



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ
ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΣΥΣΤΗΜΑΤΩΝ ΜΕΤΑΔΟΣΗΣ ΠΛΗΡΟΦΟΡΙΑΣ ΚΑΙ
ΤΕΧΝΟΛΟΓΙΑΣ ΥΛΙΚΩΝ

Ανάπτυξη, Εφαρμογή και Σύγκριση Τεχνικών Ευθυγράμμισης Βασισμένων σε Γεωμετρικούς Περιγραφείς

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Erjon A. Hoxha

Επιβλέπων : Γιώργος Ματσόπουλος
Επ. Καθηγητής Ε.Μ.Π.

Αθήνα, Απρίλιος 2015



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ
ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΣΥΣΤΗΜΑΤΩΝ ΜΕΤΑΔΟΣΗΣ ΠΛΗΡΟΦΟΡΙΑΣ ΚΑΙ
ΤΕΧΝΟΛΟΓΙΑΣ ΥΛΙΚΩΝ

Ανάπτυξη, Εφαρμογή και Σύγκριση Τεχνικών Ευθυγράμμισης Βασισμένων Σε Γεωμετρικούς Περιγραφείς

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Erjon A. Hoxha

Επιβλέπων : Γιώργος Ματσόπουλος
Επ. Καθηγητής Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 3/4/2015

.....
Γ. Ματσόπουλος
Επ. Καθηγητής Ε.Μ.Π.

.....
Ν. Ουζούνογλου
Καθηγητής Ε.Μ.Π.

.....
Δ. Κουτσούρης
Καθηγητής Ε.Μ.Π.

Αθήνα, Απρίλιος 2015

.....
Erjon A. Hoxha

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

Copyright © Erjon A. Hoxha, 2015

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.

Περίληψη

Το αντικείμενο της παρούσας διπλωματικής εργασίας είναι η ανάπτυξη και η παρουσίαση τεχνικών ευθυγράμμισης τρισδιάστατων νεφών σημείων που βασίζονται σε ζεύγη αντίστοιχων σημείων. Η διαδικασία της ευθυγράμμισης έχει ως στόχο την εύρεση κατάλληλου γεωμετρικού μετασχηματισμού μέσω του οποίου εξομαλύνονται οι γεωμετρικές αποκλίσεις ανάμεσα σε διαφορετικές απεικονίσεις του ίδιου αντικειμένου.

Αρχικά, πραγματοποιείται μια επισκόπηση των τεχνικών ευθυγράμμισης που συναντώνται στην βιβλιογραφία και κατόπιν ταξινομούνται βάσει συγκεκριμένων κριτηρίων. Παρουσιάζονται οι τεχνικές λήψης των δεδομένων εισόδου και αναλύεται ο αισθητήρας Microsoft Kinect.

Στις συγκεκριμένες μεθόδους, ο προσδιορισμός των αντίστοιχων ζευγών σημείων μεταξύ των δυο συνόλων δεδομένων γίνεται συγκρίνοντας τα σημεία με βάση τους γεωμετρικούς περιγραφείς. Ένας γεωμετρικός περιγραφέας κωδικοποιεί τις γεωμετρικές ιδιότητες μίας καθορισμένης περιοχής γύρω από ένα σημείο, με στόχο την μοναδική περιγραφή και αναγνώρισή του σε όλο το σύνολο των σημείων, χωρίς να επηρεάζεται από το θόρυβο ή την πυκνότητα του νέφους σημείων. Ο καθένας από αυτούς χρησιμοποιεί διαφορετικά γεωμετρικά χαρακτηριστικά και υπολογίζεται με διαφορετικό τρόπο.

Ο υπολογισμός των γεωμετρικών περιγραφέων για κάθε σημείο αποτελεί ένα βήμα της πλήρους μεθοδολογίας. Στα επόμενα βήματα, τα λανθασμένα αντίστοιχα σημεία απορρίπτονται και από τα υπόλοιπα ζεύγη αντίστοιχων σημείων υπολογίζεται ένας ολικός γεωμετρικός μετασχηματισμός που στην συνέχεια χρησιμοποιείται για την ευθυγράμμιση των δυο νεφών σημείων.

Οι προτεινόμενες τεχνικές ευθυγράμμισης εφαρμόστηκαν σε ένα σύνολο διαθέσιμων τρισδιάστατων νεφών σημείων, τα οποία λήφθηκαν με τον αισθητήρα Microsoft Kinect. Η αποδοτικότητα και η ακρίβειά τους αξιολογήθηκε τόσο ποιοτικά όσο και ποσοτικά. Οι μέθοδοι με βάση τους γεωμετρικούς περιγραφείς συγκρίνονται μεταξύ τους καθώς και με την μέθοδο ευθυγράμμισης που βασίζεται στον επαναληπτικό αλγόριθμο πλησιέστερων σημείων ICP. Τέλος, σχολιάζεται ο καθορισμός των παραμέτρων για τις οποίες η κάθε μέθοδος επιτυγχάνει ευθυγράμμιση με την ελάχιστη απόκλιση μεταξύ των δυο απεικονίσεων.

Λέξεις Κλειδιά

Γεωμετρικός περιγραφέας, Fast Point Feature Histogram FPFH, κάθετα διανύσματα, Signature of Histograms of Orientation SHOT, 3D Shape Context 3DSC, Unique Shape Context USC, αλγόριθμος πλησιέστερων σημείων ICP, ευθυγράμμιση, Point Cloud Library PCL, νέφος σημείων, συμπαγής μετασχηματισμός, αισθητήρες βάθους, Microsoft Kinect.

Abstract

The subject of this thesis is the development and presentation of registration techniques of 3D point clouds based on pairs of corresponding points. Such a method's goal is to find the relative positions and orientations of the separately acquired views in a global coordinate framework, such that the intersecting areas between them overlap perfectly.

Initially, an introduction to the registration techniques found in the literature is performed followed by a classification according to specific criteria. The input data acquisition methods and depth sensors are presented, furthermore Microsoft Kinect sensor is explicitly analyzed.

The proposed methods are based on matching point descriptors in order to identify corresponding points between the two datasets. 3D descriptors encode information about the surrounding geometry of a point in order to uniquely describe and identify them across multiple points in the cloud, no matter the noise, resolution or transformation. Each one uses different features and has its own method of computing unique values of a point.

The calculation of a geometrical descriptor represents only one step of the complete methodology. In the next steps, wrong correspondence pairs are rejected in order to improve the overall accuracy and the remaining correspondences are used to estimate the final global rigid transformation, which is then used to align the two point clouds.

The proposed registration methods were applied to a set of available 3D point clouds which were obtained by the Microsoft Kinect sensor. Their efficiency and accuracy were assessed both qualitatively and quantitatively. They were compared to each other as well as to the method based on ICP algorithm.

Finally, a set of parameters for which each method achieves the best score is analyzed.

Keywords

3D descriptor, Fast Point Feature Histogram FPFH, Normals, Signature of Histograms of Orientation SHOT, 3D Shape Context 3DSC, Unique Shape Context USC, Iterative Closest Point ICP, pairwise registration, Point Cloud Library PCL, Point Cloud, Rigid Transformation, Depth Sensors, Microsoft Kinect.

Ευχαριστίες

Αρχικά, θα ήθελα να εκφράσω ιδιαίτερες ευχαριστίες στον επιβλέποντα Επ. Καθηγητή Ε.Μ.Π. κ. Γιώργο Ματσόπουλο που μου εμπιστεύτηκε αυτή την διπλωματική εργασία, δίνοντας μου την ευκαιρία να ασχοληθώ με ένα τόσο ενδιαφέρον θέμα.

Ευχαριστώ επίσης τους Καθηγητές του Ε.Μ.Π. κ. Νικόλαο Ουζούνογλου και κ. Δημήτρη Κουτσούρη για την συμμετοχή τους στην τριμελή εξεταστική επιτροπή.

Θα ήθελα να ευχαριστήσω τον Δρ. Θεόδωρο Οικονομόπουλο για την πολύτιμη βοήθεια και την καθοδήγηση που μου παρείχε κατά την εκπόνηση αυτής της εργασίας.

Τέλος, ευχαριστώ θερμά την οικογένειά μου για την βοήθεια και την στήριξη καθ'όλη την διάρκεια των σπουδών μου.

Περιεχόμενα

Περίληψη	i
Abstract	iii
Ευχαριστίες	v
Περιεχόμενα	vii
Κατάλογος Εικόνων	ix
Κατάλογος Πινάκων	xi
1 Εισαγωγή	1
1.1 Σκοπός Διπλωματικής Εργασίας	1
1.2 Διάρθρωση Διπλωματικής Εργασίας	2
2 Επισκόπηση των Μεθόδων Ευθυγράμμισης Δεδομένων	4
2.1 Εισαγωγή	4
2.2 Ορισμός ευθυγράμμισης Δεδομένων	5
2.3 Ταξινόμηση μεθόδων ευθυγράμμισης δεδομένων	6
2.3.1 Διάσταση μετασχηματισμού	7
2.3.2 Είδος γεωμετρικού μετασχηματισμού	7
2.3.3 Η βάση της ευθυγράμμισης	10
2.3.4 Περιοχή εφαρμογής του μετασχηματισμού	14
2.3.5 Είδος των δεδομένων	15
2.3.6 Το υποκείμενο της ευθυγράμμισης	16
2.3.7 Το αντικείμενο της ευθυγράμμισης	16
2.3.8 Αυτοματοποίηση και διαδραστικότητα	16
2.3.9 Μέθοδοι βελτιστοποίησης και τεχνικές αναζήτησης	16
2.4 Κριτήρια αξιολόγησης τεχνικών ευθυγράμμισης	17
2.4.1 Συνάρτηση επιτυχίας	18
2.4.2 Αποτίμηση τεχνικών ευθυγράμμισης	20
3 Αισθητήρες βάθους	22
3.1 Στερεοσκοπικές κάμερες	22
3.2 Time of flight	23
3.3 Το δομημένο φως - Structured light	23
3.3.1 Δισδιάστατη ανάλυση	24
3.3.2 Τρισδιάστατη ανάλυση	25
3.4 Microsoft Kinect	26
3.4.1 Μαθηματικό μοντέλο υπολογισμού της απόστασης	27
3.4.2 Βαθμονόμηση του αισθητήρα	28
3.4.3 Πρόσθεση του χρώματος στο νέφος σημείων	29
3.4.4 Ακρίβεια	29
3.4.5 Ανάλυση του βάθους και πυκνότητα νέφους σημείων	30
4 Μέθοδοι Ευθυγράμμισης με Γεωμετρικούς Περιγραφείς	32
4.1 Εισαγωγή	32
4.2 Λήψη δεδομένων	33

4.3	Μεθοδολογία ευθυγράμμισης βασισμένης στα γεωμετρικά χαρακτηριστικά (Feature based Registration).....	34
4.3.1	Υποδειματοληψία.....	35
4.3.2	Γεωμετρικοί περιγραφείς (feature descriptors)	35
4.3.3	Εύρεση αντιστοιχιών σημείων - Correspondence estimation	47
4.3.4	Απόρριψη λανθασμένων αντιστοιχιών σημείων - Correspondence rejection.....	49
4.3.5	Υπολογισμός και εφαρμογή του μετασχηματισμού - Transform estimation and application.....	51
4.3.6	Συνάρτηση μέτρησης σφάλματος.....	52
4.4	Μεθοδολογία ευθυγράμμισης βασισμένης σε επαναλήψεις (Iteration based Registration).	52
4.4.1	Υποδειματοληψία.....	53
4.4.2	Υπολογισμός αντιστοιχιών σημείων	54
4.4.3	Απόρριψη λανθασμένων αντιστοιχιών.....	54
4.4.4	Υπολογισμός του μετασχηματισμού.....	54
4.4.5	Εφαρμογή του μετασχηματισμού.....	55
4.4.6	Κριτήρια τερματισμού.....	55
5	Αποτελέσματα.....	56
5.1	Ποιοτικά αποτελέσματα	56
5.2	Ποσοτικά αποτελέσματα	60
6	Σχολιασμός παραμέτρων εισόδου	63
6.1	Συντελεστής υποδειματοληψίας.....	63
6.2	Ακτίνα γειτονικής περιοχής κάθετων διανυσμάτων.....	65
6.3	Ακτίνα γειτονικής περιοχής γεωμετρικών περιγραφών	66
7	Συμπεράσματα, Συζήτηση και Μελλοντικές Βελτιώσεις και Επεκτάσεις.....	69
7.1	Συμπεράσματα της διπλωματικής εργασίας	69
7.2	Συζήτηση	70
7.3	Μελλοντικές βελτιώσεις και επεκτάσεις	71
	Βιβλιογραφία	72

Κατάλογος Εικόνων

Εικόνα 2.1. Διάγραμμα ροής για την διαδικασία ευθυγράμμισης δυο εικόνων, εικόνας αναφοράς (reference image) και εικόνας προς ευθυγράμμιση (floating image).....	6
Εικόνα 2.2. (α) Στερεοστατικό πλαίσιο. (β) Επεμβατική τοποθέτηση των fiducial markers στο κεφάλι του ασθενή (χρυσές σφαίρες).....	11
Εικόνα 2.3. Ευθυγράμμιση βασισμένη σε σημεία (Landmark based Registration). Αριστερά εικόνα από CT, μεσαία εικόνα από PET, δεξιά ευθυγραμμισμένες εικόνες.....	13
Εικόνα 2.4. Ευθυγράμμιση με βάση την κατάτμηση (Segmentation based Registration). Εικόνες LGE-MR που ελήφθησαν σε διαφορετικές χρονικές στιγμές (αριστερά και μεσαία). Δεξιά, ευθυγραμμισμένες εικόνες.....	14
Εικόνα 2.5. Παραδείγματα εφαρμογής διαστάσεων ολικών και τοπικών μετασχηματισμών στην αρχική εικόνα	15
Εικόνα 3.1. Παράδειγμα μιας στερεοσκοπικής κάμερας.....	22
Εικόνα 3.2. Σχηματική αναπαράσταση του μέτρησης βάθους με την ToF τεχνική.....	23
Εικόνα 3.3. Σχηματική αναπαράσταση του τρόπου μέτρησης βάθους σε δυο διαστάσεις με την χρήση structured light αισθητήρων	24
Εικόνα 3.4. Σχηματική αναπαράσταση του τρόπου μέτρησης βάθους σε 3D με την χρήση structured light αισθητήρων.....	25
Εικόνα 3.5. Microsoft Kinect.....	26
Εικόνα 3.6. Αναπαράσταση μέτρησης απόστασης ενός σημείου k του αντικειμένου από την κάμερα με χρήση των τριγωνομετρικών ιδιοτήτων	27
Εικόνα 4.1. Διάγραμμα ροής της μεθοδολογίας ευθυγράμμισης βασισμένης σε γεωμετρικούς περιγραφείς (descriptors)	34
Εικόνα 4.2. Κάθετα διανύσματα (normals). Για λόγους ευκρινείας απεικονίζεται μόνο ένα κάθετο διάνυσμα για κάθε 60 σημεία.....	36
Εικόνα 4.3. (α) Αριστερά και μεσαία: κάθετα διανύσματα με απροσδιόριστη φορά. Δεξιά: η αντίστοιχη σφαίρα κάθετων διανυσμάτων (normal sphere). (β) Αριστερά και μεσαία: κάθετα διανύσματα ομοιόμορφα προσανατολισμένα προς το σημείο παρατήρησης. Δεξιά: η αντίστοιχη σφαίρα κάθετων διανυσμάτων.....	37
Εικόνα 4.4. Γραφική αναπαράσταση του σταθερού συστήματος συντεταγμένων και των γωνιών α, φ, θ για ένα ζευγος σημειων p_s και p_t και τα αντίστοιχά τους κάθετα διανύσματα n_s και n_t	39
Εικόνα 4.5. Διάγραμμα της περιοχής της επιρροής για το Fast Point Feature Histogram (FPFH). Το προς εξέταση σημείο p_q (κόκκινο χρώμα) συνδέεται μόνο με τους k -γείτονες εντός της ακτίνας r (κύκλος με γκρι χρώμα, συνδέσεις με κόκκινο χρώμα). Κάθε γειτονικό σημείο συνδέεται με τα δικά του γειτονικά σημεία και τα SPFH ιστογράμματα που προκύπτουν πολλαπλασιασμένα με έναν συντελεστή βάρους μαζί με το SPFH ιστογράμματα του p_q σχηματίζουν το τελικό FPFH. Με έντονο μαύρο παρουσιάζονται οι συνδέσεις που συμμετέχουν δυο φορές.....	40

Εικόνα 4.6. Χωρισμός του σφαιρικού πλέγματος με κέντρο το σημείο p σε 32 τομείς και υπολογισμός του ιστογράμματος ενός τομέα της σφαίρας με βάση την γωνία που σχηματίζουν τα κάθετα διανύσματα np και np_i	41
Εικόνα 4.7. Γραφική αναπαράσταση των τομών της σφαίρας του περιγραφέα 3DSC που αντιστοιχούν στα διαστήματα του τελικού ιστογράμματος	42
Εικόνα 4.8. Σύγκριση (α) ιστογραμμάτων FPFH σωστών αντίστοιχων σημείων (β) ιστογραμμάτων FPFH λανθασμένων αντίστοιχων σημείων ή σημείων που δεν ταιριάζουν .	45
Εικόνα 4.9. Σύγκριση (α) υπογραφής SHOT σωστών αντίστοιχων σημείων (β) υπογραφής SHOT λανθασμένων αντίστοιχων σημείων ή σημείων που δεν ταιριάζουν.....	46
Εικόνα 4.10. Σύγκριση (α) ιστογράμματος USC σωστών αντίστοιχων σημείων (β) ιστογράμματος USC λανθασμένων αντίστοιχων σημείων ή σημείων που δεν ταιριάζουν ...	47
Εικόνα 4.11. (α) Γραφική αναπαράσταση διαχωρισμού δισδιάστατου χώρου με την kd-tree μέθοδο. Σημεία εισόδου (2,3), (5,4), (9,6), (4,7), (8,1), (7,2). (β) Το kd-tree σε μορφή δέντρου που προκύπτει από το (α).....	48
Εικόνα 4.12. (α) Αναπαράσταση αντιστοιχιών σημείων για δυο νέφη σημείων προσδιορισμένες με βάση την σύγκριση των γεωμετρικών περιγραφέν τους. (β) Σωστές αντιστοιχίες που έχουν απομείνει μετά την απόρριψη των λανθασμένων αντιστοιχιών με εφαρμογή της μεθόδου RANSAC.....	50
Εικόνα 4.13. Διάγραμμα ροής της μεθοδολογίας ευθυγράμμισης βασισμένης σε επαναλήψεις.	53
Εικόνα 4.14. Αναπαράσταση λανθασμένων αντιστοιχιών σημείων από ακραία σημεία κατά την ταύτιση τους με το πλησιέστερο σημείο με βάση την σχετική απόσταση.....	54
Εικόνα 5.1. Αποτελέσματα ευθυγράμμισης ενός εκτυπωτή. (α) Νέφος σημείων αναφοράς και προς ευθυγράμμιση πριν την ευθυγράμμιση. Αποτελέσματα με βάση (β) ICP αλγόριθμο (γ) κάθετα διανύσματα και περιγραφείς (δ) FPFH (ε) SHOT (στ) 3DSC (ζ) USC.....	57
Εικόνα 5.2. Αποτελέσματα ευθυγράμμισης μιας οθόνης. (α) Νέφος σημείων αναφοράς και προς ευθυγράμμιση πριν την ευθυγράμμιση. Αποτελέσματα με βάση: (β) ICP αλγόριθμο (γ) κάθετα διανύσματα και περιγραφείς (δ) FPFH (ε) SHOT (στ) 3DSC (ζ) USC.....	58
Εικόνα 5.3. Αποτελέσματα ευθυγράμμισης ενός τραπεζιού. (α) Νέφος σημείων αναφοράς και προς ευθυγράμμιση πριν την ευθυγράμμιση. Αποτελέσματα με βάση (β) ICP αλγόριθμος (γ) κάθετα διανύσματα και περιγραφείς (δ) FPFH (ε) SHOT (στ) 3DSC (ζ) USC.....	59
Εικόνα 6.1. Μέσο τετραγωνικό σφάλμα MSE συναρτήσεως του ποσοστού των σημείων	63
Εικόνα 6.2. Χρόνος εκτέλεσης των μεθόδων συναρτήσεως του ποσοστού των σημείων	64
Εικόνα 6.3. Μέσο τετραγωνικό σφάλμα MSE κάθετων διανυσμάτων συναρτήσεως της βοηθητικής ακτίνας.....	65
Εικόνα 6.4. Κάθετα διανύσματα. Αριστερά: μικρή ακτίνα γειτονικής περιοχής. Δεξιά: μεγάλη ακτίνα της γειτονικής περιοχής	66
Εικόνα 6.5. Μέσο τετραγωνικό σφάλμα MSE συναρτήσεως της ακτίνας της βοηθητικής περιοχής για διάφορους γεωμετρικούς περιγραφείς (α) FPFH (β) SHOT (γ) 3DSC (δ) USC....	67
Εικόνα 6.6. Μέσο τετραγωνικό σφάλμα MSE συναρτήσεως του αριθμού των επαναλήψεων για τον αλγόριθμο ICP.....	67

Κατάλογος Πινάκων

Πίνακας 2.1. Πίνακας συναρτήσεων επιτυχίας.....	20
Πίνακας 5.1. Μέσο τετραγωνικό σφάλμα (MSE) τεχνικών ευθυγράμμισης με χρήση των γεωμετρικών περιγραφών USC, 3DSC, SHOT, FPFH, κάθετων διανυσμάτων και του αλγορίθμου ICP.....	61
Πίνακας 5.2. Συνολικός χρόνος εκτέλεσης των τεχνικών ευθυγράμμισης με χρήση των γεωμετρικών περιγραφών USC, 3DSC, SHOT, FPFH, κάθετων διανυσμάτων και του αλγορίθμου ICP.....	62

1 Εισαγωγή

1.1 Σκοπός Διπλωματικής Εργασίας

Στην παρούσα διπλωματική εργασία παρουσιάζονται μέθοδοι ευθυγράμμισης τρισδιάστατων δεδομένων. Συγκεκριμένα, νεφών σημείων οι οποίες είναι αναπαραστάσεις της επιφάνειας των πραγματικών αντικειμένων στον τρισδιάστατο χώρο μέσω ενός συνόλου σημείων τα οποία χαρακτηρίζονται από συντεταγμένες x, y, z . Τα νέφη σημείων αποτελούν τα δεδομένα εισόδου των μεθοδολογιών αυτών. Με είσοδο δυο νέφη σημείων, το αποτέλεσμα μιας τεχνικής ευθυγράμμισης είναι ένας γεωμετρικός μετασχηματισμός ο οποίος περιστρέφει και μετακινεί το νέφος σημείων προς ευθυγράμμιση ως προς το νέφος σημείων αναφοράς έτσι ώστε να ελαχιστοποιηθούν οι γεωμετρικές αποκλίσεις μεταξύ τους. Η επιτυχία των μεθόδων ευθυγράμμισης που βασίζονται στα σημεία εξαρτάται από την ποιότητα των αντίστοιχων σημείων που προσδιορίζονται μεταξύ των δυο συνόλων δεδομένων. Υπάρχουσες μέθοδοι όπως ο επαναληπτικός αλγόριθμος πλησιέστερων σημείων ICP, παρέχουν ικανοποιητικά αποτελέσματα αλλά όταν τα δεδομένα διαφέρουν αρκετά μεταξύ τους ή παρουσία θορύβου αποτυγχάνουν να ευθυγραμμίσουν με επιτυχία. Στις μεθόδους που παρουσιάζονται η διαδικασία αντιστοίχισης σημείων πραγματοποιείται με την σύγκριση των γεωμετρικών περιγραφών. Για κάθε σημείο ενός νέφους σημείων υπολογίζεται ένας γεωμετρικός περιγραφέας ο οποίος είναι μια οντότητα στην οποία κωδικοποιούνται οι γεωμετρικές ιδιότητες μιας καθορισμένης περιοχής γύρω από ένα σημείο με στόχο την μοναδική περιγραφή και αναγνώριση του σε όλο το σύνολο των σημείων. Οι γεωμετρικοί περιγραφείς μπορούν να περιγράψουν μοναδικά ένα σημείο ανεξάρτητα από το σύστημα συντεταγμένων στο οποίο αυτό βρίσκεται ή από την παρουσία θορύβου.

Στα επόμενα βήματα της πλήρους μεθοδολογίας, τα λανθασμένα αντίστοιχα σημεία απορρίπτονται και από τα υπόλοιπα ορθά ζεύγη αντίστοιχων σημείων υπολογίζεται ο γεωμετρικός μετασχηματισμός.

Οι μέθοδοι οι οποίες βασίζονται σε γεωμετρικούς περιγραφείς μπορούν να χρησιμοποιηθούν ως εναλλακτικές των μεθόδων που χρησιμοποιούν τις τιμές της φωτεινότητας εφόσον τα υπό επεξεργασία δεδομένα έχουν μοντελοποιηθεί ως νέφη σημείων. Συνήθως σε αυτή την περίπτωση οι μέθοδοι με γεωμετρικούς περιγραφείς μπορούν να ευθυγραμμίσουν τα δεδομένα με μεγαλύτερη ταχύτητα. Επίσης η εφαρμογή τους είναι αναγκαία όταν τα δεδομένα προς ευθυγράμμιση που προέρχονται από

διαφορετικές απεικονιστικές τεχνικές έχουν σημαντικές χρωματικές αποκλίσεις και όπου οι τεχνικές με βάση την φωτεινότητα αποτυγχάνουν.

1.2 Διάρθρωση Διπλωματικής Εργασίας

Στο **κεφάλαιο 2** γίνεται μια επισκόπηση των μεθόδων ευθυγράμμισης δεδομένων. Δίνεται ο ορισμός της ευθυγράμμισης και στην συνέχεια παρουσιάζονται οι κατηγορίες στις οποίες διακρίνονται οι μέθοδοι αυτές. Τα κριτήρια με βάση τα οποία ταξινομούνται αυτές οι τεχνικές είναι η διάσταση του μετασχηματισμού, το είδος του γεωμετρικού μετασχηματισμού, η βάση της ευθυγράμμισης, η περιοχή εφαρμογής του μετασχηματισμού, το είδος των δεδομένων, το υποκείμενο και αντικείμενο της ευθυγράμμισης, η αυτοματοποίηση και η διαδραστικότητα της διαδικασίας και τέλος οι μέθοδοι βελτιστοποίησης και οι τεχνικές αναζήτησης. Στην τελευταία ενότητα παρουσιάζονται οι συναρτήσεις επιτυχίας και τα κριτήρια που πρέπει να ικανοποιεί μια τεχνική ευθυγράμμισης ώστε να εφαρμοστεί στην πράξη.

Στο **κεφάλαιο 3** περιγράφονται οι αισθητήρες βάθους και οι τεχνικές απεικόνισης βάσει των οποίων ταξινομούνται. Δίνεται έμφαση στην τεχνική του δομημένου φωτός στην οποία ανήκει ο αισθητήρας Microsoft Kinect με τον οποίον λαμβάνουμε τα δεδομένα που χρησιμοποιούνται στις μεθοδολογίες ευθυγράμμισης που παρουσιάζονται και αναλύονται στην παρούσα εργασία. Περιγράφονται διάφορα χαρακτηριστικά αυτού του αισθητήρα όπως ο τρόπος λήψης δεδομένων βάθους, η βαθμονόμηση του αισθητήρα, πιθανές πηγές σφαλμάτων στην απεικόνιση των αντικειμένων, το μοντέλο θορύβου και η πυκνότητα των νεφών σημείων.

Στο **κεφάλαιο 4** παρουσιάζονται οι μέθοδοι ευθυγράμμισης τρισδιάστατων νεφών σημείων βασισμένων σε γεωμετρικούς περιγραφείς. Κατόπιν, γίνεται μια συνοπτική περιγραφή μιας μεθοδολογίας βασισμένης στον αλγόριθμο ICP η οποία έχει έναν διαφορετικό τρόπο λειτουργίας, προκειμένου να πραγματοποιηθεί σύγκριση με τις πιο πάνω μεθόδους. Περιγράφεται ο τρόπος λήψης των δεδομένων τα οποία χρησιμοποιούνται στις τεχνικές αυτές. Στην συνέχεια αναλύονται λεπτομερώς όλα τα επιμέρους βήματα της κάθε μεθοδολογίας τα οποία είναι η υποδειγματοληψία, ο υπολογισμός των γεωμετρικών περιγραφών, η εύρεση αντίστοιχων σημείων, η απόρριψη λανθασμένων αντίστοιχων σημείων, ο υπολογισμός και η εφαρμογή του μετασχηματισμού. Περιγράφονται αναλυτικά οι γεωμετρικοί περιγραφείς, ο τρόπος κατασκευής τους και αναπαρίσταται γραφικά ο τρόπος εύρεσης αντίστοιχων σημείων με βάση τους περιγραφείς τους.

Στο **κεφάλαιο 5** παρουσιάζονται τα ποιοτικά και ποσοτικά αποτελέσματα της εφαρμογής των τεχνικών ευθυγράμμισης με γεωμετρικούς περιγραφείς σε ένα σύνολο διαθέσιμων δεδομένων. Η αποδοτικότητα και η ακρίβεια των τεχνικών αυτών αξιολογείται τόσο ποιοτικά μέσω της οπτικής εξέτασης όσο και ποσοτικά με την βοήθεια της συνάρτησης μέσου τετραγωνικού σφάλματος.

Στο **κεφάλαιο 6** υπολογίζονται και ελέγχονται οι βέλτιστες παράμετροι για τις οποίες η ευθυγράμμιση επιτυγχάνει την μικρότερη τιμή της συνάρτησης σφάλματος. Αυτές είναι ο συντελεστής δειγματοληψίας, η ακτίνα της γειτονικής περιοχής των κάθετων διανυσμάτων,

η ακτίνα της γειτονικής περιοχής των γεωμετρικών περιγραφών και ο αριθμός των επαναλήψεων για την μέθοδο που βασίζεται στον αλγόριθμο ICP.

Στο **κεφάλαιο 7** παρουσιάζονται τα συμπεράσματα της διπλωματικής εργασίας και πιθανές επεκτάσεις και μελλοντικές βελτιώσεις των συγκεκριμένων τεχνικών.

2 Επισκόπηση των Μεθόδων Ευθυγράμμισης Δεδομένων

2.1 Εισαγωγή

Η ευθυγράμμιση δεδομένων είναι ένα θεμελιώδες πρόβλημα το οποίο συναντάται σε ένα ευρύ φάσμα τομέων της ερευνητικής κοινότητας όπως στην μηχανική, στη ρομποτική, στην ιατρική, στην όραση υπολογιστών και στην επεξεργασία εικόνας. Η ευθυγράμμιση διαδραματίζει σημαντικό ρόλο στον τομέα της ιατρικής απεικόνισης όπου η συνεχής πρόοδος των απεικονιστικών τεχνικών συμπεριλαμβανομένων την μαγνητική τομογραφία MRI, την αξονική τομογραφία CT και την τομογραφία εκπομπής ποζιτρονίων PET, δίνει λεπτομερή in vivo ανατομική και λειτουργική πληροφορία.

Σε πολλούς τομείς της ιατρικής απαιτείται η εκτίμηση των διαφορών μεταξύ εικόνων ή δομών που απεικονίζουν ανατομικές πληροφορίες (CT, MRI) και οι οποίες ελήφθησαν από τον ίδιο ασθενή σε διαφορετικές χρονικές στιγμές. Η ποσοτική εκτίμηση των διαφορών αυτών είναι αναγκαία για την παρακολούθηση της εξέλιξης συγκεκριμένων ασθενειών ή για την αξιολόγηση της αποτελεσματικότητας μιας θεραπείας.

Η ευθυγράμμιση χρησιμοποιείται κυρίως στην ιατρική για να συνδυαστεί πληροφορία που προέρχεται από διαφορετικές απεικονιστικές τεχνικές. Η σύνθεση ή σύντηξη των ευθυγραμμισμένων δεδομένων προσφέρει την μέγιστη δυνατή διαγνωστική πληροφορία, διότι μια απεικονιστική τεχνική παρέχει είτε λειτουργική είτε ανατομική πληροφορία. Συγκεκριμένα, η αξονική τομογραφία CT, η μαγνητική τομογραφία MRI, οι υπέρηχοι και οι ακτίνες X, παρέχουν ανατομική πληροφορία αλλά όχι λειτουργική. Οι ιατρικές απεικονιστικές διατάξεις όπως η τομογραφία εκπομπής φωτονίων SPECT και η τομογραφία εκπομπής ποζιτρονίων PET, παρέχουν πληροφορίες για τη μεταβολική δραστηριότητα των μελετώμενων ιστών και οργάνων, σε αντίθεση με τις υπόλοιπες απεικονιστικές μεθόδους που παρέχουν πληροφορίες μόνο για τις ανατομικές λεπτομέρειες.

Η ανάγκη για ευθυγράμμιση μπορεί να προκύψει επίσης όταν υπάρχει χωρική μετατόπιση μεταξύ εικόνων η οποία συνήθως οφείλεται σε κινήσεις των ασθενών, σε ακούσιες κινήσεις του σώματος π.χ. αναπνοή, καρδιακή λειτουργία ή σε διατάξεις απεικονιστικών συστημάτων που δεν είναι σταθερές καθώς και η ανάπτυξη ανωμαλιών που προκαλούν παραμορφώσεις.

Τέλος, σε αρκετές περιπτώσεις η ευθυγράμμιση εφαρμόζεται για την απόκτηση μιας μεγαλύτερης αναπαράστασης σε 2D ή 3D ενός αντικειμένου ή μιας σκηνής από τις επιμέρους εικόνες ή σαρώσεις.

2.2 Ορισμός Ευθυγράμμισης Δεδομένων

Ως ευθυγράμμιση ορίζεται η διαδικασία εύρεσης κατάλληλου γεωμετρικού μετασχηματισμού, μέσω του οποίου εξομαλύνονται οι γεωμετρικές αποκλίσεις ανάμεσα σε διαφορετικές απεικονίσεις του ίδιου αντικειμένου ή ανάμεσα σε απεικονίσεις διαφορετικών αντικειμένων. Για την ευθυγράμμιση δυο εικόνων η μια επιλέγεται ως εικόνα προς ευθυγράμμιση, $I_F(r), r \in D_1 \subset \mathbb{R}^N$ και η άλλη ως εικόνα αναφοράς, $I_R(r), r \in D_2 \subset \mathbb{R}^N$ όπου N είναι η διάσταση των εικόνων. Η κάθε μέθοδος ευθυγράμμισης παράγει έναν μετασχηματισμό $T: D_1 \rightarrow \mathbb{R}^N$ τέτοιον ώστε η μετασχηματισμένη εικόνα $I_T(r)$

$$I_T(T(r)) = I_R(r) \text{ αν } T(r) \in D_2 \quad (2.1)$$

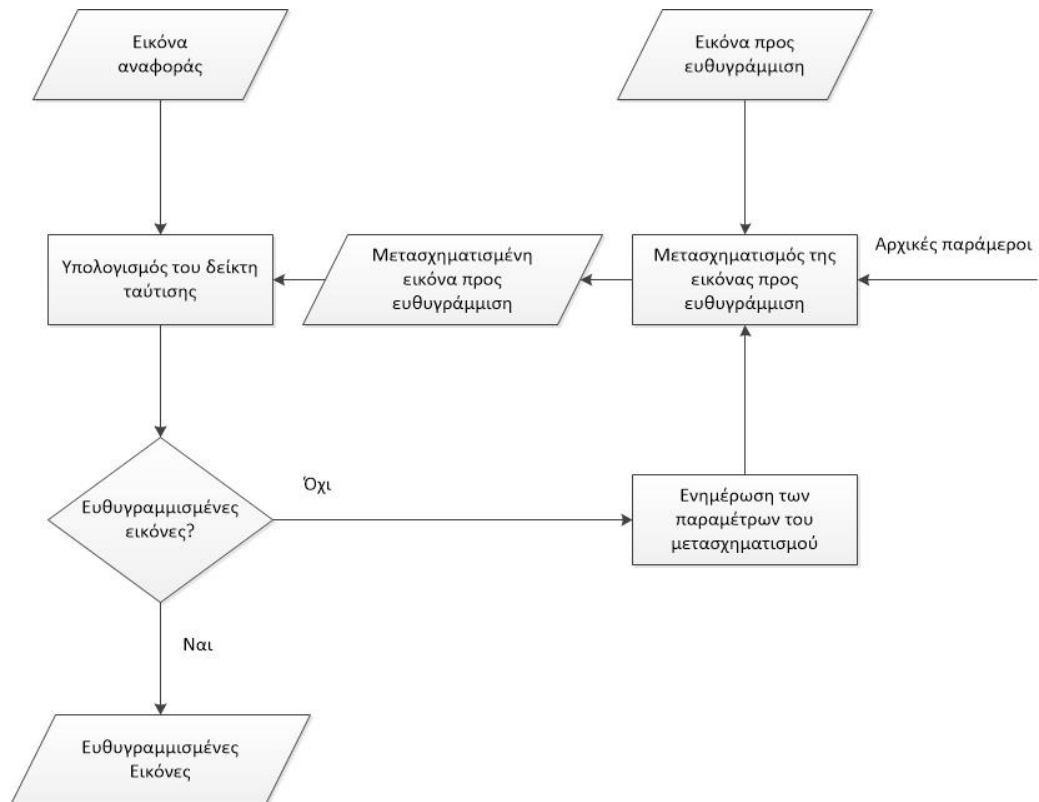
και η εικόνα αναφοράς $I_R(r)$ να περιέχουν την ίδια ανατομική πληροφορία στην θέση $r \in D_1$.

