



ΠΑΝΕΠΙΣΤΗΜΙΟ  
ΠΑΤΡΩΝ  
UNIVERSITY OF PATRAS

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ  
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ Η/Υ ΚΑΙ ΠΛΗΡΟΦΟΡΙΚΗΣ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

---

ΣΥΝΕΝΩΣΗ ΠΟΛΛΑΠΛΩΝ  
ΤΡΙΣΔΙΑΣΤΑΤΩΝ  
ΠΑΡΑΤΗΡΗΣΕΩΝ ΣΚΗΝΗΣ

---

ΣΥΝΤΑΚΤΗΣ

ΔΙΟΝΤΣΣΟΣ ΚΟΥΝΑΔΗΣ ΜΠΑΣΤΙΑΝ

ΕΠΙΒΛΕΠΩΝ: ΕΜΜΑΝΟΥΗΛ Ζ. ΨΑΡΑΚΗΣ, ΕΠΙΚΟΥΡΟΣ ΚΑΘΗΓΗΤΗΣ

ΠΑΤΡΑ ΝΟΕΜΒΡΙΟΣ 2013



## ΕΥΧΑΡΙΣΤΙΕΣ

Αρχικά θα ήθελα να ευχαριστήσω τον καθηγητή μου κ. Εμμανουήλ Ψαράκη για τις ιδέες και τον τρόπο σκέψης που μου μετάδωσε. Η περάτωση αυτής της διπλωματικής οφείλεται στην τελειοθηρεία και το πάθος του για την διεύρυνση της γνώσης της ανθρωπότητας.

Επίσης θα ήθελα να ευχαριστήσω τον φίλο και συνεργάτη κ. Ευαγγελίδη Γεώργιο, συνεπιβλέπωντα αυτής της διπλωματικής, για την πολύτιμη βοήθεια που μου παρείχε.

Ακόμα Θα ήθελα να ευχαριστίσω τον κ. Radu Patrice Horaud, διευθυντή της ομάδας PERCEPTION του ερευνητικού ινστιτούτου INRIA στην Grenoble της Γαλλίας για την εμπειρία που απέκτησα ως μέλος της ομάδας αυτής.

Τέλος θα ήθελα να ευχαριστήσω τον κ. Ευστράτιο Γαλλόπουλο. Οι κατευθύνσεις του οποίου, κατά το τρίτο έτος της φοίτησης μου έδωσαν κίνητρο για να συνεχίσω τις σπουδές μου στο τμήμα Μηχανικών Η/Υ.



# ΠΕΡΙΕΧΟΜΕΝΑ

<b>1</b>	<b>ΕΙΣΑΓΩΓΗ</b>	<b>7</b>
1.1	Διατύπωση του προβλήματος . . . . .	7
1.2	Στοιχειώδες Πρόβλημα: Αντιστοίχιση δύο νεφών, Προκύπτοντα υποπροβλήματα . . . . .	8
1.3	Η περίπτωση πολλαπλών επικαλυπτόμενων συνόλων . . . . .	8
1.4	Οργάνωση της παρούσας διπλωματικής . . . . .	8
<b>2</b>	<b>ΣΑΡΩΤΕΣ ΒΑΘΟΥΣ</b>	<b>11</b>
2.1	Τεχνολογίες σαρωτών βάθους . . . . .	12
2.2	Σαρωτές χρόνου πτήσης . . . . .	15
<b>3</b>	<b>ΜΑΘΗΜΑΤΙΚΟ ΥΠΟΒΑΘΡΟ</b>	<b>19</b>
3.1	Γεωμετρικοί μετασχηματισμοί . . . . .	19
3.1.1	Μητρώα περιστροφής . . . . .	19
3.1.2	Οικογένιες 2D γεωμετρικών μετασχηματισμών . . . . .	22
3.1.3	Οικογένιες 3D γεωμετρικών μετασχηματισμών . . . . .	27
3.2	Στοιχεία Ανάλυσης Μητρώων . . . . .	29
3.2.1	Πολική Διάσπαση και το Λήμμα Μέγιστου Τχνους . . . . .	29
3.3	Στοιχεία Θεωρίας Βελτιστοποίησης . . . . .	30
3.3.1	Βελτιστοποίηση κυρτών συναρτήσεων . . . . .	30
3.3.2	Πολλαπλασιαστές Lagrange . . . . .	31
3.4	Μοντέλα βασισμένα σε μίγματα Γκαουσιανών κατανομών . . . . .	33
3.4.1	Μείγματα Κανονικών Κατανομών . . . . .	33
3.4.2	Η σύγκλιση του EM υπο την ύπαρξη αγνώστων μεταβλητών . . . . .	34
<b>4</b>	<b>ΑΝΤΙΣΤΟΙΧΙΣΗ ΖΕΥΓΟΥΣ ΤΡΙΣΔΙΑΣΤΑΤΩΝ ΝΕΦΩΝ ΣΗΜΕΙΩΝ</b>	<b>37</b>
4.1	Εκ των προτέρων γνωστές αντιστοιχίες . . . . .	38
4.1.1	Το πρόβλημα του απόλυτου προσανατολισμού . . . . .	38
4.1.2	Επίλυση βασισμένη σε ορθομοναδιαία τετραδόνια . . . . .	39
4.1.3	Επίλυση βασισμένη σε ορθοκανωνικά μητρώα . . . . .	40
4.1.4	Η τροποποιημένη επίλυση του Shinji Umeyama . . . . .	41
4.1.5	Η συνέχεια Lipschitz και άλλες μετρικές σφάλματος . . . . .	45
4.2	Άγνωστες αντιστοιχίες . . . . .	46
4.2.1	Ο αλγόριθμος ICP και οι παραλλαγές του . . . . .	46
4.2.2	Ο αλγοριθμος NDT . . . . .	52
4.2.3	Ο αλγόριθμος ECM-PR . . . . .	56

---

4.3	Σύντομη επισκόπηση μερικών μεθόδων αντιστοίχισης εύκαμπτων και αρθρωτών σωμάτων . . . . .	64
<b>5</b>	<b>ΣΥΝΔΙΑΣΜΕΝΗ ΑΝΤΙΣΤΟΙΧΙΣΗ ΠΟΛΛΑΠΛΩΝ ΝΕΦΩΝ</b>	<b>67</b>
5.1	Εκ των προτέρων γνωστές αντιστοιχίες . . . . .	68
5.1.1	Απόλυτος προσανατολισμός πολλαπλών όψεων . . . . .	68
5.2	Άγνωστες αντιστοιχίες . . . . .	74
5.2.1	Ο αλγόριθμος ICP πολλαπλών όψεων . . . . .	74
<b>6</b>	<b>ΠΡΟΤΕΙΝΟΜΕΝΗ ΤΕΧΝΙΚΗ</b>	<b>77</b>
6.1	Διατύπωση του προβλήματος . . . . .	77
6.2	Μαθηματική Θεμελίωση του Αλγορίθμου . . . . .	78
6.2.1	Εκτίμηση του $\theta^*$ στην περίπτωση ισοτροπικών διασπορών . . . . .	81
6.2.2	Μία πιθανή υλοποίηση . . . . .	84
<b>7</b>	<b>ΠΕΙΡΑΜΑΤΑ</b>	<b>87</b>
7.1	Πειράματα αντιστοίχισης ζεύγους . . . . .	87
7.1.1	Με χρήση πραγματικών δεδομένων . . . . .	88
7.1.2	Με χρήση τεχνητών δεδομένων . . . . .	92
7.2	Πειράματα αντιστοίχισης πολλαπλών νεφών . . . . .	96
7.2.1	Οπτικά αποτελέσματα . . . . .	96
7.2.2	Ποσοτικοποίηση των αποκλίσεων . . . . .	97
	<b>ΠΑΡΑΡΤΗΜΑ</b>	<b>97</b>
	<b>Α΄ ΥΛΟΠΟΙΗΣΕΙΣ</b>	<b>101</b>
A.1	Υλοποιήσεις αντιστοίχισης ζεύγους . . . . .	103
A.2	Υλοποιήσεις πολλαπλής αντιστοίχισης . . . . .	125

# Κεφάλαιο 1

## ΕΙΣΑΓΩΓΗ

ΣΚΟΠΟΣ της αντιστοίχισης επιφανειών είναι να μετασχηματίσει πολλαπλά τρισδιάστατα σύνολα δεδομένων σε ένα κοινό σύστημα αναφοράς έτσι ώστε να ευθυγραμμίσει τμήματα επικαλυπτόμενων επιφανειών που συνυπάρχουν σε περισσότερα του ενός συνόλων. Αυτά τα σύνολα δεδομένων απαρτίζονται από σημεία που έχουν προέλθει από σάρωση-(scanning) επιφανειών τρισδιάστατων σχημάτων ή αντικειμένων. Η τεχνολογία τρισδιάστατης σάρωσης βρίσκεται ακόμα στην άνθιση της το οποίο συνεπάγεται ότι οι δυνατότητες των οργάνων μέτρησης (acquisition-devices) είναι περιορισμένες. Οι περιορισμοί αυτοί δημιουργούν την ανάγκη για πολλαπλές σαρώσεις της ίδιας φυσικής σχηματικής-αντικειμένου από διαφορετικές οπτικές γωνίες με σκοπό αυτές να επανασυντεθούν, με χρήση κατάλληλων τεχνικών, ώστε τελικά να έχουμε σαν αποτέλεσμα μια ολική ανακατασκευή της φυσικής σχηματικής. Η τελική ανακατασκευή ονομάζεται Άτλαντας και είναι πολύ χρήσιμος σε ποικίλες εφαρμογές όπως (medical image analysis, reverse engineering, mechanical engineering, object tracking, animation, etc.).

### 1.1 Διατύπωση του προβλήματος

Τα σύνολα σημείων που έχουν προέλθει από σάρωση επιφανειών, συνήθως, συναντώνται στη βιβλιογραφία ως νέφη σημείων (point clouds). Αφαιρετικά θα μπορούσαμε να τα φανταστούμε ως σύνολα σημείων που έχουν προκύψει από τη δειγματοληψία μιας επιφάνειας. Η αντιστοίχιση νέφων σημείων έχει μελετηθεί εκτενώς, από την δεκαετία του 1980, με θεωρήσεις ποικίλων παραμορφώσεων μεταξύ τους. Οι κυριότερες θεωρήσεις που έχουν γίνει μέχρι σήμερα υποθέτουν ότι η σχέση μεταξύ τους περιγράφεται πλήρως από έναν ολικό μετασχηματισμό άκαμπτου υλικού (rigid body transformation) Δηλαδή ότι το ένα εκ των δύο συνόλων (point clouds) έχει προέλθει από την εφαρμογή ενός Ευκλείδειου [3.1] μετασχηματισμού στο αρχικό νέφος σημείων. Τα τελευταία χρόνια έχουν δημοσιευτεί ποικίλες εργασίες, οι οποίες ενδιαφέρονται με την αντιστοίχιση τέτοιων συνόλων θεωρώντας ότι η μεταξύ τους σχέση δεν είναι πλέον Ευκλείδεια αλλά ότι τα σημεία μπορεί να έχουν υποστεί τοπικές Ευκλείδειες παραμορφώσεις (articulated body transformation)<sup>[1]</sup> ή ακόμα και μή Ευκλείδειες<sup>[3]</sup>.

## 1.2 Στοιχειώδες Πρόβλημα: Αντιστοίχιση δύο νεφών, Προκύπτοντα υποπροβλήματα

Ακόμα και η απλούστερη περίπτωση, του Ευκλείδειου μετασχηματισμού, εγείρονται ποικίλα υποπροβλήματα που πρέπει να επιλυθούν. Για παράδειγμα ας υποθέσουμε δύο σύνολα σημείων τα οποία έχουν προέλθει από την αποτύπωση της ίδιας φυσικής σκηνής μέσω του ίδιου συστήματος μέτρησης (acquisition device) που για χάριν ευκολίας από εδώ και στο εξής θα ονομάζεται κάμερα βάθους **Depth-camera**. Αφώτου αποτυπώσαμε το πρώτο μετακινήσαμε την Depth-camera στο χώρο, και στη συνέχεια αποτυπώσαμε το δεύτερο. Οπότε λόγω αποκλειόμενων περιοχών (**occlusions**)<sup>1</sup> ένα υποσύνολο των σημείων της πρώτης αποτύπωσης μπορεί να μην περιέχεται στη δεύτερη και αντίστροφα. Ως δεύτερο επακόλουθο θα ήταν ορθό να λάβουμε υπόψη μας ότι η κάμερα βάθους προσθέτει θόρυβο με τη μορφή διασποράς των πραγματικών σημείων αλλά και εικονικών σημείων που δεν υπάρχουν στη φυσική σκηνή. Επαγωγικά σκεπτόμενοι, το πρώτο συστατικό που χρειαζόμαστε για να λύσουμε το πρόβλημα της αντιστοίχισης είναι να γνωρίζουμε τα αντίστοιχα σημεία μεταξύ των δύο αποτυπώσεων καθώς και να έχουμε τη δυνατότητα να αφαιρούμε τα εικονικά σημεία (**spurious points - outliers**). Το πρόβλημα αυτό στη βιβλιογραφία αναφέρεται ως πρόβλημα αντιστοίχισης (**correspondence problem**) και θα χρησιμοποιείται αυτός ο όρος στη συνέχεια. Όπως θα δούμε παρακάτω για την εύρεση του βέλτιστου μετασχηματισμού, αφού γνωρίζουμε τις αντιστοιχίες μεταξύ των σημείων (**correspondences**), υπάρχουν κλειστού τύπου λύσεις που εγγυώνται βέλτιστη λύση ως προς κάποια μετρική σφάλματος.

## 1.3 Η περίπτωση πολλαπλών επικαλυπτόμενων συνόλων

Το πρόβλημα της αντιστοίχισης γίνεται ακόμα πιο δύσκολο όταν θέλουμε να συνθέσουμε περισσότερα των δύο νεφών σημείων. Η απλοϊκή προσέγγιση θα ήταν να τα συνθέσουμε ανά δύο και στη συνέχεια να συνθέσουμε τα ήδη αντιστοιχισμένα ζεύγη σε ένα κοινό σύστημα αναφοράς έως ότου καταλήξουμε σε ένα νέφος-άτλαντα που περιέχει όλα τα αρχικά υπο-νέφη. Το μειονέκτημα αυτής της προσέγγισης είναι ότι κάθε αντιστοίχιση των νεφών ανά δύο συσσωρεύει σφάλματα από τις προηγούμενες αντιστοιχίσεις τα οποία είναι πολύ πιθανό να καταστήσουν ατελέσφορο τον τελικό άτλαντα.

## 1.4 Οργάνωση της παρούσας διπλωματικής

Σκοπός μας στην παρούσα διπλωματική εργασία είναι η σε βάθος μελέτη των υπαρχόντων μεθόδων αντιστοίχισης, νεφών σημείων, και η υλοποίηση των βασικών τεχνικών ανά δύο [κεφάλαιο 4] και στη συνέχεια να επεκταθεί στις υπάρχουσες μεθόδους ταυτόχρονης αντιστοίχισης πολλαπλών νέφων [κεφάλαιο 5]. Στο [κεφάλαιο 6] προτείνεται ένας συμμετρικός αλγόριθμος ταυτόχρονης αντιστοίχισης πολλαπλών νεφών, όπου με τον όρο συμμετρικός εννοούμε ότι δεν εξαναγκάζει το κοινό σύστημα αναφοράς να βρισκείται σε κάποιο από τα προς αντιστοίχιση σύνολα

<sup>1</sup> Δηλαδή περιοχές ορατές μόνο από συγκεκριμένες οπτικές γωνίες.



