



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΣΥΣΤΗΜΑΤΩΝ ΜΕΤΑΔΟΣΗΣ ΠΛΗΡΦΟΡΙΑΣ
ΚΑΙ ΤΕΧΝΟΛΟΓΙΑ ΥΛΙΚΩΝ

Ανάπτυξη Συστήματος Ευθυγράμμισης Τρισδιάστατων Επιφανειών Με Χρήση Αισθητήρα KINECT

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Χριστόδουλος Α. Σκουρουμούνης

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

Αθήνα, Μάιος 2014



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΣΥΣΤΗΜΑΤΩΝ ΜΕΤΑΔΟΣΗΣ ΠΛΗΡΦΟΡΙΑΣ
ΚΑΙ ΤΕΧΝΟΛΟΓΙΑ ΥΛΙΚΩΝ

Ανάπτυξη Συστήματος Ευθυγράμμισης Τρισδιάστατων Επιφανειών Με Χρήση Αισθητήρα KINECT

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Χριστόδουλος Α. Σκουρουμούνης

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

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 16η Μαΐου 2014.

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

.....
Ν. Ουζούνογλου
Καθηγητής Ε.Μ.Π.

.....
Δ. Κουτσούρης
Καθηγητής Ε.Μ.Π.

Αθήνα, Μάιος 2014

.....

Χριστόδουλος Α. Σκουρουμούνης

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

Copyright © Χριστόδουλος Α. Σκουρουμούνης, 2014

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

Σκοπός της παρούσας διπλωματικής εργασίας είναι η παρουσίαση μιας μεθόδου αυτόματης ευθυγράμμισης τρισδιάστατων ιατρικών δεδομένων. Η διαδικασία της ευθυγράμμισης έχει ως στόχο την διόρθωση των σχετικών μετατοπίσεων στο χώρο έτσι ώστε στα ευθυγραμμισμένα δεδομένα οι ίδιες ανατομικές περιοχές να συμπίπτουν χωρικά.

Η παρούσα μέθοδος μπορεί να χρησιμοποιηθεί για την ευθυγράμμιση τρισδιάστατων δεδομένων τα οποία λαμβάνονται από τον αισθητήρα KINECT. Μια τέτοια ευθυγράμμιση μπορεί να εφαρμοστεί και στον τομέα της ιατρικής. Στην ακτινοθεραπεία, πριν από την κάθε θεραπεία, ο ασθενής θα πρέπει να ευθυγραμμιστεί με την τομογραφία δεδομένων (CT). Για την ευθυγράμμιση αυτή χρειάζεται ένα σύστημα επαλήθευσης του προσανατολισμού του ασθενή για τον ακριβή προσδιορισμό του ασθενή και την προσαρμογή του θεραπευτικού κρεβατιού για την βελτιστοποίηση τη θεραπείας. Το παρόν σύστημα υλοποίησης αυτής της διαδικασίας απαιτεί χειροκίνητη μετακίνηση του θεραπευτικού κρεβατιού με χρήση laser και προσδιορισμός ορόσημων σημείων στο δέρμα του ασθενή. Η ανάγκη για ευθυγράμμιση προκύπτει από την χωρική μετατόπιση μεταξύ των εικόνων καθώς και στην ανάπτυξη ανωμαλιών που προκαλούν παραμορφώσεις (π.χ. κύστες, όγκοι)

Η παρούσα διπλωματική προτείνει μια μέθοδο χωρίς καθορισμό ορόσημων σημείων και είναι πλήρως αυτοματοποιημένη για την εύρεση της μετακίνηση του θεραπευτικού κρεβατιού. Για την εύρεση του πίνακα μετασχηματισμού, εισάγεται στον αλγόριθμο η παρούσα θέση και μια θέση αναφοράς και με βάση των ζευγαριών σημείων στις εικόνες που προκύπτουν από τα διάφορα βήματα του αλγορίθμου, υπολογίζονται απευθείας οι παράμετροι του μετασχηματισμού με προσέγγιση ελαχίστων τετραγώνων. Εγγενώς, οι μέθοδοι ευθυγράμμισης οι οποίες βασίζονται στα σημεία (point-based), είναι ικανές για ευθυγράμμιση δεδομένων με ακαθόριστες αρχικές συνθήκες των επιφανειών και με μερική ταύτιση αυτών.

Λέξεις Κλειδιά

Ευθυγράμμιση τρισδιάστατων ιατρικών δεδομένων, επεξεργασία εικόνας, Τμηματοποίηση εικόνας, ανάλυση εικόνας, ευθυγράμμιση εικόνας, RANSAC, Kinect, Συμπαγής ICP, Μη Συμπαγής ICP, Point Cloud

Abstract

The purpose of this thesis is to present a method of automatic alignment of three-dimensional medical data. The alignment process is aimed at correcting for relative displacements in space so that the data is aligned and same anatomical regions match spatially.

This method can be used to align the three dimensional data obtained from the KINECT sensor. Such an alignment technique can be applied in medicine. In radiotherapy, prior to each treatment, the patient should be aligned with the tomography data (CT). This alignment requires a system for verifying the orientation of the patient to accurately identify the patient and adapting the treatment bed to optimize the treatment. In practice, this process requires manual movement of the therapeutic bed using a laser source for scanning fixed landmark points on the skin of the patient. Alignment is required due to the positional displacement between images and also due to the development of irregularities that cause deformations (e.g. cysts, tumors).

This thesis proposes a fully automated method for determining the movement of the therapeutic bed, which does not required manual definition of landmark points. The requested transformation matrix is estimated from several point correspondences formed by aligning the points of the examined bed position with respect to a predetermined reference position. A least-squares approximation approach is used for matching the resulting point correspondences.. Inherently, alignment methods based on points (point-based),are able to align data with arbitrary initial conditions of surfaces or partially overlapping surfaces.

Keywords

3D medical image registration, Image Processing, Image Segmentation, Image Analysis, Image Alignment, RANSAC, Kinect, Linear ICP, Non-Linear ICP, Medical Image, Point Cloud

Ευχαριστίες

Η διπλωματική αυτή εργασία εκπονήθηκε κατά το ακαδημαϊκό έτος 2013-2014 υπό την επίβλεψη του κ. Γεώργιου Ματσόπουλου, καθηγητή της Σχολής Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών του Ε.Μ.Π., στον οποίο οφείλω ιδιαίτερες ευχαριστίες για την ανάθεσή της και το άριστο κλίμα συνεργασίας καθ' όλη τη διάρκεια εκπόνησής της. Επίσης, θα ήθελα να ευχαριστήσω τον Θεόδωρο Οικονομόπουλο, υποψήφιο διδάκτορα της σχολής, για τη βοήθεια και την καθοδήγησή του. Τέλος, ευχαριστώ τους γονείς, την οικογένεια και τους φίλους μου για τη στήριξη και τη βοήθειά τους κατά τη διάρκεια των σπουδών μου.

ΠΕΡΙΛΗΨΗ	I
ABSTRACT	II
ΕΥΧΑΡΙΣΤΙΕΣ	III
CONTENTS	V
LIST OF FIGURES	X
LIST OF TABLES	XIII
1 ΕΙΣΑΓΩΓΗ	1
1.1 ΣΚΟΠΟΣ ΔΙΠΛΩΜΑΤΙΚΗΣ ΕΡΓΑΣΙΑΣ	1
1.2 ΔΙΑΡΘΡΩΣΗ ΔΙΠΛΩΜΑΤΙΚΗΣ ΕΡΓΑΣΙΑΣ	2
2 ΕΠΙΣΚΟΠΗΣΗ ΜΕΘΟΔΩΝ ΕΥΘΥΓΡΑΜΜΙΣΗΣ ΕΙΚΟΝΩΝ	5
2.1 ΕΦΑΡΜΟΓΕΣ ΤΕΧΝΙΚΩΝ ΕΥΘΥΓΡΑΜΜΙΣΗΣ ΣΤΗΝ ΙΑΤΡΙΚΗ	5
2.2 ΟΡΙΣΜΟΣ ΕΥΘΥΓΡΑΜΜΙΣΗΣ ΔΕΔΟΜΕΝΩΝ	6
2.3 ΤΑΞΙΝΟΜΗΣΗ ΜΕΘΟΔΩΝ ΕΥΘΥΓΡΑΜΜΙΣΗΣ	8
2.3.1 ΔΙΑΣΤΑΣΗ ΜΕΤΑΣΧΗΜΑΤΙΣΜΟΥ	9
2.3.2 ΦΥΣΗ ΤΟΥ ΜΕΤΑΣΧΗΜΑΤΙΣΜΟΥ	10
2.3.2.1 Συμπαγής Μετασχηματισμός τρισδιάστατων δεδομένων	10

2.3.2.2 Μετασχηματισμός τύπου affine	10
2.3.2.3 Προβολικός Μετασχηματισμός	11
2.3.2.4 Καμπυλόγραμμοι Μετασχηματισμοί τρισδιάστατων δεδομένων	12
2.3.3 ΑΥΤΟΜΑΤΟΠΟΙΗΣΗ ΚΑΙ ΔΙΑΔΡΑΣΤΙΚΟΤΗΤΑ	13
2.3.4 ΤΕΧΝΙΚΕΣ ΑΝΑΖΗΤΗΣΗΣ	14
2.4. ΚΡΙΤΗΡΙΑ ΑΞΙΟΛΟΓΗΣΗΣ ΜΕΘΟΔΩΝ ΕΥΘΥΓΡΑΜΜΙΣΗΣ	15
2.4.1. ΣΥΝΑΡΤΗΣΗ ΕΠΙΤΥΧΙΑΣ	15
2.4.1.1. Απόλυτη διαφορά χρωματικών Πυκνοτήτων	16
2.4.1.2. Συντελεστής Συσχέτισης	17
2.4.1.3. Συντελεστής αμοιβαίας πληροφορίας	18
2.4.2 ΑΠΟΤΙΜΗΣΗ ΜΕΘΟΔΩΝ ΕΥΘΥΓΡΑΜΜΙΣΗΣ	18
2.4.2.1 Ακρίβεια	18
2.4.2.2 Ευρωστία και Ευστάθεια	19
2.4.2.3 Αξιοπιστία	19
2.4.2.4 Απαιτήσεις και Ανάγκες	19
2.4.2.5 Αλγοριθμική πολυπλοκότητα	19
2.4.2.6 Κλινική Χρησιμότητα	20
2.5 ΕΙΔΟΣ ΕΥΘΥΓΡΑΜΜΙΣΗΣ	20
2.5.1 ΕΥΘΥΓΡΑΜΜΙΣΗ ΒΑΣΙΣΜΕΝΗ ΣΕ ΕΞΩΓΕΝΗ ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ	20
2.5.2. ΕΥΘΥΓΡΑΜΜΙΣΗ ΒΑΣΙΣΜΕΝΗ ΣΕ ΕΝΔΟΓΕΝΗ ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ	22
2.5.2.1. Voxel property-based	22
2.5.2.2. Landmark property-based	23
2.5.2.3. Segmentation property-based	24
2.6 ΔΙΑΔΙΚΑΣΙΑ ΣΥΝΘΕΣΗΣ ΙΑΤΡΙΚΩΝ ΔΕΔΟΜΕΝΩΝ	25

2.6.1 ΜΕ ΧΡΗΣΗ ΛΟΓΙΚΩΝ ΤΕΛΕΣΤΩΝ	25
2.6.2 ΜΕ ΧΡΗΣΗ ΧΑΡΤΗ ΨΕΥΔΟ-ΧΡΩΜΑΤΟΣ	26
2.6.3 ΤΕΧΝΙΚΕΣ ΤΑΞΙΝΟΜΗΣΗΣ	26
3 ΣΥΛΛΟΓΗ 3D ΝΕΦΩΝ ΣΗΜΕΙΩΝ (POINT CLOUD)	27
3.1 3D POINT CLOUD	27
3.2 ΑΠΕΙΚΟΝΙΣΗ ΒΑΘΟΥΣ	27
3.2.1 TIME OF FLIGHT	28
3.2.2 STRUCTURED LIGHT	28
3.2.2.1 2D Ανάλυση	29
3.2.2.2 3D Ανάλυση	29
3.2 ΣΥΣΚΕΥΕΣ ΣΑΡΩΣΗΣ 3D POINT CLOUD	31
3.2.1 ΣΑΡΩΤΗΣ ΧΕΙΡΟΣ (HAND HELD)	31
3.2.2 ΚΙΝΗΤΟΙ ΣΑΡΩΤΕΣ (MOBILE SCANNER)	31
3.2.3 ΕΠΙΓΕΙΟΙ ΣΑΡΩΤΕΣ	32
3.3 ΣΑΡΩΤΗΣ KINECT	33
3.3.1 ΠΑΡΟΥΣΙΑΣΗ ΤΜΗΜΑΤΩΝ ΑΙΣΘΗΤΗΡΑ KINECT	33
3.3.1.1 Έγχρωμη Κάμερα	33
3.3.1.2 Πομπός και Αισθητήρας Υπερύθρων (IR emitter and IR depth sensor)	34
3.3.1.3 Κινητήρας Κλίσης (Tilt motor)	35
3.3.2 ΛΗΨΗ ΔΕΔΟΜΕΝΩΝ ΒΑΘΟΥΣ	35
3.3.2.1 Συσχέτιση των στοιχείων	36
3.3.2.2 Μαθηματικό Μοντέλο Εύρεσης Βάθους	39
3.3.2.3 Βαθμονόμηση Αισθητήρα	40

3.3.2.4 Πηγή Σφαλμάτων	41
3.3.2.5 Μοντέλο Θορύβου	41
3.3.2.6 Πυκνότητα Νέφους Σημείων	42
4 ΜΕΘΟΔΟΛΟΓΙΑ ΑΥΤΟΜΑΤΗΣ ΕΥΘΥΓΡΑΜΜΙΣΗΣ ΕΙΚΟΝΩΝ	43
4.1 ΕΙΣΑΓΩΓΗ	43
4.2 ΣΥΜΠΑΓΗΣ ICP	44
4.2.1 ΕΠΙΛΟΓΗ ΣΗΜΕΙΩΝ	45
4.2.2 ΤΑΥΤΙΣΗ ΣΗΜΕΙΩΝ	46
4.2.3 ΒΑΡΥΤΗΤΑ ΣΗΜΕΙΩΝ	48
4.2.4 ΑΠΟΡΡΙΨΗ ΖΕΥΓΑΡΙΩΝ	49
4.2.4.1 Μέθοδος RANSAC	49
4.2.5 ΣΥΝΑΡΤΗΣΗ ΜΕΤΡΗΣΗΣ ΣΦΑΛΜΑΤΟΣ	51
4.2.6 ΕΛΑΧΙΣΤΟΠΟΙΗΣΗ ΣΥΝΑΡΤΗΣΗΣ ΣΦΑΛΜΑΤΟΣ	52
4.2.7 ΣΥΝΘΗΚΗ ΤΕΡΜΑΤΙΣΜΟΥ	53
4.3 ΜΗ ΣΥΜΠΑΓΗΣ ICP	53
4.3.1 LEVENBERG-MARQUARDT ICP	54
4.3.2 ΥΠΟΛΟΓΙΣΜΟΣ ΠΑΡΑΓΩΓΩΝ ΑΛΓΟΡΙΘΜΟΥ	57
4.3.2.1 Μετασχηματισμός Απόστασης	57
4.3.2.2 Quaternions	57
5 ΑΠΟΤΕΛΕΣΜΑΤΑ	59
5.1 ΚΑΘΟΡΙΣΜΟΣ ΒΕΛΤΙΣΤΩΝ ΠΑΡΑΜΕΤΡΩΝ ΕΥΘΥΓΡΑΜΜΙΣΗΣ	59

5.2 ΕΥΘΥΓΡΑΜΜΙΣΗ ΕΙΚΟΝΩΝ ΜΕ ΓΝΩΣΤΗ ΠΑΡΑΜΟΡΦΩΣΗ	60
5.3 ΕΥΘΥΓΡΑΜΜΙΣΗ ΠΡΑΓΜΑΤΙΚΩΝ ΕΙΚΟΝΩΝ	67
6 ΑΞΙΟΛΟΓΗΣΗ	81
6.1 ΑΞΙΟΛΟΓΗΣΗ ΤΗΣ ΠΡΟΤΕΙΝΟΜΕΝΗΣ ΜΕΘΟΔΟΥ ΕΥΘΥΓΡΑΜΜΙΣΗΣ	81
6.1.1 ΦΩΤΙΣΜΟΣ	81
6.1.2 ΑΠΟΣΤΑΣΗ	82
6.2 ΜΕΛΛΟΝΤΙΚΕΣ ΒΕΛΤΙΩΣΕΙΣ ΚΑΙ ΕΠΕΚΤΑΣΕΙΣ	83

List of Figures

<i>Εικόνα 1.1 :</i>	<i>2</i>
<i>Εικόνα 2.1 :</i>	<i>8</i>
<i>Εικόνα 2.2 :</i>	<i>13</i>
<i>Εικόνα 2.3 :</i>	<i>21</i>
<i>Εικόνα 2.4 :</i>	<i>21</i>
<i>Εικόνα 2.5 :</i>	<i>23</i>
<i>Εικόνα 2.6 :</i>	<i>24</i>
<i>Εικόνα 2.7 :</i>	<i>25</i>
<i>Εικόνα 3.1 :</i>	<i>28</i>
<i>Εικόνα 3.2 :</i>	<i>29</i>
<i>Εικόνα 3.3 :</i>	<i>30</i>
<i>Εικόνα 3.4 :</i>	<i>31</i>
<i>Εικόνα 3.5 :</i>	<i>32</i>
<i>Εικόνα 3.6 :</i>	<i>33</i>
<i>Εικόνα 3.7 :</i>	<i>34</i>
<i>Εικόνα 3.8:</i>	<i>34</i>
<i>Εικόνα 3.9 :</i>	<i>35</i>
<i>Εικόνα 3.10 :</i>	<i>35</i>
<i>Εικόνα 3.11:</i>	<i>36</i>
<i>Εικόνα 3.12 :</i>	<i>36</i>
<i>Εικόνα 3.13 :</i>	<i>37</i>
<i>Εικόνα 3.14 :</i>	<i>37</i>

Εικόνα 3.15 :	38
Εικόνα 3.16 :	38
Εικόνα 3.17 :	40
Εικόνα 4.1 :	45
Εικόνα 4.2 :	46
Εικόνα 4.3 :	47
Εικόνα 4.4 :	48
Εικόνα 4.5 :	51
Εικόνα 4.6 :	56
Εικόνα 5.1 :	59
Εικόνα 5.2 :	60
Εικόνα 5.3 :	61
Εικόνα 5.4 :	65
Εικόνα 5.5 :	65
Εικόνα 5.6 :	66
Εικόνα 5.7 :	66
Εικόνα 5.8 :	66
Εικόνα 5.9 :	71
Εικόνα 5.10 :	71
Εικόνα 5.11 :	72
Εικόνα 5.12 :	72
Εικόνα 5.13 :	73
Εικόνα 5.14 :	73
Εικόνα 5.15 :	74
Εικόνα 5.16 :	74
Εικόνα 5.17 :	77

Εικόνα 5.18 :	77
Εικόνα 5.19 :	78
Εικόνα 5.20 :	78
Εικόνα 5.21 :	79
Εικόνα 5.22 :	79
Εικόνα 5.23 :	80
Εικόνα 5.24 :	80
Εικόνα 6.1 :	82
Εικόνα 6.2 :	83
Εικόνα 6.3 :	83

List of Tables

<i>Πίνακας 2.1:</i>	<hr/>	<i>16</i>
<i>Πίνακας 5.1:</i>	<hr/>	<i>63</i>
<i>Πίνακας 5.1:</i>	<hr/>	<i>69</i>
<i>Πίνακας 5.3:</i>	<hr/>	<i>75</i>

1 Εισαγωγή

1.1 Σκοπός Διπλωματικής Εργασίας

Οι πιο συνήθεις διαδικασίες στις μοντέρνες μεθόδους ακτινοθεραπείας στοχεύουν στο να εξειδικεύσουν την κατανομή της δόσης ανάλογα με το σχήμα του όγκου ενώ ελαχιστοποιούν την δόση στον γειτονικό υγιή ιστό. Η προσέγγιση αυτή όμως καθιστά ακριβής με τις προϋποθέσεις ότι η θέση και ο προσανατολισμός τους ασθενή είναι ακριβές και ρυθμιζόμενη. Έτσι καθίσταται αναγκαία η επίτευξη της τοποθέτησης του ασθενή σε σωστή θέση για να πετύχουν την βέλτιστη θεραπεία.

Σε αυτή την διπλωματική εργασία παρουσιάζεται μια μέθοδος αυτόματης ευθυγράμμισης τρισδιάστατων ιατρικών δεδομένων. Με τον αλγόριθμο της ευθυγράμμισης τρισδιάστατων επιφανειών επιδιώκουμε την αναπαράσταση τρισδιάστατων ιατρικών εικόνων από διάφορες χρονικές στιγμές του ανθρώπινου σώματος στο ίδιο σύστημα συντεταγμένων. Ακολουθώς επιδιώκουμε την εύρεση της μετατόπισης και περιστροφή της εξεταστικής τράπεζας του ακτινογραφικού μηχανήματος, ώστε αυτή να συμπίπτει με την τρισδιάστατη ιατρική εικόνα αναφοράς για την επίτευξη της βέλτιστης θεραπείας του ασθενή.

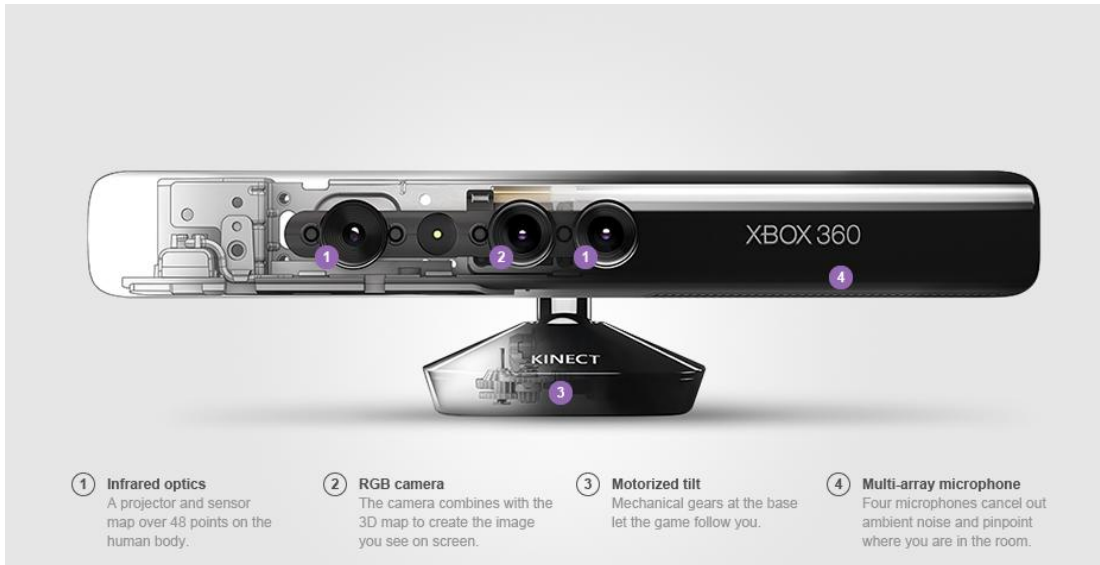
Μέχρι στιγμής υπάρχει μια πληθώρα από τεχνικές για ευθυγράμμιση επιφανειών όπου θα αναφερθούν και λεπτομερέστερα σε επόμενο κεφάλαιο. Αρχικός σκοπός μας όμως ήταν η χρήση ενός οπτικού αισθητήρα, χαμηλού κόστους για να μας δώσει την τρισδιάστατη πληροφορία μιας περιοχής του ανθρώπινου σώματος. Για τους σκοπούς αυτής της διπλωματικής εργασίας, χρησιμοποιήσαμε τον οπτικό αισθητήρα της Microsoft, KINECT ο οποίος δεν έχει χρησιμοποιηθεί ακόμη στον συγκεκριμένο τομέα.

Εισαγωγικά, ο αισθητήρας KINECT αποτελεί ένα χαμηλού κόστους οπτικό αισθητήρα τρισδιάστατης αναπαράστασης νέφους σημείων, με σχετικά πολύ καλή ακρίβεια και τα λογισμικά αυτού διατίθενται χωρίς κανένα κόστος. Αποτελείται από μια διάταξη τριών καινοτόμων υλικών τα οποία συνυπάρχουν και συνεργάζονται για την λήψη, επεξεργασία και παρουσίαση του τελικού αποτελέσματος. Τα τρία αυτά στοιχεία υλικού είναι:

Έγχρωμη VGA κάμερα: Βοηθά στην αναγνώριση προσώπου και άλλων χαρακτηριστικών με την ανίχνευση των τριών χρωματικών καναλιών (κόκκινο, πράσινο και μπλε).

Αισθητήρας βάθους: Είναι ένας προβολέας υπερύθρων και μία μονοχρωματική CMOS όπου συνεργάζονται για την «παρακολούθηση» του χώρου, ανεξάρτητα από τις συνθήκες φωτισμού.

Multi-array microphone: Πρόκειται για μία σειρά από τέσσερα μικρόφωνα που μπορούν να απομονώσουν τις διάφορες φωνές από τον θόρυβο στο χώρο.



Εικόνα 1.1 : Απεικόνιση τμημάτων του συστήματος KINECT

Οι διάφορες τεχνικές και αλγόριθμοι οι οποίοι ήδη υπάρχουν και εφαρμόζονται, απαιτούν γνώση αντίστοιχων σημείων ή ανατομικών δομών, που ανιχνεύονται είτε αυτόματα ή ορίζονται από τον χρήστη και με βάση αυτών προσδιορίζουν τον κατάλληλο μετασχηματισμό που συνδέει τα συστήματα συντεταγμένων των δύο απεικονίσεων. Σε αντίθεση με αυτό, στον αλγόριθμο που παρουσιάζεται σε αυτή την διπλωματική εργασία, η μέθοδος ευθυγράμμισης τρισδιάστατων ιατρικών εικόνων, γίνεται εντελώς αυτόματα, χωρίς την ανάγκη εισαγωγής από τον χρήστη σημείων ή συγκεκριμένων ανατομικών δομών και χωρίς να απαιτείται κατάτμηση εικόνων.

1.2 Διάρθρωση Διπλωματικής Εργασίας

Στο **Κεφάλαιο 2** γίνεται μια επισκόπηση των μεθόδων ευθυγράμμισης των ιατρικών δεδομένων τις οποίες συναντούμε στη βιβλιογραφία. Επίσης γίνεται μια εισαγωγή του αναγνώστη στην επιστήμη της ευθυγράμμισης δεδομένων παρουσιάζοντας τις διάφορες κατηγορίες στις οποίες διαχωρίζεται ο τομέας αυτός. Οι κατηγορίες διαχωρίζονται με βάση την διάσταση του μετασχηματισμού, την φύση του μετασχηματισμού (Συμπαγής, Affine, Προβολικός, Καμπυλόγραμμοι), την αυτοματοποίηση και διαδραστικότητα των τεχνικών και τέλος με βάση τις τεχνικές αναζήτησης. Τέλος, αναφέρονται οι μέθοδοι αξιολόγησης μιας τεχνικής ευθυγράμμισης και τα κριτήρια τα οποία πρέπει να ικανοποιεί μια τέτοια τεχνική.

Στο **Κεφάλαιο 3** εισάγεται η έννοια του νέφους σημείων και διαφόρων ορισμών της ιδέας αυτής. Αναφερόμαστε στους τρόπους απεικόνισης του βάθους όπου δίνεται βαρύτητα στην μέθοδο του δομημένου φωτός όπου και μας ενδιαφέρει. Στην συνέχεια αναφέρονται διάφορες συσκευές σάρωσης τρισδιάστατων νεφών σημείων, μία εκ των οποίων είναι και ο αισθητήρας KINECT, για τον οποίο εμβαθύνουμε. Αναφερόμαστε στα διάφορα μέρη του αισθητήρα όπως επίσης και στα διάφορα χαρακτηριστικά αυτού, όπως

είναι ο τρόπος λήψης δεδομένων βάθους, η βαθμονόμηση του αισθητήρα, πιθανές πηγές σφαλμάτων στην απεικόνιση αντικειμένων, το μοντέλο θορύβου που θεωρούμε ότι υπάρχει και τέλος η πυκνότητα των σημείων του νέφους.

Στο **Κεφάλαιο 4** παρουσιάζεται η μέθοδος αυτόματης ευθυγράμμισης εικόνων. Αρχικά αναλύεται ο συμπαγής αλγόριθμος Iterative Closest Point (ICP) ο οποίος εφαρμόζεται στην παρούσα διπλωματική εργασία και όλα τα επιμέρους τμήματα αυτού όπως είναι η επιλογή σημείων, η ταύτιση σημείων, η βαρύτητα σημείων, η απόρριψη ζευγαριών σημείων με την μέθοδο RANdom Sample Consensus (RANSAC), η επιλογή συνάρτησης σφάλματος όπως και η μεθοδολογία ελαχιστοποίησης της και οι συνθήκες τερματισμού του αλγορίθμου. Τέλος, αναλύεται ο μη συμπαγής ICP που εφαρμόζεται επίσης στην παρούσα διπλωματική εργασία και συγκεκριμένα ο αλγόριθμος Levenberg-Marquardt ICP.

Στο **Κεφάλαιο 5** παρουσιάζεται ο προτεινόμενος αλγόριθμος ευθυγράμμισης και υπολογίζονται οι παράμετροι με τις οποίες επιτυγχάνεται μια ισορροπία μεταξύ χρόνου εκτέλεσης και ακρίβειας αποτελεσμάτων. Εν συνεχεία ελέγχεται η απόδοση αυτού σε γνωστές παραμορφώσεις εικόνων και τέλος εφαρμόζεται σε πραγματικές εικόνες όπου και παρουσιάζονται τα αποτελέσματα αυτού.

Στο **Κεφάλαιο 6** αξιολογείται η προτεινόμενη μέθοδος και παρουσιάζονται πιθανές επεκτάσεις και μελλοντικές βελτιώσεις αυτής.

2 Επισκόπηση Μεθόδων Ευθυγράμμισης Εικόνων

2.1 Εφαρμογές Τεχνικών Ευθυγράμμισης στην Ιατρική

Τα τελευταία χρόνια παρουσιάζεται αυξημένη χρήση των ιατρικών εικόνων για τη διάγνωση, το σχεδιασμό και την καθοδήγηση θεραπειών και για την επίβλεψη της εξέλιξης μίας ασθένειας. Η παρακολούθηση των διαφόρων αντικειμένων γίνεται με την συνεχή αναπαράσταση τους σε διάφορες χρονικές στιγμές και συχνά συνδυάζοντας εικόνες με διαφορετικές απεικονιστικές τεχνικές και άρα διαφορετική πληροφορία, δηλαδή είτε ανατομική είτε λειτουργική πληροφορία. Η σύνθεση αυτών των δύο απεικονιστικών τεχνικών μεγιστοποιεί τη διαγνωστική διαθέσιμη πληροφορία. Επιπρόσθετα, ο όγκος δεδομένων που παράγονται από κάθε μελλοντική γενιά απεικονιστικών συστημάτων είναι μεγαλύτερος από τα δεδομένα τα οποία αντλούνταν από τα προηγούμενης γενιάς απεικονιστικά συστήματα. Λόγω αυτής της τάσης, δηλαδή, να έχουμε ακόμη καλύτερα δεδομένα σε κάθε νέα γενιά απεικονιστικών συστημάτων και για το γεγονός ότι με το συνδυασμό διαφόρων τεχνικών λαμβάνουμε μια σχετικά πλήρη πληροφορία του υπό εξέταση οργάνου, υπάρχουν διάφοροι συνδυασμοί ιατρικών απεικονιστικών τεχνικών(1). Από την μία μεριά, είναι τα απεικονιστικά συστήματα τα οποία μας δίνουν λειτουργική πληροφορία όπως είναι οι ιατρικές απεικονιστικές διατάξεις εκπομπής μονού φωτονίου (SPECT) και τομογραφίας εκπομπής ποζιτρονίων (PET). Οι δύο αυτές τεχνικές μπορούν να απεικονίσουν λειτουργικές πληροφορίες ακόμη και για τα αρχικά στάδια εξέλιξης των καρκινικών όγκων, αλλά δεν αποδίδουν αξιόπιστα τα ανατομικά χαρακτηριστικά του υπό εξέταση οργάνου. Από την άλλη μεριά, είναι οι απεικονιστικές τεχνικές οι οποίες μας παρέχουν ανατομική πληροφορία, όπως είναι η αξονική (CT) και η μαγνητική (MR) τομογραφία, οι υπέρηχοι και οι ακτίνες-Χ. Για την σωστή σύνθεση μιας εικόνας όπου εμπεριέχει διάφορες πληροφορίες από διαφορετικές τεχνικές, θα πρέπει τα σημεία των δύο ή περισσότερων εικόνων να αντιπροσωπεύουν το ίδιο σημείο στο όργανο υπό μελέτη. Έτσι, η ευθυγράμμιση καλείται να επιτελέσει αυτή την σωστή αντιστοίχιση των εικόνων.

Επιπρόσθετα, η ευθυγράμμιση επιφανειών χρησιμοποιείται για την ανάδειξη μεταβολών μεταξύ δεδομένων που έχουν ανακτηθεί σε διαφορετικές χρονικές στιγμές. Η χρήση αυτής της τεχνικής παρουσιάζεται κυρίως κατά την μέτρηση της οστικής υποστήριξης προς τα εμφυτεύματα με χρήση οδοντιατρικών ακτινογραφιών (1), (2). Τέλος, η ευθυγράμμιση ιατρικών δεδομένων βρίσκει εφαρμογή σε περιπτώσεις όπου χρησιμοποιούνται δεδομένα από ανατομικούς άτλαντες σε συνδυασμό με πραγματικά κλινικά δεδομένα, καθώς και σε μελέτες επί πληθυσμού ασθενών (3).

Συνοψίζοντας, η ευθυγράμμιση ιατρικών δεδομένων είναι ένα βασικό εργαλείο για λόγους εξουδετέρωσης σφαλμάτων και απόκτησης περισσότερης

πληροφορίας από διάφορες απεικονιστικές τεχνικές. Όπως αναφέρθηκε και πιο πάνω, επιχειρείται η σύνδεση δεδομένων από διαφορετικές απεικονιστικές τεχνικές, ακόμη και από την ίδια απεικονιστική τεχνική σε διαφορετικές χρονικές στιγμές για την απόκτηση μιας πληρέστερης πληροφορίας. Επίσης, η ευθυγράμμιση δεδομένων παρουσιάζεται και σε θέματα χωρικών μεταβολών όπου δημιουργούνται λόγω ακούσιων κινήσεων του ασθενή (π.χ. καρδιακή λειτουργία, αναπνοή), σε απεικονιστικά συστήματα τα οποία δεν έχουν την απαραίτητη σταθερότητα λόγω κινητών μερών τους (π.χ. CCD) όπως και στην ανάπτυξη ανωμαλιών που προκαλούν παραμορφώσεις (π.χ. όγκοι, κύστες).

2.2 Ορισμός Ευθυγράμμισης Δεδομένων

Ευθυγράμμιση εικόνων επιτυγχάνεται μεταξύ εικόνων του ίδιου ή διαφορετικού αντικειμένου, τα οποία λαμβάνονται με την ίδια ή διαφορετική απεικονιστική τεχνική. Ο όρος ευθυγράμμιση έχει δύο ερμηνείες ελαφρώς διαφορετικές. Η πρώτη ερμηνεία της ευθυγράμμισης αναφέρεται στην τοπική ευθυγράμμιση όπου είναι ο μετασχηματισμός της θέσης x από την μία εικόνα σε μία άλλη, ή από την εικόνα στο σύστημα συντεταγμένων του θεραπευτικού μηχανήματος. Χρησιμοποιούμε το σύμβολο \mathcal{T} για τον τοπικό μετασχηματισμό ευθυγράμμισης. Η δεύτερη ερμηνεία της ευθυγράμμισης αφορά την χαρτογράφηση της θέσης και της αντίστοιχης έντασης φωτεινότητας στην συγκεκριμένη θέση. Χρησιμοποιούμε το σύμβολο \mathcal{T} για την περιγραφή της δεύτερης έννοιας, η οποία ενσωματώνει τις έννοιες της επαναδειγματοληψίας και παρεμβολής(1).

Χρησιμοποιώντας γεωμετρικούς όρους, ο μετασχηματισμός ευθυγράμμισης αναφέρεται και σαν χαρτογράφηση. Για την χαρτογράφηση \mathcal{T} , η οποία μετατρέπει την θέση x από μία εικόνα σε μία άλλη, ή από μία εικόνα στο σύστημα συντεταγμένων του θεραπευτικού μηχανήματος, ισχύει:

$$\mathcal{T}: x_A \mapsto x_B \Leftrightarrow \mathcal{T}(x_A) = x_B \quad (1.1)$$

από το οποίο συμπεράνουμε ότι η χαρτογράφηση \mathcal{T} είναι τοπική χαρτογράφηση. Επιβάλλεται όμως και η χαρτογράφηση του \mathcal{T} , όπου εμπεριέχει τον μετασχηματισμό της θέσης και της αντίστοιχης έντασης φωτεινότητας από την εικόνα A στην εικόνα B. Ως εκ τούτου, το \mathcal{T} χαρτογραφεί μια εικόνα σε μια εικόνα, ενώ το \mathcal{T} χαρτογραφεί μεταξύ συντεταγμένων. Αν θέλουμε την επικάλυψη δύο εικόνων ή την αφαίρεση της μίας από την άλλη, τότε θα χρειαστούμε να γνωρίζουμε το \mathcal{T} και όχι το \mathcal{T} . Το \mathcal{T} ορίζεται μόνο στην περιοχή όπου επικαλύπτονται τα οπτικά πεδία των εικόνων και εξαρτάται από την δειγματοληψία και ανάλυση της εικόνας. Η $A(x_A)$ ορίζεται ως η ένταση της εικόνας A στην θέση x_A και αντίστοιχα για την εικόνα B. Επίσης υποθέσουμε ότι οι εικόνες έχουν πεπερασμένο οπτικό πεδίο, όπως ισχύει στις ιατρικές εικόνες που αναπαριστούν ένα μέρος του ανθρώπινου σώματος.

Μπορούμε να υποθέσουμε τις εντάσεις δύο εικόνων οι οποίες χαρτογραφούνται από σημεία του ασθενή μέσα στα αντίστοιχα οπτικά τους πεδία (Ω):

$$A : x_A \in \Omega_A \mapsto A(x_A) \quad (1.2)$$

$$B : x_B \in \Omega_B \mapsto B(x_B) \quad (1.3)$$

Η ύπαρξη δύο διαφορετικών οπτικών πεδίων Ω_A και Ω_B οφείλεται στο γεγονός ότι μπορεί οι δύο εικόνες να αναφέρονται σε δύο διαφορετικά τμήματα του ασθενή.

