμε τις 3D συντεταγμένες ενός μεγάλου πλήθους σημείων πάνω στην επιφάνεια του αντικειμένου. Εφαρμόζοντας τον αλγόριθμο ICP για το πλήθος των μετρούμενων σημείων μπορούμε να υπολογίσουμε τον μετασχηματισμό

συντεταγμένων \mathbf{M}_n που συνδέει το σύστημα συντεταγμένων του CAD μοντέλου στην ονομαστική του θέση με το αντίστοιχο των μετρήσεων της συστοιχίας μέτρησης (βλ. Σχ. 5.9).

Έστω ότι το αντικείμενο μετακινείται από την ονομαστική του θέση σε μια άλλη τυχαία θέση (*πραγματική θέση*) μέσα στο πεδίο όρασης της συστοιχίας μέτρησης. Προφανώς η θέση και ο προσανατολισμός του CAD συστήματος συντεταγμένων θα μεταβληθεί ακολουθώντας την κίνηση του αντικειμένου, σε αντίθεση με το σύστημα συντεταγμένων μέτρησης που παραμένει σταθερό, καθώς το άκρο του βραχίονα άρα και η συστοιχία μέτρησης παραμένουν στην αρχική (ονομαστική) τους θέση. Επαναλαμβάνοντας την διαδικασία τρισδιάστατης ανακατασκευής της επιφάνειας του αντικειμένου για την νέα του θέση και εφαρμόζοντας εκ νέου τον αλγόριθμο ICP θα πάρουμε έναν διαφορετικό μετασχηματισμό συντεταγμένων \mathbf{M}_a (βλ. Σχ. 5.9).

Στόχος μας είναι να μεταφράσουμε τη μεταβολή της θέσης και προσανατολισμού του



Σχήμα 5.9 Συστήματα και μετασχηματισμοί συντεταγμένων που χρησιμοποιούνται στην παρακολούθηση τροχιάς.

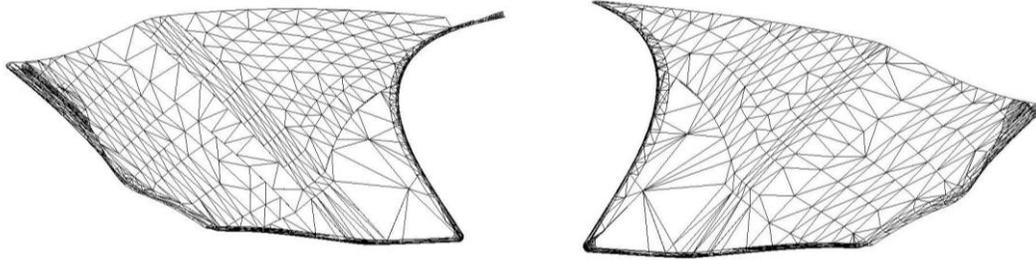
CAD συστήματος συντεταγμένων σε μια αντίστοιχη κίνηση του άκρου του βραχίονα, ώστε μετά την εκτέλεσή της η συστοιχία μέτρησης να «βλέπει» το αντικείμενο ακριβώς με τον ίδιο τρόπο όπως και στην ονομαστική του θέση. Μαθηματικά αυτό μεταφράζεται ότι επιθυμούμε η συστοιχία μέτρησης να βρεθεί σε μια τελική θέση, από την οποία πραγματοποιώντας εκ νέου τρισδιάστατη ανακατασκευή της επιφάνειας του αντικειμένου κι εφαρμόζοντας εκ νέου τον αλγόριθμο ICP να πάρουμε τον ίδιο μετασχηματισμό \mathbf{M}_n με εκείνον της ονομαστικής θέσης του

αντικειμένου. Ουσιαστικά πρέπει να βρούμε μια σχετική μετατόπιση \mathbf{B}_d του άκρου του βραχίονα ως προς την αρχική του θέση \mathbf{B}_n , έτσι ώστε η καινούρια του διορθωμένη θέση $\mathbf{B}_a = \mathbf{B}_n \mathbf{B}_d$ να τοποθετεί την συστοιχία μέτρησης στην επιθυμητή θέση και προσανατολισμό. Από το Σχ. 5.9 είναι προφανές ότι ο μετασχηματισμός συντεταγμένων που αντιστοιχεί στην απαιτούμενη διόρθωση θέσης και προσανατολισμού του άκρου του βραχίονα δίνεται από την εξίσωση:

$$\mathbf{B}_d = \mathbf{T}_n \mathbf{M}_d \mathbf{T}_a^{-1} = \mathbf{X} \mathbf{M}_d \mathbf{M}_n^{-1} \mathbf{X}^{-1} \quad (5.6)$$