αλλά μπορεί να είναι οπουδήποτε στο χώρο. Στο [κεφάλαιο 7] παρατίθενται πειράματα και μετρήσεις των δυνατοτήτων του προτεινόμενου αλγορίθμου καθώς και κάποιων εκ των υπολοίπων μεθόδων που υλοποιήθηκαν στη διαδικασία κατασκευής του τελικού αλγορίθμου. Συμπληρωματικά στο [κεφάλαιο 3] επεξηγούνται και αποδεικνύονται αναλυτικά διάφορες από τις μαθηματικές έννοιες και θεωρήματα που χρησιμοποιήθηκαν στην παρούσα διπλωματική. Ειδικότερα αναλύονται οι γεωμετρικοί μετασχηματισμοί, οι οποίοι αποτελούν ίσως το σημαντικότερο εργαλείο για την μαθηματική μοντελοποίηση και επίλυση προβλημάτων της υπολογιστικής όρασης. Επίσης παρουσιάζονται διάφορα πιθανοτικά μοντέλα, μετρικές, και παραμετροποιήσεις αλγορίθμων που χρησιμοποιήθηκαν στις μεθόδους αντιστοίχισης. Τέλος, στο παράρτημα της διπλωματικής δίνεται ο κώδικας των αλγορίθμων και των πειραμάτων του [κεφαλαίου 7]. Στην επόμενο [κεφάλαιο 2] γίνεται μια παρουσίαση των τεχνολογιών τρισδιάστατων σαρωτών καθώς και μία εκτενέστερη ανάλυση των σαρωτών ‘χρόνου πτήσης’ και συγκεκριμένα του μοντέλου SR 4000 που χρησιμοποιήθηκε για την απόκτηση πραγματικών δεδομένων.



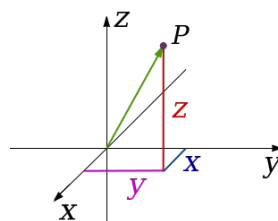
## Κεφάλαιο 2

# ΣΑΡΩΤΕΣ ΒΑΘΟΥΣ

Οι σαρωτές βάθους (3D scanners) είναι όργανα που χρησιμοποιούνται για την παραγωγή πρωταρχικών νεφών σημείων (raw point clouds) από γεωμετρικά δείγματα (geometric samples) της επιφάνειας ενός αντικειμένου. Αυτά τα σημεία μπορούν να χρησιμοποιηθούν στη συνέχεια στην διαδικασία της ολοκλήρωσης ή ανακατασκευής της μορφής-Σχήματος του αντικειμένου. Η διαδικασία αυτή ονομάζεται επίσης και ανακατασκευή (reconstruction). Οι σαρωτές βάθους έχουν πολλά κοινά στοιχεία με τις κλασικές φωτογραφικές μηχανές, όπως το κωνικό οπτικό πεδίο (conic field of view) και το ότι μπορούν να συλλέξουν πληροφορία μόνο για φωτεινές επιφάνειες (non obscured surfaces). Σε αναλογία με τις φωτογραφικές μηχανές οι οποίες συλλέγουν πληροφορία χρώματος για τις επιφάνειες που βρίσκονται στο οπτικό τους πεδίο οι σαρωτές βάθους συλλέγουν-μετρούν την απόσταση της αντίστοιχης επιφάνειας που φαίνεται στο εκάστοτε σημείο της 'εικόνας'. Η λειτουργία αυτή επιτρέπει την ανάκτηση της θέσης του εκάστοτε σημείου στο χώρο (three dimensional position of the point).

Όπως αναφέρθηκε και προηγουμένως, για τις περισσότερες εφαρμογές, μια μοναδική σάρωση δεν μπορεί να δώσει ένα πλήρες μοντέλο (3D model) του αντικειμένου. Πολλαπλές, οι οποίες αριθμούνται και σε εκατοντάδες, σαρώσεις μπορεί να χρειαστούν, κάθε μία από διαφορετική οπτική γωνία, για να μπορέσουμε να έχουμε, τελικά, πληροφορία για όλες τις πλευρές-πτυχές του αντικειμένου. Όστε κατόπιν να προχωρήσουμε στην διαδικασία αντιστοίχισης-σύνθεσης (registration-fusion process).

Αξίζει να σημειώσουμε έδω ότι οι σαρωτές επαφής επιστρέφουν αυτούσιες τις τρισδιάστατες συντεταγμένες των σημείων της επιφάνειας σε αντίθεση με τους σαρωτές χωρίς επαφή οι οποίοι αντί του βάθους  $z$ , η μέτρηση που επιστρέφουν είναι μια εκτίμηση της απόστασης του αντικειμένου από το οπτικό κέντρο (κέντρο του αισθητήρα ή το σημείο από το οποίο εκπέμπεται η φωτεινή ακτίνα). Η μέτρηση της απόστασης χρησιμοποιείται στη συνέχεια σε συνδιασμό με την οριζόντια και κατακόρυφη συντεταγμένη (lateral coordinates)  $x, y$  για την εξαγωγή της τιμής του βάθους  $z$ . Αυτό επιτυγχάνεται με τη βοήθεια βασικής τριγωνομετρίας.



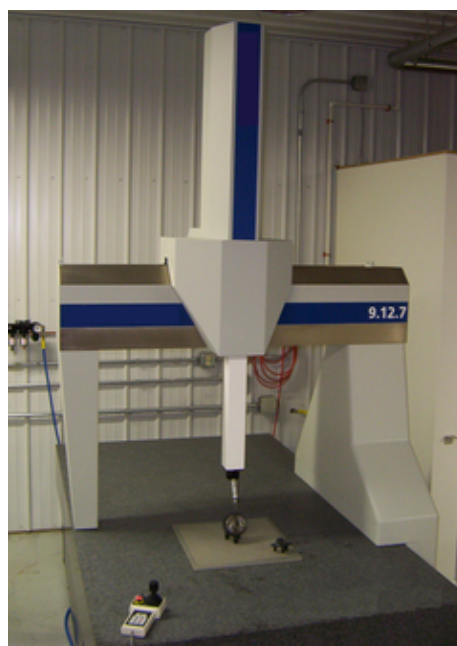
Σχήμα 2.1:  $P$  είναι η μέτρηση απόστασης ενώ  $Z$  το εξαγόμενο βάθος.

## 2.1 Τεχνολογίες σαρωτών βάθους

Οι υπάρχουσες τεχνολογίες σαρωτών βάθους χωρίζονται σε δύο μεγάλες κατηγορίες. Του σαρωτές επαφής (**contact scanners**) και τους σαρωτές χωρίς επαφή (**non-contact scanners**). Η δεύτερη κατηγορία χωρίζεται σε δύο υποκατηγορίες αυτούς με φωτεινή πηγή (**active**) και αυτούς που βασίζονται σε εξωτερική πηγή φωτός (**passive**) σαρωτές.

### A. Σαρωτές επαφής

Οι σαρωτές επαφής ιχνηλατούν (**probe**) το αντικείμενο χρησιμοποιώντας φυσική επαφή μέσω μίας ακίδας καθώς το αντικείμενο είναι τοποθετημένο σε μία επίπεδη πλάκα ακριβείας (**precision flat surface**). Η οποία, προηγουμένως έχει υποστεί λείανση με γνωστό μέγιστο εναπομένοντα βαθμό αδρότητας. Σε περιπτώσεις όπου το αντικείμενο δέν είναι επίπεδο ή δέν μπορεί να τοποθετηθεί σταθερά πάνω σε μία επιφάνεια συνήθως υποστηρίζεται μέσω κάποιας μέγγενης (**fixture tool - vise**). Οι σαρωτές επαφής, όπως αυτός που φαίνεται στην διπλανή εικόνα, συνήθως αποτελούνται από δύο κάθετους μεταξύ τους βραχίονες που έχουν τη δυνατότητα να ολισθαίνουν σε ράγα και επιπλέον έναν αρθρωτό βραχίονα με συμπαγείς ακίδες και υψηλής ακριβείας αισθητήρες γωνίας (**angular sensors**). Οι σαρωτές επαφής χαρακτηρίζονται από πολύ υψηλή ακρίβεια ( $\mu m$ ) αλλά εκτός του ότι είναι πολύ αργοί εξαιτίας των βαρέων μηχανικών μερών τους, εμπεριέχουν τον κίνδυνο καταστροφής του αντικειμένου. Συμπληρωματικά να αναφέρουμε ότι έχουν τον περιορισμό του μεγέθους των προς σάρωση αντικειμένων και ότι δέν μπορούν να χρησιμοποιηθούν για σάρωση εξωτερικών χώρων.



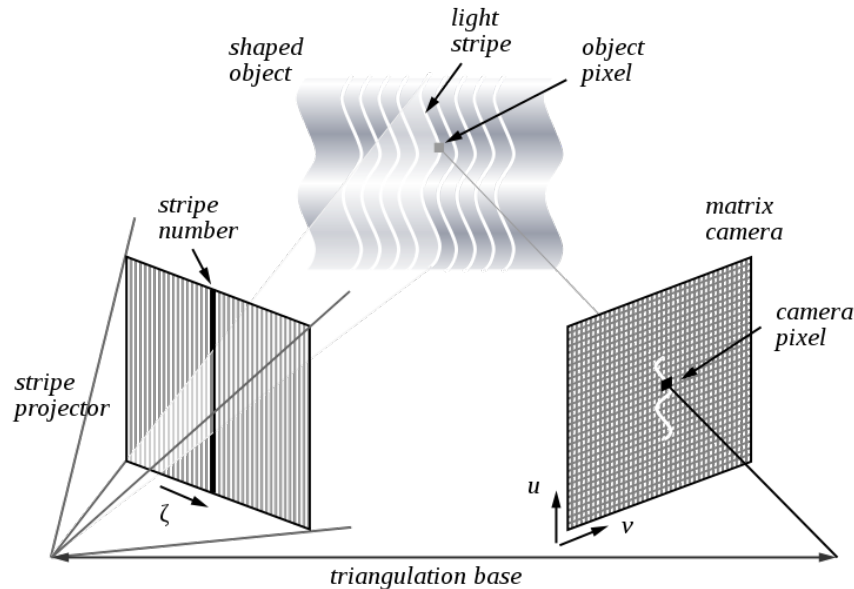
Σχήμα 2.2: Contact type 3D scanner - Coordinate measuring machine

### B. Ενεργοί Σαρωτές Χωρίς Επαφή

Οι ενεργοί σαρωτές χωρίς επαφή στέλνουν μία φωτεινή ακτίνα συνήθως laser και μετρούν τον χρόνο πτήσης ή τη γωνία διάθλασης αυτής μετά την πρόσκρουση της στο προς σάρωση αντικείμενο. Μέσω αυτών των ιδιοτήτων προσπαθούν να ανακτήσουν την απόσταση του σαρωτή από το αντικείμενο. Οι σαρωτές που καταμετρούν τον χρόνο πτήσης ονομάζονται **Time of Flight cameras** και θα αναλυθούν στην επόμενη [υποενότητα 2.2]. Οι σαρωτές που μετρούν τα χαρακτηριστικά της διάθλασης της φωτεινής ακτίνας χωρίζονται σε τρεις κατηγορίες, τους σαρωτές δομημένου φωτός (**structured light scanners**), τους σαρωτές τριγωνοποίησης (**triangulation scanners**) και τους σαρωτές διαμορφωμένου φωτός (**Modulated li-**

ght scanners).

### B-I. Σαρωτές δομημένου φωτός



Σχήμα 2.3: Η αρχή λειτουργίας ενός σαρωτή βάθους δομημένου φωτός.

Οι σαρωτές δομημένου φωτός (structured light) αποτελούνται από έναν εκπομπό δομημένου φωτός (LCD projector) και έναν αισθητήρα (CCD-CMOS typical camera sensor). Ο εκπομπός (light emitter) βομβαρδίζει το αντικείμενο με ακτίνες φωτός διαμορφωμένες σύμφωνα με κάποιο πρότυπο (pattern) και στη συνέχεια αναχτά το βάθος της επιφάνειας του αντικείμενου αναλύοντας το συχνοτικό περιεχόμενο της 'εικόνας' που αποτύπωσε ο αισθητήρας ή χρησιμοποιώντας άλλες τεχνικές μηχανικής μάθησης (Machine Learning).

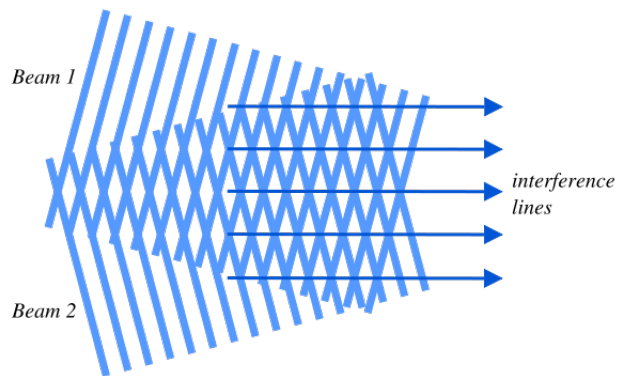
Η ακρίβεια αυτών των σαρωτών είναι της τάξης των  $10\mu m$ .

### B-II. Σαρωτές τριγωνοποίησης

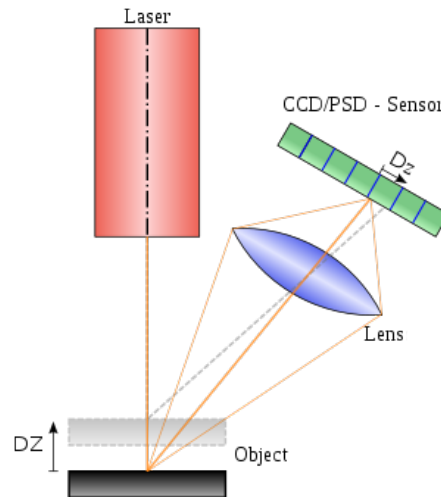
Οι σαρωτές τριγωνοποίησης (Triangulation scanners) αποτελούνται από μία φωτεινή πηγή, συνήθως laser και έναν αισθητήρα CCD-CMOS. Η αρχή στην οποία βασίζονται είναι ότι η ανακλώμενη ακτίνα ανάλογα με την απόσταση της επιφάνειας στην οποία προσέκρουσε θα 'πέσει' σε διαφορετικό σημείο στον αισθητήρα. Οπότε βρίσκοντας το σημείο στο οποίο 'βλέπει' ο αισθητήρας την φωτεινή ακτίνα, εκτιμάται η απόσταση της επιφάνειας του αντικείμενου.

### B-III. Σαρωτές διαμορφωμένου φωτός

Οι σαρωτές διαμορφωμένου φωτός (Modulated Light Scanners) εκπέμπουν ένα συνεχώς μεταβαλλόμενο 'φως' προς το αντικείμενο. Συνήθως η φωτεινή πηγή μεταβάλλει με ημιτονικό τρόπο το πλάτος (Amplitude) του εκπεμπόμενου φωτός.



Σχήμα 2.4: Η αρχή του δομημένου φωτός: δύο light patterns φαίνονται διαφορετικά από διαφορετικές οπτικές γωνίες λόγω υπέρθεσης και διάθλασης.



Σχήμα 2.5: Triangulation scanner principle: Find the dot!

Στη συνέχεια ένας φωτογραφικός αισθητήρας ανιχνεύει κατά πόσο ολισθήθηκε το ημιτονοειδές πρότυπο (*sinusoidal pattern*) και μέσω αυτού καθορίζει την απόσταση που ταξίδεψε το φως. Το πλεονέκτημα αυτής της τεχνολογίας είναι ότι επιτρέπει στον σαρωτή να αγνοήσει το 'φως' προερχόμενο από άλλες πηγές εκτός του laser οπότε δεν υπάρχει παρεμβολή (*interference*).

### C. Παθητικοί σαρωτές χωρίς επαφή

Οι non contact passive σαρωτές δεν εκπέμπουν κάποιο είδος ακτινοβολίας αλλά βασίζονται στην ακτινοβολία περιβάλλοντος (*ambient light*). Οι σαρωτές αυτού του τύπου ομαδοποιούνται σε δύο κατηγορίες, τους στερεοσκοπικούς και τους φωτομετρικούς.