Για την σύγκριση των εικόνων A και B, θέλουμε να οριστούν και οι δύο στην τοποθεσία x_A έτσι ώστε να μπορούμε να έχουμε $B(x_A)$. Αυτό όμως είναι λάθος καθώς η εικόνα B δεν ορίζεται στην τοποθεσία x_A , έτσι θα πρέπει να εισαχθεί η έννοια του B^T για τον μετασχηματισμό της εικόνας B με δεδομένη τη χαρτογράφηση \mathcal{T} . Αν το \mathcal{T} εγγράφει ακριβώς τις εικόνες, τότε τα $A(x_A)$ και $B^T(x_A)$ θα αναπαριστούν την ίδια τοποθεσία στο αντικείμενο με κάποιο σφάλμα το οποίο εξαρτάται από το \mathcal{T} αν η παρεμβολή δεν υπήρχε στην εικόνα τότε θα μπορούσαμε να γράφαμε $B^T(x_A)$ αλλά λόγω τις διακριτής φύσης των ιατρικών εικόνων, η παρεμβολή είναι απαραίτητη για τον πραγματικό μετασχηματισμό.

Καθώς οι εικόνες A και B, οι οποίες λήφθηκαν από την ίδια ή διαφορετική (τυπικότητα) απεικονιστική τεχνική, αναπαριστούν ένα αντικείμενο X, υπάρχει μια σχέση μεταξύ των χωρικών συντεταγμένων στις εικόνες A και B. Η εικόνα A είναι τέτοια ώστε η θέση $x \in X$ χαρτογραφείται σαν x_A και η εικόνα B χαρτογραφεί το x σε x_B . Η διαδικασία ευθυγράμμισης περιέχει την ανάκτηση του τοπικού μετασχηματισμού \mathcal{T} ο οποίος χαρτογραφεί το x_A σε x_B σε όλη την περιοχή η οποία ενδιαφέρει, η οποία είναι η τομή των δύο οπτικών πεδίων Ω_A και Ω_B , όπου συμβολίζεται ως $\Omega_{A,B}^T$. Είναι εμφανές ότι η τομή των οπτικών πεδίων εξαρτάται από τα δύο οπτικά πεδία των εικόνων A και B, όπως επίσης και από τον τοπικό μετασχηματισμό \mathcal{T} και ορίζεται ως:

$$\Omega_{A,B}^T = \{x_A \in \Omega_A \mid \mathcal{T}^{-1}(x_A) \in \Omega_B\} \quad (1.4)$$

Αλγόριθμοι ευθυγράμμισης που κάνουν χρήση των γεωμετρικών χαρακτηριστικών των εικόνων όπως τα σημεία, τις γραμμές και τις επιφάνειες, καθορίζουν το μετασχηματισμό \mathcal{T} . Ο υπολογισμός γίνεται με το προσδιορισμό των συνόλων από στοιχεία των εικόνων (x_A και x_B) που αντιστοιχούν στην ίδια φυσική οντότητα και είναι ορατή και στις δύο εικόνες και υπολογίζουν τον μετασχηματισμό \mathcal{T} . Όταν οι αλγόριθμοι αυτοί είναι επαναληπτικοί, επαναληπτικά προσδιορίζουν τον μετασχηματισμό \mathcal{T} και όταν ο αλγόριθμος συγκλίνει, τότε θα έχουμε υπολογίσει το \mathcal{T} .

Από την άλλη, οι αλγόριθμοι ευθυγράμμισης οι οποίοι βασίζονται στην ένταση φωτεινότητας της εικόνας δουλεύουν διαφορετικά. Επαναληπτικά υπολογίζεται ο μετασχηματισμός \mathcal{T} ο οποίος βελτιστοποιεί ένα μέτρο ομοιότητας των εικονοστοιχείων. Πολλά από αυτά τα μέτρα ομοιότητας περιλαμβάνουν ανάλυση συνόλων από ίσης έντασης σημεία της εικόνας. Για μία εικόνα A, ένα σύνολο από ίσης έντασης σημεία με τιμή έντασης a είναι ένα σύνολο από εικονοστοιχεία σε μια υπό-περιοχή της εικόνας A, και καθορίζεται ως εξής:

$$\Omega_a = \{x_A \in \Omega_A \mid A(x_A) = a\} \quad (1.5)$$

Κάποιοι αλγόριθμοι δεν δουλεύουν σε σύνολα από περιοχές οι οποίες έχουν μια καθορισμένη ένταση, αλλά σε σύνολα από περιοχές οι οποίες περιέχουν σημεία της εικόνας και έχουν ένταση ορισμένου εύρους τιμών.

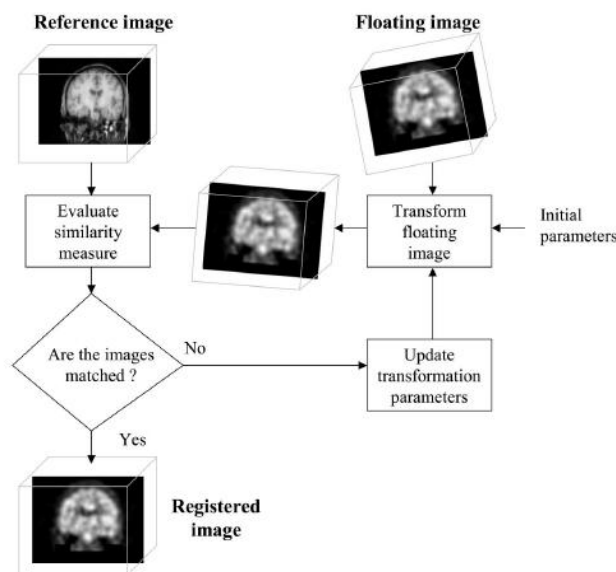
Όμως, το Ω_α είναι το σύνολο ίσης έντασης σημείων μέσα στην εικόνα A, η οποία έχει οπτικό πεδίο το Ω_A . Έτσι, όπως έχουμε προαναφέρει, για ευθυγράμμιση δύο εικόνων με μέτρο ομοιότητας των εικονοστοιχείων, πρέπει να εργαστούμε στην τομή των δύο οπτικών τους πεδίων $\Omega_{A,B}^T$. Τα σύνολα με τις ίδιες εντάσεις σε αυτό το πεδίο είναι μια συνάρτηση του μετασχηματισμού T και ορίζεται ως εξής:

$$\Omega_\alpha^T = \{x_A \in \Omega_{A,B}^T | A(x_A) = \alpha\} \quad (1.6)$$

Παρόμοια μπορούμε να υπολογίσουμε και τα σύνολα που αντιστοιχούν στην εικόνα B αλλά με κάποιες διαφορές αφού πάντα υποθέτουμε ότι η εικόνα B είναι προς ευθυγράμμιση και άρα έχει μετασχηματιστεί. Άρα τα σύνολα με τις ίδιες εντάσεις καθορίζονται ως τα σύνολα των εικονοστοιχείων στον χώρο A τα οποία έχουν ένταση φωτεινότητας στην εικόνα B^T :

$$\Omega_b^T = \{x_A \in \Omega_{A,B}^T | B^T(x_A) = b\} \quad (1.7)$$

Το πιο κάτω διάγραμμα αναπαριστά όλη τα επιμέρους βήματα τα οποία θα πρέπει να εκτελεστούν για να πετύχουμε ευθυγράμμιση και τα οποία θα αναλύσουμε εκτενέστερα στα πιο κάτω κεφάλαια:



Εικόνα 2.1 : Σχηματικό διάγραμμα αλγορίθμου ευθυγράμμισης

2.3 Ταξινόμηση Μεθόδων Ευθυγράμμισης

Με βάση το πρόβλημα το οποίο μελετάται, η διαδικασία της ευθυγράμμισης διαφοροποιείται. Με βάση τις (5),(6), (7)και (8)όπου γίνεται μια ανασκόπηση και ταξινόμηση των μεθόδων ευθυγράμμισης, η διαδικασία της ευθυγράμμισης μπορεί να ταξινομηθεί με βάση τα κριτήρια που διατυπώθηκαν από τον Van den

Elsen (8). Τα κριτήρια σε μια μέθοδο ευθυγράμμισης επιλέγονται με βάση το πρόβλημα το οποίο πρόκειται να αναλυθεί όπου αυτό καθορίζεται από την τεχνική απεικόνισης των δεδομένων, το συσχετισμό που παρουσιάζουν τα δεδομένα και την προέλευσή τους.

Στην Ιατρική, η ευθυγράμμιση τις περισσότερες φορές πρέπει να γίνει σε δεδομένα τα οποία περιέχουν διαφορετική πληροφορία, δηλαδή λαμβάνονται από διαφορετικές απεικονιστικές τεχνικές. Σε αντίθεση με την σύνθεση όμοιων απεικονιστικών τεχνικών, στη σύνθεση διαφορετικών τεχνικών μπορεί είναι δύσκολη η αντιστοίχιση των τιμών των χρωματικών πυκνοτήτων από την μια εικόνα στην άλλη. Επίσης, δεν είναι αναγκαίο μία από τις δύο εικόνες να είναι πραγματική. Μπορεί η απεικόνιση ενός τμήματος του ασθενή, να συσχετιστεί με ένα μοντέλο φυσιολογικού τμήματος, έτσι ώστε να εντοπιστούν οι τυχόν ανωμαλίες που υπάρχουν (4), (5), (6). Επίσης, η ευθυγράμμιση επιφανειών βρίσκει εφαρμογή στην σωστή τοποθέτηση του ασθενή με βάση την αρχική ακτινοβολή του για βελτιστοποίηση της ακτινοθεραπείας και αποφυγή ανεπιθύμητων καταστάσεων στον ασθενή. Αυτή η ευθυγράμμιση γίνεται μέσω CTδεδομένων (7), (8), ή με βάση τον αλγόριθμο ο οποίος παρουσιάζεται στην παρούσα διπλωματική εργασία, δηλαδή με οπτική ευθυγράμμιση.

Τα κριτήρια στα οποία βασίζεται η ταξινόμηση των μεθόδων ευθυγράμμισης είναι τα εξής:

- Διάσταση Μετασχηματισμού
- Φύση του μετασχηματισμού
 - Συμπαγής
 - Affine
 - Προβολικός
 - Καμπυλόγραμμος
- Αυτοματοποίηση και διαδραστικότητα
- Τεχνικές αναζήτησης

2.3.1 Διάσταση μετασχηματισμού

Στην ιατρική επιστήμη έχουμε μια πληθώρα από ιατρικές εικόνες και απεικονιστικές τεχνικές και κατά συνέπεια ο μετασχηματισμός πρέπει να έχει τις ανάλογες διαστάσεις. Αυτό το κριτήριο χωρίζεται σε δύο υποκατηγορίες. Στην πρώτη κατηγορία βρίσκονται οι εικόνες οι οποίες δεν είναι συνάρτηση του χρόνου, ενώ στην άλλη υποκατηγορία είναι οι εικόνες που λαμβάνονται με την πάροδο του χρόνου. Σε εικόνες όπου προέρχονται από προβολές ή τομές του ίδιου επιπέδου, ο μετασχηματισμός μπορεί να είναι μονοδιάστατος ή δισδιάστατος. Αν έχουμε σύνολο από τρισδιάστατα δεδομένα ή ακόμη και μια αλληλουχία δισδιάστατων εικόνων και σαν τρίτη διάσταση τον χρόνο, τότε ο μετασχηματισμός πρέπει να είναι τρισδιάστατος. Τέλος, μπορεί να παρουσιαστεί και τεσσάρων διαστάσεων μετασχηματισμός όπου θα ευθυγραμμίζει χρονικές ακολουθίες τρισδιάστατων συνόλων από δεδομένα.

2.3.2 Φύση του μετασχηματισμού

Ένας μετασχηματισμός ονομάζεται ολικός εάν εφαρμόζεται σε ολόκληρη την εικόνα και τοπικός εάν η εφαρμογή του γίνεται σε όλα τα προκαθορισμένα σε διαστάσεις μικρά τμήματα της εικόνας. Με την εφαρμογή ενός τοπικού μετασχηματισμού στα διάφορα τμήματα της εικόνας, στα όρια αυτών των περιοχών παρουσιάζονται ασυνέχειες και για τον λόγο αυτό σπανίως εφαρμόζονται σε μία εικόνα χωρίς να τεθεί σε ένα είδος προ-επεξεργασίας για την εξάλειψη της όποιας ασυνέχειας. Η δημιουργία μιας ασυνέχειας εξαρτάται από τον βαθμό ελαστικότητας του μετασχηματισμού και από την τήρηση συγκεκριμένων περιορισμών κατά την εφαρμογή του. Με βάση τον βαθμό ελαστικότητας του μετασχηματισμού ευθυγράμμισης, γίνεται ένας διαχωρισμός σε τέσσερις υποκατηγορίες.

2.3.2.1 Συμπαγής Μετασχηματισμός τρισδιάστατων δεδομένων

Ο μετασχηματισμός αυτός έχει σαν γνώρισμα του το γεγονός ότι περιγράφει μόνο τις μετατοπίσεις και περιστροφές του αντικειμένου στο χώρο και όχι το σχήμα του αντικειμένου. Για την ανάλυση της περιστροφή ενός αντικειμένου γύρω από ένα άξονα (x) με γωνία ϕ (φ_x) χρειαζόμαστε την μήτρα περιστροφής (R_x) ως προς αυτό τον άξονα. Στην περίπτωση αυτή αναλύουμε την μήτρα περιστροφής ως προς όλους τους άξονες, έτσι η μήτρα περιστροφής (R) αποτελείται από το γινόμενο τριών μητρών περιστροφής (R_x, R_y, R_z) οι οποίες αντιπροσωπεύουν την περιστροφή γύρω από τον αντίστοιχο άξονα.

$$R = R_x \cdot R_y \cdot R_z$$
$$\Rightarrow R = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \varphi_x & -\sin \varphi_x \\ 0 & \sin \varphi_x & \cos \varphi_x \end{pmatrix} \begin{pmatrix} \cos \varphi_y & 0 & \sin \varphi_y \\ 0 & 1 & 0 \\ -\sin \varphi_y & 0 & \cos \varphi_y \end{pmatrix} \begin{pmatrix} \cos \varphi_z & -\sin \varphi_z & 0 \\ \sin \varphi_z & \cos \varphi_z & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (2.1)$$

Η μετατόπιση ως προς τους κύριους άξονες αναπαριστάται με το διάνυσμα:

$$\mathbf{d} = \begin{pmatrix} d_x \\ d_y \\ d_z \end{pmatrix} \quad (2.2)$$

Έτσι, στον συμπαγή μετασχηματισμό, ένα σημείο $\mathbf{r} = (x, y, z)$ μετασχηματίζεται στο σημείο $\mathbf{r}' = (x', y', z')$ σύμφωνα με την σχέση (14):

$$\mathbf{r}' = R \mathbf{r} + \mathbf{d} \quad (2.3)$$

Με βάση τις (2.1) και (2.2), η (2.3) γράφεται ως:

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = R \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} d_x \\ d_y \\ d_z \end{pmatrix} \quad (2.4)$$

2.3.2.2 Μετασχηματισμός τύπου affine

Ένας μετασχηματισμός ορίζεται ως μετασχηματισμός affine όταν για κάθε γραμμή/επιφάνεια σε μία εικόνα ο μετασχηματισμός αυτός την μετασχηματίζει σε γραμμή/επιφάνεια με ταυτόχρονη διατήρηση της σχέσεως παραλληλίας. Ο

μετασχηματισμός τύπου affine μετατρέπει όλα τα τρισδιάστατα σημεία σε ομογενείς συντεταγμένες:

$$\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \Rightarrow \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (2.5)$$

Όπως επίσης και οι διάφοροι μετασχηματισμοί (scaling, translation, rotation) μετατρέπονται σε affine μετασχηματισμούς:

Scaling:

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} s_x & 0 & 0 & 0 \\ 0 & s_y & 0 & 0 \\ 0 & 0 & s_z & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (2.6)$$

Translation:

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & d_x \\ 0 & 1 & 0 & d_y \\ 0 & 0 & 1 & d_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (2.7)$$

Rotation: ως προς τον άξονα (α) όπου μπορεί να είναι ο x, y ή z , και αντικαθιστούμε το R_α με μήτρα περιστροφής (R_x, R_y, R_z) που αντιστοιχεί στον άξονα που θέλουμε να περιστραφεί

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} & & & 0 \\ & R_\alpha & & 0 \\ & & & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (2.8)$$

Στην γενικότερη περίπτωση οι νέες συντεταγμένες μετά την εφαρμογή μετασχηματισμού affine δίνονται από:

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} a & b & c & d \\ e & f & g & h \\ i & j & k & l \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

$$\begin{aligned} x' &= ax + by + cz + d \\ y' &= ex + fy + gz + h \\ z' &= ix + jy + kz + l \end{aligned} \quad (2.9)$$

Ο μετασχηματισμός affine λόγω του ότι δεν διατηρεί τα μήκη και τις γωνίες μπορεί να αλλάζει το σχήμα του αντικειμένου που απεικονίζεται. Ο μετασχηματισμός τύπου affine εφαρμόζεται κυρίως σε εικόνες στερεών αντικειμένων οι οποίες έχουν υποστεί παραμόρφωση ή όταν δεν υπάρχουν επαρκείς πληροφορίες που αφορούν στον τρόπο ανάκτησης των εικόνων.

2.3.2.3 Προβολικός Μετασχηματισμός

Σε αντίθεση με τον μετασχηματισμό τύπου affine, ο προβολικός μετασχηματισμός, απεικονίζει οποιαδήποτε ευθεία γραμμή στην πρώτη εικόνα,

σε ευθεία γραμμή στην δεύτερη εικόνα χωρίς να είναι απαραίτητη η διατήρηση της παραλληλίας της. Επίσης γίνεται η χρήση των ομογενών συντεταγμένων.

$$\begin{aligned} \begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} &= \begin{bmatrix} a & b & c & d \\ e & f & g & h \\ i & j & k & l \\ m & n & p & q \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \\ x' &= \frac{ax + by + cz + d}{mx + ny + pz + q} \\ y' &= \frac{ex + fy + gz + h}{mx + ny + pz + q} \\ z' &= \frac{ix + jy + kz + l}{mx + ny + pz + q} \end{aligned} \quad (2.10)$$

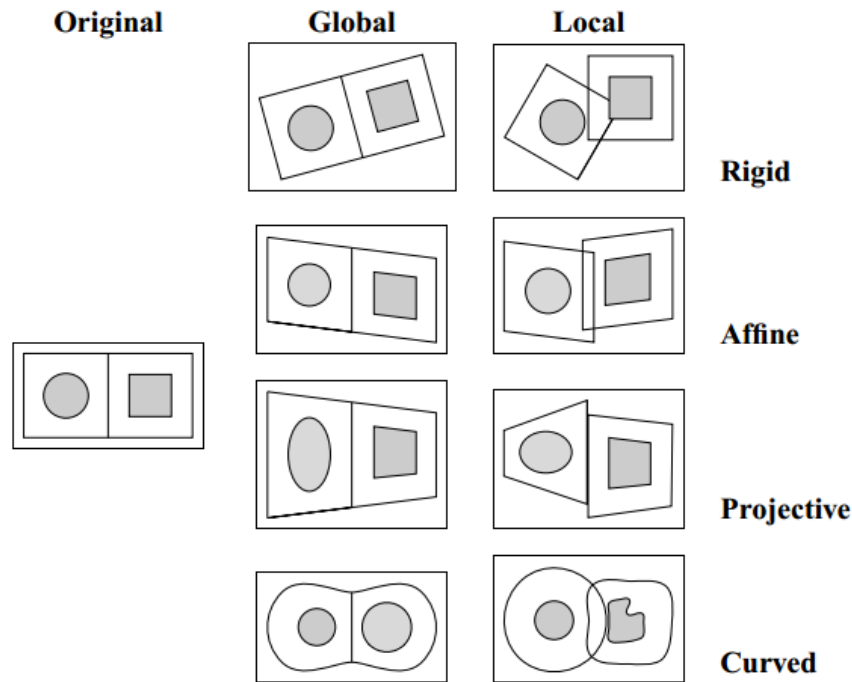
Ο προβολικός μετασχηματισμός επίσης αλλοιώνει το σχήμα του απεικονιζόμενου αντικειμένου. Είναι κατάλληλος για την ευθυγράμμιση εικόνων προβολών με σύνολα τομογραφικών δεδομένων τα οποία προβάλλονται στις δύο διαστάσεις εφαρμόζοντας σε αυτά ένα τρισδιάστατο μετασχηματισμό. Πέραν από την χρήση του για ευθυγράμμιση, χρησιμοποιείται επίσης ως ένας ελαστικός μετασχηματισμός με περιορισμούς όταν ο πλήρης ελαστικός μετασχηματισμός αποτυγχάνει ή παρουσιάζει μεγάλη πολυπλοκότητα λόγω του πλήθους των παραμέτρων του (9).

2.3.2.4 Καμπυλόγραμμοι Μετασχηματισμοί τρισδιάστατων δεδομένων

Οι καμπυλόγραμμοι μετασχηματισμοί σε αντίθεση με τους προηγούμενους μετασχηματισμούς περιγράφονται από την εξίσωση $(x', y', z') = \mathbf{F}(x, y, z)$ και με την μορφή σταθερών πινάκων. \mathbf{F} είναι η συνάρτηση η οποία μετασχηματίζει τις συντεταγμένες της πρώτης εικόνας σε συντεταγμένες της δεύτερης εικόνας. Χαρακτηριστικό γνώρισμα αυτών των μετασχηματισμών είναι ότι ευθείες γραμμές στην πρώτη εικόνα μετατρέπονται σε καμπύλες στην δεύτερη εικόνα. Ο πιο γνωστός καμπυλόγραμμικός μετασχηματισμός είναι ο μετασχηματισμός πολυωνυμικού τύπου:

$$\begin{aligned} x' &= \sum_{k=0}^K \sum_{l=0}^L \sum_{m=0}^M a_{klm} x^k y^l z^m \\ y' &= \sum_{k=0}^K \sum_{l=0}^L \sum_{m=0}^M b_{klm} x^k y^l z^m \\ z' &= \sum_{k=0}^K \sum_{l=0}^L \sum_{m=0}^M c_{klm} x^k y^l z^m \end{aligned} \quad (2.11)$$

όπου a, b, c παράμετροι και K, L, M σταθερές οι οποίες ορίζουν την τάξη των πολυωνύμων. Για να πετύχουμε αύξηση της ακρίβειας, θα πρέπει το πολυώνυμο να περιέχει περισσότερες παραμέτρους, άρα θα πρέπει να αυξηθεί η τάξη του πολυωνύμου. Με την αύξηση του πολυωνύμου όμως θα έχουμε και αύξησης του υπολογιστικού κόστους.



Εικόνα 2.2 : Παραδείγματα εφαρμογής δισδιάστατων ολικών και τοπικών μετασχηματισμών στην αρχική εικόνα.

2.3.3 Αυτοματοποίηση και διαδραστικότητα

Ανάλογα με τον βαθμό τον οποίο επεμβαίνει ο χρήστης στην διαδικασία της ευθυγράμμισης, οι μέθοδοι ευθυγράμμισης διαχωρίζονται σε τρεις κατηγορίες. Στην πρώτη κατηγορία, ο χρήστης επεμβαίνει με τον μικρότερο βαθμό και οι μέθοδοι που εμπεριέχονται σε αυτή την κατηγορία ονομάζονται αυτόματες μέθοδοι ευθυγράμμισης. Η μοναδική αλληλεπίδραση του χρήστη με τους αλγόριθμους ευθυγράμμισης είναι να εισάγει στον αλγόριθμο τα δεδομένα ή και τις πληροφορίες οι οποίες είναι σχετικές με την ανάκτηση των δεδομένων. Στη δεύτερη κατηγορία, ο χρήστης έχει τον μεγαλύτερο βαθμό αλληλεπίδρασης, όπου η ευθυγράμμιση γίνεται από το χρήστη και οι μέθοδοι που εμπεριέχονται σε αυτή την κατηγορία ονομάζονται διαδραστικές (10) ή χειροκίνητες μέθοδοι ευθυγράμμισης. Η τελευταία κατηγορία είναι οι μέθοδοι στις οποίες ο χρήστης πρέπει να επέμβει στην διαδικασία της ευθυγράμμισης σε μία από τις δύο περιπτώσεις. Η πρώτη περίπτωση είναι για να αρχικοποιήσει τον αλγόριθμο, π.χ. με κατάτμηση των δεδομένων, και η άλλη περίπτωση είναι ο χρήστης να καθοδηγήσει τον αλγόριθμο, π.χ. με την απόρριψη ή αποδοχή προτεινόμενων λύσεων. Οι μέθοδοι που εμπεριέχονται σε αυτή την κατηγορία ονομάζονται ημιαυτόματες μέθοδοι ευθυγράμμισης.

2.3.4 Τεχνικές Αναζήτησης

Υπάρχουν δύο τρόποι για τον υπολογισμό των παραμέτρων του μετασχηματισμού. Στη πρώτη περίπτωση, οι παράμετροι μπορούν να υπολογιστούν άμεσα, δηλαδή, καθορίζονται με σαφή τρόπο από τα διαθέσιμα δεδομένα, ή βάσει αναζήτησης, δηλαδή, καθορίζονται από την εύρεση ενός βέλτιστου από κάποια συνάρτηση που ορίζεται στο χώρο των παραμέτρων. Στην πρώτη περίπτωση, ο τρόπος υπολογισμού των παραμέτρων του μετασχηματισμού γίνεται με τη χρήση της προσέγγισης ελαχίστων τετραγώνων με επικρατούσα τεχνική την ανάλυση σε ανώμαλες ιδιοτιμές (singular value decomposition).

Στην περίπτωση της αναζήτησης με μεθόδους βελτιστοποίησης, οι μέθοδοι έχουν την δυνατότητα να διαμορφώνουν το μοντέλο σε μια τυπική μαθηματική συνάρτηση των παραμέτρων μετασχηματισμού με σκοπό τη βελτιστοποίηση της συνάρτησης αυτής. Αυτή η συνάρτηση επιχειρεί την ποσοτικοποίηση της ομοιότητας, όπως υπαγορεύεται από το μοντέλο μεταξύ δύο εικόνων, δεδομένου ενός μετασχηματισμού. Τέτοιες λειτουργίες είναι λιγότερο πολύπλοκες στη περίπτωση όπου οι δύο εικόνες προέρχονται από την ίδια απεικονιστική διάταξη, δεδομένου ότι η ομοιότητα είναι πιο εύκολο να προσδιοριστεί. Εάν όμως η συνάρτηση αυτή δεν έχει ομαλή μαθηματικά συμπεριφορά, θα πρέπει να γίνει εξαντλητική αναζήτηση σε όλο το εύρος τιμών της για την εύρεση των ιδανικών παραμέτρων όπου θα την βελτιστοποιούν. Για την αποφυγή αυτής της χρονοβόρας τεχνικής για εύρεση του ολικού μεγίστου ή ελαχίστου όπου θα βελτιστοποιήσει την συνάρτηση, οδηγούμαστε σε μια βελτιστοποίηση η οποία ψάχνει τοπικά ελάχιστα αντί του ολικού. Ξεκινώντας από τυχαία σημεία του χώρου των ανεξάρτητων μεταβλητών, βρίσκουμε ένα πλήθος από υψηλότερα τοπικά ελάχιστα. Επίσης μπορεί να εφαρμοστεί μια ακόμη μέθοδος στην οποία μετατοπίζεται η θέση ενός σημείου που αντιστοιχεί σε τοπικό ακρότατο εκτελώντας βήματα πεπερασμένου πλάτους γύρω από αυτό, ενώ παράλληλα ελέγχεται αν η συνάρτηση σε κάθε νέο σημείο επιστρέφει καλύτερη τιμή ή αν πάντα υπερσχύει το αρχικό ακρότατο (11).

Γενικά οι μέθοδοι ολικής βελτιστοποίησης, δεν παρουσιάζουν προβλήματα εγκλωβισμού σε τοπικά ακρότατα και βρίσκουν με ικανοποιητική ακρίβεια την περιοχή του ολικού ακρότατου. Επίσης, οι μέθοδοι τοπικής βελτιστοποίησης, αν έχουν μια καλή αρχική εκτίμηση του ολικού ακρότατου, συγκλίνουν με επιτυχία σε αυτό. Συμπερασματικά, για την απόκτηση με ακρίβεια του ολικού ακρότατου χωρίς προβλήματα εγκλωβισμού σε τοπικά ακρότατα, συνδυάζουμε τις δύο μεθόδους. Χρησιμοποιούμε τις μεθόδους ολικής βελτιστοποίησης για εύρεση της περιοχής του ολικού ακρότατου και τις μεθόδους τοπικής βελτιστοποίησης στην περιοχή αυτή για την εύρεση με ακρίβεια του ολικού ακρότατου.

Μέθοδοι ολικής βελτιστοποίησης είναι οι γενετικοί αλγόριθμοι (genetic algorithm), η μέθοδος απαγορευμένης αναζήτησης (tabu search), η μέθοδος διακλάδωσης και φραγής (branch and bound) και η μέθοδος της προσομοίωσης ανόπτησης, (12), (simulated annealing) όπου χρησιμοποιείτε σε συναρτήσεις με μεγάλο βαθμό μη γραμμικότητας και σε συναρτήσεις με οριακές συνθήκες. Μέθοδοι τοπικής βελτιστοποίησης είναι οι Downhill Simplex, (13), (14), η μέθοδος του Powell (13), η μέθοδος του Brent, η βελτιστοποίηση Levenberg-

Marquardt, ο επαναληπτικός αλγόριθμος Newton-Raphson, στοχαστικές μέθοδοι αναζήτησης και μέθοδος της καθοδικής κλίσης (gradient descent)

2.4 Κριτήρια Αξιολόγησης Μεθόδων Ευθυγράμμισης

2.4.1 Συνάρτηση Επιτυχίας

Οι συναρτήσεις προς βελτιστοποίηση οι οποίες αναφέρονται πιο πάνω θα πρέπει να οριστούν ανάλογα με τις ιδιότητες των εικόνων προς ευθυγράμμιση. Σε αυτό το υποκεφάλαιο αναφέρονται οι βασικές συναρτήσεις αξιολόγησης και οι περιπτώσεις όπου αυτές εφαρμόζονται. Οι συναρτήσεις βελτιστοποίησης μπορεί να είναι είτε συναρτήσεις ομοιότητας, ή συναρτήσεις μέτρησης απόστασης.

Οι συναρτήσεις ομοιότητας (O), στην ιδανική περίπτωση, γίνονται μέγιστες όταν οι δύο εικόνες είναι πλήρως ευθυγραμμισμένες και μειώνεται με την μεταξύ τους απομάκρυνση.

$$T = \arg \max_T [O(I_T(\mathbf{r}), I_R(\mathbf{r}))] \quad (2.12)$$

Όπου I_T η μετασχηματισμένη εικόνα, I_R εικόνα αναφοράς και $\mathbf{r} = (x, y, z)$ σημείο της εικόνας.

Από την άλλη, οι συναρτήσεις μέτρησης απόστασης, γίνονται ελάχιστες όταν οι δύο εικόνες είναι πλήρως ευθυγραμμισμένες και συνήθως υλοποιείται με την χρήση της προσέγγισης ελάχιστων τετραγώνων, (15), οπότε και ελαχιστοποιείται το μέσω τετραγωνικό σφάλμα:

$$MSE = \frac{1}{n} \sum_{i=1}^n \|\mathbf{r}'_i - T(\mathbf{r}_i)\|^2 \quad (2.13)$$

Οι πιο δημοφιλείς συναρτήσεις ομοιότητας παρουσιάζονται στον πιο κάτω πίνακα:

Συναρτήσεις Ομοιότητας

Συνάρτηση	Χρήση	Μαθηματικός Τύπος
Απόλυτη διαφορά	SAD Single-Modality	$S_{SAD}(I_R, I_T) = \sum_{\mathbf{r}} I_R(\mathbf{r}) - I_T(\mathbf{r}) $
Άθροισμα των Διαφορών τετραγώνων	SSD Single-Modality	$S_{SSD}(I_R, I_T) = \sum_{\mathbf{r}} I_R(\mathbf{r}) - I_T(\mathbf{r}) ^2$
Συσχέτιση	CC Single-Modality	$S_{CC}(I_R, I_T) = \sum_{\mathbf{r} \in \Omega} \frac{(I_R(\mathbf{r}) - \mathbb{E}[I_R(\mathbf{r})])(I_T(\mathbf{r}) - \mathbb{E}[I_T(\mathbf{r})])}{\sigma(I_R)\sigma(I_T)}$
Κανονικοποιημένη Συσχέτιση	NCC Single-Modality	$S_{NCC}(I_R, I_T) = \sum_{\mathbf{r} \in \Omega} \frac{I_R(\mathbf{r})I_T(\mathbf{r})}{\sqrt{\mathbb{E}[I_T(\mathbf{r})^2]}\sqrt{\mathbb{E}[I_R(\mathbf{r})^2]}}$
Κλίση Συσχέτισης	GC Single-Modality	$S_{GC}(I_R, I_T) = \frac{1}{d} \sum_{i=1}^d S_{CC} \left(\frac{\partial I_R}{\partial x_i}, \frac{\partial I_T}{\partial x_i} \right)$
Αμοιβαία Πληροφορία	MI Multi-Modality	$S_{MI}(I_R, I_T) = \sum_i \sum_j p_{I_R I_T}(i, j) \log \frac{p_{I_R I_T}(i, j)}{p_{I_R}(i, j)p_{I_T}(i, j)}$
Κανονικοποιημένη αμοιβαία πληροφορία	NMI Multi-Modality	$S_{NMI}(I_R, I_T) = \frac{2 S_{MI}(I_R, I_T)}{H(I_R) + H(I_T)}$

Πίνακας 2.1 : Πίνακας συναρτήσεων ομοιότητας με τις αντίστοιχες δυνατότητές τους και μαθηματικούς τύπους

Όπως ακριβώς διαφορετικές κατηγορίες μετασχηματισμών, είναι κατάλληλες για την μοντελοποίηση διαφορετικών γεωμετρικών παραμορφώσεων μεταξύ των εικόνων, έτσι και εδώ, διαφορετικές συναρτήσεις ομοιότητας χρησιμοποιούνται για τις διάφορες αλλοιώσεις στην ένταση μεταξύ των εικόνων. Επίσης, οι συναρτήσεις ομοιότητας μπορούν να κατηγοριοποιηθούν με βάση την καταλληλότητά τους σε ευθυγράμμιση με μία απεικονιστική τεχνική (mono-modality) ή σε ευθυγράμμιση με πολλές απεικονιστικές τεχνικές (multi-modality).

2.4.1.1 Απόλυτη διαφορά χρωματικών Πυκνοτήτων

Η απόλυτη διαφορά χρωματικών πυκνοτήτων περιγράφεται από την σχέση (16), (17), (18):

$$S_{SAD}(I_R, I_T) = \sum_{\mathbf{r}} |I_R(\mathbf{r}) - I_T(\mathbf{r})| \quad (2.14)$$

Η τεχνική αυτή όμως παρουσίαζε προβλήματα σε περιπτώσεις όπου οι εικόνες προέρχονταν από διαφορετικές απεικονιστικές τεχνικές ή αν τα δεδομένα παρουσίαζαν τοπικές χρωματικές μεταβολές. Επίσης ένα μειονέκτημα

αυτής της συνάρτησης ομοιότητας είναι το γεγονός ότι είναι ευαίσθητη στον θόρυβο αν αυτός δεν ακολουθεί κανονική κατανομή.

Παραλλαγή αυτής είναι το άθροισμα των διαφορών τετραγώνων:

$$S_{SSD}(I_R, I_T) = \sum_r |I_R(\mathbf{r}) - I_T(\mathbf{r})|^2 \quad (2.15)$$

Εφαρμογές αυτών των συναρτήσεων ομοιότητας παρουσιάζονται σε ακτινογραφία με εικόνες φλουροσκοπίας (19), οδοντιατρικές ακτινογραφίες (20) και Μαγνητική τομογραφία (21), (22)

2.4.1.2 Συντελεστής Συσχέτισης

Ο συντελεστής συσχέτισης των χρωματικών πυκνοτήτων περιγράφεται από την σχέση:

$$S_{CC}(I_R, I_T) = \sum_{r \in \Omega} \frac{(I_R(\mathbf{r}) - \mathbb{E}[I_R(\mathbf{r})])(I_T(\mathbf{r}) - \mathbb{E}[I_T(\mathbf{r})])}{\sigma(I_R)\sigma(I_T)} \quad (2.16)$$

$$\sigma(I_R) = \sqrt{\sum_r [I_R(\mathbf{r}) - \mathbb{E}[I_R(\mathbf{r})]]^2}, \sigma(I_T) = \sqrt{\sum_r [I_T(\mathbf{r}) - \mathbb{E}[I_T(\mathbf{r})]]^2} \quad (2.17)$$

Όπου:

$\mathbb{E}[I_R(\mathbf{r})]$, $\mathbb{E}[I_T(\mathbf{r})]$: μέση τιμή φωτεινότητας στην εικόνα αναφοράς και στην μετασχηματισμένη εικόνα αντίστοιχα.

Το εύρο τιμών που μπορεί να λάβει ο συντελεστής συσχέτισης είναι [0,1], όπου η τιμή του καθώς τείνει στην μονάδα, σημαίνει ότι οι δύο εικόνες συσχετίζονται περισσότερο. Μία σημαντική ιδιότητα του συντελεστή συσχέτισης είναι το ότι παραμένει αναλλοίωτος σε γραμμικούς μετασχηματισμούς, δηλαδή:

$$S_{CC}(\alpha_1 I_R + \beta_1, \alpha_2 I_T + \beta_2) = S_{CC}(I_R, I_T) \quad (2.18)$$

Αυτό καθιστά τον συντελεστή συσχέτισης μη-ευαίσθητο στην αλλαγή της φωτεινότητας ή/και στην αντίθεση των εικόνων.

Μια άλλη παραλλαγή του συντελεστή συσχέτισης είναι ο συντελεστής κανονικοποιημένης συσχέτισης:

$$S_{NCC}(I_R, I_T) = \sum_{r \in \Omega} \frac{I_R(\mathbf{r})I_T(\mathbf{r})}{\sqrt{\mathbb{E}[I_T(\mathbf{r})^2]\mathbb{E}[I_R(\mathbf{r})^2]}} \quad (2.19)$$

Επίσης μια άλλη παραλλαγή του συντελεστή συσχέτισης, είναι η κλίση της συσχέτισης που ορίζεται ως:

$$S_{GC}(I_R, I_T) = \frac{1}{d} \sum_{i=1}^d S_{CC} \left(\frac{\partial I_R}{\partial x_i}, \frac{\partial I_T}{\partial x_i} \right) \quad (2.20)$$

Εφαρμογές αυτών των συναρτήσεων ομοιότητας παρουσιάζονται σε Ευθυγράμμιση τρισδιάστατων δεδομένων CT, MR και PET εγκεφάλου (23), ευθυγράμμιση δεδομένων CT μέσω υπολογισμού ψηφιακά ανακατασκευασμένης ακτινογραφίας (24), ευθυγράμμιση ακτινογραφιών επιβεβαίωσης πεδίων ακτινοβολήσης (25), καθώς και για την ευθυγράμμιση δεδομένων αξονικής τομογραφίας και ακτινογραφικών φιλμ (26).