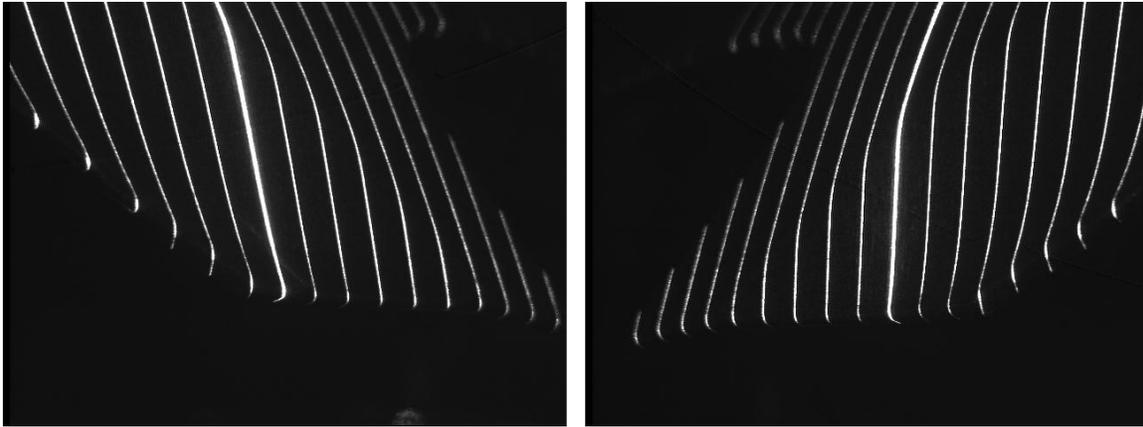
Οδηγώντας το άκρο του βραχίονα στην καινούρια του θέση \mathbf{B}_a διαπιστώνουμε πράγματι ότι το ρομπότ ακολουθεί την κίνηση του αντικειμένου, με αποτέλεσμα η συστοιχία μέτρησης να «βλέπει» το αντικείμενο ακριβώς με τον ίδιο τρόπο όπως και στην αρχική του θέση.

Η παραπάνω τεχνική ελέγχου περιγράφει τον υπολογισμό ενός και μόνο μετασχηματισμού διόρθωσης θέσης και προσανατολισμού του άκρου του βραχίονα, ώστε να παρακολουθήσει μία μόνο μετακίνηση ενός παρατηρούμενου αντικειμένου. Εάν το αντικείμενο διαγράφει μια συνεχή τροχιά στον χώρο τότε ο βραχίονας θα πρέπει να είναι σε θέση να το παρακολουθεί συνεχώς, μεταβάλλοντας ανάλογα τη θέση και τον προσανατολισμό του άκρου του. Στην περίπτωση αυτή θα πρέπει να δημιουργηθεί ένα σύστημα αυτομάτου ελέγχου κλειστού βρόχου και διακριτού χρόνου, όπως αυτό που απεικονίζεται στο Σχ. 1.2. Ο υπολογισμός του μετασχηματισμού διόρθωσης \mathbf{B}_d , με βάση τις μετρήσεις τρισδιάστατης ανακατασκευής που θα πραγματοποιεί η συστοιχία μέτρησης, θα πρέπει να επαναλαμβάνεται σε περιοδικές και όσο το δυνατόν πυκνότερες χρονικές στιγμές, έτσι ώστε το σύστημα να μπορεί να λειτουργήσει σε πραγματικό χρόνο (real-time). Σημαντική παράμετρος για την ευστάθεια και ακρίβεια του προτεινόμενου συστήματος αυτομάτου ελέγχου διακριτού χρόνου είναι η περίοδος του (χρονική καθυστέρηση που απαιτείται για τους απαραίτητους υπολογισμούς) να είναι αρκετά μικρή ώστε το άκρο του βραχίονα να προλαβαίνει να παρακολουθεί τις μεταβολές θέσης του παρατηρούμενου αντικειμένου.