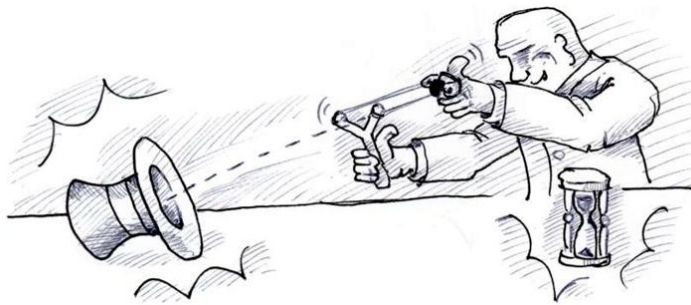
### C-I. Στερεοσκοπικοί Σαρωτές

Οι στερεοσκοπικοί σαρωτές ([stereoscopic passive scanners](#)) χρησιμοποιούν δύο κάμερες σε μικρή απόσταση μεταξύ τους ώστε να δημιουργήσουν στερεοσκοπικό οπτικό σύστημα και στη συνέχεια ανακτούν την απόσταση του αντικειμένου βασιζόμενοι στις αρχές της ανθρώπινης όρασης.

### C-II. Φωτομετρικοί σαρωτές

Οι φωτομετρικοί σαρωτές (Photometric Passive Scanners) χρησιμοποιούν μία κάμερα αλλά λαμβάνουν πολλαπλές 'φωτογραφίες' της ίδιας φυσικής σκηνής υπό διαφορετικές συνθήκες φωτισμού ([shutter speed , apperature or Exposure Value EV variaton](#)) και στη συνέχεια εξάγουν πληροφορία για την κατευθυντικότητα (orientation) της επιφάνειας σε κάθε εικονοστοιχείο-pixel.

## 2.2 Σαρωτές χρόνου πτήσης



Σχήμα 2.6: Time Of Flight (TOF) camera Principle.

Η αρχή στην οποία στηρίζεται η λειτουργία της κάμερας χρόνου πτήσης, ([Time Of Flight cameras \(TOF\)](#)), είναι η χρονομέτρηση του ταξιδιού του φωτός από τη στιγμή που αυτό φύγει από την κάμερα μέχρις ότου επιστρέψει, μετά την ανάκλαση του στην επιφάνεια του αντικειμένου. Το παραπάνω σκίτσο ίσως να θεωρηθεί αδόκιμο διότι η εκσφενδονιζόμενη πέτρα πιθανώς να μην επιστρέψει στον ρίπτη και επιπλέον η ταχύτητα της δέν είναι γνωστή (αντίθετα με αυτή του φωτός), παραταύτα παραμένει επεξηγηματικό. Η κατηγορία των TOF καμερών ανήκει στα [scannerless LIDARS](#)<sup>1</sup> με την έννοια οτι δέν χρησιμοποιεί μία μηχανοτρονικά κατευθυνόμενη ακτίνα για όλο το οπτικό πεδίο σαρώνοντας το σημείο πρὸς σημείο αλλά σε έναν, συνήθως υπέρυθρο, φωτοεκπομπό (light emitter) ο οποίος εκπέμπει βομβαρδίζει ταυτόχρονα όλα τα σημεία του οπτικού πεδίου της κάμερας. Οι κάμερες χρόνου πτήσης άρχισαν να εμφανίζονται γύρω στο 2000, όπου η τεχνολογία των ημιαγωγών κατάφερε να υποστηρίξει τόσο γρήγορες διαδικασίες όπως αυτή της χρονομέτρησης της 'πτήσης' του φωτός σε κοντινές αποστάσεις, με κάποια σχετική ακρίβεια. Η δυναμική περιοχή αυτών των καμερών ξεκινάει απο μερικά μέτρα και φτάνει εως και τα 60 μέτρα. Η ακρίβεια της μέτρησης κυμαίνεται γύρω

<sup>1</sup> LIDAR από το Laser-Radar ονομάζεται η τεχνολογία radar τα οποία στέλνουν μια ακτίνα πρὸς το στόχο (illuminating the target) και στη συνέχεια μετρούν την απόσταση του ανακλώμενο φῶς.

στο 1 τετραγωνικό εκατοστό. Η ανάλυση του κάδρου (lateral resolution) είναι αρκετά μικρότερη συγκριτικά με τις 2D κάμερες, συνήθως γύρω στα  $320 \times 240$  εικονοστοιχεία. Συγκριτικά με άλλες μεθόδους απόκτησης τρισδιάστατων 'εικόνων' αυτές οι κάμερες λειτουργούν πολύ γρήγορα παρέχοντας έως και 100 εικόνες ανά δευτερόλεπτο. Τα κύρια μέρη των TOF καμερών είναι τρία:

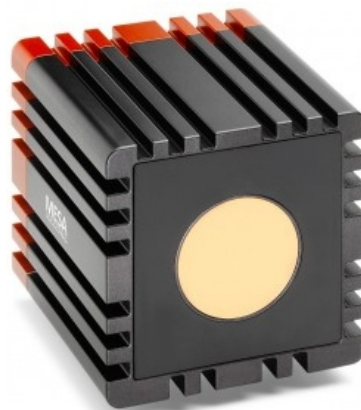
Α. Μια μονάδα φωτισμού της σκηνής (**illumination unit**) η οποία, εξαιτίας ότι το φως πρέπει να διαμορφωθεί σε συχνότητες τουλάχιστον  $100 \text{ MHz}$ , μπορεί να κατασκευαστεί μόνο από *LED* ή διόδους *Laser*. Συνήθως η ακτινοβολία είναι στο εύρος των υπερύθρων.

Β. Έναν φωτογραφικό αισθητήρα για τη 'μέτρηση' του χρόνου πτήσης. Οι κυριότερες τεχνικές με τις οποίες επιτυγχάνεται αυτό είναι η *RF Modulated light sources with phase detectors* και η *Range Gated Imagers* και οι οποίες παρουσιάζονται στη συνέχεια.

Γ. Τα κυκλώματα ελέγχου του αισθητήρα και της μονάδας φωτισμού ελέγχονται από υψίσυχνα σήματα. Τα σήματα αυτά πρέπει να είναι απόλυτα συγχρονισμένα ώστε να επιτευχθεί μια ικανοποιητική ακρίβεια. Για παράδειγμα εάν το σήμα ελέγχου της μονάδας φωτισμού και του αισθητήρα έχουν διαφορά φάσης  $10 \text{ picosecond}$  το λάθος στη μέτρηση της απόστασης αυξάνει κατά  $1.5 \text{ millimeter}$ . Συγκριτικά, οι υπάρχοντες μικροεπεξεργαστές λειτουργούν περίπου στα  $3 \text{ GHz}$ , το οποίο αντιστοιχεί σε περιόδους των  $300 \text{ picosecond}$  όπου η αντίστοιχη ακρίβεια της μέτρησης κυμαίνεται στα  $45 \text{ millimeters}$ .

### I. RF Modulated light sources with phase detectors:

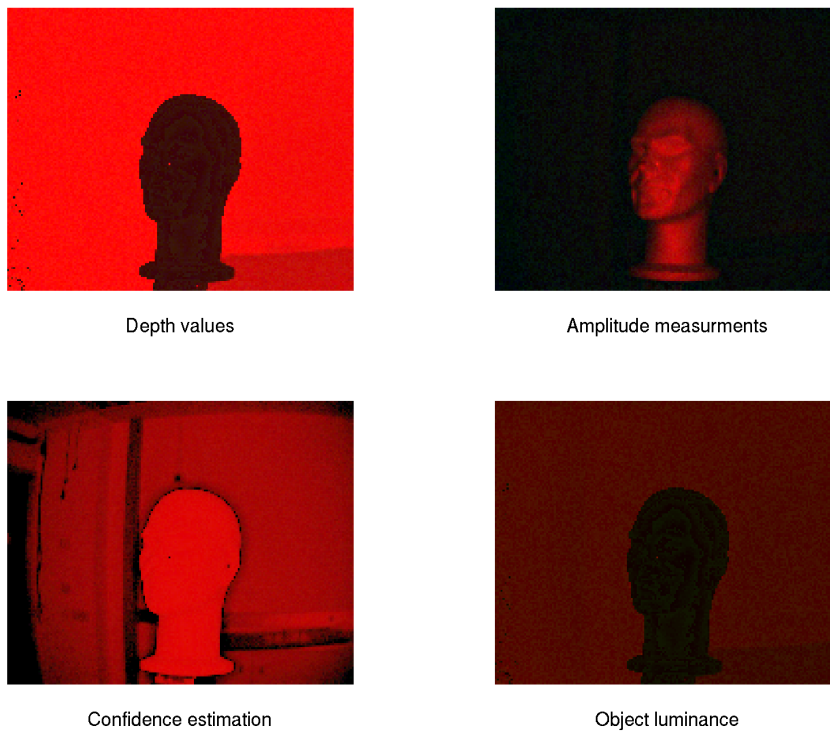
Αυτή η κατηγορία λειτουργεί διαμορφώνοντας την εξερχόμενη ακτίνα φωτός πάνω σε μία φέρουσα στην περιοχή των ραδιοσυχνοτήτων (*RF carrier*), και στη συνέχεια μετρώντας την ολίσθηση φάσης της φέρουσας στην πλευρά του δέκτη. Αυτή η προσέγγιση έχει δυναμική περιοχή εξαρτώμενη από το μήκος κύματος της φέρουσας, διότι η ολίσθηση φάσης επαναλαμβάνεται για χρόνους πτήσης μεγαλύτερους της μίας περιόδου. Η κάμερα η οποία χρησιμοποιήθηκε για τα πειράματα της παρούσας διπλωματικής είναι μία *SwissRanger MESA 4000* η οποία υπάγεται σε αυτή την κατηγορία με τη διαφορά ότι χρησιμοποιεί παλμικά (*pulsed*) *LEDs* αντί για πομπό *laser*. Η ανάλυση της συγκεκριμένης κάμερας είναι  $176 \times 144$  εικονοστοιχεία και η δυναμική της περιοχή κυμαίνεται από 5 έως 10 μέτρα. Η *SR 4000* επιστρέφει τέσσε-



Σχήμα 2.8: SR4000 by MESA

ρις εξόδους: μέτρηση φάσης, τις συντεταγμένες ( $x, y, z$ ) το πλάτος (**amplitude**) που είναι ανάλογο του ποσοστού του φωτός που ανακλάστηκε και τέλος μία εκτίμηση εμπιστοσύνης (**confidence**) των μετρήσεων για κάθε εικονοστοιχείο. Το λογισμικό που χρησιμοποιήσαμε μας έδινε σαν έξοδο μία εικόνα της μορφής του Σχήματος 7. Οι τέσσερις υποεικόνες παρέχουν, από πάνω προς τα κάτω, την απόσταση (**distance**) του κάθε σημείου (απο την οποία εξάγαμε τις συντεταγμένες που χρησιμοποιούσαμε στα πειράματα), το πλάτος (**amplitude**), την τιμή εμπιστο-





Σχήμα 2.7: Raw data provided by the MESA SR4000 in a single image (segm. for visualization)

σύνης (confidence value) και τέλος μια εκτίμηση της φωτεινότητας (luminance) του ίδιου του σημείου.

**I. Range Gated Imagers:** Οι συσκευές αυτού του τύπου έχουν ενσωματωμένο ένα κλείστρο-φωτοφράκτη (*shutter*) μπροστά από τον αισθητήρα ο οποίος ανοιγοκλείνει με τον ίδιο ρυθμό με τον οποίο εξέρχονται οι φωτεινοί παλμοί (light pulses). Εξαιτίας του ότι ένα μέρος του επιστρέφοντος φωτός δεν φτάνει στον αισθητήρα λόγω της ύπαρξης του φωτοφράκτη, το ποσοστό του φωτός που τελικά φτάνει στον αισθητήρα είναι ανάλογο της απόστασης που διάνυσε ο φωτεινός παλμός.

Οι κάμερες χρόνου πτήσης **πλεονεκτούν** σε σχέση με τις υπόλοιπες τεχνολογίες σαρωτών βάθους στο ότι είναι απλοϊκές στην κατασκευή τους και επιπλέον δεν απαιτούν μηχανικά μέρη όπως τα **Lidars**<sup>1</sup> αφού εκπέμπουν φως προς όλα τα σημεία της επιφάνειας ταυτόχρονα. Εξαιτίας της ταυτόχρονης μέτρησης όλης της σκηνής επιτυγχάνουν πολύ υψηλούς ρυθμούς (frame rates) έως και 100 πλαίσια ανα δευτερολεπτο. Από την άλλη πλευρά **μειονεκτούν** στο ότι είναι επιρρεπείς στην ύπαρξη μεγάλης φωτεινής έντασης από το αντικείμενο προς την κάμερα. Επιπλέον μειονεκτούν σε σχέση με τα **Lidars**<sup>1</sup> στο ότι φωτίζοντας όλη τη σκηνή ταυτόχρονα δημιουργούνται πολλαπλές ανακλάσεις<sup>2</sup> που μπορεί τελικά να κάνουν την μετρούμενη απόσταση, του σημείου, να είναι μεγαλύτερη από την πραγματική.

<sup>2</sup>διότι το φως μπορεί να φτάσει ένα σημείο του αντικειμένου με πολλαπλούς τρόπους.



## Κεφάλαιο 3

# ΜΑΘΗΜΑΤΙΚΟ ΥΠΟΒΑΘΡΟ

### 3.1 Γεωμετρικοί μετασχηματισμοί

Όπως τέθηκε από τον Felix Klein<sup>[8]</sup> το 1939: “Γεωμετρία είναι η μελέτη ιδιοτήτων αμετάβλητων κάτω από δεδομένες ομάδες μετασχηματισμών (*groups of transformations*)”. Στην περίπτωση της υπολογιστικής όρασης, ειδικότερα είναι η μελέτη ιδιοτήτων των σημείων του επιπέδου (στις δύο διαστάσεις) και κατ’ επέκταση του χώρου (στις 3 διαστάσεις) που παραμένουν αμετάβλητες κάτω από ομάδες μετασχηματισμών. Στην περίπτωση που οι μετασχηματισμοί είναι γραμμικοί<sup>1</sup> μπορούν να περιγραφούν με τον πολλαπλασιασμό του διανύσματος συντεταγμένων του σημείου με ένα μητρώο το οποίο περιγράφει τον μετασχηματισμό. Το μητρώο αυτό ονομάζεται μητρώο του μετασχηματισμού (*transformation matrix*). Ποιά φορμαλιστικά: Αν  $\mathcal{T}$  είναι ένας γραμμικός μετασχηματισμός και  $\vec{x}$  είναι το διάνυσμα συντεταγμένων ενός σημείου, ο μετασχηματισμός μπορεί να γραφεί ως  $\mathcal{T}(\vec{x}) = \mathbf{T}\vec{x}$  όπου  $\mathbf{T}$  είναι το αντίστοιχο μητρώο του μετασχηματισμού και ο πολλαπλασιασμός είναι η πράξη του συνόλου. Οι οικογένειες μετασχηματισμών που θα αναφέρουμε στη συνέχεια περιγράφουν κυρίως γραμμικούς μετασχηματισμούς, αναφέρονται επίσης και κάποιες περιπτώσεις απλών μη γραμμικών μετασχηματισμών (π.χ. προβολικοί). Μη γραμμικοί μετασχηματισμοί που επιφέρουν παραμορφώσεις, όπως καμπυλώσεις ευθειών, είναι πέρα από τα πλαίσια αυτής της διπλωματικής διότι είναι μη αντιστρέψιμοι και είναι σχεδόν αδύνατο να ανακτηθούν.