Το μεγάλο πλήθος προβλημάτων και των εφαρμογών τους έχουν οδηγήσει στην ανάπτυξη ενός μεγάλου αριθμού ανεξάρτητων μεθοδολογιών ευθυγράμμισης. Η κάθε τεχνική είναι συνήθως σχεδιασμένη για μια συγκεκριμένη εφαρμογή και όχι απαραίτητα για ένα συγκεκριμένο τύπο προβλημάτων ή δεδομένων. Επομένως, είναι αδύνατο να υλοποιηθεί μια γενική μέθοδος που θα εφαρμοστεί σε όλες τις περιπτώσεις. Παρ'όλα αυτά, όλες οι τεχνικές ευθυγράμμισης έχουν κάποια κοινά στοιχεία, ανεξάρτητα από την κατηγορία στην οποία ανήκουν. Πιο συγκεκριμένα, το σύνολο των τεχνικών ευθυγράμμισης απαρτίζεται από τα ακόλουθα τρία βασικά συστατικά τα οποία αναλύονται περαιτέρω στις ενότητες που ακολουθούν:

- Έναν γεωμετρικό μετασχηματισμό
- Μια συνάρτηση ομοιότητας ή σφάλματος
- Μια διαδικασία βελτιστοποίησης

Η ευθυγράμμιση των δεδομένων είναι μια επαναληπτική διαδικασία η οποία ευθυγραμμίζει την εικόνα I_F σε σχέση με την εικόνα αναφοράς I_R . Σε κάθε επανάληψη της διαδικασίας η υπό ευθυγράμμιση εικόνα I_F μετασχηματίζεται μέσω ενός γεωμετρικού μετασχηματισμού του οποίου οι παράμετροι παράγονται από μια μέθοδο βελτιστοποίησης. Οι νέες παράμετροι του μετασχηματισμού παράγονται μετά από τον υπολογισμό του κριτηρίου ομοιότητας ή σφάλματος μεταξύ εικόνας αναφοράς I_R και της μετασχηματισμένης εικόνας, I'_F . Η επαναληπτική διαδικασία τερματίζεται όταν μεγιστοποιηθεί η τιμή της συνάρτησης ομοιότητας ή αντίστοιχα όταν ελαχιστοποιηθεί η τιμή της συνάρτησης σφάλματος. Οι βέλτιστες παράμετροι του μετασχηματισμού εφαρμόζονται τότε στην υπό ευθυγράμμιση εικόνα I_F παράγοντας την τελική ευθυγραμμισμένη εικόνα.

Το κατωθι διάγραμμα αναπαριστά όλα τα επιμέρους βήματα που θα πρέπει να εκτελεστούν έτσι ώστε να πραγματοποιηθεί η ευθυγράμμιση των εικόνων και τα οποία αναλύονται εκτενέστερα στα παρακάτω κεφάλαια.



Εικόνα 2.1. Διάγραμμα ροής για την διαδικασία ευθυγράμμισης δύο εικόνων: εικόνας αναφοράς (reference image) και εικόνας προς ευθυγράμμιση (floating image).

2.3 Ταξινόμηση μεθόδων ευθυγράμμισης δεδομένων

Η διαδικασία της ευθυγράμμισης διαφοροποιείται με βάση το πρόβλημα που πρέπει να αντιμετωπιστεί καθώς και από την εφαρμοζόμενη μέθοδο. Μια ανασκόπηση γίνεται στις εργασίες [1] [2] [3] [4]. Η ταξινόμηση των μεθόδων ευθυγράμμισης που παρουσιάζονται σε αυτή την ενότητα βασίζεται στα κριτήρια που έχουν διατυπωθεί από τους van den Elsen, Pol και Viergever [5]. Τα κριτήρια είναι τα εξής:

- Διάσταση γεωμετρικού μετασχηματισμού
2D, 3D, πολυδιάστατος
- Είδος γεωμετρικού μετασχηματισμού
 - Συμπαγής
 - Affine
 - Προβολικός
 - Καμπυλόγραμμος

- Βάση της ευθυγράμμισης
 - Ευθυγράμμιση βασισμένη σε εξωγενή χαρακτηριστικά
 - Ευθυγράμμιση βασισμένη σε ενδογενή χαρακτηριστικά
 - Μέθοδοι με βάση τα σημεία
 - Μέθοδοι τμηματοποίησης
 - Μέθοδοι με βάση τη φωτεινότητα
- Περιοχή εφαρμογής του μετασχηματισμού
 - Μέθοδοι ολικού μετασχηματισμού
 - Μέθοδοι τοπικού μετασχηματισμού
- Είδος των δεδομένων
 - Μονοτροπικά δεδομένα (monomodal)
 - Πολυτροπικά δεδομένα (multimodal)
- Υποκείμενο της ευθυγράμμισης
- Αντικείμενο της ευθυγράμμισης
- Αυτοματοποίηση και διαδραστικότητα
 - Αυτόματες μέθοδοι
 - Ημιαυτόματες μέθοδοι
 - Χειροκίνητες μέθοδοι
- Μέθοδοι βελτιστοποίησης και τεχνικές αναζήτησης

2.3.1 Διάσταση μετασχηματισμού

Ανάλογα με την διάσταση των δεδομένων προς ευθυγράμμιση ο μετασχηματισμός μπορεί να είναι μονοδιάστατος, δισδιάστατος ή πολυδιάστατος. Όταν τα δεδομένα εισόδου είναι εικόνες προβολών ή τομές του ίδιου επιπέδου ο μετασχηματισμός είναι δισδιάστατος. Για ευθυγράμμιση τρισδιάστατων συνόλων δεδομένων, ο μετασχηματισμός χαρακτηρίζεται ως τρισδιάστατος. Τα δεδομένα εισόδου μπορούν να είναι χρονικές ακολουθίες δισδιάστατων ή τρισδιάστατων εικόνων, όπου ο χρόνος αποτελεί μια άλλη διάσταση. Σε αυτή την περίπτωση ο μετασχηματισμός χαρακτηρίζεται ως τρισδιάστατος και ως τετραδιάστατος αντίστοιχα.

2.3.2 Είδος γεωμετρικού μετασχηματισμού

Ο γεωμετρικός μετασχηματισμός ορίζει το μαθηματικό μοντέλο σύμφωνα με το οποίο οι υπό ευθυγράμμιση εικόνες θα μεταβληθούν έτσι ώστε η γεωμετρική τους απόκλιση να ελαχιστοποιηθεί. Οι γεωμετρικοί μετασχηματισμοί γενικά χωρίζονται σε άκαμπτους, συμπαγείς και εύκαμπτους ή ελαστικούς. Ανάλογα με τον βαθμό ελαστικότητας του, ένας μετασχηματισμός μπορεί να είναι συμπαγής (rigid), τύπου affine, προβολικός ή καμπυλόγραμμος. Στη συνέχεια παρουσιάζονται οι βασικοί γεωμετρικοί μετασχηματισμοί προσαρμοσμένοι σε τρεις διαστάσεις.

2.3.2.1 Συμπαγής μετασχηματισμός

Ο συμπαγής μετασχηματισμός διατηρεί αναλλοίωτο το μέγεθος και το σχήμα των αντικειμένων. Σε έναν συμπαγή μετασχηματισμό οι διαστάσεις και οι γωνίες δεν αλλάζουν.

Αναλύεται σε δυο συνιστώσες, την περιστροφή και την μετατόπιση. Σε τρεις διαστάσεις ο πίνακας περιστροφής είναι ένας 3×3 πίνακας που προκύπτει από το γινόμενο τριών πινάκων, R_x, R_y, R_z οι οποίοι αντιπροσωπεύουν την περιστροφή γύρω από τους x, y, z άξονες αντίστοιχα.

$$R = R_x \cdot R_y \cdot R_z \quad (2.2)$$

$$R = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \varphi_x & -\sin \varphi_x \\ 0 & \sin \varphi_x & \cos \varphi_x \end{pmatrix} \begin{pmatrix} \cos \varphi_y & 0 & \sin \varphi_y \\ 0 & 1 & 0 \\ -\sin \varphi_y & 0 & \cos \varphi_y \end{pmatrix} \begin{pmatrix} \cos \varphi_z & -\sin \varphi_z & 0 \\ \sin \varphi_z & \cos \varphi_z & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (2.3)$$

όπου $\varphi_x, \varphi_y, \varphi_z$ είναι οι γωνίες περιστροφής γύρω από τους άξονες x, y, z .

Η μετατόπιση ορίζεται από έναν 3×1 πίνακα ως εξής:

$$d = \begin{pmatrix} d_x \\ d_y \\ d_z \end{pmatrix} \quad (2.4)$$

όπου d_x, d_y, d_z η μετατόπιση στους άξονες x, y, z [6].

Έτσι, στον συμπαγή μετασχηματισμό, το σημείο $r = (x, y, z)$ μετασχηματίζεται στο σημείο $r' = (x', y', z')$ σύμφωνα με την σχέση

$$r' = Rr + d \quad (2.5)$$

Από τις (2.3) και (2.4) η (2.5) μπορεί να γραφτεί ως:

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = R \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} d_x \\ d_y \\ d_z \end{pmatrix} \quad (2.6)$$

Οι ολικοί συμπαγείς μετασχηματισμοί είναι οι πιο διαδεδομένοι στη βιβλιογραφία. Σε σχέση με άλλους μετασχηματισμούς, είναι πιο εύκολα και γρήγορα υλοποιήσιμοι καθώς μόνο 6 παράμετροι χρειάζεται να υπολογιστούν, τρεις παράμετροι για την περιστροφή και τρεις παράμετροι για την μετατόπιση. Ένα άλλο χαρακτηριστικό τους είναι ότι ένας ολικός συμπαγής μετασχηματισμός μπορεί να υπολογιστεί από μια μικρή περιοχή ενδιαφέροντος της αρχικής εικόνας και στη συνέχεια να εφαρμοστεί σε ολόκληρη την εικόνα.

2.3.2.2 Μετασχηματισμός τύπου affine

Ένας μετασχηματισμός τύπου affine διατηρεί τα σημεία, τις ευθείες γραμμές και τα επίπεδα. Οι ευθείες γραμμές μετασχηματίζονται σε ευθείες γραμμές διατηρώντας την παραλληλία τους. Όμως, ο μετασχηματισμός αυτός, δεν διατηρεί απαραίτητα τις γωνίες και

τα μήκη. Επομένως, το σχήμα του αντικειμένου που μετασχηματίζεται μπορεί να αλλάξει. Ο μετασχηματισμός τύπου affine μπορεί να αναλυθεί σε έναν γραμμικό μετασχηματισμό και μια μετατόπιση. Στις τρεις διαστάσεις ορίζεται από την εξίσωση:

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} d_x \\ d_y \\ d_z \end{pmatrix} \quad (2.7)$$

όπου $\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$ είναι ένας πίνακας με πραγματικές τιμές.

Ο μετασχηματισμός τύπου affine χρησιμοποιείται κυρίως στην ευθυγράμμιση των εικόνων που προέρχονται από τη μαγνητική τομογραφία MRI, οι οποίες εγγενώς περιέχουν γεωμετρικές παραμορφώσεις.

2.3.2.3 Προβολικός μετασχηματισμός

Σε έναν προβολικό μετασχηματισμό μια ευθεία γραμμή μετασχηματίζεται σε μια ευθεία γραμμή χωρίς απαραίτητα να διατηρείται η σχέση παραλληλίας τους [7]. Ο προβολικός μετασχηματισμός μπορεί να αναπαρασταθεί με έναν γραμμικό μετασχηματισμό με μια παραπάνω διάσταση. Στις τρεις διαστάσεις εκφράζεται με την μορφή:

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} x_1/w \\ y_1/w \\ z_1/w \end{pmatrix}, \text{ και } \begin{pmatrix} x_1 \\ y_1 \\ z_1 \\ w \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} \quad (2.8)$$

Το w αντιστοιχεί στην επιπλέον ομογενή συντεταγμένη. Συνήθως $a_{44} = 1$.

Η χρήση των προβολικών μετασχηματισμών στην βιβλιογραφία είναι περιορισμένη. Εφαρμόζονται κυρίως στην ευθυγράμμιση εικόνων 2D-3D, π.χ. στην ευθυγράμμιση εικόνων ακτινών X και τρισδιάστατων συνόλων δεδομένων τομογραφίας. Οι προβολικοί μετασχηματισμοί μπορούν να χρησιμοποιηθούν ως μια ειδική περίπτωση των καμπυλόγραμμων μετασχηματισμών όταν αυτοί αποτυγχάνουν ή παρουσιάζουν μεγάλη πολυπλοκότητα λόγω του πλήθους των παραμέτρων.

2.3.2.4 Καμπυλόγραμμοι μετασχηματισμοί

Σε αντίθεση με τους προηγούμενους μετασχηματισμούς οι καμπυλόγραμμοι μετασχηματισμοί δεν περιγράφονται με την μορφή σταθερών πινάκων αλλά εκφράζονται αλγεβρικά με την εξίσωση $(x', y', z') = F(x, y, z)$ όπου F είναι μια συνάρτηση μετασχηματισμού. Σε τέτοιους μετασχηματισμούς οι ευθείες γραμμές μετατρέπονται σε

καμπύλες. Γνωστοί καμπυλόγραμμοι μετασχηματισμοί είναι οι μετασχηματισμοί πολυωνυμικού τύπου:

$$\begin{aligned}x' &= \sum_{k=0}^K \sum_{l=0}^L \sum_{m=0}^M a_{klm} x^k y^l z^m \\y' &= \sum_{k=0}^K \sum_{l=0}^L \sum_{m=0}^M b_{klm} x^k y^l z^m \\z' &= \sum_{k=0}^K \sum_{l=0}^L \sum_{m=0}^M c_{klm} x^k y^l z^m\end{aligned}\tag{2.9}$$

όπου a, b, c παράμετροι και K, L, M σταθερές οι οποίες ορίζουν την τάξη των πολυωνύμων. Καθώς αυξάνει η τάξη των πολυωνύμων αυξάνει και ο αριθμός των παραμέτρων που πρέπει να προσδιοριστούν, με αποτέλεσμα την δυνατότητα για μεγαλύτερη ακρίβεια αλλά και αύξηση του υπολογιστικού κόστους. [8].

Οι καμπυλόγραμμοι μετασχηματισμοί χρησιμοποιούνται κυρίως στην αντιμετώπιση των ανατομικών διαφορών που μπορούν να παρουσιάσουν οι διαφορετικές εικόνες και περιορίζονται σε μεθόδους ευθυγράμμισης οι οποίες βασίζονται σε ενδογενή χαρακτηριστικά.

2.3.3 Η βάση της ευθυγράμμισης

Το κριτήριο αυτό κατηγοριοποιεί τις μεθόδους ευθυγράμμισης βάσει των χαρακτηριστικών που χρησιμοποιούνται για την ταύτιση των εικόνων. Βάση ευθυγράμμισης μπορεί να αποτελέσει η αντιστοίχιση συγκεκριμένων υποσυνόλων σημείων από τις δυο εικόνες, η αντιστοίχιση της μέσης τοπικής φωτεινότητας ή αντιστοίχιση γραμμών ή επιπέδων. Στην περίπτωση που η ευθυγράμμιση βασίζεται σε εξωγενή χαρακτηριστικά, ως βάση της διαδικασίας ευθυγράμμισης μπορούν να χρησιμοποιηθούν τεχνητά αντικείμενα τα οποία συνδυάζονται με την ανατομική πληροφορία των απεικονιζόμενων περιοχών [9] [10]. Διακρίνονται δυο κύριες κατηγορίες: στην πρώτη ανήκουν οι μέθοδοι βασισμένες στα εξωγενή χαρακτηριστικά και στην δεύτερη ανήκουν οι μέθοδοι βασισμένες στα ενδογενή χαρακτηριστικά.

2.3.3.1 Ευθυγράμμιση βασισμένη σε εξωγενή χαρακτηριστικά

Οι μέθοδοι με βάση τα εξωγενή χαρακτηριστικά ή εξωτερικά υποβοηθούμενες τεχνικές, βασίζονται σε τεχνητά αντικείμενα τα οποία τοποθετούνται στον ασθενή πριν από την διαδικασία της απεικόνισης. Τέτοια αντικείμενα που συνήθως αναφέρονται και ως fiducial markers [11], θα πρέπει να είναι πολύ προσεκτικά σχεδιασμένα έτσι ώστε να είναι ορατά και να μπορούν να ανιχνευτούν εύκολα και με ακρίβεια στις εικόνες ανεξάρτητα από την απεικονιστική τεχνική που θα χρησιμοποιηθεί. Τα σημάδια που αφήνουν τα αντικείμενα αυτά στις εικόνες χρησιμοποιούνται στην συνέχεια στην διαδικασία της ευθυγράμμισης.

Ίσως η πιο γνωστή μορφή ενός fiducial marker είναι το στερεοστατικό πλαίσιο [12]. Η συσκευή αυτή βιδώνεται στο κρανίο του ασθενή και είναι από τις τεχνικές που παρέχουν

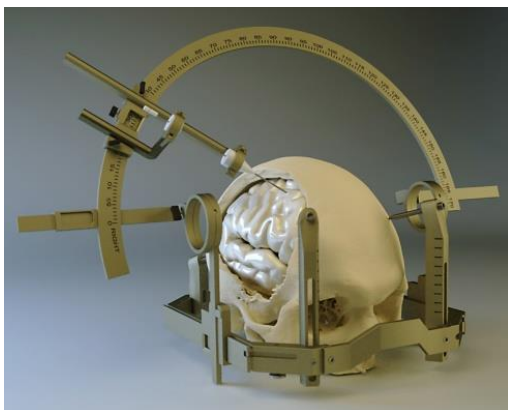
μεγάλη ακρίβεια ευθυγράμμισης. Το στερεοστατικό πλαίσιο χρησιμοποιείται επίσης σε μεγάλο βαθμό στην νευροχειρουργική για καθοδήγηση.

Οι εξωτερικά υποβοηθούμενες τεχνικές ευθυγράμμισης είναι σχετικά ακριβείς και πολύ γρήγορες. Η ευθυγράμμιση πραγματοποιείται με την ταύτιση των αντίστοιχων fiducial markers. Οι παράμετροι της ευθυγράμμισης μπορούν να υπολογιστούν εύκολα χωρίς την απαίτηση πολύπλοκων αλγορίθμων βελτιστοποίησης οι οποίοι επιβραδύνουν σημαντικά την διαδικασία. Συνήθως οι μέθοδοι που βασίζονται στα εξωγενή χαρακτηριστικά περιορίζονται στον συμπαγή μετασχηματισμό όπου επιτρέπονται μόνο περιστροφές και μετατοπίσεις.

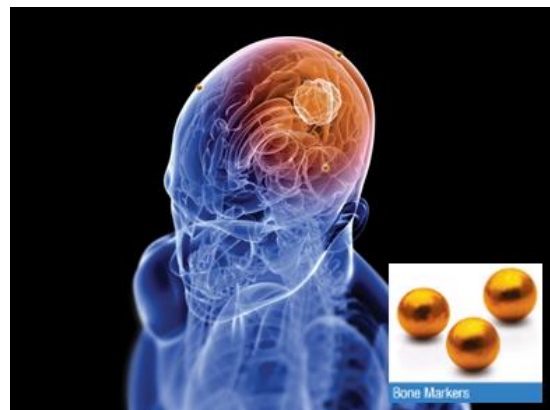
Μειονέκτημα των μεθόδων που βασίζονται σε εξωγενή χαρακτηριστικά είναι ότι χρειάζεται κάποια προετοιμασία και αλληλεπίδραση με τον χρήστη σαν πρώτο στάδιο καθώς και ο επεμβατικός τρόπος τοποθέτησης των αντικειμένων.

Ωστόσο, σε αυτή την κατηγορία υπάρχουν άλλες τεχνικές οι οποίες στηρίζονται στη μη επεμβατική τοποθέτηση των αντικειμένων ή σημάδιων. Τέτοια αντικείμενα μπορεί να είναι σημάδια τα οποία προσκολλούνται πάνω στο δέρμα [13], πλαίσια συγκράτησης κεφαλιού, εξατομικευμένα πλαίσια αφρού ή ακόμα και οδοντιατρικοί προσαρμογείς [14]. Γενικά είναι λιγότερο ακριβείς.

Ένας σημαντικός παράγοντας στην επιλογή του υλικού ενός fiducial marker είναι η συμβατότητα του με την απεικονιστική διάταξη. Π.χ. τα μεταλλικά αντικείμενα δεν είναι κατάλληλα στην απεικόνιση μαγνητικής τομογραφίας λόγω των υψηλής έντασης μαγνητικών πεδίων που εφαρμόζονται.



(α)



(β)

Εικόνα 2.2. (α) Στερεοστατικό πλαίσιο. (β) Επεμβατική τοποθέτηση των fiducial markers στο κεφάλι του ασθενή (χρυσές σφαίρες).

2.3.3.2 Ευθυγράμμιση βασισμένη σε ενδογενή χαρακτηριστικά

Οι μέθοδοι ευθυγράμμισης με χρήση ενδογενών χαρακτηριστικών βασίζονται μόνο στη πληροφορία που περιέχεται στις εικόνες και όχι στην εισαγωγή ή χρήση κάποιου τεχνητού αντικειμένου κατά την διαδικασία λήψης της εικόνας. Με βάση τα χαρακτηριστικά που

χρησιμοποιούνται για την ταύτιση, αυτές οι μέθοδοι ευθυγράμμισης χωρίζονται σε τρεις κατηγορίες [15]:

- Μέθοδοι με βάση τα σημεία
- Μέθοδοι τμηματοποίησης
- Μέθοδοι με βάση τη φωτεινότητα

(α) Μέθοδοι με βάση τα σημεία - Landmark based methods

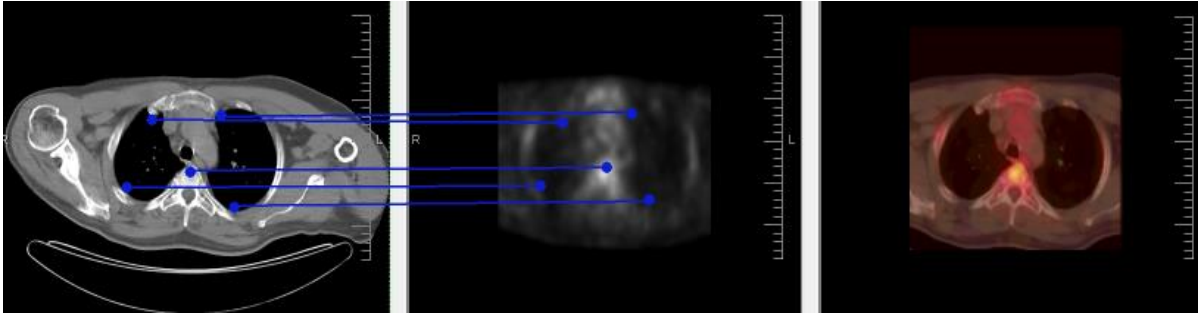
Οι μέθοδοι με βάση τα σημεία βασίζονται σε διακριτικά σημεία ή ορόσημα ως χαρακτηριστικά τα οποία ταυτίζονται στις εικόνες. Τα ορόσημα είναι σημεία εντός των εικόνων τα οποία μπορούν να εντοπιστούν από τον χρήστη συνήθως διαδραστικά και τα οποία αντιστοιχούν σε εμφανή και με ακρίβεια εντοπιζόμενα σημεία της μορφολογίας της ορατής ανατομίας [16]. Ορόσημα αποτελούν επίσης γεωμετρικά χαρακτηριστικά σημεία όπως τοπικά ακρότατα, γωνίες κλπ και γενικά εντοπίζονται με αυτόματο τρόπο.

Τέτοιες μέθοδοι περιορίζονται στον προσδιορισμό ενός συμπαγούς μετασχηματισμού λόγω του μικρού αριθμού των σημείων που μπορούν να χρησιμοποιηθούν ως σταθερά ανατομικά διακριτικά σημεία. Αν ο αριθμός των σημείων ήταν αρκετά μεγάλος θα μπορούσε να υλοποιηθεί κάποιος πιο πολύπλοκος μετασχηματισμός. Εφόσον ο αριθμός των διαθέσιμων σημείων σε αυτή την διαδικασία είναι περιορισμένος (μικρότερος σε σύγκριση με την αρχική εικόνα), οι μέθοδοι βελτιστοποίησης είναι αποτελεσματικές και ταχύτερες και συνήθως ελαχιστοποιούν απλά μεγέθη όπως η μέση απόσταση μεταξύ αντίστοιχων ορόσημων.

Τα ανατομικά ορόσημα συχνά χρησιμοποιούνται και σε συνδυασμό με άλλες μεθόδους, για παράδειγμα σε αλγόριθμους ευθυγράμμισης που είναι επιρρεπείς σε τοπικά ακρότατα (μέγιστα ή ελάχιστα) κατά την διάρκεια της βελτιστοποίησης [17]. Περιορίζοντας τον χώρο αναζήτησης, σύμφωνα με τις θέσεις των ανατομικών σημείων είναι δυνατή η αποφυγή αστοχιών, τοπικών ασυνεχών καθώς και η επιτάχυνση της διαδικασίας αναζήτησης.

Αρκετές άλλες μέθοδοι πραγματοποιούν ευθυγράμμιση με βάση τα ορόσημα σημεία απλώς δοκιμάζοντας έναν αριθμό από υποθετικούς μετασχηματισμούς που μπορεί να υπολογιστεί, για παράδειγμα από τρία τυχαία ζεύγη αντίστοιχων σημείων μεταξύ συνόλων σημείων. Οι συνήθεις μέθοδοι βελτιστοποίησης σε αυτή την περίπτωση έχουν πολύ μεγάλο υπολογιστικό κόστος.

Γενικά, οι μέθοδοι με βάση τα σημεία, σε θεωρητικό επίπεδο μπορούν να εφαρμοστούν σε κάθε εικόνα ανεξάρτητα από το αντικείμενο προς απεικόνιση ή το πρόβλημα της ευθυγράμμισης.



Εικόνα 2.3. Ευθυγράμμιση βασισμένη σε σημεία (Landmark based Registration). Αριστερά εικόνα από CT, μεσαία εικόνα από PET, δεξιά ευθυγραμμισμένες εικόνες

(β) Μέθοδοι τμηματοποίησης - Segmentation based methods

Οι μέθοδοι τμηματοποίησης βασίζονται σε αντίστοιχα τμήματα μεταξύ εικόνων τα οποία εξάγονται μέσω μιας διαδικασίας τμηματοποίησης, η οποία αποτελεί το πρώτο στάδιο της ευθυγράμμισης. Η τμηματοποίηση είναι συνήθως μια αυτόματη διαδικασία, αλλά υπάρχουν και περιπτώσεις στις οποίες ο χρήστης καθοδηγεί το βήμα αυτό. Τα εξαγόμενα χαρακτηριστικά μπορούν να είναι σημεία, καμπύλες ή επιφάνειες.

Οι μέθοδοι στις οποίες τα εξαγόμενα χαρακτηριστικά είναι καμπύλες ή επιφάνειες χωρίζονται σε δυο κατηγορίες:

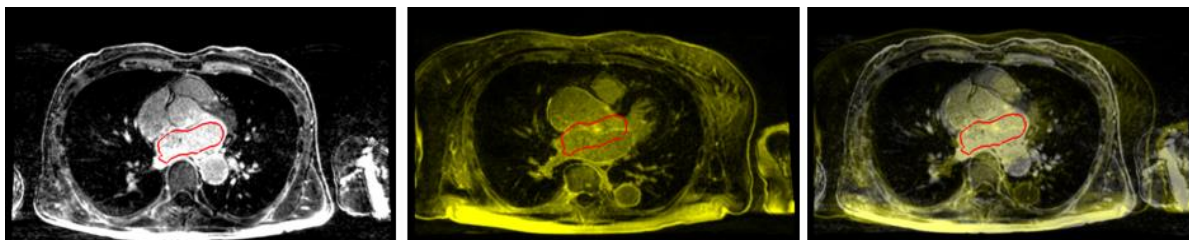
Στην πρώτη κατηγορία ανήκουν οι μέθοδοι που βασίζονται σε ένα άκαμπτο μοντέλο [18] όπου ανατομικά ίδιες δομές (κυρίως επιφάνειες) εξάγονται από τις δυο εικόνες υπό ευθυγράμμιση. Στην δεύτερη κατηγορία ανήκουν οι μέθοδοι που βασίζονται σε ένα εύκαμπτο μοντέλο, [19] όπου μια εξαγόμενη δομή από την μια εικόνα παραμορφώνεται ελαστικά έτσι ώστε να ταιριάζει στην δεύτερη εικόνα. Η διαδικασία της παραμόρφωσης γίνεται πάντα επαναληπτικά και με μικρές παραμορφώσεις κάθε φορά.

Οι μέθοδοι βασισμένοι σε άκαμπτο μοντέλο είναι και οι πιο διαδεδομένες στην κλινική πρακτική, ειδικά όταν τα εξαγόμενα τμήματα είναι επιφάνειες. Η ευρεία χρήση τους οφείλεται στη σχετικά εύκολη εκτέλεση του βήματος της κατάτμησης εικόνας, στο χαμηλό υπολογιστικό κόστος καθώς και στην επιτυχία της μεθόδου “Head-Hat” και της γρήγορης μεθόδου ταύτισης “Chamfer” [20] - [21].

Σε αντίθεση με την ευθυγράμμιση που βασίζεται σε άκαμπτα μοντέλα, που κυρίως χρησιμοποιείται για ευθυγράμμιση εικόνων της ίδιας πηγής (intrasubject), τα εύκαμπτα μοντέλα είναι θεωρητικώς κατάλληλα για ευθυγράμμιση εικόνων διαφορετικών πηγών (intersubject) ή για ευθυγράμμιση με άτλαντα, όπου η μια εικόνα είναι κατασκευασμένη από μια βάση συλλογής απεικονιστικών δεδομένων

Τα εύκαμπτα μοντέλα είναι κατάλληλα για τον προσδιορισμό καμπυλόγραμμων μετασχηματισμών μεταξύ εικόνων αλλά όχι για τον προσδιορισμό ολικών, συμπαγών, ή τύπου affine μετασχηματισμών.

Μειονέκτημα των μεθόδων τμηματοποίησης είναι ότι η ακρίβεια της ευθυγράμμισης περιορίζεται στην ακρίβεια του σταδίου τμηματοποίησης. Παρ’όλα αυτά μπορούν να χρησιμοποιηθούν σχεδόν για όλες τις ανατομικές περιοχές και απεικονιστικές τεχνικές και είναι συνήθως αυτοματοποιημένες.



Εικόνα 2.4. Ευθυγράμμιση με βάση την κατάτμηση (Segmentation based Registration). Εικόνες LGE-MR που ελήφθησαν σε διαφορετικές χρονικές στιγμές (αριστερά και μεσαία). Δεξιά, ευθυγραμμισμένες εικόνες

(γ) Μέθοδοι με βάση την φωτεινότητα - *Voxel Property-based, Intensity based methods*

Οι μέθοδοι αυτής της κατηγορίας διαφέρουν σημαντικά από τις μεθόδους των δυο παραπάνω κατηγοριών διότι βασίζονται άμεσα στις τιμές της φωτεινότητας των εικονοστοιχείων της εικόνας, χωρίς να πραγματοποιηθεί προηγούμενη μείωση πληροφορίας, ποσοστού δεδομένων ή τμηματοποίηση της εικόνας.

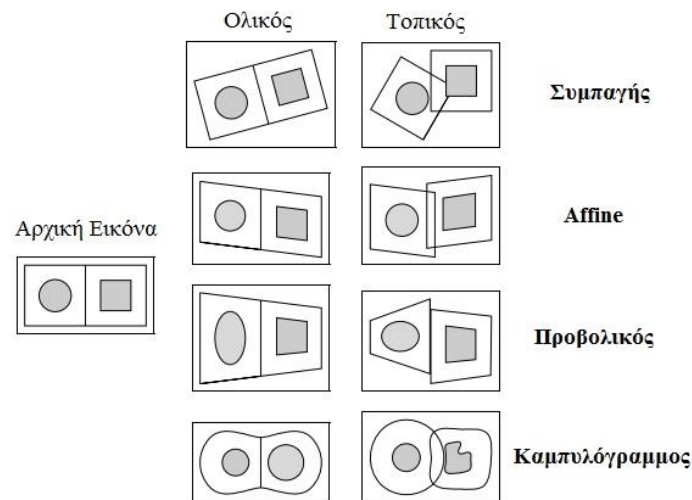
Υπάρχουν δυο διαφορετικές προσεγγίσεις: Στην πρώτη τα επίπεδα του γκρι της εικόνας μειώνονται ή περιορίζονται σε ένα σύνολο διανυσμάτων. Δυο μέθοδοι που χρησιμοποιούν την πρώτη προσέγγιση είναι η μέθοδος των κυρίων αξόνων και η μέθοδος των ροπών. Η ευθυγράμμιση των εικόνων πραγματοποιείται με την ευθυγράμμιση του κέντρου βάρους των εικόνων και των κυρίων αξόνων τα οποία υπολογίζονται από τις ροπές μηδενικής και πρώτης τάξης για την κάθε εικόνα [22]. Γενικά τέτοιες μέθοδοι δεν είναι πολύ ακριβείς αλλά είναι αυτόματες, γρήγορες και εύκολα υλοποιήσιμες και επομένως χρησιμοποιούνται ευρέως σε προβλήματα ευθυγράμμισης όπου η μεγάλη ακρίβεια δεν είναι σημαντικός παράγοντας ή σαν ένα στάδιο προεθυγράμμισης σε άλλες μεθόδους [23]. Σε πολλές εφαρμογές είναι απαραίτητη η τμηματοποίηση της εικόνας για την επίτευξη αποδεκτών αποτελεσμάτων.

Οι μέθοδοι της δεύτερης προσέγγισης, σε αντίθεση με τις άλλες μεθόδους, δεν χρειάζονται κάποια προεπεξεργασία ή αφαίρεση της πληροφορίας αλλά χρησιμοποιούν όλη την διαθέσιμη πληροφορία που περιέχεται στις εικόνες. Αν και η χρήση τους παλιά σε εφαρμογές 3D/3D ήταν περιορισμένη λόγω του μεγάλου υπολογιστικού κόστους οι απαιτήσεις διάφορων ιατρικών εφαρμογών για ευθυγράμμιση ιατρικών δεδομένων με μεγάλη ακρίβεια και η ανάπτυξη γρηγορότερων υπολογιστικών συστημάτων επέτρεψαν τις μεθόδους πλήρους περιεχομένου της εικόνας να χρησιμοποιούνται στην κλινική πρακτική. Σε σύγκριση με τις υπόλοιπες μεθόδους ευθυγράμμισης, οι μέθοδοι της δεύτερης κατηγορίας είναι οι πιο ακριβείς αλλά ταυτόχρονα και οι πιο χρονοβόρες.

2.3.4 Περιοχή εφαρμογής του μετασχηματισμού

Αν ο μετασχηματισμός εφαρμόζεται στο σύνολο της εικόνας καλείται ολικός ενώ αν μετασχηματίζει μόνο κάποια περιοχή της εικόνας καλείται τοπικός. Οι τοπικοί μετασχηματισμοί, ειδικά οι συμπαγείς, οι τύπου affine και οι προβολικοί μετασχηματισμοί, σπανίως εφαρμόζονται απευθείας σε μια εικόνα διότι στα όρια των περιοχών εφαρμογής τους δημιουργούνται ασυνέχειες οι οποίες εξαρτώνται από τον βαθμό ελαστικότητας του

μετασχηματισμού. Π.χ. η μετασχηματισμένη εικόνα μπορεί να περιέχει κενά ή σχισμές. Οι τοπικοί μετασχηματισμοί συνήθως είναι μη αντιστρέψιμοι που σημαίνει ότι με την εφαρμογή του αντίστροφου μετασχηματισμού στην μετασχηματισμένη εικόνα η αρχική εικόνα είναι αδύνατο να αναπαραχθεί.