Σχήμα 5.10 CAD μοντέλα δύο μεταλλικών επιφανειών.

Είναι αρκετά δύσκολο να ελέγξουμε την ακρίβεια του προτεινόμενου συστήματος παρακολούθησης τροχιάς, καθώς αρκετές διακριτές διαδικασίες χρησιμοποιούνται για την επίτευξη του επιθυμητού στόχου. Για την εξαγωγή των πειραματικών αποτελεσμάτων χρησιμοποιείται ένα βιομηχανικό ρομπότ της KUKA τύπου KR-15/1 και μια συστοιχία αποτελούμενη από μια ψηφιακή κάμερα Sony XC-HR50 σε συνδυασμό με μια συσκευή εκπομπής φωτός λέιζερ, τα οποία είναι σταθερά στερεωμένα στο άκρο του βραχίονα. Αρχικά πραγματοποιείται η βαθμονόμηση της κάμερας και των επιπέδων των γραμμών λέιζερ, χρησιμοποιώντας τις τεχνικές που περιγράφονται στο Κεφ. 5.2, καθώς και η επίλυση του προβλήματος της βαθμονόμησης βραχίονα – μηχανικής όρασης ακολουθώντας την διαδικασία του Κεφ. 3.2. Για τα πειράματά μας χρησιμοποιούνται δύο μεταλλικές επιφάνειες με γνωστά CAD δεδομένα (Σχ. 5.10) των οποίων η τοπολογία δεν είναι ιδιαίτερα πολύπλοκη, καθώς θέλουμε ν' αποφύγουμε προβλήματα σκιάσεων και διαθλάσεων των γραμμών λέιζερ που προσπίπτουν πάνω τους. Χρησιμοποιώντας ένα εκ των δύο μεταλλικών αντικειμένων του Σχ. 5.10, τοποθετούμε το άκρο του βραχίονα σε μια αρχική (ονομαστική) θέση ώστε η συστοιχία μέτρησης να μπορεί να παρατηρεί την επιφάνεια του αντικειμένου (Σχ. 5.11). Πραγματοποιώντας μέτρηση τρισδιάστατης ανακατασκευής ενός πλήθους σημείων της επιφάνειας του αντικειμένου κι εφαρμόζοντας τον αλγόριθμο ICP μπορεί να υπολογιστεί ο μετασχηματισμός συντεταγμένων \mathbf{M}_n που συνδέει το σύστημα συντεταγμένων μέτρησης της συστοιχίας και του CAD μοντέλου στην ονομαστική (αρχική) θέση του αντικειμένου. Εκτός απ' τον μετασχηματισμό συντεταγμένων \mathbf{M}_n , ο αλγόριθμος ICP παρέχει επίσης μια τιμή μέσου τετραγωνικού σφάλματος d_n , η οποία αντιστοιχεί στην ακρίβεια «ταιριάσματος» μεταξύ του πλήθους



Σχήμα 5.11 Ψηφιακές εικόνες των μεταλλικών επιφανειών του Σχ. 5.10.

των μετρούμενων σημείων ανακατασκευής και του πλέγματος των τριγώνων του CAD μοντέλου. Προφανώς ένα τέλειο ταίριασμα ($d_n=0$) δεν είναι δυνατό στην πράξη, λόγω ανακριβειών στους αλγορίθμους επεξεργασίας εικόνας και βαθμονόμησης της κάμερας. Στον πίνακα 1 φαίνονται τα πειραματικά αποτελέσματα της εφαρμογής του αλγορίθμου ICP χρησιμοποιώντας και τα δύο αντικείμενα του Σχ. 5.10. Μια πρώτη παρατήρηση είναι ότι η χρήση των κατανομών KD-tree για τον υπολογισμό των ελαχίστων αποστάσεων έχει σαν αποτέλεσμα την επιτάχυνση του αλγορίθμου ICP κατά 30-40 φορές. Επίσης το μέσο τετραγωνικό σφάλμα «ταιριάσματος» ($<0.1\text{mm}^2$) είναι πολύ ικανοποιητικό για πραγματικές εφαρμογές και ανήκει στην ίδια κλίμακα μεγέθους με τις ακρίβειες των αλγορίθμων βαθμονόμησης κάμερας κι επεξεργασίας εικόνας.