#### 3.1.1 Μητρώα περιστροφής

Ας υποθέσουμε ένα σημείο  $\mathbf{v}$ , όπως αυτό που φαίνεται στο Σχήμα 3.1, το οποίο περιστρέφεται (αντιρολογιακά) κατά γωνία  $\theta$  στο σημείο  $\mathbf{v}^{rot}$ . Η περιστροφή αυτή μπορεί να περιγραφεί με δύο τρόπους. Είτε ως περιστροφή του διανύσματος συντεταγμένων του  $\mathbf{v} = [x \ y]^T$  κατά  $\theta$  ως προς την αρχή των αξόνων (*dashed line*) είτε ως περιστροφή του ίδιου του συστήματος συντεταγμένων κατά γωνία  $\theta$  (*basis vector arrows*). Για να υπολογίσουμε τις συντεταγμένες του σημείου  $\mathbf{v}^{rot}$

<sup>1</sup>Υπό τον αυστηρό ορισμό, γραμμικός μετασχηματισμός είναι μια συνάρτηση μεταξύ δυο διανυσματικών χώρων η οποία διατηρεί τις πράξεις πρόσθεσης και βαθμωτού πολλαπλασιασμού μεταξύ διανυσμάτων αυτών των χώρων.

αρχικά, χρησιμοποιώντας βασική τριγωνομετρία, υπολογίζουμε τις συντεταγμένες των περιστραμμένων διανυσμάτων  $\tilde{\mathbf{e}}_x, \tilde{\mathbf{e}}_y$  της τυπικής βάσης του  $\mathbb{R}^2$ :

$$\begin{aligned} e_x &= [1 \ 0]^T \stackrel{\text{ccw rotation by } \theta}{\circlearrowleft} \tilde{\mathbf{e}}_x = [\cos(\theta) \ \sin(\theta)]^T \\ e_y &= [0 \ 1]^T \stackrel{\text{ccw rotation by } \theta}{\circlearrowleft} \tilde{\mathbf{e}}_y = [-\sin(\theta) \ \cos(\theta)]^T. \end{aligned} \quad (3.1)$$

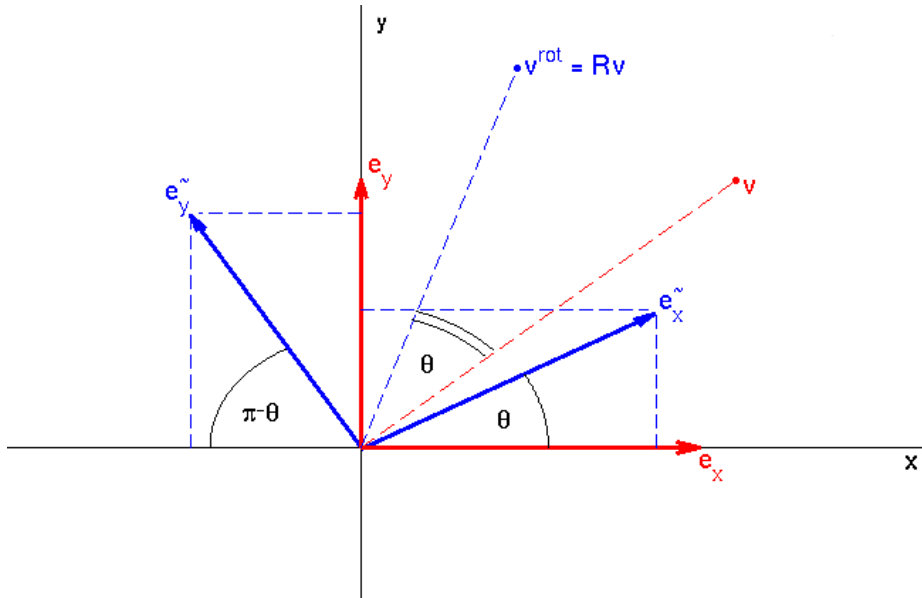
Όπως γνωρίζουμε από τη γραμμική άλγεβρα η προβολή ενός διανύσματος στο χώρο στηλών ενός μητρώου προκύπτει από τον πολλαπλασιασμό του διανύσματος με το αντίστοιχο μητρώο. Στην περίπτωση μας θέλουμε να υπολογίσουμε την προβολή  $\mathbf{v}^{rot}$  του  $\mathbf{v}$  στο χώρο που δημιουργούν τα  $\tilde{\mathbf{e}}_x, \tilde{\mathbf{e}}_y$ :

$$\mathbf{v}_{rot} = \begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{e}}_x & \tilde{\mathbf{e}}_y \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}. \quad (3.2)$$

Το μητρώο που κατασκευάσαμε επιτυχάνει αντιρολογιακή περιστροφή ενός σημείου-διανύσματος του επιπέδου κατά γωνία  $\theta$  και ονομάζεται μητρώο περιστροφής  $\mathbf{R}$ :

$$\mathbf{R} = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix}. \quad (3.3)$$

Εύκολα μπορούμε να παρατηρήσουμε ότι μητρώα αυτής της μορφής είναι ορθοκανονικά και η ορίζουσα τους είναι  $\pm 1$ . Το σύνολο (set) αυτών των μητρώων μαζί με την πράξη του πολλαπλασιασμού κατασκευάζουν το σύνολο  $SO(2)$ , (Special<sup>2</sup> Orthogonal group on 2 dimensions).



Σχήμα 3.1: From human perception to  $SO(2)$

<sup>2</sup>Τυπικά μόνο τα μητρώα που έχουν ορίζουσα  $|\mathbf{R}| = +1$  (proper rotations) δημιουργούν το σύνολο. Διότι για παράδειγμα σύνθεση δύο μητρώων με  $|\mathbf{R}| = -1$  (improper rotations - reflections) δίνει μητρώο με  $|\mathbf{R}| = +1$ , και επαγωγικά αλλοιώνει μία από τις ιδιότητες του συνόλου.

### Expanding our view: from the plane to 3D Euclidean space

Αν επεκτείνουμε τη συλλογιστική μας μπορούμε να δούμε ότι στην πραγματικότητα στο Σχήμα 3.1 περιγράφουμε την περιστροφή του τρισδιάστατου σημείου  $\mathbf{v} = [x \ y \ 0]^T$  στο χώρο ως προς γωνία  $\theta$  γύρω από τον άξονα βάθους  $z$ . Με αυτό το σκεπτικό θα μπορούσαμε να ξαναγράψουμε την Σχέση (3.2), επαυξάνοντας το μητρώο  $\mathbf{R}$  ως:

$$\mathbf{v}_{rot} = \begin{bmatrix} x' \\ y' \\ 0 \end{bmatrix} = \begin{bmatrix} \cos(\theta) & \sin(\theta) & 0 \\ -\sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 0 \end{bmatrix}. \quad (3.4)$$

Το επαυξημένο αυτό μητρώο επιτυγχάνει τρισδιάστατη περιστροφή κατά γωνία  $\theta$  ως προς τον άξονα  $z$ . Εναλλάσσοντας τη θέση της στήλης του κλειδωμένου<sup>3</sup> άξονα μπορούμε να κατασκευάσουμε τα αντίστοιχα μητρώα περιστροφής ως προς  $x$  και  $y$ . Συνοψίζοντας:

$$\begin{aligned} \mathbf{R}_x &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & \sin(\theta) \\ 0 & -\sin(\theta) & \cos(\theta) \end{bmatrix}, \\ \mathbf{R}_y &= \begin{bmatrix} \cos(\theta) & 0 & -\sin(\theta) \\ 0 & 1 & 0 \\ \sin(\theta) & 0 & \cos(\theta) \end{bmatrix}, \\ \mathbf{R}_z &= \begin{bmatrix} \cos(\theta) & \sin(\theta) & 0 \\ -\sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix}. \end{aligned} \quad (3.5)$$

Όπως και στις δύο διαστάσεις, η συλλογή των μητρώων  $\mathbf{R}_x, \mathbf{R}_y, \mathbf{R}_z \forall \theta$  μαζί<sup>4</sup> με την πράξη του πολλαπλασιασμού ορίζουν το αντίστοιχο σύνολο που συμβολίζεται  $SO(3)$ . Όπως γνωρίζουμε κάθε περιστροφή στον Ευκλείδειο χώρο (3D Euclidean Space) μπορεί να κατασκευαστεί από τρεις περιστροφές, μία ως προς κάθε άξονα ( $x, y, z$ ) της τυπικής βάσης. Στην περίπτωση που χρησιμοποιούμε μητρώα η αντίστοιχη περιστροφή μπορεί να επιτευχθεί μέσω σύνθεσης-πολλαπλασιασμού των τριών στοιχειώδων μητρώων  $\mathbf{R}_x, \mathbf{R}_y, \mathbf{R}_z$  με γωνίες  $\theta_x, \theta_y$  και  $\theta_z$  αντίστοιχα.

### Γωνίες Euler και η αναπαράσταση άξονα-γωνίας:

Το 1775 ο Leonhard Euler απέδειξε ότι οποιαδήποτε περιστροφή ενός άκαμπτου σώματος στην οποία τουλάχιστον ένα σημείο του σώματος παραμένει αμετάβλητο, μπορεί να αναπαρασταθεί ως μία περιστροφή γύρω από έναν άξονα που περνάει από αυτό το σημείο κατά μία συγκεκριμένη γωνία. Η αναπαράσταση αυτή ονομάζεται αναπαράσταση άξονα-γωνίας ([axis angle representation](#))<sup>5</sup>.

#### Μετατροπή μητρώου περιστροφής σε αναπαράσταση άξονα γωνίας:

Δεδομένου ενός μητρώου περιστροφής  $\mathbf{R}$  ο άξονας (Euler axis)  $\mathbf{u}$  μπορεί να βρε-

<sup>3</sup>Κλειδωμένος άξονας ονομάζεται ο άξονας ως προς τον οποίο γίνεται η περιστροφή.

<sup>4</sup>Μία ομάδα (group) ορίζεται ως μία συλλογή (set) εφοδιασμένη με μια συνολοθεωρητική πράξη τέτοια ώστε η εφαρμογή της στα στοιχεία του συνόλου δημιουργεί στοιχεία που επίσης ανήκουν στη σύνολο.

<sup>5</sup>Το πλεονέκτημα της αναπαράστασης άξονα-γωνίας έναντι των Euler angles είναι ότι αποφεύγει το φαινόμενο gimbal lock.

θεί ως το ιδιοδιάνυσμα του  $\mathbf{R}$  που αντιστοιχεί στην ιδιοτιμή<sup>6</sup>  $\lambda = 1$ . Στη συνέχεια η γωνία περιστροφής  $\phi$  μπορεί να βρεθεί ως:

$$\phi = \arccos \frac{1 - \text{tr}(\mathbf{R})}{2} \quad (3.6)$$

όπου το  $\text{tr}(\cdot)$  συμβολίζει τον τελεστή ίχνους.

Μετατροπή αναπαράστασης άξονα γωνίας σε μητρώο περιστροφής:

Αντίστοιχα δεδομένου του άξονα  $\mathbf{u}$  και της γωνίας περιστροφής  $\phi$  το μητρώο περιστροφής  $\mathbf{R}$  μπορεί να υπολογιστεί βάσει της ακόλουθης σχέσης του Rodriguez ως:

$$\mathbf{R} = \cos(\phi)\mathbf{I} + \sin(\phi)[\mathbf{u}]_x + (1 - \cos(\phi))\mathbf{u}\mathbf{u}^T, \quad (3.7)$$

όπου το  $[\mathbf{u}]_x$  ορίζεται ως:

$$[\mathbf{u}]_x = \begin{bmatrix} 0 & -u_z & -u_y \\ u_z & 0 & -u_x \\ u_y & u_x & 0 \end{bmatrix}. \quad (3.8)$$

συμβολίζει το cross product matrix του  $\mathbf{u}$ .

### 3.1.2 Οικογένειες $2\Delta$ γεωμετρικών μεσασχηματισμών

Στην περίπτωση σημείων στις 2 διαστάσεις ( $2\Delta$ ) οι γεωμετρικοί μεσασχηματισμοί χωρίζονται σε τέσσερις βασικές κατηγορίες. Οι δυνατότητες και οι περιορισμοί τους αναλύονται στη συνέχεια και ταυτόχρονα δίνονται οπτικά παραδείγματα από την εφαρμογή τους σε πραγματικές εικόνες.

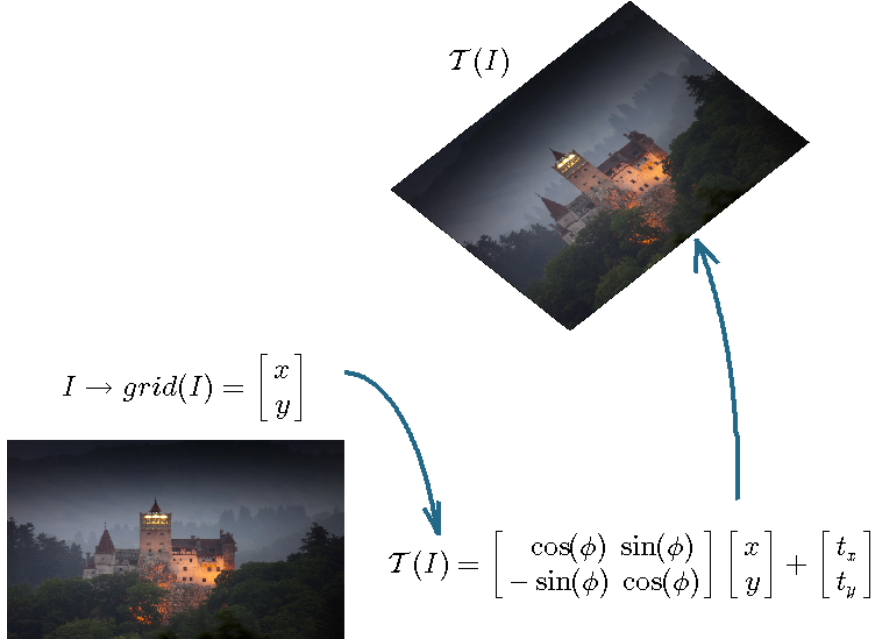
#### I. Ευκλείδειοι μεσασχηματισμοί (Euclidean transformations)

Αποτελούν την πιο περιοριστική κατηγορία προβολικών μεσασχηματισμών (ονομάζονται επίσης και μεσασχηματισμοί ισομετρίας) και χαρακτηρίζονται από 3 βαθμούς ελευθερίας. Οι παραμορφώσεις που μπόρουν να επιτύχουν χαρακτηρίζονται από περιστροφές και ολισθήσεις των σημείων στο επίπεδο. Η σύνθεση (*composition*) τέτοιων μεσασχηματισμών αποτελεί επίσης Ευκλείδειο μεσασχηματισμό. Το μεσασχηματίζων μητρώο της οικογένειας αυτής έχει την ακόλουθη μορφή.

$$\mathbf{T}_e = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix} \quad (3.9)$$

Το υπομητρώο  $[t_x \ t_y]^T$  περιέχει τους 2 από τους 3 βαθμούς ελευθερίας του μεσασχηματισμού και επιτυγχάνει ολίσθηση του σημείου στο επίπεδο. Ο τρίτος βαθμός ελευθερίας ( $\theta$ ) διαμορφώνει το  $2 \times 2$  κύριο υπομητρώο του  $\mathbf{T}_e$ , το οποίο είναι το μητρώο περιστροφής της σχέσης (3.3). Αξίζει να σημειωθεί ότι το σύνολο (set) των Ευκλείδειων μεσασχηματισμών, μαζί με την πράξη του πολλαπλασιασμού ορίζει το σύνολο  $E(2)$  ή  $ISO(2)$  (**2 Dimensional Euclidean group**). Επιπλέον στην περίπτωση που το μητρώο περιστροφής δεν εσωκλείει ανάκλαση το σύνολο των μεσασχηματισμών ονομάζεται  $SE(2)$  (**Special Euclidean group**).