Εικόνα 2.5. Παραδείγματα εφαρμογής δισδιάστατων ολικών και τοπικών μετασχηματισμών στην αρχική εικόνα

2.3.5 Είδος των δεδομένων

Οι τεχνικές ευθυγράμμισης μπορούν να κατηγοριοποιηθούν ανάλογα με το είδος των δεδομένων που επεξεργάζονται. Οι κατηγορίες είναι τρεις, ανάλογα με το αν τα δεδομένα προέρχονται από την:

- ίδια πηγή (mono-modality)
- διαφορετικές πηγές (multimodality)
- το ένα σύνολο δεδομένων είναι ένας ιατρικός άτλαντας (modality to atlas)

Υπάρχουν τεχνικές οι οποίες μπορούν να ευθυγραμμίσουν αποκλειστικά δεδομένα του ίδιου τύπου ή της ίδιας πηγής (CT-CT ή MRI-MRI). Αυτού του είδους ευθυγράμμιση, μπορεί να χρησιμοποιηθεί για την παρακολούθηση της εξέλιξης συγκεκριμένων ασθενειών, για την αξιολόγηση αποτελεσματικότητας μιας θεραπείας και σε τεχνικές ψηφιακής αφαιρετικής ακτινογραφίας. Άλλες τεχνικές ευθυγράμμισης μπορούν να επεξεργαστούν δεδομένα που έχουν ληφθεί χρησιμοποιώντας διαφορετικές μεθόδους λήψης (CT-MRI ή MRI-PET). Η συνηθέστερη περίπτωση στην οποία χρησιμοποιούνται είναι η σύνθεση ή σύντηξη ιατρικών δεδομένων. Οι τεχνικές ευθυγράμμισης της πρώτης κατηγορίας ονομάζονται μονοτροπικές (monomodality), ενώ αυτές που ανήκουν στην δεύτερη κατηγορία ονομάζονται πολυτροπικές (multimodality). Υπάρχουν επίσης τεχνικές ευθυγράμμισης στις οποίες το ένα σύνολο δεδομένων προς ευθυγράμμιση είναι ένας ιατρικός άτλαντας, δηλαδή ένα σύνολο

δεδομένων σχηματισμένο από μια βάση συλλογής απεικονιστικών δεδομένων από διαφορετικές πηγές.

2.3.6 Το υποκείμενο της ευθυγράμμισης

Το υποκείμενο της διαδικασίας ευθυγράμμισης σχετίζεται με την προέλευση των υπό ευθυγράμμιση δεδομένων. Τα δεδομένα εισόδου μιας τεχνικής ευθυγράμμισης μπορούν να προέρχονται από τον ίδιο ασθενή (intrasubject) ή από διαφορετικούς ασθενείς (intersubject). Υπάρχει επίσης μια περίπτωση στην οποία το ένα σύνολο δεδομένων προέρχεται από τον ασθενή και το δεύτερο σύνολο δεδομένων είναι ένας ιατρικός άτλαντας (subject to atlas).

2.3.7 Το αντικείμενο της ευθυγράμμισης

Το αντικείμενο της ευθυγράμμισης αποτελεί η συγκεκριμένη περιοχή της ανατομίας του ασθενούς η οποία πρέπει να ευθυγραμμιστεί. Υπάρχουν αρκετές τεχνικές ευθυγράμμισης οι οποίες στοχεύουν σε συγκεκριμένες ανατομικές περιοχές ή έχουν βελτιστοποιηθεί για συγκεκριμένες ανατομικές περιοχές.

2.3.8 Αυτοματοποίηση και διαδραστικότητα

Ανάλογα με τον βαθμό αλληλεπίδρασης με τον χρήστη, οι τεχνικές ευθυγράμμισης χωρίζονται σε τρεις κατηγορίες: χειροκίνητες, ημιαυτόματες και αυτόματες.

Οι μέθοδοι στις οποίες η ευθυγράμμιση πραγματοποιείται από τον χρήστη χαρακτηρίζονται ως χειροκίνητες (manual). Σε αυτή την περίπτωση ο χρήστης βασίζεται σε κατάλληλο λογισμικό το οποίο εμφανίζει την τρέχουσα κατάσταση της ευθυγράμμισης. Στις ημιαυτόματες μεθόδους ευθυγράμμισης (semi-automatic), ο χρήστης έχει μικρότερο βαθμό αλληλεπίδρασης και μπορεί να αρχικοποιήσει ή και να καθοδηγήσει τον αλγόριθμο ευθυγράμμισης, π.χ. με την απόρριψη ή την αποδοχή προτεινόμενων λύσεων. Στις αυτόματες μεθόδους ευθυγράμμισης ο χρήστης εισάγει στους αλγόριθμους μόνο τα δεδομένα ή και πληροφορίες σχετικές με την ανάκτηση των δεδομένων.

2.3.9 Μέθοδοι βελτιστοποίησης και τεχνικές αναζήτησης

Η μέθοδος βελτιστοποίησης καθορίζει την διαδικασία παραγωγής νέων τιμών για τις παραμέτρους του μετασχηματισμού βάσει της τιμής της συνάρτησης ομοιότητας ή συνάρτησης σφάλματος. Τελικός στόχος της βελτιστοποίησης είναι ο υπολογισμός τιμών για τις παραμέτρους του μετασχηματισμού οι οποίες, εφαρμοζόμενες στα υπό ευθυγράμμιση δεδομένα, παράγουν την καλύτερη δυνατή ακρίβεια της ευθυγράμμισης.

Σε αυτή την ενότητα παρουσιάζονται οι διαδικασίες βελτιστοποίησης οι οποίες χρησιμοποιούνται από τους αλγόριθμους ευθυγράμμισης ώστε να υπολογιστούν οι ζητούμενες παράμετροι μετασχηματισμού. Ο προσδιορισμός των παραμέτρων του μετασχηματισμού είναι το πιο βασικό βήμα στην διαδικασία της ευθυγράμμισης καθώς επηρεάζει άμεσα την ακρίβεια της διαδικασίας. Ο ολικός συμπαγής μετασχηματισμός απαιτεί μόνο έξι παραμέτρους για τον προσδιορισμό του μετασχηματισμού ενώ ένας ελαστικός μετασχηματισμός για παράδειγμα μπορεί να χρειαστεί εκατοντάδες ή ακόμα και

χιλιάδες παραμέτρους. Οι μέθοδοι βελτιστοποίησης μπορούν να χωριστούν σε δυο κατηγορίες:

- Απευθείας μέθοδοι
- Προσεγγιστικές μέθοδοι ή μέθοδοι αναζήτησης

Στις απευθείας μεθόδους οι παράμετροι μπορούν να υπολογιστούν άμεσα από τα διαθέσιμα δεδομένα και εξαρτώνται μόνο από την φύση του μετασχηματισμού.

Στις μεθόδους της δεύτερης κατηγορίας οι παράμετροι υπολογίζονται με βάση την αναζήτηση των σημείων ακρότατου της συνάρτησης επιτυχίας της ευθυγράμμισης. Όταν η συνάρτηση επιτυχίας δεν έχει ομαλή μαθηματική συμπεριφορά οι παράμετροι του μετασχηματισμού υπολογίζονται με εξαντλητική αναζήτηση σε όλο το χώρο των παραμέτρων. Μια τέτοια μέθοδος έχει πολύ μεγάλο υπολογιστικό κόστος και είναι χρονοβόρα διαδικασία καθώς εξετάζονται όλες οι τιμές της κάθε παραμέτρου της ευθυγράμμισης. Για την αποφυγή μιας τέτοιας τεχνικής συνήθως χρησιμοποιείται κάποια μέθοδος βελτιστοποίησης η οποία αναζητεί τα ακρότατα της συνάρτησης επιτυχίας. Ο αλγόριθμος βελτιστοποίησης μπορεί να εγκλωβιστεί σε τοπικό ακρότατο αντί ολικού. Για την αντιμετώπιση τέτοιου προβλήματος οι μέθοδοι ολικής βελτιστοποίησης προσεγγίζουν την βέλτιστη λύση με δυο τεχνικές. Στην πρώτη τεχνική, ξεκινώντας από τυχαία σημεία του χώρου των ανεξάρτητων μεταβλητών υπολογίζεται ένα πλήθος τοπικών ακρότατων και επιλέγεται το ακρότατο με την χαμηλότερη ή την υψηλότερη τιμή ανάλογα με το αν η συνάρτηση επιτυχίας είναι συνάρτηση μέτρησης σφάλματος ή συνάρτηση ομοιότητας. Στην δεύτερη τεχνική, επιλέγεται ένα τοπικό ακρότατο, η θέση του οποίου μετατοπίζεται με βήμα πεπερασμένου πλάτους και σε κάθε βήμα ελέγχεται αν η τιμή της συνάρτησης επιτυχίας είναι καλύτερη [24].

Μειονέκτημα των μεθόδων ολικής βελτιστοποίησης είναι ότι δεν συγκλίνουν με ακρίβεια στην περιοχή του ολικού ακρότατου. Συνήθως, για να υπολογιστεί η βέλτιστη λύση μετά την μέθοδο ολικής βελτιστοποίησης εφαρμόζεται μια μέθοδος τοπικής βελτιστοποίησης.

Παραδείγματα μεθόδων ολικής βελτιστοποίησης αναφέρονται παρακάτω:

- Μέθοδος προσομοίωσης ανάκτησης (simulated annealing) [25]
- Εξελικτικοί αλγόριθμοι (evolutionary algorithms)
- Γενετικοί αλγόριθμοι (genetic algorithms)
- Μέθοδος απαγορευμένης αναζήτησης
- Μέθοδος διακλάδωσης και φραγής (branch and bound).

Μερικά παραδείγματα μεθόδων τοπικής βελτιστοποίησης αναφέρονται παρακάτω:

- Μέθοδος Downhill-Simplex [26]
- Μέθοδος του Powell [27]
- Μέθοδος συζυγών κλίσεων (Conjugate Gradient)
- Μέθοδος Quasi-Newton [8].

2.4 Κριτήρια αξιολόγησης τεχνικών ευθυγράμμισης

Σε κάθε βήμα της επαναληπτικής διαδικασίας ευθυγράμμισης, η προς ευθυγράμμιση εικόνα μετασχηματίζεται χρησιμοποιώντας συγκεκριμένες τιμές των παραμέτρων του μετασχηματισμού. Έπειτα, η μετασχηματισμένη εικόνα συγκρίνεται με την εικόνα

αναφοράς βάσει προκαθορισμένων κριτηρίων. Συνήθως τα κριτήρια αυτά αποτελούνται από συναρτήσεις ομοιότητας ή σφάλματος, οι οποίες ποσοτικοποιούν το βαθμό ομοιότητας ή απόκλισης ανάμεσα στην ευθυγραμμισμένη εικόνα και την εικόνα αναφοράς.

Η προτιμώμενη συνάρτηση ομοιότητας ή σφάλματος εξαρτάται από την εφαρμογή της διαδικασίας ευθυγράμμισης και από το είδος των υπό επεξεργασία δεδομένων. Υπάρχει ένα μεγάλο πλήθος τέτοιων συναρτήσεων όπου οι περισσότερες προέρχονται από στατιστικές εφαρμογές.

2.4.1 Συνάρτηση επιτυχίας

Οι συναρτήσεις επιτυχίας μπορεί να είναι είτε συναρτήσεις ομοιότητας είτε συναρτήσεις μέτρησης απόστασης ανάλογα με το είδος της ευθυγράμμισης.

Όταν η ευθυγράμμιση βασίζεται στα ζεύγη αντίστοιχων σημείων, ως μέτρο επιτυχίας της ευθυγράμμισης χρησιμοποιείται μια συνάρτηση μέτρησης απόστασης όπου στόχος είναι η ελαχιστοποίηση της. Συνήθως υλοποιείται με την χρήση προσέγγισης ελαχίστων τετραγώνων όπου ελαχιστοποιείται το μέσο τετραγωνικό σφάλμα MSE:

$$MSE = \frac{1}{n} \sum_{i=1}^n \|r_i - T(r_i)\|^2 \quad (2.10)$$

Με $r_i = (x_i, y_i, z_i)$ και $r_i' = (x_i', y_i', z_i')$ συμβολίζονται τα αντίστοιχα ζεύγη σημείων στο προς ευθυγράμμιση σύνολο δεδομένων και στο σύνολο δεδομένων αναφοράς αντίστοιχα.

Όταν η ευθυγράμμιση βασίζεται στις τιμές χρωματικών πυκνοτήτων των στοιχείων της εικόνας, ως μέτρο επιτυχίας της ευθυγράμμισης χρησιμοποιείται μια συνάρτηση ομοιότητας η οποία εκφράζει ένα μέτρο συσχετισμού μεταξύ εικόνων. Ο στόχος μιας τέτοιας συνάρτησης είναι η μεγιστοποίηση της.

$$T = \arg \max_T [O(I_T(r), I_R(r))] \quad (2.11)$$

Οι πιο συνηθισμένες συναρτήσεις ομοιότητας που βρίσκουν εφαρμογή στην επίλυση δισδιάστατων και τρισδιάστατων προβλημάτων παρουσιάζονται παρακάτω.

2.4.1.1 Άθροισμα της απόλυτης διαφοράς χρωματικών πυκνοτήτων

Το άθροισμα της απόλυτης διαφοράς των χρωματικών πυκνοτήτων έχει την μορφή [28]:

$$C_{AD}(r) = \sum_r |I_R(r) - I_T(r)| \quad (2.12)$$

Η συνάρτηση αυτή δεν είναι κατάλληλη για δεδομένα που παρουσιάζουν τοπικές χρωματικές μεταβολές ή προέρχονται από διαφορετικές απεικονιστικές τεχνικές. Η συνάρτηση CAD είναι ευαίσθητη στον θόρυβο αν αυτό δεν ακολουθεί κανονική κατανομή. Μια παραλλαγή αυτής της συνάρτησης είναι το άθροισμα των διαφόρων τετραγώνων:

$$C_{SAD}(r) = \sum_r |I_R(r) - I_T(r)|^2 \quad (2.13)$$

2.4.1.2 Συντελεστής συσχέτισης

Ο συντελεστής συσχέτισης χρωματικών πυκνοτήτων περιγράφεται από την σχέση:

$$CC(I_R, I_T) = \frac{\left| \frac{\sum_r [I_T(r) - \bar{I}_T][I_R(r) - \bar{I}_R]}{\sqrt{\sum_r [I_T(r) - \bar{I}_T]^2} \sqrt{\sum_r [I_R(r) - \bar{I}_R]^2}} \right|}{\quad} \quad (2.14)$$

όπου \bar{I}_T, \bar{I}_R είναι η μέση τιμή των εικόνων $I_T(r)$ και $I_R(r)$ αντίστοιχα. Ο συντελεστής συσχέτισης παίρνει τιμές στο διάστημα $[0,1]$. Η συσχέτιση δυο εικόνων μεγιστοποιείται καθώς ο συντελεστής συσχέτισης πλησιάζει την μονάδα. Ο συντελεστής συσχέτισης παραμένει αναλλοίωτος σε γραμμικούς μετασχηματισμούς δηλαδή: $CC(a_1 I_T + b_1, a_2 I_R + b_2) = CC(I_T, I_R)$. Επομένως, δεν επηρεάζεται από μεταβολές της φωτεινότητας ή της αντίθεσης των εικόνων. Στη περίπτωση που τα δεδομένα εισόδου προέρχονται από διαφορετικές απεικονιστικές τεχνικές, η ευθυγράμμιση με βάση τον συντελεστή συσχέτισης αποτυγχάνει λόγω της διαφορετικής τιμής της χρωματικής πυκνότητας που αποδίδεται στις εικόνες από την κάθε τεχνική. Για να μπορεί να χρησιμοποιηθεί η συνάρτηση αυτή, θα πρέπει να πραγματοποιηθεί ένα στάδιο προεπεξεργασίας των δεδομένων.

2.4.1.3 Εντροπία διαφοράς

Εάν η τιμή των χρωματικών πυκνοτήτων I των στοιχείων μιας εικόνας θεωρηθεί τυχαία μεταβλητή με κατανομή $P(I=k)$, η εντροπία (H) της εικόνας της διαφοράς των χρωματικών πυκνοτήτων μεταξύ της εικόνας αναφοράς και της μετασχηματισμένης εικόνας, I_D , υπολογίζεται από την σχέση:

$$H(I_D) = -\sum_{k=0}^{G-1} P(I_D = k) \log_2 P(I_D = k) \quad (2.15)$$

Με G συμβολίζεται το πλήθος των επιπέδων του γκρι, που συνήθως είναι 256. Η εντροπία ανήκει στα κριτήρια της ευθυγράμμισης που βασίζονται στα ιστογράμματα των εικόνων.

2.4.1.4 Συντελεστής αμοιβαίας πληροφορίας

Ο συντελεστής αμοιβαίας πληροφορίας είναι ένας δείκτης του βαθμού εξάρτησης των δυο υπό σύγκριση εικόνων I_T και I_R και δίνεται από την σχέση:

$$MI(I_T, I_R) = \sum_{k=0}^{G-1} \sum_{l=0}^{G-1} P(I_T = k, I_R = l) \log_2 \frac{P(I_T = k, I_R = l)}{P(I_T = k) P(I_R = l)} \quad (2.16)$$

Ο συντελεστής αμοιβαίας πληροφορίας δείχνει το ποσό της πληροφορίας που η I_R περιέχει για την I_T .

Χαρακτηριστικό του συντελεστή αμοιβαιότητας είναι το ότι δεν επηρεάζεται από το θόρυβο και τις μεταβολές της φωτεινότητας ή αντίθεσης των εικόνων. Μπορεί να εφαρμοστεί σε δισδιάστατα ή τρισδιάστατα δεδομένα τα οποία μπορεί να προέρχονται από διαφορετικές απεικονιστικές τεχνικές. Η ευθυγράμμιση με βάση τον συντελεστή αμοιβαίας πληροφορίας είναι πλήρως αυτόματη. Συνήθως βρίσκει εφαρμογή σε προβλήματα στα οποία χρησιμοποιείται συμπαγής μετασχηματισμός αλλά μπορεί να εφαρμοστεί και σε προβλήματα όπου χρησιμοποιείται ένας μη συμπαγής μετασχηματισμός [8].

Συνοπτικά, οι σημαντικότερες συναρτήσεις επιτυχίας παρουσιάζονται στον παρακάτω πίνακα :

Πίνακας 2.1. Πίνακας συναρτήσεων επιτυχίας

Συνάρτηση	Μαθηματικός Τύπος
Μέσο τετραγωνικό σφάλμα	MSE $MSE = \frac{1}{n} \sum_{i=1}^n \ r_i - T(r_i)\ ^2$
Άθροισμα της απόλυτης διαφοράς χρωματικών πυκνοτήτων	$C_{AD}(r)$ $C_{AD}(r) = \sum_r I_R(r) - I_T(r) $
Άθροισμα των διαφορών τετραγώνων	$C_{SAD}(r)$ $C_{SAD}(r) = \sum_r I_R(r) - I_T(r) ^2$
Συντελεστής συσχέτισης	$CC(I_R, I_T)$ $CC(I_R, I_T) = \frac{\sum_r [I_T(r) - \bar{I}_T][I_R(r) - \bar{I}_R]}{\sqrt{\sum_r [I_T(r) - \bar{I}_T]^2} \sqrt{\sum_r [I_R(r) - \bar{I}_R]^2}}$
Εντροπία διαφοράς	$H(I_D)$ $H(I_D) = -\sum_{k=0}^{G-1} P(I_D = k) \log_2 P(I_D = k)$
Συντελεστής αμοιβαίας πληροφορίας	$MI(I_T, I_R)$ $MI(I_T, I_R) = \sum_{k=0}^{G-1} \sum_{l=0}^{G-1} P(I_T = k, I_R = l) \log_2 \frac{P(I_T = k, I_R = l)}{P(I_T = k)P(I_R = l)}$

2.4.2 Αποτίμηση τεχνικών ευθυγράμμισης

Προκειμένου να αξιολογηθεί η καταλληλότητα μιας μεθόδου ευθυγράμμισης θα πρέπει να μελετηθεί το κατά πόσο μια μέθοδος πληρεί ένα σύνολο κριτηρίων τα οποία αναλύονται παρακάτω [8]:

2.4.2.1 Ακρίβεια (precision and accuracy)

Η ακρίβεια (precision) αναφέρεται στο συστηματικό σφάλμα που λαμβάνεται όταν ο αλγόριθμος της ευθυγράμμισης εφαρμόζεται σε συγκεκριμένη είσοδο και είναι ιδιότητα του συστήματος. Το συστηματικό σφάλμα μπορεί να είναι σφάλμα ολοκλήρου του συστήματος ευθυγράμμισης ή των επιμέρους παραγόντων όπως σφάλματα της μεθόδου

βελτιστοποίησης. Η ακρίβεια (accuracy) αναφέρεται στο άμεσο σφάλμα. Το άμεσο σφάλμα μπορεί να είναι ποιοτικό ή ποσοτικό, είναι ανάλογο της περίπτωσης και δεν είναι ιδιότητα του συστήματος. Το ποιοτικό σφάλμα μπορεί να υπολογιστεί με οπτικό έλεγχο από τους ειδικούς. Ωστόσο, ο υπολογισμός του ποσοτικού σφάλματος βασίζεται σε μαθηματικές ή στατιστικές τεχνικές.

2.4.2.2 Ευρωστία – ευστάθεια

Ο όρος ευρωστία (robustness), αναφέρεται στην δυνατότητα μιας μεθόδου να παράγει ικανοποιητικά αποτελέσματα υπό την παρουσία θορύβου. Πολλές φορές μπορεί να χαρακτηριστεί και σαν ευστάθεια, που σημαίνει ότι μικρές διαταραχές στην είσοδο μπορούν να προκαλούν αντίστοιχα μικρές διακυμάνσεις στην έξοδο. Για παράδειγμα, στην περίπτωση ευθυγράμμισης εικόνων, μικρή παραμόρφωση ή μετακίνηση της προς ευθυγράμμιση εικόνας πρέπει να έχει ως αποτέλεσμα την σύγκλιση του αλγορίθμου σε σχεδόν ίδιο αποτέλεσμα.

2.4.2.3 Αξιοπιστία

Η αξιοπιστία είναι ένας όρος που σημαίνει το κατά πόσο ένας αλγόριθμος συμπεριφέρεται όπως αναμένεται όταν εφαρμόζεται σε ένα λογικό εύρος κλινικών δεδομένων εισόδου.

2.4.2.4 Απαιτήσεις και ανάγκες-αλγοριθμική πολυπλοκότητα

Αυτά τα δυο κριτήρια συνδέονται άμεσα μεταξύ τους. Όταν για παράδειγμα, ένας πολύπλοκος αλγόριθμος απαιτείται για ευθυγράμμιση δεδομένων με πολύπλοκες παραμορφώσεις, προφανώς οι απαιτήσεις και οι ανάγκες για την εκτέλεση του θα είναι μεγαλύτερες απ'ότι στην περίπτωση ευθυγράμμισης βάσει του συμπαγή μετασχηματισμού. Επίσης, αυτά τα κριτήρια σχετίζονται σε μεγάλο βαθμό με την κλινική εφαρμογή. Αν ένας αλγόριθμος υλοποιηθεί με σκοπό την χρήση του σε real time εφαρμογές θα πρέπει να είναι γρήγορος και αποδοτικός και να μην είναι απαιτητικός σε πόρους.

2.4.2.5 Κλινική χρησιμότητα

Μια μέθοδος ευθυγράμμισης θα πρέπει να εξεταστεί ως προς την κλινική χρησιμότητα και τη δυνατότητα υλοποίησης κλινικών εφαρμογών αλλά και ως προς τα πλεονεκτήματά της συγκριτικά με τις υπάρχουσες τεχνικές. Έτσι, κάθε φορά πρέπει να εξετάζεται το εξής ερώτημα: θα ήταν στη συγκεκριμένη περίπτωση κάποια άλλη μέθοδος αποτελεσματικότερη, ευκολότερη ή φθηνότερη;

Είναι αδύνατον μια μέθοδος ευθυγράμμισης να ικανοποιεί όλα τα παραπάνω κριτήρια. Ωστόσο, τα κριτήρια που πρέπει να πληρεί μια μέθοδος εξαρτώνται πάντα από την εφαρμογή και η επιλογή της είναι θέμα προσωπικής κρίσης.

3 Αισθητήρες βάθους

Οι 3D αισθητήρες ή οι αισθητήρες βάθους παρέχουν πληροφορία σχετικά με την απόσταση των σημείων ενός αντικειμένου από την κάμερα. Το σύστημα λήψης σχηματίζει μια εικόνα όπου το κάθε εικονοστοιχείο (pixel) έχει σαν τιμή την απόσταση του σημείου του αντικειμένου που αντιστοιχεί στο συγκεκριμένο pixel από τον αισθητήρα. Από την θέση του κάθε εικονοστοιχείου στην εικόνα βάθους και από την τιμή του μπορούν να υπολογιστούν μέσω συγκεκριμένων τεχνικών οι συντεταγμένες (x, y, z) στον χώρο ενός σημείου του αντικειμένου. Το σύνολο των σημείων αυτών ονομάζεται νέφος σημείων. Το νέφος σημείων είναι μια αναπαράσταση της εξωτερικής επιφάνειας πραγματικών αντικειμένων ως ένα σύνολο από σημεία στον τρισδιάστατο χώρο όπου το κάθε σημείο χαρακτηρίζεται από τις συντεταγμένες (x, y, z) και αναλογα με την περίπτωση μπορεί να έχει και άλλη πληροφορία όπως χρώμα και υφή. Ανάλογα με την τεχνική που χρησιμοποιείται για τον σχηματισμό των εικόνων βάθους υπάρχουν διάφοροι αισθητήρες που χωρίζονται στις παρακάτω κατηγορίες.

3.1 Στερεοσκοπικές κάμερες

Αποτελούνται από δυο ολίδιες κάμερες που βρίσκονται συνήθως σε απόσταση μερικών εκατοστών μεταξύ τους και λαμβάνουν ελαφρώς διαφορετικές εικόνες. Επεξεργάζοντας τις διαφορές μεταξύ των εικόνων είναι δυνατόν να υπολογιστούν οι αποστάσεις των σημείων του χώρου από την κάμερα. Τέτοιες κάμερες χρειάζονται άριστη βαθμονόμηση κάτι το οποίο είναι πρακτικά αδύνατο και επίσης καλό φωτισμένο περιβάλλον. Όταν τα αντικείμενα προς εξέταση έχουν απλή υφή ή η σκηνή δεν περιέχει πολλά αντικείμενα ο αλγόριθμος αδυνατεί να βρει σημεία ενδιαφέροντος και στην συνέχεια αντιστοιχίες σημείων μεταξύ των εικόνων δίνοντας έτσι λανθασμένα αποτελέσματα. Αυτές οι αιτίες κάνουν τις στερεοσκοπικές κάμερες τις λιγότερο ακριβείς σε όλες τις κατηγορίες [29].



Εικόνα 3.1. Παράδειγμα μιας στερεοσκοπικής κάμερας

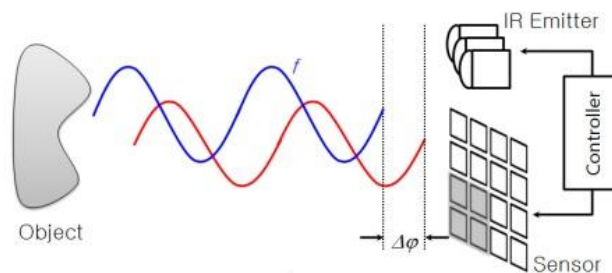
3.2 Time of flight

Η τεχνική αυτή βασίζεται στον χρόνο που χρειάζεται μια δέσμη φωτός ή ένας παλμός για να διανύσει μια απόσταση. Ένας πομπός εκπέμπει ένα κύμα το οποίο ανακλάται από το αντικείμενο. Το ανακλώμενο κύμα στην συνέχεια ανιχνεύεται από έναν αισθητήρα. Συνήθως, στην πράξη μετριέται η διαφορά φάσης μεταξύ των δυο κυμάτων και από αυτή έπειτα μπορεί να υπολογιστεί ο χρόνος τ σύμφωνα με την σχέση (3.1) [30].

$$\tau = \frac{\Delta\varphi}{2\pi f} \quad (3.1)$$

Δεδομένου ότι η ταχύτητα του φωτός, c , είναι μια γνωστή σταθερά η απόσταση του αντικειμένου από το σύστημα υπολογίζεται εύκολα ως εξής:

$$d = \frac{c}{2} \frac{\Delta\varphi}{2\pi f} \quad (3.2)$$



Εικόνα 3.2. Σχηματική αναπαράσταση του μέτρησης βάθους με την ToF τεχνική

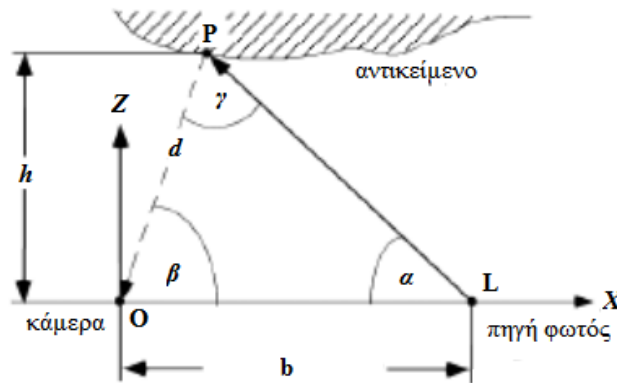
Τέτοιοι αισθητήρες δεν επηρεάζονται από τις συνθήκες φωτισμού και δίνουν πολύ ακριβής πληροφορία. Το Microsoft Kinect v2 είναι ένα παράδειγμα ενός αισθητήρα που βασίζεται στην τεχνική ToF.

3.3 Το δομημένο φως - Structured light

Οι αισθητήρες με βάση την τεχνική του δομημένου φωτός, χρησιμοποιούν αντί για δεύτερη κάμερα όπως στην περίπτωση των στερεοσκοπικών καμερών, έναν προβολέα (πηγή φωτός) ο οποίος προβάλλει πάνω στα αντικείμενα της σκηνής ένα πολύ καλά καθορισμένο σχέδιο υπερέυθρου φωτός (π.χ. ένα πλέγμα, κουκκίδες ή οριζόντιες γραμμές) [31]. Το σχέδιο αυτό όταν παρατηρείται από άλλη οπτική γωνία, διάφορη του προβολέα εμφανίζεται παραμορφωμένο. Αναλύοντας αυτή την παραμόρφωση μπορεί να αποκτηθεί η απόσταση των σημείων του αντικειμένου από την κάμερα επιστρέφοντας ένα σύνολο από σημεία με συντεταγμένες (x, y, z) ή αλλιώς ένα νέφος σημείων στον τρισδιάστατο χώρο. Το μαθηματικό μοντέλο υπολογισμού της απόστασης της συγκεκριμένης τεχνικής αναλύεται σε δυο και στις τρεις διαστάσεις όπως παρουσιάζεται παρακάτω:

3.3.1 Δισδιάστατη ανάλυση

Το σύστημα σε δυο διαστάσεις έχει την εξής διάταξη:



Εικόνα 3.3. Σχηματική αναπαράσταση του τρόπου μέτρησης απόστασης σε δυο διαστάσεις με την χρήση structured light αισθητήρων

Η κάμερα βρίσκεται στο σημείο O , η πηγή φωτός στο σημείο L και το σημείο παρατήρησης είναι το σημείο P . Η απόσταση μεταξύ πηγής και κάμερας είναι σταθερή και γνωστή, b . Επίσης είναι γνωστή και η γωνιά α . Η γωνιά β μπορεί να υπολογιστεί με χρήση της προβολικής γεωμετρίας. Αναλύοντας το τρίγωνο OLP , η απόσταση d του σημείου P από την κάμερα μπορεί να υπολογιστεί από τον νόμο των ημιτόνων ως εξής:

$$\frac{d}{\sin \alpha} = \frac{b}{\sin \gamma} \quad (3.3)$$

Από την $\gamma = \pi - (\alpha + \beta)$ και $\sin(\pi - \gamma) = \sin \gamma$ προκύπτει:

$$\frac{d}{\sin(\alpha)} = \frac{b}{\sin(\pi - \gamma)} = \frac{b}{\sin(\alpha + \beta)} \quad (3.4)$$

Επομένως η απόσταση d είναι:

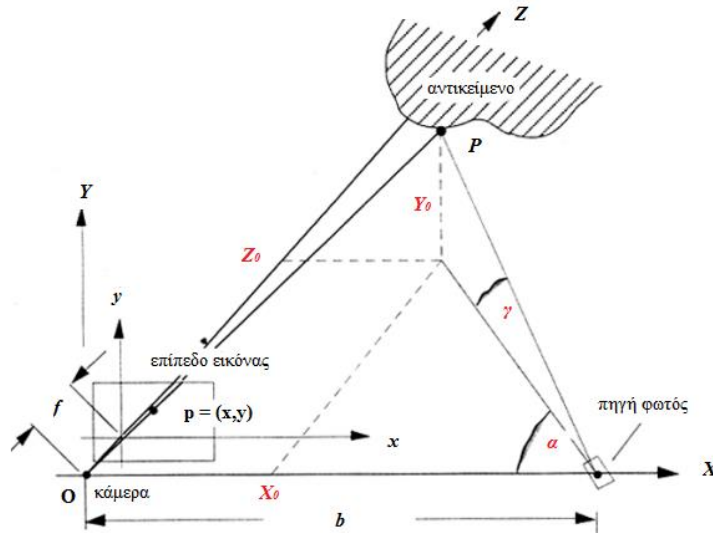
$$d = \frac{b \cdot \sin \alpha}{\sin(\alpha + \beta)} \quad (3.5)$$

Τελικά από την απόσταση d προκύπτει και η θέση του σημείου $P(X_0, Z_0)$ με συντεταγμένες

$$\begin{aligned} X_0 &= d \cdot \cos \beta \\ Z_0 &= h = d \cdot \sin \beta \end{aligned} \quad (3.6)$$

3.3.2 Τρισδιάστατη ανάλυση

Το σύστημα σε τρεις διαστάσεις έχει την εξής διάταξη:



Εικόνα 3.4. Σχηματική αναπαράσταση του τρόπου μέτρησης απόστασης σε τρεις διαστάσεις με την χρήση structured light αισθητήρων

Η κάμερα βρίσκεται στο σημείο O και έχει σύστημα συντεταγμένων XYZ . Το επίπεδο εικόνας όπως και προηγουμένως είναι το επίπεδο $Z=f$. Η απόσταση b και η γωνία α είναι γνωστά. Το σημείο του αντικείμενου $P(X_0, Y_0, Z_0)$ προβάλλεται στο σημείο $p(x, y)$ στο επίπεδο εικόνας. Ο άξονας της κάμερας είναι ο άξονας Z και το οπτικό κέντρο της πηγής φωτός βρίσκεται στον X άξονα. Από τα όμοια τρίγωνα στο επίπεδο εικόνας και στον τρισδιάστατο χώρο μπορεί να γραφτεί

$$\frac{X_0}{x} = \frac{Z_0}{f} = \frac{Y_0}{y} \quad (3.7)$$

Από τα ορθογώνια τρίγωνα

$$\tan \alpha = \frac{Z_0}{b - X_0} \quad (3.8)$$

Έτσι προκύπτει

$$Z_0 = \frac{X_0}{x} f = (b - X_0) \cdot \tan \alpha \quad (3.9)$$

$$X_0 \left(\frac{f}{x} + \tan \alpha \right) = b \cdot \tan \alpha \quad (3.10)$$

Τέλος, η θέση του σημείου $P(X_0, Y_0, Z_0)$ προσδιορίζεται από τις συντεταγμένες

$$X_0 = \frac{b \cdot x \cdot \tan \alpha}{f + x \cdot \tan \alpha}, \quad Y_0 = \frac{b \cdot y \cdot \tan \alpha}{f + x \cdot \tan \alpha}, \quad Z_0 = \frac{b \cdot f \cdot \tan \alpha}{f + x \cdot \tan \alpha} \quad (3.11)$$

Η γωνία γ όπως παρατηρείται δεν παίζει ρόλο στον υπολογισμό της θέσης του σημείου P επειδή η δέσμη φωτός θεωρήθηκε επίπεδη και κάθετη ως προς το επίπεδο XZ [32].

3.4 Microsoft Kinect

Το Kinect αρχικά σχεδιάστηκε ως ένα τηλεχειριστήριο για την φυσική αλληλεπίδραση σε ένα περιβάλλον ηλεκτρονικού παιχνιδιού. Ο Microsoft Kinect είναι ένας αισθητήρας βάθους ο οποίος βασίζεται στην τεχνική του δομημένου φωτός και λαμβάνει εικόνες βάθους και χρώματος. Η ενοποίηση της πληροφορίας βάθους και χρώματος έχει σαν αποτέλεσμα ένα νέφος σημείων όπου το κάθε σημείο περιέχει πληροφορία σχετικά με την θέση του στον χώρο δηλαδή συντεταγμένες x, y, z και χρώμα. Το οπτικό πεδίο του αισθητήρα στο οριζόντιο επίπεδο είναι 57 μοίρες και 43 μοίρες στο κατακόρυφο επίπεδο. Ο αισθητήρας μπορεί να καταγράφει ταυτόχρονα πληροφορία βάθους και χρώματος με ένα μέγιστο 30 πλαίσια το δευτερόλεπτο (frames per second). Το επίπεδο εικόνας έχει ανάλυση 640x480 που σημαίνει ότι ο μέγιστος αριθμός των σημείων για το οποίο μπορεί να υπολογιστεί η απόσταση τους από την κάμερα είναι 307200 [33].