Στην συνέχεια μετακινούμε την χρησιμοποιούμενη μεταλλική επιφάνεια σε μια νέα θέση μέσα στο πεδίο όρασης της συστοιχίας μέτρησης, και χρησιμοποιώντας τις τεχνικές που περιγράφονται παραπάνω υπολογίζουμε τον μετασχηματισμό διόρθωσης θέσης και προσανατολισμού του άκρου του βραχίονα, σύμφωνα με την εξίσωση (5.6). Εάν όλοι οι υπολογισμοί ήταν απόλυτα ακριβείς και η κίνηση του βραχίονα χαρακτηριζόταν από απόλυτη ακρίβεια τότε στην τελική της θέση η συστοιχία μέτρησης θα «έβλεπε» την επιφάνεια του παρατηρούμενου αντικειμένου με τον ίδιο ακριβώς τρόπο όπως και στην αρχική του θέση. Αυτό σημαίνει ότι αν επαναλαμβάναμε την διαδικασία για την τελική αυτή θέση της συστοιχίας μέτρησης θα είχαμε ιδανικά $\mathbf{M}_a=\mathbf{M}_n \Rightarrow \mathbf{B}_d=\mathbf{0}$.

Γνωρίζοντας ότι σε πρακτικές εφαρμογές δεν υπάρχει απόλυτη ακρίβεια, δημιουργούμε ένα σύστημα κλειστού βρόχου, όπως αυτό του Σχ. 1.2, και μετακινώντας

το αντικείμενο από την ονομαστική στην τελική του θέση αφήνουμε το άκρο του βραχίονα να συγκλίνει σε μια νέα θέση τροφοδοτώντας τον ρομποτικό ελεγκτή με διορθώσεις θέσεις και προσανατολισμού σε κάθε επανάληψη. Στον πίνακα 2 φαίνονται οι διορθώσεις θέσης και προσανατολισμού του άκρου του βραχίονα συναρτήσει του αριθμού των επαναλήψεων πραγματοποίησης μετρήσεων τρισδιάστατης ανακατασκευής και υπολογισμού μετασχηματισμού διόρθωσης σύμφωνα με την εξίσωση (5.6). Είναι προφανές ότι ο βραχίονας πραγματοποιεί μια σχετικά μεγάλη κίνηση στην πρώτη επανάληψη, ακολουθώντας την μετατόπιση του παρατηρούμενου σώματος, και στην συνέχεια περιφέρεται γύρω από την νέα του θέση ισορροπίας με μετατοπίσεις της τάξης των 0.05-0.25mm και γωνίες Euler της τάξης των 0.01-0.03°. Τα μεγέθη αυτά είναι αντιπροσωπευτικά για όλα τα πειράματα που πραγματοποιήθηκαν χρησιμοποιώντας και τα δύο μεταλλικά αντικείμενα του Σχ. 5.10, εφόσον η μετακίνηση του αντικειμένου δεν είναι τόσο μεγάλη ώστε να βγει εκτός του πεδίου όρασης της συστοιχίας μέτρησης. Η ακρίβεια της προτεινόμενης μεθόδου είναι πολύ καλή, και φαίνεται ότι περιορίζεται περισσότερο από παράγοντες όπως η ανάλυση της κάμερας και η ανακρίβειες επεξεργασίας εικόνας, παρά από την ακρίβεια βαθμονόμησης της κάμερας και υπολογισμού του μετασχηματισμού βραχίονα – μηχανικής όρασης (Hand/Eye).

Πίνακας 1. Πειραματικά δεδομένα του αλγορίθμου ICP.

	3D σημεία × 3D τρίγωνα	επαναλήψεις ICP	Χρόνος σύγκλισης	Χρόνος σύγκλισης (K-D trees)	Μέσο τετραγωνικό σφάλμα (d_k)
Αντικείμενο 1	722×1423	15	4750 ms	110 ms	0.085782
Αντικείμενο 2	665×1172	7	1953 ms	63 ms	0.072229

Πίνακας 2. Διορθώσεις θέσης και προσανατολισμού του άκρου του βραχίονα συναρτήσει του αριθμού επαναλήψεων του συστήματος ελέγχου.