2.4.1.3 Συντελεστής αμοιβαίας πληροφορίας

Ο συντελεστής αμοιβαίας πληροφορίας είναι ο δείκτης ο οποίος μας δείχνει τον βαθμό εξάρτησης των υπό σύγκριση εικόνων I_R, I_T . Ο συντελεστής αυτός δίνεται από:

$$S_{MI}(I_R, I_T) = \sum_i \sum_j p_{I_R I_T}(i, j) \log \frac{p_{I_R I_T}(i, j)}{p_{I_R}(i, j)p_{I_T}(i, j)} \quad (2.21)$$

Όπου:

$$p_{I_R I_T}(i, j) = p(I_T = i, I_R = j) \text{ και } p_{I_R}(i, j) = p(I_R = i), p_{I_T}(i, j) = p(I_T = j)$$

Ο συντελεστής αυτός εξαρτάται από την υπό συνθήκη πιθανότητα $p_{I_R I_T}(i, j)$ η οποία υποδηλώνει το ποσό της πληροφορίας που η I_R περιέχει για την I_T .

Όπως φαίνεται και από τον πίνακα, ο συντελεστής αυτός είναι εφαρμόσιμος σε περιπτώσεις όπου οι δύο εικόνες προέρχονται από διαφορετικές απεικονιστικές τεχνικές. Χαρακτηριστικό αυτής της συνάρτησης ομοιότητας είναι το γεγονός ότι δεν επηρεάζεται από τον θόρυβο και από αλλαγές στην φωτεινότητα ή/και στην αντίθεση των εικόνων.

Μια παραλλαγή του συντελεστή αυτού είναι ο κανονικοποιημένος συντελεστής αμοιβαίας πληροφορίας ο οποίος δίνεται από την :

$$S_{NMI}(I_R, I_T) = \frac{2 S_{MI}(I_R, I_T)}{H(I_R) + H(I_T)} \quad (2.22)$$

Όπου:

$$\text{Εντροπία : } H(I) = \sum_i p_I(i) \log \frac{1}{p_I(i)} \quad (2.23)$$

Όπου $p_I(i)$ είναι η πιθανότητα εμφάνισης του $i \in I$ στοιχείου.

Εφαρμογές αυτών των συναρτήσεων ομοιότητας παρουσιάζονται σε ευθυγράμμιση ιατρικών δισδιάστατων δεδομένων (19), (27), και τρισδιάστατων δεδομένων (28), (29), (30), από διαφορετικά απεικονιστικά συστήματα, σε προβλήματα με συμπαγή μετασχηματισμό (31) καθώς και σε μη συμπαγείς μετασχηματισμούς (32), (33).

2.4.2 Αποτίμηση μεθόδων ευθυγράμμισης

Για τον προσδιορισμό του κατά πόσο μια μέθοδος είναι κατάλληλη για ευθυγράμμιση των συγκεκριμένων εικόνων θα πρέπει να μελετήσουμε το κατά πόσο η μέθοδος αυτή μας προσφέρει τα ακόλουθα χαρακτηριστικά

2.4.2.1 Ακρίβεια

Στην ιδανική περίπτωση, όταν σε μία μέθοδο εισάγουμε δύο εικόνες οι οποίες είναι οι ίδιες, θα περιμέναμε το αποτέλεσμα της μεθόδου να είναι ίσο με την μονάδα αν η συνάρτηση προς βελτιστοποίηση είναι συνάρτηση ομοιότητας και μηδέν αν είναι συνάρτηση μέτρησης απόστασης. Η οποιαδήποτε απόκλιση από αυτές τις δύο τιμές αντιστοίχως θεωρείται ως συστηματικό σφάλμα (precision) το οποίο εξαρτάται αποκλειστικά από το σύστημα ευθυγράμμισης.

Αιτίες της παρουσίας αυτών των σφαλμάτων είναι λόγο κίνησης του προς απεικόνιση αντικειμένου, σφάλματα στην υλοποίηση της βελτιστοποίησης της συνάρτησης ή ακόμη και στο συνολικό σύστημα ευθυγράμμισης. Μια επίσης σημαντική αιτία είναι η ευθυγράμμιση αντικειμένων που έχουν πραγματικές διαφορές, όπως για παράδειγμα η ευθυγράμμιση δύο οδοντιατρικών εικόνων από τον ίδιο ασθενή, ο οποίος έχει υποστεί κάποια επέμβαση στην απεικονιζόμενη περιοχή. Τότε, θα υπάρχει πάντα σφάλμα, ακόμα και αν η ευθυγράμμιση έχει πλήρη επιτυχία. Η μέτρηση του συστηματικού σφάλματος, γίνεται με χρήση εξωγενών οδηγών σημείων κατά την ανάκτηση της εικόνας.

Ένα δεύτερο σφάλμα το οποίο μπορεί να εισέλθει στην υλοποίηση της ευθυγράμμισης είναι το άμεσο σφάλμα (accuracy) το οποίο μπορεί να είναι ποιοτικό ή ποσοτικό και το οποίο αναλόγως με την περίπτωση διαφοροποιείται. Για τον υπολογισμό του ποιοτικού σφάλματος απαιτείται χρήση απλών εργαλείων οπτικοποίησης. Σε αντίθεση, για τον υπολογισμό του ποσοτικού σφάλματος θα πρέπει να γίνει χρήση ενός προτύπου αναφοράς το οποίο δεν υπάρχει στην κλινική πράξη έτσι δεν μπορεί να αποτιμηθεί εκτός αν οριστεί ως προς κάποιο κριτήριο.

2.4.2.2 Ευρωστία και Ευστάθεια

Στην ευθυγράμμιση, η ευρωστία/ευστάθεια είναι η ιδιότητα όπου όταν στην είσοδο του συστήματος εισάγονται μικρές διαταραχές, θα πρέπει στην έξοδο να λαμβάνουμε επίσης μικρές διαταραχές. Για τον έλεγχο της ευρωστίας/ευστάθειας του συστήματος, τοποθετούμε δύο εικόνες οι οποίες βρίσκονται αρχικά σε διαφορετική θέση και βρίσκουμε τις παραμέτρους της ευθυγράμμισης. Στην συνέχεια μετακινούμε ελαφρά την μία εικόνα και βρίσκουμε επίσης τις παραμέτρους της ευθυγράμμισης. Τώρα συγκρίνοντας τις παραμέτρους των δύο ευθυγραμμίσεων που έγιναν, ορθής και μετακινημένης, μπορούμε να πάρουμε μία ιδέα για την ευρωστία του συστήματος.

2.4.2.3 Αξιοπιστία

Ένας αλγόριθμος θεωρείται αξιόπιστος όταν εφαρμοζόμενος σε διαφορετικά σύνολα δεδομένων παράγει αποτελέσματα με σφάλμα εντός του αποδεκτού εύρους τιμών.

2.4.2.4 Απαιτήσεις και Ανάγκες

Ένας αλγόριθμος θα πρέπει να αξιολογηθεί κατά πόσο η κλινική έρευνα και το όφελος που προσφέρει, είναι ανάλογα των απαιτήσεων του, της χρήσης των συστημάτων που χρησιμοποιεί και της δυσκολίας υλοποίησης του

2.4.2.5 Αλγοριθμική πολυπλοκότητα

Στις περιπτώσεις όπου ο χρόνος εκτέλεσης του αλγορίθμου συμβαδίζει με την αλγοριθμική πολυπλοκότητα, ο χρόνος που απαιτείται για να φέρει σε

πέρας την λειτουργία του θα πρέπει να είναι μέσα σε ικανοποιητικά όρια τα οποία τίθενται από το κλινικό περιβάλλον.

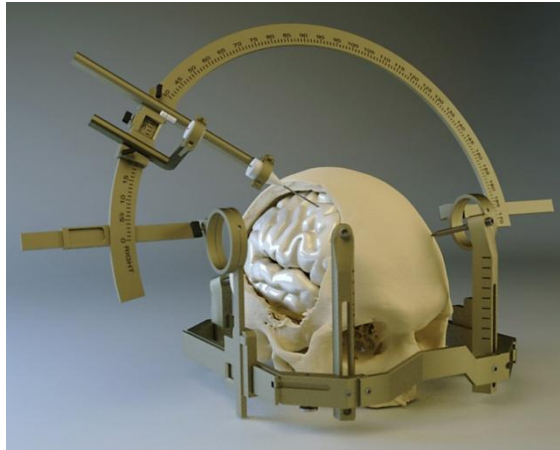
2.4.2.6 Κλινική Χρησιμότητα

Θα πρέπει να αξιολογηθεί κατά πόσο η μέθοδος η οποία μελετάται, όχι μόνο έχει κάποια κλινική χρησιμότητα, αλλά και τι περισσότερο προσφέρει από ήδη υπάρχουσες τεχνικές και μεθοδολογίες.

2.5 Είδος Ευθυγράμμισης

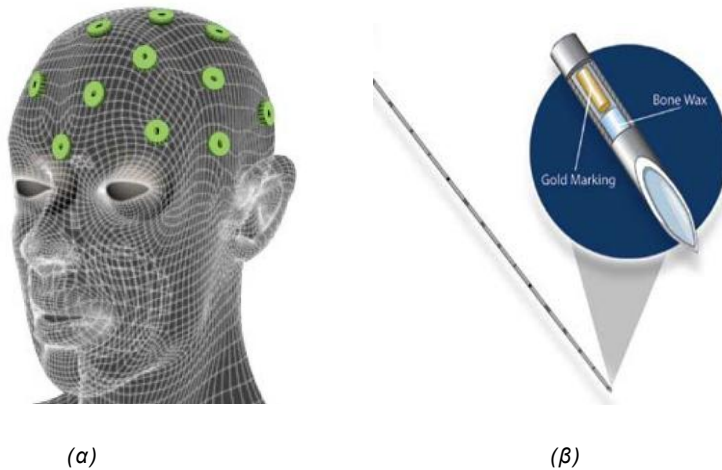
2.5.1 Ευθυγράμμιση βασισμένη σε εξωγενή χαρακτηριστικά

Οι μέθοδοι με βάση εξωγενή χαρακτηριστικά βασίζονται σε τεχνητά αντικείμενα τα οποία συνδέονται με τον ασθενή και τα οποία έχουν σχεδιαστεί έτσι ώστε να είναι ορατά και με μεγάλη ακρίβεια ανιχνεύσιμα σε όλες τις πιθανές θέσεις όπου ο ασθενής μπορεί να βρίσκεται. Έτσι, η εύρεση σημείων, τα οποία θα καθορίσουν τον μετασχηματισμό για ευθυγράμμιση, είναι συγκριτικά εύκολη, γρήγορη, είναι εφικτή και δεδομένου ότι οι παράμετροι της ευθυγράμμισης μπορούν να υπολογιστούν με σαφήνεια, δεν απαιτούνται πολύπλοκοι αλγόριθμοι βελτιστοποίησης. Η τρόπος τοποθέτησης των διατάξεων, χωρίζει το κριτήριο αυτό σε δύο υποκατηγορίες. Στην πρώτη κατηγορία, είναι η επεμβατική τοποθέτηση του αντικειμένου η οποία μας δίνει περισσότερη ακρίβεια και η δεύτερη υποκατηγορία είναι η μη επεμβατική η οποία μας δίνει λιγότερη πληροφορία. Μια πολύ συχνή εφαρμογή ευθυγράμμισης βασισμένης με εξωγενή χαρακτηριστικά γίνεται με το *stereotactic frame* το οποίο χρησιμοποιείται ως οδηγός ο οποίος βιδώνεται σταθερά στην εξωτερική επιφάνεια του κρανίου του ασθενή. Οι συντεταγμένες των σημάδιων που προσδιορίζονται με την απεικόνιση του κεφαλιού σε τομές, μπορεί να εμπεριέχουν σφάλματα λόγω διαταραχών και είναι αναγκαία η προεπεξεργασία των εικόνων για την διόρθωσή τους. Ακολούθως, οι συντεταγμένες των σημάδιων χρησιμοποιούνται για τον υπολογισμό του μετασχηματισμού του συνολικού όγκου του κεφαλιού σε σχέση με το σύστημα αναφοράς του τρισδιάστατου στερεοστατικού πλαισίου (34), (35). Σε αυτές τις περιπτώσεις χρησιμοποιούμε μετασχηματισμό *affine* και διγραμμικό μετασχηματισμό. Η τεχνική αυτή μέχρι πρόσφατα είχε τον τίτλο «χρυσός κανόνας» για την ακρίβεια που παρείχε, αφού η ακρίβεια της είναι της τάξης του 1mm στο εσωτερικό μιας τομής για την περίπτωση των υπολογιστικών και μαγνητικών τομογραφιών, και περίπου ίση με το μισό του πάχους των τομών κάθετα στο επίπεδο σάρωσης.



Εικόνα 2.3 :Ευθυγράμμιση βασισμένη σε εξωγενή χαρακτηριστικά με την βοήθεια του στερεοστατικού πλαισίου

Τέτοιου είδους πλαίσια χρησιμοποιούνται για τον εντοπισμό και την καθοδήγηση στην νευροχειρουργική. Επίσης, ένα άλλο αντικείμενο που χρησιμοποιείται σαν οδηγός και εφαρμόζεται με επέμβαση είναι τα fiducials markers.



(α)

(β)

Εικόνα 2.4 : (α) Σημεία στην περιοχή της κεφαλής του ασθενή στα οποία θα εισαχθούν fiducials markers

(β) Μηχανισμός τοποθέτησης των fiducials markers

Μη επεμβατική τεχνική είναι και η τοποθέτηση δεικτών πάνω στο δέρμα του ασθενή. Συνήθως το πλήθος των σημαδιών είναι μικρό με συνέπεια ο μετασχηματισμός που υπολογίζεται με βάση αυτά να είναι ολικός και συμπαγής ή τύπου affine. Άλλες μη επεμβατικές τεχνικές με μεγαλύτερες συσκευές που εφαρμόζονται πάνω στον ασθενή είναι τα εξατομικευμένα καλούπια αφρού (foam molds), καλούπια συγκράτησης κεφαλής και οδοντιατρικοί προσαρμοστές (10).

2.5.2 Ευθυγράμμιση βασισμένη σε ενδογενή χαρακτηριστικά

Οι μέθοδοι με χρήση ενδογενών χαρακτηριστικών στηρίζονται μόνο στο περιεχόμενο της εικόνας όπου δημιουργείται από τον ασθενή (36). Επίσης αποκαλούνται ανασκοπικές για τον λόγο ότι δεν χρειάζονται κάποια προφύλαξη κατά την στιγμή της λήψης της εικόνας. Αυτές οι τεχνικές διαχωρίζονται σε τρεις επιμέρους κατηγορίες.

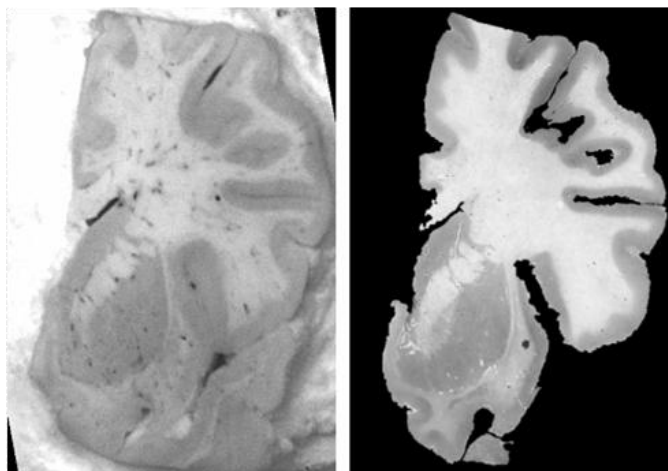
2.5.2.1 Voxel property-based

Η πρώτη κατηγορία αφορά την ευθυγράμμιση με βάση τα εικονοστοιχεία. Οι τεχνικές που ανήκουν στην κατηγορία αυτή ξεχωρίζουν από τις άλλες μεθόδους με ενδογενή χαρακτηριστικά από το γεγονός ότι λειτουργούν άμεσα στις τιμές του γκρι της εικόνας, χωρίς προηγούμενη αναγωγή των δεδομένων από τον χρήστη ή κατάτμηση. Υπάρχουν δύο διαφορετικές προσεγγίσεις: η πρώτη είναι να μειώσει αμέσως τα επίπεδα του γκρι στα περιεχόμενα της εικόνας σε ένα αντιπροσωπευτικό σύνολο διανυσμάτων και προσανατολισμών και η δεύτερη είναι να χρησιμοποιήσει το πλήρες περιεχόμενο της εικόνας σε όλη τη διαδικασία ευθυγράμμισης.

Μέσα σε αυτές τις μεθόδους το κέντρο βαρύτητας της εικόνας και οι κύριοι άξονες της υπολογίζονται από τις πρώτες μόλις απεικονίσεις. Εν συνεχεία η ευθυγράμμιση των εικόνων πραγματοποιείται με την ευθυγράμμιση του κέντρου βάρους και των κύριων αξόνων τους. Το αποτέλεσμα δεν είναι συνήθως πολύ ακριβές, και η μέθοδος δεν είναι κατάλληλη για να χειριστεί διαφορές στον σαρωμένο όγκο. Παρά τα μειονεκτήματά της, χρησιμοποιείται ευρέως σε προβλήματα ευθυγράμμισης που δεν απαιτούν υψηλή ακρίβεια, λόγω της αυτόματης και πολύ γρήγορης φύσης της, καθώς και την εύκολη εφαρμογή της. Η μέθοδος αυτή χρησιμοποιείται κυρίως στην επανευθυγράμμιση των μελετών καρδιακών σπινθηρογραφημάτων. Οι μέθοδοι της συγκεκριμένης κατηγορίας χρησιμοποιούν σαν είσοδό τους, κατακερματισμένες εικόνες ή εικόνες με δυαδικά δεδομένα. Σε πολλές εφαρμογές της μεθόδου, είναι αναγκαία η κατάτμηση της εικόνας πριν την οποιαδήποτε άλλη ενέργεια για να παράγει αποδεκτά αποτελέσματα.

Μέθοδοι που χρησιμοποιούν το πλήρες περιεχόμενο της εικόνας είναι οι πιο ενδιαφέρουσες μέθοδοι που ερευνούνται σήμερα. Θεωρητικά, αυτές είναι οι πιο ευέλικτες μέθοδοι ευθυγράμμισης, για τον λόγο ότι σε αντίθεση με όλες τις άλλες μεθόδους που αναφέρονται, δεν ξεκινούν με τη μείωση του επιπέδου του γκρι της εικόνας σε σχετικά αραιές εξαγόμενες πληροφορίες, αλλά χρησιμοποιούν όλο τον διαθέσιμο όγκο πληροφοριών κατά την διαδικασία ευθυγράμμισης. Αν και οι μέθοδοι αυτοί είναι αρκετά παλιές, η χρήση τους σε εκτεταμένες κλινικές εφαρμογές 3D/3D ήταν αρχικά περιορισμένη λόγω του υψηλού υπολογιστικού τους κόστους. Όμως οι αυξανόμενες απαιτήσεις διάφορων ιατρικών εφαρμογών για ευθυγράμμιση ιατρικών δεδομένων με μεγάλη ακρίβεια, καθώς και η ανάπτυξη ολοένα και πιο γρήγορων υπολογιστών

με μεγάλες εσωτερικές μνήμες, επέτρεψαν στις μεθόδους επεξεργασίας πλήρους περιεχομένου της εικόνας να χρησιμοποιούνται στην κλινική πρακτική.



Εικόνα 2.5 : *Voxel property-based*, κρυστομή, ιστολογική τομή μυελίνης

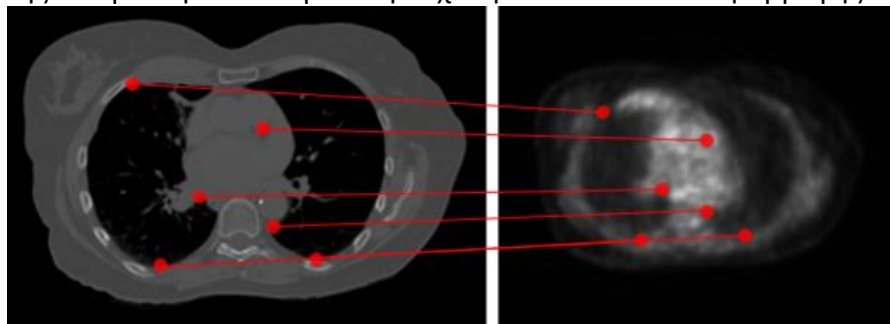
2.5.2.2 Landmark property-based

Η τελευταία κατηγορία είναι οι μέθοδοι που βασίζονται σε σημεία οδηγούς/ορόσημα. Τα ορόσημα μπορεί να είναι ανατομικά, δηλαδή εμφανή και με ακρίβεια καθορισμένα σημεία της μορφολογίας της ορατής ανατομίας τα οποία προσδιορίζονται συνήθως διαδραστικά από τον χρήστη, ή γεωμετρικά, δηλαδή τα σημεία τα οποία βρίσκονται στην θέση ενός βέλτιστου γεωμετρικού χαρακτηριστικού, π.χ. τοπική καμπυλότητα, ακρότατα, γωνίες, κλπ, και γενικά εντοπίζονται με αυτόματο τρόπο.

Οι τεχνικές αυτές είναι ευέλικτες, με την έννοια ότι τουλάχιστον στη θεωρία μπορούν να εφαρμοστούν σε οποιαδήποτε εικόνα χωρίς να παίζει ρόλο ποιο είναι το αντικείμενο όπου θα απεικονιστεί. Οι συγκεκριμένοι μέθοδοι ως επί το πλείστον χρησιμοποιούνται για την εύρεση του συμπαγή ή τύπου affine μετασχηματισμού. Αν τα σύνολα των σημείων είναι αρκετά μεγάλα, μπορούν θεωρητικά να χρησιμοποιηθούν και για πιο πολύπλοκους μετασχηματισμούς. Τα ανατομικά ορόσημα χρησιμοποιούνται συχνά με μια εντελώς διαφορετική βάση ευθυγράμμισης: μέθοδοι που βασίζονται σε βελτιστοποίηση του παραμετροποιημένου χώρου όπου δεν είναι κυρτός, είναι επιρρεπείς μερικές φορές να κολλήσουν σε τοπικά βέλτιστα, το οποίο ενδέχεται να οδηγήσει σε μεγάλη αστοχία. Με τον περιορισμό του χώρου αναζήτησης, σύμφωνα με ανατομικά σημεία, τέτοιες αστοχίες είναι λιγότερο πιθανό να συμβούν. Επιπλέον, η διαδικασία αναζήτησης μπορεί να επιταχυνθεί σημαντικά. Ένα μειονέκτημα είναι ότι στις πλείστες περιπτώσεις απαιτείται η αλληλεπίδραση του χρήστη για την αναγνώριση των οδηγών σημείων .

Στις μεθόδους τις συγκεκριμένης κατηγορίας, το σύνολο των προσδιορισμένων σημείων είναι πιο αραιό σε σύγκριση με το αρχικό περιεχόμενο της εικόνας, έτσι καθίστανται ιδανικές για σχετικά γρήγορες διαδικασίες βελτιστοποίησης. Τέτοιου είδους αλγόριθμοι ελαχιστοποιούν τη μέτρηση της μέσης απόστασης ανάμεσα σε κάθε ορόσημο και το πιο κοντινό ομόλογό του. Για τη βελτιστοποίηση της προηγούμενης μέτρησης, ο

επαναληπτικός αλγόριθμος πλησιέστερου σημείου (ICP) είναι ο πιο δημοφιλής. Η δημοτικότητά του οφείλεται στην ευελιξία του να μπορεί να χρησιμοποιηθεί για σύνολα δεδομένων, και με βάση αυτού μπορούν να οριστούν άμεσα ή έμμεσα καμπύλες, επιφάνειες και όγκοι για το υπό εξέταση σύνολο σημείων, όπως επίσης και για την υπολογιστική ταχύτητα και ευκολία εφαρμογής του.



Εικόνα 2.6 : Landmark based, εικόνα από CT-PET

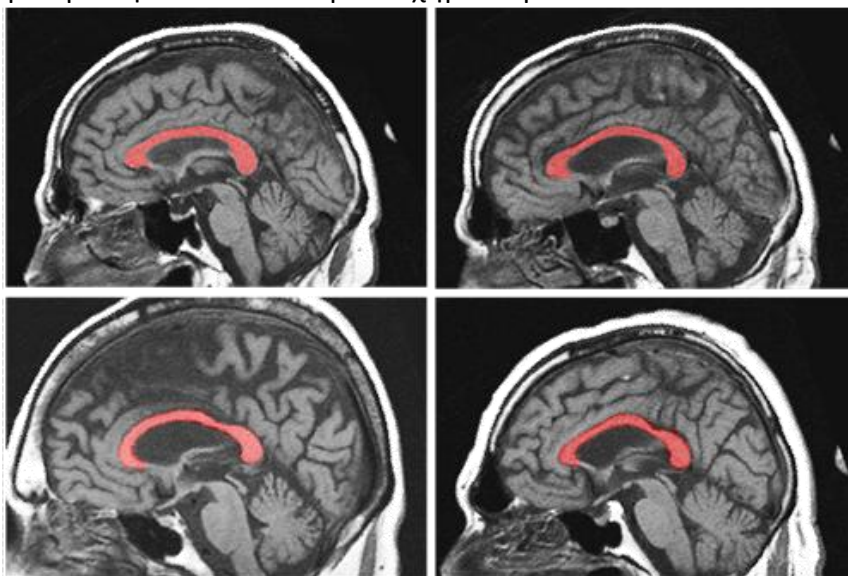
2.5.2.3 Segmentation property-based

Η τελευταία κατηγορία είναι οι μέθοδοι που βασίζονται σε τμήματα της εικόνας. Οι μέθοδοι αυτοί χωρίζονται σε δύο υποκατηγορίες. Η πρώτη περίπτωση είναι να βασίζονται σε ένα άκαμπτο μοντέλο, όπου ανατομικά ίδιες δομές (κυρίως επιφάνειες) εξάγονται από τις δύο εικόνες για να ευθυγραμμιστούν και χρησιμοποιείται ως μοναδική είσοδος για τη διαδικασία ευθυγράμμισης. Η δεύτερη περίπτωση είναι να βασίζονται σε ένα εύκαμπτο μοντέλο, όπου μια δομή εξάχθηκε (επίσης κυρίως επιφάνειες και καμπύλες) από μία εικόνα και παραμορφώνεται ελαστικά για να προσαρμοστεί στη δεύτερη εικόνα.

Οι μέθοδοι της πρώτης περίπτωσης είναι πιθανώς οι πιο δημοφιλείς μέθοδοι που χρησιμοποιούνται κλινικά σήμερα. Καθώς η κατάτμηση των πλείστων εικόνων είναι μια εύκολη επεξεργασία για να εκτελεστεί, και η υπολογιστική πολυπλοκότητα είναι σχετικά χαμηλή (εξαρτάται από την μέθοδο, για παράδειγμα σε μεθόδους όπως το region growing η κατάτμηση μπορεί να πάρει περισσότερο χρόνο από την ευθυγράμμιση), η μέθοδος έχει παραμείνει δημοφιλής, και πάρα πολλές έρευνες έχουν δημοσιευθεί για την αυτοματοποίηση του τμήματος της κατάτμησης, ή αλλιώς την επέκταση της μεθόδου. Ένα μειονέκτημα των μεθόδων τμηματοποίησης είναι ότι η ακρίβεια εγγραφής περιορίζεται ανάλογα με την ακρίβεια του σταδίου τμηματοποίησης.

Στη δεύτερη κατηγορία, το κριτήριο βελτιστοποίησης είναι διαφορετικό: ορίζεται και υπολογίζεται πάντα τοπικά, και η παραμόρφωση περιορίζεται από τους περιορισμούς του ελαστικού μοντέλου (υπό ένα όρο κανονικοποίησης) όπου επιβάλλονται επάνω στην εξαγμένη καμπύλη ή επιφάνεια. Για να διευκολυνθεί η φυσική μοντελοποίηση, η δομή δεδομένων του παραμορφωμένου μοντέλου δεν λαμβάνεται ως ένα σύνολο σημείων. Αντί αυτού, η δομή δεδομένων του παραμορφωμένου μοντέλου συχνά αναπαρίσταται ως ανιχνεύσιμες συναρτήσεις όπως σφήνες (splines). Η διαδικασία παραμόρφωσης γίνεται πάντοτε επαναληπτικά, με μικρές παραμορφώσεις σε κάθε στιγμή. Οι μέθοδοι αυτοί βασίζονται σε ένα πρότυπο

μοντέλο που πρέπει να οριστεί σε μια εικόνα. Μετά από αυτό, μπορούν να προσδιοριστούν δύο είδη προσεγγίσεων: το πρότυπο είτε παραμορφώνεται για να ταιριάζει με μια κατακερματισμένη δομή στη δεύτερη εικόνα, ή η δεύτερη εικόνα χρησιμοποιείται ως έχει χωρίς τμηματοποίηση. Ένα μειονέκτημα των μεθόδων αυτών είναι ότι συχνά χρειάζονται μια καλή αρχική θέση, προκειμένου να συγκλίνουν σωστά, η οποία μπορεί να επιτευχθεί μέσω κατάλληλης προεπεξεργασίας των υπό ευθυγράμμιση εικόνων. Ένα άλλο μειονέκτημα είναι ότι η τοπική παραμόρφωση του προτύπου μπορεί να είναι απρόβλεπτα ασταθής εάν η δομή στόχος διαφέρει επαρκώς από τη δομή του προτύπου. Τα ελαστικά μοντέλα είναι τα πλέον κατάλληλα για την εύρεση των τοπικών μετασχηματισμών καμπυλών μεταξύ των εικόνων, και λιγότερο για την εύρεση (ολικών)συμπαγών ή τύπου affine μετασχηματισμών.



Εικόνα 2.7 : Segmentation Based, εικόνα από το μεσολόβιο

2.6 Διαδικασία Σύνθεσης Ιατρικών Δεδομένων

Η διαδικασία της σύνθεσης των δύο συνόλων από δεδομένα, αποτελεί το τελικό στάδιο της επεξεργασίας τους με σκοπό την παρουσίαση των ευθυγραμμισμένων πλέον συνόλων σε ένα κοινό οπτικό πεδίο. Πιο κάτω παρουσιάζονται οι δημοφιλέστεροι τρόποι σύνθεσης εικόνων.

2.6.1 Με χρήση λογικών τελεστών

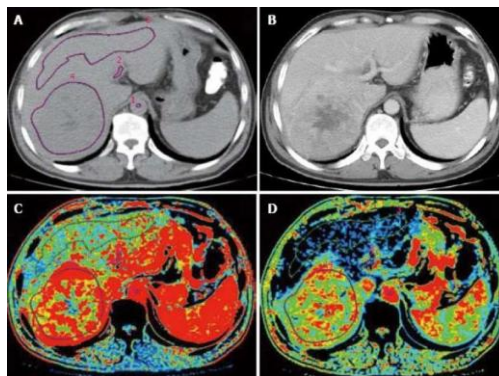
Υπάρχουν μεθοδολογίες σύνθεσης εικόνας με την χρήση λογικών τελεστών. Ας υποθέσουμε ότι η μία εικόνα είναι η εικόνα αναφοράς I_R η οποία δεν επεξεργάζεται και η δεύτερη εικόνα η οποία είναι η προς ευθυγράμμιση εικόνα I_T . Τότε από την δεύτερη εικόνα καθορίζεται μια περιοχή ενδιαφέροντος και οι πληροφορίες από τις δύο εικόνες μετά συνδυάζονται. Μια απλή μέθοδος συνδυασμού των πληροφοριών των δύο εικόνων είναι με τον λογικό τελεστή XOR, με βάση την εξίσωση:

$$I_{COMBINED}(\mathbf{r}) = I_R(\mathbf{r})(1 - M(\mathbf{r})) + I_T(\mathbf{r})M(\mathbf{r}) \quad (2.24)$$

Όπου $M(\mathbf{r})$ είναι η μάσκα Βοoleanόπου προσδιορίζει το κάθε pixel το οποίο αντικαθίσταται από την προς ευθυγράμμιση εικόνα $I_T(\mathbf{r})$ στην συνδυασμένη εικόνα $I_{COMBINED}(\mathbf{r})$

2.6.2 Με χρήση χάρτη ψευδό-χρώματος

Με βάση αυτή την τεχνική σύνθεσης, η ευθυγραμμισμένη εικόνα δημιουργείται με την χρήση κλίμακας ψευδοχρώματος και με διαφάνεια επικαλύπτει την εικόνα αναφοράς. Ο χάρτης ψευδοχρώματος καθορίζεται με βάση την τριπλέτα R,G,Bσε κάθε διακριτή τιμή pixel.



2.6.3 Τεχνικές ταξινόμησης

Σε αυτές τις τεχνικές, η σύνθεση πραγματοποιείται με την επεξεργασία και των δύο εικόνων με σκοπό την δημιουργία μιας συντηγμένης εικόνας με κατάλληλο pixelταξινόμησης. Η μέθοδος αυτή χρησιμοποιεί ένα διπλό ιστογράμμο $P(x, y)$ των δύο ευθυγραμμισμένων εικόνων το οποίο ορίζεται ως η πιθανότητα του pixel (i, j) να έχει τιμή y στην υπό ευθυγράμμιση εικόνα, δεδομένου ότι το ίδιο pixelέχει τιμή x στην εικόνα αναφοράς:

$$P(x, y) = P(I_T(i, j) = y | I_A(i, j) = x) \quad (2.25)$$

3 Συλλογή 3D νεφών σημείων (Point Cloud)

3.1 3D Point Cloud

Στον τομέα της όρασης υπολογιστών, σε πολλές περιπτώσεις η τρισδιάστατη απεικόνιση αντικειμένων βασίζεται σε σύνολα σημείων για την εξαγωγή των απαραίτητων πληροφοριών. Βάσει αυτών των πληροφοριών και με κατάλληλους αλγόριθμους, μας δίνεται η δυνατότητα για την αναπαραγωγή του φυσικού αντικειμένου σε ένα ψηφιακό τρισδιάστατο επίπεδο. Τα σημεία τα οποία σχετίζονται με το συγκεκριμένο μοντέλο προσδιορίζονται ως προς τους βασικούς άξονες Χ,Υκαι Ζ. Όλες οι συντεταγμένες των σημείων που σχετίζονται με ένα μοντέλο, περιλαμβάνονται σε μία ιδέα/οντότητα η οποία ονομάζεται νέφος σημείων (point cloud). Το νέφος σημείων περιέχει όλες τις συντεταγμένες οι οποίες βοηθούν στην επεξήγηση και αναπαράσταση της εξωτερικής και εσωτερικής επιφάνειας ενός τρισδιάστατου αντικειμένου ανάλογα με την μέθοδο λήψης των σημείων.

Βιομηχανίες όπως η κατασκευαστική, χρησιμοποιεί το νέφος σημείων για αντίστροφη μηχανική (reverse engineer), διασφάλιση της ποιότητας, παραμετροποίηση και σχεδίαση προϊόντων. Χρησιμοποιώντας υπολογιστικές τεχνικές (computer-aided design - CAD), ένα φυσικό αντικείμενο μπορεί να σαρωθεί και να αναλυθεί για την διαπίστωση της ακρίβειας ή τυποποίησής του. Στην ιατρική τεχνολογία, μπορούν να βοηθήσουν στην προσαρμογή των ιατρικών συσκευών για να καλύψουν τις ανάγκες των διάφορων ασθενών. Στην ιατρική απεικόνιση χρησιμοποιούνται τα νέφη σημείων, π.χ. στην τρισδιάστατη αξονική τομογραφία (CAT) για την συμπίεση δεδομένων και την μετατροπή τους σε τρισδιάστατα ή τετραδιάστατα υπολογιστικά μοντέλα.

Η τρισδιάστατη ανακατασκευή ενός αντικειμένου δεν είναι η μόνη χρήση του νέφους σημείων. Όταν απαθανατίζεται η επιφάνεια ενός αντικειμένου, μπορούν να αποθηκευτούν και άλλες πληροφορίες στο νέφος σημείων όπως είναι το χρώμα της επιφάνειας, η υφή της, κλπ, δημιουργώντας έτσι ένα νέφος σημείων μεγαλύτερης διάστασης. Αυτή η επιπρόσθετη πληροφορία επιτρέπει στον χρήστη να έχει μεγαλύτερη ευελιξία όσον αφορά την επεξεργασία των ανακατασκευασμένων εικόνων και επιφανειών. Η ανακατασκευή μιας επιφάνειας με ένα νέφος σημείων το οποίο εμπεριέχει και το χρώμα της επιφάνειας, δίνει την δυνατότητα μιας πιο ρεαλιστικής ανακατασκευής του αντικειμένου στον τρισδιάστατο ψηφιακό χώρο.

Για την συλλογή ενός νέφους σημείων από την σάρωση ενός αντικειμένου, χρειάζεται ένα εξειδικευμένο εργαλείο ώστε να παρέχονται οι αναλυτικές πληροφορίες που απαιτούνται για την ανακατασκευή.

3.2 Απεικόνιση Βάθους

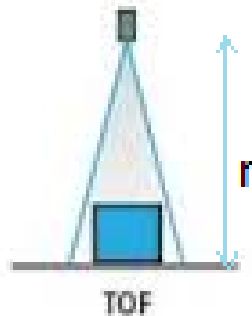
Η απεικόνιση βάθους είναι η διαδικασία σχηματισμού μιας εικόνας στην οποία το χαρακτηριστικό όπου ενδιαφέρει τον χρήστη είναι η απόσταση των σημείων από την

κάμερα. Πρέπει να αναφέρουμε ότι στην απεικόνιση βάθους, η τιμή του κάθε εικονοστοιχείου (pixel) είναι η απόσταση του σημείου του αντικειμένου που αντιπροσωπεύεται από το συγκεκριμένο pixel σε σχέση με το σύστημα λήψης της απεικόνισης. Ο τρόπος σχηματισμού αυτών των εικόνων βάθους είναι ένα κριτήριο διαχωρισμού των τεχνικών απόκτησης δεδομένων βάθους. Οι βασικότερες κατηγορίες με βάση τον τρόπο σχηματισμού των εικόνων βάθους είναι οι μέθοδοι με το χρόνο πτήσης (time-of-flight) και του δομημένου φωτός (structured light) όπου στηρίζεται στην στερεοσκοπική τεχνική.