<sup>6</sup>Κάθε μητρώο περιστροφής έχει μια τέτοια ιδιοτιμή



Σχήμα 3.2: Αποτέλεσμα από την εφαρμογή Ευκλείδειου μετασχηματισμού, με ολίσθηση και περιστροφή κατά  $\phi = \pi/4$ .

**Αμετάβλητες ιδιότητες.** Η εφαρμογή Ευκλείδειων μετασχηματισμών που ανήκουν στο  $SE(2)$  διατρέπει τον προσανατολισμό (orientation), ενώ σε αντίθετη περίπτωση, αρνητικής ορίζουσας, τον αντιστρέφει. Επίσης αφήνει αναλλοίωτες τις γωνίες μεταξύ ευθειών, τα εμβαδά και τα μήκη των σχημάτων<sup>7</sup>.

## II. Μετασχηματισμοί ομοιότητας (similarity transformations)

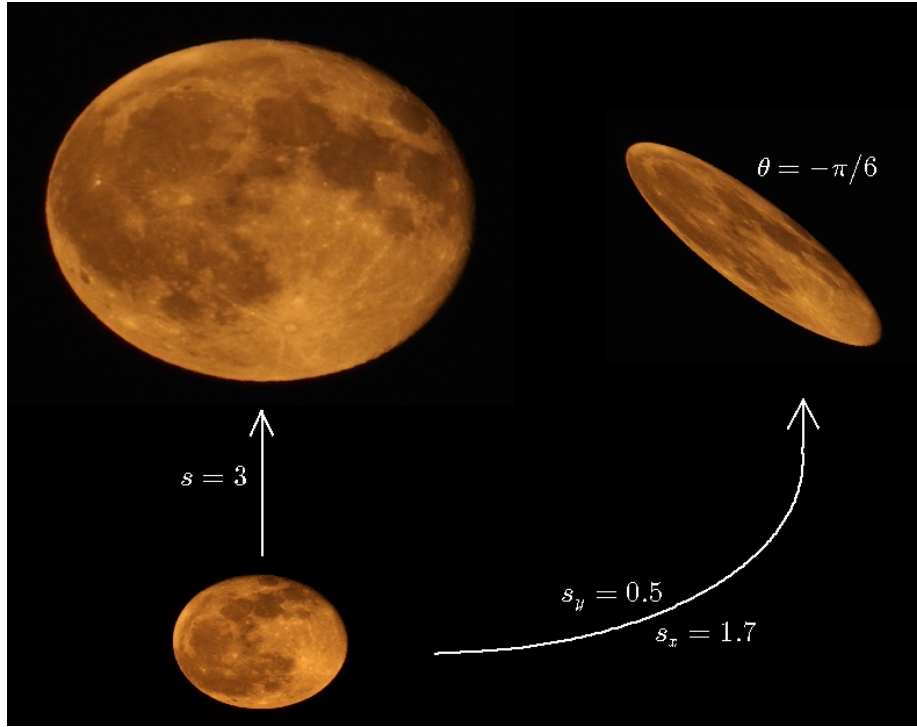
Η οικογένεια αυτή επεκτείνει τους Ευκλείδειους μετασχηματισμούς επιτρέποντας, επιπλέον, την ισοτροπική κλιμάκωση (scaling) της εικόνας. Οι μετασχηματισμοί ομοιότητας χαρακτηρίζονται από 4 βαθμούς ελευθερίας. Τρεις για το Ευκλείδειο τμήμα  $(\theta, t_x, t_y)$  και ένας επιπλέον για την κλιμάκωση  $(s)$ .

Το μετασχηματίζων μητρώο της οικογένειας αυτής κατασκευάζεται με τον πολλαπλασιασμό του υπομητρώου περιστροφής με μια βαθμωτή ποσότητα.

$$\mathbf{T}_s = \begin{bmatrix} s \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} & \begin{bmatrix} t_x \\ t_y \end{bmatrix} \\ 0 & 0 & 1 \end{bmatrix} \quad (3.10)$$

Στο ακόλουθο Σχήμα φαίνεται το αποτέλεσμα μιας μεγένθυσης (similarity) και μιας ανισοτροπικής κλιμάκωσης (ανεξάρτητης κλιμάκωσης σε κάθε διάσταση). Η

<sup>7</sup>Στην ουσία ο μετασχηματισμός εφαρμόζεται σε ένα σύνολο σημείων. Η αναφορά σε μήκη και εμβαδά προϋποθέτει την ύπαρξη σχημάτων. Αφαιρετικά σκεπτόμενοι μπορούμε να θεωρήσουμε μια εικόνα ως ένα σύνολο σημείων εφοδιασμένων με υφή (textured point set) όπου το καθένα εκτός από τις συντεταγμένες του ορίζει και ένα εικονοστοιχείο (pixel) με κέντρο το σημείο και δεδομένη φωτεινότητα (intensity).



Σχήμα 3.3: Ισοτροπική (left) και ανισοτροπική (right) κλιμάκωση

αρχική εικόνα φαίνεται στο κάτω αριστερά.

**Αμετάβλητες ιδιότητες.** Η εφαρμογή μετασχηματισμών ομοιότητας διατηρεί τους λόγους των μηκών καθώς και τις γωνίες μεταξύ ευθυγράμμων τμημάτων. Επιπλέον δεν επηρεάζει τα κυκλωτερή σημεία.

### III. Μετασχηματισμοί συγγένειας (affine transformations)

Πρόκειται για μη ιδιάζοντες μετασχηματισμούς που προκύπτουν εισάγοντας δύο επιπλέον βαθμούς ελευθερίας στους μετασχηματισμούς ομοιότητας. Συνολικά χαρακτηρίζονται από 6 βαθμούς ελευθερίας και το μετασχηματίζων μητρώο έχει την ακόλουθη μορφή.

$$\mathbf{T}_{af} = \begin{bmatrix} a_{11} & a_{12} & t_x \\ a_{21} & a_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix} \quad (3.11)$$

Για να γίνουν αντιληπτές οι παραμορφωτικές δυνατότητες των μετασχηματισμών συγγένειας εφαρμόζουμε διάσπαση QR στο  $2 \times 2$  κύριο υπομητρώο του  $\mathbf{T}_{af}$ :

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \underbrace{\begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix}}_{\text{orthonormal}} \underbrace{\begin{bmatrix} \text{sgn}(r_{11}) & 0 \\ 0 & \text{sgn}(r_{22}) \end{bmatrix}}_{\text{rotation by } \theta} \underbrace{\begin{bmatrix} \frac{1}{|r_{11}|} & 0 \\ 0 & \frac{1}{|r_{22}|} \end{bmatrix}}_{\text{scaling}(s_x, s_y)} \underbrace{\begin{bmatrix} 1 & r_{12} \\ 0 & 1 \end{bmatrix}}_{\text{shear}(\gamma)} \quad (3.12)$$



ή SVD (διάσπαση ιδιζουσών τιμών):

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \underbrace{\begin{bmatrix} \cos(\phi) & \sin(\phi) \\ -\sin(\phi) & \cos(\phi) \end{bmatrix}}_{U(\text{orthonormal})} \underbrace{\begin{bmatrix} s_x & 0 \\ 0 & s_y \end{bmatrix}}_S \underbrace{\begin{bmatrix} \cos(\eta) & \sin(\eta) \\ -\sin(\eta) & \cos(\eta) \end{bmatrix}}_{V^T(\text{orthonormal})}. \quad (3.13)$$

Απο τις παραπάνω διασπάσεις γίνεται εύκολα αντιληπτό ότι ένας μετασχηματισμός συγγένειας μπορεί να επιτευχθεί με στρέβλωση<sup>8</sup> ακολουθούμενο από μια ανισοτροπική κλιμάκωση και μια περιστροφή (3.12) ή από δύο περιστροφές μεταξύ των οποίων παρεμβάλεται μια ανισοτροπική κλιμάκωση όπως φαίνεται στη σχέση (3.13). Οι δυο παραπάνω διασπάσεις δεν είναι μοναδικές, για παράδειγμα θα μπορούσαμε να εφαρμόσουμε διάσπαση LU στο υπομητρώ του  $\mathbf{T}$  όπου ο μετασχηματισμός θα επιτυγχανόταν μέσω δυο διαδοχικών shearing. Αξίζει να σημειωθεί ότι στην περίπτωση που το κύριο υπομητρώ του  $\mathbf{T}_{af}$  είναι ιδιάζων ο μετασχηματισμός αποδομείται σε μια μη αντιστρέψιμη παραμόρφωση η οποία καταστέλει τη μία ή και τις δύο διαστάσεις του συνόλου σημείων. Η διάσπαση ιδιζουσών τιμών δείχνει με σαφή τρόπο τι συμβαίνει σε αυτή την περίπτωση.



Σχήμα 3.4: Παράδειγμα μη ιδιάζοντος μετασχήματ. συγγένειας

**Αμετάβλητες ιδιότητες.** Οι μετασχηματισμοί συγγένειας διατηρούν τις παραλληλίες, τους λόγους των εμβαδών, την ιδανική ευθεία<sup>9</sup> όπως επίσης και τους λόγους των μηκών μεταξύ συγγραμμικών ή παράλληλων σημείων.

<sup>8</sup>στρέβλωση shear ονομάζεται γραμμική μετατόπιση (displacement) ενός σημείου ως προς μία δεδομένη κατεύθυνση σε ποσοστό ανάλογο μιας προσημασμένης απόστασης **signed distance**. Για παράδειγμα μετατόπιση της τετμημένης ενός σημείου σε ποσοστό ανάλογο της τεταγμένης του : *i.e.*,  $x' = x + \gamma y$

<sup>9</sup>The line at infinity  $I_\infty$  is an ideal line of the projective plane, on which, parallel lines are assumed to be intersecting.

### III. Μετασχηματισμοί ομογραφίας (homography)

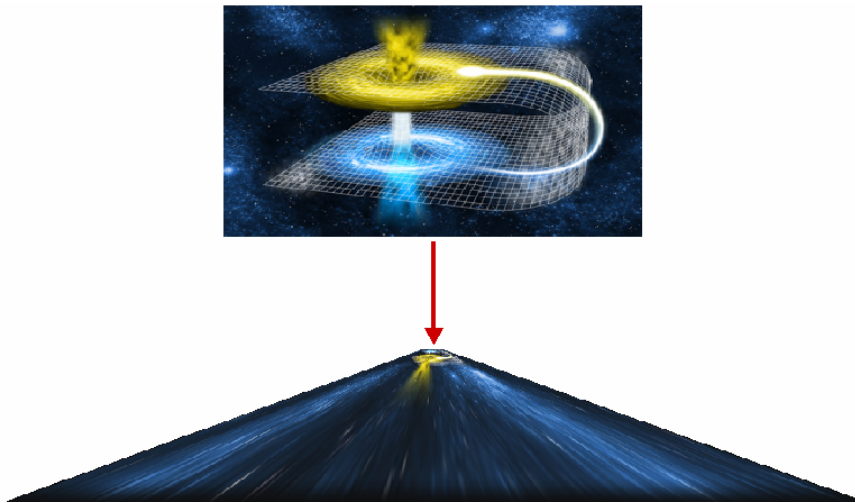
Η οικογένεια αυτή είναι η πιο γενική κατηγορία γραμμικών μετασχηματισμών. Χαρακτηρίζεται από οκτώ βαθμούς ελευθερίας.

$$\mathbf{T}_{hom} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & 1 \end{bmatrix}. \quad (3.14)$$

Η εφαρμογή ενός προβολικού μετασχηματισμού πρέπει να γίνει στις ομογενείς συντεταγμένες των σημείων και στη συνέχεια απαιτείται επανομογενοποίηση των νέων συντεταγμένων. Δηλαδή επαναπροβολή των σημείων στο επίπεδο, το οποίο επιτυγχάνεται με τη διαίρεση των νέων συντεταγμένων με τη συνιστώσα βάθους.

$$\mathbf{v} = \begin{bmatrix} x \\ y \end{bmatrix} \stackrel{\hat{\mathbf{v}}_{\text{homogenous}}}{=} \xrightarrow{\mathcal{T}_h(\cdot)} \begin{bmatrix} \hat{x} \\ \hat{y} \\ \hat{z} \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \triangleq \hat{\mathbf{v}} = \begin{bmatrix} \hat{x} \\ \hat{y} \\ \hat{z} \end{bmatrix}. \quad (3.15)$$

Ενας μετασχηματισμός ομογραφίας μπορεί να γίνει κατανοητός ως ένας μετασχηματισμός συγγένειας σε συνδυασμό με μια ολίσθηση στον 3Δ χώρο. Η επαναπροβολή των σημείων στο επίπεδο ευθύνεται για τις επιπλέον "προβολικές" παραμορφώσεις.



Σχήμα 3.5: Παράδειγμα προβολικού μετασχηματισμού

**Αμετάβλητες ιδιότητες.** Η εφαρμογή προβολικών μετασχηματισμών σε ένα σύνολο σημείων αφήνει αναλλοίωτους τους προσανατολισμούς των ευθειών. Ειδικότερα, συνευθειακά σημεία παραμένουν συνευθειακά, επίσης τεμνόμενες όπως και παράλληλες ευθείες διατηρούν την μορφολογία τους δηλαδή τις μεταξύ τους γωνίες και το πλήθος σημείων τομής. Επιπλέον οι μετασχηματισμοί αυτού ÷ του είδους διατηρούν ιδιομορφίες της εικόνας όπως σημεία κάμψης (points of inflection), επαπτομενικές ασυνέχειες (tangent discontinuities) και λόγους των λόγων

μήκων ευθύγραμμων τμημάτων (cross ratio)<sup>10</sup>.

### 3.1.3 Οικογένιες 3D γεωμετρικών μετασχηματισμών

Η κατηγορίες μετασχηματισμών στο 3D είναι οι ίδιες με αυτές του επιπέδου, απλώς στην περίπτωση τρισδιάστατων σημείων η φυσική υπόσταση των ομογενών συντεταγμένων είναι πιο δυσνόητη. Στην περίπτωση των δύο διαστάσεων μπορούσαμε εύκολα να φανταστούμε ότι η εφαρμογή του μετασχηματισμού γίνεται στον τρισδιάστατο χώρο (ομογενείς συντεταγμένες) και ότι τελικά το αποτέλεσμα του επαναπροβάλλεται στις δυο διαστάσεις (διαίρεση των συντεταγμένων με την συνιστώσα βάθους). Αφαιρετικά θα μπορούσαμε να πούμε ότι και στις τρεις διαστάσεις η παραμόρφωση επιτυγχάνεται σε μια τέταρτη διάσταση και το αποτέλεσμα της επαναπροβάλλεται στον 3D Ευκλείδιο χώρο.

Στη συνέχεια δίνονται τα μετασχηματιστικά μητρώα των προαναφερθέντων οικογενειών τροποποιημένα για τρισδιάστατα σημεία σε ομογενείς συντεταγμένες καθώς και ένα παράδειγμα εφαρμογής τους.