Επαναλήψεις διόρθωσης	X (mm)	Y (mm)	Z (mm)	A (°)	B (°)	C (°)
1	50.33	-35.25	-47.62	2.57	-1.75	2.05
2	0.31	-0.24	0.16	0.035	0.028	0.032
3	-0.23	0.16	0.2	0.021	-0.017	0.023
4	0.14	-0.10	-0.08	0.014	0.011	0.016
5	0.10	0.07	0.13	-0.010	0.008	-0.009

6 ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΔΗΜΟΣΙΕΥΣΕΙΣ

6.1 Συμπεράσματα διατριβής

Η παρούσα διατριβή πραγματεύεται το πρόβλημα της οπτικής παρακολούθησης τροχιάς ενός κινούμενου στο χώρο αντικειμένου, από το άκρο ενός βιομηχανικού ρομποτικού βραχίονα ο οποίος φέρει μια συσκευή μηχανικής όρασης. Η συσκευή μηχανικής όρασης μπορεί να είναι μια μονή κάμερα, μια στερεοσκοπική κεφαλή ή μια συστοιχία κάμερας και συσκευής λέιζερ. Η διατριβή ασχολείται με την παρακολούθηση τροχιάς κινούμενων αντικειμένων οι επιφάνειες των οποίων δεν περιλαμβάνουν εμφανή χαρακτηριστικά γεωμετρίας ή υφής που θα διευκόλυναν την οπτική αναγνώρισή τους, όπως για παράδειγμα γωνίες, σπές ή ακμές. Γι' αυτό το λόγο για την οπτική παρακολούθηση των αντικειμένων έχει επιλεγεί μια συσκευή μηχανικής όρασης που αποτελείται από μια μονή κάμερα και μια συσκευή λέιζερ. Με τη βοήθεια κατάλληλων αλγορίθμων που αναλύονται στο Κεφ. 5, οι ψηφιακές εικόνες που λαμβάνονται από την κάμερα και απεικονίζουν τις προβολές των γραμμών λέιζερ πάνω στην επιφάνεια του παρατηρούμενου αντικειμένου, μετατρέπονται σε μια πλήρως ανακατασκευασμένη τρισδιάστατη επιφάνεια. Η ανακατασκευασμένη αυτή επιφάνεια χρησιμοποιείται ως βάση για τον υπολογισμό του μετασχηματισμού συντεταγμένων του άκρου του βραχίονα, ώστε αυτός να παρακολουθήσει τη στροφή και τη μετατόπιση του κινούμενου στο χώρο αντικειμένου.

Οι περισσότερες εργασίες που έχουν γίνει πάνω στο αντικείμενο της οπτικής παρακολούθησης τροχιάς προϋποθέτουν την ύπαρξη ενός ή περισσότερων διακριτών χαρακτηριστικών στην επιφάνεια του παρατηρούμενου αντικειμένου. Σε αυτές τις περιπτώσεις το πρόβλημα της επεξεργασίας εικόνας και της τρισδιάστατης ανακατασκευής του παρατηρούμενου διακριτού χαρακτηριστικού του αντικειμένου είναι αρκετά εύκολη υπόθεση. Η πρωτοτυπία της παρούσας διατριβής έγκειται στο γεγονός ότι δεν προϋποθέτει την ύπαρξη εμφανών γεωμετρικών χαρακτηριστικών στην επιφάνεια του παρατηρούμενου αντικειμένου, γεγονός που καθιστά την προτεινόμενη μέθοδο γενικής χρήσης και ευρέως εφαρμόσιμη σε πρακτικές εφαρμογές. Η εφικτότητα και η

ακρίβεια του προτεινόμενου συστήματος παρακολούθησης τροχιάς επαληθεύονται από τα πειραματικά αποτελέσματα.