3.2.1 Time of Flight

Η βασική ιδέα πίσω από αυτή την τεχνική απεικόνισης βάθους είναι ο χρόνος που χρειάζεται το φως να διανύσει μία άγνωστη απόσταση(42). Συσκευές με αυτή την τεχνική απεικόνισης βάθους, εκπέμπουν ένα παλμό φωτός, ορατό ή μη, σε μία γνωστή χρονική στιγμή t_1 , και μετράνε το χρόνο t_2 που κάνει το φως για να ανιχνευτεί από ένα αισθητήρα. Έχοντας ως γνωστό το μέτρο της ταχύτητας του φωτός, c , τότε η απόσταση του αντικειμένου όπου ανακλάτε ο παλμός φωτός βρίσκεται σε βάθος:

$$r = \frac{t_2 - t_1}{2c} = \frac{\Delta t}{2c} \quad (3.1)$$



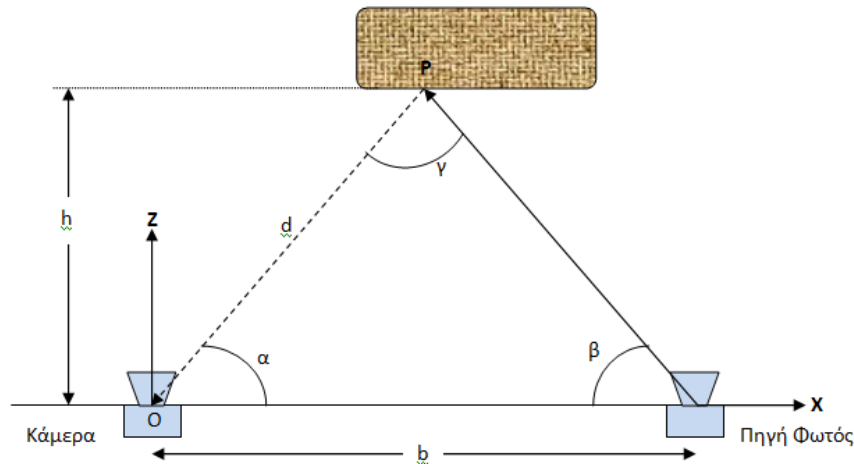
Εικόνα 3.1 : Σχηματική αναπαράσταση του τρόπου μέτρησης βάθους στην μέθοδο time-of-flight

3.2.2 Structured Light

Η απεικόνιση του βάθους με δομημένο φως είναι μια παραλλαγή της στερεοσκοπικής μεθόδου. Η διαφορά με την στερεοσκοπική μέθοδο όπου χρησιμοποιεί δύο κάμερες, είναι ότι σε αυτή τη μέθοδο ένα καλά καθορισμένο πρότυπο (σημεία ή γραμμές) προβάλλεται πάνω στην σκηνή από μία θέση, και στην συνέχεια απεικονίζεται από μία άλλη θέση με μία συγκεκριμένη σχέση μεταξύ των δύο θέσεων.

3.2.2.1 2D Ανάλυση

Το σύστημα στο επίπεδο έχει την εξής διάταξη:



Εικόνα 3.2 : Σχηματική αναπαράσταση του τρόπου μέτρησης βάθους στην μέθοδο structured light στον 2D χώρο

Όπου O είναι το κέντρο της κάμερας και το οπτικό κέντρο του συστήματος συντεταγμένων του επιπέδου της εικόνας. Η απόσταση μεταξύ πηγής και κάμερας είναι σταθερή, γνωστή και ίση με b . Η απόσταση d μεταξύ της κάμερας και του σημείου $P = (X_0, Y_0)$ του αντικειμένου μπορεί να υπολογιστεί με τον νόμο ημιτόνων:

$$\frac{d}{\sin(\alpha)} = \frac{b}{\sin(\gamma)} \quad (3.2)$$

όπου $\gamma = \pi - (\alpha + \beta)$, $\sin(\pi - \gamma) = \sin(\gamma)$, έτσι καταλήγουμε στην:

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

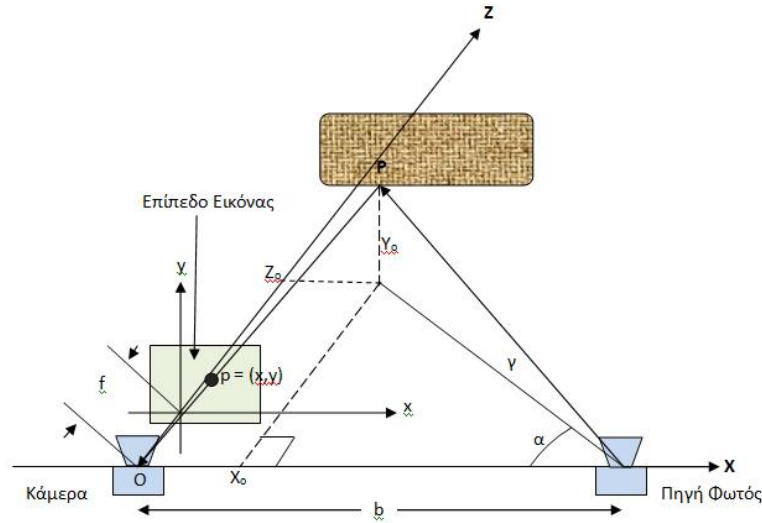
Έτσι η τοποθεσία του σημείου $P = (X_0, Y_0)$ μπορεί να αναπαρασταθεί στο σύστημα συντεταγμένων της κάμερας με δύο διαστάσεων πολικές συντεταγμένες (d, β) . Η μετατροπή τους σε καρτεσιανές συντεταγμένες υπολογίζεται ως:

$$X_0 = d \cos(\beta), Z_0 = h = d \sin(\beta) \quad (3.4)$$

Ο άξονας Z συμπίπτει με τον οπτικό άξονα της κάμερας και το επίπεδο της εικόνας βρίσκεται στο $Z = f$. Έτσι η απόσταση μπορεί να καθοριστεί από την γωνία β (43).

3.2.2.2 3D Ανάλυση

Στην γενική περίπτωση του τρισδιάστατου χώρου, έχουμε την εξής διάταξη:



Εικόνα 3.3 : Σχηματική αναπαράσταση του τρόπου μέτρησης βάθους στην μέθοδο structured light στον 3D χώρο

όπου η κάμερα έχει σύστημα συντεταγμένων το XYZ και το επίπεδο της εικόνας βρίσκεται στο επίπεδο $Z = f$ όπως και προηγουμένως. Το σημείο του αντικείμενο, $\mathbf{P} = (X_0, Y_0, Z_0)$, προβάλλεται στο σημείο $\mathbf{p} = (x, y)$ του επιπέδου της εικόνας.

Στην περίπτωση του τρισδιάστατου χώρου, η κάμερα και η πηγή του φωτός μπορούν να προσδιοριστούν αυθαίρετα. Επίσης, ο άξονας Z είναι ο οπτικός άξονας της κάμερας. Υποθέτουμε ότι το οπτικό κέντρο της πηγής φωτός είναι στον άξονα X.

Οι γωνίες α και γ καθορίζουν την διεύθυνση της ακτίνας φωτός στον τρισδιάστατο χώρο. Με βάση τα όμοια τρίγωνα στον τρισδιάστατο χώρο και στο επίπεδο της εικόνας έχουμε:

$$\frac{X_0}{x} = \frac{Z_0}{f} = \frac{Y_0}{y} \quad (3.5)$$

με την βοήθεια της τριγωνομετρίας στο τρίγωνο από την πλευρά της πηγής έχουμε:

$$\tan(\alpha) = \frac{Z_0}{b - X_0} \quad (3.6)$$

και:

$$\begin{aligned} Z_0 &= \frac{X_0}{x} f = (b - X_0) \tan(\alpha) \\ \Rightarrow X_0 \left(\frac{f}{x} + \tan(\alpha) \right) &= b \tan(\alpha) \end{aligned} \quad (3.7)$$

Έτσι η τρισδιάστατη θέση του σημείου $\mathbf{P} = (X_0, Y_0, Z_0)$ προσδιορίζεται από τις:

$$X_0 = \frac{x b \tan(\alpha)}{f + x \tan(\alpha)}, Y_0 = \frac{y b \tan(\alpha)}{f + x \tan(\alpha)}, Z_0 = \frac{f b \tan(\alpha)}{f + x \tan(\alpha)} \quad (3.8)$$

Παρατηρούμε ότι η γωνία γ δεν παίζει ρόλο στον υπολογισμό του σημείου $\mathbf{P} = (X_0, Y_0, Z_0)$. Ο λόγος είναι γιατί ο δέσμη φωτός υποθέσαμε ότι είναι επίπεδη όταν

υπολογίσαμε το σημείο $P = (X_0, Y_0, Z_0)$. Το επίπεδο της δέσμης είναι κάθετο στο επίπεδο ΧΖ και το επίπεδο Χ με το επίπεδο της δέσμης φωτός δημιουργούν την γωνία α .

3.2 Συσκευές Σάρωσης 3D Point Cloud

3.2.1 Σαρωτής Χειρός (Hand Held)

Ο αισθητήρας χειρός χρησιμοποιείται για εφαρμογές όπου γίνονται σαρώσεις με μικρή απόσταση από την επιφάνεια προς σάρωση. Χρησιμοποιείται σε ένα ευρύ φάσμα εφαρμογών αντίστροφης μηχανικής οι οποίες απαιτούν πολύ μεγαλύτερες ανοχές από τα επίγεια συστήματα σάρωσης (Terrestrial). Η ακρίβεια που προσφέρει το σύστημα σε σαρώσεις μικρής εμβέλειας είναι της τάξης του 0.1 mm. Συχνά παρατηρείται να προσάπτεται σε ένα αυτόματο ρομποτικό βραχίονα ο οποίος έχει σκοπό την πιο ακριβή και σταθερή σάρωση της επιφάνειας από κοντινές αποστάσεις. Επίσης σε κάποιες περιπτώσεις παρατηρείται η στήριξή του σε ένα σύστημα τρίποδα όπως τον τρίποδα *surphaser*.

Οι βασικές χρήσεις του σαρωτή χειρός είναι:

- Βιομηχανικός σχεδιασμός
- Τεχνικά σχέδια
- 3Dσάρωση σε CAD
- Ψηφιοποίηση μικρών μοντέλων (π.χ. από πηλό, μακέτες)



Εικόνα 3.4 : Σαρωτής Χειρός

3.2.2 Κινητοί Σαρωτές (Mobile Scanner)

Οι σαρωτές αυτοί προσφέρουν ένα σύστημα χαρτογράφησης μεγάλης κλίμακας με μικρό κόστος, αυξημένη ακρίβεια και πολύ γρήγορη συλλογή δεδομένων. Η ακρίβεια των συστημάτων αυτών κυμαίνεται από τα $\pm 20 \sim \pm 100$ mm. Υπάρχουν κινητοί σαρωτές ενσωματωμένοι σε οχήματα οι οποίοι χαρτογραφούν ολόκληρες πόλεις, σιδηροδρομικά δίκτυα και δίκτυα αυτοκινητόδρομων. Η δυνατότητα χαρτογράφησης δικτύων όπως είναι οι σιδηρόδρομοι και οι αυτοκινητόδρομοι οφείλεται στο γεγονός ότι οι σαρωτές μπορούν

να ανακτήσουν τα δεδομένα ακόμα και όταν το όχημα στο οποίο βρίσκεται έχει ταχύτητα 80km/h(44), (45). Οι κινητοί σαρωτές συνήθως έχουν ενσωματωμένο και ένα σύστημα GPSγια άμεσο προσανατολισμό του συστήματος όταν το όχημα δεν είναι επανδρωμένο. Οι κινητοί σαρωτές μπορούν επίσης να τοποθετηθούν και σε βάρκα έτσι ώστε να χαρτογραφήσει και την ακτογραμμή.

Οι βασικές χρήσεις του κινητού σαρωτή είναι:

- Χαρτογράφηση Πόλης
- Παράκτιες μελέτες
- Μελέτη και χαρτογράφηση αυτοκινητόδρομου και σιδηρόδρομου
- Χρήση σε λατομία (37)
- Ψηφιακά μοντέλα εδάφους



Εικόνα 3.5 : Κινητός Σαρωτής

3.2.3 Επίγειοι Σαρωτές

Οι επίγειοι σαρωτές, γνωστοί και ως LiDAR (Light detection and ranging) τείνουν να γίνουν οι πιο προτιμώμενοι σαρωτές για γρήγορη και ακριβή συλλογή 3D δεδομένων. Τα επίγεια LiDAR αποτελούνται από τον σαρωτή πάνω σε ένα τρίποδο. Τα συστήματα αυτά έχουν μεγάλη ακρίβεια με μέσο όρο περίπου +/- 3mm ανά σημείο σάρωσης για στατική σάρωση. Οι επίγειοι σαρωτές μπορούν να χωριστούν σε σαρωτές βασισμένους στη φάση και σε σαρωτές βασισμένους στην ώρα επιστροφής του σήματος(time off light). Οι σαρωτές που είναι βασισμένοι στην φάση μετράνε την διαφορά της φάσης της ακτίνας laser. Το πλεονέκτημα αυτής της τεχνολογίας είναι ότι αυξάνεται σημαντικά ο όγκος πληροφορίας που αποκτάται (>1,000,000 σημεία το δευτερόλεπτο)(47). Οι σαρωτές στην δεύτερη κατηγορία εκπέμπουν ένα παλμό φωτός και μετρούν τον χρόνο ο οποίος θα χρειαστεί για να επιστρέψει πίσω στον σαρωτή. Χρησιμοποιούνται τόσο σε αεροπλάνα για την χαρτογράφηση του εδάφους, όσο και σε πλοία για την χαρτογράφηση του βυθού. Το πλεονέκτημα αυτών των σαρωτών είναι ότι μπορούν να λάβουν πληροφορία από πιο μακρινές αποστάσεις από ότι οι σαρωτές που βασίζονται στη φάση.



Εικόνα 3.6 : Επίγειος Σαρωτής

3.3 Σαρωτής KINECT

Γνωστό και ως Project Natal, ο αισθητήρας KINECT είναι ένας αισθητήρας κίνησης ο οποίος αρχικά παρουσιάστηκε ως ένα τηλεχειριστήριο παιχνιδιών. Στην συνέχεια όμως, με την κάμερα βάθους που παρέχει μια πολύ καλή ακρίβεια και με το χαμηλό κόστος του, κέντρισε το ενδιαφέρον των ερευνητών για νέες καινοτόμες εφαρμογές ή για την αντικατάσταση παλιών μεθόδων με μεθόδους όπου έχουν χαμηλότερο κόστος.

Το KINECT αποτελείται από δύο τμήματα. Το μηχανικό τμήμα το οποίο είναι η βάση του και η οποία έχει μηχανισμό που επιτρέπει στο KINECT να αλλάζει την οριζόντια κλίση του. Το τμήμα αυτό στηρίζει το κυρίως κομμάτι του αισθητήρα στο οποίο βρίσκονται οι αισθητήρες και οι κάμερες. Πιο αναλυτικά, περιέχει:

- RGB κάμερα
- Πομπό υπέρυθρων (IR emitter)
- Υπέρυθρο αισθητήρα βάθους
- Μια σειρά από μικρόφωνα
- LED

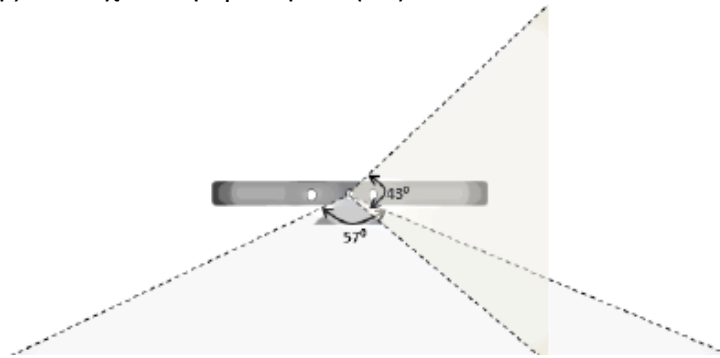
Στα πιο κάτω υποκεφάλαια αναλύονται τα μέρη του αισθητήρα KINECT τα οποία μας ενδιαφέρουν για την οπτική λήψη δεδομένων και αναπαράσταση και όχι για την ακουστική.

3.3.1 Παρουσίαση Τμημάτων Αισθητήρα KINECT

3.3.1.1 Έγχρωμη Κάμερα

Η έγχρωμη κάμερα είναι υπεύθυνη για την λήψη και μετάδοση των έγχρωμων δεδομένων. Η λειτουργία της είναι η ανίχνευση των βασικών χρωμάτων, κόκκινου, πράσινου και μπλε από την σκηνή η οποία θα απεικονιστεί. Η ροή των δεδομένων η οποία μεταδίδεται από την έγχρωμη κάμερα είναι μια διαδοχή των καρέ της εικόνας. Έχει την δυνατότητα ταχύτητας λήψης 30 καρέ το δευτερόλεπτο με ανάλυση εικόνας 640 x 480 pixels, ή 12 καρέ το δευτερόλεπτο το οποίο δίνει και την μέγιστη ανάλυσή του η οποία είναι 1280 x 960 pixels.

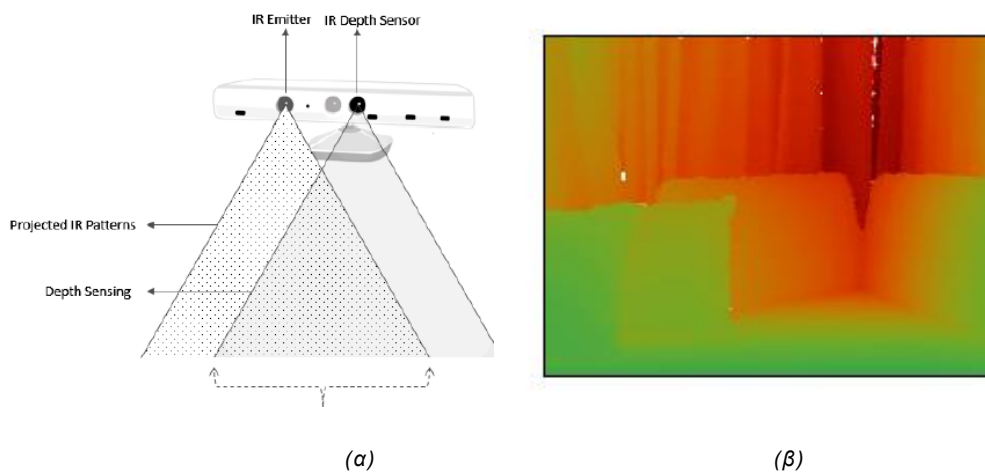
Το οπτικό πεδίο της έγχρωμης κάμερας είναι 43° κάθετα και 57° οριζόντια με σταθερό τον μηχανισμό κλίσης που έχει στην βάση του (48).



Εικόνα 3.7 : Αναπαράσταση 3D Εύρους αισθητήρα KINECT

3.3.1.2 Πομπός και Αισθητήρας Υπερύθρων (IR emitter and IR depth sensor)

Για τον υπολογισμό του βάθους των σημείων ενός αντικειμένου, ο αισθητήρας KINECT χρησιμοποιεί τις δύο αυτές συσκευές υπέρυθρων οι οποίες συνεργάζονται με τρόπο ο οποίος θα περιγραφεί λεπτομερέστερα σε πιο κάτω κεφάλαιο. Γενικά, ο πομπός υπέρυθρων, εκπέμπει ένα πλαίσιο από υπέρυθρες κουκκίδες σε όλο το οπτικό πεδίο του αισθητήρα KINECT. Οι υπέρυθρες αυτές κουκκίδες είναι αόρατες σε εμάς αλλά για τον υπέρυθρο αισθητήρα βάθους είναι ορατές, ο οποίος μετά την ανάκλασή τους στο αντικείμενο, τις λαμβάνει και με βάση αυτές μας παρέχει πληροφορίες για το βάθος του αντικειμένου από όπου είχαν ανακλαστεί. Το πεδίο των υπέρυθρων καθώς και ένα αποτέλεσμα απεικόνισης με αυτούς τους αισθητήρες φαίνονται πιο κάτω (48).

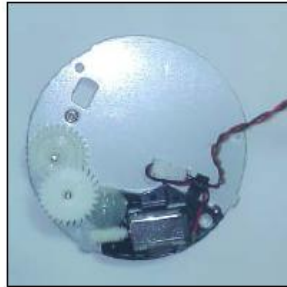


Εικόνα 3.8: (α) Αναπαράσταση 3D προβολής του IR pattern και εύρος λήψης της υπέρυθρης κάμερας
(β) Παράδειγμα αποτελέσματος της υπέρυθρης κάμερας όπου με κόκκινο αναπαριστούνται τα μακρινά και με πράσινο τα κοντινά προς την κάμερα αντικείμενα

Τα δεδομένα βάθους μπορούν να έχουν ανάλυση 640×480 , 320×240 ή 80×60 pixels, και το οπτικό πεδίο του αισθητήρα καθώς και της έγχρωμης κάμερας παραμένει το ίδιο.

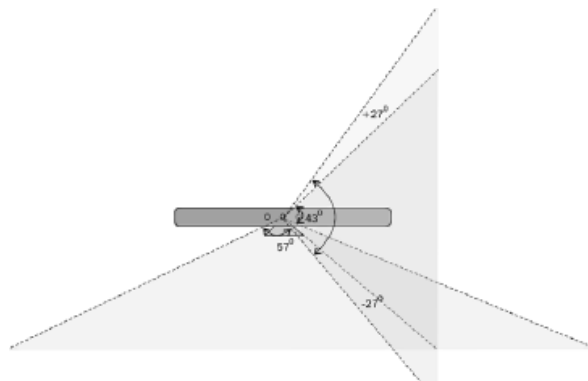
3.3.1.3 Κινητήρας Κλίσης (Tilt motor)

Η βάση και το σώμα του αισθητήρα KINECT ενώνονται με ένα μικρό κινητήρα όπως φαίνεται στην εικόνα πιο κάτω. Η χρήση του είναι για να πετύχει ο αισθητήρας την σωστή κατακόρυφη κλίση έτσι ώστε το αντικείμενο που θα απεικονιστεί να είναι μέσα στο οπτικό πεδίο του αισθητήρα(48).



Εικόνα 3.9 : Μηχανισμός κλίσης του αισθητήρα KINECT που βρίσκεται στην βάση του

Ο κινητήρας αυτός έχει την δυνατότητα να δώσει κλίση στον αισθητήρα 27° κάθετα προς τα πάνω και προς τα κάτω. Έτσι το δυνατό οπτικό πεδίο του αισθητήρα KINECT όταν αυτός παραμένει σταθερός σε μια συγκεκριμένη θέση αυξάνεται κατά 27° προς τα πάνω και 27° προς τα κάτω, άρα σύνολο 54° στην κάθετο διεύθυνση όπως φαίνεται και στο πιο κάτω σχήμα.

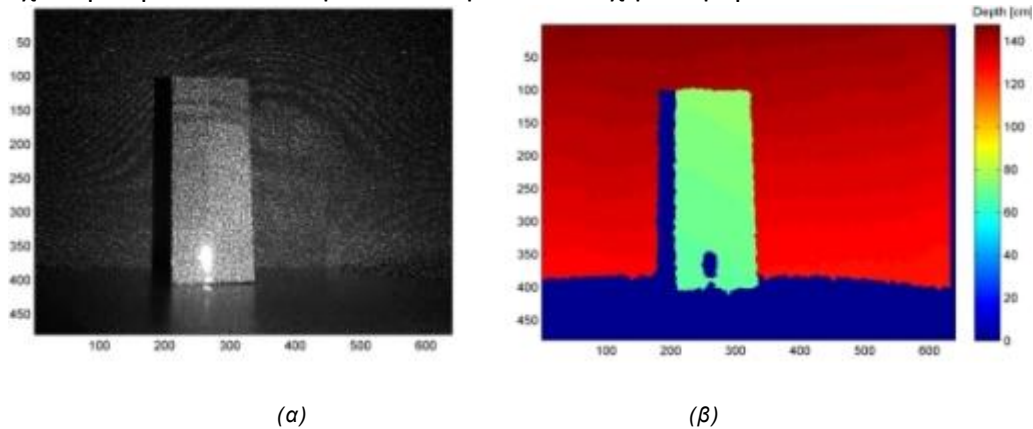


Εικόνα 3.10 : Αναπαράσταση 3D Εύρους αισθητήρα KINECT λαμβάνοντας υπόψη και την δυνατότητα που δίνει ο μηχανισμός κλίσης

3.3.2 Λήψη Δεδομένων Βάθους

Για την λήψη των δεδομένων βάθους των διάφορων σημείων, ο αισθητήρας KINECT χρησιμοποιεί την μέθοδο του δομημένου φωτός(49). Ο αισθητήρας KINECT περιέχει ένα υπέρυθρο πομπό ακτινών laser, μια υπέρυθρη κάμερα και μία έγχρωμη κάμερα. Ο πομπός laser εκπέμπει μία δέσμη laser η οποία διαχωρίζεται σε πολλές δέσμες μετά την διάθλασή της για την δημιουργία ενός σταθερού μοτίβου από κουκκίδες οι οποίες προβάλλονται στην σκηνή. Αυτό το μοτίβο απεικονίζεται από την υπέρυθρη κάμερα και συσχετίζεται με

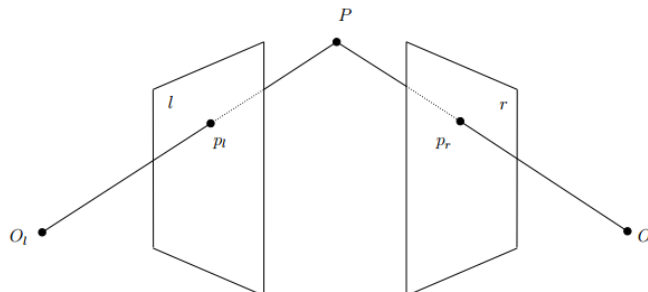
ένα μοτίβο αναφοράς. Το μοτίβο αναφοράς αποκτάται από την απεικόνιση του χώρου σε μία γνωστή απόσταση από τον αισθητήρα και αποθηκεύεται στην μνήμη του αισθητήρα. Όταν μία κουκκίδα προβληθεί στο αντικείμενο προς απεικόνιση τότε αν η απόσταση (βάθος) της είναι μικρότερη ή μεγαλύτερη από την απόσταση όπου αποκτήθηκε το μοτίβο αναφοράς, η θέση της κουκκίδας στην υπέρυθρη εικόνα θα μετακινηθεί στην διεύθυνση προς την νοητή γραμμή που ενώνει την κάμερα με την πηγή υπέρυθρων. Αυτές οι μετακινήσεις όλων των θέσεων στις κουκκίδες μετρούνται με ένα αλγόριθμο συσχέτισης όμοιο με αυτό που χρησιμοποιείται στην στερεοσκοπική τεχνική, του οποίου το αποτέλεσμα είναι μία εικόνα ανισότητας (50),(51). Έτσι η απόσταση για κάθε εικονοστοιχείο μπορεί να ανακτηθεί από την αντίστοιχη διαφορά.



Εικόνα 3.11 (α) Απεικόνιση IR patternπου εκπέμπει ο IR πομπός του αισθητήρα στο αντικείμενο στο χώρο
(β) Το αποτέλεσμα της απεικόνισης βάθους

3.3.2.1 Συσχέτιση των στοιχείων

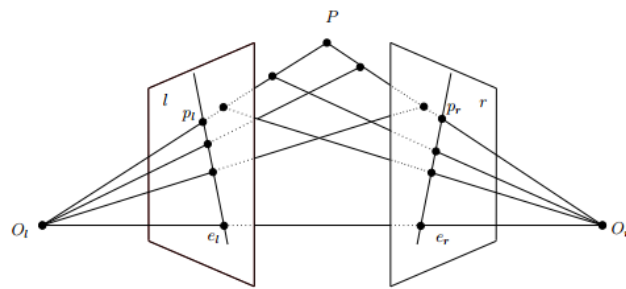
Σε αυτό το υποκεφάλαιο θα μελετήσουμε τον αλγόριθμο συσχέτισης όπως είναι στην στερεοσκοπική τεχνική ο οποίος χρησιμοποιείται από το λογισμικό της Prime Sense και για τον αισθητήρα KINECT. Υποθέτουμε την δεξιά και αριστερή προβολή ενός σημείου σε δύο επίπεδα εικόνας όπως φαίνεται στο σχήμα.



Εικόνα 3.12 : Αναπαράσταση δύο προοπτικών επιπέδων και των προβολών του σημείου στο χώρο πάνω στα επίπεδα αυτά

Τα προβολικά κέντρα των δύο επιπέδων εικόνας αναγράφονται ως O_r και O_l , αντιστοίχως. Ονομάζουμε την νοητή ευθεία η οποία ενώνει αυτά τα δύο κέντρα ως ευθεία βάσης. Το σημείο P προβάλλεται στα δύο επίπεδα εικόνας στα σημεία p_r και p_l αντίστοιχα, τα οποία βρίσκονται από την τομή του επιπέδου με την ευθεία που ενώνει το προβολικό κέντρο με το σημείο P (52).

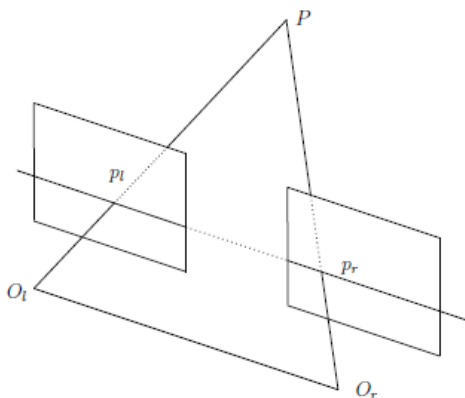
Ορίζουμε ως επιπολικό επίπεδο (epipolar plane), το επίπεδο το οποίο σχηματίζεται από τα σημεία των προβολικών κέντρων και του σημείου που απεικονίζεται. Σαν αριστερό επιπολικό ονομάζουμε την προβολή του δεξιού προβολικού κέντρου στο αριστερό επίπεδο εικόνας και συμβολίζεται με e_l και αντίστοιχα για το δεξιό επιπολικό (e_r). Για κάθε σημείο του χώρου, μπορούμε να ορίσουμε δύο επιπολικές γραμμές. Οι επιπολικές γραμμές ορίζονται ως η τομή του επιπολικού επιπέδου με το προβολικό επίπεδο (επίπεδα εικόνας).



Εικόνα 3.13 : Αναπαράσταση Επιπολικής Γεωμετρίας, επιπολικό επίπεδο (PO_lO_r), δεξιό και αριστερό επιπολικό (e_r, e_l), και επιπολικές γραμμές (p_r, e_r, p_l, e_l)

Η επιπολική γεωμετρία περιγράφει την σχέση μεταξύ δεξιάς και αριστερής προβολής του σημείου P .

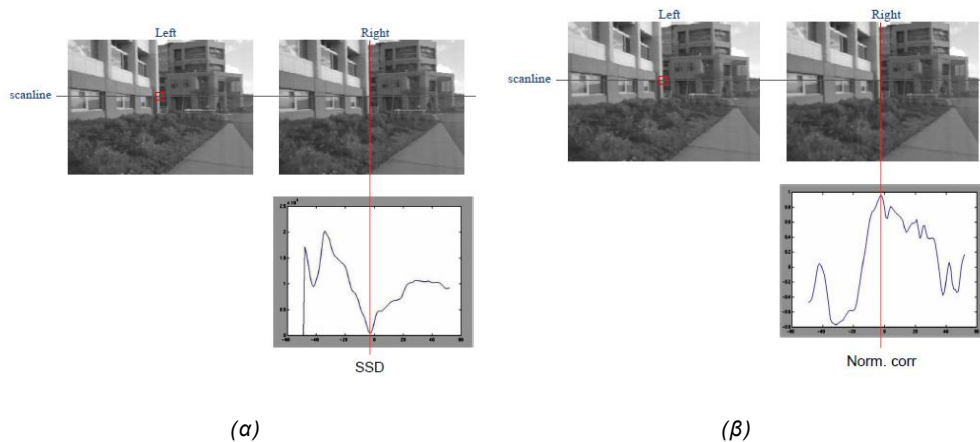
Στην περίπτωση όπου η διαμόρφωση των δύο επιπέδων χαρακτηρίζεται ως κανονική, δηλαδή έχουν παράλληλα τα επίπεδα εικόνων του, όπως συμβαίνει και στον αισθητήρα KINECT:



Εικόνα 3.14 : Παράλληλη διάταξη των δύο προοπτικών επιπέδων

η ευθεία βάσης είναι παράλληλη στα δύο επίπεδα εικόνας και οι επιπολικές ευθείες συμπίπτουν. Έτσι, οι ομογενείς συντεταγμένες των επιπολικών σημείων p_r και p_l βρίσκονται στην ίδια ευθεία αλλά σε διαφορετική θέση.

Συνεπώς, η συσχέτιση των στοιχείων που γίνεται, είναι για κάθε εικονοστοιχείο της μίας εικόνας, προσδιορίζεται η επιπολική ευθεία με βάση την γεωμετρία του συστήματος και των συντεταγμένων του στοιχείου στον χώρο και το επίπεδο της εικόνας. Στην ίδια επιπολική ευθεία, στην δεύτερη εικόνα, όπου είναι η ίδια με αυτή της πρώτης εικόνας (λόγω κανονικότητας συστήματος), εξετάζονται όλα τα εικονοστοιχεία της με βάση μία συνάρτηση. Η συνάρτηση μπορεί να είναι άθροισμα τετραγώνων διαφοράς (SSD) ή κανονικοποιημένη συσχέτιση. Στην κάθε περίπτωση, βρίσκουμε το σημείο/εικονοστοιχείο το οποίο ελαχιστοποιεί ή μεγιστοποιεί την συνάρτηση αντίστοιχα, όπως φαίνεται πιο κάτω.

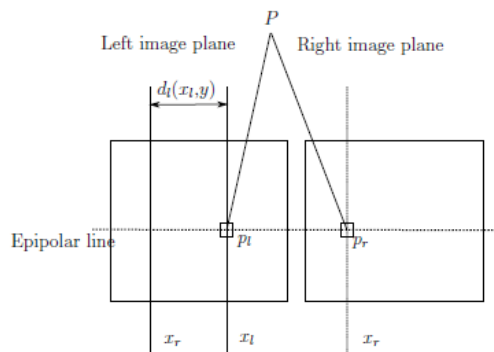


Εικόνα 3.15 : Συσχέτιση των εικόνων και εύρεση του ριχελιπού την δεύτερη εικόνα που ταυτίζεται με το υπό έλεγχο ριχελιστην πρώτη εικόνα με:

(α) Ελαχιστοποίηση αθροίσματος τετραγώνων διαφοράς (SSD)

(β) Μεγιστοποίηση κανονικοποιημένης συσχέτισης (Norm. Corr.)

Τα δύο επίπεδα εικόνας είναι όπως φαίνονται πιο κάτω:



Εικόνα 3.16 : Αναπαράσταση ανισότητας μεταξύ δύο όμοιων ριχελιστις δύο εικόνες

Το εικονοστοιχείο p_r στο αριστερό επίπεδο εικόνας φαίνεται να είναι μετακινημένο προς τα αριστερά σε σύγκριση με το εικονοστοιχείο p_l στο αριστερό επίπεδο εικόνας. Αυτή η μετακίνηση στις ομογενείς συντεταγμένες των δύο προβολών του σημείου P στα δύο επίπεδα ονομάζεται disparity και ορίζεται ως:

$$d_l(x_l, y) = x_l - x_r \quad (3.9)$$

Και με βάση αυτό, όπως θα αποδειχθεί και στο επόμενο κεφάλαιο, μπορούμε να υπολογίσουμε το βάθος του αντίστοιχου στοιχείου στο χώρο από την:

$$Z_k = \frac{fD}{d} \quad (3.10)$$

Όπου:

Z_k : Απόσταση (βάθος) του σημείου k στο αντικείμενο από τον αισθητήρα

d : παρατηρούμενη διαφορά (disparity)

f : εστιακό μήκος της υπέρυθρης κάμερας

D : η μετατόπιση του σημείου k

3.3.2.2 Μαθηματικό Μοντέλο Εύρεσης Βάθους

Στο κεφάλαιο αυτό θα παρουσιαστεί το μαθηματικό μοντέλο υπολογισμού του βάθους κάθε στοιχείου του νέφους. Στην πιο κάτω εικόνα παρουσιάζεται η σχέση μεταξύ της απόστασης ενός αντικειμένου από τον αισθητήρα και της επιφάνειας με τις κουκκίδες αναφοράς. Η απόσταση (d) του επιπέδου στο οποίο βρίσκεται το αντικείμενο με το επίπεδο αναφοράς, είναι η διαφορά η οποία θέλουμε να υπολογίσουμε. Ο υπολογισμός γίνεται στον τρισδιάστατο χώρο με αρχή των αξόνων στο κέντρο της κάμερας το οποίο συμπίπτει με το προβολικό κέντρο. Ο άξονας Z είναι κάθετος στο επίπεδο της εικόνας και εκτείνεται προς το αντικείμενο και ο άξονας X είναι παράλληλος με την νοητή γραμμή η οποία ενώνει την κάμερα με τον πομπό υπέρυθρων.

Υποθέτουμε ότι η κουκκίδα βρίσκεται στο επίπεδο αναφοράς με απόσταση Z_0 από τον αισθητήρα και η απεικόνισή της απεικονίζεται στο επίπεδο της εικόνας. Αν η κουκκίδα μετακινηθεί ως προς τον αισθητήρα, τότε η απεικόνισή της στο επίπεδο της εικόνας θα παρουσιάσει μια μετακίνηση στον άξονα X. αυτή την μετακίνηση μπορούμε να την υπολογίσουμε με την βοήθεια όμοιων τριγώνων:

$$\frac{D}{d} = \frac{Z_0 - Z_k}{Z_0} \quad (3.11)$$

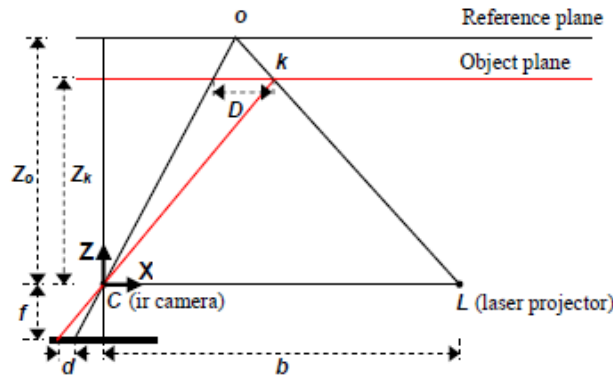
Και

$$\frac{d}{f} = \frac{D}{Z_k} \quad (3.12)$$

Όπου Z_k υποδηλώνει την νέα απόσταση την οποία έχει η κουκκίδα και κατά συνέπεια το επίπεδο του αντικειμένου, βείναι η απόσταση μεταξύ της κάμερας από τον πομπό υπέρυθρων, f είναι η εστιακή απόσταση της κάμερας, D είναι η μετακίνηση της κουκκίδας στο επίπεδο του αντικειμένου και d είναι η μετακίνησης της κουκκίδας. Διαιρώντας το Δαπό τις (3.11) και (3.12) και λύνοντας ως προς το Z_k έχουμε:

$$Z_k = \frac{Z_0}{1 + \frac{Z_0}{fb} d} \quad (3.13)$$

Η εξίσωση (3.13) είναι το μαθηματικό μοντέλο για την εξαγωγή του βάθους από την παρατηρούμενη διαφορά με όρους οι οποίοι είναι σταθεροί (Z_0, f και b) και μπορούν να υπολογιστούν.