#### I. 3D Ευκλείδιος Μετασχηματισμός

$$\mathbf{T}_{3\Delta \text{ euclidean}} = \begin{bmatrix} \mathbf{R}_{3\Delta} & \mathbf{t}_{3\Delta} \\ \mathbf{0}_{3\Delta}^T & 1 \end{bmatrix} \quad (3.16)$$

#### II. 3D Μετασχηματισμός ομοιότητας

$$\mathbf{T}_{3\Delta \text{ similarity}} = \begin{bmatrix} s\mathbf{R}_{3\Delta} & \mathbf{t}_{3\Delta} \\ \mathbf{0}_{3\Delta}^T & 1 \end{bmatrix} \quad (3.17)$$

#### III. 3D Μετασχηματισμός συγγένειας

$$\mathbf{T}_{3\Delta \text{ affinity}} = \begin{bmatrix} \mathbf{A}_{3\Delta} & \mathbf{t}_{3\Delta} \\ \mathbf{0}_{3\Delta}^T & 1 \end{bmatrix} \quad (3.18)$$

#### IV. 3D Μετασχηματισμός ομογραφίας

$$\mathbf{T}_{3\Delta \text{ homography}} = \begin{bmatrix} \mathbf{A}_{3\Delta} & \mathbf{t}_{3\Delta} \\ \mathbf{v}_{3\Delta}^T & \alpha \end{bmatrix}, \quad (3.19)$$

όπου το  $\mathbf{R}_{3\Delta}$  συμβολίζει ένα τρισδιάστατο μητρώο περιστροφής, το  $\mathbf{t}_{3\Delta} = [t_x t_y t_z]^T$  ένα  $3 \times 1$  διάνυσμα ολίσθησης, το  $\mathbf{A}_{3\Delta}$  ένα  $3 \times 3$  αντιστρέψιμο μητρώο, το  $\mathbf{v}_{3\Delta}$  ένα τυχαίο  $3 \times 1$  διάνυσμα και τέλος  $\alpha$  είναι μια βαθμωτή ποσότητα.

**Αμετάβλητες ιδιότητες.** Οι ευκλείδειοι μετασχηματισμοί (I) διατηρούν τους όγκους, οι μετασχηματισμοί ομοιότητας (II) διατηρούν την απολυτή κωνική  $\Omega_\infty$ .<sup>11</sup> Οι μετασχηματισμοί συγγένειας (III) της παραλληλίας μεταξύ επιπέδων, τους λόγους των όγκων καθώς και το επίπεδο στο άπειρο (plane at infinity)  $\pi_\infty$ . Τέλος, οι μετασχηματισμοί ομογραφίας (IV) διατηρούν τις τόμες επιφανειών καθώς και την καμπυλότητα κατά gauss<sup>12</sup>.

<sup>10</sup> Ο λόγος των  $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4$  ορίζεται ως  $\frac{|\mathbf{v}_1 \mathbf{v}_2| |\mathbf{v}_3 \mathbf{v}_4|}{|\mathbf{v}_1 \mathbf{v}_3| |\mathbf{v}_2 \mathbf{v}_4|}$ , όπου  $|\mathbf{v}_i \mathbf{v}_j| = \det \begin{bmatrix} v_{ix} & v_{jx} \\ v_{iy} & v_{jy} \end{bmatrix}$ .

<sup>11</sup> The absolute conic  $\Omega_\infty$  is a point (conic) on the plane at infinity  $\pi_\infty$  which remains unchanged under similarity transformations.

<sup>12</sup> Σχετίζεται με τις Gaussian Images.

<sup>14</sup> Ο χρωματισμός των σημείων είναι τεχνητός καθώς δεν έχουμε πληροφορία χρώματος παρά μόνο συντεταγμένων. Επίσης εξαιτίας του μικρού πλήθους των σημείων παρατηρούνται λευκά-κενά στο νέφος.



Σχήμα 3.6: Παράδειγμα εφαρμογής Ευκλείδειου μετασχηματισμού σε τρισδιάστατο νέφος σημείων<sup>14</sup>



Σχήμα 3.7: Παράδειγμα ισομετρίας με ισοτροπική κλιμάκωση (similarity transform)



Σχήμα 3.8: Παράδειγμα 3D μετασχηματισμού συγγένειας



Σχήμα 3.9: Παράδειγμα 3D προβολικού μετασχηματισμού

## 3.2 Στοιχεία Ανάλυσης Μητρώων

Στην υποενότητα αυτή παρουσιάζονται και αποδεικνύονται λήμματα της Γραμμικής Άλγεβρας που χρησιμοποιήθηκαν σε κάποιες από τις μεθόδους αντιστοίχισης. Για μια εκτεταμένη έρευνα σε τύπους και θεωρήματα της γραμμικής άλγεβρας προτείνουμε στον αναγνώστη να ανατρέξει στο *The Matrix Cookbook*.

### 3.2.1 Πολική Διάσπαση και το Λήμμα Μέγιστου Ίχνους

Η πολική διάσπαση (**Polar Decomposition**) ενός μητρώου είναι μια παραγοντοποίηση ανάλογη με την πολική αναπαράσταση βαθμωτού μιγαδικού αριθμού, ως  $z = re^{j\theta}$  όπου  $r$  είναι το μέτρο και  $e^{j\theta}$  το μιγαδικό πρόσημο (**complex sign**) του  $z$ . Στην περίπτωση τετραγωνικού μιγαδικού μητρώου  $\mathbf{A}$  η πολική διάσπαση έχει τη μορφή  $\mathbf{A} = \mathbf{Q}\mathbf{P}$ , όπου το  $\mathbf{Q}$  αναπαριστά ένα ορθομοναδιαίο (**Unitary**) και το  $\mathbf{P}$  ένα Ερμιτιανό συμμετρικά θετικά ημιορισμένο (**Hermitian Positive semidefinite**) μητρώο αντίστοιχα. Διαισθητικά η πολική διάσπαση διαχωρίζει το μητρώο  $\mathbf{A}$  σε μια συνιστώσα (component)  $\mathbf{P}$  η οποία τανύει το χώρο που δημιουργεί ένα σύνολο ορθογωνίων αξόνων και μία συνιστώσα περιστροφής  $\mathbf{Q}$ . Η διάσπαση αυτή υπάρχει πάντα, και επιπλέον στην περίπτωση που το  $\mathbf{A}$  είναι αντιστρέψιμο, το  $\mathbf{P}$  είναι θετικά ορισμένο. Η πολική διάσπαση μπορεί να υπολογιστεί χρησιμοποιώντας την διάσπαση ιδιαιζουσών τιμών του  $\mathbf{A}$  ως  $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$  με  $\mathbf{P}, \mathbf{Q}$  τα:

$$\mathbf{P} = \mathbf{V}\mathbf{\Sigma}\mathbf{V}^H, \quad \mathbf{Q} = \mathbf{U}\mathbf{V}^H. \quad (3.20)$$

Είναι εύκολο να δούμε ότι ισχύει  $\mathbf{Q} = \mathbf{A}\mathbf{P}^{-1}$ .

**Λήμμα:** Έστω μητρώο  $\mathbf{C} \in \mathbb{R}^{n \times n}$ , το βέλτιστο ορθογώνιο μητρώο  $\mathbf{R}$  που μεγιστοποιεί το ίχνος του γινομένου  $\mathbf{R}\mathbf{C}$  ισούται με την ανάστροφη ορθογώνια συνιστώσα  $\mathbf{Q}^T$  που προκύπτει από την πολική διάσπαση του  $\mathbf{C}$  ως  $\mathbf{C} = \mathbf{Q}\mathbf{P}$ .

Απόδειξη: Αρχεί να δείξουμε ότι:

$$\text{tr}(\mathbf{R}\mathbf{C}) = \text{tr}(\mathbf{R}\mathbf{Q}\mathbf{P}) \leq \text{tr}(\mathbf{P}). \quad (3.21)$$

Θέτουμε  $\mathbf{B} = \mathbf{RQ}$  και επομένως πετυχαίνουμε την ισότητα με  $\mathbf{RQ} = \mathbf{I}$  και άρα  $\mathbf{R} = \mathbf{Q}^T$ , επίσης αναλύουμε το μητρώο  $\mathbf{P}$  ως:

$$\mathbf{P} = \mathbf{V}\mathbf{\Sigma}\mathbf{V}^T = (\mathbf{V}\mathbf{\Sigma}^{\frac{1}{2}})(\mathbf{\Sigma}^{\frac{1}{2}}\mathbf{V}^T) = \mathbf{\Xi}\mathbf{\Xi}^T = \sum_{i=1}^n \xi_i \xi_i^T, \quad (3.22)$$

όπου  $\xi_i$  είναι η  $i$ -οστή στήλη του μητρώου  $\mathbf{\Xi}$ .

$$\text{tr}(\mathbf{RQP}) = \text{tr}(\mathbf{B}\mathbf{\Xi}\mathbf{\Xi}^T) = \text{tr}(\mathbf{\Xi}^T\mathbf{B}\mathbf{\Xi}) = \sum_{i=1}^n \xi_i^T \mathbf{B} \xi_i. \quad (3.23)$$

Στη συνέχεια επιλύουμε την  $\max \sum_{i=1}^n \xi_i^T \mathbf{B} \xi_i$ , με συνθήκη  $\mathbf{B}\mathbf{B}^T = \mathbf{I}$ , αφού το  $\mathbf{B}$  είναι ορθογώνιο ως γινόμενο ορθογωνίων. Απο την ανισότητα Cauchy-Shwarz έχουμε ότι το εσωτερικό γινόμενο δύο διανυσμάτων είναι πάντα μικρότερο ή ίσο απο τη ρίζα του γινομένου των μέτρων τους:

$$\xi_i^T \mathbf{B} \xi_i \leq \sqrt{(\xi_i^T \xi_i)(\xi_i^T \mathbf{B}^T \mathbf{B} \xi_i)} = \xi_i^T \xi_i \quad \forall i, \quad (3.24)$$

και επομένως συμπεραίνουμε ότι  $\text{tr}(\mathbf{RQP}) \leq \text{tr}(\mathbf{P})$ .

### 3.3 Στοιχεία Θεωρίας Βελτιστοποίησης

Ός πεδίο των μαθηματικών, η βελτιστοποίηση ([optimization - mathematical programming](#)) ασχολείται με προβλήματα εύρεσης του ‘καλύτερου’ στοιχείου που ικανοποιεί κάποια δεδομένα κριτήρια. Στην απλούστερη μορφή του, ένα πρόβλημα βελτιστοποίησης περιγράφει τη μεγιστοποίηση ή ελαχιστοποίηση μίας πραγματικής συνάρτησης. Συνήθως η προς βελτιστοποίηση συνάρτηση καλείται αντικειμενική ([objective function](#)) και ειδικότερα συνάρτηση κόστους ή απώλειας ([cost or loss function](#)) όταν πρόκειται για πρόβλημα ελαχιστοποίησης ή συνάρτηση οφέλους ([utility function](#)) όταν πρόκειται για πρόβλημα μεγιστοποίησης. Συνήθως χρησιμοποιούνται οι παρακάτω συμβολισμοί

$$\min_{x \in \mathfrak{A}} f(x) \quad \text{ή} \quad \max_{x \in \mathfrak{A}} f(x),$$

δηλαδή ‘Υπολόγισε το ελάχιστο/μέγιστο της  $f$  όταν το  $x \in \mathfrak{A}$ ’, και οι ακόλουθοι ισοδύναμοι

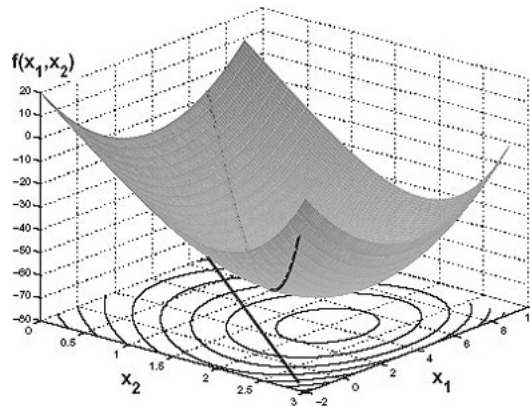
$$\arg \min_{x \in \mathfrak{A}} f(x) \quad \text{ή} \quad \arg \max_{x \in \mathfrak{A}} f(x),$$

δηλαδή ‘Βρες την τιμή του ορίσματος  $x$  η οποία ελαχιστοποιεί/μεγιστοποιεί την  $f$ . Όπως θα δούμε στη συνέχεια, το πρόβλημα της αντιστοίχισης νεφών σημείων τελικά καταλήγει σε ένα πρόβλημα βελτιστοποίησης-εύρεσης του ‘καλύτερου’ ορίσματος μιας αντικειμενικής συνάρτησης.

#### 3.3.1 Βελτιστοποίηση κυρτών συναρτήσεων

Ένας απο τους κυριότερους κλάδους της θεωρίας βελτιστοποίησης αφορά την ελαχιστοποίηση κυρτών συναρτήσεων. ([Convex Minimization](#))<sup>15</sup>

<sup>15</sup>Κάθε πρόβλημα μεγιστοποίησης μιας κοίλης (concave) συνάρτησης  $f$  μπορεί να επανατέθει σαν ένα πρόβλημα ελαχιστοποίησης της  $-f$  η οποία είναι κυρτή.



Σχήμα 3.10: Παράδειγμα 2Δ κυρτής συνάρτησης με μοναδικό ολικό ελάχιστο.

Κυρτές συναρτήσεις ονομάζονται αυτές που έχουν μοναδικό ολικό ελάχιστο. Η ιδιότητα που κάνει αυτού του τύπου προβλήματα ευκολότερα σε σχέση με τα υπόλοιπα προβλήματα βελτιστοποίησης είναι ότι, ακόμα και να μην υπάρχει κλειστός τύπος για την εύρεση του ζητούμενου ακροτάτου, οποιαδήποτε επαναληπτική μέθοδος θα μπορούσε να προσεγγίσει το ακρότατο. Το πρόβλημα που προκύπτει συνήθως είναι ότι στην πράξη οι αντικειμενικές συναρτήσεις εκτός του ότι είναι μη γραμμικές, οπότε συνήθως δεν επιδέχονται κλειστού τυπου λύση, είναι τμηματικά κυρτές (*locally convex or quasiconvex*). Το οποίο χαρακτηριστικό κάνει τις, καθώς αρχικοποιημένες, επαναληπτικές μεθόδους να εγκλωβίζονται σε κάποιο από τα τοπικά ελάχιστα με αποτέλεσμα να αποτυγχάνουν να βρουν το ολικό ελάχιστο.

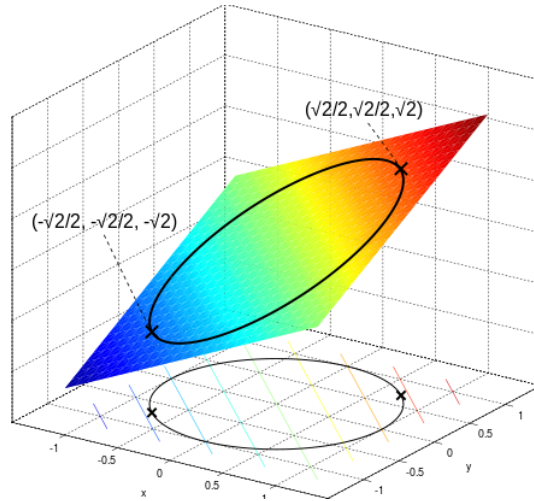
### 3.3.2 Πολλαπλασιαστές Lagrange

Η μέθοδος των πολλαπλασιαστών Lagrange είναι μια στρατηγική για την εύρεση τοπικών ακροτάτων συναρτήσεων υπό την ύπαρξη επιπλέον περιορισμών ισότητας (*equality constraints*). Η βασική αρχή της στρατηγικής αυτής είναι ότι εάν ένα υποψήφιο ακρότατο (candidate extremum) δεν ικανοποιεί τους επιπλέον περιορισμούς τότε θα πληρώσει ένα επιπλέον ‘πρόστιμο’ με τη μορφή μιας ποσότητας που προστίθεται (*at a minimization*) ή αφαιρείται (*at a maximization*) από την υπό βελτιστοποίηση συνάρτηση κόστους ή οφέλους. Για παράδειγμα εάν επιθυμούμε να βρούμε το μέγιστο μιας συνάρτησης  $f(x)$  το οποίο πρέπει να υπακούει στον περιορισμό  $g(x) = c$  τότε η προς μεγιστοποίηση συνάρτηση θα πρέπει να τροποποιηθεί ως  $\Lambda(x, \lambda) = f(x) - \lambda(g(x) - c)$ , όπου το  $\lambda$  είναι μια παράμετρος (*i.e., the multiplier*) η οποία καλείται πολλαπλασιαστής Lagrange και η οποία ρυθμίζει την τιμή της  $\Lambda(x, \lambda)$  έτσι ώστε όσο περισσότερο απέχει η τιμή ενός ακροτάτου  $x_0$  από τον περιορισμό  $f(x_0) = c$ , τόσο μικρότερη να γίνεται η τιμή της επαυξημένης συνάρτησης  $\Lambda(x, \lambda)$ . Αυτό που απέδειξε ο Lagrange είναι ότι υπάρχει πάντα (τουλάχιστον ένα)  $\lambda_0$  τέτοιο ώστε το σημείο  $(x_0, \lambda_0)$  να είναι σημείο στασιμότητας της συνάρτησης  $\Lambda(x, \lambda)$  το οποίο βάσει του θεωρήματος του Fermat<sup>16</sup> μεταφράζεται στο ότι  $\nabla \Lambda(x_0, \lambda_0) = \vec{0}$ . Να σημειώσουμε εδώ ότι η μέθοδος των πολλαπλασιαστών Lagrange δίνει μια αναγκαία αλλά όχι ικανή συνθήκη για την εύρεση των ζητούμενων ακροτάτων, με την έννοια ότι κάποια από τα στάσιμα σημεία που θα προκύψουν από την λύση της  $\nabla \Lambda(x, \lambda) = \vec{0}$  μπορεί να μην παρέχουν

<sup>16</sup>Not the last one!