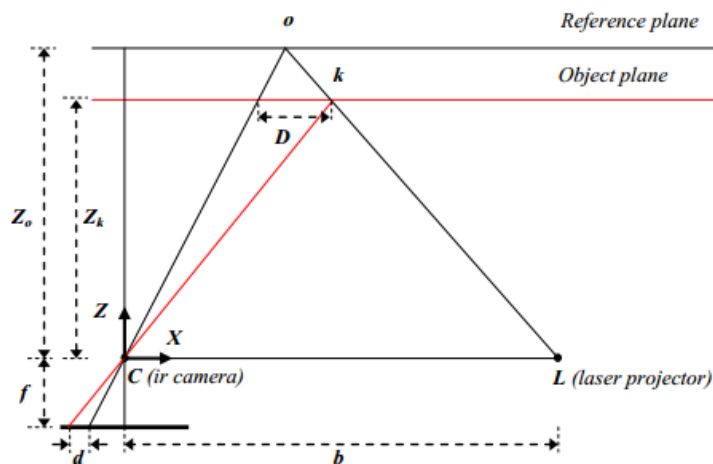
Εικόνα 3.5. Microsoft Kinect.

Ο αισθητήρας Kinect αποτελείται από έναν πομπό υπερύθρου φωτός, μια κάμερα ή αισθητήρα υπερύθρων και από μια κάμερα RGB για πρόσθεση πληροφορίας σχετικά με το χρώμα. Η πηγή laser εκπέμπει μια δέσμη φωτός η οποία χωρίζεται σε πολλαπλές δέσμες μέσω του φαινομένου της περίθλασης για να προβάλλει ένα προκαθορισμένο σχέδιο από κουκκίδες πάνω στο αντικείμενο. Αυτό το σχέδιο συλλαμβάνεται από την κάμερα υπερύθρων και συσχετίζεται έναντι ενός προτύπου αναφοράς. Το πρότυπο αναφοράς λαμβάνεται με την σύλληψη ενός επιπέδου σε μια γνωστή απόσταση από τον αισθητήρα και αποθηκεύεται στην μνήμη της συσκευής. Όταν μια κουκκίδα προβάλλεται σε ένα αντικείμενο η απόσταση του οποίου από τον αισθητήρα είναι μικρότερη ή μεγαλύτερη από την απόσταση του επιπέδου αναφοράς, η θέση της προβαλλόμενης κουκκίδας στο επίπεδο της εικόνας μετατοπίζεται ως προς τον X άξονα. Αυτές οι μετατοπίσεις υπολογίζονται εύκολα για κάθε κουκκίδα. Στη συνέχεια για κάθε εικονοστοιχείο (pixel) στο επίπεδο της

εικόνας, η απόσταση ενός σημείου του αντικειμένου k μπορεί να υπολογιστεί από την αντίστοιχη μετατόπιση στον X άξονα.

3.4.1 Μαθηματικό μοντέλο υπολογισμού της απόστασης

Στην εικόνα 2.11 παρουσιάζεται η σχέση μεταξύ της απόστασης ενός σημείου k του αντικειμένου από τον αισθητήρα, σε σχέση με έναν επίπεδο αναφοράς (reference plane) και της μετρούμενης μετατόπισης d . Για την αναπαράσταση ενός σημείου με συντεταγμένες (x, y, z) στον χώρο ορίζεται ένα σύστημα συντεταγμένων με X, Y, Z άξονες και αρχή των αξόνων το κέντρο της κάμερας υπερύθρου φωτός. Ο άξονας Z επιλέγεται να είναι κάθετος ως προς το επίπεδο της εικόνας και με κατεύθυνση ως προς το αντικείμενο. Ο άξονας X επιλεγεται να είναι παραλληλος με την ευθεία που ενώνει την κάμερα με τον πομπό του υπερύθρου φωτός και ο άξονας Y επομένως είναι κάθετος με τους X, Z άξονες.



Εικόνα 3.6. Αναπαράσταση μέτρησης απόστασης ενός σημείου k του αντικειμένου από την κάμερα με χρήση τριγωνομετρικών ιδιοτήτων

Όπως προαναφέρθηκε, η σχετική απόσταση ενός σημείου από το επίπεδο αναφοράς συνεπάγεται μια μετατόπιση ως προς τον X άξονα της προβαλλόμενης κουκκίδας στο επίπεδο της εικόνας. Αυτή η μετατόπιση μπορεί να υπολογιστεί με την βοήθεια όμοιων τριγώνων:

$$\frac{D}{b} = \frac{Z_o - Z_k}{Z_o} \quad (3.12)$$

$$\frac{d}{f} = \frac{D}{Z_k} \quad (3.13)$$

όπου Z_k είναι η απόσταση του σημείου k από τον αισθητήρα, b είναι η απόσταση του αισθητήρα από τον πομπό, f είναι η εστιακή απόσταση της κάμερας υπερύθρων, D είναι η μετατόπιση του σημείου k στον χώρο, d είναι η παρατηρούμενη μετατόπιση στο επίπεδο

της εικόνας και Z_o είναι η απόσταση στον χώρο του επιπέδου αναφοράς. Με αντικατάσταση της μεταβλητής D στην εξίσωση 3.12 προκύπτει η απόσταση στον χώρο του σημείου k από τον αισθητήρα ως εξής:

$$Z_k = \frac{Z_o}{1 + \frac{Z_o}{fb}d} \quad (3.14)$$

Η εξίσωση είναι το βασικό μαθηματικό μοντέλο από το οποίο προκύπτει η απόσταση ενός σημείου στον χώρο με βάση την παρατηρούμενη μετατόπιση στο επίπεδο της εικόνας, δεδομένου ότι οι σταθερές παράμετροι Z_o, f, b προσδιορίζονται με την βαθμονόμηση. Οι συντεταγμένες X_k, Y_k του σημείου k υπολογίζονται από τις x, y συντεταγμένες του σημείου στο επίπεδο της εικόνας:

$$X_k = -\frac{Z_k}{f}(x_k - x_o + \delta x) \quad (3.15)$$

$$Y_k = -\frac{Z_k}{f}(y_k - y_o + \delta y) \quad (3.16)$$

όπου x_k, y_k είναι οι συντεταγμένες του σημείου στο επίπεδο της εικόνας, x_o, y_o είναι οι συντεταγμένες του σημείου αναφοράς και $\delta x, \delta y$ είναι οι διορθώσεις για την παραμόρφωση του φακού. Τελικά μετά από αυτή την ανάλυση το σημείο k στον χώρο έχει συντεταγμένες (X_k, Y_k, Z_k) .

3.4.2 Βαθμονόμηση του αισθητήρα

Οι παράμετροι βαθμονόμησης που περιλαμβάνονται στο μαθηματικό μοντέλο για τον υπολογισμό των συντεταγμένων των σημείων από τις μετρήσεις στο επίπεδο της εικόνας είναι:

- Το εστιακό μήκος, f
- Οι συντεταγμένες του σημείου αναφοράς (x_o, y_o)
- Οι συντελεστές παραμόρφωσης των φακών της κάμερας υπερύθρων $(\delta x, \delta y)$
- Μήκος βάσης, η απόσταση μεταξύ του αισθητήρα και του πομπού, b
- Η απόσταση του επιπέδου αναφοράς, Z_o

Στις μετρήσεις θεωρήθηκε ότι ο άξονας X του συστήματος συντεταγμένων της εικόνας πρέπει να είναι παράλληλος με την ευθεία βάσης. Ωστόσο, τυχόν αποκλίσεις μεταξύ τους μπορούν να αποφευχθούν εάν το σύστημα συντεταγμένων της εικόνας και το σύστημα συντεταγμένων στον χώρο είναι παράλληλα. Από τις παραπάνω παραμέτρους οι πρώτες τρεις μπορούν να προσδιοριστούν με απλή βαθμονόμηση της κάμερας υπερύθρων. Ο προσδιορισμός της απόστασης βάσης, b και της απόστασης αναφοράς Z_o είναι μια λίγο πιο πολύπλοκη διαδικασία. Στην πράξη, οι παρατηρούμενες μετατοπίσεις στο επίπεδο της εικόνας κανονικοποιούνται και κβαντίζονται στο διάστημα [0...2047]. Επομένως η

μεταβλητής d στην εξίσωση (3.14) πρέπει να αντικατασταθεί με $md' + n$ όπου d' η κανονικοποιημένη παρατηρούμενη μετατόπιση και m, n είναι οι παράμετροι μιας γραμμικής κανονικοποίησης. Λαμβάνοντας αυτά υπόψη η εξίσωση (3.14) δίνει:

$$Z_k^{-1} = \left(\frac{m}{fb} \right) d' + \left(Z_o^{-1} + \frac{n}{fb} \right) \quad (3.17)$$

Η εξίσωση αυτή αναπαριστά μια γραμμική σχέση μεταξύ της αντίστροφης απόστασης ενός σημείου και της αντίστοιχης κανονικοποιημένης παρατηρούμενης μετατόπισής του. Θεωρώντας τις κανονικοποιημένες παρατηρούμενες μετατοπίσεις για ένα πλήθος σημείων του αντικειμένου με γνωστές αποστάσεις ως προς τον αισθητήρα, οι συντελεστές αυτής της γραμμικής σχέσης μπορούν να υπολογιστούν με προσέγγιση ελαχίστων τετραγώνων. Η συμπερίληψη των κανονικοποιημένων παραμέτρων ωστόσο δεν επιτρέπει τον υπολογισμό των b και Z_o χωριστά. Οι παραπάνω παράμετροι βαθμονόμησης προσδιορίζουν την σχέση μεταξύ μετρούμενων μεγεθών στο επίπεδο της εικόνας (x, y, d') και των συντεταγμένων των σημείων του αντικειμένου στο χώρο (X, Y, Z) . Ο υπολογισμός τους για κάθε σημείο έχει ως αποτέλεσμα την δημιουργία ενός νέφους σημείων.

3.4.3 Πρόσθεση του χρώματος στο νέφος σημείων

Έκτος από τις (x, y, z) συντεταγμένες, τα σημεία μπορούν να περιέχουν και άλλη πληροφορία όπως χρώμα. Ο αισθητήρας Kinect περιλαμβάνει επίσης μια κάμερα RGB μέσω της οποίας υπολογίζονται οι τιμές του κάθε χρώματος κόκκινο, πράσινο και μπλε. Για να μπορεί να πραγματοποιηθεί αυτό αρχικά οι δυο κάμερες θα πρέπει να βαθμονομηθούν και θα πρέπει να προσδιοριστούν οι παράμετροι που περιγράφουν την σχετική θέση τους στον χώρο.

3.4.4 Ακρίβεια

Πιθανά σφάλματα στα δεδομένα που λαμβάνονται από το Kinect προέρχονται από τρεις κύριες πηγές: από τον αισθητήρα, από τις συνθήκες στις οποίες λαμβάνονται τα δεδομένα και από τις ιδιότητες της επιφάνειας του αντικειμένου.

Τα σφάλματα που προέρχονται από τον αισθητήρα οφείλονται κυρίως στην μη σωστή βαθμονόμηση και σε λανθασμένες μετρήσεις της μετατόπισης στο επίπεδο της εικόνας. Αυτές οι αιτίες οδηγούν σε συστηματικό σφάλμα στις συντεταγμένες των σημείων του αντικειμένου. Τέτοιο συστηματικό σφάλμα μπορεί να αποφευχθεί με κατάλληλη βαθμονόμηση όπως περιγράφηκε στην ενότητα 3.4.2.

Τα σφάλματα που σχετίζονται με τις συνθήκες μέτρησης οφείλονται στις συνθήκες φωτισμού και στην γεωμετρία της απεικόνισης. Οι συνθήκες φωτισμού επηρεάζουν την συσχέτιση και την μέτρηση των μετατοπίσεων στο επίπεδο της εικόνας. Σε ισχυρό φωτισμό οι κουκκίδες λέιζερ εμφανίζονται με χαμηλή αντίθεση στην υπέρυθη εικόνα και επομένως οδηγούν σε λανθασμένα δεδομένα ή σε κενό στο προκύπτον νέφος σημείων. Οι λείες και γυαλιστερές επιφάνειες εμφανίζονται υπερεκτεθειμένες στην υπέρυθη εικόνα και

εμποδίζουν την μέτρηση των μετατοπίσεων στο επίπεδο της εικόνας με αποτέλεσμα το νέφος σημείων να παρουσιάσει επίσης κενά σε αυτή την περιοχή.

Η γεωμετρία της απεικόνισης περιλαμβάνει την απόσταση του αντικειμένου από τον αισθητήρα και τον προσανατολισμό της επιφάνειας του αντικειμένου. Το εύρος λειτουργίας του αισθητήρα είναι μεταξύ 0.5 και 5 μέτρα σύμφωνα με τις προδιαγραφές και το τυχαίο σφάλμα στις μετρήσεις αυξάνεται με αύξηση της απόστασης του αντικειμένου από την κάμερα. Οι πειραματικές μετρήσεις έχουν δείξει ότι το τυχαίο σφάλμα στην μέτρηση του βάθους είναι ανάλογο του τετραγώνου της απόστασης του αντικειμένου από την κάμερα. Σε απόσταση 0.5m από τον αισθητήρα το τυχαίο σφάλμα είναι της τάξης μερικών χιλιαστών, σε απόσταση 1m περίπου 0.25cm και περίπου 0.5cm σε απόσταση 1.5m από τον αισθητήρα. Το μέγιστο σφάλμα μπορεί να φτάσει την τιμή 4cm στην μέγιστη απόσταση των 5m.

Υποθέτοντας ότι στην αξίωση (3.17) οι παράμετροι βαθμονόμησης έχουν προσδιοριστεί με ακρίβεια και d' είναι μια τυχαία μεταβλητή που ακολουθεί κανονική κατανομή η διασπορά των μετρήσεων της απόστασης των σημείων του αντικειμένου μπορεί να προκύψει από την διασπορά των μετρήσεων της μετατόπισης στο επίπεδο της εικόνας:

$$\sigma_z^2 = \left(\frac{\partial Z}{\partial d} \right)^2 \sigma_{d'}^2 \quad (3.18)$$

Και με απλοποίηση προκύπτει η τυπική απόκλιση της απόστασης σημείων αντικειμένου

$$\sigma_z = \left(\frac{m}{fb} \right) Z^2 \sigma_{d'} \quad (3.19)$$

όπου $\sigma_{d'}$ και σ_z είναι η τυπική απόκλιση των κανονικοποιημένων μετατοπίσεων στο επίπεδο εικόνας και η τυπική απόκλιση των μετρούμενων αποστάσεων των σημείων του αντικειμένου από τον αισθητήρα. Η εξίσωση (3.19) δείχνει ότι το τυχαίο σφάλμα μέτρησης απόστασης είναι ανάλογο του τετραγώνου της απόστασης του αντικειμένου από την κάμερα.

3.4.5 Ανάλυση του βάθους και πυκνότητα νέφους σημείων

Η απόσταση των σημείων στον XY επίπεδο (κάθετο στον άξονα της κάμερας) καθορίζεται από την ανάλυση της κάμερας υπερύθρων ή ακριβέστερα από το μέγεθος του εικονοστοιχείου στο επίπεδο της εικόνας. Εφόσον κάθε εικόνα βάθους περιέχει ένα σταθερό αριθμό εικονοστοιχείων, 640x480 εικονοστοιχεία, η πυκνότητα σημείων μειώνεται καθώς η απόσταση του αντικειμένου από τον αισθητήρα αυξάνεται. Θεωρώντας την πυκνότητα σημείων ως τον αριθμό των σημείων ανά μονάδα επιφάνειας, καθώς ο αριθμός των σημείων παραμένει σταθερός, η επιφάνεια είναι ανάλογη του τετραγώνου της απόστασης των σημείων από τον αισθητήρα. Επομένως, η πυκνότητα σημείων στο επίπεδο XY είναι αντιστρόφως ανάλογη του τετραγώνου της απόστασης του αντικειμένου από την κάμερα.

Η ανάλυση ή ευκρίνεια βάθους αναφέρεται στην ελάχιστη διαφορά απόστασης που μπορεί να μετρηθεί και καθορίζεται από τον αριθμό των bits ανά εικονοστοιχείο που χρησιμοποιείται για την αποθήκευση των μετρούμενων μετατοπίσεων στο επίπεδο της εικόνας. Το Kinect χρησιμοποιεί 10 bits και επομένως 1024 επίπεδα για τις μετατοπίσεις και ένα bit για να σημειωθεί ένα εικονοστοιχείο για το οποίο δεν μετρήθηκε η μετατόπιση [33].

4 Μέθοδοι Ευθυγράμμισης με Γεωμετρικούς Περιγραφείς

4.1 Εισαγωγή

Συνήθως, οι τεχνικές ευθυγράμμισης υψηλής ακρίβειας λαμβάνουν υπ' όψιν τους όλο το σύνολο των σημείων στις υπό ευθυγράμμιση δομές δεδομένων. Τέτοιες τεχνικές ονομάζονται καθολικές τεχνικές ευθυγράμμισης και έχουν ως βάση είτε την σχετική θέση του σημείου στα δεδομένα είτε την τιμή του σημείου η οποία στις ψηφιακές εικόνες είναι η ένταση της φωτεινότητας του.

Παρά το γεγονός ότι οι καθολικές τεχνικές ευθυγράμμισης είναι ιδιαίτερα ακριβείς και αποδοτικές σε δισδιάστατα δεδομένα το αυξημένο μέγεθος των δεδομένων που εμπλέκονται στην τρισδιάστατη απεικόνιση καθιστά την χρήση καθολικών τεχνικών ευθυγράμμισης μη πρακτική από άποψη χρόνου και υπολογιστικού κόστους. Τα τρισδιάστατα απεικονιστικά δεδομένα αποτελούνται συνήθως από πλήθος επίπεδων εικόνων τομογραφίας, αυξάνοντας σημαντικά τον χρόνο επεξεργασίας που απαιτείται για δεδομένα αυτού του τύπου σε σχέση με τα δισδιάστατα δεδομένα.

Επίσης, όταν απαιτείται ευθυγράμμιση δεδομένων που προέρχονται από διαφορετικές απεικονιστικές τεχνικές οι μέθοδοι ευθυγράμμισης που βασίζονται στις τιμές της φωτεινότητας αποτυγχάνουν λόγω του ότι οι τιμές της φωτεινότητας σε κάθε σύνολο δεδομένων μπορούν να παρουσιάσουν τελείως διαφορετική πληροφορία. Ακόμα και όταν τα δεδομένα προς ευθυγράμμιση προέρχονται από το ίδιο απεικονιστικό σύστημα σε διαφορετικές χρονικές στιγμές το ένα σύνολο μπορεί να έχει ληφθεί σε συνθήκες ισχυρότερου φωτισμού και συνεπώς να διαφέρει αρκετά από το άλλο σύνολο δεδομένων. Σε αυτή την περίπτωση η ευθυγράμμιση με βάση τις τιμές φωτεινότητας είναι μια δύσκολη διαδικασία έως και αδύνατη.

Στις περιπτώσεις αυτές, θα μπορούσε να εξεταστεί η δυνατότητα της αξιοποίησης μόνο ενός μικρού υποσυνόλου των σημείων των δεδομένων που εξάγονται από αυτά τα σύνολα για την επιτυχή ευθυγράμμιση τους. Η ιδέα αυτή οδήγησε στην ανάπτυξη σημειακών τεχνικών τρισδιάστατης ευθυγράμμισης.

Για τα τρισδιάστατα δεδομένα τα εξαχθέντα σημεία μπορούν να μοντελοποιηθούν ως νέφη σημείων στον τρισδιάστατο χώρο. Ένα νέφος διακριτών σημείων εξάγεται από κάθε δομή δεδομένων. Επομένως η ευθυγράμμιση των δυο συνόλων τρισδιάστατων δεδομένων μπορεί να επιτευχθεί με την ευθυγράμμιση των νεφών σημείων τους. Στην συνέχεια αντιστοιχίες σημείων προσδιορίζονται με σκοπό τον υπολογισμό ενός γεωμετρικού μετασχηματισμού για την ταύτιση των δυο δομών δεδομένων ελαχιστοποιώντας την χωρική μεταξύ τους απόκλιση.

Σε αυτό το κεφάλαιο παρουσιάζονται μέθοδοι ευθυγράμμισης τρισδιάστατων νεφών σημείων βασισμένων σε γεωμετρικούς περιγραφείς. Με σκοπό την σύγκριση τους με μια λίγο διαφορετική στον τρόπο λειτουργίας τεχνική παρουσιάζεται επίσης συνοπτικά και μια μεθοδολογία βασισμένη στον αλγόριθμο ICP. Τα βήματα των εξετασθέντων μεθόδων ευθυγράμμισης βάσει γεωμετρικών περιγραφέων, σε κάθε περίπτωση είναι τα ίδια με την διαφορά ότι ένας διαφορετικός περιγραφέας χρησιμοποιείται κάθε φορά.

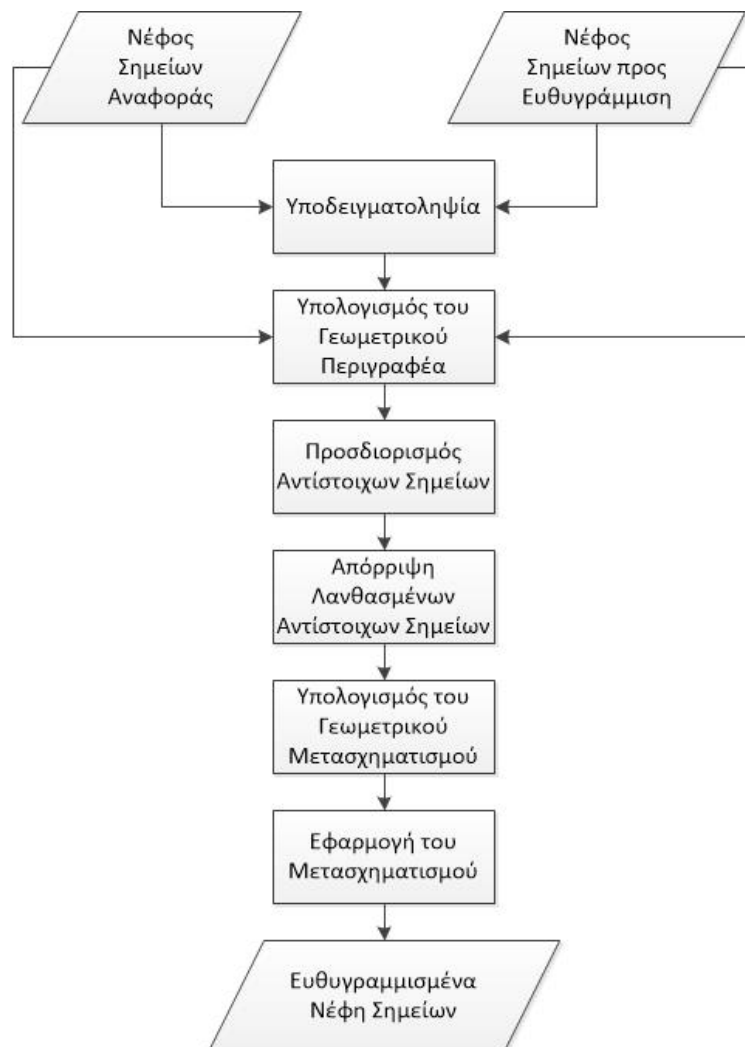
Για την μελέτη των τεχνικών αυτών τα δεδομένα που χρησιμοποιήθηκαν έχουν ληφθεί με την βοήθεια του αισθητήρα βάθους, Microsoft Kinect. Όπως περιγράφηκε και στο κεφάλαιο 3 αυτά τα νέφη σημείων είναι σύνολα από σημεία με (x,y,z) συντεταγμένες και αποτελούν μια αναπαράσταση των αντικειμένων στον τρισδιάστατο χώρο. Το πρώτο στάδιο αυτών των τεχνικών είναι ο προσδιορισμός των αντιστοιχιών σημείων μεταξύ των δυο συνόλων και στην συνέχεια ο προσδιορισμός ενός μετασχηματισμού ο οποίος ευθυγραμμίζει το νέφος σημείων προς ευθυγράμμιση με το νέφος σημείων αναφοράς. Η χρήση μόνο των συντεταγμένων (x,y,z) των σημείων δεν αρκεί για τον υπολογισμό των ζευγών αντιστοιχιών σημείων. Για να μπορούν να συγκριθούν, τα σημεία πρέπει να περιγραφούν με έναν σαφή και μοναδικό τρόπο. Έτσι, εισάγεται η έννοια του γεωμετρικού περιγραφέα. Ένας γεωμετρικός περιγραφέας κωδικοποιεί τις γεωμετρικές ιδιότητες μιας καθορισμένης περιοχής γύρω από ένα σημείο με σκοπό την μοναδική περιγραφή και αναγνώριση του στο σύνολο των σημείων. Οι γεωμετρικές αυτές ιδιότητες μπορεί να είναι γωνίες μεταξύ κάθετων διανυσμάτων, αποστάσεις ή μια σχετική θέση ως προς το εξεταζόμενο σημείο. Υπάρχει ένα μεγάλο πλήθος γεωμετρικών περιγραφέων οι οποίοι χρησιμοποιούν συγκεκριμένα χαρακτηριστικά της γειτονικής περιοχής. Αυτοί εξετάζονται με λεπτομέρεια στην ενότητα 4.3.2.

4.2 Λήψη δεδομένων

Τα δεδομένα που χρησιμοποιήθηκαν στις μεθόδους ευθυγράμμισης σε αυτή την εργασία είναι τρισδιάστατα σύνολα δεδομένων που ονομάζονται νέφη σημείων. Το νέφος σημείων είναι μια αναπαράσταση της επιφάνειας των πραγματικών αντικειμένων ως ένα σύνολο από σημεία στον τρισδιάστατο χώρο όπου το κάθε σημείο χαρακτηρίζεται από τις συντεταγμένες (x, y, z) και ανάλογα με την περίπτωση μπορεί να έχει και άλλη πληροφορία όπως χρώμα και υφή. Όλα τα ζεύγη δεδομένων έχουν ληφθεί με χρήση της κάμερας Microsoft Kinect υπό όχι αυστηρώς ελεγχόμενες συνθήκες φωτισμού. Αρχικά συλλέγεται το πρώτο νέφος σημείων το οποίο ονομάζεται νέφος σημείων αναφοράς. Με μικρή μετακίνηση της κάμερας από την αρχική θέση αποκτάται και το δεύτερο νέφος που αποτελεί το νέφος σημείων προς ευθυγράμμιση. Η μετακίνηση της κάμερας και πιθανές αλλαγές στις συνθήκες διάθλασης του φωτός έχουν σαν αποτέλεσμα τα δυο νέφη σημείων του ίδιου αντικειμένου να διαφέρουν χωρικά κατά κάποιο βαθμό. Το πλήθος των σημείων τους κυμαίνεται ανάλογα με την περίπτωση και μπορεί να φτάσει έως 307200 σημεία ανά νέφος. Ακόμα και για το ίδιο αντικείμενο δυο διαφορετικά νέφη σημείων μπορούν να περιέχουν διαφορετικό αριθμό σημείων. Στα συγκεκριμένα σετ τα σημεία εκτός από τις συντεταγμένες (x, y, z) περιείχαν επίσης πληροφορία σχετικά με το χρώμα η οποία απορρίφθηκε αφού δεν χρησιμοποιείται στις μεθόδους που παρουσιάζονται.

4.3 Μεθοδολογία ευθυγράμμισης βασισμένης στα γεωμετρικά χαρακτηριστικά (Feature based Registration)

Η μεθοδολογία ευθυγράμμισης βασισμένης σε γεωμετρικούς περιγραφείς αποτελείται από πολλά διαφορετικά βήματα τα οποία εκτελούνται διαδοχικά. Στην εικόνα 4.1 παρουσιάζεται το διάγραμμα ροής της διαδικασίας. Σαν είσοδος του αλγορίθμου είναι δυο τρισδιάστατα νέφη σημείων από τα οποία το ένα καλείται νέφος σημείων προς ευθυγράμμιση και το άλλο καλείται νέφος σημείων αναφοράς. Ο σκοπός είναι η περιστροφή και η μετακίνηση του νέφους σημείων προς ευθυγράμμιση ως προς το νέφος σημείων αναφοράς, έτσι ώστε αυτά να συμπίπτουν χωρικά με την μικρότερη απόσταση ή διαφορά μεταξύ τους. Σε περίπτωση που κάποιο από αυτά έχει σημαντικά μεγάλο μέγεθος μπορεί να εφαρμοστεί πάνω του υποδειγματοληψία με σκοπό την μείωση του υπολογιστικού κόστους στα επόμενα βήματα και επομένως την μείωση του χρόνου εκτέλεσης, χωρίς όμως να επηρεάζεται σημαντικά η ακρίβεια της μεθόδου.



Εικόνα 4.1. Διάγραμμα ροής της μεθοδολογίας ευθυγράμμισης βασισμένης σε γεωμετρικούς περιγραφείς (descriptors)

Ένας γεωμετρικός περιγραφέας υπολογίζεται για κάθε σημείο των δυο συνόλων. Συγκρίνοντας τους περιγραφείς των δυο νεφών με έναν “έξυπνο” τρόπο (k Nearest

Neighbour kNN), υπολογίζονται ζεύγη αντίστοιχων σημείων μεταξύ του νέφους προς ευθυγράμμιση και του νέφους αναφοράς. Όμως, δεν είναι όλα τα αντίστοιχα σημεία που προκύπτουν ορθά. Τα λανθασμένα αντίστοιχα σημεία (outliers), απορρίπτονται με χρήση μιας επαναληπτικής μεθόδου γνωστή ως RANSAC (Random Sample Consensus). Από τις υπόλοιπες αντιστοιχίες (inliers) υπολογίζονται οι παράμετροι ενός ολικού συμπαγή μετασχηματισμού, ο οποίος ορίζεται από έναν πίνακα περιστροφής και έναν πίνακα μετακίνησης. Ο ολικός μετασχηματισμός τελικά εφαρμόζεται στο προς ευθυγράμμιση νέφος σημείων (floating point cloud) ώστε να το ευθυγραμμίσει με το νέφος σημείων αναφοράς.

4.3.1 Υποδειγματοληψία

Η υποδειγματοληψία είναι ένα προαιρετικό βήμα και έχει ως στόχο τον περιορισμό του πλήθους των σημείων των αρχικών νεφών. Ωστόσο, η εφαρμογή της είναι προτεινόμενη για μεγάλο μέγεθος δεδομένα όπως αυτά της παρούσας εργασίας ώστε να μειωθεί το υπολογιστικό κόστος της διαδικασίας της ευθυγράμμισης καθώς επίσης και ο χρόνος εκτέλεσης. Για την υποδειγματοληψία μπορεί να χρησιμοποιηθεί ένα απλό 3D Voxel Grid φίλτρο το οποίο χωρίζει το χώρο του νέφους σημείων σε voxels (το αντίστοιχο εικονοστοιχείο στις τρεις διαστάσεις). Όλα τα σημεία που ανήκουν σε ένα voxel προσεγγίζονται από ένα σημείο το οποίο αντιστοιχεί στο κέντρο βάρους όλων των σημείων του voxel. Αντί για το κέντρο βάρους του voxel θα μπορούσε να χρησιμοποιηθεί πιο εύκολα και γρήγορα το σημείο που βρίσκεται ακριβώς στο γεωμετρικό κέντρο του κύβου αλλά αυτή η προσέγγιση είναι λιγότερο ακριβής. Σημαντικός παράγοντας είναι ο συντελεστής υποδειγματοληψίας που υποδηλώνει το κατά πόσο θα μειωθεί το πλήθος των σημείων έτσι ώστε να μην χαθεί μεγάλο ποσοστό της πληροφορίας. Η βέλτιστη επιλογή του συντελεστή υποδειγματοληψίας εξετάζεται στο κεφάλαιο 6.

4.3.2 Γεωμετρικοί περιγραφείς (feature descriptors)

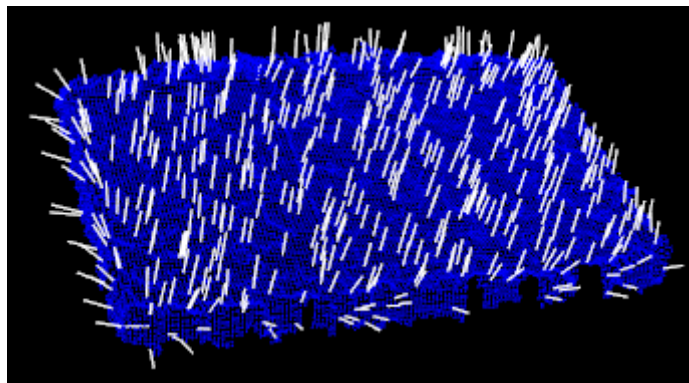
Η βάση της μεθοδολογίας ευθυγράμμισης είναι η εύρεση ενός σετ αντίστοιχων σημείων μεταξύ των δυο νεφών. Για να επιτευχθεί αυτό θα πρέπει τα σημεία να περιγραφούν με έναν σαφή και μοναδικό τρόπο. Παριστάνοντας τα σημεία απλά με τις (x, y, z) συντεταγμένες τους δεν αρκεί. Σε δυο διαδοχικές σάρωσεις π.χ., αν δυο σημεία μπορεί να έχουν τις ίδιες (x, y, z) συντεταγμένες, αυτά μπορεί να ανήκουν σε δυο διαφορετικές επιφάνειες και να αναπαριστούν τελείως διαφορετική πληροφορία όταν εξετάζονται ως προς τα γειτονικά τους σημεία. Μερικές συσκευές σάρωσης νεφών σημείων μπορεί να εισάγουν και άλλη πληροφορία όπως χρώμα, ένταση ή υφή αλλά ούτε αυτό δεν είναι αρκετό για να λυθεί το πρόβλημα και η σύγκριση των σημείων παραμένει ένα δυσεπίλυτο πρόβλημα.

Μια καλή προσέγγιση του προβλήματος είναι να περιγραφεί το κάθε σημείο σε σχέση με τον γειτονικό του χώρο. Έτσι, εισάγεται η έννοια του τοπικού γεωμετρικού περιγραφέα (local feature descriptor). Οι γεωμετρικές ιδιότητες ή τα γεωμετρικά χαρακτηριστικά όπως απόσταση, γωνίες κλπ μιας περιοχής πεπερασμένου αριθμού γειτονικών σημείων γύρω από το προς εξέταση σημείο κωδικοποιούνται έτσι ώστε να δημιουργήσουν τον γεωμετρικό περιγραφέα. Έτσι μας δίνεται μια εικόνα της μορφής της περιοχής γύρω από το προς

εξέταση σημείο. Οι γεωμετρικοί περιγραφείς είναι επομένως μια σύνθετη και αρκετά πιο ακριβής περιγραφή ενός σημείου και μπορεί επίσης να θεωρηθεί ως μια υπογραφή που ξεχωρίζει ένα σημείο ανάμεσα σε άλλα σημεία του συνόλου. Ο σκοπός του είναι η μοναδική αναγνώριση ενός σημείου ανάμεσα σε πολλά νέφη σημείων χωρίς να επηρεάζεται από το θόρυβο, από την πυκνότητα του νέφους σημείων και από το γεγονός ότι πάνω στο νέφος σημείων έχει εφαρμοστεί συμπαγής μετασχηματισμός.

Υπάρχει ένας μεγάλος αριθμός διαθέσιμων γεωμετρικών περιγραφών όπου ο καθένας χρησιμοποιεί διαφορετικά γεωμετρικά χαρακτηριστικά και υπολογίζεται με διαφορετικό τρόπο. Οι γεωμετρικοί περιγραφείς χωρίζονται σε δυο μεγάλες κατηγορίες. Στην πρώτη κατηγορία ανήκουν οι ολικοί γεωμετρικοί περιγραφείς όπου ένας περιγραφέας υπολογίζεται για όλο το νέφος σημείων και χρησιμοποιούνται κυρίως στην 3D αναγνώριση αντικειμένων (3D object recognition), γεωμετρική ανάλυση κλπ. Στην δεύτερη κατηγορία ανήκουν οι τοπικοί γεωμετρικοί περιγραφείς οι οποίοι υπολογίζονται για κάθε σημείο του νέφους και χρησιμοποιούνται συνήθως σε μεθόδους ευθυγράμμισης αλλά και στην 3D αναγνώριση αντικειμένων. Μερικοί από τους τοπικούς γεωμετρικούς περιγραφείς αναλύονται στις επόμενες ενότητες.