Εικόνα 3.17 : Αναπαράσταση της ανισότητας και τοπολογία για τον μαθηματικό υπολογισμό της ανισότητας

Τότε οι συντεταγμένες του κάθε σημείου του αντικειμένου μπορούν να υπολογιστούν:

$$X_k = \frac{Z_k}{f} (x_k - x_0 + \delta x) \quad (3.14)$$

$$Y_k = \frac{Z_k}{f} (y_k - y_0 + \delta y)$$

όπου x_k και y_k είναι οι συντεταγμένες στην εικόνα του σημείου, x_0 και y_0 είναι οι συντεταγμένες του κέντρου της εικόνας στο επίπεδο εικόνας, δx και δy είναι οι διορθώσεις για την διαταραχή των φακών του αισθητήρα.

3.3.2.3 Βαθμονόμηση Αισθητήρα

Στις εξισώσεις (3.14) παρατηρούμε ότι οι παράγοντες βαθμονόμησης εμπεριέχονται στο μαθηματικό μοντέλο υπολογισμού του βάθους ενός σημείου. Επίσης, πρέπει να εξετάσουμε το γεγονός ότι μπορεί να υπάρχει μια γωνία μεταξύ του άξονα X και της ευθείας βάσης η οποία στον υπολογισμό του βάθους θεωρήθηκε μηδενική. Ωστόσο αυτό δεν επηρεάζει τον υπολογισμό των συντεταγμένων του αντικειμένου αν ορίσουμε το σύστημα συντεταγμένων του βάθους να είναι παράλληλο με σύστημα συντεταγμένων της εικόνας αντί με την ευθεία της βάσης. Έτσι μπορούμε να αγνοήσουμε την πιθανή ύπαρξη εσφαλμένης γωνίας μεταξύ της βάσης και του άξονα X.

Οι παράμετροι εστιακό βάθος, τα κεντρικά σημεία της εικόνας στο επίπεδο εικόνας και η διαταραχή των φακών του αισθητήρα μπορούν να βαθμονομηθούν εύκολα από την βαθμονόμηση της υπέρυθρης κάμερας. Σε αντίθεση με αυτά, η ευθεία βάσης και η απόσταση στην οποία αποκτήθηκε το πρότυπο αναφοράς είναι δύσκολο να βαθμονομηθούν για τον λόγο ότι δεν μπορούμε να υπολογίσουμε ακριβώς τις διάφορες ανισότητες (disparities) λόγω του περιορισμού του εύρους τιμών. Αντί αυτού, οι ανισότητες αυτές κανονικοποιούνται μεταξύ του 0 και του 2047 μεταδίδονται με 11 bits. Έτσι στην εξίσωση (3.13), οι ανισότητες (d) αντικαθίστανται από τον όρο $md' + n$, όπου

το d' είναι οι κανονικοποιημένες ανισότητες και m , πείναι οι παράμετροι της γραμμικής κανονικοποίησης. Έτσι η εξίσωση (3.13) μετατρέπεται σε:

$$Z_k^{-1} = \frac{m}{fb} d' + \left(Z_0^{-1} + \frac{n}{fb} \right) \quad (3.15)$$

όπου μας εκφράζει την γραμμική σχέση μεταξύ του βάθους και της αντίστοιχης κανονικοποιημένης ανισότητας. Έτσι παρατηρούμε ότι, με την απεικόνιση ενός αριθμού από σημεία γνωστής απόστασης από τον αισθητήρα, οι παράμετροι της γραμμικής σχέσης μπορούν να υπολογιστούν με την χρήση της μεθόδου ελαχίστων τετραγώνων.

3.3.2.4 Πηγή Σφαλμάτων

Τα σφάλματα στην απεικόνιση του αισθητήρα μπορούν να προέλθουν από διάφορες πηγές όπως είναι από τον ίδιο τον αισθητήρα, από τις συνθήκες στις οποίες κάνουμε μέτρηση και από τις ιδιότητες της επιφάνειας του αντικειμένου προς απεικόνιση.

Τα εγγενή σφάλματα του αισθητήρα οφείλονται στη λανθασμένη βαθμονόμηση και στη λανθασμένη μέτρηση των ανισοτήτων. Υπό αυτές τις συνθήκες θα οδηγείται ο χρήστης σε συστηματικά σφάλματα στις συντεταγμένες του αντικειμένου. Τα συστηματικά σφάλματα αντιμετωπίζονται με μια σωστή βαθμονόμηση όπως περιγράφηκε στο πιο πάνω κεφάλαιο.

Σφάλματα από τις συνθήκες στις οποίες βρίσκεται ο αισθητήρας μας οφείλονται κυρίως στο φωτισμό και τη γεωμετρία της απεικόνισης. Ο φωτισμός επηρεάζει τον αλγόριθμο συσχέτισης που περιγράφηκε πιο πάνω και κατά συνέπεια και την μέτρηση των ανισοτήτων. Σε κατάσταση ισχυρού φωτισμού, οι κουκκίδες του πομπού laser έχουν χαμηλότερη αντίθεση στην υπέρυθη εικόνα, έτσι μπορεί να οδηγήσει σε παρουσία κενού στην εικόνα λόγω μη εντοπισμού της κουκκίδας. Η γεωμετρία στην απεικόνιση επηρεάζει στην ύπαρξη ή όχι σφάλματος στο αποτέλεσμα. Ο κύριος λόγος ύπαρξης σφαλμάτων από την γεωμετρία της απεικόνισης είναι λόγω του εύρους στο οποίο ο αισθητήρας μπορεί να κάνει μετρήσεις, το οποίο με βάση τις προδιαγραφές του είναι μεταξύ 0.5 m και 5 m, και όπως θα αναφέρουμε και πιο κάτω, τα σφάλματα στις μετρήσεις αυξάνονται με την αύξηση της απόστασης στην οποία βρίσκεται το αντικείμενο.

Επίσης, η υφή της επιφάνειας του αντικειμένου μπορεί να σχετίζεται με τη δημιουργία σφαλμάτων. Για παράδειγμα, γυαλιστερές και λείες επιφάνειες εμποδίζουν την μέτρηση των ανισοτήτων, έτσι στο νέφος σημείων παρουσιάζεται ένα κενό σε κάποια σημεία.

3.3.2.5 Μοντέλο Θορύβου

Υποθέτοντας ότι στην εξίσωση (3.15), οι παράμετροι βαθμονόμησης είναι με ακρίβεια προσδιορισμένοι και ότι η παράμετρος d είναι τυχαία μεταβλητή που ακολουθεί κανονική κατανομή, μπορούμε να επεξεργαστούμε τη διακύμανση των μετρήσεων ανισότητας για την λήψη της διακύμανσης των μετρήσεων βάθους ως εξής:

$$\sigma_{Z_k}^2 = \left(\frac{\partial Z_k}{\partial d} \right)^2 \sigma_{d'}^2 \quad (3.16)$$

και λαμβάνουμε την έκφραση της τυπικής απόκλισης του βάθους:

$$\sigma_{Z_k} = \left(\frac{m}{fb} \right) Z_k^2 \sigma_{d'} \quad (3.17)$$

όπου $\sigma_{d'}$ και σ_{Z_k} είναι αντιστοίχως η τυπική απόκλιση των κανονικοποιημένων μετρήσεων των ανισοτήτων και η τυπική απόκλιση του υπολογισμένου βάθους. Εδώ βλέπουμε ότι το σφάλμα στις μετρήσεις βάθους στην εικόνα (σ_{Z_k}) είναι ευθέως ανάλογο του τετραγώνου της απόστασης του αντικειμένου από τον αισθητήρα (Z_k^2).

3.3.2.6 Πυκνότητα Νέφους Σημείων

Η ευκρίνεια της υπέρυθρης κάμερας καθορίζει και την απόσταση όπου θα έχουν τα σημεία στο επίπεδο της εικόνας. Καθώς κάθε εικόνα βάθους περιέχει σταθερά 640x 480 pixels, τότε, η πυκνότητα των σημείων θα μειώνεται καθώς η απόσταση του αντικειμένου από την κάμερα αυξάνεται. Αν θεωρήσουμε την πυκνότητα σημείων ως τον αριθμό των σημείων σε μία περιοχή, καθώς ο αριθμός των σημείων παραμένει σταθερός, η περιοχή είναι ανάλογη ως προς την απόσταση στο τετράγωνο των σημείων από την κάμερα. Κατά συνέπεια, η πυκνότητα των σημείων σε ένα νέφος σημείων είναι αντιστρόφως ανάλογη της τετραγωνικής απόστασης από τον αισθητήρα:

$$\rho \propto \frac{1}{Z^2} \quad (3.18)$$

Η διακριτική ικανότητα του βάθους σε έναν αισθητήρα καθορίζεται από τον αριθμό των bits που χρησιμοποιεί για την αναπαράσταση ενός pixel, και συγκεκριμένα στον αισθητήρα KINECT είναι 11 bits. Ωστόσο το ένα bit από αυτά χρησιμοποιείται για να στιγματίσει κατά πόσο σε ένα pixel μετρήθηκε ή όχι ανισότητα, έτσι το KINECT μας παρέχει $2^{10} = 1024$ επίπεδα για την αναπαράσταση της ανισότητας. Καθώς το βάθος είναι επίσης αντιστρόφως ανάλογο των τιμών των ανισοτήτων, η ευκρίνεια του βάθους επίσης είναι αντιστρόφως ανάλογη των επιπέδων των ανισοτήτων.

4 Μεθοδολογία Αυτόματης Ευθυγράμμισης Εικόνων

4.1 Εισαγωγή

Η βέλτιστη ευθυγράμμιση νεφών σημείων έχει σαν σκοπό την εφαρμογή μετασχηματισμού σε ένα εκ των δύο συνόλων σημείων για την μετακίνησή του στο άλλο σύνολο σημείων με την μικρότερη δυνατή απόσταση. Σε αυτή την εργασία θα αναφερόμαστε στο σύνολο αναφοράς ως $\mathbb{Q} = \{q_1, q_2, \dots, q_N\}$ και στο σύνολο προς ευθυγράμμιση ως $\mathbb{P} = \{p_1, p_2, \dots, p_N\}$. Με σκοπό την ποσοτικοποίηση της ακρίβειας αυτής της μεθόδου, θα χρειαστούμε μία συνάρτηση σφάλματος (E). Η δημοφιλέστερη συνάρτηση έκφρασης σφάλματος είναι το σφάλμα αθροίσματος τετραγώνων, όπου το σφάλμα ορίζεται ως η απόσταση μεταξύ των σημείων του μοντέλου από τον κοντινότερο γείτονά του στο σύνολο προς ευθυγράμμιση μετά την ευθυγράμμιση (τ):

$$E = \sum_{i=1}^N \|\tau(p_i) - q_i\|^2 \quad (4.1)$$

Οι μετασχηματισμοί μπορούν να είναι συμπαγείς ή μη συμπαγείς, ανάλογα με το πρόβλημα, τις παραδοχές και τις απαιτήσεις μας. Και στις δύο περιπτώσεις, οι μεθοδολογίες αυτές μπορούν να χειριστούν και τους έξι βαθμούς ελευθερίας του χώρου (μετακίνηση και περιστροφή). Η συνάρτηση σφάλματος (4.1), εκφράζοντας τον μετασχηματισμό (τ) με μήτρες περιστροφής (\mathbf{R}) και μετακίνησης (T), μπορεί να γραφτεί ως εξής:

$$E = \sum_{i=1}^N \|\mathbf{R}p_i + T - q_i\|^2 \quad (4.2)$$

Σε δεδομένα τα οποία προέρχονται από τεχνικές δομημένου φωτός, δεν έχουμε την πληροφορία της συσχέτισης των σημείων. Αυτό μας οδηγεί στο να μην είναι γνωστό το κοντινότερο σημείο, q_i , του σημείου το οποίο ψάχνουμε, p_i , μετά τον μετασχηματισμό του. αυτή η πληροφορία της συσχέτισης των σημείων μπορεί να υπολογιστεί πριν από την εύρεση και εφαρμογή του μετασχηματισμού. Ο αναδρομικός αλγόριθμος πλησιέστερου σημείου (Iterative Closest Point - ICP) κάνει αυτό ακριβώς. Η μεθοδολογία του ICP έχει την ικανότητα να συγκλίνει μονοσήμαντα στο κοντινότερο τοπικό ακρότατο της μέσης τετραγωνικής απόστασης με μεγάλη ακρίβεια και ταχύτητα. Στην παρούσα εργασία αναπτύσσονται και οι δύο τεχνικές μετασχηματισμού (συμπαγής και μη-συμπαγής).

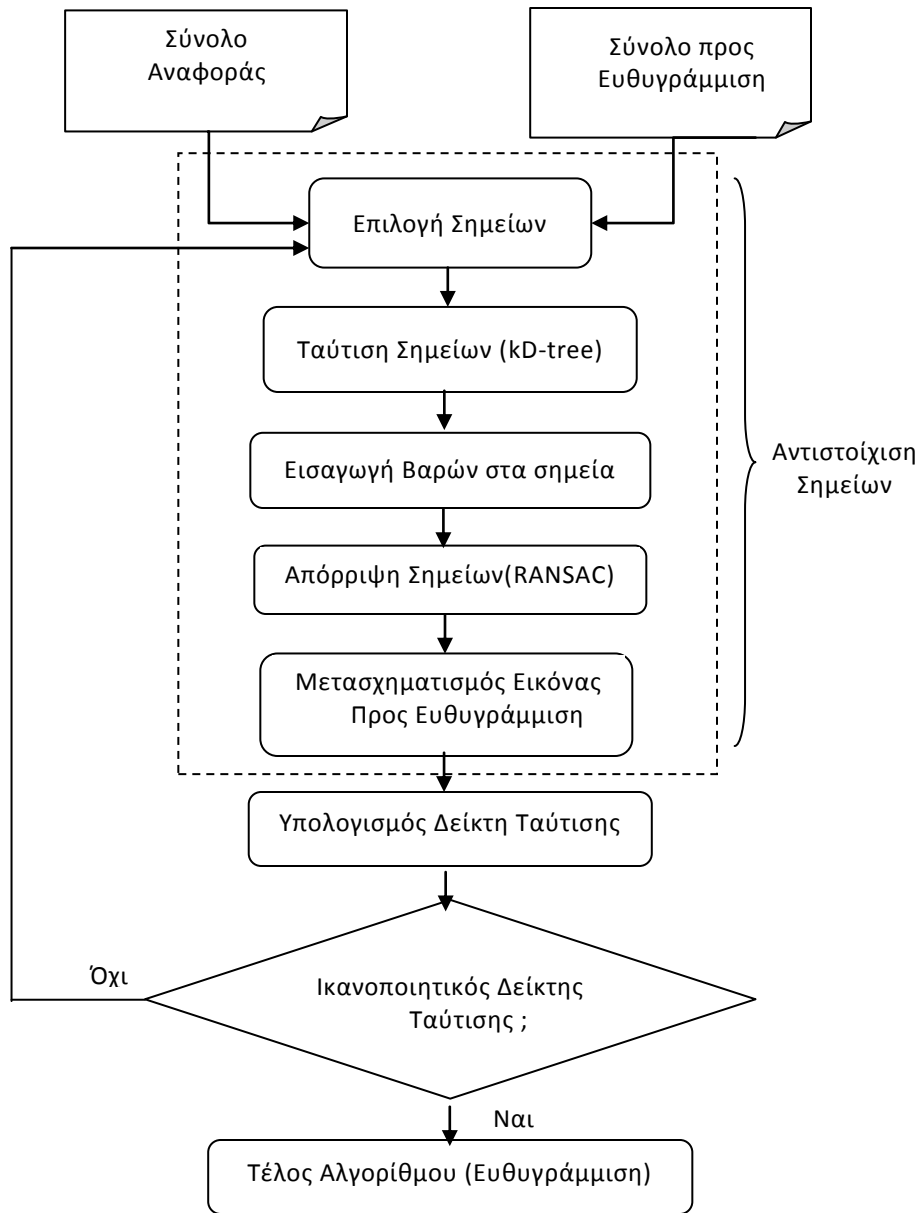
Αρχικό βήμα είναι η επισημοποίηση των υποθέσεων για το είδος της σχέσης (τ) μεταξύ των δύο τρισδιάστατων συνόλων δεδομένων, το οποίο είναι κατάλληλο για την χαρτογράφηση ενός σημείου από το σύνολο αναφοράς στο σύνολο προς ευθυγράμμιση. Στα περισσότερα προβλήματα ευθυγράμμισης, το (τ) είναι ένας μετασχηματισμός συνόλων μεταξύ απεικονιστικών τεχνικών, είτε με τα ίδια είτε με διαφορετικά χαρακτηριστικά απεικόνισης σε διαφορετικές χρονικές στιγμές. Σε αυτό το πλαίσιο, ένας συμπαγής μετασχηματισμός μπορεί να εφαρμοστεί υπό την προϋπόθεση ότι οι ελαστικές

παραμορφώσεις που υπάρχουν είναι αμελητέες σε σύγκριση με την απαιτούμενη ακρίβεια του μετασχηματισμού. Εάν οι παραμορφώσεις μεταξύ των δύο επιφανειών που αναπαριστούν τα τρισδιάστατα σύνολα είναι σημαντικές, και κυρίως αν είναι παραμορφώσεις πέραν από την πιθανή ύπαρξη θορύβου και παραμορφώσεων λόγω της απεικονιστικής τεχνικής, θα πρέπει να εφαρμοστεί ένας μη συμπαγής μετασχηματισμός

4.2 Συμπαγής ICP

Με βάση την μετακίνηση και περιστροφή των σημείων του νέφους υπολογίζεται η μήτρα του μετασχηματισμού η οποία υπολογίζεται με την μέθοδο SVD (Singular Value Decomposition). Έχοντας την μήτρα μετασχηματισμού, εφαρμόζεται στο νέφος σημείων προς ευθυγράμμιση και υπολογίζεται ένας δείκτης ταύτισης των δύο εικόνων, όπου με βάση αυτόν θα καθοριστεί κατά πόσο θα εκτελεστεί ξανά ο αλγόριθμος. Συνθήκη τερματισμού μπορεί να είναι ο δείκτης αυτός να φτάσει σε μία επιθυμητή τιμή ή/και ο αριθμός των επαναλήψεων ή/και η διαφορά της παρούσας μήτρας μετασχηματισμού να είναι κατά μια ποσότητα (ϵ) διαφορετική από την προηγούμενη μήτρα μετασχηματισμού.

Το διάγραμμα ροής δεδομένων του αλγορίθμου είναι το εξής:



Εικόνα 4.1 : Λογικό Διάγραμμα προτεινόμενου αλγορίθμου με συμπαγή μετασχηματισμό ευθυγράμμισης τρισδιάστατων νεφών σημείων

4.2.1 Επιλογή Σημείων

Είναι σημαντική η αυτόματη επιλογή περιορισμένου πλήθους σημείων προς ευθυγράμμιση από τα δύο νέφη σημείων και όχι η εφαρμογή του αλγορίθμου σε όλα τα σημεία του νέφους. Με αυτή την επιλογή ενός πλήθους σημείων είναι δυνατή η ταχύτερη εκτέλεση του αλγορίθμου.

Η μείωση των σημείων επιτυγχάνεται με την κατάργηση σημείων τα οποία είναι σχετικά μακριά από άλλα σημεία. Επίσης, για την μείωση της υπολογιστικής πολυπλοκότητας, η ποσότητα των σημείων μπορεί να μειωθεί με μια τυχαία ή

ομοιόμορφη υπό-δειγματοληψία. Η ιδέα πίσω από τη τυχαία δειγματοληψία είναι η διαφορετική δειγματοληψία σε κάθε επανάληψη του αλγορίθμου με σκοπό να αποτρέψει όποια τάση προς ακραίες τιμές (38). Επίσης, επίπεδες περιοχές στο νέφος σημείων μπορεί να περιέχουν περιττές πληροφορίες οι οποίες ορισμένες από αυτές μπορούν να αποκοπούν.

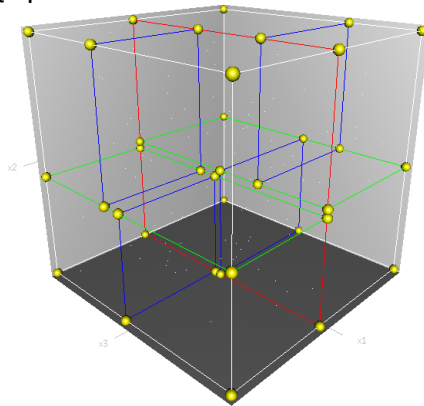
Μια επίσης καλή τεχνική για περισσότερη ακρίβεια στο αποτέλεσμα της ευθυγράμμισης, είναι η προσθήκη σημείων όπου πιθανόν να περιέχουν πλούσια πληροφορία. Σημεία όπως ακμές, γωνίες και καμπυλότητες στο νέφος σημείων μπορούν να δώσουν πολύ σημαντική πληροφορία. Έτσι μία καλή στρατηγική είναι η λήψη αυτών των σημείων με μεγαλύτερη βαρύτητα. Μία τεχνική για την μεγαλύτερη βαρύτητα αυτών των σημείων είναι η προσθήκη σημείων κοντά σε αυτά τα σημεία με υπέρ-δειγματοληψία στην περιοχή αυτή.

Ο αλγόριθμος που παρουσιάζεται στην παρούσα εργασία είναι βασισμένος στην αναφορά του Besl, όπου λαμβάνει υπόψη όλα τα διαθέσιμα σημεία, έτσι και ο αλγόριθμος αυτός δεν κάνει επιλογή σημείων.

4.2.2 Ταύτιση Σημείων

Μετά την επιλογή (ή μη επιλογή) των σημείων στα δύο σύνολα δεδομένων, πρέπει να εφαρμοστεί μία μέθοδος ταύτισης των σημείων από το ένα νέφος σημείων με τα ομόλογα σημεία στο δεύτερο νέφος σημείων. Η εύρεση των κοντινότερων γειτόνων ενός σημείου είναι συνήθως το πιο χρονοβόρο στάδιο του αλγορίθμου ICP του οποίου ο χρόνος εκτέλεσής του μπορεί να φτάσει και μέχρι το 95 % του συνολικού χρόνου εκτέλεσης όλου του αλγορίθμου (39). Η τεχνική ταιριάσματος σημείων που χρησιμοποιείται εδώ είναι με βάση την μέθοδο έρευνας με kD -trees (kD – Dimensional tree).

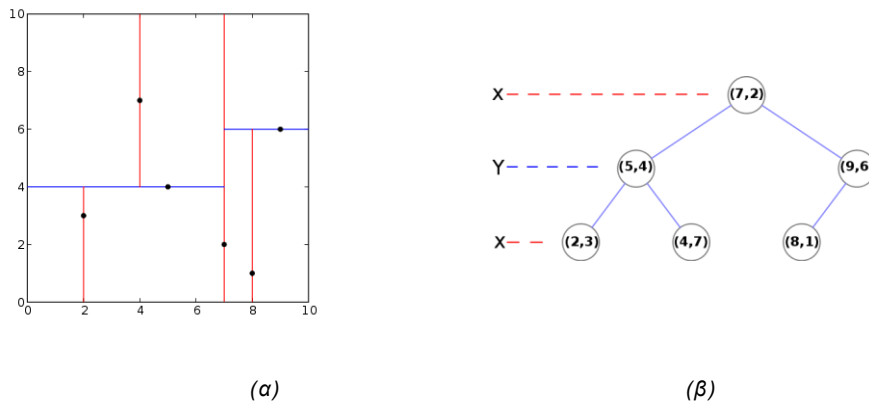
Η δομή kD -tree, είναι μια δομή δεδομένων για την οργάνωση των σημείων τα οποία βρίσκονται σε ένα k -διάστατο χώρο.



Εικόνα 4.2 : Τρισδιάστατη αναπαράσταση kD tree

Το δέντρο kD είναι ένα δυαδικό δέντρο στο οποίο ο κάθε κόμβος αντιπροσωπεύει ένα σημείο k -διαστάσεων. Έστω ότι βρισκόμαστε σε ένα χώρο k -διαστάσεων με τις τιμές των σημείων να έχουν k -διαστάσεις, δηλαδή, $(X,Y,Z,...,K)$. Η δημιουργία μια δομής σαν αυτή

γίνεται με τον συνεχή διαχωρισμό του χώρου σε μικρότερους υποχώρους με σημείο διαχωρισμού να είναι η συντεταγμένη ενός άξονα όπου είναι η μεσαία τιμή από όλες τις συντεταγμένες των σημείων αυτού του άξονα. Στην περίπτωση μας, βρίσκουμε το σημείο με την μεσαία τιμή με βάση τον άξονα X από όλα τα σημεία του νέφους. Όσα σημεία έχουν για τιμή συντεταγμένης X μεγαλύτερη από την μεσαία τιμή, ομαδοποιούνται σε ένα υπόχωρο, και όσα σημεία έχουν για τιμή συντεταγμένης X μικρότερη από την μεσαία τιμή, ομαδοποιούνται σε ένα άλλο υπόχωρο. Στην συνέχεια, ο κάθε ένας υπόχωρος, διαχωρίζεται με τον ίδιο τρόπο αλλά με μεσαία τιμή βάσει του επόμενου άξονα. Στην περίπτωση μας, σε κάθε υπόχωρο, βρίσκουμε το σημείο με την μεσαία τιμή με βάση τον άξονα Y από όλα τα σημεία του υποχώρου. Όσα σημεία έχουν για τιμή συντεταγμένης Y μεγαλύτερη από την μεσαία τιμή, ομαδοποιούνται σε ένα υπόχωρο, και όσα σημεία έχουν για τιμή συντεταγμένης Y μικρότερη από την μεσαία τιμή, ομαδοποιούνται σε ένα άλλο υπόχωρο. Η διαδικασία διαχωρισμού συνεχίζεται μέχρι όλοι οι υποχώροι να περιέχουν μόνο ένα σημείο. Όταν φτάσουμε στον τελευταίο άξονα των συντεταγμένων των σημείων, στην περίπτωση μας K, και κάποιος υποχώρος αποτελείται από περισσότερα από ένα σημεία, τότε συνεχίζουμε τον διαχωρισμό του υποχώρου, με βάση την πρώτη συντεταγμένη των σημείων X και κάνοντας την ίδια διαδικασία. Η μεσαία τιμή του κάθε υποχώρου, ορίζεται και ως η ρίζα του υπό-δέντρου του υποχώρου που αντιπροσωπεύει. Ένα δισδιάστατο παράδειγμα φαίνεται στο πιο κάτω σχήμα.



Εικόνα 4.3 : (α) Δισδιάστατος χώρος ο οποίος διαχωρίζεται με την μεθοδολογία της δημιουργίας *kD tree* κυκλικά ως προς την διάσταση X, Y και μετά πάλι από την αρχή

(β) Το *kD tree* σε μορφή γράφου το οποίο εξάγεται από την εφαρμογή του αλγορίθμου της δημιουργίας *kD tree* στο (α)

Η δομή αυτή εφαρμόζεται σε ένα από τα δύο σύνολα σημείων και για τα σημεία του άλλου συνόλου επιδιώκουμε να καθορίσουμε σε πια γειτονία του πρώτου συνόλου βρίσκονται. Η μεθοδολογία του πλησιέστερου γείτονα (Nearest Neighbor (NN) search) έχει σαν σκοπό την εύρεση του σημείου στην πιο πάνω δομή το οποίο είναι πλησιέστερο σε ένα σημείο το οποίο εισάγεται. Αυτή η έρευνα μπορεί να διεκπεραιωθεί αποτελεσματικά χρησιμοποιώντας τις ιδιότητες της δομής δέντρου για τον αποκλεισμό μεγάλων τμημάτων αυτού.

Τα «φύλλα» μίας ρίζας ενός δέντρου (ή υπό-δέντρου), διαχωρίζονται με βάση ένα άξονα του χώρου (X, Y ή Z). Σε κάθε ρίζα ενός δέντρου (ή υπό-δέντρου), ελέγχουμε την

συντεταγμένη του σημείου στον άξονα όπου αναφέρεται η υπό έρευνα ρίζα. Με βάση την τιμή της συντεταγμένης αυτής και με την τιμή της υπό εξέταση ρίζας, η έρευνα αυτή συνεχίζεται αναδρομικά στον δεξιό ή αριστερό κλάδο του δέντρου (ή υπό-δέντρου). Η έρευνα αυτή σταματάει όταν φτάσει σε κόμβο του δέντρου ο οποίος δεν έχει «φύλλα», δηλαδή ο κόμβος αυτός είναι «φύλλο», και το σημείο αυτό είναι και ο πλησιέστερος γείτονας του σημείου εισόδου.

4.2.3 Βαρύτητα Σημείων

Στο παρόν υποκεφάλαιο, μελετάται προσθήκη διαφορετικών τιμών βαρύτητας στα συσχετισμένα σημεία των δύο νεφών όπου βρέθηκαν από την προηγούμενη μέθοδο. Οι τεχνικές αυτές είναι:

- Σταθερή Βαρύτητα σε όλα τα σημεία
- Ανάλογα με την μεταξύ τους απόσταση. Μεγάλη βαρύτητα στα συσχετισμένα σημεία των οποίων η μεταξύ τους απόσταση είναι μικρή, ενώ αντίθετα, στα συσχετισμένα σημεία με μεγάλη απόσταση εισάγεται μικρή βαρύτητα.

$$Weight = 1 - \frac{Dist(p_i, q_i)}{\max Distance} \quad (4.3)$$

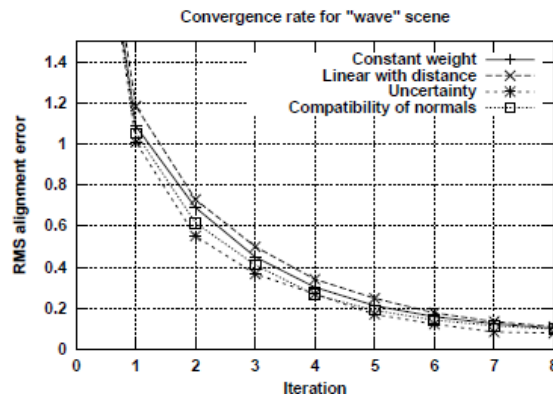
- Ανάλογα με την συμβατότητα των διανυσμάτων τους, δηλ. :

$$Weight = n_1 \cdot n_2 \quad (4.4)$$

Όπου n_i , το κάθετο διάνυσμα του i σημείου

- Στάθμιση με βάση την αναμενόμενη επίδραση του θορύβου σάρωσης στην αβεβαιότητα του σφάλματος.

Στην αναφορά (40), προστέθηκε επιπρόσθετος θόρυβος σε σύνολα δεδομένων με σκοπό την διαφορά μεταξύ των τεχνικών εισαγωγής βαρύτητας στα διάφορα συσχετισμένα σημεία και παρουσιάστηκε η αποτελεσματικότητα της ευθυγράμμισης σε κάθε τεχνική όπως φαίνεται στην Εικόνα 4.4.



Εικόνα 4.4 : Γραφική αναπαράσταση του σφάλματος ευθυγράμμισης για διάφορες μεθόδους εφαρμογής βαρών στα διάφορα σημεία του συνόλου

Όπως φαίνεται και στην Εικόνα 4.4, ακόμη και με την προσθήκη του επιπλέον θορύβου, όλες οι τεχνικές στάθμισης έχουν παρόμοιες επιδόσεις. Έτσι, η τεχνική που επιλέγεται για την εισαγωγή βαρύτητας στον αλγόριθμο που παρουσιάζεται σε αυτή την εργασία, επιλέγουμε να αναθέσουμε σε όλα σταθερό βάρος.

4.2.4 Απόρριψη Ζευγαριών

Η απόρριψη συσχετισμένων ζευγαριών σημείων είναι μια κατάλληλη μέθοδος για την αφαίρεση πιθανών σφαλμάτων στην συσχέτιση ζευγαριών. Το μέρος της επεξεργασίας αυτής γίνεται πριν από την εφαρμογή του αναδρομικού αλγόριθμου πλησιέστερων σημείων για την μείωση του πλήθους των συσχετισμένων σημείων έτσι ώστε να μειωθεί το υπολογιστικό κόστος της ευθυγράμμισης.

Στην περίπτωση όπου τα νέφη σημείων λαμβάνονται από κάμερες βασισμένες στην μέθοδο του δομημένου φωτός, τα νέφη τα οποία λαμβάνονται από διαφορετική τοποθεσία της κάμερας, δεν είναι πλήρως αλληλοκαλυπτώμενα. Για την αποφυγή σφαλμάτων λόγω μερικής αλληλοκάλυψης των νεφών σημείων, εφαρμόζουμε απόρριψη ζευγαριών τα οποία αναφέρονται σε σημεία της περιφέρειας κάθε νέφους σημείων.

Επίσης, αναλόγως της απόστασης που έχουν τα συσχετισμένα σημεία των δύο συνόλων, αυτά απορρίπτονται αν η απόστασή τους είναι μεγαλύτερη από ένα προκαθορισμένο κατώφλι. Στην περίπτωσή μας το κατώφλι είναι (0.1), η οποία τιμή αποφασίστηκε μετά από δοκιμές οι οποίες αναφέρονται στο κεφάλαιο 5, και άρα:

$$Rejected\ Pairs \rightarrow Dist(p_i, q_i) > threshold = 0.1 \forall i$$

Τέλος, για την απόρριψη συσχετισμένων ζευγαριών χρησιμοποιούμε την μέθοδο RANSAC (RANDOM Sample Consensus) η οποία εξηγείται πιο κάτω.

4.2.4.1 Μέθοδος RANSAC

Η μέθοδος RANSAC, είναι μια επαναληπτική μέθοδος για υπολογισμό παραμέτρων ενός μοντέλου από ένα σύνολο από δεδομένα παρατήρησης το οποίο περιέχει εσφαλμένες τιμές. Είναι μια μη-ντετερμινιστική τεχνική υπό την έννοια ότι παράγει ένα λογικό αποτέλεσμα μόνο με μία ορισμένη πιθανότητα, όπου αυτή η πιθανότητα αυξάνεται καθώς περισσότερες επαναλήψεις εκτελούνται.

Η βασική παραδοχή που γίνεται στην μέθοδο RANSAC, είναι ότι τα δεδομένα αποτελούνται από τιμές οι οποίες είναι σωστές, δηλαδή, τιμές δεδομένων οι οποίες μπορούν να αναπαριστούν παραμέτρους ενός μοντέλου αν και μπορεί να περιέχουν θόρυβο στις τιμές τους, όπως επίσης και εσφαλμένες τιμές που είναι δεδομένα που δεν ταιριάζουν στο μοντέλο. Τέτοια σφάλματα μπορούν να προέλθουν από μεγάλες τιμές θορύβου, από εσφαλμένες μετρήσεις ή και λανθασμένες υποθέσεις σχετικά με την ερμηνεία του μοντέλου. Η μεθοδολογία RANSAC επίσης θεωρεί ότι λαμβάνοντας υπόψη ένα μικρό σύνολο από σωστές τιμές, υπάρχει μια διαδικασία που μπορεί να εκτιμήσει τις παραμέτρους του μοντέλου το οποίο βέλτιστα εξηγεί ή ταιριάζει αυτά τα δεδομένα.

Σαν είσοδος στον αλγόριθμο RANSAC χρησιμοποιείται το σύνολο σημείων προς ευθυγράμμιση, το σύνολο αναφοράς και οι παράμετροι εμπιστοσύνης. Η μέθοδος

RANSAC επιτυγχάνει τον σκοπό της με την επαναληπτική επιλογή τυχαίων ελάχιστων υποσυνόλων από το μοντέλο αναφοράς. Στην περίπτωση μας, ο ελάχιστος αριθμός δειγμάτων είναι οχτώ, όπου αυτά τα δεδομένα είναι υποθετικά σωστά και αυτή η υπόθεση δοκιμάζεται ως εξής:

1. Αρχικά, το υποσύνολο με τον ελάχιστο αριθμό δειγμάτων ορίζει ένα υποθετικό μοντέλο
2. Ακολούθως, όλα τα άλλα δείγματα δοκιμάζονται με το προσαρμοσμένο μοντέλο, και αυτά τα σημεία τα οποία προσαρμόζονται καλά στο εκτιμώμενο μοντέλο, θεωρούνται ως σωστά δεδομένα, και όλα όσα δεν προσαρμόζονται καλά, ορίζονται σαν εσφαλμένα δεδομένα. Η απόσταση από το εκτιμώμενο μοντέλο με την οποία θα θεωρούνται ως σωστά δεδομένα, εισάγεται σαν παράμετρος στον αλγόριθμο
3. Τα βήματα 1 και 2 εκτελούνται για N επαναλήψεις, όπου ο αριθμός επαναλήψεων ορίζεται με βάση την ακρίβεια την οποία επιθυμούμε και θα οριστεί πιο κάτω. Συγκρίνοντας τα σωστά και τα εσφαλμένα σημεία σε κάθε επανάληψη, επιλέγουμε το μοντέλο το οποίο έχει τα περισσότερα σωστά στοιχεία
4. Επιπρόσθετα, μπορεί να εφαρμοστεί για ακόμη μια φορά ο αλγόριθμος με είσοδο μόνο τα δεδομένα τα οποία θεωρήθηκαν σωστά από τη βέλτιστη επανάληψη για περαιτέρω ακρίβεια

Ο αριθμός των επαναλήψεων N του αλγορίθμου μπορεί να προσδιοριστεί με βάση την πιθανότητα (p), η οποία δηλώνει την πιθανότητα, σε κάποια επανάληψη του αλγορίθμου, όλα τα δείγματα (n) να είναι δείγματα με σωστή πληροφορία. Αν υποθέσουμε ότι η πιθανότητα (w) είναι η πιθανότητα ένα σημείο να είναι σημείο με σωστή πληροφορία από όλα τα σημεία του συνόλου, δηλαδή:

$$w = \frac{\text{αριθμός σωστών πληροφοριών}}{\text{πλήθος όλων των σημείων στο σύνολο}} \quad (4.5)$$

Μπορεί η πιθανότητα w να μην είναι δυνατό να υπολογιστεί, έτσι μπορούμε να θεωρήσουμε μια προσέγγιση της πιθανότητας. Υποθέτοντας ότι τα n σημεία για τον υπολογισμό του μοντέλου επιλέγονται ανεξάρτητα, τότε w^n , είναι η πιθανότητα όλα τα σημεία αυτά να είναι σωστή πληροφορία και $1 - w^n$, είναι η πιθανότητα τουλάχιστο ένα σημείο να είναι εσφαλμένη πληροφορία. Η πιθανότητα αυτή, υψωμένη στην δύναμη N , είναι η πιθανότητα ο αλγόριθμος να μην επιλέξει ποτέ ένα σύνολο από σημεία τα οποία να είναι όλα σωστές πληροφορίες και πρέπει να ισούται με $1 - p$, άρα:

$$1 - p = (1 - w^n)^N \quad (4.6)$$

Και λύνοντας ως προς τον αριθμό των επαναλήψεων έχουμε:

$$\Rightarrow N = \frac{\log(1 - p)}{\log(1 - w^n)} \quad (4.7)$$

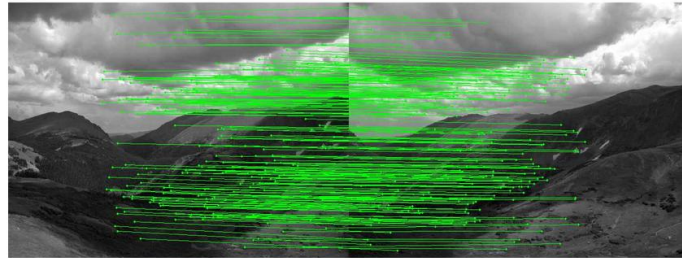
Αν υποθέσουμε ότι τα διανύσματα x και x' αναπαριστούν τα διανύσματα των συσχετισμένων σημείων στα δύο νέφη εικόνων, η σχέση αυτών με την θεμελιώδη μήτρα (F) είναι:

$$x'^T F x = 0 \quad (4.8)$$

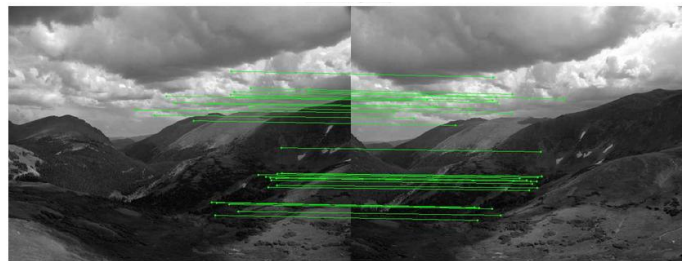
Έτσι ο επαναληπτικός αλγόριθμος RANSAC εκτελεί N φορές:

1. Επιλογή ενός τυχαίου υποσυνόλου από 8 δείγματα από το σύνολο αναφοράς τα οποία έχουν σημεία συσχέτισης στο σύνολο προς ευθυγράμμιση
2. Υπολογισμός της θεμελιώδους μήτρας F και υπολογισμός του υπερεπιπέδου το οποίο δημιουργείται από την θεμελιώδη μήτρα
3. Εύρεση του πλήθους των συσχετισμένων ζευγαριών τα οποία ανήκουν στο μοντέλο που υπολογίζεται από τη θεμελιώδη μήτρα

Στο τέλος επιλέγεται το υποσύνολο σημείων το οποίο έχει το μεγαλύτερο πλήθος σωστής πληροφορίας. Στην μέθοδο αυτή θέτουμε σαν ελάχιστη απαίτηση να βρει τουλάχιστο τρία συσχετισμένα ζεύγη σημείων. Αποτελέσματα της μεθόδου φαίνονται στην Εικόνα 4.5.