λύση για το αρχικό πρόβλημα. Ακολουθεί ένα απλό παράδειγμα της μεθόδου των πολλαπλασιαστών Lagrange:



Σχήμα 3.11: Οπτικοποίηση της αντικειμενικής συνάρτησης και του περιορισμού.

Ας υποθέσουμε ότι θέλουμε να μεγιστοποιήσουμε την  $f(x, y) = x + y$  υπο τον περιορισμό  $x^2 + y^2 = 1$ , ο οποίος υποχρεώνει τη λύση να ανήκει στο μοναδιαίο κύκλο (feasible set). Αρχικά περιγράφουμε την ικανοποίηση του περιορισμού ως μηδενισμό της συνάρτησης  $g(x, y) - c = x^2 + y^2 - 1$  και στην συνέχεια κατασκευάζουμε την Lagrangian ως:

$$\Lambda(x, y, \lambda) = f(x, y) + \lambda(g(x, y) - c) = x + y + \lambda(x^2 + y^2 - 1) \quad (3.25)$$

Υπολογίζοντας τις μερικές παραγώγους  $\nabla \Lambda(x, \lambda) = \vec{0}$  έχουμε:

$$\frac{\partial \Lambda(x, \lambda)}{\partial x} = 1 + 2\lambda x = 0 \Rightarrow x = \frac{-1}{2\lambda}, \quad (3.26)$$

$$\frac{\partial \Lambda(x, \lambda)}{\partial y} = 1 + 2\lambda y = 0 \Rightarrow y = \frac{-1}{2\lambda}, \quad (3.27)$$

$$\frac{\partial \Lambda(x, \lambda)}{\partial \lambda} = x^2 + y^2 - 1 = 0. \quad (3.28)$$

Επιλύοντας τώρα την (3.28) με χρήση των (3.26) και (3.27) έχουμε:

$$\frac{1}{4\lambda^2} + \frac{1}{4\lambda^2} = 1 \Rightarrow \lambda = \mp \frac{1}{\sqrt{2}}, \quad (3.29)$$

απο τις οποίες βρίσκουμε:

$$f\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) = \sqrt{2}, \quad f\left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right) = -\sqrt{2}, \quad (3.30)$$

και επομένως καταλήγουμε ότι το ζητούμενο μέγιστο επιτυγχάνεται στο σημείο

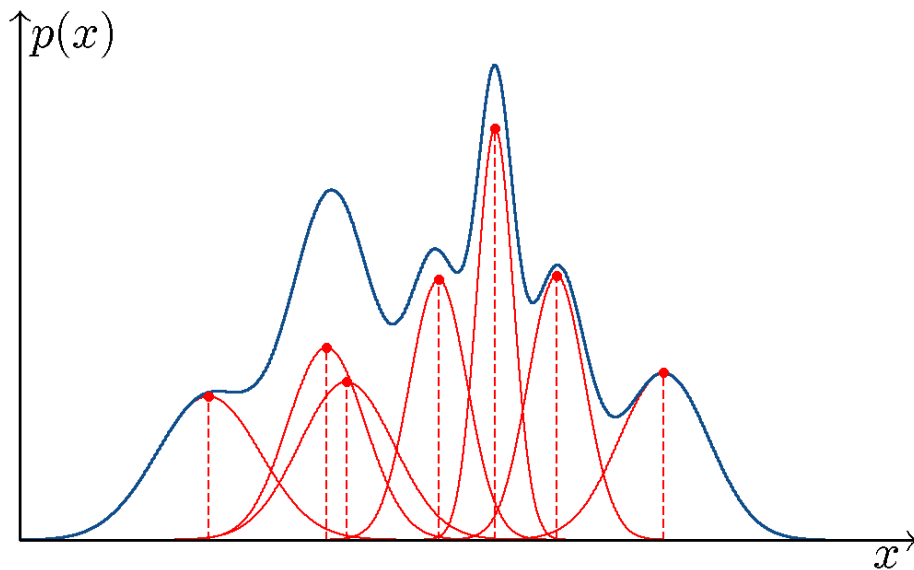
$$(x_0, y_0) = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right). \quad (3.31)$$



### 3.4 Μοντέλα βασισμένα σε μίγματα Γκαουσιανών κατανομών

Σε αυτό το σημείο θα παρουσιάσουμε δύο μαθηματικά εργαλεία από το πεδίο της στατιστικής. Συγκεκριμένα τα μοντέλα που βασίζονται σε μίγματα κανονικών κατανομών (**Gaussian mixture models, GMM**) και τον αλγόριθμο μεγιστοποίησης αναμενόμενης πιθανοφάνειας (**Expectation Maximization algorithm, EM**). Σε προβλήματα εκτίμησης παραμέτρων ενός στατιστικού μοντέλου στα οποία υπεισέρχονται λανθάνουσες μεταβλητές (**latent variables**), τα δύο αυτά εργαλεία, συνδυαζόμενα, δίνουν μια πολύ ισχυρή μέθοδο επαναληπτικής βελτιστοποίησης αυτών των παραμέτρων, την οποία θα την χρησιμοποιήσουμε για την μοντελοποίηση του προβλήματος αντιστοίχισης ζευγούς στην [υποενότητα 4.2.3] και επίσης στην μέθοδο που θα αναπτύξουμε στο [κεφάλαιο 6] η οποία αποτελεί επέκταση της προηγούμενης. Σκοπός της παρούσας υποενότητας είναι να κάνει μια γενική παρουσίαση των εργαλείων καθώς και να δώσει την απόδειξη σύγκλισης του αλγορίθμου μεγιστοποίησης αναμενόμενης πιθανοφάνειας στην περίπτωση αγνώστων μεταβλητών.

#### 3.4.1 Μείγματα Κανονικών Κατανομών



Σχήμα 3.12: Οπτικοποίηση της συνάρτησης πυκνότητας πιθανότητας (—) ενός GMM η οποία προκύπτει από την υπέρθεση επτά ισοβαρών συνιστωσών (—).

Τα GMM είναι υπερθέσεις gaussian κατανομών διαφορετικών παραμέτρων, (Σχήμα 3.12), και αποτελούν ένα τρόπο μοντελοποίησης αυθαίρετων κατανομών πυκνότητας πιθανότητας με την βοήθεια κατανομών γνωστών οικογενειών. Υπάρχουν πολλά είδη μοντέλων ανάμειξης κατανομών. Στην παρούσα διπλωματική ενδιαφερόμαστε για τα πεπερασμένα GMM διότι βασίζονται σε gaussian κατανομές οι οποίες κατά βάθος εξαρτώνται από το τετράγωνο μιας Mahalanobis απόστασης η οποία θα αποδειχθεί πολύ χρήσιμη στη συνέχεια. Υποθέτουμε ότι μπορούμε να

δημιουργήσουμε ένα GMM από υπέρθεση απείρου πλήθους κανονικών κατανομών τότε αυτό το GMM θα μπορούσε να περιγράψει οποιαδήποτε αυθαίρετη κατανομή. Εδώ θα αρκестούμε σε GMM με πεπερασμένες συνιστώσες. Φορμαλιστικά ένα GMM με  $n$  συνιστώσες ορίζει την κατανομή  $p(x)$  ενός ενδεχομένου  $x$  ως

$$p(x) = \sum_{i=1}^n \pi_i \mathcal{N}(x|\mu_i, \sigma_i), \quad (3.32)$$

όπου  $\pi_i$  είναι το βάρος της αντίστοιχης συνιστώσας  $\mathcal{N}(x|\mu_i, \sigma_i)$ . Εφόσον το παρακάτω άθροισμα πρέπει να περιγράφει συνάρτηση πυκνότητας πιθανότητας τα βάρη πρέπει να δημιουργούν έναν κυρτό συνδιασμό (convex combination) των συνιστωσών, με άλλα λόγια

$$\sum_{i=1}^n \pi_i = 1. \quad (3.33)$$

Σε αυτό το σημείο εισάγουμε μια  $n$ -διάστατη δυαδική τυχαία μεταβλητή  $\mathbf{z}$ , σε κάθε υλοποίηση της οποίας ένα μοναδικό στοιχείο  $Z_i$  είναι μονάδα ενώ τα υπόλοιπα είναι μηδέν. Υπο αυτή τη θεώρηση οι τιμές των  $Z_i \in \{0, 1\}$ . Παρατηρούμε επίσης ότι  $\sum_i Z_i = 1$  και επίσης ότι υπάρχουν  $n$  δυνατές καταστάσεις του τυχαίου διάνυσματος  $\mathbf{z}$  ανάλογα με το ποιο στοιχείο του είναι μονάδα. Έχοντας ορίσει το τυχαίο διάνυσμα  $\mathbf{z}$  μπορούμε να κατασκευάσουμε την από κοινού συνάρτηση πυκνότητας πιθανότητας  $p(x, \mathbf{z})$  ως προς την περιθώρια κατανομή  $p(\mathbf{z})$  και μία κατανομή πιθανοφάνειας  $p(x|\mathbf{z})$  ως. Για να το πετύχουμε αυτό αρχικά ορίζουμε την  $p(\mathbf{z})$  μέσω των βαρών  $\pi_i$  ως  $p(Z_i = 1) = \pi_i$ . Επίσης λόγω της ειδικής δομής της  $\mathbf{z}$  μπορούμε να γράψουμε ότι

$$p(\mathbf{z}) = \prod_{i=1}^n \pi_i^{Z_i}, \quad (3.34)$$

και να εξάγουμε ότι η πιθανοφάνεια ενός ενδεχομένου  $x$  (δεδομένης της μεταβλητής  $\mathbf{z}$ ) ισούται με

$$p(x|Z_i = 1) = \mathcal{N}(x|\mu_i, \sigma_i) = \prod_{i=1}^n \mathcal{N}(x|\mu_i, \sigma_i)^{Z_i}. \quad (3.35)$$

Βάσει των σχέσεων (3.34) και (3.35) μπορούμε να ορίσουμε την περιθώρια κατανομή ενός ενδεχομένου  $x$  ως

$$p(x) = \sum_{\mathbf{z}} p(x, \mathbf{z}) = \sum_{\mathbf{z}} p(\mathbf{z})p(x|\mathbf{z}) = \sum_{i=1}^n \pi_i \mathcal{N}(x|\mu_i, \sigma_i), \quad (3.36)$$

και τέλος να κατασκευάσουμε την εκ των υστέρων πιθανότητα μιας μεταβλητής  $\mathbf{z}$  δεδομένου του ενδεχομένου  $x$ , την οποία θα συμβολίζουμε  $a_{i|x}$ , ως

$$a_{i|x} = p(Z_i = 1|x) = \frac{p(Z_i = 1)p(x|Z_i = 1)}{p(x)} = \frac{\pi_i \mathcal{N}(x|\mu_i, \sigma_i)}{\sum_{k=1}^n \pi_k \mathcal{N}(x|\mu_k, \sigma_k)}. \quad (3.37)$$

### 3.4.2 Η σύγκλιση του EM υπο την ύπαρξη αγνώστων μεταβλητών

Ο αλγόριθμος μεγιστοποίησης πιθανοφάνειας (Expectation Maximization, EM) προτάθηκε από τους Dempster et.al<sup>[21]</sup> το 1977 και χρησιμοποιείται για την εκτίμηση των βέλτιστων με την έννοια της μέγιστης πιθανοφάνειας, παραμέτρων

στοχαστικών μοντέλων<sup>17</sup> υπο την ύπαρξη αγνώστων μεταβλητών (latent variables).

Ας υποθέσουμε ένα στοχαστικό μοντέλο που περιγραφεται απο ένα σύνολο παραμέτρων  $\theta$  για το οποίο έχουμε ένα σύνολο παρατηρήσεων  $\mathcal{Y}$  καθώς και ένα σύνολο αγνώστων λανθανόντων μεταβλητών  $\mathcal{Z}$ . Στόχος του EM είναι να μεγιστοποιήσει την συνάρτηση πιθανοφάνειας

$$p(\mathcal{Y}; \theta) = \sum_{\mathcal{Z}} p(\mathcal{Y}, \mathcal{Z}; \theta). \quad (3.38)$$

Χωρίς απώλεια της γενικότητας κάνουμε την υπόθεση ότι η  $\mathcal{Z}$  είναι διακριτή τυχαία μεταβλητή· σε περιπτώσεις συνεχών τ.μ. μπορούμε απλώς να αντικαταστήσουμε την άθροιση με ολοκλήρωση.