4.3.2.1 Κάθετα διανύσματα (normals)



Εικόνα 4.2. Κάθετα διανύσματα (normals). Για λόγους ευκρινείας απεικονίζεται μόνο ένα κάθετο διάνυσμα για κάθε 60 σημεία

Το κάθετο διάνυσμα (normal) ενός σημείου είναι το πιο απλό παράδειγμα ενός γεωμετρικού περιγραφέα και αποτελεί επίσης ένα βασικό στοιχείο άλλων πιο σύνθετων γεωμετρικών περιγραφών που αναλύονται πιο κάτω. Ο υπολογισμός του βασίζεται στην επιλογή ενός κατάλληλου συνόλου γειτονικών σημείων P^k , τα οποία προσδιορίζουν και την γεωμετρία της επιφάνειας γύρω από το σημείο p_q [34]. Για να βρεθεί το κάθετο διάνυσμα ενός σημείου αρκεί να υπολογιστεί το κάθετο διάνυσμα του επιπέδου Π , εφαπτόμενο προς την επιφάνεια και περιέχει το προς εξέταση σημείο. Το επίπεδο μπορεί να αναπαρασταθεί από ένα σημείο x και από το κάθετο διάνυσμα \vec{n} και είναι ο βέλτιστος με την έννοια ότι ελαχιστοποιεί το άθροισμα των τετραγώνων των αποστάσεων των σημείων $p_i \in P^k$ από το επίπεδο Π ,

$$\min \sum_{i=1}^k \text{dist}(p_i, \Pi)^2$$

όπου η απόσταση d_i ορίζεται ως $d_i = (p_i - x) \cdot \vec{n}$ και k είναι το πλήθος των γειτονικών σημείων του p_q . Λαμβάνοντας ως x το κέντρο βάρους του συνόλου P^k ,

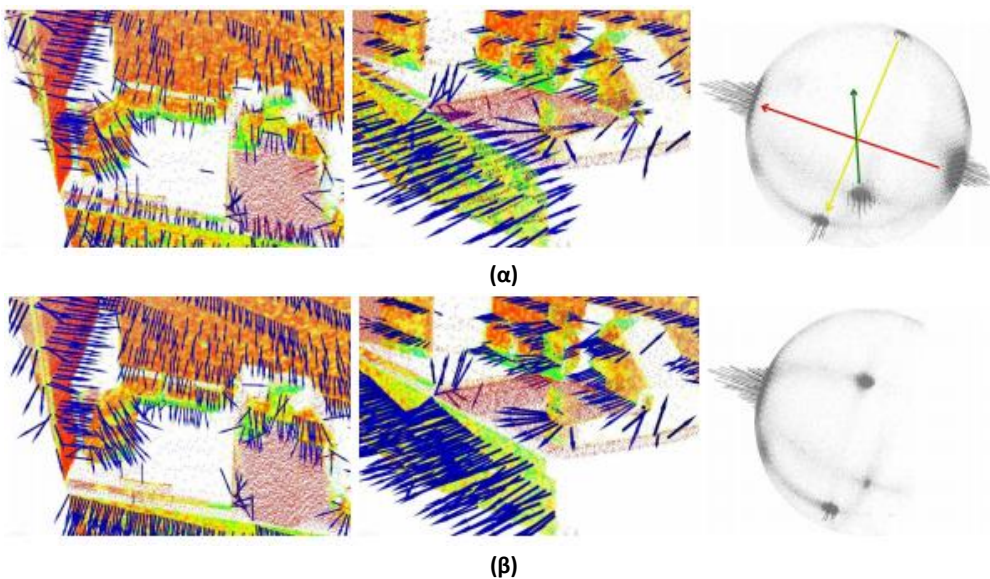
$$x = \bar{p} = \frac{1}{k} \cdot \sum_{i=1}^k p_i \quad (4.1)$$

το διάνυσμα \vec{n} υπολογίζεται αναλύοντας τις ιδιοτιμές και τα ιδιοδιανύσματα του πίνακα $C \in \mathbb{R}^{3 \times 3}$ του P^k ,

$$C = \frac{1}{k} \sum_{i=1}^k \xi_i \cdot (p_i - \bar{p}) \cdot (p_i - \bar{p})^T \quad (4.2)$$

$$C \cdot \vec{v}_j = \lambda_j \cdot \vec{v}_j, j \in \{0, 1, 2\}$$

Ο όρος ξ_i αντιστοιχεί στο βάρος του σημείου p_i και συνήθως ισούται με 1. Ο πίνακας C είναι συμμετρικός και οι ιδιοτιμές του είναι πραγματικοί αριθμοί $\lambda_j \in \mathbb{R}$. Εάν $0 \leq \lambda_0 \leq \lambda_1 \leq \lambda_2$, το ιδιοδιάνυσμα \vec{v}_0 που αντιστοιχεί στην μικρότερη ιδιοτιμή λ_0 είναι η λύση για το μοναδιαίο κάθετο διάνυσμα $+\vec{n} = \{n_x, n_y, n_z\}$ ή $-\vec{n}$. Γενικά, δεν υπάρχει μαθηματικός τρόπος για τον προσδιορισμό του προσήμου του κάθετου διανύσματος \vec{n} . Επομένως τα διανύσματα που υπολογίζονται μέσω της ανάλυσης κυρίων συνιστωσών (Principal Component Analysis PCA), όπως παραπάνω, έχουν αόριστη φορά και δεν είναι ομοιόμορφα προσανατολισμένα σε όλο το νέφος σημείων. Στην εικόνα 4.3α αριστερά παρουσιάζονται δυο τμήματα ενός νέφους σημείων όπου η φορά των κάθετων διανυσμάτων είναι απροσδιόριστη και δεξιά η σφαίρα κάθετων διανυσμάτων (normal sphere ή Extended Gaussian Image) η οποία περιγράφει την φορά όλων των \vec{n}_i .



Εικόνα 4.3. (α) Αριστερά και μεσαία: κάθετα διανύσματα με απροσδιόριστη φορά. Δεξιά: η αντίστοιχη σφαίρα κάθετων διανυσμάτων (normal sphere). (β) Αριστερά και μεσαία: κάθετα διανύσματα ομοιόμορφα προσανατολισμένα προς το σημείο παρατήρησης. Δεξιά: η αντίστοιχη σφαίρα κάθετων διανυσμάτων.

Η λύση του προβλήματος προσδιορισμού της κατεύθυνσης των διανυσμάτων είναι εύκολη εάν είναι γνωστό το σημείο παρατήρησης v_p . Για να προσανατολιστούν όλα τα διανύσματα \vec{n}_i ομοιόμορφα προς το σημείο παρατήρησης πρέπει να ικανοποιηθεί η εξίσωση:

$$\vec{n}_i \cdot (v_p - p_i) > 0 \quad (4.3)$$

Στην εικόνα 4.3β παρουσιάζονται τα ίδια τμήματα νέφους σημείων με προσανατολισμένα κάθετα διανύσματα. Η διαδικασία αυτή είναι πολύ σημαντική στην ευθυγράμμιση νεφών σημείων όπου χρειάζεται μια πολύ ακριβής περιγραφή των σημείων με σκοπό την εύρεση ορθών αντιστοιχιών οι οποίες στην συνέχεια θα χρησιμοποιηθούν για τον προσδιορισμό του μετασχηματισμού.

Στην τελική του μορφή ο συγκεκριμένος γεωμετρικός περιγραφέας είναι ένας πίνακας 3x1 με τιμές τις συντεταγμένες του κάθετου διανύσματος.

$$\vec{n}_i = \begin{pmatrix} n_x \\ n_y \\ n_z \end{pmatrix}$$

4.3.2.2 Fast Point Feature Histograms (FPFH)

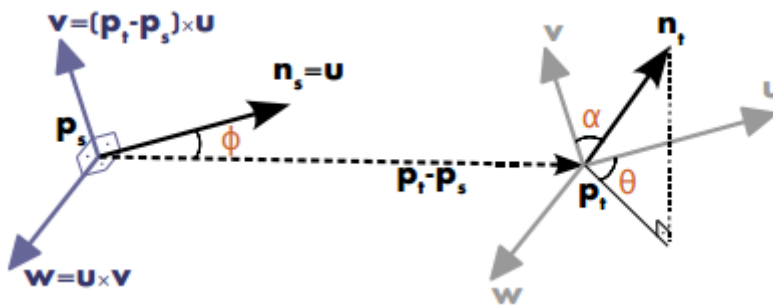
Τα κάθετα διανύσματα, σε σύγκριση με άλλους γεωμετρικούς περιγραφείς είναι σε κάποιο βαθμό στοιχειώδη στην αναπαράσταση της γεωμετρίας γύρω από ένα συγκεκριμένο σημείο. Αν και υπολογίζονται εξαιρετικά γρήγορα και εύκολα, δεν περιγράφουν με αρκετή λεπτομέρεια και ακρίβεια την γεωμετρία του P^k καθώς αυτή προσεγγίζεται με ελάχιστες τιμές. Επομένως, πολλά διαφορετικά σημεία θα παρουσιάσουν ίδιες ή αρκετά συγκρίσιμες τιμές των κάθετων διανυσμάτων μειώνοντας την περιγραφική πληροφορία των σημείων. Αυτό έχει σαν αποτέλεσμα λανθασμένες αντιστοιχίες σημείων και στην συνέχεια μεγάλες αποκλίσεις στην ευθυγράμμιση. Συνεπώς είναι αναγκαία τα σημεία να περιγραφούν με πιο σύνθετους και ακριβείς περιγραφείς.

Ο γεωμετρικός περιγραφέας FPFH κωδικοποιεί τις γεωμετρικές ιδιότητες μιας καθορισμένης περιοχής γύρω από ένα σημείο p_i , η οποία είναι μια σφαίρα ακτίνας r και κέντρο το εξεταζόμενο σημείο, σε ένα πολυδιάστατο ιστόγραμμα. Αυτό επιτυγχάνεται αναλύοντας την διαφορά των διευθύνσεων των κάθετων διανυσμάτων του σημείου p_i και των γειτονικών του σημείων. Η διαφορά αυτή μπορεί να εκφραστεί από ένα σετ γωνιών α, φ, θ οι οποίες αποτελούν και τα χαρακτηριστικά με τα οποία θα σχηματιστεί το τελικό ιστόγραμμα [34] - [35].

Πιο συγκεκριμένα, για το κάθε ζεύγος σημείων ορίζεται αρχικά ένα σταθερό σύστημα συντεταγμένων u, v, w :

$$\begin{cases} u = n_s \\ v = u \times \frac{(p_t - p_s)}{\|p_t - p_s\|_2} \\ w = u \times v \end{cases} \quad (4.4)$$

Ένα από τα σημεία ορίζεται ως σημείο πηγής (source) p_s και το άλλο ως σημείο προορισμού (target) p_t . Το σημείο πηγής επιλέγεται έτσι ώστε η γωνία που σχηματίζεται μεταξύ κάθετου διανύσματος πάνω σε αυτό και της ευθείας που ενώνει τα δυο σημεία να είναι η μικρότερη δυνατή.



Εικόνα 4.4. Γραφική αναπαράσταση του σταθερού συστήματος συντεταγμένων και των γωνιών α, φ, θ για ένα ζευγος σημειων p_s και p_t και τα αντίστοιχά τους κάθετα διανύσματα n_s και n_t .

Με βάση το σταθερό σύστημα συντεταγμένων uvw , η διαφορά μεταξύ των δυο κάθετων διανυσμάτων n_s και n_t εκφράζεται από ένα σετ γωνιών ως εξής:

$$\begin{aligned} \alpha &= v \cdot n_t \\ \varphi &= u \cdot \frac{(p_t - p_s)}{d} \\ \theta &= \arctan(w \cdot n_t, u \cdot n_t) \end{aligned} \quad (4.5)$$

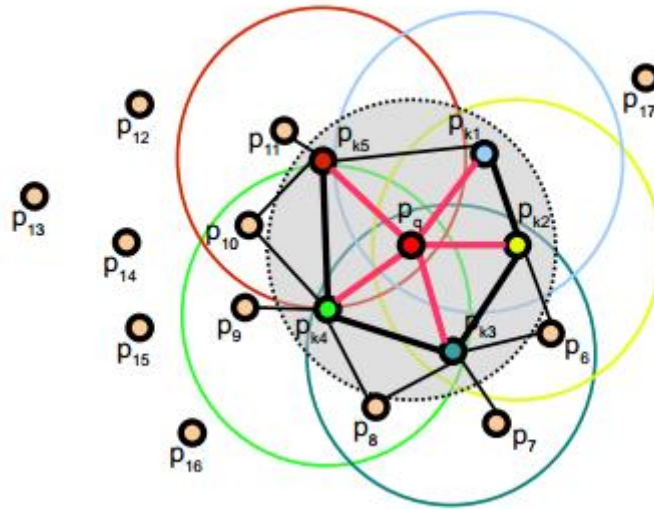
όπου d είναι η Ευκλείδεια απόσταση μεταξύ των σημείων p_s και p_t , $d = \|p_t - p_s\|_2$. Συνήθως, το τέταρτο χαρακτηριστικό, η απόσταση d , δεν θεωρείται σημαντικός παράγοντας, λόγω του ότι η απόσταση μεταξύ γειτονικών σημείων αυξάνεται με την απομάκρυνση από το σημείο παρατήρησης.

Όπως παρουσιάζεται και στην εικόνα 4.4, για κάθε προς εξέταση σημείο p_q , υπολογίζεται ένα σετ γωνιών $\langle \alpha, \varphi, \theta \rangle$ σύμφωνα με την εξίσωση 4.5 μόνο για τα ζεύγη (p_q, p_k) όπου p_k είναι τα γειτονικά σημεία εντός της ακτίνας r . Αυτά τα σετ θα σχηματίσουν το ιστόγραμμα Simplified Point Feature Histogram (SPFH) του p_q .

Στην συνέχεια, για κάθε γειτονικό σημείο p_k , προσδιορίζονται τα νέα του γειτονικά σημεία και με τον ίδιο τρόπο υπολογίζεται το δικό τους ιστόγραμμα SPFH το οποίο θα χρησιμοποιηθεί, πολλαπλασιασμένο με έναν συντελεστή βάρους, στον σχηματισμό του τελικού ιστογράμματος του p_q που καλείται (FPFH) ως εξής:

$$FPFH(p_q) = SPFH(p_q) + \frac{1}{k} \sum_{i=1}^k \frac{1}{\omega_k} \cdot SPFH(p_k) \quad (4.6)$$

Ο συντελεστής βάρους ω_k , αναπαριστά την απόσταση μεταξύ του προς εξέταση σημείου p_q και ενός γειτονικού σημείου p_k εκτός εάν οριστεί αλλιώς.



Εικόνα 4.5. Διάγραμμα της περιοχής της επιρροής για το Fast Point Feature Histogram (FPFH). Το προς εξέταση σημείο p_q (κόκκινο χρώμα) συνδέεται μόνο με τους k -γείτονες εντός της ακτίνας r (κύκλος με γκρι χρώμα, συνδέσεις με κόκκινο χρώμα). Κάθε γειτονικό σημείο συνδέεται με τα δικά του γειτονικά σημεία και τα SPFH ιστογράμματα που προκύπτουν πολλαπλασιασμένα με έναν συντελεστή βάρους μαζί με το SPFH ιστόγραμμα του p_q σχηματίζουν το τελικό FPFH. Με έντονο μαύρο παρουσιάζονται οι συνδέσεις που συμμετέχουν δυο φορές.

Για να δημιουργηθεί ένα ιστόγραμμα SPFH για ένα σημείο p_q , το εύρος τιμών του κάθε χαρακτηριστικού (γωνίες α, φ, θ) χωρίζεται στον ίδιο αριθμό ίσων διαστημάτων. Το ένα διάστημα του ιστογράμματος SPFH αντιστοιχεί στον αριθμό των εμφανίσεων ενός συγκεκριμένου συνδυασμού των τιμών των γωνιών α, φ, θ . Αν ο αριθμός των χαρακτηριστικών είναι d και το εύρος τιμών κάθε χαρακτηριστικού χωρίζεται σε b ίσα διαστήματα, το ιστόγραμμα θα έχει συνολικά b^d διαστήματα. Ο FPFH στην βιβλιοθήκη PCL, η οποία χρησιμοποιήθηκε για την υλοποίηση του εν λόγω περιγραφέα παράγει ιστογράμματα 33 διαστημάτων.

4.3.2.3 SHOT (Signature of Histograms of Orientations)

Ο γεωμετρικός περιγραφέας SHOT είναι ένας συνδυασμός των περιγραφέν με βάση την υπογραφή και των περιγραφέν με βάση το ιστόγραμμα [36] - [37]. Χρησιμοποιώντας τα πλεονεκτήματα των δυο κατηγοριών δημιουργείται ένας περιγραφέας μοναδικός, περισσότερο ακριβής, αποτελεσματικός και ανθεκτικός στο θόρυβο, καθώς και ανεξάρτητος από τις περιστροφές και μετακινήσεις. Ο SHOT κωδικοποιεί μια υπογραφή ιστογραμμάτων τα οποία αναπαριστούν γεωμετρικά χαρακτηριστικά της περιοχής ενός σημείου. Για τον υπολογισμό του αρχικά σε κάθε σημείο p του νέφους σημείων υπολογίζεται ένα τοπικό

σύστημα συντεταγμένων (local reference frame). Τα n γειτονικά σημεία p_i του σημείου p χρησιμοποιούνται για τον υπολογισμό ενός πίνακα συνδιασποράς C (weighted covariance matrix) πολλαπλασιασμένο με έναν συντελεστή βάρους:

$$C = \frac{1}{n} \sum_{i=1}^n (r - \|p_i - p\|) \cdot (p_i - p) \cdot (p_i - p)^T \quad (4.7)$$

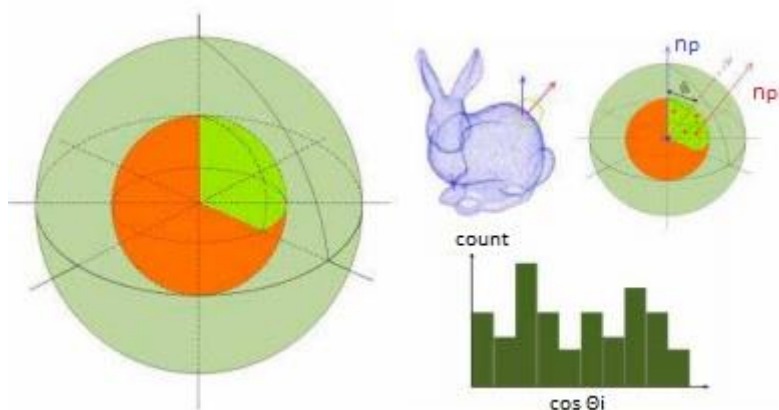
όπου r είναι η ακτίνα της γειτονικής περιοχής του σημείου p . Με ανάλυση του πίνακα συνδιασποράς προκύπτουν τρία ιδιοδιανύσματα τα οποία προσδιορίζουν το τοπικό σύστημα συντεταγμένων του σημείου p . Τα ιδιοδιανύσματα σύμφωνα με τις ιδιοτιμές τους ταξινομούνται σε φθίνουσα σειρά ως v_1, v_2, v_3 και αντιπροσωπεύουν τους X, Y, Z άξονες του τοπικού συστήματος συντεταγμένων. Η φορά του X άξονα προσδιορίζεται από την διεύθυνση των διανυσμάτων με αρχή το p και τέλος τα p_i ως εξής:

$$X = \begin{cases} v_1 & ,if \ |S_x^+| \geq |S_x^-| \\ -v_1 & ,otherwise. \end{cases} \quad (4.8)$$

$$S_x^+ = \{p_i \mid (p_i - p) \cdot v_1 \geq 0\} \quad (4.9)$$

$$S_x^- = \{p_i \mid (p_i - p) \cdot v_1 < 0\}$$

Η φορά του Y άξονα προσδιορίζεται με τον ίδιο τρόπο ενώ η φορά του Z άξονα προσδιορίζεται από τους X, Y άξονες με την προϋπόθεση ότι πρέπει να είναι κάθετοι μεταξύ τους. Το τοπικό σύστημα συντεταγμένων χρησιμοποιείται για την σωστή τοποθέτηση ενός σφαιρικού πλέγματος με κέντρο το σημείο p το οποίο χωρίζει την γειτονική περιοχή σε έναν συγκεκριμένο αριθμό τομών.



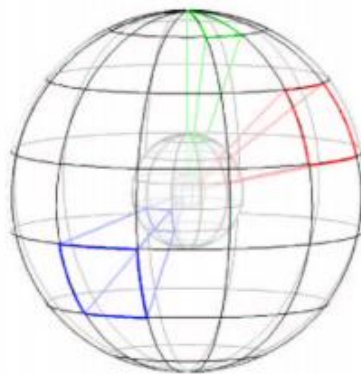
Εικόνα 4.6. Χωρισμός του σφαιρικού πλέγματος με κέντρο το σημείο p σε 32 τομές και υπολογισμός του ιστογράμματος ενός τομέα της σφαίρας με βάση την γωνία που σχηματίζουν τα κάθετα διανύσματα n_p και n_{p_i}

Στο επόμενο βήμα, σε κάθε τομέα της σφαίρας υπολογίζεται η γωνία θ_i που σχηματίζουν τα κάθετα διανύσματα του σημείου p (σημείου του κέντρου της σφαίρας) και των γειτονικών

σημείων p_i . Στη πράξη, χρησιμοποιείται το $\cos \theta_i$ το οποίο υπολογίζεται γρήγορα από την σχέση $\cos \theta_i = n_p \cdot n_{p_i}$. Με όλες τις γωνίες που πρόεκυψαν σε έναν τομέα κατασκευάζεται ένα ιστόγραμμα με b διαστήματα όπου στο καθένα προστίθεται ο αριθμός των σημείων που παρουσιάζουν μια συγκεκριμένη τιμή της γωνίας θ_i . Αν το σφαιρικό πλέγμα περιέχει k τομείς και το κάθε ιστόγραμμα περιέχει b διαστήματα τότε το τελικό ιστόγραμμα θα έχει $k \cdot b$ τιμές. Η βιβλιοθήκη PCL χωρίζει το σφαιρικό πλέγμα σε 32 τομείς και το ιστόγραμμα του κάθε τομέα σε 11 διαστήματα. Η αζιμουθιακή συνιστώσα χωρίζεται σε 8 τμήματα, η συνιστώσα ανύψωσης χωρίζεται σε 2 τμήματα και η ακτινική συνιστώσα επίσης σε 2 τμήματα. Ο περιγραφέας για το σημείο p κατασκευάζεται ενώνοντας με την σειρά όλα τα ιστογράμματα των τομών του σφαιρικού πλέγματος με βάση το τοπικό σύστημα συντεταγμένων. Ο περιγραφέας στην τελική του μορφή είναι μια υπογραφή με $32 \cdot 11 = 352$ τιμές συν 9 τιμές για το τοπικό σύστημα συντεταγμένων οι οποίες όμως δεν χρησιμοποιούνται στην σύγκριση μεταξύ των περιγραφέων, παρά μόνο για την σωστή τοποθέτηση του σφαιρικού πλέγματος στο σημείο προς εξέταση. Η υπογραφή αυτή κανονικοποιείται έτσι ώστε το άθροισμα όλων των διαστημάτων να είναι 1 με σκοπό την αντιμετώπιση διαφορετικής πυκνότητας σημείων σε διαφορετικά νέφη.

4.3.2.4 3DSC (3D Shape Context)

Ο 3DSC είναι μια προέκταση του 2DSC [38], στις τρεις διαστάσεις. Η περιοχή που στηρίζεται ο υπολογισμός του είναι μια σφαίρα με κέντρο το σημείο p (το σημείο για το οποίο υπολογίζεται ο περιγραφέας) και με τον βόριο πόλο της να συμπίπτει με το κάθετο διάνυσμα \vec{n} του p . Η σφαίρα χωρίζεται σε τομείς και στη συνέχεια υπολογίζεται ένα άθροισμα όπως περιγράφεται παρακάτω.



Εικόνα 4.7. Γραφική αναπαράσταση των τομέων της σφαίρας του περιγραφέα 3DSC που αντιστοιχούν στα διαστήματα του τελικού ιστογράμματος

Η αζιμουθιακή συνιστώσα και η συνιστώσα ανύψωσης χωρίζονται σε ίδια μεταξύ τους διαστήματα ενώ η ακτινική συνιστώσα σε λογαριθμική κλίμακα. Έτσι η ακτινική συνιστώσα έχει $J+1$ τμήματα $R = \{R_0 \dots R_J\}$, η συνιστώσα ανύψωσης έχει $K+1$ τμήματα $\Theta = \{\Theta_0 \dots \Theta_K\}$ και η αζιμουθιακή συνιστώσα έχει $L+1$ τμήματα $\Phi = \{\Phi_0 \dots \Phi_L\}$. Το κάθε δοχείο ή τομέας της σφαίρας αντιστοιχεί σε ένα στοιχείο του $J \times K \times L$ διανύσματος. Το

πρώτο ακτινικό τμήμα R_0 έχει ακτίνα r_{\min} ενώ το R_j υποδηλώνει την μέγιστη ακτίνα της σφαίρας r_{\max} . Οι διαφορετικές ακτίνες υπολογίζονται σύμφωνα με τον τύπο:

$$R_j = \exp \left\{ \ln(r_{\min}) + \frac{j}{J} \ln \left(\frac{r_{\max}}{r_{\min}} \right) \right\} \quad (4.10)$$

Οι τομείς που βρίσκονται πιο κοντά στο κέντρο είναι μικρότεροι. Επομένως επιλέγεται μια ακτίνα $r_{\min} > 0$ έτσι ώστε οι πολύ μικρές μεταβολές σχήματος πολύ κοντά στο κέντρο να μην έχουν μεγάλο βάρος στον υπολογισμό του περιγραφέα. Για κάθε σημείο p_i που βρίσκεται εντός του τομέα ή δοχείου (j, k, l) , δηλαδή οι σφαιρικές συντεταγμένες του βρίσκονται στα διαστήματα $[R_j, R_{j+1})$, $[\Phi_k, \Phi_{k+1})$ και $[\Theta_l, \Theta_{l+1})$ υπολογίζεται ένας συντελεστής βάρους $w(p_i)$ όπως στην εξίσωση 4.11, ο οποίος προστίθεται κάθε φορά στην τιμή του συγκεκριμένου τομέα και αντιστοιχεί σε ένα διάστημα του τελικού ιστογράμματος.

$$w(p_i) = \frac{1}{\rho_i \sqrt[3]{V(j, k, l)}} \quad (4.11)$$

Το $V(j, k, l)$ είναι ο όγκος του τομέα και ρ_i είναι η τοπική πυκνότητα σημείων γύρω από το p_i , δηλαδή ο αριθμός των σημείων εντός μιας σφαίρας με κέντρο το p_i και ακτίνα δ . Η τιμή του κάθε διαστήματος του ιστογράμματος αντιστοιχεί σε ένα τομέα της σφαίρας και είναι επομένως το άθροισμα όλων των σημείων του τομέα πολλαπλασιασμένων με τον συντελεστή βάρους $w(p_i)$. Το τελικό ιστόγραμμα κανονικοποιείται για να μην υπάρχει εξάρτηση από την πυκνότητα των σημείων [39].

Ωστόσο υπάρχει ένας βαθμός ελευθερίας ως προς την αξιμουθιακή συνιστώσα ο οποίος πρέπει να αφαιρεθεί για να είναι δυνατή η σύγκριση των 3DSC περιγραφέων υπολογισμένων σε διαφορετικά συστήματα συντεταγμένων. Για να λυθεί αυτό το πρόβλημα, το σφαιρικό πλέγμα περιστρέφεται γύρω από το κάθετο διάνυσμα \vec{n} του σημείου p L φορές, αριθμός που αντιστοιχεί στα διαστήματα που χωρίζεται η αξιμουθιακή συνιστώσα και η παραπάνω διαδικασία υπολογισμού του ιστογράμματος επαναλαμβάνεται για το κάθε διάστημα δίνοντας συνολικά L περιγραφείς για αυτό το σημείο. Η βιβλιοθήκη PCL χρησιμοποιεί ιστόγραμμα του περιγραφέα 3DSC 1980 διαστημάτων.

4.3.2.5 USC-(Unique Shape Context)

Ο USC είναι μια αναβάθμιση του 3DSC με σκοπό να αποφεύγονται οι επιπλέον υπολογισμοί του περιγραφέα για κάθε τμήμα στο οποίο χωρίζεται η αξιμουθιακή συνιστώσα του [40]. Η αναπαράσταση ενός σημείου από πολλούς περιγραφείς όπως στην περίπτωση του 3DSC, όχι μόνο δεν είναι αποδοτική μέθοδος αλλά δυσκολεύει σημαντικά και το επόμενο βήμα, δηλαδή της εύρεσης αντίστοιχων σημείων μεταξύ δυο νεφών όπου πολλαπλοί περιγραφείς πρέπει να συγκριθούν μεταξύ τους αντί για δυο. Η διαδικασία αυτή είναι εξαιρετικά χρονοβόρα, ειδικά όταν η σύγκριση γίνεται με εξαντλητική αναζήτηση. Ο USC προσδιορίζει σε κάθε σημείο p , ένα τοπικό σύστημα συντεταγμένων (local reference frame) το οποίο

καθορίζει μοναδικά στον χώρο την περιοχή στήριξης για τον υπολογισμό του περιγραφέα, δηλαδή το σφαιρικό πλέγμα με κέντρο το p και ακτίνα r . Δεδομένου του σημείου p και μιας σφαιρικής γειτονικής περιοχής ακτίνας r υπολογίζεται αρχικά ένας πίνακας συνδιασποράς M πολλαπλασιασμένος με έναν συντελεστή βάρους ως εξής:

$$M = \frac{1}{Z} \sum_{i:d_i \leq r} (r - d_i)(p_i - p)(p_i - p)^T \quad (4.12)$$

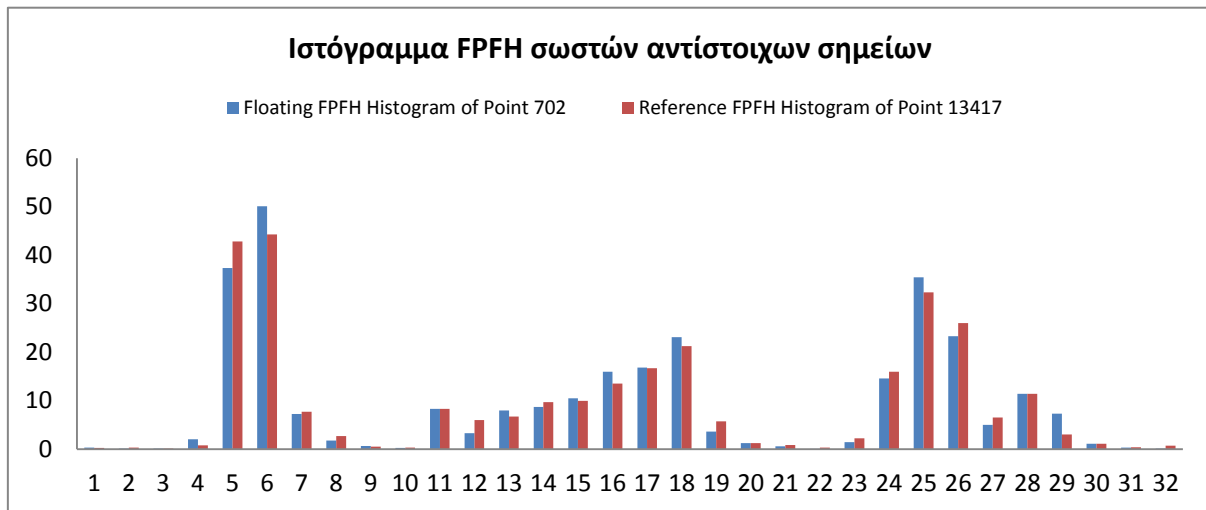
όπου p_i είναι ένα γειτονικό σημείο εντός της σφαιρικής περιοχής, $d_i = \|p_i - p\|_2$, και Z είναι ένας συντελεστής κανονικοποίησης:

$$Z = \sum_{i:d_i \leq r} (r - d_i) \quad (4.13)$$

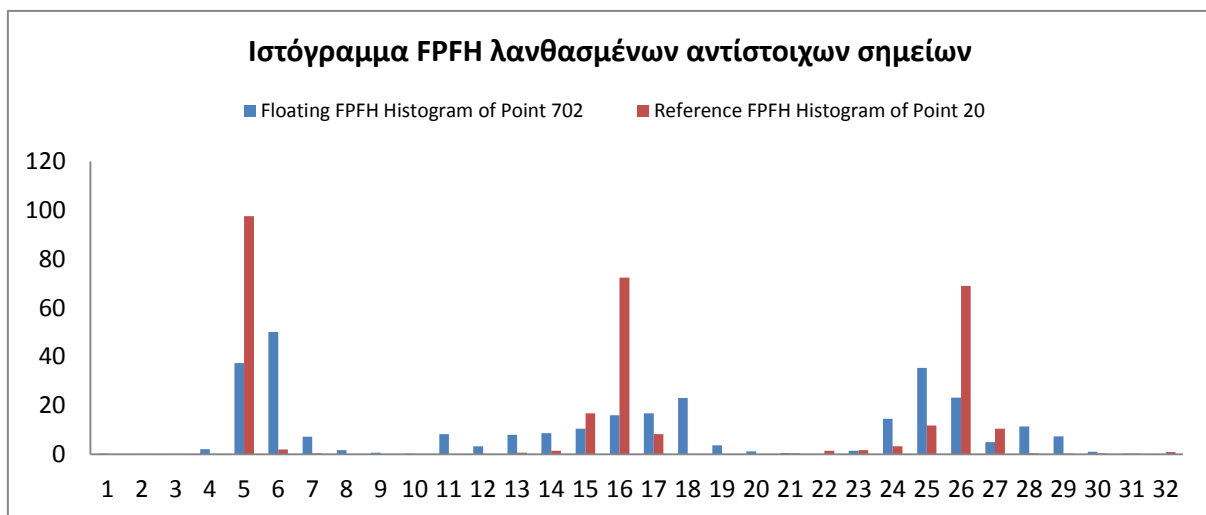
Αναλύοντας τον πίνακα M ως προς τα χαρακτηριστικά ποσά του, προκύπτουν τρία μοναδιαία διανύσματα του τοπικού συστήματος συντεταγμένων. Το πρόσημο των ιδιοδιανυσμάτων με την αντίστοιχη μικρότερη και μεγαλύτερη ιδιοτιμή υπολογίζεται όπως στις εξισώσεις (4.8) και (4.9). Το πρόσημο του τρίτου διανύσματος προκύπτει από το εξωτερικό γινόμενο των δυο προηγούμενων. Το τοπικό σύστημα συντεταγμένων προσδιορίζει μοναδικώς τον προσανατολισμό του σφαιρικού πλέγματος στον χώρο. Στην συνέχεια ο USC περιγραφέας υπολογίζεται ακριβώς με τον ίδιο τρόπο όπως ο 3DSC με την διαφορά ότι σε αυτή την περίπτωση μόνο ένα ιστόγραμμα υπολογίζεται για κάθε σημείο. Θα πρέπει να σημειωθεί ότι ο προσδιορισμός του τοπικού συστήματος συντεταγμένων λύνει το πρόβλημα της ύπαρξης ενός βαθμού ελευθερίας κατά τον υπολογισμό του 3DSC που οδηγούσε σε πολλαπλούς υπολογισμούς του περιγραφέα. Όπως και στην περίπτωση του περιγραφέα 3DSC η βιβλιοθήκη PCL χρησιμοποιεί ιστόγραμμα του περιγραφέα USC 1980 διαστημάτων συν 9 τιμές για το τοπικό σύστημα συντεταγμένων.

4.3.2.6 Γραφική αναπαράσταση γεωμετρικών περιγραφέων

Οι γεωμετρικοί περιγραφείς υπολογίζονται για κάθε σημείο του νέφους σημείων αναφοράς και προς ευθυγράμμιση. Με την σύγκριση τους μπορούν να βρεθούν αντίστοιχα σημεία μεταξύ νεφών σημείων. Στη συνέχεια παρουσιάζονται γραφικά οι γεωμετρικοί περιγραφείς FPFH, SHOT και USC δυο σημείων που αντιστοιχούν και δυο σημείων που δεν αντιστοιχούν. Ο FPFH περιγραφέας είναι ένα ιστόγραμμα 33 διαστημάτων. Για την εύρεση των αντίστοιχων σημείων τα διαστήματα συγκρίνονται ένα προς ένα με την μέθοδο εύρεσης k-Nearest Neighbor (kNN) όπως παρουσιάζεται στις επόμενες ενότητες.