(α)



(β)

Εικόνα 4.5 : (α) Αναπαράσταση της εικόνας αναφοράς και της εικόνας προς ευθυγράμμισης, μαζί με τα συσχετισμένα σημεία μεταξύ των εικόνων πριν από την εφαρμογή μεθόδου RANSAC

(β) Αναπαράσταση της εικόνας αναφοράς και της εικόνας προς ευθυγράμμισης, μαζί με τα συσχετισμένα σημεία μεταξύ των εικόνων μετά από την εφαρμογή μεθόδου RANSAC

4.2.5 Συνάρτηση μέτρησης σφάλματος

Το τελευταίο μέρος του αλγορίθμου ICP είναι η εύρεση της τιμής μια συνάρτησης μέτρησης σφάλματος για την ελαχιστοποίησή της. Η συνάρτηση σφάλματος που χρησιμοποιείται είναι το άθροισμα διαφορών τετραγώνων μεταξύ των συσχετισμένων σημείων. Η συνάρτηση αυτή χρησιμοποιείται για συμπαγή μετασχηματισμούς. Η λύση της συνάρτησης αυτής βρίσκεται με την μέθοδο Singular Value Decomposition (SVD). Η συνάρτηση έχει την μορφή:

$$E = \sum_{i=1}^N \|\mathbf{R}p_i + T - q_i\|^2 \quad (4.9)$$

4.2.6 Ελαχιστοποίηση Συνάρτησης Σφάλματος

Ας υποθέσουμε ότι p_i και q_i υποδηλώνουν τα N ζευγάρια ταιριασμένων σημείων. Υπολογίζουμε το κέντρο μάζας του συνόλου αναφοράς και του συνόλου προς ευθυγράμμιση ως εξής:

$$\bar{p} = \frac{1}{m} \sum p \quad (4.10)$$

$$\bar{q} = \frac{1}{n} \sum q \quad (4.11)$$

Όπου m και n είναι ο αριθμός των σημείων του μοντέλου αναφοράς και του προς ευθυγράμμιση μοντέλου αντίστοιχα. Οι αποκλίσεις των σημείων από τα κέντρα μάζας στα δύο σύνολα είναι:

$$p'_i = p_i - \bar{p} \quad (4.12)$$

$$q'_i = q_i - \bar{q} \quad (4.13)$$

Σκοπός είναι η εύρεση της μήτρας περιστροφής και το διάνυσμα μετακίνησης, έτσι ώστε να ελαχιστοποιηθεί η συνάρτηση σφάλματος:

$$E = \sum_{i=1}^N \|\mathbf{R}p_i + T - q_i\|^2 \quad (4.14)$$

Χρησιμοποιώντας τις εξισώσεις (αποκλίσεις), η συνάρτηση σφάλματος μπορεί να γραφτεί ως:

$$E = \sum_{i=1}^N \|\mathbf{R}(p'_i + \bar{p}) + T - (q'_i + \bar{q})\|^2 = \sum_{i=1}^N \|\mathbf{R}p'_i - q'_i + (\mathbf{R}\bar{p} - \bar{q} + T)\|^2 \quad (4.15)$$

Για την ελαχιστοποίηση της συνάρτησης σφάλματος, το διάνυσμα μετακίνησης θα πρέπει να επιλεγεί έτσι ώστε να μετακινεί τα περιστρεμμένα δεδομένα στο κέντρο μάζας του μοντέλου αναφοράς, άρα:

$$T = \bar{q} - \mathbf{R}\bar{p} \quad (4.16)$$

Έτσι η συνάρτηση σφάλματος γίνεται:

$$E = \sum_{i=1}^N \|\mathbf{R}p'_i - q'_i\|^2 = \mathbf{R}\mathbf{R}^T \sum_{i=1}^N \|p'_i\|^2 - 2 \cdot \text{tr} \left(\mathbf{R} \sum_{i=1}^N p'_i q'_i{}^T \right) + \sum_{i=1}^N \|q'_i\|^2 \quad (4.17)$$

$$\Rightarrow E = \sum_{i=1}^N \|p'_i\|^2 - 2 \cdot \text{tr} \left(\mathbf{R} \sum_{i=1}^N p'_i q'_i{}^T \right) + \sum_{i=1}^N \|q'_i\|^2 \quad (4.18)$$

Αν θέσουμε $\mathbf{N} = \sum_{i=1}^N p'_i q'_i{}^T$, για να ελαχιστοποιηθεί το σφάλμα E , το ίχνος $\text{tr}(\mathbf{R}\mathbf{N})$ ($\text{tr}(A) = \sum_{i=1}^N a_{ii}$) θα πρέπει να μεγιστοποιηθεί. Αν συμβολίσουμε τις στήλες του \mathbf{N} και τις γραμμές του \mathbf{R} με c_i και r_i αντίστοιχα όπου $i \in \{1,2,3\}$, τότε με βάση την ανισότητα του Cauchy-Schwarz, υπολογίζεται το ίχνος:

$$\text{tr}(\mathbf{RN}) = \sum_{i=1}^3 r_i \cdot c_i \leq \sum_{i=1}^3 \|r_i\| \cdot \|c_i\| \quad (4.19)$$

Έχοντας υπόψη ότι η μήτρα μετασχηματισμού είναι ορθογωνική από τον ορισμό της, τα διανύσματα γραμμής της θα έχουν μοναδιαίο μήκος, έτσι:

$$\text{tr}(\mathbf{RN}) \leq \sum_{i=1}^3 \sqrt{c_i^T c_i} = \text{tr}(\sqrt{\mathbf{N}^T \mathbf{N}}) \quad (4.20)$$

Αν θεωρήσουμε την SVD της μήτρας \mathbf{N} :

$$\mathbf{N} = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^T \quad (4.21)$$

Όπου \mathbf{U} και \mathbf{V} είναι ορθογωνικές μήτρες για τις οποίες ισχύουν οι εξής σχέσεις:

$$\mathbf{U}^T = \mathbf{U}^{-1}, \mathbf{V}^T = \mathbf{V}^{-1} \quad (4.22)$$

Οι στήλες της μήτρας \mathbf{V} περιέχουν τα ιδιοδιανύσματα της μήτρας $\mathbf{N}^T \mathbf{N}$ και οι στήλες της μήτρας \mathbf{U} περιέχουν τα ιδιοδιανύσματα της μήτρας $\mathbf{N} \mathbf{N}^T$.

Επιλέγοντας τότε την μήτρα περιστροφής:

$$\mathbf{R} = \mathbf{V} \mathbf{U}^T \quad (4.23)$$

Το ίχνος γίνεται:

$$\text{tr}(\mathbf{V} \mathbf{U}^T \mathbf{U} \mathbf{\Sigma} \mathbf{V}^T) = \text{tr}(\mathbf{V} \mathbf{\Sigma} \mathbf{V}^{-1}) = \text{tr}(\sqrt{\mathbf{V} \mathbf{\Sigma}^T \mathbf{\Sigma} \mathbf{V}^{-1}}) = \text{tr}(\sqrt{\mathbf{N}^T \mathbf{N}}) \quad (4.24)$$

Όπου με βάση την (4.20) είναι το μέγιστο δυνατό.

4.2.7 Συνθήκη Τερματισμού

Ο αλγόριθμος εκτελεί επαναλήψεις των πιο πάνω βημάτων μέχρι να ικανοποιηθούν κάποιες συνθήκες για τον τερματισμό του. οι συνθήκες στις οποίες ο αλγόριθμος τερματίζει τον επαναληπτικό του χαρακτήρα είναι οι:

1. Μέγιστος αριθμός επαναλήψεων: Μπορεί να εισαχθεί από τον χρήστη μέσω του γραφικού περιβάλλοντος ή τήρηση της προεπιλεγμένης τιμής των 100 επαναλήψεων
2. Τιμή συναρτήσεως Σφάλματος: Όταν η τιμή της συνάρτησης σφάλματος είναι μικρότερη από ένα κατώφλι το οποίο ορίζεται ως το άθροισμα των διαφορών των συσχετισμένων σημείων διαιρεμένο με τον αριθμό των συσχετισμένων σημείων, τότε ο αλγόριθμος σταματά
3. Μεταβολή Μήτρας Μετασχηματισμού: Το άθροισμα των απόλυτων διαφορών μεταξύ της παρούσας μήτρας μετασχηματισμού και της προηγούμενης μήτρας μετασχηματισμού, όταν φτάσει σε μία ελάχιστη τιμή η οποία καθορίζεται από τον χρήστη, τότε σταματάει ο αλγόριθμος

4.3 Μη Συμπαγής ICP

Σε αυτό το κεφάλαιο θα περιγραφεί η μέθοδος ευθυγράμμισης με μη συμπαγή μετασχηματισμό ICP ο οποίος χρησιμοποιεί την μέθοδο Levenberg-Marquardt. Γενικά ο μη συμπαγής μετασχηματισμός εκτελείται σε δύο βήματα. Αρχικά δημιουργούνται τα συσχετισμένα σημεία εφαρμόζοντας τον συμπαγή μετασχηματισμό, και στην συνέχεια

υπολογίζεται η μήτρα μετασχηματισμού για την ελαχιστοποίηση της Ευκλείδειας απόστασης μεταξύ των συσχετισμένων σημείων. Αυτά τα δύο στάδια επαναλαμβάνονται μέχρι η ακρίβεια της ευθυγράμμισης να φτάσει σε επιθυμητά επίπεδα. Η μέθοδος αυτή μοντελοποιεί την ευθυγράμμιση σαν ένα γενικό πρόβλημα βελτιστοποίησης και προτάθηκε η μέθοδος αυτή με σκοπό την ελαχιστοποίηση του σφάλματος ευθυγράμμισης με την εφαρμογή μη γραμμικών διαδικασιών βελτιστοποίησης. Το πλεονέκτημα της μεθόδου αυτής είναι η ευελιξία στον ορισμό της συνάρτησης βελτιστοποίησης προκειμένου να ληφθούν υπόψη διάφορες πτυχές της ταξινόμησης, όπως η απόρριψη ακραίων τιμών, και η ταχύτητα.

4.3.1 Levenberg-Marquardt ICP

Το γενικό πρόβλημα διαμορφώνεται όπως περιγράφηκε και στον συμπαγή ICP. Η συνάρτηση σφάλματος είναι μια μη γραμμική συνάρτησης ελαχίστων τετραγώνων και μπορεί να γραφεί ως άθροισα των N_d τετραγώνων υπολοίπων:

$$E(\mathbf{a}) = \sum_{i=1}^{N_d} \|\mathbf{R}p_i + T - q_i\|^2 = \sum_{i=1}^{N_d} (e_i(\mathbf{a}))^2 \quad (4.25)$$

Ορίζοντας το διάνυσμα υπολοίπων ως:

$$\mathbf{e}(\mathbf{a}) = \{e_i(\mathbf{a})\}_{i=1}^{N_d} \quad (4.26)$$

Έτσι μπορούμε να ξαναγράψουμε την (4.25) :

$$E(\mathbf{a}) = \|\mathbf{e}(\mathbf{a})\|^2 \quad (4.27)$$

Ο αλγόριθμος Levenberg-Marquardt συνδυάζει την τεχνική «κλίσης καθόδου» (gradient descent) και την Gauss – Newton. Σκοπός της κάθε επανάληψης είναι η επιλογή μίας ανανέωσης \mathbf{x} της παρούσας εκτίμησης \mathbf{a}_k έτσι ώστε η επόμενη εκτίμηση $\mathbf{a}_{k+1} = \mathbf{a}_k + \mathbf{x}$ να μειώνει το σφάλμα ευθυγράμμισης.

Αρχικά ορίζουμε την ανανέωση της τεχνικής Gauss – Newton. Αναπτύσσοντας το $E(\mathbf{a} + \mathbf{x})$ σε δεύτερης τάξης έχουμε:

$$E(\mathbf{a} + \mathbf{x}) = E(\mathbf{a}) + (\nabla E(\mathbf{a})\mathbf{x}) + \frac{1}{2!}((\nabla^2 E(\mathbf{a})\mathbf{x})\mathbf{x}) + \text{high order terms (h. o. t.)} \quad (4.28)$$

Και ορίζοντας τα:

$$E(\mathbf{a}) = \mathbf{e}^T \mathbf{e} \quad (4.29)$$

$$\nabla E(\mathbf{a}) = 2(\nabla \mathbf{e})^T \mathbf{e} \quad (4.30)$$

$$\nabla^2 E(\mathbf{a}) = 2(\nabla^2 \mathbf{e})\mathbf{e} + 2(\nabla \mathbf{e})^T \nabla \mathbf{e} \quad (4.31)$$

Ορίζοντας την $N_d \times p$ Jacobian ως:

$$J = \nabla \mathbf{e} \quad (4.32)$$

$$J_{i,j} = \frac{\partial E_i}{\partial \mathbf{a}_j} \quad (4.33)$$

Όπου p είναι ο αριθμός των στοιχείων στο \mathbf{a} . Έτσι η προσέγγιση Gauss – Newton αμελώντας τα $(\nabla^2 \mathbf{e})\mathbf{e}$ γίνεται:

$$E(\mathbf{a} + \mathbf{x}) \approx \mathbf{e}^T \mathbf{e} + \mathbf{x}^T J^T \mathbf{e} + \mathbf{x}^T J^T J \mathbf{x} \quad (4.34)$$

Και διαφοροποιώντας ως προς το \mathbf{x} και μηδενίζοντας την παράγωγο έχουμε:

$$\nabla_{\mathbf{x}} E(\mathbf{a} + \mathbf{x}) = J^T \mathbf{e} + J^T J \mathbf{x} = 0 \quad (4.34)$$

Το οποίο μας δίνει την ανανέωση της μεθόδου Gauss – Newton:

$$\mathbf{x}_{GN} = -(\mathbf{J}^T \mathbf{J})^{-1} \mathbf{J}^T \mathbf{e} \quad (4.35)$$

Η μέθοδος Gauss – Newton είναι συνήθως γρήγορη στα μη γραμμικά προβλήματα αλλά δεν εγγυάται την σύγκλιση.

Στην συνέχεια ορίζουμε την ανανέωση της τεχνικής «κλίσης- καθόδου». Λόγω του ότι αντιμετωπίζουμε πρόβλημα ελαχίστων τετραγώνων, η ανανέωση αυτής της τεχνικής είναι:

$$\mathbf{x}_{GD} = -\lambda^{-1} \mathbf{J}^T \mathbf{e} \quad (4.36)$$

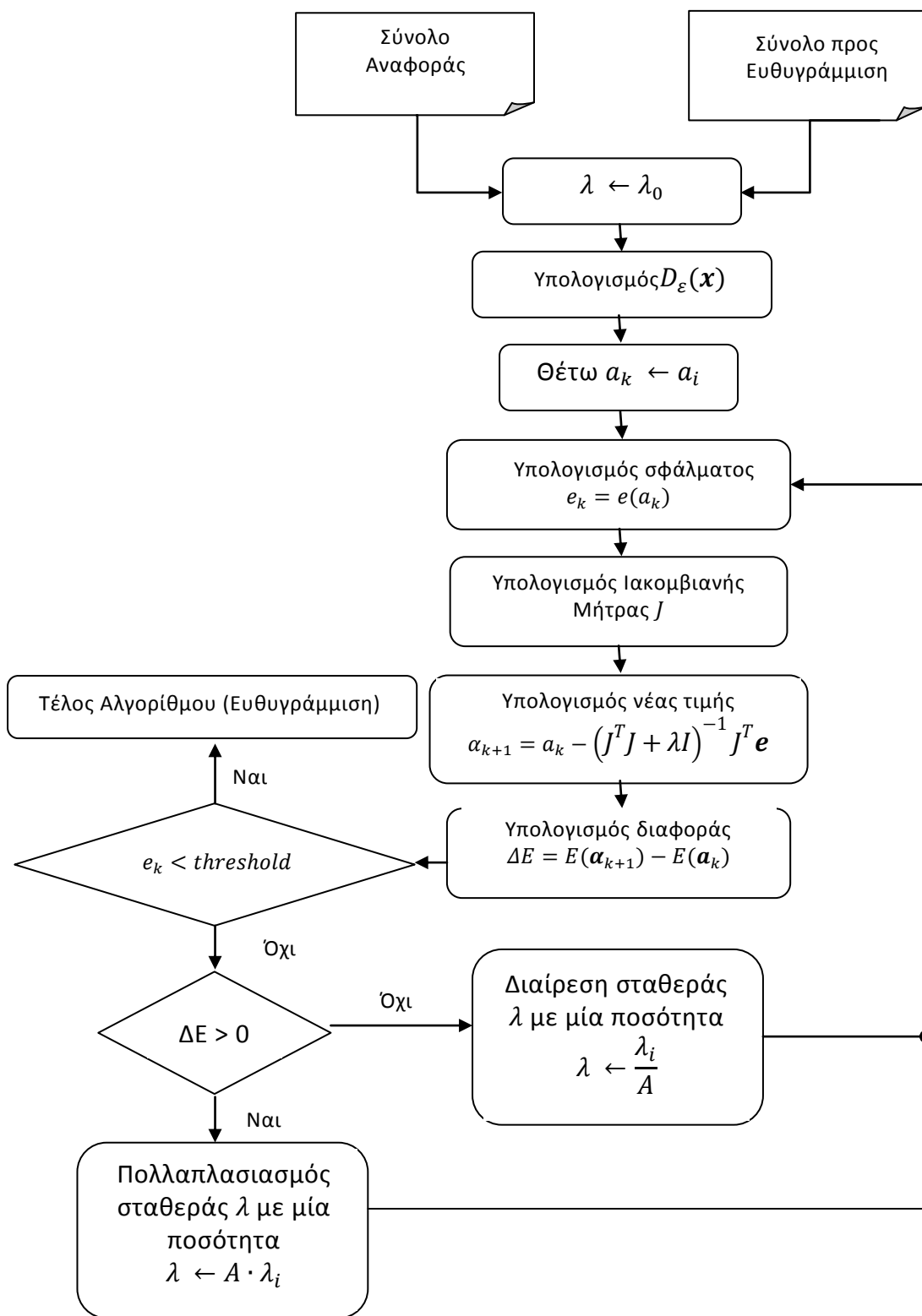
Όπου το λ είναι το αντίστροφο του μήκους βήματος. Η μέθοδος αυτή έχει σαν ιδιότητα, αν δεν φτάσει σε τοπικό ελάχιστο, μπορεί πάντα να μειώνει το σφάλμα με την μείωση του μήκος βήματος. Από την άλλη όμως, η μέθοδος αυτή είναι αργή.

Η ευθυγράμμιση Levenberg-Marquardt συνδυάζει τις ανανεώσεις και των δύο τεχνικών και έχει σαν ανανέωση:

$$\mathbf{x}_{LM} = -(\mathbf{J}^T \mathbf{J} + \lambda \mathbf{I})^{-1} \mathbf{J}^T \mathbf{e} \quad (4.37)$$

Για μεγάλες τιμές της παραμέτρου λ , έχουμε μικρό μήκος βήματος στην μέθοδο «κλίσης- καθόδου» και άρα ασφάλεια στην σύγκλιση σε ελάχιστο, ενώ για μεγάλες τιμές του λ , έχουμε μεγαλύτερα και πιο ακριβή βήματα στην τεχνική Gauss – Newton, η οποία συγκλίνει πιο γρήγορα στο ολικό ελάχιστο. Έτσι κατά την εφαρμογή του αλγορίθμου Levenberg-Marquardt γίνεται μεγάλη αύξηση του λ όταν το σφάλμα σε μία επανάληψη αυξηθεί, και αντίστοιχα μεγάλη μείωση του λ όταν το σφάλμα μειώνεται.

Το διάγραμμα ροής δεδομένων του αλγορίθμου είναι το εξής:



Εικόνα 4.6 : Λογικό Διάγραμμα προτεινόμενου αλγορίθμου με μη συμπαγή μετασχηματισμό ευθυγράμμισης τρισιδιάστατων νεφών σημείων

4.3.2 Υπολογισμός Παραγώγων Αλγόριθμου

4.3.2.1 Μετασχηματισμός Απόστασης

Η συνάρτηση απόστασης δίνεται από την εξίσωση:

$$D_\varepsilon(\mathbf{x}) = \min_j \varepsilon^2(\|\mathbf{m}_j - \mathbf{x}\|) \quad (4.38)$$

και σε συνδυασμό με την εξίσωση (4.25), η νέα μορφή του προβλήματος ευθυγράμμισης είναι:

$$E(\mathbf{a}) = \sum_{i=1}^{N_d} D_\varepsilon(T(\mathbf{a}, \mathbf{d}_i)) \quad (4.39)$$

Σκοπός μας είναι ο υπολογισμός των παραγώγων της συνάρτησης σφάλματος ($\nabla_{\mathbf{a}} E(\mathbf{a})$):

$$\nabla_{\mathbf{a}} E(\mathbf{a}) = \sum_{i=1}^{N_d} \nabla_{\mathbf{x}} D_\varepsilon(T(\mathbf{a}, \mathbf{d}_i)) \cdot \nabla_{\mathbf{a}}^T T(\mathbf{a}, \mathbf{d}_i) \quad (4.40)$$

Λόγο της διακριτότητας του προβλήματος, αρχικά οι παράγωγοι του μετασχηματισμού απόστασης ($\nabla_{\mathbf{x}} D_\varepsilon = [\frac{\partial D_\varepsilon}{\partial x}, \frac{\partial D_\varepsilon}{\partial y}, \frac{\partial D_\varepsilon}{\partial z}]$) υπολογίζονται πιο εύκολα με την χρήση διαφορών, δηλ.:

$$\frac{\partial D_\varepsilon(x, y, z)}{\partial x} = \frac{D_\varepsilon(x+1, y, z) - D_\varepsilon(x-1, y, z)}{2} \quad (4.41)$$

$$\frac{\partial D_\varepsilon(x, y, z)}{\partial y} = \frac{D_\varepsilon(x, y+1, z) - D_\varepsilon(x, y-1, z)}{2} \quad (4.42)$$

$$\frac{\partial D_\varepsilon(x, y, z)}{\partial z} = \frac{D_\varepsilon(x, y, z+1) - D_\varepsilon(x, y, z-1)}{2} \quad (4.43)$$

Παρατηρούμε ότι ο υπολογισμός του $\nabla_{\mathbf{a}}^T T(\mathbf{a}, \mathbf{d}_i)$ εξαρτάται από τις παραμέτρους του συμπαγή μετασχηματισμού. Ο υπολογισμός των παραγώγων αυτών υπολογίζονται αναλυτικά με βάση την αναφορά (56), με την βοήθεια της θεωρίας των quaternions.

4.3.2.2 Quaternions

Υποθέτουμε ότι ένα quaternion ορίζεται ως $\mathbf{q} = [s, \mathbf{v}]$, όπου s είναι ένα βαθμωτό μέγεθος και \mathbf{v} είναι ένα διάνυσμα (56). Αν υποθέσουμε ένα σημείο \mathbf{d} στο οποίο θα πρέπει να εφαρμοστεί περιστροφή $\mathbf{v} = [v_x, v_y, v_z]$, το σημείο αυτό θα πρέπει να αναπαρασταθεί στο χώρο των quaternions ως $\mathbf{r} = [0, \mathbf{d}]$. Έτσι το περιστρεμμένο σημείο λαμβάνεται από:

$$\mathbf{r}' = \mathbf{q} \mathbf{r} \mathbf{q}^{-1} = [0, s^2 \mathbf{d} + (\mathbf{d} \cdot \mathbf{v}) \cdot \mathbf{v} + 2s(\mathbf{v} \times \mathbf{d}) + \mathbf{v} \times (\mathbf{v} \times \mathbf{d})] = [0, T_x, T_y, T_z] \quad (4.44)$$

Όπου:

$$T_x = (s^2 + v_x^2 - v_y^2 - v_z^2)d_x + 2(v_x v_y - s v_z)d_y + 2(v_x v_z + s v_y)d_z \quad (4.45)$$

$$T_y = 2(v_x v_y + s v_z)d_x + (s^2 - v_x^2 + v_y^2 - v_z^2)d_y + 2(v_y v_z - s v_x)d_z \quad (4.46)$$

$$T_z = 2(v_x v_z + s v_y)d_x + 2(v_z v_y - s v_x)d_y + (s^2 - v_x^2 - v_y^2 + v_z^2)d_z \quad (4.47)$$

Με την εισαγωγή και ενός διανύσματος μετατόπισης $\mathbf{t} = [t_x, t_y, t_z]$, οι κανονικοποιημένη μορφή των εξισώσεων (4.45), (4.46), (4.47) μπορεί να οριστεί ως:

$$T_x = \frac{(s^2 + v_x^2 - v_y^2 - v_z^2)d_x}{s^2 + v_x^2 + v_y^2 + v_z^2} + \frac{2(v_x v_y - s v_z)d_y}{s^2 + v_x^2 + v_y^2 + v_z^2} + \frac{2(v_x v_z + s v_y)d_z}{s^2 + v_x^2 + v_y^2 + v_z^2} + t_x \quad (4.48)$$

$$T_y = \frac{2(v_x v_y + s v_z)d_x}{s^2 + v_x^2 + v_y^2 + v_z^2} + \frac{(s^2 - v_x^2 + v_y^2 - v_z^2)d_y}{s^2 + v_x^2 + v_y^2 + v_z^2} + \frac{2(v_y v_z - s v_x)d_z}{s^2 + v_x^2 + v_y^2 + v_z^2} + t_y \quad (4.49)$$

$$T_z = \frac{2(v_x v_z + s v_y)d_x}{s^2 + v_x^2 + v_y^2 + v_z^2} + \frac{2(v_y v_z - s v_x)d_y}{s^2 + v_x^2 + v_y^2 + v_z^2} + \frac{(s^2 - v_x^2 - v_y^2 + v_z^2)d_z}{s^2 + v_x^2 + v_y^2 + v_z^2} + t_z \quad (4.50)$$

Και ο μήτρα $\nabla_{\mathbf{a}}^T T(\mathbf{a}, \mathbf{d}_i)$ είναι της μορφής:

$$\nabla_{\mathbf{a}}^T T(\mathbf{a}, \mathbf{d}_i) = \begin{pmatrix} \frac{\partial T_x}{\partial s} & \frac{\partial T_x}{\partial v_x} & \frac{\partial T_x}{\partial v_y} & \frac{\partial T_x}{\partial v_z} & \frac{\partial T_x}{\partial t_x} & \frac{\partial T_x}{\partial t_y} & \frac{\partial T_x}{\partial t_z} \\ \frac{\partial T_y}{\partial s} & \frac{\partial T_y}{\partial v_x} & \frac{\partial T_y}{\partial v_y} & \frac{\partial T_y}{\partial v_z} & \frac{\partial T_y}{\partial t_x} & \frac{\partial T_y}{\partial t_y} & \frac{\partial T_y}{\partial t_z} \\ \frac{\partial T_z}{\partial s} & \frac{\partial T_z}{\partial v_x} & \frac{\partial T_z}{\partial v_y} & \frac{\partial T_z}{\partial v_z} & \frac{\partial T_z}{\partial t_x} & \frac{\partial T_z}{\partial t_y} & \frac{\partial T_z}{\partial t_z} \end{pmatrix} \quad (4.51)$$

Έχοντας υπολογίσει τις ποσότητες $\nabla_x D_\varepsilon(T(\mathbf{a}, \mathbf{d}_i)), \nabla_{\mathbf{a}}^T T(\mathbf{a}, \mathbf{d}_i)$ μπορούμε να υπολογίσουμε και την Ιακομβιανή μήτρα:

$$J_{i,j} = \nabla_x D_\varepsilon(T(\mathbf{a}, \mathbf{d}_i)) \cdot \nabla_{\mathbf{a}_j}^T T(\mathbf{a}, \mathbf{d}_i) \quad (4.52)$$

όπου:

$$\nabla_{\mathbf{a}_j}^T T(\mathbf{a}, \mathbf{d}_i) = \left[\frac{\partial T_x(\mathbf{a}, \mathbf{d}_i)}{\partial \alpha_j}, \frac{\partial T_y(\mathbf{a}, \mathbf{d}_i)}{\partial \alpha_j}, \frac{\partial T_z(\mathbf{a}, \mathbf{d}_i)}{\partial \alpha_j} \right] \quad (4.53)$$

5 Αποτελέσματα

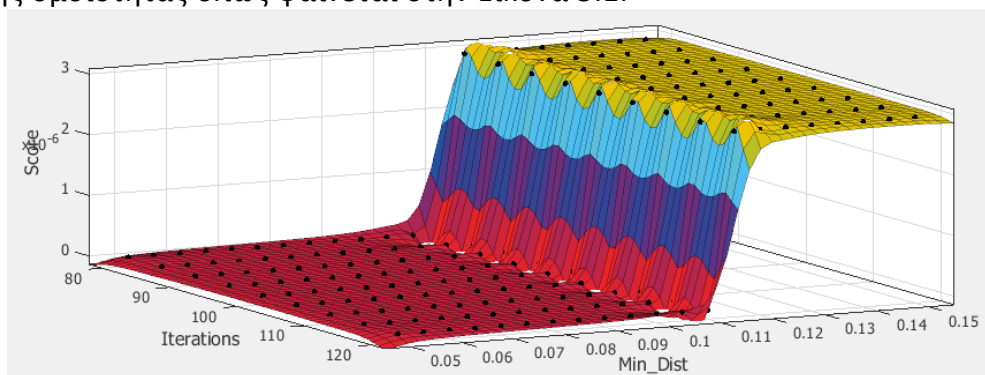
5.1 Καθορισμός Βέλτιστων Παραμέτρων Ευθυγράμμισης

Στον προτεινόμενο αλγόριθμο για ευθυγράμμιση τρισδιάστατων επιφανειών, καλείται ο χρήστης να εισάγει το πλήθος των επαναλήψεων του αλγορίθμου και την ελάχιστη επιτρεπτή απόσταση μεταξύ των σημείων όπου επιλέχθηκαν.

Με σκοπό τον προσδιορισμό των τιμών αυτών των παραμέτρων για μία ικανοποιητική ευθυγράμμιση, ο αλγόριθμος εκτελέστηκε για ένα μεγάλο εύρος από τιμές των δύο αυτών παραμέτρων με καθορισμένο πίνακα μετασχηματισμού. Οι τιμές των δύο παραμέτρων βρίσκονταν στα εξής σύνολα:

$$\begin{aligned} \text{Επαναλήψεις} &\in \{80 \sim 120\} \\ \text{Ελάχιστη Απόσταση} &\in \{0.05 \sim 0.1\} \end{aligned}$$

Σε κάθε εκτέλεση του αλγορίθμου, υπολογίζεται η συνάρτηση ομοιότητας του αλγορίθμου και στην συνέχεια δημιουργείται ένα τρισδιάστατο γράφημα του οποίου οι άξονες είναι το πλήθος των επαναλήψεων, η ελάχιστη δυνατή απόσταση και η τιμή της συνάρτησης ομοιότητας όπως φαίνεται στην Εικόνα 5.1.



Εικόνα 5.1 : Αναπαράσταση τιμής συνάρτησεως ομοιότητας για μεταβλητό αριθμό επαναλήψεων και ελάχιστη επιτρεπτή απόσταση σημείων

Καθώς η τιμή της συνάρτησης ομοιότητας μειώνεται, αυτό συνεπάγεται ότι οι δύο τρισδιάστατες επιφάνειες ευθυγραμμίζονται πλήρως. Για την πλήρη ευθυγράμμιση των δύο επιφανειών απαιτείται ένας μεγάλος αριθμός επαναλήψεων και η ελάχιστη δυνατή απόσταση να τείνει στο μηδέν, όπου αυτό θα προκαλούσε αύξηση της διάρκειας εκτέλεσης του αλγορίθμου και την αύξηση των υπολογιστικών απαιτήσεων. Σκοπός μας είναι η επιλογή των τιμών των παραμέτρων έτσι ώστε να επιτυγχάνεται μία αποδεκτή ευθυγράμμιση των δύο τρισδιάστατων επιφανειών, μέσα σε ένα λογικό και αποδεκτό με βάση τους ιατρούς χρονικό διάστημα, χωρίς την απαίτηση μεγάλων υπολογιστικών διαδικασιών.

Παρατηρώντας το γράφημα, σαν τιμές των παραμέτρων επιλέχθηκαν οι εξής:

Επαναλήψεις = 100
Ελάχιστη Απόσταση = 0.1

5.2 Ευθυγράμμιση εικόνων με γνωστή παραμόρφωση

Με σκοπό την αξιολόγηση της προτεινόμενης μεθόδου ευθυγράμμισης τρισδιάστατων ιατρικών δεδομένων, αρχικά χρησιμοποιούνται εικόνες οι οποίες έχουν μετασχηματιστεί με ένα γνωστό και δοσμένο από τον χρήστη πίνακα μετασχηματισμού. Ακολούθως επιχειρούμε την ευθυγράμμιση της παραγόμενης εικόνας με την αρχική εικόνα για την εύρεση του πίνακα μετασχηματισμού και ορίζουμε μια συνάρτηση ομοιότητας για σύγκριση των δύο μετασχηματισμών.

Αρχικά ο χρήστης καλείται να εισάγει τις περιστροφές και τις μετατοπίσεις ως προς τους τρεις άξονες (x,y,z) και με συγκεκριμένες πράξεις, όπως είδαμε και σε προηγούμενο κεφάλαιο:

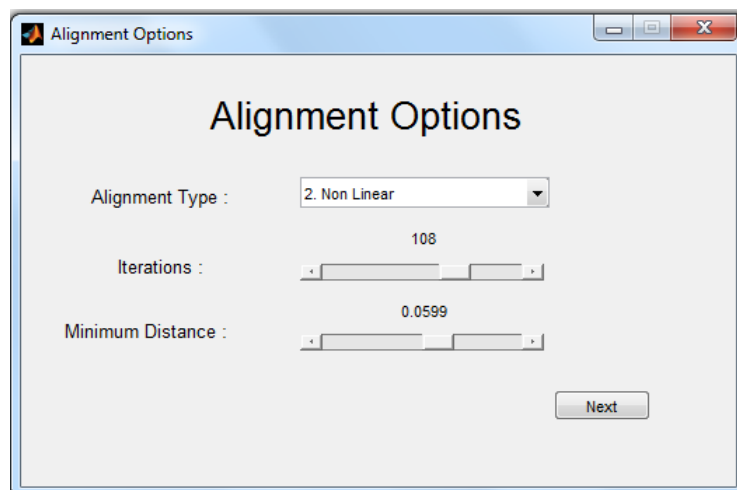
Transformation Matrix

$$= \begin{pmatrix} 1 & 0 & 0 & d_x \\ 0 & \cos \varphi_x & -\sin \varphi_x & 0 \\ 0 & \sin \varphi_x & \cos \varphi_x & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos \varphi_y & 0 & \sin \varphi_y & 0 \\ 0 & 1 & 0 & d_y \\ -\sin \varphi_y & 0 & \cos \varphi_y & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos \varphi_z & -\sin \varphi_z & 0 & 0 \\ \sin \varphi_z & \cos \varphi_z & 0 & 0 \\ 0 & 0 & 1 & d_z \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Έτσι, μπορούμε να αποκτήσουμε την μήτρα μετασχηματισμού, όπως φαίνεται πιο πάνω και στην Εικόνα 5.2.

Εικόνα 5.2 : Γραφικό Περιβάλλον για την εισαγωγή ενός δοκιμαστικού πίνακα μετασχηματισμού

Ακολούθως ο χρήστης καλείται να καθορίσει τις παραμέτρους της ευθυγράμμισης που επιθυμεί να εκτελέσει στην εικόνα, οι οποίες παράμετροι παρουσιάζονται στο γραφικό περιβάλλον όπως φαίνεται στην Εικόνα 5.3.