Για το σχεδιασμό και την υλοποίηση του προαναφερθέντος συστήματος παρακολούθησης τροχιάς ήταν απαραίτητη η επίλυση διαφόρων άλλων προβλημάτων, όπως για παράδειγμα η βαθμονόμηση της κάμερας και της συσκευής λείζερ, η βαθμονόμηση βραχίονα – μηχανικής όρασης, η επιλογή των κατάλληλων αλγορίθμων επεξεργασίας εικόνας και η τρισδιάστατη ανακατασκευή των παρατηρούμενων επιφανειών. Κατά την επίλυση των προβλημάτων της βαθμονόμησης βραχίονα – μηχανικής όρασης (Κεφ. 3) καθώς και της ταυτόχρονης βαθμονόμησης βραχίονα – μηχανικής όρασης και χώρου εργασίας (Κεφ. 4) χρησιμοποιείται μια καινούρια μέθοδος για τον υπολογισμό των άγνωστων μετασχηματισμών συντεταγμένων. Η καινοτομία της προτεινόμενης μεθόδου έγκειται στο γεγονός ότι οι μετασχηματισμοί που συνδέουν τα συστήματα συντεταγμένων της κάμερας και του μοντέλου βαθμονόμησης δεν εξάγονται απ' ευθείας από τον αλγόριθμο βαθμονόμησης της κάμερας όπως συμβαίνει με τις συμβατικές μεθόδους, αλλά από την τρισδιάστατη ανακατασκευή τεσσάρων χαρακτηριστικών σχημάτων (κύκλων) στην επιφάνεια του μοντέλου βαθμονόμησης. Η προτεινόμενη μέθοδος δίνει καλύτερα αποτελέσματα σε σχέση με τις συμβατικές μεθόδους, γεγονός που δικαιολογείται από την μεγαλύτερη ακρίβεια που προσφέρει η τεχνική της τρισδιάστατης ανακατασκευής σε σύγκριση με τους μετασχηματισμούς εξωτερικών παραμέτρων, όπως αυτοί λαμβάνονται από τον αλγόριθμο βαθμονόμησης κάμερας. Η ανωτερότητα της προτεινόμενης μεθόδου επαληθεύεται από τα πειραματικά αποτελέσματα.

6.2 Δημοσιεύσεις

Η έρευνα για την εκπόνηση της παρούσας διατριβής δημοσιεύτηκε και ανακοινώθηκε στις παρακάτω εργασίες σε επιστημονικά περιοδικά και διεθνή συνέδρια.

Περιοδικά

- D. Aristos and S.G.Tzafestas. An Extensive Analysis of the Simultaneous Robot-World and Hand-Eye Calibration Problem. *Int. Journal of Factory Automation, Robotics and Soft Computing*, Issue 1, pp. 7 –18, January 2008.

(Δημοσιεύτηκε επίσης στο βιβλίο, *Recent advances in Control Systems, Robotics and Automation – Second Edition*, Internationalsar, Palermo, Italy, ISBN: 978-88-901928-3-8.)

- D. Aristos and S.G.Tzafestas. A Method for the Registration of a Known CAD Model into the Workspace Frame of a Robot", *International Journal on Artificial Intelligence Tools*, pp.281-304, 2010.

Συνέδρια

- D. Aristos and S.G. Tzafestas. Simultaneous Robot-World and Hand-Eye Calibration Using a Stereo Head of Cameras, *Proc. of the International Mediterranean Modeling Multiconference (IMAACA 2005)*, pp. 41-46, 2005.
- D. Aristos and S.G.Tzafestas. Registration of a Known CAD Model into the Workspace Frame of a Robot. *Proc. of the European Control Conference 2007 (ECC'07), Kos, Greece (2007)*, pp. 2574-2581.
- D. Aristos and S.G.Tzafestas. Simultaneous Object Recognition and Position Tracking for Robotic Applications. *5th IEEE International Conference on Mechatronics (ICM 2009)*, Malaga, Spain, April 14-17, 2009.

ΒΙΒΛΙΟΓΡΑΦΙΑ

Ρομποτική και Μαθηματικά

Σπύρος Γ. Τζαφέστας (2003), Ρομποτική: Ανάλυση-Έλεγχος-Σχεδιασμός-Προγραμματισμός-Αίσθηση.