(α)

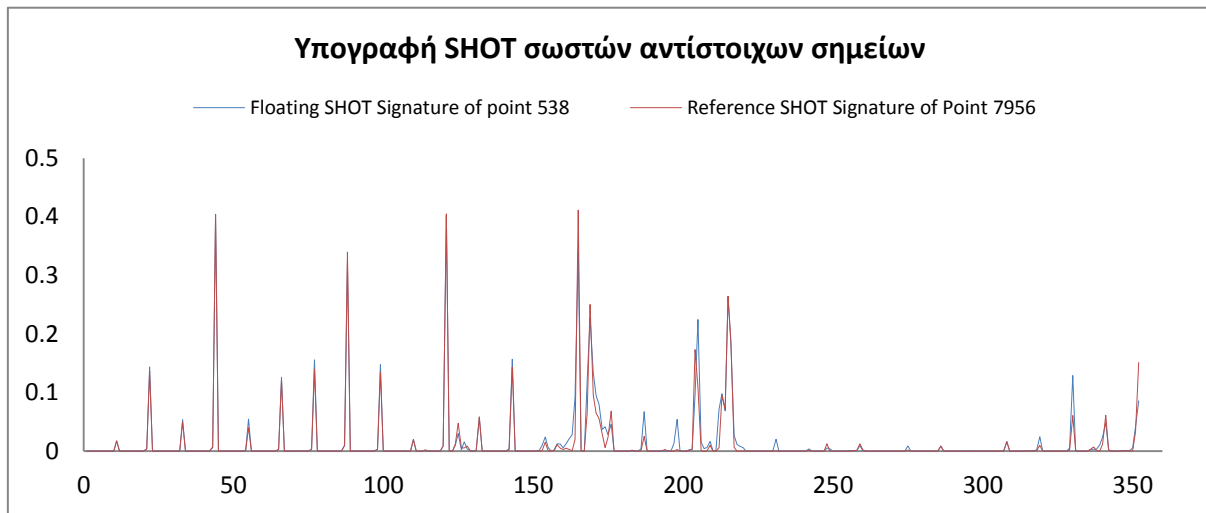


(β)

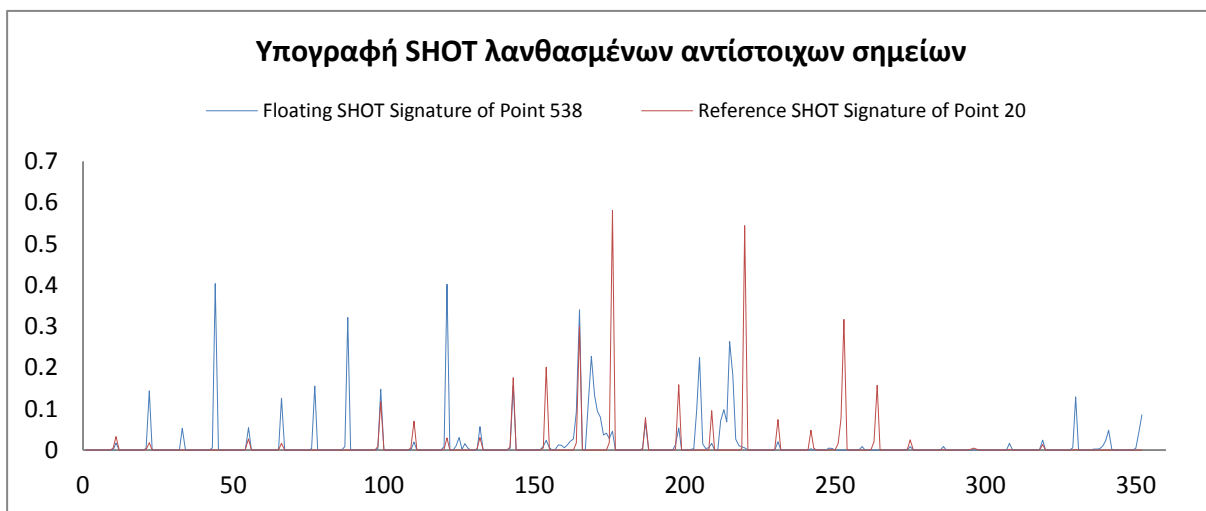
Εικόνα 4.8. Σύγκριση (α) ιστογραμμάτων FPFH σωστών αντίστοιχων σημείων (β) ιστογραμμάτων FPFH λανθασμένων αντίστοιχων σημείων ή σημείων που δεν ταιριάζουν

Ο περιγραφέας SHOT είναι μια υπογραφή που αποτελείται από 32 ιστογράμματα 11 διαστημάτων τα οποία τοποθετούνται το ένα μετά το άλλο και από 9 τιμές του τοπικού συστήματος συντεταγμένων οι οποίες βοηθούν στην σωστή τοποθέτηση της σφαίρας που ορίζει την γειτονική περιοχή αλλά δεν συμμετέχουν στην σύγκριση των περιγραφέων. Ο τελικός περιγραφέας SHOT περιέχει 352 στοιχεία. Όπως και πριν τα αντίστοιχα σημεία προσδιορίζονται με σύγκριση του κάθε διαστήματος με την μέθοδο kNN.

Λόγω του ότι για τον περιγραφέα 3DSC υπολογίζονται πολλαπλοί περιγραφείς για το ίδιο σημείο η γραφική του αναπαράσταση δεν είναι δυνατή.



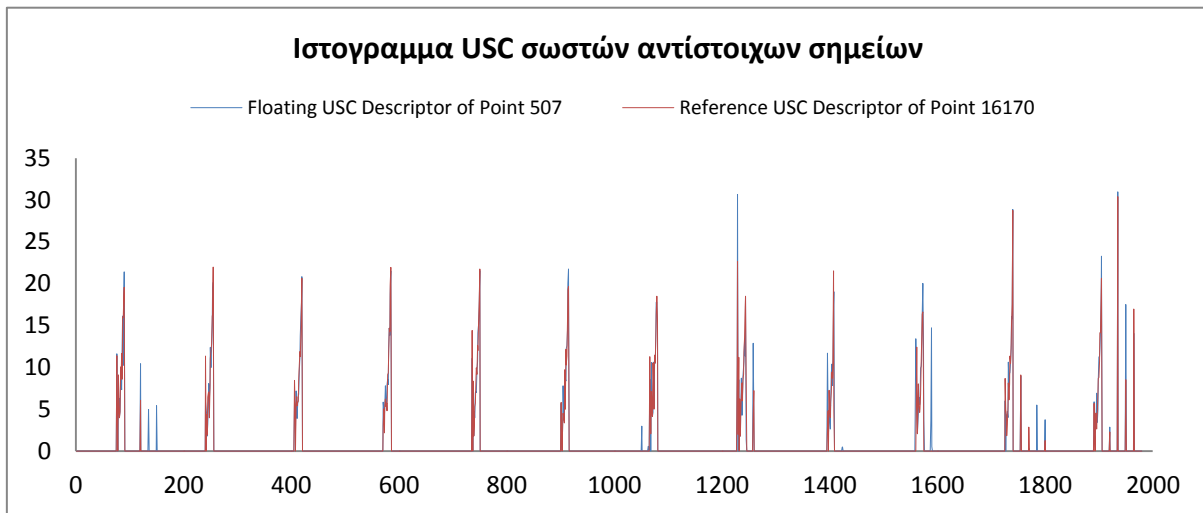
(α)



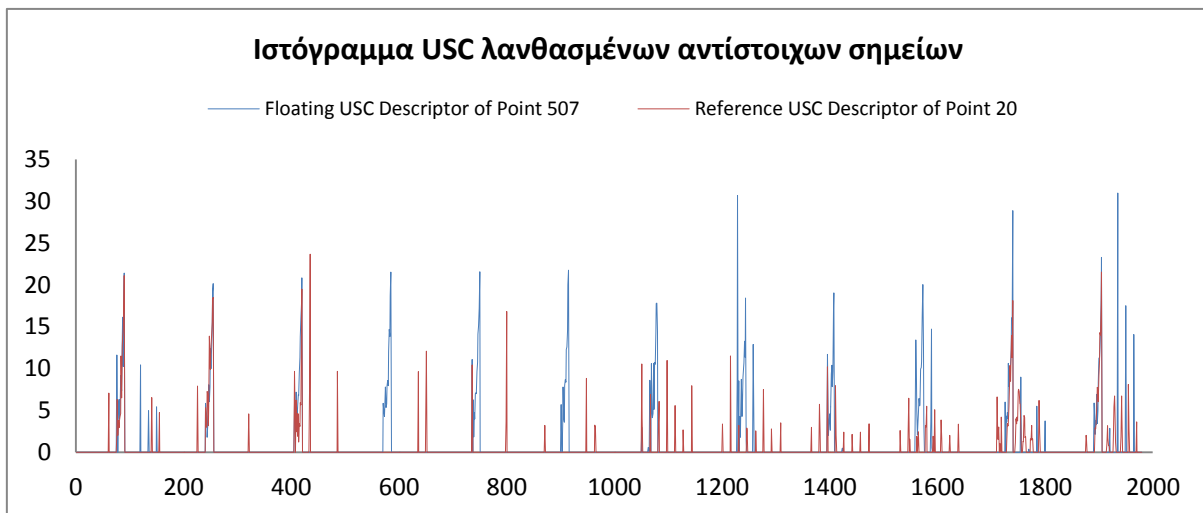
(β)

Εικόνα 4.9. Σύγκριση (α) υπογραφής SHOT σωστών αντίστοιχων σημείων (β) υπογραφής SHOT λανθασμένων αντίστοιχων σημείων ή σημείων που δεν ταιριάζουν

Τα ιστογράμματα USC παρουσιάζονται στην εικόνα 4.10. Ο περιγραφέας αυτός αποτελείται από ένα ιστόγραμμα 1980 διαστημάτων που αντιστοιχούν στις τιμές των τομέων του σφαιρικού πλέγματος και από 9 τιμές που αντιστοιχούν στο τοπικό σύστημα συντεταγμένων και που δεν συμμετέχουν στην σύγκριση μεταξύ των περιγραφών παρά μόνο για την σωστή τοποθέτηση του σφαιρικού πλέγματος και δεν απεικονίζονται.



(α)



(β)

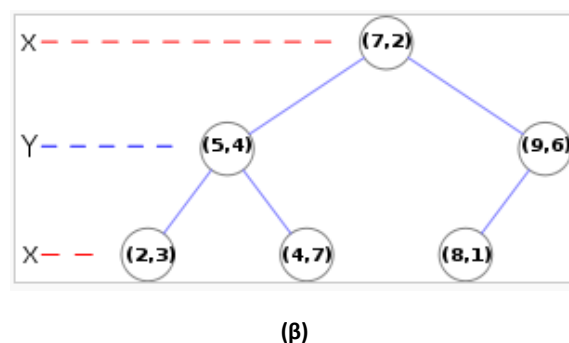
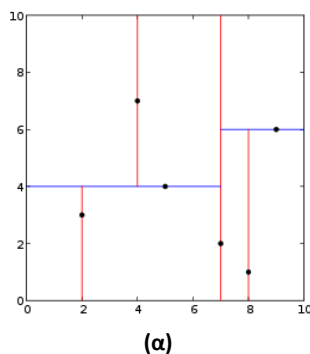
Εικόνα 4.10. Σύγκριση (α) ιστογράμματος USC σωστών αντίστοιχων σημείων (β) ιστογράμματος USC λανθασμένων αντίστοιχων σημείων ή σημείων που δεν ταιριάζουν

4.3.3 Εύρεση αντιστοιχιών σημείων - Correspondence estimation

Εφόσον έχει ολοκληρωθεί ο υπολογισμός των γεωμετρικών περιγραφών για όλα τα σημεία των νεφών σημείων το επόμενο βήμα είναι η εφαρμογή μιας μεθόδου η οποία προσδιορίζει ζεύγη σημείων μεταξύ δυο νεφών με βάση την ομοιότητα των περιγραφών τους. Στο χώρο των γεωμετρικών περιγραφών, οι πιο διαδεδομένες μέθοδοι αντιστοίχισης σημείων είναι δυο. Η πρώτη είναι η μέθοδος εξαντλητικής αναζήτησης (brute force matching) όπου ο περιγραφέας ενός σημείου στο προς ευθυγράμμιση νέφος σημείων συγκρίνεται με όλους τους περιγραφείς των σημείων του νέφους σημείων αναφοράς και επιλέγεται για συσχέτιση το σημείο με το περιγραφέα που έχει την μικρότερη διαφορά με τον αντίστοιχό του. Η εξαντλητική αναζήτηση δεν είναι εφικτή από άποψη χρόνου εκτέλεσης γι' αυτό συνήθως εφαρμόζεται μια άλλη μέθοδος η οποία είναι γνωστή ως k-Nearest Neighbor Search και βασίζεται στη χρήση δενδροειδών δομών kD-tree (k-Dimensional tree).

Ένα kD-tree είναι μια δομή δεδομένων που χρησιμοποιείται κυρίως στην επιστήμη των υπολογιστών για την οργάνωση ενός συνόλου σημείων k-διαστάσεων. Μπορεί επίσης να αναπαρασταθεί σαν ένα δυαδικό δέντρο αναζήτησης όπου ο κάθε κόμβος αντιπροσωπεύει ένα σημείο k-διαστάσεων. Ο κάθε κόμβος ο οποίος δεν είναι φύλλο (τερματικός κόμβος) χωρίζει τον χώρο σε δυο υποχώρους που αναπαρίστανται από το αριστερό και το δεξιό υποδέντρο και ομαδοποιεί τα σημεία σε ένα από αυτά ανάλογα με την τιμή μιας συντεταγμένης τους. Πιο συγκεκριμένα, έστω ένα σύνολο σημείων k-διαστάσεων σε έναν k-διάστατο χώρο. Ως ρίζα του δέντρου επιλέγεται το σημείο του οποίου η συντεταγμένη του πρώτου της άξονα είναι η μεσαία τιμή όλων των συντεταγμένων του ίδιου άξονα. Τα σημεία με την πρώτη συντεταγμένη που είναι μικρότερη της μεσαίας τιμής ομαδοποιούνται σε έναν υποχώρο και τα σημεία που έχουν πρώτη συντεταγμένη μεγαλύτερη της μεσαίας τιμής ομαδοποιούνται στον άλλο υποχώρο. Ο κάθε υποχώρος χωρίζεται ξανά σε δυο υποχώρους με βάση την μεσαία τιμή από όλες τις τιμές συντεταγμένων αλλά αυτή τη φορά του δεύτερου άξονα. Το σημείο με την μεσαία τιμή επιλέγεται ως ρίζα του δέντρου του υποχώρου που αντιπροσωπεύει και με βάση τις συντεταγμένες του δεύτερου άξονα χωρίζει τον υποχώρο με τον ίδιο τρόπο όπως πριν. Η διαδικασία διαχωρισμού συνεχίζεται μέχρι όλοι οι υποχώροι να περιέχουν μόνο ένα σημείο. Σε περίπτωση που ένας υποχώρος περιέχει περισσότερα από ένα σημεία και έχουμε φτάσει στον τελευταίο άξονα ο διαχωρισμός συνεχίζεται με βάση τον πρώτο άξονα και στην συνέχεια τον δεύτερο κλπ, εάν χρειαστεί, μέχρι ο υποχώρος να έχει ένα μοναδικό σημείο. Το κάθε επίπεδο του δέντρου αναζήτησης αντιπροσωπεύει τον άξονα ή την διάσταση με βάση την οποία γίνεται ο διαχωρισμός.

Η kd-tree δομή εφαρμόζεται στο νέφος σημείων αναφοράς. Με είσοδο ένα σημείο από το προς ευθυγράμμιση νέφος σημείων ο αλγόριθμος αναζήτησης του πλησιέστερου σημείου (nearest neighbor search NN) ψάχνει με βάση το kd-tree στο νέφος σημείων αναφοράς για το κοντινότερο σημείο. Ξεκινώντας από την ρίζα του δέντρου σε κάθε κόμβο συγκρίνεται η συντεταγμένη του σημείου εισόδου με την συντεταγμένη του κόμβου και ανάλογα με το αν είναι μικρότερο ή μεγαλύτερο η έρευνα συνεχίζεται στο αριστερό ή στο δεξιό υποδέντρο αντίστοιχα. Η συντεταγμένη ή η διάσταση που εξετάζεται κάθε φορά είναι η διάσταση του αναφέρεται ο κόμβος του δέντρου. Η αναζήτηση σταματά όταν ο κόμβος που εξετάζεται είναι κόμβος φύλλο. Το σημείο αυτό σημειώνεται και ως ο πλησιέστερος γείτονας του σημείου εισόδου.



Εικόνα 4.11. (α) Γραφική αναπαράσταση διαχωρισμού δισδιάστατου χώρου με την kNN μέθοδο. Σημεία εισόδου (2,3), (5,4), (9,6), (4,7), (8,1), (7,2). (β) Το kd-tree σε μορφή δέντρου που προκύπτει από το (α)

4.3.4 Απόρριψη λανθασμένων αντιστοιχιών σημείων - Correspondence rejection

Κατά την εκτίμηση των αντίστοιχων σημείων, είναι πιθανό ορισμένες από αυτές να είναι λανθασμένες ή μη ακριβείς (false correspondences). Επειδή οι λανθασμένες αντιστοιχίες επηρεάζουν αρνητικά στον υπολογισμό του τελικού μετασχηματισμού αυτές πρέπει να απορριφθούν. Αυτό μπορεί να επιτευχθεί με την χρήση της μεθόδου RANSAC (Random Sample Consensus).

Η RANSAC είναι μια επαναληπτική μέθοδος εκτίμησης παραμέτρων ενός μαθηματικού μοντέλου από ένα σύνολο παρατηρούμενων δεδομένων τα οποία περιέχουν και λανθασμένα δεδομένα (outliers). Η τεχνική αυτή είναι μη ντετερμινιστική με την έννοια ότι παράγει ένα λογικό αποτέλεσμα με μια ορισμένη πιθανότητα η οποία αυξάνεται καθώς αυξάνεται ο αριθμός των επαναλήψεων. Μια βασική παραδοχή της μεθόδου RANSAC είναι ότι τα δεδομένα αποτελούνται από τα σωστά δεδομένα (inliers), δηλαδή δεδομένα των οποίων η κατανομή μπορεί να προσεγγιστεί από ένα σετ παραμέτρων ενός μοντέλου και από λανθασμένα δεδομένα τα οποία δεν ταιριάζουν στο μοντέλο. Τα λανθασμένα δεδομένα μπορεί να προέρχονται από το υψηλό επίπεδο θορύβου, από εσφαλμένες μετρήσεις ή από λανθασμένες υποθέσεις σχετικά με την ερμηνεία των δεδομένων. Η δεύτερη παραδοχή είναι ότι με δεδομένο ένα μικρό σετ από inliers υπάρχει μια διαδικασία η οποία μπορεί να εκτιμήσει τις παραμέτρους ενός μοντέλου που περιγράφει βέλτιστα τα δεδομένα αυτά.

Η μέθοδος RANSAC αποτελείται από δυο βασικά βήματα τα οποία επαναλαμβάνονται. Στο πρώτο βήμα, από τα δεδομένα εισόδου επιλέγεται τυχαία ένα υποσύνολο το οποίο περιέχει έναν ελάχιστο αριθμό στοιχείων. Με τα στοιχεία μόνο αυτού του δείγματος υπολογίζεται ένα μαθηματικό μοντέλο και οι παράμετροι του. Ο αριθμός των στοιχείων του δείγματος είναι ο ελάχιστος που μπορεί να προσδιορίσει επαρκώς τις παραμέτρους του μοντέλου. Στο δεύτερο βήμα ελέγχονται όλα τα στοιχεία του συνόλου δεδομένων αν ταιριάζουν στο μοντέλο που υπολογίστηκε στο πρώτο βήμα. Τα στοιχεία που δεν ταιριάζουν στο μοντέλο και απέχουν περισσότερο από αυτό ώστε να ξεπερνούν ένα κατώφλι που ορίζεται από τον χρήστη θεωρούνται ως λανθασμένα δεδομένα ή outliers. Αυτή η διαδικασία επαναλαμβάνεται k φορές, αριθμός που προσδιορίζεται από τον χρήστη. Σε κάθε επανάληψη υπολογίζονται οι παράμετροι ενός μοντέλου και ελέγχεται εάν ο αριθμός των inliers είναι μεγαλύτερος σε σύγκριση με το προηγούμενη εκτέλεση. Μετά από N επαναλήψεις κρατιούνται οι παράμετροι του μοντέλου το οποίο έχει τα περισσότερα inliers όπως αυτό εξηγείται στο δεύτερο βήμα.

Ο αλγόριθμος δέχεται δυο παραμέτρους εισόδου: n , ο αριθμός των στοιχείων του δείγματος για το οποίο υπολογίζονται οι παράμετροι του μοντέλου και t , η τιμή του κατωφλιού. Οι παράμετροι, προσδιορίζονται από τον χρήστη συνήθως με βάση τις πειραματικές εκτιμήσεις. Ωστόσο ο αριθμός των επαναλήψεων k μπορεί να υπολογιστεί με βάση την πιθανότητα να πετύχουμε το επιθυμητό αποτέλεσμα

$$1 - p = (1 - w^n)^k \quad (4.14)$$

$$k = \frac{\log(1 - p)}{\log(1 - w^n)} \quad (4.15)$$

όπου p είναι η πιθανότητα το πλήθος στοιχείων n που επιλέγονται σε μια επανάληψη για τον προσδιορισμό των παραμέτρων του μοντέλου να είναι όλα σωστά δεδομένα ή inliers και w είναι η πιθανότητα να επιλεχθεί ένα inlier κάθε φορά που επιλέγεται ένα στοιχείο. Η p είναι επίσης η πιθανότητα η μέθοδος να παράγει ικανοποιητικό αποτέλεσμα.

$$w = \frac{\text{αριθμός των inliers στο σετ δεδομένων}}{\text{αριθμός των στοιχείων}} \quad (4.16)$$

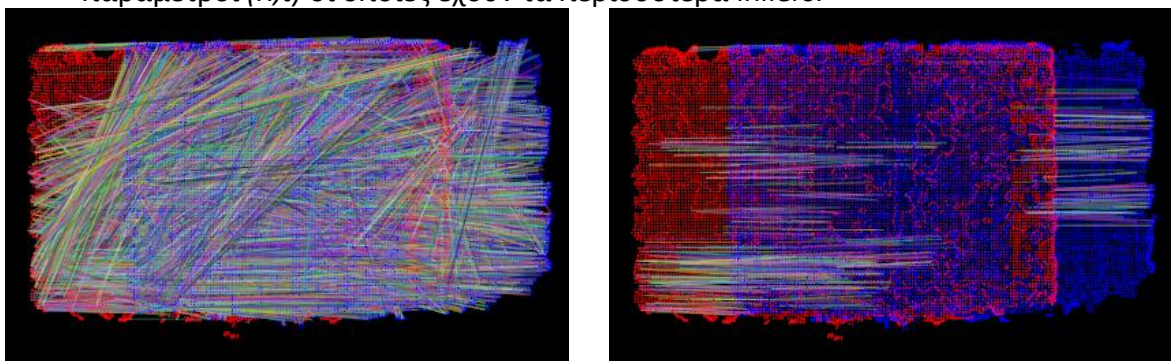
Τα δεδομένα εισόδου της μεθόδου RANSAC σε αυτή την εργασία είναι ένα σετ από αντιστοιχίες σημείων. Συγκεκριμένα για αυτά τα δεδομένα ο αλγόριθμος εκτελείται ως εξής:

1. Επιλέγονται τυχαία από όλα τα ζεύγη αντιστοιχιών τρία από αυτά με p_i ένα σημείο του νέφους σημείων προς ευθυγράμμιση και q_i το αντίστοιχο του σημείο που έχει βρεθεί στο νέφος σημείων αναφοράς
2. Για αυτά τα ζεύγη υπολογίζονται οι παράμετροι του μετασχηματισμού, (R, t) , δηλαδή οι πίνακες περιστροφής και μετατόπισης μέσω της μεθόδου SVD (Singular Value Decomposition).
3. Με αυτές τις παραμέτρους προσδιορίζονται τα inliers σύμφωνα με την σχέση:

$$((Rp_i + t) - q_i)^2 < \varepsilon \quad (4.17)$$

όπου με ε συμβολίζεται το κατώφλι κάτω από το οποίο ένα ζεύγος σημείων θεωρείται σωστό.

4. Τα παραπάνω τρία βήματα επαναλαμβάνονται k φορές και κρατούνται οι παράμετροι (R, t) οι οποίες έχουν τα περισσότερα inliers.



(α)

(β)

Εικόνα 4.12. (α) Αναπαράσταση αντιστοιχιών σημείων για δυο νέφη σημείων προσδιορισμένες με βάση την σύγκριση των γεωμετρικών περιγραφών τους. (β) Σωστές αντιστοιχίες που έχουν απομείνει μετά την απόρριψη των λανθασμένων αντιστοιχιών με εφαρμογή της μεθόδου RANSAC.

4.3.5 Υπολογισμός και εφαρμογή του μετασχηματισμού - Transform estimation and application

Με δοσμένα δυο σύνολα αντίστοιχων σημείων $P = \{p_1, \dots, p_n\}$ και $Q = \{q_1, \dots, q_m\}$ όπου p_i και q_i αντίστοιχα σημεία, οι βέλτιστες παράμετροι του συμπαγή μετασχηματισμού, (R, t) , προκύπτουν όταν η συνάρτηση μέτρησης σφάλματος ελαχιστοποιείται:

$$E(R, t) = \frac{1}{n} \sum_{i=1}^n \|(Rp_i + t) - q_i\|^2 \quad (4.18)$$

Αρχικά υπολογίζεται το κέντρο μάζας των δυο συνόλων:

$$\begin{aligned} \mu_p &= \frac{1}{n} \sum_{i=1}^n p_i \\ \mu_q &= \frac{1}{m} \sum_{i=1}^m q_i \end{aligned} \quad (4.19)$$

Για κάθε σημείο ενός συνόλου αφαιρείται το κέντρο μάζας ως εξής:

$$\begin{aligned} P' &= \{p_i - \mu_p\} = \{p_i'\} \\ Q' &= \{q_i - \mu_q\} = \{q_i'\} \end{aligned} \quad (4.20)$$

Σκοπός είναι η εύρεση του πίνακα περιστροφής R και του πίνακα μετατόπισης t έτσι ώστε να ελαχιστοποιηθεί η συνάρτηση σφάλματος.

Έστω ο πίνακας

$$W = \sum_{i=1}^n q_i' p_i'^T \quad (4.21)$$

Η Singular Value Decomposition του πίνακα W είναι η διάσπαση του σε τρεις πίνακες:

$$W = U \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \end{bmatrix} V^T \quad (4.22)$$

όπου $U, V \in \mathbb{R}^{3 \times 3}$ είναι ορθογώνιοι πίνακες και ισχύει $U^T U = I$, $V^T V = I$ και $\sigma_1 \geq \sigma_2 \geq \sigma_3$. Τα διαγώνια στοιχεία του πίνακα S είναι οι ιδιοτιμές του πίνακα W . Οι στήλες του πίνακα V περιέχουν τα ιδιοδιανύσματα του πίνακα $W^T W$ και οι στήλες του πίνακα U περιέχουν τα ιδιοδιανύσματα του πίνακα $W W^T$. Αν ο βαθμός του πίνακα W είναι $rank(W)=3$, η βέλτιστη λύση του $E(R, t)$ είναι μοναδική και δίνεται από τους παρακάτω τύπους:

$$\begin{aligned} R &= UV^T \\ t &= \mu_q - R\mu_p \end{aligned} \quad (4.23)$$

Οι παράμετροι του μοντέλου που υπολογίζονται στην παραπάνω μέθοδο είναι οι παράμετροι ενός συμπαγούς μετασχηματισμού, ένας 3×3 πίνακας περιστροφής R και ένας 3×1 πίνακας μετατόπισης, t . Σύμφωνα με τον μετασχηματισμό αυτό, το κάθε σημείο f_j του νέφους σημείων προς ευθυγράμμιση (floating point cloud) με συντεταγμένες (x, y, z) μετασχηματίζεται στο σημείο f_j' με συντεταγμένες (x', y', z') για να ευθυγραμμιστεί με το νέφος σημείων αναφοράς (reference point cloud).

$$f_j' = R \cdot f_j + t \quad (4.24)$$

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = R \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \\ t_z \end{pmatrix} \quad (4.25)$$

4.3.6 Συνάρτηση μέτρησης σφάλματος

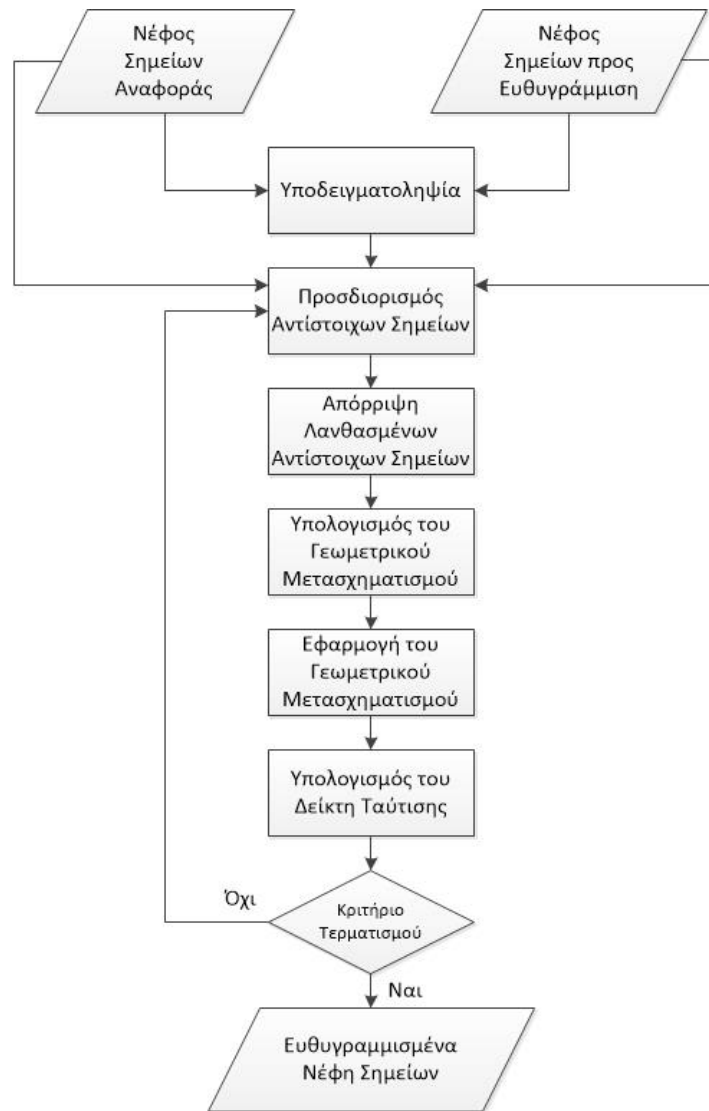
Ο συμπαγής μετασχηματισμός εφαρμόζεται στο προς ευθυγράμμιση νέφος σημείων με σκοπό την βέλτιστη ευθυγράμμιση του με το νέφος σημείων αναφοράς. Στην παρούσα μελέτη το μέγεθος που ποσοτικοποιεί την ακρίβεια της ευθυγράμμισης είναι η συνάρτηση του μέσου τετραγωνικού σφάλματος MSE:

$$MSE = \frac{1}{N_f} \sum_{i=0}^{N_f-1} \|p_f - p_r\|^2 \quad (4.26)$$

Το MSE υπολογίζει το σφάλμα μεταξύ σημείων του νέφους σημείων αναφοράς και του μετασχηματισμένου νέφους σημείων (aligned point cloud). Ο αριθμός των σημείων των νεφών αναφοράς και προς ευθυγράμμιση του ίδιου αντικειμένου που συλλέγονται με τον αισθητήρα Microsoft Kinect αλλάζει και επομένως τα αντίστοιχα σημεία δεν είναι γνωστά εξ'αρχής. Επομένως, για κάθε σημείο $p_f = (x_f, y_f, z_f)^T$ του μετασχηματισμένου νέφους σημείων προσδιορίζεται ο πλησιέστερος γείτονας $p_r = (x_r, y_r, z_r)^T$ στο νέφος σημείων αναφοράς με την βοήθεια της μεθόδου εύρεσης kNN με βάση τις (x, y, z) συντεταγμένες. Το άθροισμα των αποστάσεων στο τετράγωνο όλων των ζευγών σημείων, πολλαπλασιασμένο με $1/N_f$ όπου N_f το πλήθος των σημείων του μετασχηματισμένου νέφους, δίνει το μέσο τετραγωνικό σφάλμα ή το σφάλμα της ευθυγράμμισης. Όσο πιο μικρή η τιμή του MSE τόσο πιο επιτυχής είναι η ευθυγράμμιση.

4.4 Μεθοδολογία ευθυγράμμισης βασισμένης σε επαναλήψεις (Iteration based Registration)

Η μέθοδος ευθυγράμμισης βασισμένη σε επαναλήψεις είναι μια επαναληπτική μέθοδος η οποία υπολογίζει και ενημερώνει έναν συμπαγή μετασχηματισμό σε κάθε επανάληψη με σκοπό την ελαχιστοποίηση της απόστασης δυο νεφών σημείων προς ευθυγράμμιση. Μια τέτοια μέθοδος είναι η επαναληπτική μέθοδος πλησιέστερου σημείου (Iterative Closest Point ICP) [41].



Εικόνα 4.13. Διάγραμμα ροής της μεθοδολογίας ευθυγράμμισης βασισμένης σε επαναλήψεις.

Σε αντίθεση με την μέθοδο ευθυγράμμισης βασισμένης σε γεωμετρικούς περιγραφείς στην μέθοδο ICP τα σημεία δεν περιγράφονται από έναν γεωμετρικό περιγραφέα. Οι αντιστοιχίες σημείων προσδιορίζονται με βάση την σχετική τους απόσταση στον τρισδιάστατο χώρο. Αν και τα περισσότερα βήματα του αλγορίθμου είναι παρόμοια με τα βήματα της προηγούμενης μεθόδου, υπάρχουν σημαντικές διαφορές στον τρόπο λειτουργίας του που αναλύονται πιο κάτω.

Το διάγραμμα ροής της μεθοδολογίας παρουσιάζεται στην εικόνα 4.13.

4.4.1 Υποδειματοληψία

Αν και ο αλγόριθμος πλησιέστερου σημείου είναι σχετικά ένας γρήγορος αλγόριθμος, όταν τα δεδομένα εισόδου είναι πολύ μεγάλου μεγέθους μπορεί να εφαρμοστεί πάνω τους υποδειματοληψία με σκοπό την επιτάχυνση της διαδικασίας. Αυτό επιτυγχάνεται με

εφαρμογή ενός Voxel Grid φίλτρου όπως αυτό που παρουσιάζεται στην ενότητα 4.3.1. Η υποδειγματοληψία είναι ένα προαιρετικό στάδιο.

4.4.2 Υπολογισμός αντιστοιχιών σημείων

Το επόμενο βήμα είναι η εύρεση αντιστοιχιών σημείων μεταξύ των δυο νέφους όπου για κάθε σημείο p_i του νέφους σημείων προς ευθυγράμμιση προσδιορίζεται ο πλησιέστερος γείτονας q_j στο νέφος σημείων αναφοράς μέσω της μεθόδου kNN ο τρόπος λειτουργίας της οποίας παρουσιάζεται με λεπτομέρεια στην παράγραφο 4.3.3. Ο πλησιέστερος γείτονας ορίζεται με βάση την Ευκλείδεια απόσταση ως:

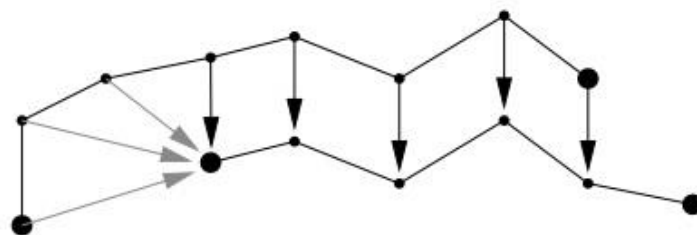
$$\hat{i} = \arg \min_i \|p_i - q_j\|^2 \quad (4.27)$$

όπου $i \in [0, 1, \dots, N]$ και N είναι το πλήθος των σημείων του νέφους σημείων προς ευθυγράμμιση.

Εφόσον τα σύνολα δεδομένων είναι τρισδιάστατα και τα σημεία έχουν τρεις συντεταγμένες οι kd-tree δομές που χρησιμοποιούνται σε αυτή την μέθοδο είναι τρισδιάστατες. Μια ιδιότητα της ICP μεθόδου είναι ότι ένα σετ αντίστοιχων σημείων υπολογίζεται σε κάθε επανάληψη μεταξύ του νέφους σημείων αναφοράς και του μετασχηματισμένου νέφους σημείων της προηγούμενης εκτέλεσης.