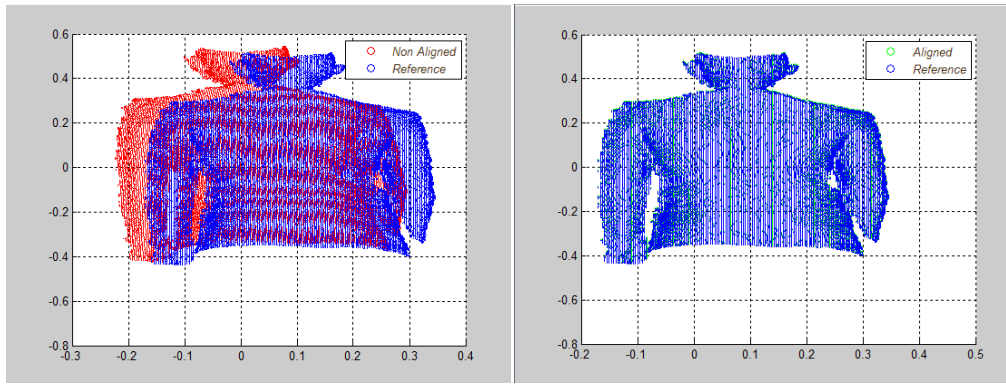
Εικόνα 5.3 : Γραφικό Περιβάλλον για την εισαγωγή των επιθυμητών παραμέτρων ευθυγράμμισης

Τέλος, αφού οι επιφάνειες ευθυγραμμιστούν, βρίσκουμε τον νέο πίνακα μετασχηματισμού ($Q'(i, j)$) από τον οποίο βρίσκονται οι υπολογισμένες μετατοπίσεις και περιστροφές περί των τριών αξόνων οι οποίες συγκρίνονται με τις αρχικές. Η σύγκριση των δοσμένων και των υπολογισμένων μετατροπών του τρισδιάστατου συνόλου γίνεται με μία συνάρτηση ομοιότητας. Σαν συνάρτηση ομοιότητας επιλέχθηκε η ρίζα της μέσης τετραγωνικής διαφοράς (root mean squared difference - RMSD) μεταξύ των διάφορων μεταβλητών (v_i^{given} των μεταβλητών που εισάγονται και $v_i^{calculated}$, των μεταβλητών που υπολογίζονται) μετατροπής του νέφος σημείων και δίνεται από την εξίσωση:

$$RMSD = \sqrt{\frac{1}{N} \sum_{i=1}^N \|v_i^{given} - v_i^{calculated}\|^2}$$

Όπου $N(=6)$ είναι ο αριθμός των διάφορων μεταβλητών μετατροπής του νέφος σημείων.

Ακολουθούν μερικές δοκιμές επαλήθευσης του αλγόριθμου με τυχαίες τιμές όπου παρουσιάζονται τόσο οι τιμές της συνάρτησης ομοιότητας, όσο και το αποτέλεσμα της ευθυγράμμισης.

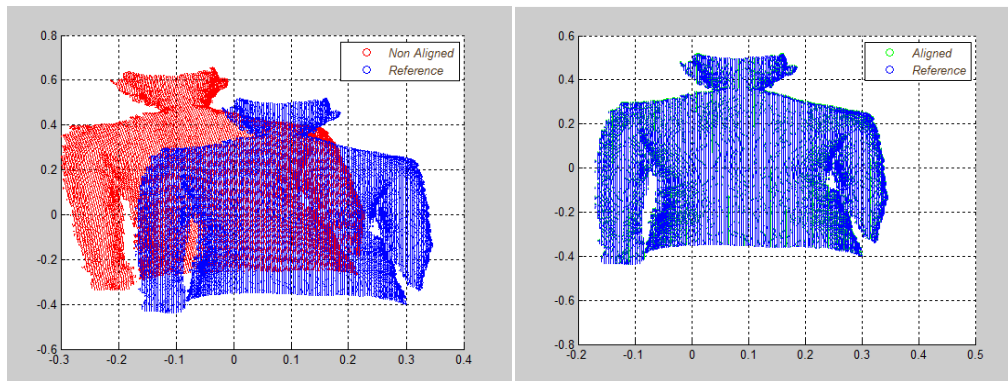


(α)

(β)

Εικόνα 5.4 : (α) Εικόνα αναφοράς (μπλε) και εικόνα μετασχηματισμένη (κόκκινο) κατά $(r_x, r_y, r_z, d_x, d_y, d_z) = (1,3,2,0,0,0)$

(β) Αποτέλεσμα ευθυγράμμισης (πράσινο) σε σχέση με την εικόνα αναφοράς (μπλε)

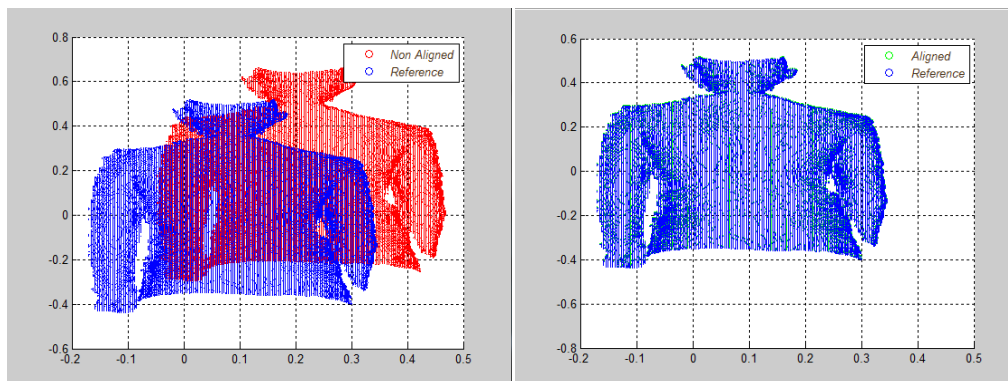


(α)

(β)

Εικόνα 5.5 : (α) Εικόνα αναφοράς (μπλε) και εικόνα μετασχηματισμένη (κόκκινο) κατά $(r_x, r_y, r_z, d_x, d_y, d_z) = (6,7,5,0,0,0)$

(β) Αποτέλεσμα ευθυγράμμισης (πράσινο) σε σχέση με την εικόνα αναφοράς (μπλε)

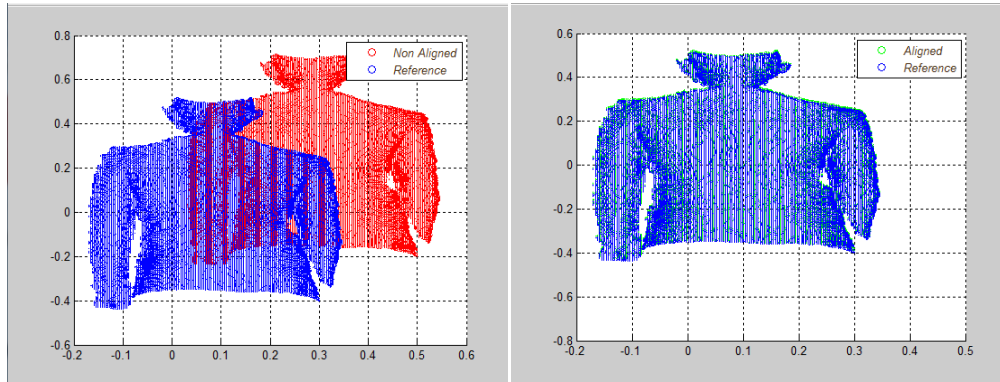


(α)

(β)

Εικόνα 5.6 : (α) Εικόνα αναφοράς (μπλε) και εικόνα μετασχηματισμένη (κόκκινο) κατά $(r_x, r_y, r_z, d_x, d_y, d_z) = (0, 0, 0, 0.123, 0.145, 0.115)$

(β) Αποτέλεσμα ευθυγράμμισης (πράσινο) σε σχέση με την εικόνα αναφοράς (μπλε)

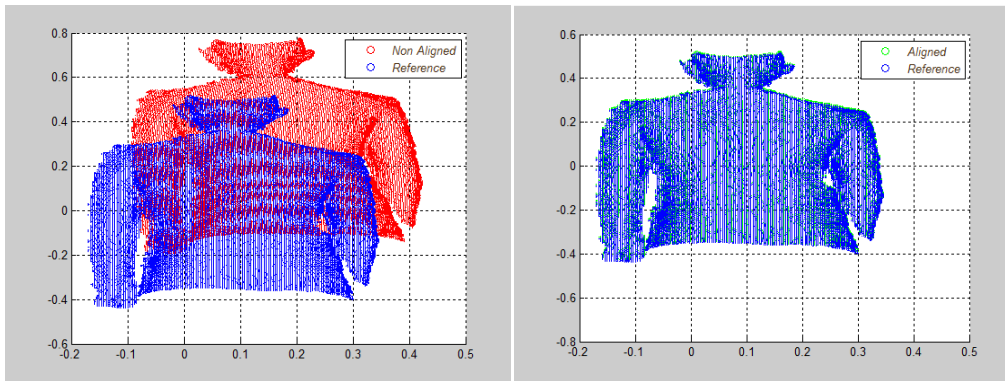


(α)

(β)

Εικόνα 5.7 : (α) Εικόνα αναφοράς (μπλε) και εικόνα μετασχηματισμένη (κόκκινο) κατά $(r_x, r_y, r_z, d_x, d_y, d_z) = (0, 0, 0, 0.2, 0.2, 0.2)$

(β) Αποτέλεσμα ευθυγράμμισης (πράσινο) σε σχέση με την εικόνα αναφοράς (μπλε)



(α)

(β)

Εικόνα 5.8 : (α) Εικόνα αναφοράς (μπλε) και εικόνα μετασχηματισμένη (κόκκινο) κατά $(r_x, r_y, r_z, d_x, d_y, d_z) = (1, 3, 2, 0.123, 0.234, 0.102)$

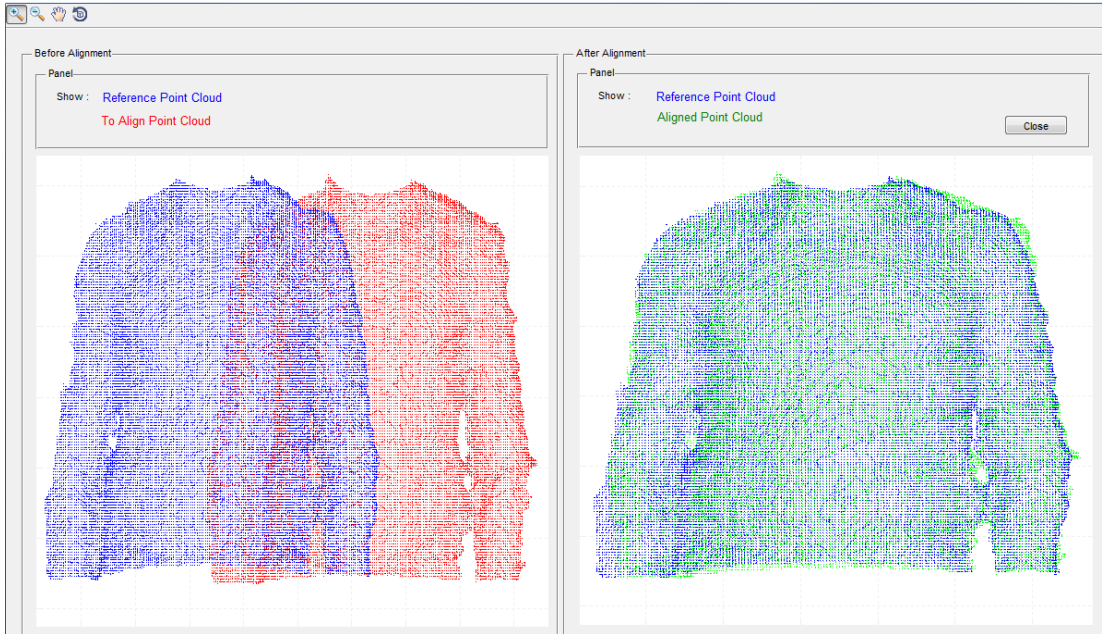
(β) Αποτέλεσμα ευθυγράμμισης (πράσινο) σε σχέση με την εικόνα αναφοράς (μπλε)

Παρατηρούμε ότι στην περίπτωση όπου ο πίνακας μετασχηματισμού είναι ο μοναδιαίος πίνακας, δηλαδή ότι σαν είσοδο του αλγόριθμου ευθυγράμμισης εισάγουμε δύο όμοιες εικόνες, παρουσιάζεται ένα σφάλμα. Το σφάλμα αν και πολύ μικρό, στην ιδανική περίπτωση θα έπρεπε να είναι μηδενικό. Η παρουσία αυτού του σφάλματος οφείλεται στο γεγονός ότι ο αλγόριθμος δεν αναγνωρίζει ότι οι εισόδους είναι ίδιες και κάνει μία επανάληψη του αλγόριθμου ευθυγράμμισης. Στην επανάληψη αυτή, λόγω πεπερασμένης ακρίβειας, βρίσκει μια μικρή παραμόρφωση των εισόδων, όπου και δημιουργείται το σφάλμα στην έξοδο.

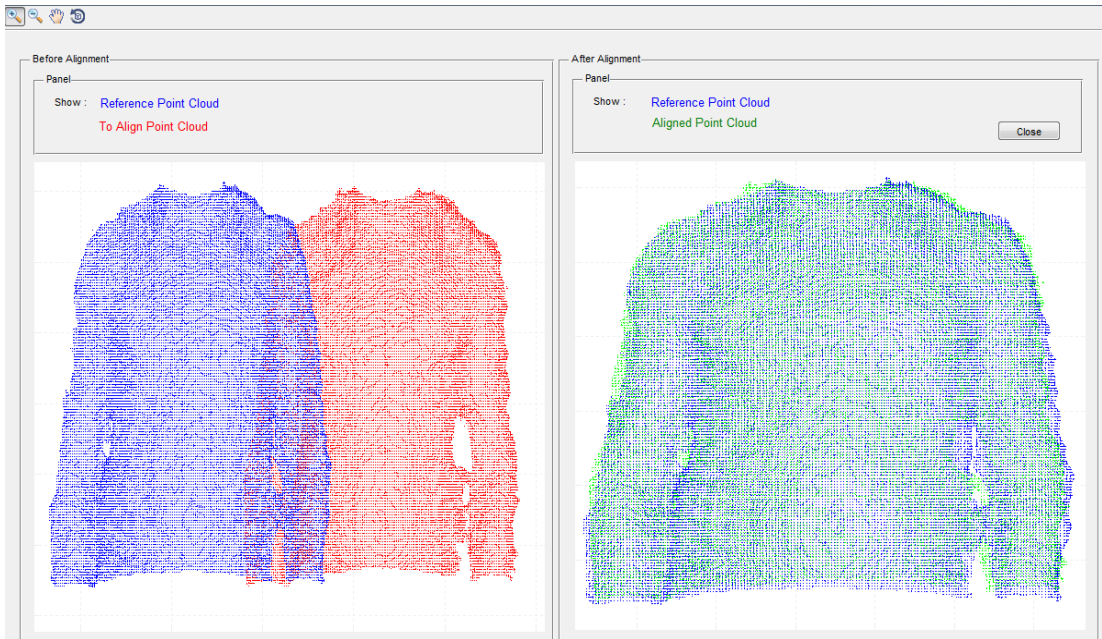
5.3 Ευθυγράμμιση πραγματικών εικόνων

Στις πραγματικές εικόνες δεν είναι δυνατό να αξιολογήσουμε την αντιστοίχιση των σημείων οπότε βασιζόμαστε στην τιμή του δείκτη ταύτισης. Τα τρισδιάστατα δεδομένα πάρθηκαν από λήψεις οι οποίες έγιναν στο εργαστήριο για διάφορες μετακινήσεις και περιστροφές του αντικειμένου ως προς τους τρεις άξονες για διάφορες τιμές παραμόρφωσης. Τα αποτελέσματα παρουσιάζονται στον Πίνακα 5.1 για γραμμική ευθυγράμμιση και ακολούθως στα παραγόμενα τρισδιάστατα ευθυγραμμισμένα δεδομένα εφαρμόζεται μη γραμμικός μετασχηματισμός για περαιτέρω ακρίβεια και περαιτέρω μείωση της μεταξύ τους απόσταση των δύο επιφανειών (Πίνακα 5.2). Ακολούθως παρουσιάζονται τα αποτελέσματα ευθυγράμμισης κάθε μεθόδου. Για την μη γραμμική ευθυγράμμιση, μειώνουμε ακόμη περισσότερο την ελάχιστη επιτρεπτή απόσταση από 0.1 σε 0.001.

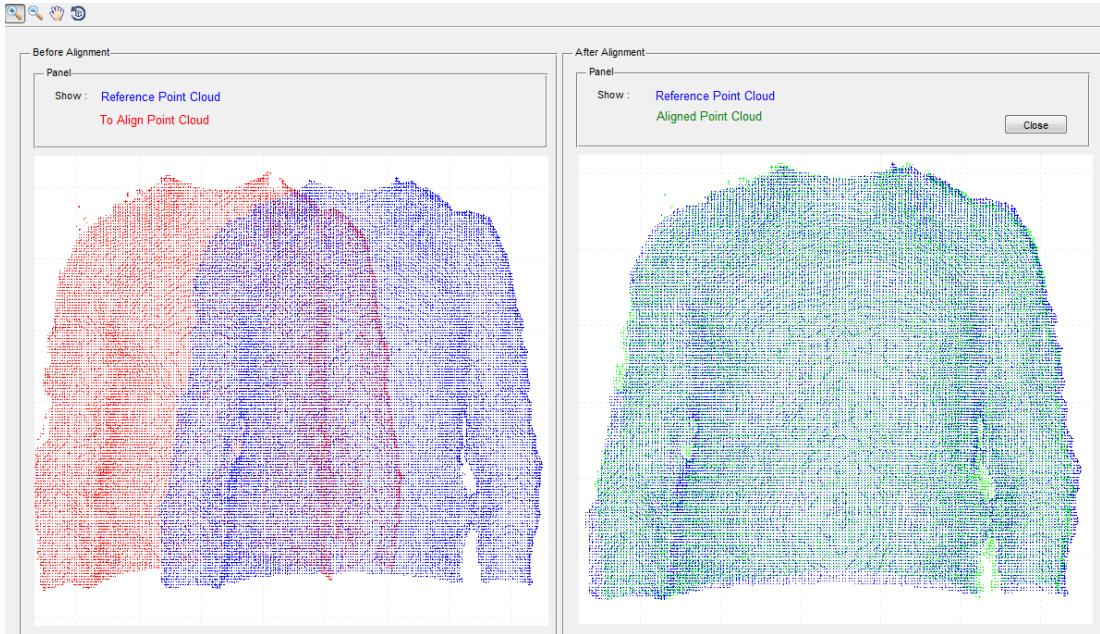
Ευθυγράμμιση με γραμμικό μετασχηματισμό



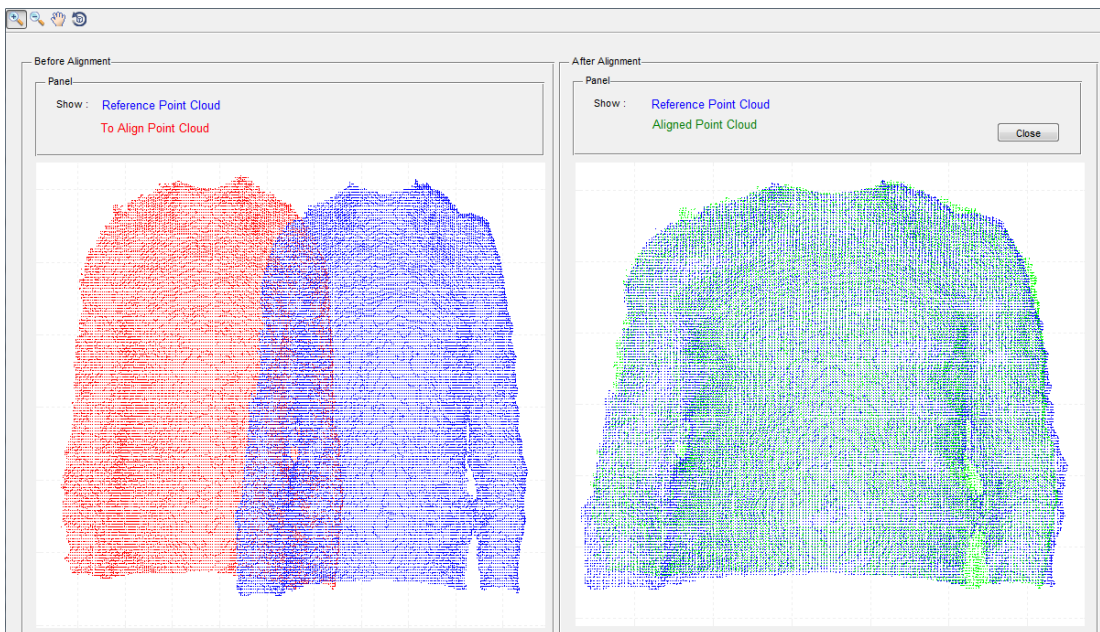
Εικόνα 5.9 : Ευθυγράμμιση παραμορφωμένης εικόνας ως προς τα δεξιά με μικρή τιμή παραμόρφωσης



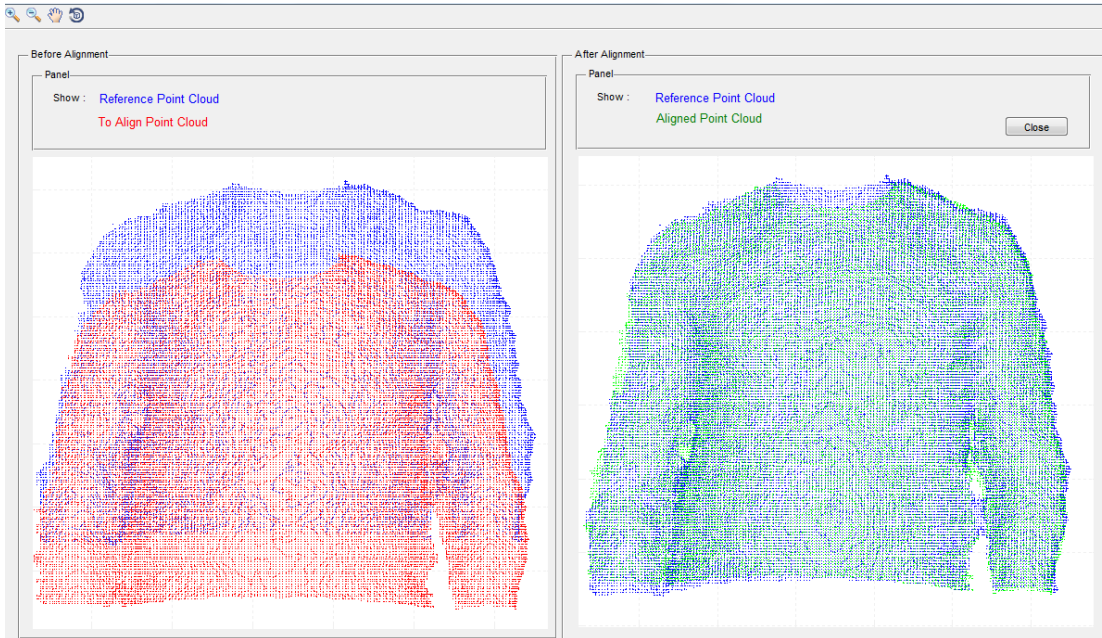
Εικόνα 5.10 : Ευθυγράμμιση παραμορφωμένης εικόνας ως προς τα δεξιά με μεγάλη τιμή παραμόρφωσης



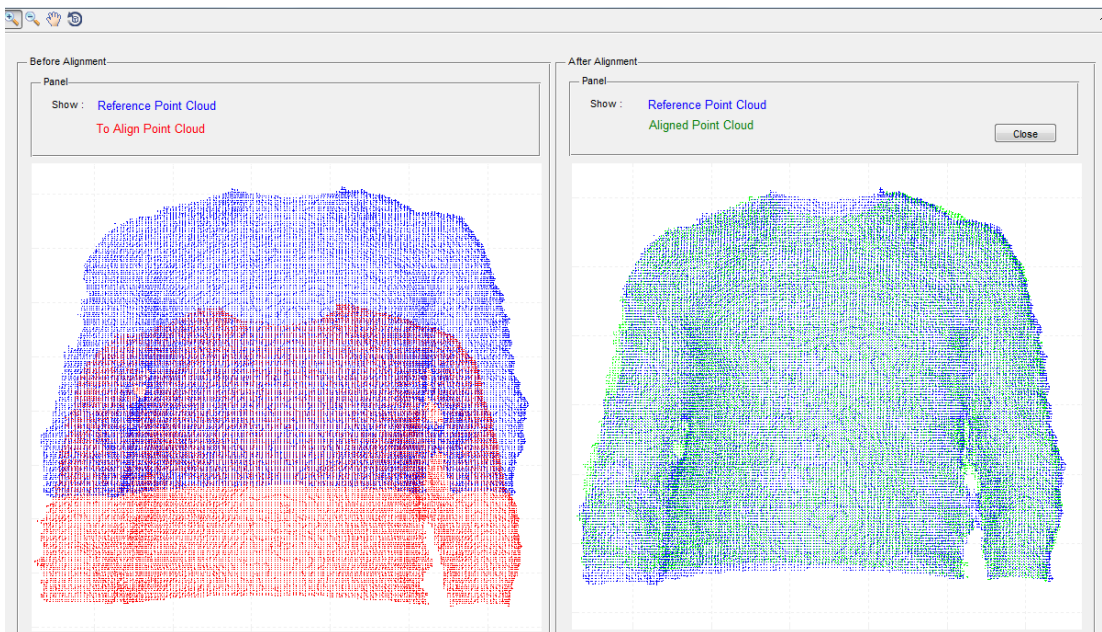
Εικόνα 5.11 : Ευθυγράμμιση παραμορφωμένης εικόνας ως προς τα αριστερά με μικρή τιμή παραμόρφωσης



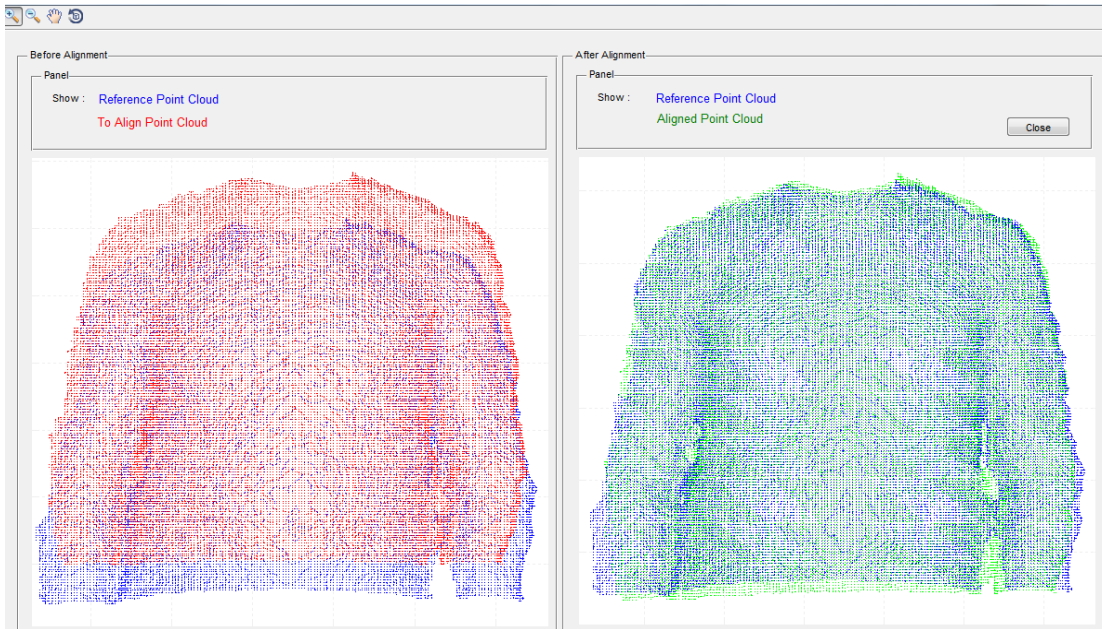
Εικόνα 5.12 : Ευθυγράμμιση παραμορφωμένης εικόνας ως προς τα αριστερά με μεγάλη τιμή παραμόρφωσης



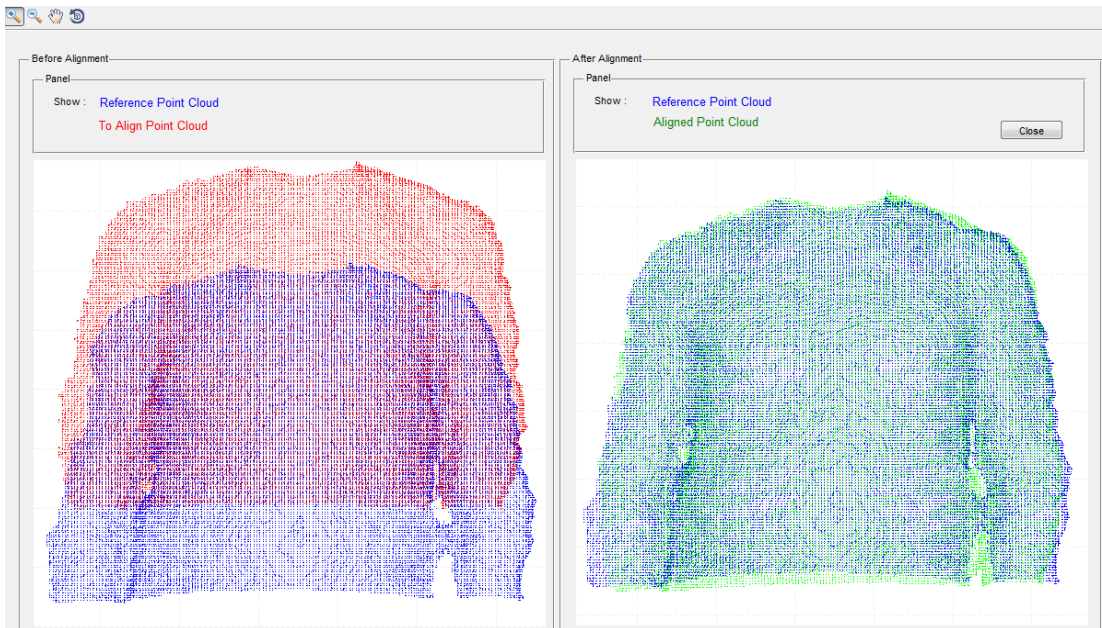
Εικόνα 5.13 : Ευθυγράμμιση παραμορφωμένης εικόνας ως προς τα κάτω με μικρή τιμή παραμόρφωσης



Εικόνα 5.14 : Ευθυγράμμιση παραμορφωμένης εικόνας ως προς τα κάτω με μεγάλη τιμή παραμόρφωσης

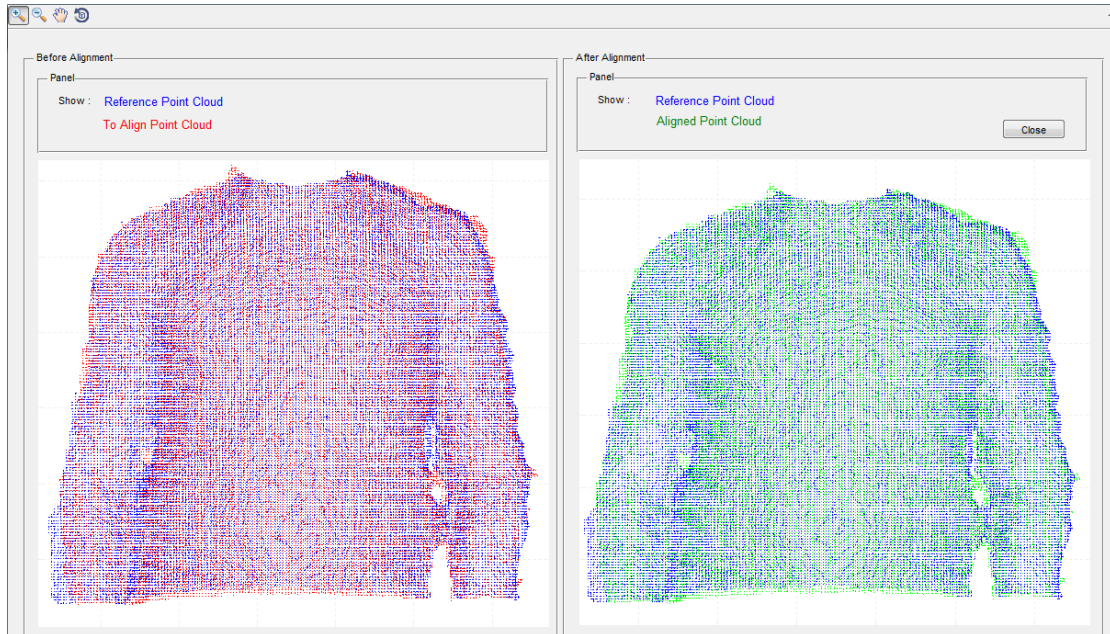


Εικόνα 5.15 : Ευθυγράμμιση παραμορφωμένης εικόνας ως προς τα πάνω με μικρή τιμή παραμόρφωσης

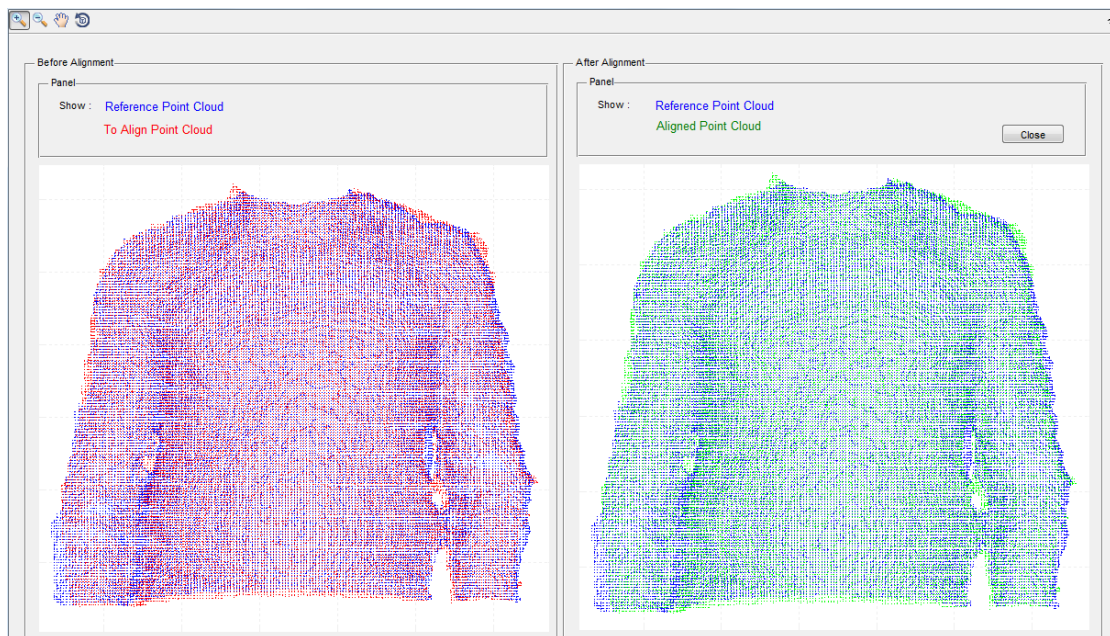


Εικόνα 5.16 : Ευθυγράμμιση παραμορφωμένης εικόνας ως προς τα πάνω με μεγάλη τιμή παραμόρφωσης

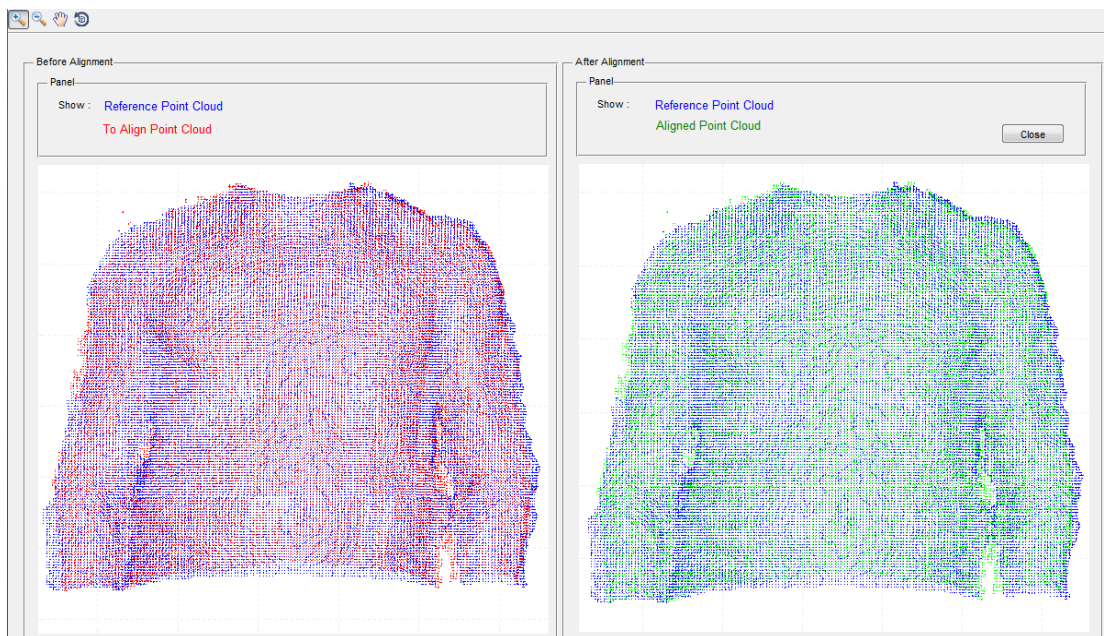
Ευθυγραμμίσεις με μη γραμμικό μετασχηματισμό



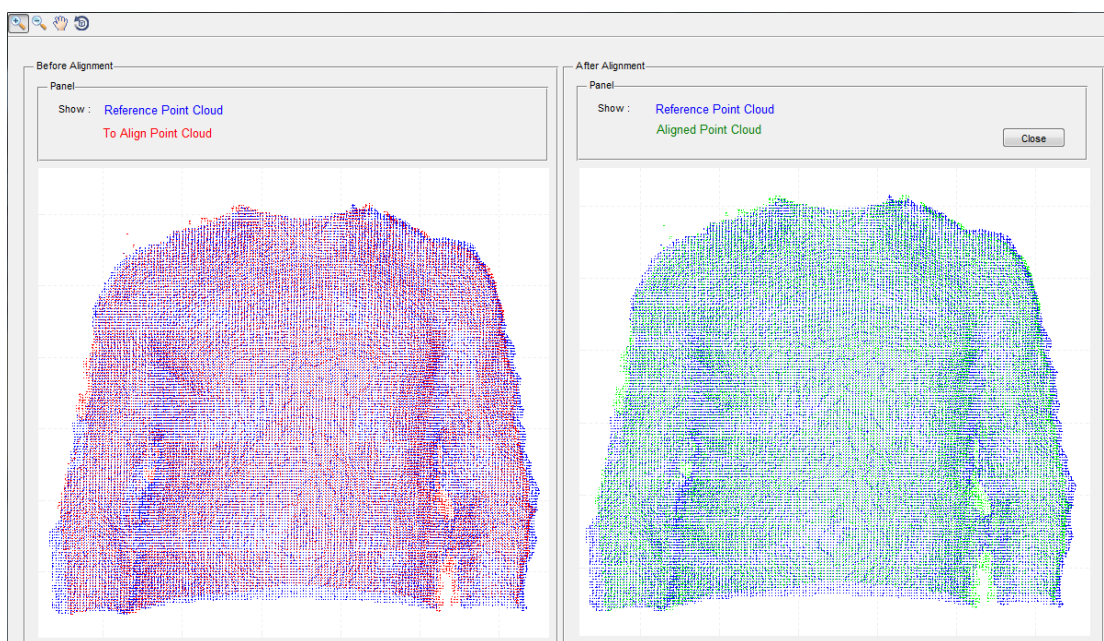
Εικόνα 5.17 : Ευθυγράμμιση παραμορφωμένης εικόνας ως προς τα δεξιά με μικρή τιμή παραμόρφωσης