K. Fu, R. Gonzalez and C. Lee (1987). *Robotic: Control, Sensing, Vision and Intelligence. McGraw-Hill.*

R. Fletcher (1990). *Practical Methods of Optimization. John Wiley & Sons.*

P. E. Gill, W. Murray and M. H. Wright (1989). *Practical Optimization. Academic Press, London.*

J. J. Moré, D. C. Sorensen, K. E. Hillstrom, and B. S. Garbow (1984). The MINPACK Project, in *Sources and Development of Mathematical Software. W. J. Cowell, ed., Prentice-Hall*, pages 88-111.

J. J. Moré, B. S. Garbow, and K. E. Hillstrom (1980). *User Guide for MINPACK-1. Argonne National Laboratory Report ANL-80-74, Argonne, Ill.*

W. Press, B. Flannery, S. Teukolsky and W. Vetterling (1988). *Numerical Recipes in C : The Art of Scientific Computing. Cambridge University Press.*

G.H. Golub and C.F. Van Loan (1983). *Matrix Computations (3rd ed.). Baltimore, MD: Johns Hopkins, Univ. Press.*

Vladislav Kravchenko (2003), *Applied Quaternionic Analysis. Helderermann Verlag ISBN 3-88538-228-8.*

Βαθμονόμηση κάμερας

Zhengyou Zhang (2000). Flexible Camera Calibration By Viewing a Plane From Unknown Orientations. *IEEE Transactions on Robotics and Automation*, Vol. 22, No. 11, pp. 1330-1334.

R. Y. Tsai (1987). A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf tv cameras and lenses. *IEEE Journal of Robotics and Automation*, Vol. 3, No. 4, pp. 323-344.

Berthold K.P.Horn (2000). Tsai's Camera Calibration Method Revisited. Department of Electrical Engineering and Computer Science Massachusetts Institute of Technology.

O. Faugeras and G. Toscani (1986). The calibration problem in stereo. Proc. of the *IEEE Conference on Computer Vision and Pattern Recognition*, Miami Beach, FL, pp. 15-20.

J. González, J. Gámez, C. Artal and A. Cabrera (2005). Stability Study of Camera Calibration Methods. *VI Workshop en Agentes Físicos, WAF'2005, I Congreso Español de Informática (CEDI'2005)*, September 13-16, Granada, Spain.

W. Sun, J. R. Cooperstock (2006). An empirical evaluation of factors influencing camera calibration accuracy using three publicly available techniques. *Machine Vision and Applications*, 17-1, pp. 51-56.

Σαρωτές λέιζερ

F. Blais (2004). Review of 20 Years of Range Sensor Development. *Journal of Electronic Imaging* 13, 1, pp. 231–243.

Josep F. Collado (2004). New Methods for Triangulation-based Shape Acquisition using Laser Scanners. Ph.D. Thesis, ISBN: 84-689-3091-1.

Srdan Tocovic (2002). Adaptive 3D modeling of objects by combining shape from silhouette and shape from structured light. Diploma Thesis, PRIP-TR-68.

C. Brenner, J. Böhm, J. Gühring (1999). Photogrammetric calibration and accuracy evaluation of a cross-pattern stripe projector. *Proc. of the SPIE: Videometrics VI*, San Jose, USA, Vol. 3641, pp. 164-172.

Επεξεργασία εικόνας

Jain, R. Katsuri, B. Schunck (1995). *Machine Vision*. McGraw-Hill.

O. D. Faugeras (1993). *Three Dimensional Computer Vision. A Geometric View-point*. MIT Press, Boston.

M.R. Shortis, T.A. Clarke and T. Short (1994). *A comparison of some techniques for the subpixel location of discrete target images*. *Proc. of the SPIE: Videometrics III*, Boston, MA, USA, Vol. 2350, pp. 239-250.

R. Fisher and D. Naidu (1996). A Comparison of Algorithms for Subpixel Peak Detection. in Sanz (ed.) *Advances in Image Processing, Multimedia and Machine Vision*, Springer-Verlag, Heidelberg.

ICP / KD-trees

Paul J. Besl and Neil D. McKay (1992). A method for registration of 3-D shapes. *IEEE Transactions on Pattern analysis and Machine Intelligence*, Vol. 14, No. 2, pp. 239-256.

B. Horn (1987). Closed-form solution of absolute orientation using unit quaternions. *Journal of the Optical Society of America*, Vol. 4, No. 4, pp. 629-642.