4.4.3 Απόρριψη λανθασμένων αντιστοιχιών

Οι αντιστοιχίες σημείων περιέχουν επίσης και λανθασμένα ζεύγη τα οποία επηρεάζουν αρνητικά τον υπολογισμό του μετασχηματισμού και επομένως πρέπει να απορριφθούν. Συνήθως λανθασμένες αντιστοιχίες εισάγουν τα ακραία σημεία όπου αν και ταυτίζονται με το πλησιέστερο σημείο αυτό δεν είναι πάντα το σωστό. Μια τέτοια περίπτωση παρουσιάζεται στην εικόνα 4.14.



Εικόνα 4.14. Αναπαράσταση λανθασμένων αντιστοιχιών σημείων από ακραία σημεία κατά την ταύτιση τους με το πλησιέστερο σημείο με βάση την σχετική απόσταση

Οι λανθασμένες αντιστοιχίες σημείων μπορούν να απορριφθούν με την μέθοδο RANSAC.

4.4.4 Υπολογισμός του μετασχηματισμού

Από τις αντιστοιχίες σημείων (correspondences) που απομένουν υπολογίζονται οι παράμετροι ενός συμπαγή μετασχηματισμού, ο οποίος ορίζεται από ένα πίνακα

περιστροφής R και ένας πίνακας μετατόπισης t . Οι πίνακες αυτοί υπολογίζονται με βάση την μέθοδο SVD (Singular Value Decomposition) ώστε να ελαχιστοποιηθεί η τετραγωνική απόσταση των σημείων:

$$\hat{R}, \hat{t} = \arg \min_{R, t} \sum_{i=1}^N \|(Rp_i + t) - q_i\|^2 \quad (4.28)$$

4.4.5 Εφαρμογή του μετασχηματισμού

Ο μετασχηματισμός που προκύπτει εφαρμόζεται στο προς ευθυγράμμιση νέφος σημείων. Όπως και στις μεθόδους ευθυγράμμισης βάσει γεωμετρικών περιγραφών, η συνάρτηση επιτυχίας που επιλέχθηκε είναι η MSE.

$$MSE = \frac{1}{N} \sum_{i=1}^N \|(Rp_i + t) - q_i\|^2 \quad (4.29)$$

4.4.6 Κριτήρια τερματισμού

Ο αλγόριθμος εκτελεί επαναλήψεις των παραπάνω βημάτων μέχρι να ικανοποιηθεί μια από τις συνθήκες για τον τερματισμό του. Οι συνθήκες στις οποίες ο αλγόριθμος τερματίζει τις επαναλήψεις είναι:

1. Μέγιστος αριθμός επαναλήψεων. Είναι μια παράμετρος που καθορίζεται από τον χρήστη.
2. Η μέθοδος συγκλίνει, δηλαδή ο μετασχηματισμός δεν αλλάζει σημαντικά. Το άθροισμα των διαφορών μεταξύ του προηγούμενου και του τρέχοντος μετασχηματισμού είναι μικρότερο από ένα προκαθορισμένο από τον χρήστη κατώφλι
3. Ύπαρξη μιας λύσης. Η τιμή της συνάρτησης του σφάλματος είναι μικρότερη από ένα κατώφλι το οποίο καθορίζεται από τον χρήστη.

5 Αποτελέσματα

Σε αυτό το κεφάλαιο παρουσιάζονται τα αποτελέσματα της εφαρμογής των μεθόδων ευθυγράμμισης που περιγράφηκαν, στα διαθέσιμα τρισδιάστατα νέφη σημείων. Η αποδοτικότητα και η ακρίβεια των τεχνικών αυτών αξιολογήθηκε τόσο ποιοτικά όσο και ποσοτικά. Τα αποτελέσματα του ποιοτικού και του ποσοτικού ελέγχου παρουσιάζονται στις ακόλουθες ενότητες.

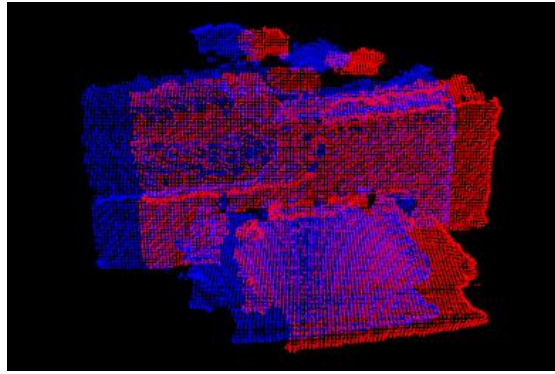
5.1 Ποιοτικά αποτελέσματα

Η ποιοτική αποτίμηση των ευθυγραμμισμένων δεδομένων έγινε μέσω οπτικής εξέτασης. Προκειμένου να γίνει η οπτική αξιολόγηση των μεθόδων αυτών, με την βοήθεια της βιβλιοθήκης PCL αναπτύχθηκε λογισμικό το οποίο απεικονίζει τα νέφη σημείων αναφοράς (reference point cloud) και προς ευθυγράμμιση (floating point cloud) πριν και μετά την ευθυγράμμιση. Μερικά παραδείγματα παρουσιάζονται στις εικόνες 5.1, 5.2, 5.3. Αρχικά απεικονίζονται τα δυο νέφη σημείων πριν την ευθυγράμμιση και στην συνέχεια τα ευθυγραμμισμένα νέφη για τις διάφορες τεχνικές και γεωμετρικούς περιγραφείς.

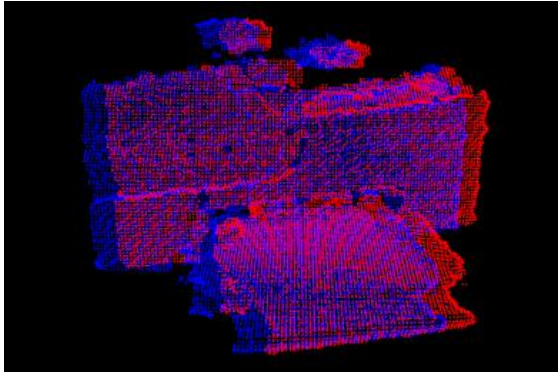
Σε κάθε περίπτωση, με κόκκινο χρώμα αναπαρίσταται το νέφος σημείων αναφοράς και με μπλε χρώμα το νέφος σημείων προς ευθυγράμμιση.

Στην εικόνα 5.1 τα δυο νέφη σημείων αναπαριστούν έναν εκτυπωτή και ελήφθησαν με την χρήση ενός μη βαθμονομημένου αισθητήρα Microsoft Kinect σε διαφορετικές χρονικές στιγμές και με μετακίνηση του αντικειμένου προς τα αριστερά. Όπως φαίνεται, για το συγκεκριμένο σετ δεδομένων, ο αλγόριθμος ICP και τα κάθετα διανύσματα (normals) αποτυγχάνουν στην ευθυγράμμιση των δεδομένων. Η μέγιστη ακρίβεια επιτυγχάνεται με την χρήση των FPFH, SHOT, 3DSC και USC γεωμετρικών περιγραφέων.

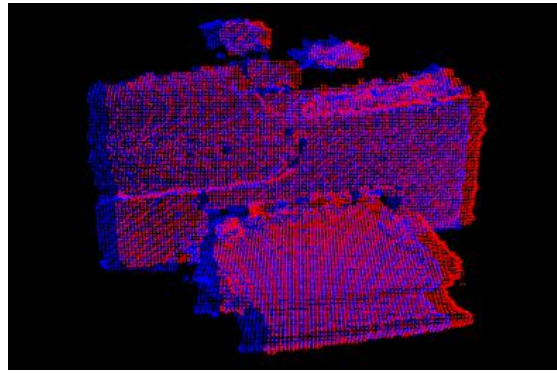
Στην εικόνα 5.2 τα δυο νέφη σημείων προς ευθυγράμμιση αναπαριστούν μια οθόνη, αντικείμενο για το οποίο τα νέφη σημείων ελήφθησαν όπως περιγράφεται στην ενότητα 4.2. Για το συγκεκριμένο σετ δεδομένων ο αλγόριθμος ICP, ο περιγραφέας FPFH και τα κάθετα διανύσματα αποτυγχάνουν να ευθυγραμμίσουν τα δεδομένα. Ο περιγραφέας SHOT επιτυγχάνει μέτρια ακρίβεια. Οι USC και 3DSC, καταφέρνουν να ευθυγραμμίσουν με επιτυχία τα δυο σύνολα και επιτυγχάνουν την μέγιστη ακρίβεια σε σύγκριση με τις υπόλοιπες μεθόδους. Παρατηρώντας επίσης τις τιμές της συνάρτησης σφάλματος για τα δεδομένα αυτού του αντικειμένου το συμπέρασμα αυτό επιβεβαιώνεται. Πιο συγκεκριμένα η τιμή του MSE είναι 0.001220m για τα μη ευθυγραμμισμένα δεδομένα, 0.000017m και 0.000019m για τους USC και 3DSC αντίστοιχα, 0.000106m για τον SHOT 0.003376m για τον FPFH, 0.00014m για τα κάθετα διανύσματα και 0.00017m για τον ICP.



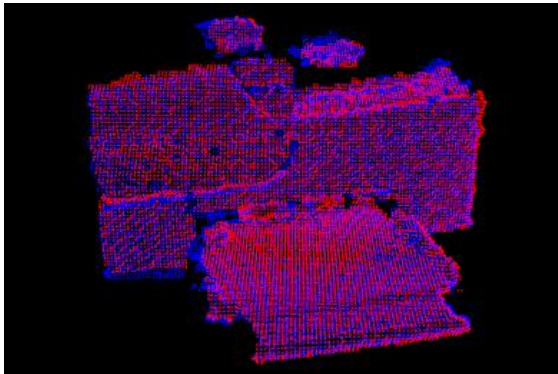
(α)



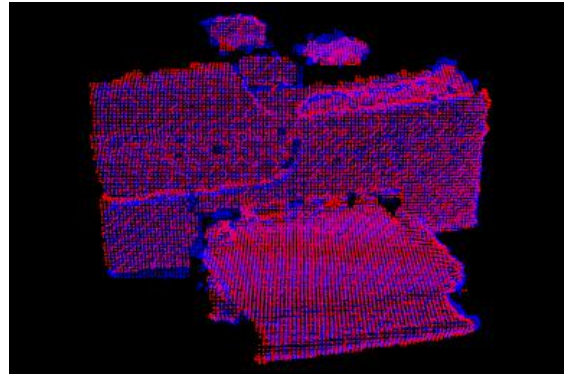
(β)



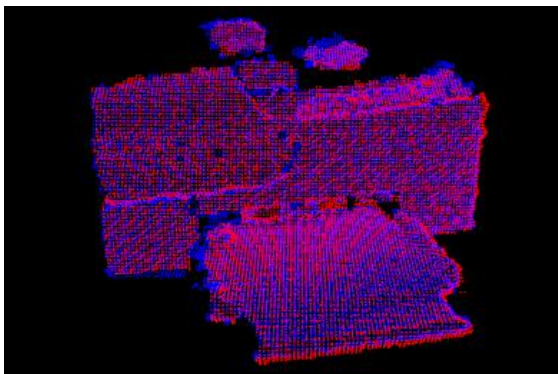
(γ)



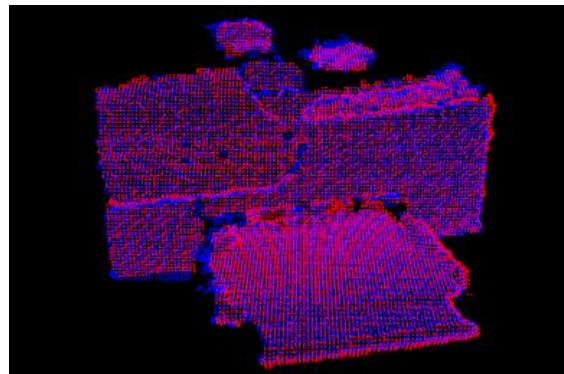
(δ)



(ε)

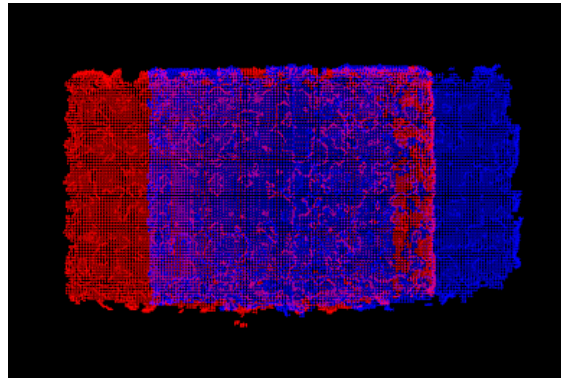


(στ)

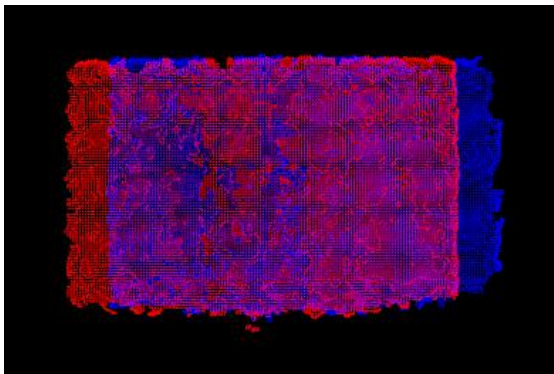


(ζ)

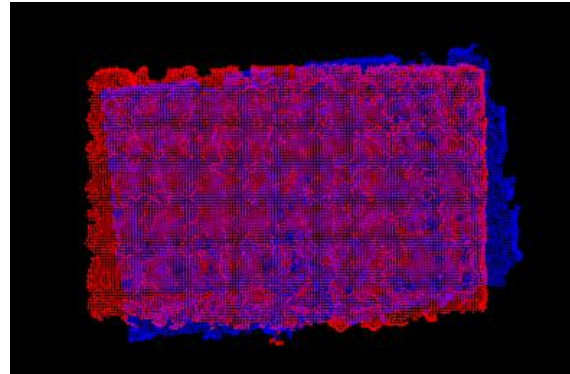
Εικόνα 5.1. Αποτελέσματα ευθυγράμμισης ενός εκτυπωτή. (α) Νέφος σημείων αναφοράς και προς ευθυγράμμιση πριν την ευθυγράμμιση. Αποτελέσματα με βάση (β) ICP αλγόριθμο (γ) κάθετα διανύσματα και περιγραφείς (δ) FPFH (ε) SHOT (στ) 3DSC (ζ) USC



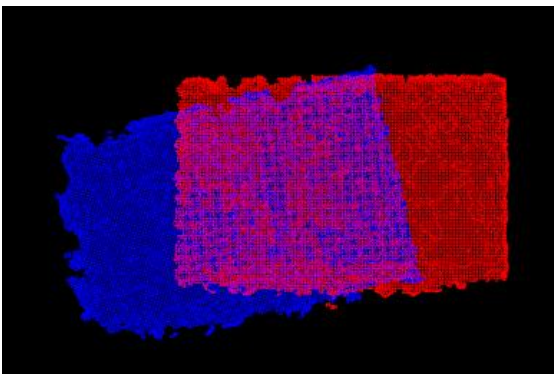
(α)



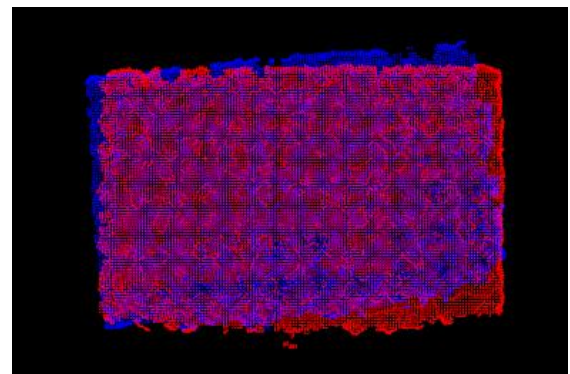
(β)



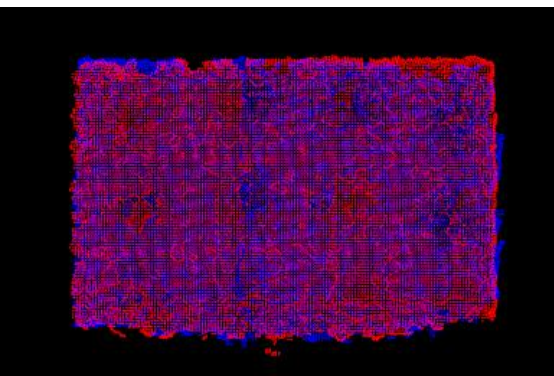
(γ)



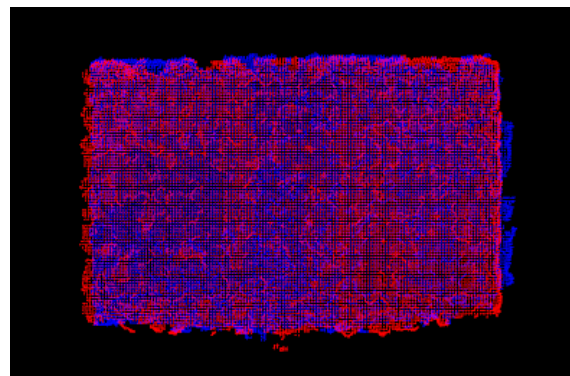
(δ)



(ε)

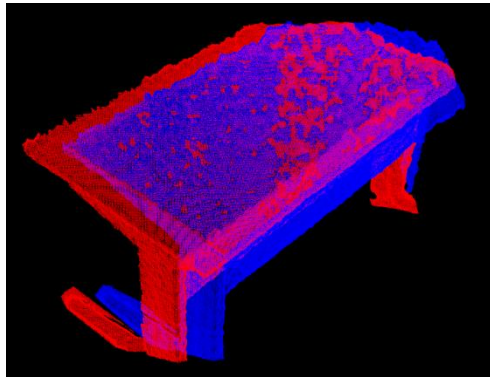


(στ)

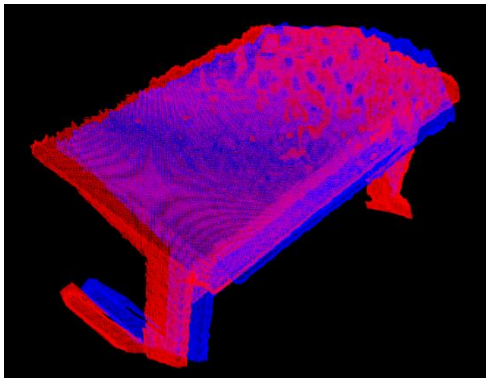


(ζ)

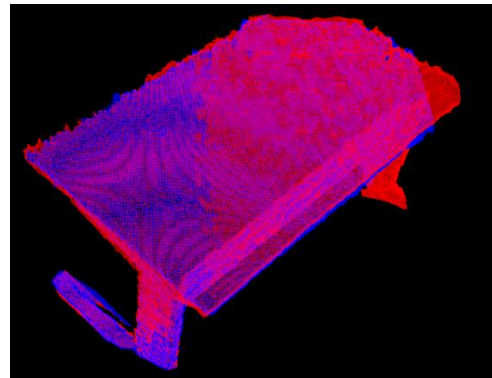
Εικόνα 5.2. Αποτελέσματα ευθυγράμμισης μιας οθόνης. (α) Νέφος σημείων αναφοράς και προς ευθυγράμμιση πριν την ευθυγράμμιση. Αποτελέσματα με βάση: (β) ICP αλγόριθμο (γ) κάθετα διανύσματα και περιγραφείς (δ) FPFH (ε) SHOT (στ) 3DSC (ζ) USC



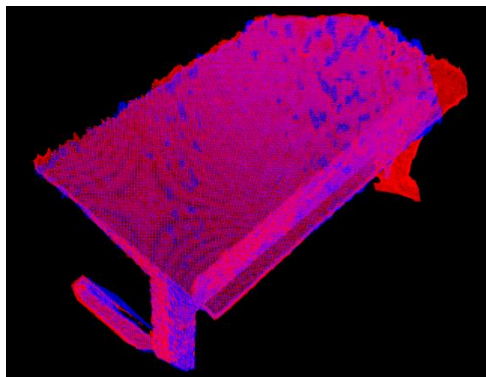
(α)



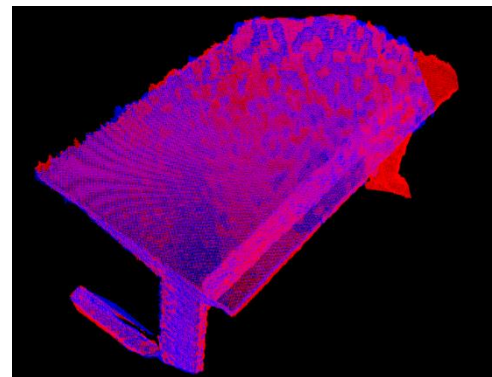
(β)



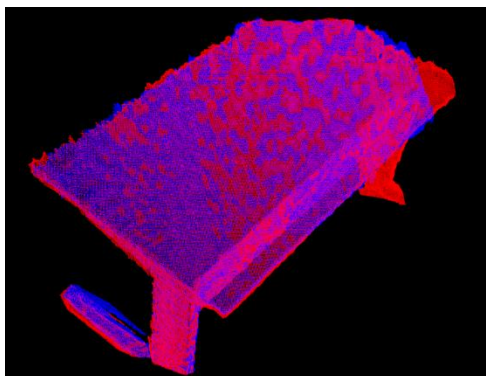
(γ)



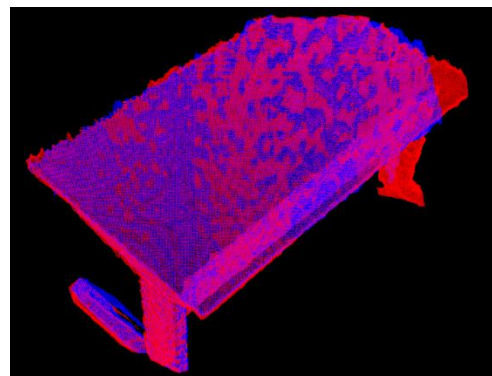
(δ)



(ε)



(στ)



(ζ)

Εικόνα 5.3. Αποτελέσματα ευθυγράμμισης ενός τραπεζιού. (α) Νέφος σημείων αναφοράς και προς ευθυγράμμιση πριν την ευθυγράμμιση. Αποτελέσματα με βάση (β) ICP αλγόριθμος (γ) κάθετα διανύσματα και περιγραφείς (δ) FPFH (ε) SHOT (στ) 3DSC (ζ) USC

Στην εικόνα 5.3 τα νέφη σημείων αναπαριστούν ένα τραπέζι. Όπως παρατηρείται ο αλγόριθμος ICP αποτυγχάνει να ευθυγραμμίσει τα δυο σύνολα δεδομένων. Τα κάθετα διανύσματα καταφέρνουν ευθυγράμμιση με μέτρια ακρίβεια ενώ οι τέσσερις γεωμετρικοί περιγραφείς, FPFH, SHOT, 3DSC και USC ευθυγραμμίζουν τα δεδομένα επιτυχώς.

Προφανώς, μόνο η ποιοτική αποτίμηση των αποτελεσμάτων δεν μπορεί να οδηγήσει σε ασφαλή συμπεράσματα σχετικά με την απόδοση των τεχνικών ευθυγράμμισης που παρουσιάστηκαν. Στις εικόνες 5.1 και 5.3 για παράδειγμα, η διαφορά στην ακρίβεια μεταξύ των FPFH, SHOT, 3DSC και USC δεν μπορεί να παρατηρηθεί εύκολα καθώς όλοι οι περιγραφείς ευθυγραμμίζουν σχεδόν άριστα τα δεδομένα αυτά. Για μια καλύτερη αξιολόγηση των παραπάνω τεχνικών χρησιμοποιήθηκε η συνάρτηση μέτρησης σφάλματος MSE η οποία ποσοτικοποιεί την επιτυχία της ευθυγράμμισης. Τα ποσοτικά αυτά αποτελέσματα παρουσιάζονται στην επομένη ενότητα.

5.2 Ποσοτικά αποτελέσματα

Έκτος από την ποιοτική εξέταση των ευθυγραμμισμένων δεδομένων, πραγματοποιήθηκε επίσης και ποσοτικός έλεγχος των αποτελεσμάτων με την βοήθεια της συνάρτησης μέτρησης σφάλματος *MSE* σύμφωνα με την εξίσωση 4.26. Η τιμή της συνάρτησης σφάλματος υπολογίστηκε για την κάθε τεχνική για 17 διαφορετικά σετ δεδομένων. Τα αποτελέσματα παρουσιάζονται στον πίνακα 5.1. Στην δεύτερη στήλη του πίνακα παρουσιάζονται οι τιμές της συνάρτησης σφάλματος για κάθε ζεύγος νεφών σημείων πριν την ευθυγράμμιση και στις ακόλουθες στήλες οι τιμές *MSE* μετά την ευθυγράμμιση για τους διάφορους γεωμετρικούς περιγραφείς καθώς και για τον αλγόριθμο ICP. Στην τελευταία σειρά του πίνακα 5.1 υπολογίζεται η μέση τιμή του μέσου τετραγωνικού σφάλματος (μ) και η τυπική απόκλιση (σ) όλων των ζευγών δεδομένων.

Όπως παρατηρείται, στον πίνακα 5.1, η μέθοδος ευθυγράμμισης που βασίζεται στον USC γεωμετρικό περιγραφέα είναι η πιο ακριβής με μέση τιμή συνάρτησης σφάλματος 0.000100883m. Συγκριτική ακρίβεια παρουσιάζει ο 3DSC περιγραφέας με μέση τιμή της συνάρτησης σφάλματος 0.000107229. Ακολουθεί ο περιγραφέας SHOT η μέση τιμή του *MSE* του οποίου είναι 0.000124509. Ο γεωμετρικός περιγραφέας FPFH και τα κάθετα διανύσματα σε σύγκριση με τους προηγούμενους περιγραφείς επιτυγχάνουν μέτρια ακρίβεια ευθυγράμμισης. Η λιγότερη ακριβής τεχνική σε αυτή την μελέτη είναι η μέθοδος που βασίζεται στο αλγόριθμο ICP η οποία σε σχέση με τους USC και 3DSC περιγραφείς έχει μέσο σφάλμα 6 φορές μεγαλύτερο.

Στο σύνολο των 17 σετ δεδομένων εκτός από 2 ζεύγη όπου οι απόκλιση είναι σχετικά μεγαλύτερη για τον 3DSC, στα υπόλοιπα ζεύγη δεδομένων τα αποτελέσματα μεταξύ των USC και 3DSC περιγραφέων είναι συγκρίσιμα. Το ίδιο συμπέρασμα προκύπτει και για τον SHOT που σε σύγκριση με τον 3DSC έχει μεγάλη απόκλιση σε 2 μόνο ζεύγη δεδομένων.

Εκτός από το σφάλμα της ευθυγράμμισης, για όλα τα σετ δεδομένων μετρήθηκε και ο συνολικός χρόνος εκτέλεσης της κάθε μεθόδου και οι τιμές καταχωρήθηκαν στον πίνακα 5.2. Στην τελευταία σειρά του πίνακα υπολογίζεται ο μέσος χρόνος εκτέλεσης για όλα τα ζεύγη νεφών σημείων ανά μέθοδο. Παρατηρείται ότι η ακρίβεια είναι αντιστρόφως ανάλογη του χρόνου εκτέλεσης της μεθόδου. Η πιο χρονοβόρες μέθοδοι είναι οι μέθοδοι με βάση τους περιγραφείς USC και 3DSC, ο μέσος χρόνος εκτέλεσης των οποίων είναι 1017.94 και 975.17 δευτερόλεπτα αντίστοιχα. Σημαντική μείωση του χρόνου εκτέλεσης παρουσιάζει ο περιγραφέας SHOT με μόλις 58.94 δευτερόλεπτα και περίπου 17 φορές μικρότερο από τους προηγούμενους 2 περιγραφείς. Ο αλγόριθμος ICP και τα κάθετα διανύσματα επιτυγχάνουν σημαντικά μικρότερο χρόνο εκτέλεσης με την λιγότερη όμως ακρίβεια.

Πίνακας 5.1. Μέσο τετραγωνικό σφάλμα (MSE) τεχνικών ευθυγράμμισης με χρήση των γεωμετρικών περιγραφών USC, 3DSC, SHOT, FPFH, κάθετων διανυσμάτων και του αλγορίθμου ICP

Ζεύγος δεδομένων	Πριν την ευθυγράμμιση	Μέθοδος Ευθυγράμμισης					
		USC	3DSC	SHOT	FPFH	Normals	ICP
1	0.002046790	0.000060465	0.000051941	0.000070307	0.000055118	0.000057669	0.000325544
2	0.005617340	0.000124052	0.000121823	0.000063904	0.000312984	0.000116751	0.000162806
3	0.003396220	0.000399475	0.000385993	0.000374704	0.000398547	0.000392308	0.000317698
4	0.000189146	0.000027664	0.000012741	0.000012554	0.000012696	0.000017615	0.000047507
5	0.016375100	0.000115599	0.000115002	0.000078045	0.000074511	0.000280176	0.006088490
6	0.004215360	0.000183521	0.000337454	0.000306088	0.000200741	0.000174630	0.000209687
7	0.001345050	0.000046568	0.000011631	0.000057324	0.000017620	0.000017147	0.000018319
8	0.000303232	0.000136033	0.000084188	0.000329966	0.000197336	0.000389069	0.000077145
9	0.000083936	0.000026206	0.000023998	0.000040425	0.000024222	0.000054384	0.000026563
10	0.000176236	0.000038655	0.000030550	0.000038420	0.000050489	0.000072870	0.000034589
11	0.001852380	0.000014917	0.000011100	0.000012165	0.000014304	0.000022224	0.000168650
12	0.000647387	0.000212916	0.000228935	0.000220150	0.000297704	0.000368276	0.000097761
13	0.001536650	0.000214722	0.000231198	0.000287196	0.000354224	0.001360290	0.000190054
14	0.000273208	0.000016105	0.000016514	0.000020195	0.000017838	0.000038091	0.000067943
15	0.001227220	0.000017864	0.000019814	0.000106221	0.003376130	0.000023755	0.000172219
16	0.002794280	0.000049192	0.000100081	0.000074644	0.000088929	0.003105780	0.001097860
17	0.002142420	0.000031055	0.000039932	0.000024349	0.000029751	0.000103443	0.000724610
	0.002601291	0.000100883	0.000107229	0.000124509	0.000324891	0.000387910	0.000578085
$\mu \pm \sigma$	\pm	\pm	\pm	\pm	\pm	\pm	\pm
	0.003875585	0.000103651	0.00011867	0.000125012	0.000797476	0.000771651	0.001447308

Πίνακας 5.2. Συνολικός χρόνος εκτέλεσης των τεχνικών ευθυγράμμισης με χρήση των γεωμετρικών περιγραφών USC, 3DSC, SHOT, FPFH, κάθετων διανυσμάτων και του αλγορίθμου ICP

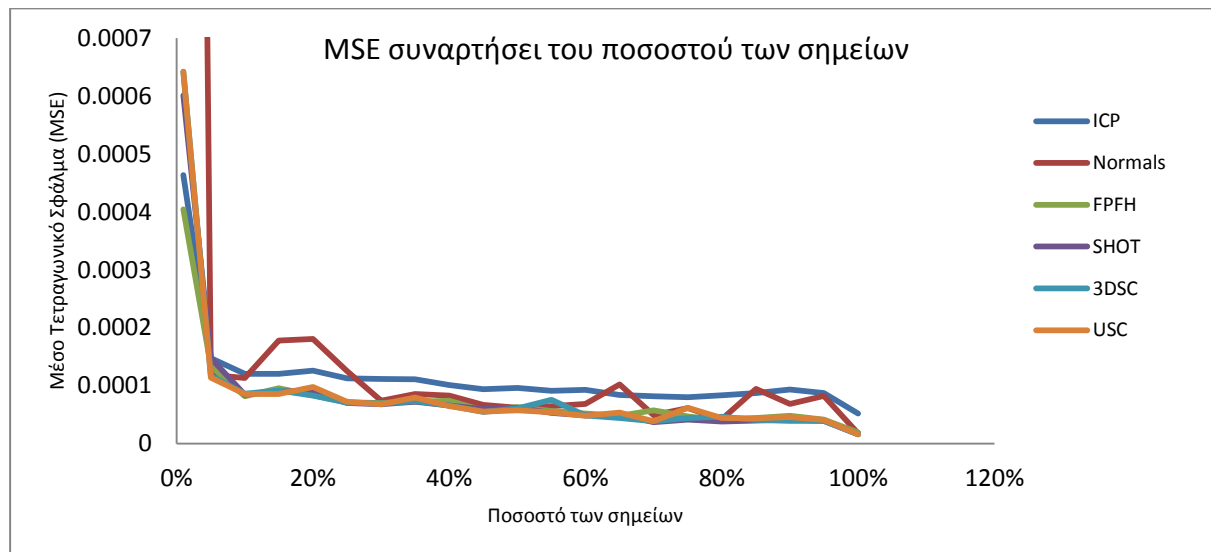
Ζεύγος δεδομένων	Μέθοδος Ευθυγράμμισης					
	USC	3DSC	SHOT	FPFH	Normals	ICP
1	2169	2110	84	26	11	7
2	778	574	38	10	3	1
3	3618	3663	353	41	17	13
4	865	890	38	11	4	3
5	1761	2030	137	29	6	2
6	129	107	12	4	1	1
7	95	106	10	3	1	1
8	148	117	14	4	1	1
9	44	51	5	2	1	1
10	42	47	5	2	1	1
11	2136	2107	66	25	11	8
12	88	98	7	3	1	1
13	84	94	8	3	1	1
14	1026	528	27	10	2	1
15	558	529	15	7	2	1
16	1820	1860	93	16	9	4
17	1944	1667	90	15	7	3
Μέσος Χρόνος	1017.94	975.17	58.94	12.41	4.64	2.94

6 Σχολιασμός παραμέτρων εισόδου

Στις προτεινόμενες μεθόδους ευθυγράμμισης καλείται ο χρήστης να εισάγει μερικές παραμέτρους ο προσδιορισμός των οποίων επηρεάζει σημαντικά την ακρίβεια της κάθε μεθόδου.

6.1 Συντελεστής υποδειγματοληψίας

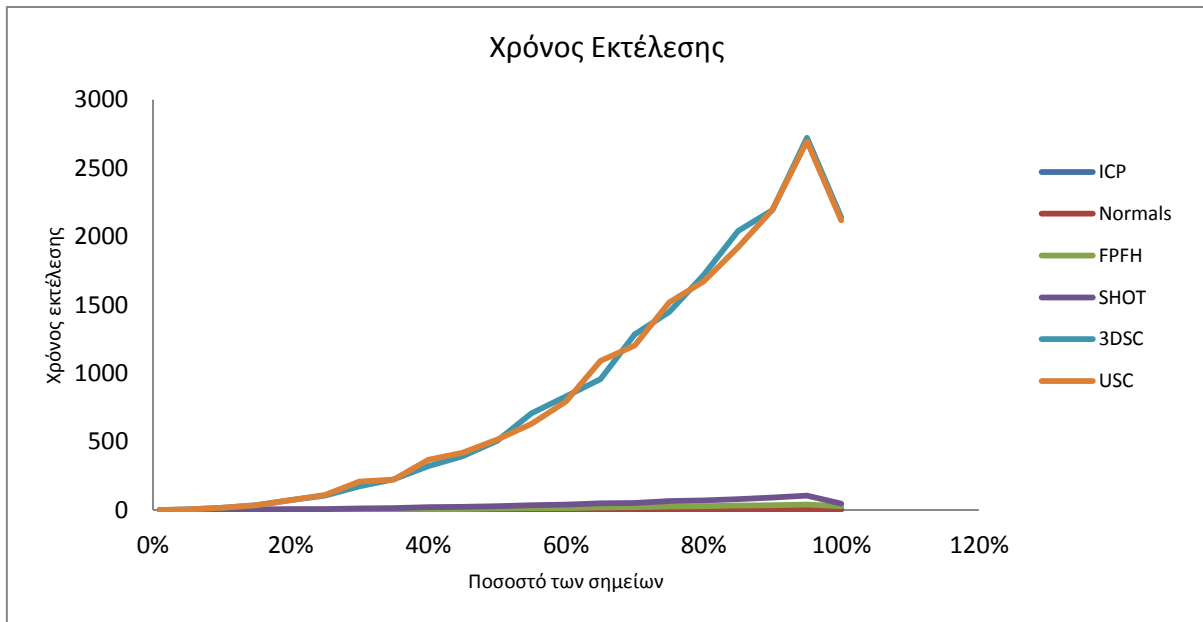
Η δειγματοληψία είναι ένα προαιρετικό βήμα. Ωστόσο, η εφαρμογή της σε δεδομένα εισόδου μεγάλου πλήθους σημείων μειώνει τον χρόνο εκτέλεσης και το υπολογιστικό κόστος. Ο δείκτης δειγματοληψίας προσδιορίζει το ποσοστό των σημείων του αρχικού νέφους σημείων που χρειάζονται για την επιτυχή ευθυγράμμιση.