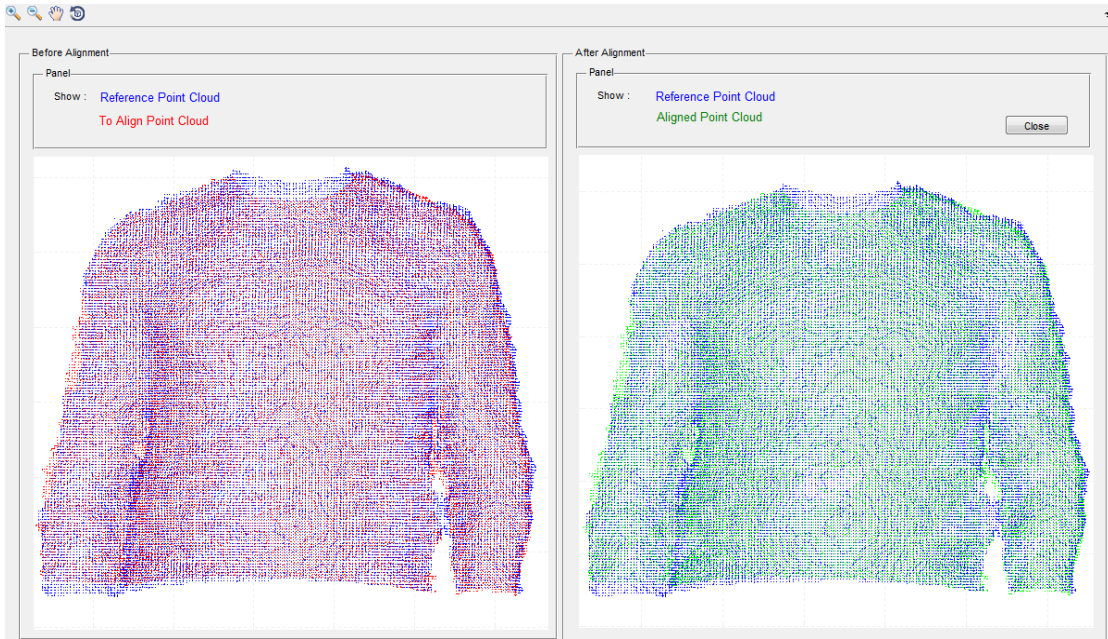
Εικόνα 5.18 : Ευθυγράμμιση παραμορφωμένης εικόνας ως προς τα δεξιά με μεγάλη τιμή παραμόρφωσης



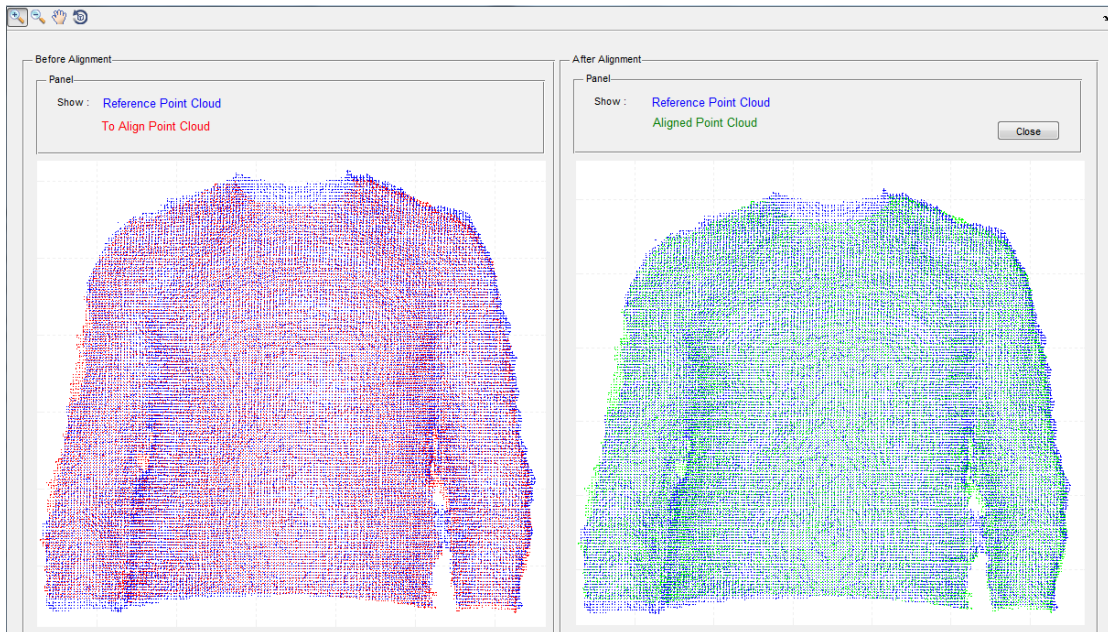
Εικόνα 5.19 : Ευθυγράμμιση παραμορφωμένης εικόνας ως προς τα αριστερά με μικρή τιμή παραμόρφωσης



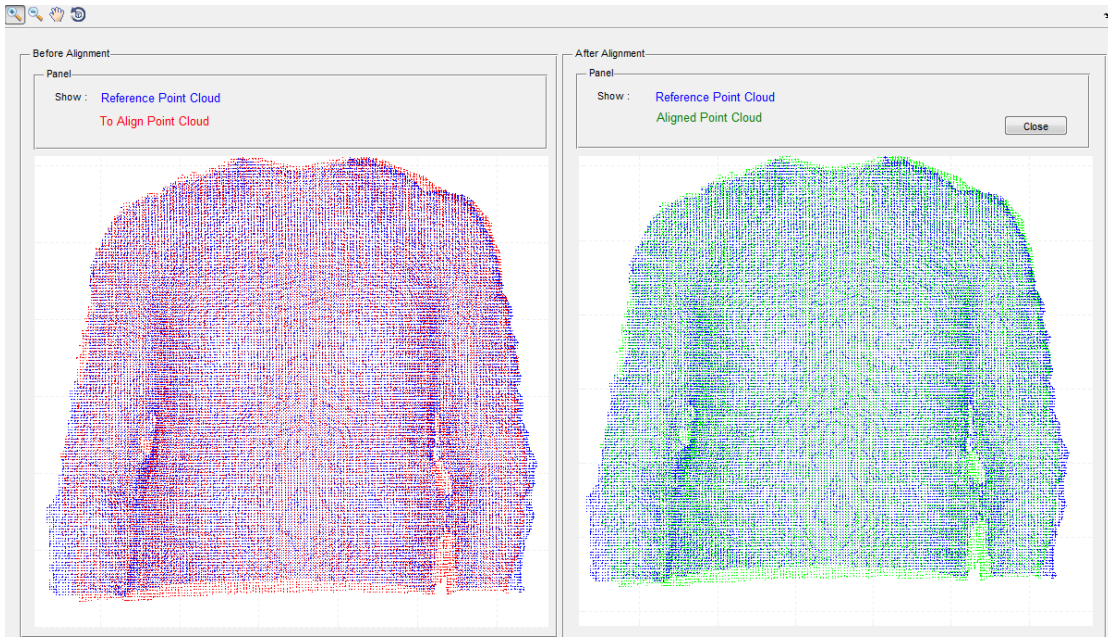
Εικόνα 5.20 : Ευθυγράμμιση παραμορφωμένης εικόνας ως προς τα αριστερά με μεγάλη τιμή παραμόρφωσης



Εικόνα 5.21 : Ευθυγράμμιση παραμορφωμένης εικόνας ως προς τα κάτω με μικρή τιμή παραμόρφωσης



Εικόνα 5.22 : Ευθυγράμμιση παραμορφωμένης εικόνας ως προς τα κάτω με μεγάλη τιμή παραμόρφωσης



Εικόνα 5.23 : Ευθυγράμμιση παραμορφωμένης εικόνας ως προς τα πάνω με μικρή τιμή παραμόρφωσης



Εικόνα 5.24 : Ευθυγράμμιση παραμορφωμένης εικόνας ως προς τα πάνω με μεγάλη τιμή παραμόρφωσης

6 Αξιολόγηση

6.1 Αξιολόγηση της προτεινόμενης μεθόδου ευθυγράμμισης

Το αποτέλεσμα της προτεινόμενης μεθόδου, εξαρτάται σημαντικά από τον τρόπο λήψης των τρισδιάστατων δεδομένων. Για το βέλτιστο επιθυμητό αποτέλεσμα θα πρέπει το περιβάλλον στο οποίο γίνεται η λήψη των δεδομένων να ικανοποιεί τους περιορισμούς οι οποίοι εισάγονται από τον αισθητήρα KINECT. Βάση προδιαγραφών, η βέλτιστη λήψη δεδομένων από τον αισθητήρα KINECT εξαρτάται από τον φωτισμό και από την απόσταση του αισθητήρα από το αντικείμενο.

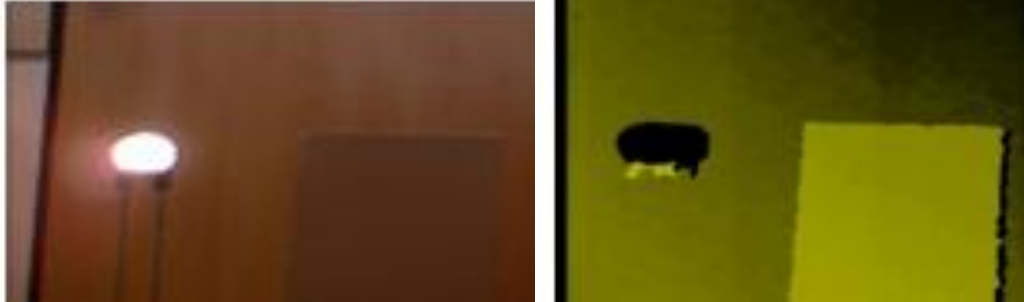
Επίσης, τόσο το αποτέλεσμα της ευθυγράμμισης επιφανειών όσο και ο χρόνος υλοποίησης της ευθυγράμμισης αυτών, εξαρτάται σημαντικά από το ποσοστό της μεταβολής μεταξύ των δύο επιφανειών. Στην ιατρική περίπτωση όμως, όταν τηρείται το κλινικό πρωτόκολλο, οι περιστροφές και οι μετατοπίσεις του ασθενή στο κλινικό κρεβάτι είναι σχετικά μικρές, έτσι δεν επηρεάζεται σημαντικά ο χρόνος και το αποτέλεσμα της ευθυγράμμισης. Στις περιπτώσεις μεγάλων μετατοπίσεων και περιστροφών, ιδανικό θα ήταν να προστεθεί ένα αρχικό βήμα το οποίο θα έκανε μια αρχική, γρήγορη ευθυγράμμιση περιορισμένης ακρίβειας.

Τέλος, τόσο ο αλγόριθμος ταύτισης σημείων, όσο και ο αλγόριθμος απόρριψης σημείων μπορούν να παραλληλοποιηθούν μοιράζοντας έτσι τον υπολογιστικό φόρτο σε περισσότερους από ένα επεξεργαστή. Η παράλληλη επεξεργασία δεδομένων είναι πλέον εφικτή στους σύγχρονους πολυπύρηνους υπολογιστές.

Ένα πολύ καλό υπολογιστικό εργαλείο με μεγάλο χρόνο εκτέλεσης χαρακτηρίζεται ως ένα ανούσιο εργαλείο πλέον στην σημερινή επιστήμη της ιατρικής. Έτσι σκοπός μας είναι η εύρεση της χρυσής τομής μεταξύ ταχύτητας και ακρίβειας των αποτελεσμάτων, επιλέγοντας ικανοποιητική ακρίβεια σε όρια αποδεκτά από τους ειδικούς αλλά και πρακτικούς χρόνους εκτέλεσης. Ο συγκεκριμένος αλγόριθμος μπορεί να ευθυγραμμίσει τρισδιάστατες επιφάνειες με χρόνο εκτέλεσης μερικών λεπτών, όπου το εύρος των τιμών του χρόνου εκτέλεσης εξαρτάται από την ακρίβεια του αποτελέσματος που ζητά ο χρήστης.

6.1.1 Φωτισμός

Ο αισθητήρας KINECT είναι ευαίσθητος στην υψηλή ένταση φωτεινής ροής. Αντικείμενα τα οποία έχουν υψηλή ένταση φωτεινής ροής δεν ανιχνεύονται για τον λόγο ότι τίθενται σαν ένα σημείο με αρκετά μεγάλη απόσταση, έτσι στην τελική απεικόνιση του αισθητήρα, τα αντικείμενα αυτά δεν αναπαρίστανται στην εικόνα όπως φαίνεται στην Εικόνα 6.1.



(α)

(β)

Εικόνα 6.1 : (α) Απεικόνιση Χώρου με την RGB κάμερα του αισθητήρα KINECT με την παρουσία αντικειμένου έντονης φωτεινής ροής (λάμπα)

(β) Αναπαράσταση εικόνας βάθους του χώρου με μη αναγνώριση του αντικειμένου όπου έχει έντονη φωτεινή ροή

Αντιλαμβανόμαστε ότι ο αισθητήρας KINECT θα πρέπει να λειτουργεί σε χώρο όπου οι εντάσεις φωτεινής ροής δεν ξεπερνούν ένα κατώφλι, το οποίο δεν μεταβάλλεται για τον λόγο ότι εξαρτάται από την αλληλεπίδραση των IR ακτινών του αισθητήρα με την φωτεινή ένταση. Βάση προδιαγραφών, ο αισθητήρας KINECT λειτουργεί καλά όταν το περιβάλλον περιέχει αντικείμενα τα οποία ακτινοβολούνται με εντάσεις φωτεινής ροής κάτω από 5000 lux. Για την κατανόηση του κατωφλίου αναφέρουμε ότι, η ένταση φωτεινής ροής ενός αρκετά καλά φωτισμένου γραφείου είναι περίπου 500~750 lux, και η ένταση φωτεινής ροής του ήλιου μπορεί να φτάσει μέχρι και τα 100000 lux (41).

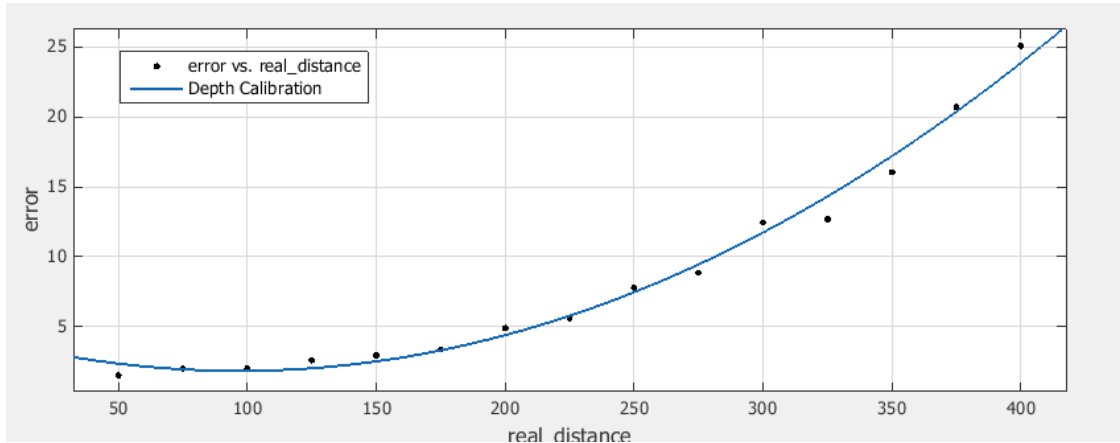
Στην περίπτωσή μας, σε ένα ομοιόμορφα φωτισμένο δωμάτιο, όπως είναι οι ιατρικές αίθουσες, ο αισθητήρας KINECT δεν παρουσιάζει πρόβλημα στις απεικονίσεις των αντικειμένων.

6.1.2 Απόσταση

Η απόσταση του αντικειμένου από τον αισθητήρα είναι ένας από τους βασικότερους παράγοντες εμφάνισης σφαλμάτων. Το σφάλμα στην απεικόνιση του αντικειμένου εμφανίζεται ως η διαφορά μεταξύ της πραγματικής απόστασης του αντικειμένου από τον αισθητήρα ως προς την μετρούμενη απόσταση από τον αισθητήρα.

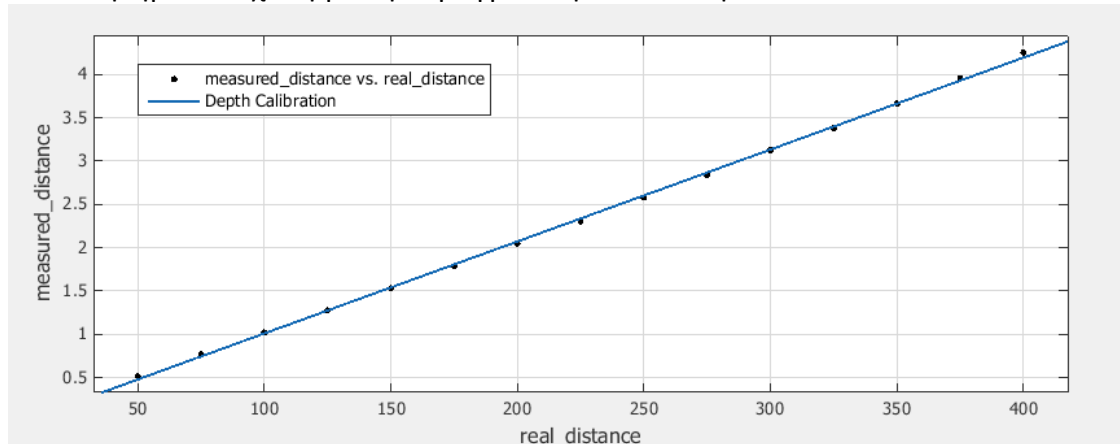
Πειραματικά η εύρεση του βέλτιστου εύρους αποστάσεων για την καλύτερη απόκτηση των δεδομένων γίνεται με την εξής πειραματική διαδικασία. Αρχικά, απεικονίζουμε μία επίπεδη επιφάνεια απέναντι από τον αισθητήρα σε γνωστή απόσταση και μετράμε την μετρούμενη απόσταση από την εικόνα την οποία λήφθηκε. Η διαφορά της απόστασης αυτής με την πραγματική απόσταση του αντικειμένου από τον αισθητήρα μας δίνει το σφάλμα το οποίο δίνει ο αισθητήρας για την γνωστή απόσταση. Ακολουθώντας αυξανόμενα σταδιακά την απόσταση του αντικειμένου και ακολουθώντας την ίδια διαδικασία μπορούμε να λάβουμε τα σφάλματα που δίνει ο αισθητήρας στις γνωστές αποστάσεις. Αναπαριστώντας αυτά τα σφάλματα συναρτήσει της απόστασης, μπορούμε να βρούμε το βέλτιστο εύρος που μας ενδιαφέρει, ανιχνεύοντας το τμήμα της καμπύλης το οποίο έχει

το ελάχιστο «ύψος» (σφάλμα). Όπως φαίνεται και από την Εικόνα 6.2, το βέλτιστο εύρος αποστάσεων μεταξύ αντικειμένου και αισθητήρα είναι μεταξύ των 1 και 2 μέτρων.



Εικόνα 6.2 : Γραφική αναπαράσταση της σχέσης μεταξύ του σφάλματος βάθους που μετράει ο αισθητήρας σε σχέση με την πραγματική απόσταση του αντικειμένου από αυτό.

Και η Εικόνα 6.3 παρουσιάζει την γραμμική σχέση μεταξύ της μετρούμενης απόστασης από τον αισθητήρα σε σχέση με την πραγματική απόσταση.



Εικόνα 6.3 : Γραφική αναπαράσταση της γραμμικής σχέσης μεταξύ της μέτρησης βάθους από τον αισθητήρα σε σχέση με το πραγματικό βάθος

6.2 Μελλοντικές βελτιώσεις και επεκτάσεις

Όπως ήδη προαναφέραμε, η ταχύτητα και η ακρίβεια του προτεινόμενου αλγορίθμου μπορούν να βελτιωθούν με την εισαγωγή μίας αρχικής ευθυγράμμισης. Λόγο της μικρής μεταβολής που μπορούν να παρουσιάζουν τα τρισδιάστατα ιατρικά δεδομένα, μπορεί να εισαχθεί μία αρχική ευθυγράμμιση μόνο ως προς την μετατόπιση, αγνοώντας τυχόν περιστροφές. Μια τέτοια αρχική ευθυγράμμιση μπορεί να υλοποιηθεί με την ευθυγράμμιση των κέντρων μάζας των εικόνων. Θέλοντας να λαμβάνουμε υπόψη και τυχόν περιστροφές από την εικόνα αναφοράς στην αρχική ευθυγράμμιση, μπορούμε να υλοποιήσουμε μια αρχική ευθυγράμμιση η οποία να δίνει εκτιμήσεις τόσο για μετατόπιση, όσο και για τυχόν περιστροφές.

Μια επιπλέον προσθήκη για την περαιτέρω βελτίωση τόσο της ακρίβειας αλλά κυρίως της ταχύτητας υλοποίησης του αλγορίθμου θα ήταν η προσθήκη αλγορίθμου για την επιλογή σημείων σε αντίθεση με την παρούσα υλοποίηση η οποία λαμβάνει υπόψη όλα τα σημεία. Με την επιλογή χαρακτηριστικών σημείων όπως είναι οι κορυφές ή σημεία τα οποία ορίζουν μια ομοιόμορφη επίπεδη περιοχή, μπορούμε επίσης να ευθυγραμμίσουμε δύο επιφάνειες με γρηγορότερη εκτέλεση των βημάτων του αλγορίθμου λόγω λιγότερης, αλλά ικανοποιητικής πληροφορίας για την συσχέτιση των επιφανειών.

Τέλος, αξίζει η σύνδεση του αλγορίθμου αυτού με το ιατρικό κρεβάτι των ακτινοθεραπευτικών μηχανημάτων. Η σύνδεση αυτή είναι εφικτή με την τροποποίηση της εξόδου του αλγορίθμου. Έτσι, αντί η έξοδος να είναι ο πίνακας μετασχηματισμού για την ευθυγράμμιση, μπορεί παράλληλα να υλοποιηθεί ειδική διεπαφή για τον έλεγχο των μηχανικών μερών του ιατρικού κρεβατιού για την αυτόματη ευθυγράμμιση του ασθενή ως προς τον τρισδιάστατο χώρο. Σκοπός της σύνδεσης αυτής είναι η ευθυγράμμιση του παρόντος προσανατολισμού του ασθενή σε σχέση με ένα προηγούμενο προσανατολισμό ο οποίος θεωρείται ως σημείο αναφοράς.

1. **Won-Jin Yi, Min-Suk Heo, Sam-Sun Lee, Soon-Chul Choi, Sun-Bok Lee, Kyung-Hoe Huh.** Automatic noise robust registration of radiographs for subtraction using strategic local correlation: an application to radiographs of dental implants. *Computers in Biology and Medicine*. Μάρτιος 2005, σσ. 247-258.
2. **Diaa Eldin M. Nassar, Hany H. Ammar.** A neural network system for matching dental radiographs. *Pattern Recognition*. Ιανουάριος 2007, σσ. 65-79.
3. **J.V. Hajnal, N. Saeed, A. Oatridge, E.J. Willimas, I.R. Young, and G.M. Bydder.** Detection of subtle brain changes using subvoxel registration and subtraction of serial MR images. 1995, σσ. 677-691.
4. **Cuisenaire O, Thiran J, Macq B, Michel C, de Volder A, Marques F.** Automatic registration of 3D MR images with a computerized brain atlas. *SPIE: Medical imaging: Image Processing*. 1996, σσ. 438-448.
5. **Jain AK, Zhong Y, Lakshmanan, S.** Object matching using deformable templates. *IEEE Trans Pattern Anal Machine Intel*. 1996, σσ. 267-277.
6. **Amit Y, Kong A.** Graphical templates for model registration. *IEEE Trans Pattern Anal Machine Intel*. 1996, σσ. 225-236.
7. **Gilhuijs K, van de Ven P, van Herk M.** Automatic three-dimensional inspection of patient setup in radiation therapy using portal images, simulator images, and computed tomography data. *Medical Physics*. 1996, σσ. 389-399.
8. **Troccaz J, Laieb N, Vassal P, Menguy Y, Cinquin P, Bolla M, Giraud JY.** Patient setup optimization for external conformal radiotherapy. *Journal of image guided surgery*. 1995, σσ. 113-120.
9. **Hsu C-T, Beuker RA.** Multiresolution feature-based image registration. *Procs SPIE: Visual Communications and Image Processing*. 2000, σσ. 1490-1498.
10. **Maitz JBA, Viergever MA.** A survey of medical image registration. *Med Image Anal*. 1998, σσ. 1-36.
11. **Press W, Flannery B, Teukolsky S, Vetterling W.** Numerical recipes in C. *Cambridge University Press*. 1992.
12. **Kirkpatrick S, Gelatt CD, Vecchi MP.** Optimization by simulated annealing. *Science*. 1983, σσ. 671-680.
13. **Jacoby SLS, Kowalik JS, Pizzo JT.** Iterative methods for nonlinear optimization problems. *Glewood Cliffs. NJ: Prentice Hall*. 1972.
14. **Maes F, Vandermeulen D, Suetens P.** Comparative evaluation of multiresolution optimization strategies for multimodality image registration by maximization of mutual information. *Med Image Anal*. 1999, σσ. 373-386.
15. **Lehmann TM, Gröndahl K, Gröndahl H-G, Schmitt W, Spitzer K.** Observer-independent registration of perspective projection prior to subtraction of in vivo radiographs. *Dentomaxillofacial Radiology*. 1998, σσ. 140-150.

16. **Chianos JY, Sullivan BJ.** Coincident bit counting - A new criterion for image registration. *IEEE Trans Med Imaging.* 1993, σσ. 30-38.
17. **Radcliffe T, Rajapekshe R, Shaler S.** Pseudocorrelation: A fast, robust, absolute, gray level image alignment algorithms. *Med Phys.* 1994, σσ. 761-769.
18. **Yu JJ-H, Hung B-N, Liou C-L.** Fast algorithm for digital retinal image alignment. *Procs IEEE Ann. Int. Conf. Engineering Medicine Biology Society, Images Twenty-First Century.* November, 1989;, σσ. 374-375.
19. **Penney GP, Weese J, Little JA, Desmedt P, Hill DLG, Hawkes DJ.** A comparison of similarity measures for use in 2D-3D medical image registration. *IEEE Trans Medical Imaging.* 1998;, σσ. 586-595.
20. **DC., Yoon.** A new method for the automated alignment of dental radiographs for digital subtraction radiography. *Dentomaxillofacial Radiology.* 2000;, σσ. 11-19.
21. **Nikou C, Heitz F, Armspach J-P.** Robust registration of dissimilar single and multimodal images. *Computer Vision – ECCV '98.* s.l. : Berlin Heidelberg: Springer-Verlag, 1998;, Τόμ. II, σσ. 51-56.
22. **Rueckert D, Hayes C, Studholme C, Summers P, Leach M, Hawkes DJ.** Non-rigid registration of breast MR images using mutual information. *Wells WM.* s.l. : Berlin Heidelberg: Springer, 1998, σσ. 1144-1152.
23. **West J, Fitzpatrick JM, Wang MY, Dawant BM, Maurer CR, Jr., Kessler RM, Maciunas RJ.** Retrospective intermodality registration techniques for images of the head: Surface-based versus volume-based. *IEEE Trans Med Imaging.* 1999;, σσ. 144-150.
24. **Lemieux L, Jagoe R, Fish DR, Kitchen ND, Thomas DGT.** A patient-to-computed-tomography image registration method based on digitally reconstructed radiographs. *Med Phys.* 1994;, σσ. 1749-1760.
25. **Dong L, Boyer AL.** A portal image alignment and patient setup verification procedure using moments and correlation techniques. *Physics in medicine and biology.* 1996;, σσ. 697-723.
26. **LMG., Brown.** Registration of planar film radiographs with computed tomography. *Procs of MMBIA.* 1996, σσ. 42-51.
27. **Collignon A, Maes F, Delaere D, Vandermeulen D, Suetens P, Marshal G.** Automated multi-modality image registration based on information theory. *Information Processing in Medical Imaging.* 1995, σσ. 263-274.
28. **Maes F, Collignon A, Vandermeulen D, Marchal G, Suetens P.** Multimodality image registration by maximization of mutual information. *IEEE Trans Med Imaging.* 1997, σσ. 187-198.
29. **Viola P, Wells III WM.** Alignment by maximization of mutual information. *J Comp Vis.* 1997, σσ. 137-154.
30. **Pluim JPW, Maintz IBA, Viergever MA.** Interpolation artefacts in mutual information-based image registration. *Computer vision and image understanding.* 2000, σσ. 211-232.
31. **Wells III WM, Viola P, Atsumi H, Nakajima S, Kikinis R.** Multi-modal volume registration by maximization of mutual information. *Med Image Anal.* 1996, σσ. 35-51.
32. **Thevenaz P, Unser M.** Optimization of mutual information for multiresolution image registration. *IEEE Trans Image Process.* 2000, σσ. 2083-2099.

33. **Meyer CR, Boes JL, Kim B, Bland PH, Zasadny KR, Kison PV, Koral K, Frey KA, Wahl RL.** Demonstration of accuracy and clinical versatility of mutual information for automatic multimodality image fusion using affine and thin-plate spline warped geometric deformations. *Med Image Anal.* 1997, σσ. 195-206.
34. **Mandava VR, Fitzpatrick JM, Maurer CR, Jr, Maciunas RJ, Allen GS.** Registration of multimodal volume head images via attached markers. *Procs SPIE: Medical Imaging VI: Image Processing.* 1992, σσ. 271-282.
35. **West J, Fitzpatrick JM, Wang MY, Dawant BM.** Comparison and evaluation of retrospective intermodality brain image registration techniques. *Comput Assisted Tomography.* 1997, σσ. 554-566.
36. **Maintz JBA, Van den Elsen PA, Viergever MA.** Comparison of feature-based matching of CT and MR brain images. *Procs Computer Vision, Virtual Reality and Robotics in Medicine.* April 1995, σσ. 219-228.
37. **Magnusson, Martin.** *3D Scan Matching for Mobile Robots with Application to Mine Mapping.* Örebro : Department of Technology at Örebro University, 2006.
38. *Registration and integration of.* **T. Masuda, K. Sakaue, and N. Yokoya.** 1996. Proceedings of 13th International Conference on Pattern Recognition. σσ. 879-883.
39. **Pankaj K. Agarwal, Boris Aronov, Sarel Har-Peled, Jeff M. Phillips, Ke Yi, Wuzhou Zhang.** Nearest Neighbor Searching Under Uncertainty II.
40. **Szymon Rusinkiewicz, Marc Levoy.** *Efficient Variants of the ICP Algorithm.* s.l. : Stanford University, 2011.
41. **György Csaba, László Somlyai, Zoltán Vámosy.** *Differences Between Kinect and Structured Lighting Sensor in Robot Navigation.* Hungary : Óbuda University, John von Neumann Faculty of Informatics, 2012.
42. **Derek L G Hill, Philipp G Batchelor, Mark Holden and David J Hawkes.** *Medical Image registration.* 2000.
43. **Zitova B, FlusserJ.** Image registration methods: a survey. *Image and Vision Computing.* 2003, σσ. 977-1000.
44. *A review of medical image registration.* **Maurer CR, Fitzpatrick JM.** s.l. : Interactive Image - Guided Neurosurgery, 1993. American Association of Neurological Surgeons. σσ. 14-44.
45. **Van den Elsen PA, Pol EJD, Viergever MA.** Medical image matching-a review with classification. *IEEE Engng Med Biol.* 1993, σσ. 26-39.
46. **Thevenaz P, Ruttimann UE, Unser M.** A pyramid approach to subpixel registration based on intensity. *IEEE Trans Image Process.* 1998, σσ. 27-41.
47. **Varun Ganapathi, Christian Plagemann, Daphne Koller, Sebastian Thrun.** *Real Time Motion Capture Using a Single Time-Of-Flight Camera.* Stanford, CA, USA : Stanford University, Computer Science Department, 2011.
48. **Reinhard Klette, Karsten Schluns, Andreas Koschan.** Structred Lighting. *Computer Vision, Three-Dimensional Data from Images.* s.l. : Springer, 9, σσ. 347-375.
49. **Jens Kremer, Graham Hunter.** *Performance of the StreetMapper Mobile LiDAR Mapping System.* s.l. : Real World, 2006.
50. **L. Zhu, A. Jaakkola, J. Hyypä.** THE USE OF MOBILE LASER SCANNING DATA AND UNMANNED AERIAL. *Remote Sensing and Spatial Information Sciences.* September 2013.

51. **Jamie Carter, Keil Schmid, Kirk Waters, Lindy Betzhold, Brian Hadley, Rebecca Mataosky, and Jennifer Halleran.** Data Produced by LiDAR Sensors. *An Introduction to Lidar Technology, Data, and Applications*. 2012, σσ. 13-25.
52. **Jana, Abhijit.** *Kinect for Windows SDK*. BIRMINGHAM - MUMBAI : PACKT Publishing, 2012.
53. **Jamie Shotton, Andrew Fitzgibbon, Mat Cook, Toby Sharp, Mark Finocchio, Richard Moore, Alex Kipman, Andrew Blake.** Real-Time Human Pose Recognition in Parts from Single Depth Images. *Microsoft Research Cambridge & Xbox Incubation*. 2011.
54. **Khoshelham, K.** Accuracy Analysis of KINECT depth data. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. 29 August 2011.
55. **DRAELOS, MARK THEODORE.** *The Kinect Up Close: Modifications for Short-Range Depth*. 2012.
56. **HALAWANA, Hachem.** *Partial demosaicing of CFA images for stereo matching*. Automatic control, Computer Engineering and Image and Signal, University Lille. Lille : s.n., 2011. PHD THESIS.
57. **Shoemake, Ken.** Quaternions. Philadelphia : Department of Computer and Information Science, University of Pennsylvania.

A/A	Παράμετροι Γνωστού Μετασχηματισμού						Εκτιμώμενοι Παράμετροι Μετασχηματισμού						RMSD
	r_x	r_y	r_z	d_x	d_y	d_z	r_x	r_y	r_z	d_x	d_y	d_z	
1	0	0	0	0	0	0	0.0004	-1.3×10^{-5}	1.73×10^{-5}	-2×10^{-7}	-7×10^{-6}	2×10^{-6}	7.2×10^{-5}
2	1	3	2	0	0	0	1.1949	2.9505	2.0569	-1×10^{-4}	-5.5×10^{-3}	2×10^{-3}	0.03484
3	6	7	5	0	0	0	6.763	6.4131	5.7465	-2×10^{-4}	-5.5×10^{-3}	2.2×10^{-3}	0.20303
4	0	0	0	0.123	0.145	0.115	0.10981	-0,022886	-0.0019344	0.12263	0.13991	0.11588	0.01871
5	0	0	0	0.2	0.2	0.2	0.1656	-0.10962	0.003374	0.19868	0.19156	0.201	0.03313
6	1	3	2	0.123	0.234	0.102	1.2631	2.8635	2.0585	0.12385	0.22224	0.10588	0.05039

Πίνακας 5.1 : Πειραματικές Επαναλήψεις αλγορίθμου ευθυγράμμισης για τον έλεγχο της εγκυρότητας του αποτελέσματος αυτού

Α/Α	Εικόνα Προς Ευθυγράμμιση	Τιμή Συνάρτησης Ταύτισης (Linear ICP)				Ποσοστό Βελτίωσης Ακρίβειας
		Πριν	Μετά	Επαναλήψεις	Χρόνος Εκτέλεσης (sec)	
1	Δεξιά Ανθρώπου	0.0216152	4.46404×10^{-5}	115	512.505	1 : 484.207
2	Αρκετά Δεξιά Ανθρώπου	0.0536637	4.46404×10^{-5}	136	670.75	1 : 1202.133
3	Αριστερά Ανθρώπου	0.011316	3.56399×10^{-5}	77	638.14	1 : 292.894
4	Αρκετά Αριστερά Ανθρώπου	0.0362262	4.00024×10^{-5}	175	721.19	1 : 905.601
5	Κάτω Ανθρώπου	0.0055438	6.30064×10^{-5}	81	320.21	1 : 87.988
6	Αρκετά Κάτω Ανθρώπου	0.0103823	6.3112×10^{-5}	110	371.24	1 : 164.506
7	Πάνω Ανθρώπου	0.0135524	5.16343×10^{-5}	76	273.68	1 : 262.469
8	Αρκετά Πάνω Ανθρώπου	0.0201391	5.16358×10^{-5}	149	701.25	1 : 390.022

Πίνακας 5.2 : Πειραματικές Επαναλήψεις συμπαγής αλγορίθμου ευθυγράμμισης με αριθμό επαναλήψεων, χρόνο εκτέλεσης και ποσοστό βελτίωσης της ακρίβειας για ελάχιστη επιτρεπτή απόσταση 0.1

Α/Α	Εικόνα Προς Ευθυγράμμιση	Τιμή Συνάρτησης Ταύτισης (Non Linear ICP)				Ποσοστό Βελτίωσης Ακρίβειας
		Πριν	Μετά	Επαναλήψεις	Χρόνος Εκτέλεσης (sec)	
1	Δεξιά Ανθρώπου	4.46404×10^{-5}	4.45958×10^{-5}	1	251.44	1 : 1.001
2	Αρκετά Δεξιά Ανθρώπου	4.46404×10^{-5}	4.46415×10^{-5}	1	238.70	1 : 0.999
3	Αριστερά Ανθρώπου	3.56399×10^{-5}	3.45378×10^{-5}	2	270.16	1 : 1.031
4	Αρκετά Αριστερά Ανθρώπου	4.00024×10^{-5}	3.48397×10^{-5}	2	215.15	1 : 1.148
5	Κάτω Ανθρώπου	6.30064×10^{-5}	6.29191×10^{-5}	1	367.09	1 : 1.001
6	Αρκετά Κάτω Ανθρώπου	6.34458×10^{-5}	6.30675×10^{-5}	2	307.29	1 : 1.006
7	Πάνω Ανθρώπου	5.16343×10^{-5}	5.16356×10^{-5}	1	453.15	1 : 0.999
8	Αρκετά Πάνω Ανθρώπου	5.16358×10^{-5}	5.16372×10^{-5}	2	258.05	1 : 0.999

Πίνακας 5.3 : Πειραματικές Επαναλήψεις μη συμπαγή αλγορίθμου ευθυγράμμισης με αριθμό επαναλήψεων, χρόνο εκτέλεσης και ποσοστό βελτίωσης της ακρίβειας για ελάχιστη επιτρεπτή απόσταση 0.1 των αποτελεσμάτων ευθυγράμμισης του συμπαγή μετασχηματισμού