J. H. Friedman, J. L. Bentley and R. A. Finkel (1977). An Algorithm for Finding Best Matches in Logarithmic Expected Time. *ACM Transactions on Mathematical Software*, Vol. 3, No. 3, pp. 209-226.

D. Cheverikov, D. Svirko and P. Krsek (2002). The trimmed iterative closest point algorithm. Proc. of the 3rd *International Conference on Pattern Recognition*, pp. 545-548.

Szymon Rusinkiewicz and Marc Levoy (2001). Efficient Variants of the ICP Algorithm. Proc. of the 3rd *International Conference on 3D Digital Imaging and Modeling*, pp. 145-152.

Michael Greenspan and Mike Yurick (2003). Approximate K-D Tree Search for Efficient ICP. Proc. of the 4th *International Conference on 3D Digital Imaging and Modeling*, pp. 442-448.

Timo Zinßer, Jochen Schmidt and Heinrich Niemann (2003). Performance Analysis of Nearest Neighbor Algorithms for ICP Registration of 3-D Point Sets. Proc. of *Vision, Modeling, and Visualization Conference*, pp. 199-206.

R. F. Sproull (1991). Refinements to Nearest-Neighbor Searching in k-Dimensional Trees. *Algorithmica*, Vol. 6, pp. 579-589.

Βαθμονόμηση βραχίονα – μηχανικής όρασης

Radu Horaud and Fadi Dornaika (1995). Hand-eye calibration. *International Journal of Robotics Research*, Vol. 14, No. 3, pp. 195-210.

N. Andreff, R. Horaud and B. Espiau (2001). Robot hand-eye calibration using structure-from-motion. *International Journal of Robotics Research*, Vol. 20, No. 3, pp. 228-248.

R. Y. Tsai and R. K. Lenz (1989). A new technique for fully autonomous and efficient 3D robotics hand/eye calibration. *IEEE Journal of robotics and automation*, Vol. 5, No. 3, pp. 345-358.

F. Park and B. Martin, *Robot sensor calibration: solving $AX=XB$ on the euclidean group*, IEEE Transactions on Robotics and Automation, Vol. 10, No. 5 (1994), pp. 717-721.

H. Zhuang, Z. Roth and R. Sudhakar (1994). Simultaneous robot/world and tool/flange calibration by solving homogeneous transformation of the form $AX=YB$. *IEEE Transactions on Robotics and Automation*, vol. 10, no. 4, pp. 549-554.

Fadi Dornaika and Radu Horaud (1998). Simultaneous robot/world and hand/eye calibration. *IEEE Transactions on Robotics and Automation*, vol. 14, no. 4, pp. 617-622.

Συστήματα οπτικής παρακολούθηση τροχιάς

S. A. Hutchinson, G. D. Hager and P. I. Corke (1996). A tutorial on visual servo control. *IEEE Transactions on Robotics and Automation*, Vol. 12, No. 5, pp. 651-670.

D. Xu, L.-K. Wang, Z.-G. Tu, M. Tan (2005). Hybrid Visual Servoing Control for Robotic Arc Welding Based on Structured Light Vision. *Journal of Acta Automatica Sinica*, Vol. 31, Part 4, pp. 596-605.

D. Khadraoui, G. Motyl, P. Martinet, J. Gallice, and F. Chaumette (1996). Visual Servoing in Robotics Scheme Using a Camera/Laser-Stripe Sensor. *IEEE Transactions on Robotics and Automation*, vol. 12, n. 5, pp. 743-750, USA.

J. Pagés, C. Collewet, F. Chaumette, J. Salvi (2005). Robust decoupled visual servoing based on structured light. *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, IROS'05 2*, pp. 2676-2681.

J. Pagés, C. Collewet, F. Chaumette, J. Salvi (2006). An approach to visual servoing based on coded light. *IEEE International Conference on Robotics and Automation*, pp. 4118-4123.

J.T. Lapreste, F. Jurie, M. Dhome, F. Chaumette (2004). An efficient method to compute the inverse Jacobian matrix in visual servoing. *Proc. of IEEE International Conference on Robotics and Automation*, Volume 1, pp. 727 – 732.