Εικόνα 6.1. Μέσο τετραγωνικό σφάλμα MSE συναρτήσει του ποσοστού των σημείων

Στην εικόνα 6.1 παρουσιάζονται τα αποτελέσματα που ελήφθησαν από ένα σετ δεδομένων με την διατήρηση ενός α% των σημείων από τα δυο νέφη σημείων, αναφοράς και προς ευθυγράμμιση. Η παράμετρος α παίρνει τιμές στο εύρος 1% έως 100% με βήμα 5%. Παρατηρείται ότι με την αύξηση του ποσοστού των σημείων που χρησιμοποιούνται για τον προσδιορισμό του μετασχηματισμού, οι τιμές του μέσου τετραγωνικού σφάλματος μειώνονται φτάνοντας την μικρότερη τιμή όταν στα νέφη σημείων δεν εφαρμόζεται

δειγματοληψία. Όταν διατηρούνται λιγότερα από 5% των σημείων όλες οι μέθοδοι αποτυγχάνουν και οι τιμές του MSE είναι πολύ μεγάλες. Από την γραφική παράσταση παρατηρείται ότι τα κάθετα διανύσματα παρουσιάζουν μεγάλες διακυμάνσεις ενώ οι υπόλοιπες μέθοδοι είναι πιο σταθερές.



Εικόνα 6.2. Χρόνος εκτέλεσης των μεθόδων συναρτήσει του ποσοστού των σημείων

Όπως παρατηρήθηκε το σφάλμα της ευθυγράμμισης είναι μικρότερο για ποσοστό σημείων ίσο με 100%. Όμως, μεγαλύτερο πλήθος σημείων σημαίνει μεγαλύτερο υπολογιστικό κόστος και αυξημένος χρόνος εκτέλεσης. Πολλά από τα διαθέσιμα νέφη σημείων περιέχουν ένα σχετικά μεγάλο αριθμό σημείων το οποίο φτάνει έως 190000.

Στην εικόνα 6.2 παρουσιάζεται ο χρόνος εκτέλεσης των μεθόδων συναρτήσει του ποσοστού των σημείων που χρησιμοποιήθηκαν. Παρατηρείται ότι οι δυο περιγραφείς, 3DSC και USC, σε σύγκριση με τις υπόλοιπες μεθόδους, έχουν σημαντικά μεγαλύτερο χρόνο εκτέλεσης για το ίδιο ποσοστό των σημείων. Γενικά, για όλες τις μεθόδους ο χρόνος εκτέλεσης αυξάνεται με την αύξηση των σημείων εφόσον περισσότεροι υπολογισμοί πρέπει να πραγματοποιηθούν για τον ορισμό των περιγραφέων. Επομένως η βέλτιστη τιμή της παραμέτρου α θα πρέπει να είναι η χρυσή τομή μεταξύ ακριβείας και χρόνου εκτέλεσης. Θα πρέπει να σημειωθεί ότι τα αποτελέσματα αφορούν μόνο ένα σετ δεδομένων. Επειδή οι συνθήκες στις οποίες τα δεδομένα ελήφθησαν ήταν οι ίδιες, θεωρείται πως και τα υπόλοιπα σετ θα έχουν την ίδια συμπεριφορά. Επειδή τα κάθετα διανύσματα στην περιοχή 10%-30% παρουσιάζουν μεγάλες τιμές του MSE σε σύγκριση με άλλες μεθόδους και επειδή όλες οι άλλες μέθοδοι εκτός από ICP και USC βασίζονται στα κάθετα διανύσματα για τον υπολογισμό των περιγραφέων τους, επιλέχθηκε ως το καλύτερο και μικρότερο δυνατό ποσοστό σημείων για την επιτυχή ευθυγράμμιση χωρίς ο χρόνος εκτέλεσης να είναι απαγορευτικός το 40% των σημείων.

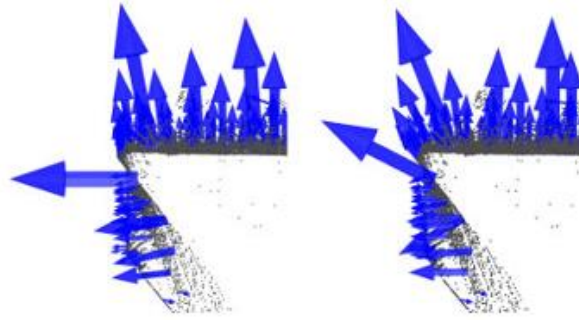
6.2 Ακτίνα γειτονικής περιοχής κάθετων διανυσμάτων

Μια άλλη σημαντική παράμετρος είναι η ακτίνα γειτονικής περιοχής για τα κάθετα διανύσματα. Η παράμετρος αυτή υπολογίζεται σε μέτρα (m) και αντιπροσωπεύει την ακτίνα μιας σφαίρας με κέντρο το σημείο για το οποίο επιθυμείται ο υπολογισμός του κάθετου διανύσματος. Για 10 σετ δεδομένων υπολογίστηκε το μέσο τετραγωνικό σφάλμα για τιμές της ακτίνας από 0.01m έως 0.4m με βήμα 0.01m. Τα αποτελέσματα της εικόνας 6.3 αφορούν την μέση τιμή των 10 μετρήσεων για την κάθε ακτίνα.



Εικόνα 6.3. Μέσο τετραγωνικό σφάλμα MSE κάθετων διανυσμάτων συναρτήσει της βοηθητικής ακτίνας

Η ακτίνα της γειτονικής περιοχής των κάθετων διανυσμάτων είναι πολύ ευαίσθητη παράμετρος και θα πρέπει να είναι τέτοια ώστε η περιοχή να έχει αρκετά σημεία για τον υπολογισμό των κάθετων διανυσμάτων αλλά ταυτόχρονα όχι πολύ μεγάλη διότι χάνεται η λεπτομέρεια και οι μικρές μεταβολές της επιφάνειας κάτι το οποίο είναι ακριβώς αυτό που θέλουμε να κωδικοποιήσουμε. Ένα τέτοιο παράδειγμα επιλογής μιας μικρής και μεγάλης ακτίνας παρουσιάζεται στην εικόνα 6.4. Η αριστερή εικόνα απεικονίζει μια σωστή επιλογή της γειτονικής ακτίνας και τα κάθετα διανύσματα είναι σχεδόν κάθετα στις επιφάνειες. Στην δεξιά εικόνα η τιμή της ακτίνας είναι αρκετά μεγαλύτερη. Για τον υπολογισμό του κάθετου διανύσματος ενός σημείου συμμετέχουν και άλλα σημεία τα οποία βρίσκονται στις γειτονικές επιφάνειες με αποτέλεσμα τα κάθετα διανύσματα να παραμορφώνονται σε αυτή την περιοχή και στην πραγματικότητα να μην είναι κάθετα στην επιφάνεια την οποία αντιπροσωπεύουν.

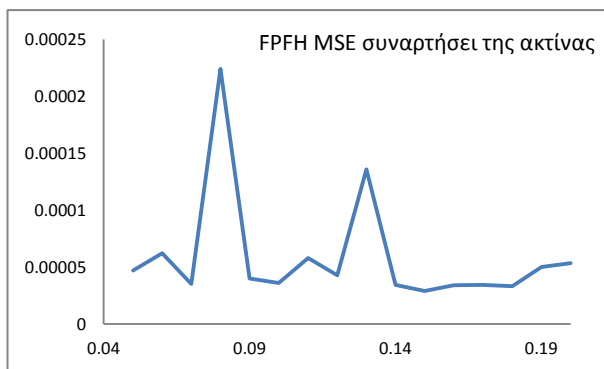


Εικόνα 6.4. Κάθετα διανύσματα. Αριστερά: μικρή ακτίνα γειτονικής περιοχής. Δεξιά: μεγάλη ακτίνα της γειτονικής περιοχής

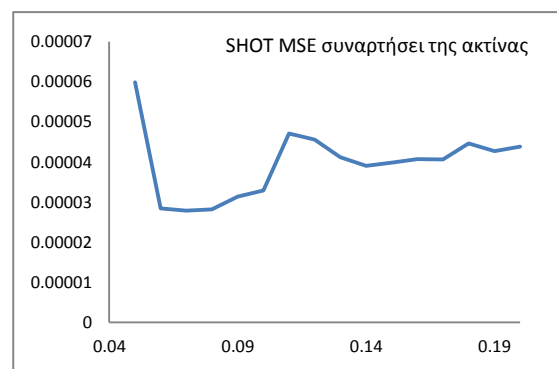
Παρατηρείται ότι με την αύξηση της ακτίνας η μέση τιμή του MSE αυξάνεται παρά τις μεγάλες διακυμάνσεις. Η μικρότερη τιμή του MSE προέκυψε για ακτίνα 0.04m

6.3 Ακτίνα γειτονικής περιοχής γεωμετρικών περιγραφών

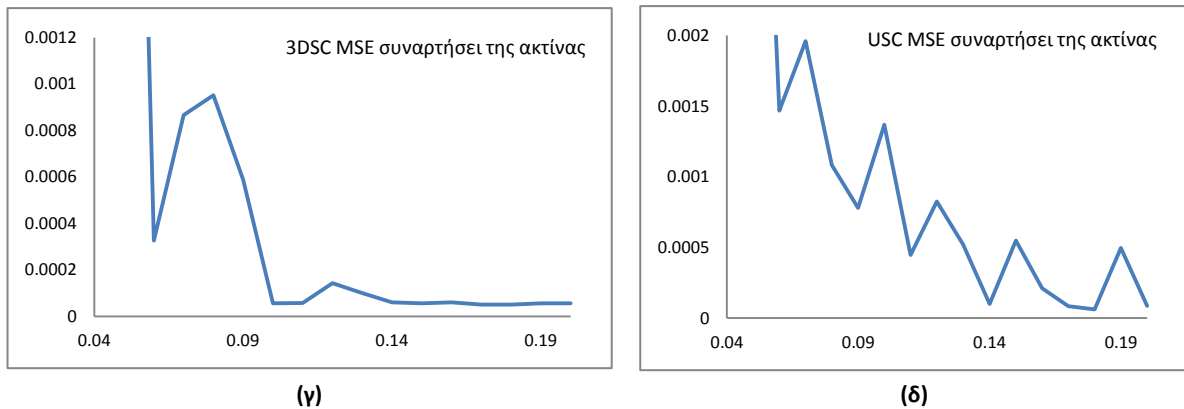
Ο υπολογισμός των γεωμετρικών περιγραφών απαιτεί επίσης τον προσδιορισμό μιας γειτονικής περιοχής ακτίνας r . Το μέγεθος αυτό αντιστοιχεί στην ακτίνα μιας σφαίρας με κέντρο το σημείο για το οποίο υπολογίζεται ο γεωμετρικός περιγραφέας. Μόνο τα σημεία που βρίσκονται εντός της σφαίρας συμμετέχουν στον υπολογισμό του. Η σωστή επιλογή της ακτίνας συνδέεται άμεσα με το σφάλμα της ευθυγράμμισης. Ο κάθε γεωμετρικός περιγραφέας έχει μια διαφορετική ακτίνα r για την οποία το MSE είναι το ελάχιστο δυνατό. Σε ένα σύνολο 5 σετ δεδομένων υπολογίστηκε το MSE για τιμές της ακτίνας r από 0.04m έως 0.2m με βήμα 0.01. Η μέση τιμή του μέσου τετραγωνικού σφάλματος αυτών των 5 μετρήσεων παρουσιάζεται για τους περιγραφείς FPFH, SHOT, 3DSC και USC στην εικόνα 6.5. Οι τρεις πρώτοι περιγραφείς χρησιμοποιούν τα κάθετα διανύσματα για τον υπολογισμό τους. Ο USC δεν χρησιμοποιεί κάθετα διανύσματα αλλά όπως και ο SHOT υπολογίζει για κάθε σημείο ένα τοπικό σύστημα συντεταγμένων για την σωστή τοποθέτηση του σφαιρικού πλέγματος στο εξεταζόμενο σημείο.



(α)



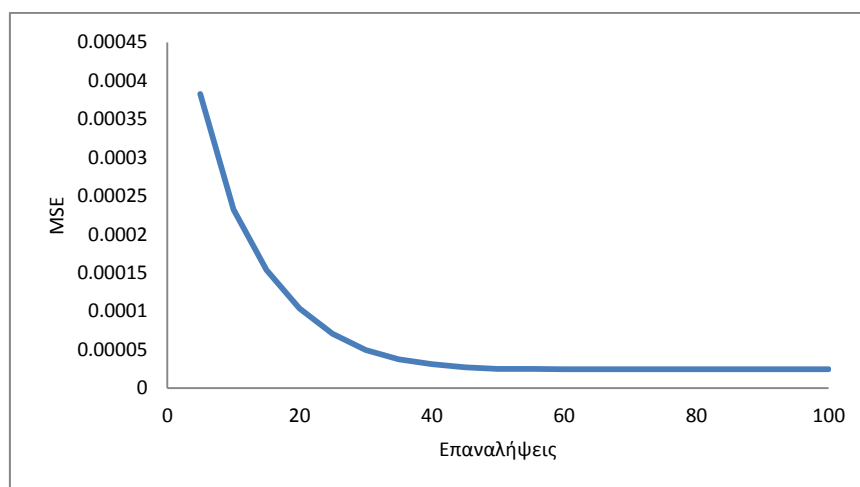
(β)



Εικόνα 6.5. Μέσο τετραγωνικό σφάλμα MSE συναρτήσει της ακτίνας της βοηθητικής περιοχής για διάφορους γεωμετρικούς περιγραφείς (α) FPFH (β) SHOT (γ) 3DSC (δ) USC

Για τις γραφικές αυτές παραστάσεις χρησιμοποιήθηκαν μόνο 5 σετ δεδομένων από 17 διαθέσιμα. Επειδή όπως περιγράφεται στην ενότητα 4.2 τα δεδομένα ελήφθησαν από το ίδιο απεικονιστικό σύστημα στις ίδιες συνθήκες φωτισμού θεωρούμε πως και στα υπόλοιπα σετ δεδομένων η ακτίνα r της γειτονικής περιοχής θα έχει την ίδια συμπεριφορά. Πιο συγκεκριμένα οι βέλτιστες τιμές της συνάρτησης σφάλματος ελήφθησαν για τιμές τις ακτίνας $r=0.15$ για τον περιγραφέα FPFH, $r=0.07$ για τον περιγραφέα SHOT και 0.18 για τους περιγραφείς 3DSC και USC.

Γενικά, η επιλογή μιας μικρής ακτίνας r έχει ως αποτέλεσμα η γειτονική περιοχή να περιέχει λίγα σημεία, ο αριθμός των οποίων δεν επαρκεί για τον υπολογισμό του περιγραφέα. Μια μεγάλη ακτίνα της γειτονικής περιοχής όμως μπορεί να εισάγει θόρυβο, δηλαδή σημεία τα οποία δεν είναι σημεία του αντικειμένου και δεν υπάρχουν στο άλλο νέφος σημείων με αποτέλεσμα δυο αντίστοιχα σημεία να έχουν διαφορετικούς περιγραφείς. Η τιμή της ακτίνας r πάντα εξαρτάται από τα δεδομένα και θα πρέπει να προσδιοριστεί ανάλογα από τις πειραματικές μετρήσεις.



Εικόνα 6.6. Μέσο τετραγωνικό σφάλμα MSE συναρτήσει του αριθμού των επαναλήψεων για τον αλγόριθμο ICP

Μια σημαντική παράμετρος στην οποία βασίζεται ο αλγόριθμος ICP είναι ο αριθμός των επαναλήψεων. Τα αποτελέσματα για την μέθοδο αυτή στο κεφάλαιο των αποτελεσμάτων πρόεκυψαν με την προεπιλεγμένη τιμή της παραμέτρου αυτής, η οποία στην βιβλιοθήκη PCL έχει την τιμή 10. Παρατηρήθηκε ότι με την αύξηση των αριθμών των επαναλήψεων ο αλγόριθμος ICP βελτιώνει αρκετά την ευθυγράμμιση. Η μέθοδος εφαρμόστηκε σε 5 σετ δεδομένων με αριθμό επαναλήψεων από 10 έως 100 και βήμα 5. Η μέση τιμή των 5 μετρήσεων υπολογίστηκε για κάθε τιμή του αριθμού των επαναλήψεων και τα αποτελέσματα παρουσιάζονται στην εικόνα 6.6. Για αριθμό επαναλήψεων 40 όπου και η συνάρτηση σφάλματος φτάνει την μικρότερη τιμή, η απόδοση της μεθόδου βελτιώνεται μέχρι και 7 φορές. Παρόλα αυτά η τεχνική αυτή εξαρτάται από τα δεδομένα. Καθώς η διαφορά μεταξύ των νεφών σημείων αυξάνει η πιθανότητα ο αλγόριθμος ICP να αποτύχει μεγαλώνει.

7 Συμπεράσματα, Συζήτηση και Μελλοντικές Βελτιώσεις και Επεκτάσεις

7.1 Συμπεράσματα της διπλωματικής εργασίας

Η παρούσα διπλωματική εργασία είχε ως αντικείμενο την ανάπτυξη και την παρουσίαση τεχνικών ευθυγράμμισης τρισδιάστατων νεφών σημείων που βασίζονται σε ζεύγη αντίστοιχων σημείων.

Αρχικά πραγματοποιήθηκε μια εισαγωγή στις τεχνικές ευθυγράμμισης που συναντούνται στην βιβλιογραφία, έγινε μια ταξινόμηση με βάση ένα σύνολο κριτηρίων και παρουσιάστηκε ο τρόπος λήψης των δεδομένων εισόδου και ο αισθητήρας με τον οποίον αποκτούνται.

Στις μεθόδους αυτές ο προσδιορισμός των αντίστοιχων ζευγών σημείων μεταξύ των δυο συνόλων δεδομένων έγινε συγκρίνοντας τα σημεία με βάση τους γεωμετρικούς περιγραφείς. Παρουσιάστηκαν ο τρόπος υπολογισμού των γεωμετρικών περιγραφέων και τα βήματα της πλήρους μεθοδολογίας. Αυτά είναι η υποδειγματοληψία, ο υπολογισμός των γεωμετρικών περιγραφέων, το οποίο είναι το μοναδικό βήμα που αλλάζει σε κάθε μεθοδολογία, ο προσδιορισμός των αντίστοιχων σημείων με βάση τους γεωμετρικούς περιγραφείς, η απόρριψη των λανθασμένων αντίστοιχων σημείων και ο υπολογισμός και η εφαρμογή του καθολικού γεωμετρικού μετασχηματισμού. Με σκοπό την σύγκριση των τεχνικών αυτών, παρουσιάστηκε συνοπτικά και μια μεθοδολογία ευθυγράμμισης βασισμένης στον αλγόριθμο ICP.

Οι προτεινόμενες τεχνικές ευθυγράμμισης εφαρμόστηκαν σε ένα σύνολο διαθέσιμων τρισδιάστατων νεφών σημείων τα οποία ελήφθησαν με τον αισθητήρα Microsoft Kinect. Η αποδοτικότητα και η ακρίβεια τους αξιολογήθηκε τόσο ποιοτικά όσο και ποσοτικά. Οι μέθοδοι με βάση τους γεωμετρικούς περιγραφείς συγκρίθηκαν μεταξύ τους καθώς και με την μέθοδο ευθυγράμμισης που βασίζεται στον αλγόριθμο ICP, όσον αφορά την ακρίβεια της κάθε μεθόδου και τον χρόνο εκτέλεσης. Η πιο ακριβείς μέθοδοι προέκυψαν οι μέθοδοι με βάση τους γεωμετρικούς περιγραφείς 3DSC και USC οι οποίες ταυτόχρονα ήταν και οι πιο χρονοβόρες. Πιο γρήγορες μέθοδοι, όπως οι μέθοδοι με βάση τον περιγραφέα FPFH και τα κάθετα διανύσματα, παρουσίασαν μικρότερη ακρίβεια σε σχέση με τις προηγούμενες. Την καλύτερη αναλογία της ακρίβειας και του χρόνου εκτέλεσης παρουσίασε ο περιγραφέας SHOT. Τέλος σχολιάστηκε ο καθορισμός των παραμέτρων για τις οποίες η

κάθε μέθοδος επιτυγχάνει ευθυγράμμιση με την ελάχιστη απόκλιση μεταξύ των δυο απεικονίσεων.

7.2 Συζήτηση

Η υλοποίηση των μεθόδων ευθυγράμμισης με βάση τους γεωμετρικούς περιγραφείς καθώς και της μεθόδου βασισμένης στον αλγόριθμο ICP βασίστηκε στην βιβλιοθήκη PCL (Point Cloud Library). Για την αξιολόγηση των μεθόδων αυτών αναπτύχθηκαν κατάλληλες εφαρμογές σε γλώσσα C++ για τους διάφορους γεωμετρικούς περιγραφείς και τον αλγόριθμο ICP, οι οποίες στην συνέχεια εφαρμόστηκαν σε 17 σύνολα δεδομένων με σκοπό την ευθυγράμμιση τους. Όλες οι εφαρμογές εκτελέστηκαν στον ίδιο Η/Υ (Intel Core i3 4000M, 2.4 GHz με μνήμη RAM 4GB και λειτουργικό σύστημα Linux 64-bit).

Από την εξέταση των ποιοτικών και ποσοτικών αποτελεσμάτων συμπεραίνεται πως την καλύτερη ακρίβεια από όλες τις μεθόδους επιτυγχάνει η μέθοδος με βάση τον γεωμετρικό περιγραφέα USC και με σχεδόν ίδια ακρίβεια η μέθοδος με βάση τον 3DSC. Ταυτόχρονα οι τεχνικές αυτές είναι και οι πιο χρονοβόρες.

Η μεγαλύτερη καθυστέρηση στις μεθόδους με βάση τους 3DSC και USC περιγραφείς παρατηρείται στο βήμα υπολογισμού του περιγραφέα λόγω του μεγάλου αριθμού των τομών του σφαιρικού πλέγματος και των υπολογισμών που πραγματοποιούνται για μια από αυτές καθώς και στο βήμα εύρεσης αντίστοιχων σημείων όπου η μεγάλη διάσταση του περιγραφέα ακόμα και με την μέθοδο kNN είναι μια χρονοβόρα διαδικασία. Για μικρές τιμές ακτίνας της γειτονικής περιοχής οι 3DSC, USC αποτυγχάνουν. Επομένως για τους συγκεκριμένους περιγραφείς χρησιμοποιείται μεγαλύτερη ακτίνα που έχει σαν αποτέλεσμα σημαντική αύξηση του χρόνου εκτέλεσης.

Στα δεδομένα που χρησιμοποιήθηκαν εφαρμόστηκε αρχικά ένα φίλτρο απόρριψης σημείων (voxel grid filter) με σκοπό την υποδειγματοληψία τους. Αρχικά τα νέφη σημείων είχαν από 6000 έως 190000 σημεία. Μετά την δειγματοληψία ο μέγιστος αριθμός σημείων ενός νέφους ήταν 25000. Για ένα τέτοιο αριθμό σημείων οι 3DSC και USC χρειάστηκαν για ακτίνα γειτονικής περιοχής 0.18m, 3600 περίπου δευτερόλεπτα ή μια ώρα. Αυτός ο χρόνος εκτέλεσης είναι σημαντικά μεγάλος και αυξάνεται αν ο αριθμός των σημείων αυξάνει.

Ο μεγάλος χρόνος εκτέλεσης τους απαγορεύει την χρήση τους σε εφαρμογές πραγματικού χρόνου (real time applications). Η ευθυγράμμιση βασισμένη στα σημεία σε αυτή την περίπτωση χάνει την ουσία της, καθώς ο χρόνος εκτέλεσης είναι συγκρίσιμος με αυτόν των τεχνικών ευθυγράμμισης που χρησιμοποιούν τις τιμές φωτεινότητας των σημείων, οι οποίες και συνήθως επιτυγχάνουν μεγαλύτερη ακρίβεια. Επομένως οι δυο αυτοί περιγραφείς μπορούν τελικά να χρησιμοποιηθούν σε τεχνικές ευθυγράμμισης όπου τα σύνολα δεδομένων περιέχουν λίγα σημεία και όπου ο χρόνος εκτέλεσης δεν είναι τόσο σημαντικός παράγοντας όσο η ακρίβεια.

Από την άλλη, οι τεχνικές που έχουν μικρό χρόνο εκτέλεσης όπως FPFH, κάθετα διανύσματα και ο αλγόριθμος ICP, δεν έχουν μεγάλη ακρίβεια και μπορούν να χρησιμοποιηθούν μόνο σε εφαρμογές όπου απαιτείται γρήγορη ευθυγράμμιση χωρίς η μεγάλη ακρίβεια να είναι αναγκαία. Η καλύτερη λύση ίσως είναι ο γεωμετρικός περιγραφέας SHOT ο οποίος παρουσιάζει αρκετά μικρό σφάλμα ευθυγράμμισης και αποδεκτό χρόνο εκτέλεσης. Σε

σύγκριση με τους 3DSC και USC η ακρίβεια του είναι μικρότερη κατά 1.2 φορές περίπου αλλά ο χρόνος εκτέλεσης μειώνεται μέχρι και 17 φορές.

Επομένως, η επιλογή του γεωμετρικού περιγραφέα ή ακόμα της μεθόδου με βάση τον αλγόριθμο ICP εξαρτάται πάντα από την περίπτωση, το πρόβλημα της ευθυγράμμισης και ακόμα από την προσωπική κρίση.

Για τις μεθόδους που παρουσιάζονται σε αυτή την διπλωματική εργασία θα μπορούσαμε να πούμε πως τη χρυσή τομή μεταξύ ακριβείας και χρόνου εκτέλεσης αποτελεί ο περιγραφέας SHOT.

7.3 Μελλοντικές βελτιώσεις και επεκτάσεις

Αν και στις προτεινόμενες μεθόδους στα νέφη σημείων πραγματοποιείται υποδειγματοληψία με σκοπό τον περιορισμό του πλήθους των σημείων και την μείωση του υπολογιστικού κόστους ο αριθμός των σημείων που απομένουν παραμένει μεγάλος. Στην παρούσα εργασία διατηρούνται 40% των αρχικών σημείων σε κάθε νέφος και για το κάθε σημείο υπολογίζεται ένας γεωμετρικός περιγραφέας. Η ταχύτητα αυτών των μεθόδων μπορεί να βελτιωθεί σημαντικά αν αντί για υποδειγματοληψία εξαχθεί από το κάθε νέφος ένα μικρό υποσύνολο των σημείων που ονομάζονται σημεία κλειδιά (keypoints) και υπολογιστούν οι γεωμετρικοί περιγραφείς μόνο για αυτό το περιορισμένο υποσύνολο. Τα σημεία κλειδιά είναι χαρακτηριστικά γεωμετρικά σημεία τα οποία βρίσκονται στην θέση ενός βέλτιστου γεωμετρικού χαρακτηριστικού όπως ακρότατα, γωνίες, συνοριακά σημεία κλπ. Η βιβλιοθήκη PCL περιέχει αυτόματες μεθόδους για τον εντοπισμό των σημείων κλειδιών όπως οι μέθοδοι NARF και ISS [42] - [43]. Ο προσδιορισμός των αντίστοιχων σημείων συνεπάγεται στην εύρεση αντίστοιχων σημείων κλειδιών με σύγκριση των γεωμετρικών περιγραφέων τους.

Μια τέτοια προσέγγιση μπορεί να εφαρμοστεί στις ακριβείς αλλά χρονοβόρες μεθόδους όπως οι μέθοδοι με βάση τους γεωμετρικούς περιγραφείς 3DSC κα USC ή SHOT.

Οι μέθοδοι ευθυγράμμισης με βάση τους γεωμετρικούς περιγραφείς FPFH και κάθετα διανύσματα που παρουσίασαν μικρότερη ακρίβεια μπορούν να χρησιμοποιηθούν σε συνδυασμό με την μέθοδο με βάση τον αλγόριθμο ICP. Έτσι, αρχικά εφαρμόζεται ευθυγράμμιση με βάση τους περιγραφείς FPFH ή τα κάθετα διανύσματα και στην συνέχεια η μέθοδος με βάση τον αλγόριθμο ICP για μια πιο ακριβής ευθυγράμμιση.

Βιβλιογραφία

- [1] B. Zitova και J. Flusser, «Image registration methods: a survey,» *ELSEVIER Image and Vision Computing*, τόμ. 21, pp. 977-1000, 2003.
- [2] J. Maintz και M. Viergever, «A survey of medical image registration,» *Elsevier Medical Image Analysis*, τόμ. 2, αρ. 1, pp. 1-36, 1998.
- [3] M. V. Wyawahare, D. P. M. Patil και H. K. Abhyankar, «Image Registration Techniques: an overview,» *International Journal of Signal Processing, Image Processing and Pattern Recognition*, τόμ. 2, αρ. 3, 2009.
- [4] C. Maurer και J. Fitzpatrick, «A review of medical image registration,» *Interactive Image-Guided Neurosurgery, American Association of Neurological Surgeons*, pp. 17-44, 1993.
- [5] P. Van den Elsen, E. Pol και M. Viergever, «Medical image matching - a review with classification,» *IEEE Engineering in Medicine and Biology Magazine*, τόμ. 12, αρ. 1, pp. 26-39.
- [6] P. Thevenaz, U. Ruttimann και M. Unser, «A pyramid approach to subpixel registration based on intensity,» *IEEE Trans Image Process*, τόμ. 7, αρ. 1, pp. 27-41, 1998.
- [7] C.-T. Hsu και R. Beuker, «Multiresolution feature-based image registration,» *Procs SPIE: Visual Communications and Image Processing*, τόμ. 4067, pp. 1490-1498, 2000.
- [8] A. N. Ουζούνογλου, «Αυτόματη Ταύτιση Βιομετρικών Χαρακτηριστικών: Εφαρμογή στα Δακτυλικά Αποτυπώματα,» 2011.
- [9] V. Mandava, J. Fitzpatrick, C. Maurer, R. Maciunas και G. Allen, «Registration of multimodal volume head images via attached markers,» *Procs SPIE: Medical Imaging VI: Image Processing*, τόμ. 1652, pp. 271-282, 1992.
- [10] J. West, M. Wang, B. Dawant και J. Fitzpatrick, «Comparison and evaluation of retrospective intermodality brain image registration techniques,» *Computational Assisted Tomography*, τόμ. 21, αρ. 4, pp. 554-566, 1997.
- [11] K. Gall και L. Verhey, «Computer-assisted positioning of radiotherapy patients using implanted radioopaque fiducials,» *Medical physics*, τόμ. 20, αρ. 4, pp. 1152-1159, 1993.
- [12] L. Lunsford, *Modern stereostatic neurosurgery*, Boston: Martinus Nijhoff, 1988.
- [13] A. Evans, S. Marrett, J. Torrescorzo, S. Ku και L. Collins, «MRI-PET correlation in three dimensions using a volume of interest (VOI) atlas,» *Journal of cerebral blood flow and metabolism*, τόμ. 11, pp. A69-A78, 1991.
- [14] L. Laitinen, B. Liliequist, M. Fagerlund και A. Eriksson, «An adapter for computer tomography guided stereotaxis,» *Surgical neurology*, τόμ. 23, pp. 559-566, 1985.
- [15] [Ηλεκτρονικό]. Available: <http://www.ariser.info/training/imgproc.php>. [Πρόσβαση 28

March 2015].

- [16] A. Evans, S. Marrett, L. Collins και T. Peters, «Anatomical-functional correlative analysis of the human brain using three dimensional imaging systems,» *Medical imaging: image processing*, τόμ. 1092, pp. 264-274, 1989.
- [17] R. Wahl, L. Quint, R. Cieslak, A. Aisen, R. Koeppe και C. Meyer, «Anamatabolic tumor maging: fusion of FDG PET with CT or MRI to localize foci of increased activity,» *Journal of nuclear medicine*, τόμ. 34, pp. 1190-1197, 1993.
- [18] C. Chen, C. Pelizzari, G. Chen, M. Cooper και D. Levin, «Image analysis of PET data with the aid of CT and MR images,» *Information processing in medical imaging*, pp. 601-611, 1987.
- [19] R. Bajcsy, R. Lieberson και M. Reivich, «A computerized system for the elastic matching of deformed radiographic images to idealized atlas images,» *Journal of computer assisted tomography*, τόμ. 7, αρ. 4, pp. 618-625, 1983.
- [20] C. Chen, C. Pelizzari, G. Chen, M. Cooper και D. Levin, «Image analysis of PET data with the aid of CT and MR images,» *Information processing in medical imaging*, pp. 601-611, 1987.
- [21] G. Borgefors, «Hierarchical chamfer matching: a parametric edge matching algorithm,» *IEEE Transactions on pattern analysis and machine intelligence*, τόμ. 10, pp. 849-865, 1988.
- [22] P. Banerjee και A. Toga, «Image alignment by integrated rotational and translational matrix,» *Physics in medicine and biology*, τόμ. 39, pp. 1969-1988, 1994.
- [23] G. Ettinger, W. Grimson και T. Lozano-Perez, «Automatic 3D image registration for medical change detection applications,» *Applications of computer vision in medical image processing*, pp. 182-185, 1994.
- [24] W. Press, B. Flannery, S. Teukolsky και W. Vetterling, «Numerical recipes in C,» *Cambridge University Press*, 1992.
- [25] S. Kirkpatrick, C. Gelatt και M. Vecchi, «Optimization by simulated annealing,» *Science*, τόμ. 220, αρ. 4598, pp. 671-680, 1983.
- [26] S. Jacoby, J. Kowalik και J. Pizzo, «Iterative methods for nonlinear optimization problems.,» Englewood Cliffs, NJ: Prentice Hall, 1972.
- [27] M. F. Vandermeulen και D. Suetens, «Comparative evaluation of multiresolution optimization strategies for multimodality image registration by maximization of mutual information,» *Med Image Anal*, τόμ. 3, αρ. 4, pp. 373-386, 1999.
- [28] T. Radcliffe, R. Rajapekshe και S. Shaler, «Pseudocorrelation: A fast, robust, absolute, gray level image alignment algorithms,» *Med Phys*, pp. 761-769, 1994.
- [29] [Ηλεκτρονικό]. Available: robotica.unileon.es. [Πρόσβαση 28 March 2015].
- [30] M. Hansard, S. Lee, O. Choi και R. Horaud, «Time of flight cameras: Principles, methods and applications,» *SpringerBriefs in Computer Science*, p. 95, 2012.
- [31] M. Rodrigues, M. Kormann, C. Schuhler και P. Tomek, «Structured light techniques for 3D surface reconstruction in robotic tasks».
- [32] G. Gerig, «Structured Lighting,» *3D Computer Vision*, 2012.
- [33] K. Khoshelham και S. O. Elberink, «Accuracy and Resolution of Kinect Depth Data for Indoor Mapping Applications,» *Sensors*, τόμ. 12, pp. 1437-1454, 2012.

- [34] R. B. Rusu, «Semantic 3D object maps for everyday manipulation in human living environments,» Munchen, 2009.
- [35] R. B. Rusu, Z. C. Marton, N. Blodow και M. Beetz, «Persistent Point Feature Histograms for 3D Point Clouds».
- [36] F. Tombari, S. Salti και L. Di Stefano, «Unique Signatures of Histograms for Local Surface Description,» Bologna, Italy.
- [37] R. Hansch, T. Weber και O. Hellwich, «Comparison of 3D interest point detectors and descriptors for point cloud fusion,» *ISPRS Annals of the Photometry, Remote Sensing and Spatial Information Sciences*, τόμ. 2, αρ. 3, 2014.
- [38] S. Belongie, J. Malik και J. Puzicha, «Shape matching and object recognition using shape contexts,» *IEEE Trans. on Pattern Analysis and Machine Intelligence*, τόμ. 24, αρ. 4, pp. 509-522, 2002.
- [39] A. Frome, D. Huber, R. Kolluri, T. Bulow και J. Malik, «Recognizing Objects in Range Data Using Regional Point Descriptors».
- [40] F. Tombari, S. Salti και L. Di Stefano, «Unique Shape Context for 3D Data Description».
- [41] P. J. Besl και N. D. McKay, «A Method for Registration of 3D Shapes,» *IEEE Transactions on pattern analysis and machine intelligence*, τόμ. 14, αρ. 2, 1992.
- [42] B. Steder, R. B. Rusu, K. Konolige και W. Burgard, «Point feature extraction on 3D range scans taking into account object boundaries».
- [43] Y. Zhong, «Intrinsic shape signatures: A shape descriptor for 3D object recognition,» *IEEE*, pp. 689-696, 2009.