Θα θεωρήσουμε επίσης ότι η απευθείας μεγιστοποίηση της πιθανοφάνειας παρατηρήσεων  $p(\mathcal{Y}; \theta)$  είναι δυσεπίλυτη σε αντίθεση με την μεγιστοποίηση της πιθανοφάνειας πλήρων δεδομένων  $p(\mathcal{Y}, \mathcal{Z}; \theta)$  η οποία τυγχάνει ευκολότερη. Εισάγοντας σε αυτό το σημείο μια κατανομή  $q(\mathcal{Z})$  η οποία περιγράφει τις εκ των προτέρων πιθανότητες της τυχαίας μεταβλητής  $\mathcal{Z}$  και εφαρμόζοντας τον κανόνα του Bayes

$$\frac{p(\mathcal{Z}, \mathcal{Y}; \theta)}{p(\mathcal{Z}|\mathcal{Y}; \theta)p(\mathcal{Y}; \theta)} = 1 \Rightarrow \mathbb{E}_{\mathcal{Z}} \left[ \log \frac{\frac{p(\mathcal{Z}, \mathcal{Y}; \theta)}{q(\mathcal{Z})}}{\frac{p(\mathcal{Z}|\mathcal{Y}; \theta)p(\mathcal{Y}; \theta)}{q(\mathcal{Z})}} \right] = \mathbb{E}_{\mathcal{Z}} [\log 1]. \quad (3.39)$$

Επικαλούμενοι το "νόμο του ασυνείδητου στατιστικού" (The law of unconscious statistician)<sup>18</sup>, την γραμμικότητα του τελεστή αναμενόμενης τιμής καθώς και την ιδιότητα του λογαρίθμου απο τη Σχέση (3.39) προκύπτει

$$\sum_{\mathcal{Z}} q(\mathcal{Z}) \log \frac{p(\mathcal{Z}, \mathcal{Y}; \theta)}{q(\mathcal{Z})} - \sum_{\mathcal{Z}} q(\mathcal{Z}) \log \frac{p(\mathcal{Z}|\mathcal{Y}; \theta)}{q(\mathcal{Z})} - \sum_{\mathcal{Z}} q(\mathcal{Z}) \log p(\mathcal{Y}; \theta) = 0. \quad (3.40)$$

Παρατηρούμε ότι ο πρώτος όρος του αριστερού μέλους της ισότητας ορίζει την αναμενόμενη λογαριθμική πιθανοφάνεια πλήρων δεδομένων την οποία θα συμβολίζουμε  $\mathcal{E}_{\mathcal{Z}}(\theta|\mathcal{Y}, \mathcal{Z})$  και η οποία θα αποτελέσει την ποσότητα της οποίας την μεγιστοποίηση, τελικά, θα επιλύσουμε. Ο δεύτερος όρος του αθροίσματος αποτελεί την απόσταση Kullback Leibler,  $\text{KL}(q \parallel p)$ , μεταξύ των κατανομών  $q$  και  $p$ . Συνεπυγμένα, η σχέση (3.40) γράφεται

$$\begin{aligned} \mathcal{E}_{\mathcal{Z}}(\theta|\mathcal{Y}, \mathcal{Z}) + \text{KL}(q \parallel p) - \sum_{\mathcal{Z}} q(\mathcal{Z}) \log p(\mathcal{Y}; \theta) &= 0 \Rightarrow \\ \mathcal{E}_{\mathcal{Z}}(\theta|\mathcal{Y}, \mathcal{Z}) + \text{KL}(q \parallel p) - \log p(\mathcal{Y}; \theta) \sum_{\mathcal{Z}} q(\mathcal{Z}) &= 0 \Rightarrow \\ \log p(\mathcal{Y}; \theta) &= \mathcal{E}_{\mathcal{Z}}(\theta|\mathcal{Y}, \mathcal{Z}) + \text{KL}(q \parallel p). \end{aligned} \quad (3.41)$$

Παρατηρώντας την σχέση (3.41) και γνωρίζοντας ότι η KL divergence είναι θετική ποσότητα συμπεραίνουμε ότι η  $\mathcal{E}_{\mathcal{Z}}(\theta|\mathcal{Y}, \mathcal{Z})$  αποτελεί ένα κάτω φράγμα για την

<sup>17</sup>Όπως για παράδειγμα τα Gaussian Mixture Models, [Ενότητα 3.4.1].

<sup>18</sup> Ο law of unconscious statistician δίνει ένα τρόπο υπολογισμού του Expectation μιας συνάρτησης  $f$  η οποία δρά σε μια τυχαία μεταβλητή  $\mathcal{Z}$  ως  $\mathbb{E}[f(\mathcal{Z})] = \sum_{\mathcal{Z}} q(\mathcal{Z})f(\mathcal{Z})$ , όπου  $q(\mathcal{Z})$  είναι η συνάρτηση μάζας πιθανότητας της τυχαίας μεταβλητής  $\mathcal{Z}$ . Αντίστοιχη Σχέση ισχύει και για συνεχής τυχαίες μεταβλητές.

$\log p(\mathcal{Y}; \theta)$  απο το οποίο καταλήγουμε οτι μεγιστοποιώντας την αναμενόμενη λογαριθμική πιθανοφάνεια πλήρων δεδομένων επιτυγχάνουμε αύξηση της λογαριθμικής πιθανοφάνειας παρατηρήσεων.

Η σχέση (3.41) αποτελεί την βασική αρχή λειτουργίας του EM. Το υποπρόβλημα που δημιουργείται είναι οτι μια μονοβηματική επίλυση της  $\mathcal{E}_{\mathcal{Z}}(\theta|\mathcal{Y}, \mathcal{Z})$  ως προς τις παραμέτρους  $\theta$  δεν θα έδινε αυτές (τις  $\theta$ ) που μεγιστοποιούν την  $\log p(\mathcal{Y}; \theta)$  απλώς θα επιτύγχανε μια αύξηση της δεύτερης το πολύ εως την τρέχουσα τιμή του  $\text{KL}(q \parallel p)$ . Για να επιτευχθεί μεγιστοποίηση της  $\log p(\mathcal{Y}; \theta)$  η διαδικασία χρειάζεται να επαναληφθεί με κάθε επανάληψη να αποτελείται απο δύο βήματα. Στο πρώτο θα υπολογίζονται οι εκ των υστέρων πιθανότητες  $p(\mathcal{Z}|\mathcal{Y}; \theta^{old})$  δεδομένου ενός συνόλου παραμέτρων  $\theta^{old}$  και στη συνέχεια η κατανομή (τιμές) των  $p(\mathcal{Z}|\mathcal{Y}; \theta^{old})$  θα θεωρούνται ως η καινούργια<sup>19</sup> κατανομή  $q(\mathcal{Z})$  και στη συνέχεια οι παράμετροι  $\theta$  επανεκτιμούνται μέσω μεγιστοποίησης της  $\mathcal{E}_{\mathcal{Z}}(\theta|\mathcal{Y}, \mathcal{Z})$  υπο τις τιμές της καινούργιας  $q(\mathcal{Z})$ .

<sup>19</sup>Εξ ορισμού η μεγιστοποίηση της  $\text{KL}(q \parallel p)$  επιτυγχάνεται οταν  $q(\mathcal{Z}) = p(\mathcal{Z}|\mathcal{Y})$ .

## Κεφάλαιο 4

# ΑΝΤΙΣΤΟΙΧΙΣΗ ΖΕΥΓΟΥΣ ΤΡΙΣΔΙΑΣΤΑΤΩΝ ΝΕΦΩΝ ΣΗΜΕΙΩΝ

Έστω δύο σύνολα-νέφη τρισδιάστατων σημείων  $\mathcal{X}$  και  $\mathcal{Y}$ , τα σημεία των οποίων, στην περίπτωση μας, υποθέτουμε ότι έχουν προέλθει από τη δειγματοληψία μιας επιφάνειας και τα οποία θέλουμε να αντιστοιχίσουμε. Δηλαδή να υπολογίσουμε τις παραμέτρους του γεωμετρικού μετασχηματισμού που τα συνδέει.

Τα τελευταία χρόνια έχουν προταθεί ποικίλες τεχνικές για την εκτίμηση αυτών των παραμέτρων. Αναλόγως με το είδος του μετασχηματισμού που θεωρείται μεταξύ των δύο συνόλων υπάρχουν και αντίστοιχες τεχνικές. Υπάρχει ένας μεγάλος αριθμός δημοσιεύσεων επιστημονικών εργασιών πάνω στο συγκεκριμένο πρόβλημα εκ των οποίων οι περισσότερες κάνουν τη θεώρηση ότι τα δύο σύνολα συνδέονται με (ολικό) μετασχηματισμό άκαμπτου σώματος (rigid body transformation). Ελάχιστες λύνουν το πρόβλημα εύκαμπτων (non-rigid) ολικών μετασχηματισμών και ακόμα λιγότερες την περίπτωση πολλαπλών τοπικών μετασχηματισμών (articulated body) ανά υποσύνολο σημείων. Αν και η περίπτωση αφθρωτών σωμάτων δεν είναι σπάνια. Για παράδειγμα αν τα δύο σημειο-σύνολα περιέχουν έναν άνθρωπο σε κίνηση θα ήταν απερίσκεπτο να θεωρήσουμε ολικό μετασχηματισμό μεταξύ τους. Αν και δέν είναι το κυρίως θέμα της παρούσας διπλωματικής στην [υποενότητα 4.3] γίνεται μια ανασκόπηση δημοσιεύσεων πάνω σε μη-συμπαγείς και τοπικούς μετασχηματισμούς.

Στην συγκεκριμένη διπλωματική γίνεται η θεώρηση ότι ο μετασχηματισμός μεταξύ των συνόλων  $\mathcal{X}$ ,  $\mathcal{Y}$  είναι ολικός και Ευκλείδειος. Οι μέθοδοι που αναλύονται στη συνέχεια, εκτός της τελευταίας [υποενότητας 4.3], επιλύουν το πρόβλημα στην περίπτωση ισομετρίας (global Euclidean transformation).

## 4.1 Εκ των προτέρων γνωστές αντιστοιχίες

Σκοπός αυτής της υποενότητας είναι να παρουσιάσει μεθόδους εύρεσης των βέλτιστων παραμέτρων μίας Ευκλείδιας παραμόρφωσης μεταξύ δύο νεφών σημείων υπό τη θεώρηση ότι αντίστοιχα σημεία είναι γνωστά εκ των προτέρων. Αυτή η τοποθέτηση του προβλήματος μοιάζει ιδεατή αλλά όπως θα δούμε στη συνέχεια οι περισσότερες εκ των τεχνικών αντιστοίχισης, αφού βρούνε ένα υποσύνολο αντίστοιχων σημείων μεταξύ των point clouds στη συνέχεια υπολογίζουν τις βέλτιστες παραμέτρους του μεταξύ τους μετασχηματισμού βελτιστοποιώντας κάποια συνάρτηση κόστους-μετρική σφάλματος<sup>1</sup> χρησιμοποιώντας αυτούσια ή τροποποιημένη κάποια από τις τεχνικές που παρουσιάζονται στην ενότητα αυτή.

### 4.1.1 Το πρόβλημα του απόλυτου προσανατολισμού

Ας υποθέσουμε δύο μητρώα  $\mathbf{W}$  και  $\mathbf{X}$ , διαστάσεων  $3 \times n$ , τα οποία περιέχουν σημεία από τα δύο επικαλυπτόμενα νέφη  $\mathcal{X}, \mathcal{Y}$  σε ένα προς ένα αντιστοιχία. Κάθε σημείο του  $\mathbf{W}, \mathbf{X}$  αναπαρίσταται ως  $\mathbf{w}_i, \mathbf{x}_i, 1 \leq i \leq n$  αντίστοιχα και οι συντεταγμένες του βρίσκονται στη  $i$ -στη στήλη του αντίστοιχου μητρώου. Ζητούμενο είναι να βρούμε έναν Ευκλείδιο μετασχηματισμό, αποτελούμενο από ένα μητρώο περιστροφής  $\mathbf{R}$  και ένα διάνυσμα ολίσθησης  $\mathbf{t}$  τέτοιο ώστε "το άθροισμα των τετραγώνων των αποστάσεων μεταξύ των αντιστοιχών σημείων των δύο νεφών να γίνει ελάχιστο" ή πιο φορμαλιστικά:

$$\{\mathbf{R}^*, \mathbf{t}^*\} = \arg \min_{\mathbf{R}, \mathbf{t}} \sum_{i=1}^n \|\mathbf{w}_i - \mathbf{R}\mathbf{x}_i - \mathbf{t}\|_2^2$$

με συνθήκες  $\mathbf{R}^T \mathbf{R} = \mathbf{I}, |\mathbf{R}| = 1,$  (4.1)

όπου  $|\mathbf{R}|$  συμβολίζει την ορίζουσα του μητρώου  $\mathbf{R}$ . Το πρόβλημα αυτό ονομάζεται ορθογώνιο πρόβλημα του Προκρούστη ή πρόβλημα του απόλυτου προσανατολισμού (**Absolute Orientation**). Η πρώτη ονομασία πιθανότατα προέρχεται από τον ομώνυμο μυθικό ληστή της δυτικής Αττικής που τάνυε ή διαμέλιζε τα θύματα του ώστε να χωρέσουν σε ένα σιδερένιο κρεβάτι. Υπό αυτό τον όρο, σήμερα, μοντελοποιούνται προβλήματα στα οποία διαφορετικά Σχήματα (shapes) ή ιδιότητες προσαρμόζονται σε ένα αυθαίρετο πρότυπο (standard).

Για το πρόβλημα του απόλυτου προσανατολισμού έχουν προταθεί ποικίλες λύσεις, οι περισσότερες εκ των οποίων το επιλύουν για μετασχηματισμό ομοιότητας αντί ισομετρίας, δηλαδή προσθέτουν μια επιπλέον παράμετρο ισοτροπικής κλιμάκωσης στο πρόβλημα. Στις ακόλουθες υποενότητες παρουσιάζονται οι τρεις κυριότερες λύσεις για την περίπτωση ισομετρίας (**Euclidean Transformation**). Οι δύο πρώτες [4.1.2], [4.1.3] παρουσιάζονται για ιστορικούς λόγους και λόγους πληρότητας καθώς είναι προγενέστερες της τρίτης [4.1.4]. Η τελευταία μέθοδος είναι αυτή την οποία, ελαφρώς τροποποιημένη, χρησιμοποιούμε και στον αλγόριθμο που θα παρουσιάσουμε στο [κεφάλαιο 6] για αυτό το λόγο αποδεικνύεται αναλυτικά και μάλιστα για την γενική περίπτωση βεβαρυμένων αποστάσεων.

Αξίζει να σημειωθεί εδώ, ότι και οι τρεις μέθοδοι καταλήγουν στην ίδια, ουσιαστικά, λύση απλά μέσω διαφορετικής μοντελοποίησης.

<sup>1</sup>Σύννητως η διαδικασία εύρεσης υποσύνολου αντιστοιχίας και εκτίμησης των παραμέτρων του μετασχηματισμού επαναλαμβάνεται με σκοπό τη βελτίωση της εκτίμησης τους.

### 4.1.2 Επίλυση βασισμένη σε ορθομοναδιαία τετραδόνια

Η μέθοδος που παρουσιάζεται σε αυτή την υποενότητα δόθηκε απο τον B.K.P. Horn<sup>[4]</sup> το 1986 και βασίζεται στην περιγραφή περιστροφών στις τρεις διαστάσεις με χρήση ορθομοναδιαίων τετραδονίων (unit quaternions)<sup>2</sup>. Για να δώσουμε μία ιδέα στον αναγνώστη, αναφέρουμε οτι τα τετραδόνια είναι μια επέκταση των μιγαδικών αριθμών και χρησιμοποιούνται στην υπολογιστική όραση για να περιγράψουν περιστροφές στις τρεις διαστάσεις, αποφεύγοντας ταυτόχρονα το φαινόμενο Gimbal lock<sup>3</sup>. Αναλυτικότερα, κάθε τετραδόνιο μπορεί να αναπαρασταθεί ως ένα  $4 \times 1$  διάνυσμα, που βέβαια υπακούει σε διαφορετική άλγεβρα. Συγκεκριμένα κάθε τετραδόνιο  $\hat{\mathbf{q}} = [q_\theta \ q_x \ q_y \ q_z]^T$  μοναδιαίου μέτρου περιγράφει μια περιστροφή στον  $SO(3)$  όπου το πραγματικό του μέρος  $q_\theta$  δηλώνει την γωνία και το φανταστικό μέρος  $[q_x \ q_y \ q_z]^T$  τον άξονα περιστροφής, όπως ορίζεται απο το θεώρημα περιστροφής του Euler [3.1.1]. Η μέθοδος που παρουσιάζεται εδώ αρχικά υπολογίζει το βέλτιστο μοναδιαίο τετραδόνιο  $\hat{\mathbf{q}} = [q_\theta \ q_x \ q_y \ q_z]^T$  και μέσω αυτού εξάγει τον βέλτιστο μετασχηματισμό ισομετρίας  $\mathbf{R}^*$ ,  $\mathbf{t}^*$ .

Τα  $\mathbf{R}^*$  και  $\mathbf{t}^*$  που προκύπτουν απο την ελαχιστοποίηση της σχέσης (4.1) υπολογίζονται ως εξής. Αρχικά υπολογίζεται το μητρώο συνδιασπορών  $\mathbf{M}$  των συνόλων  $\mathbf{W}$  και  $\mathbf{X}$  ως:

$$\mathbf{M} = \text{COV}(\mathbf{W}, \mathbf{X}) = \mathbf{W}\mathbf{X}^T = \begin{bmatrix} \nu_{xx} & \nu_{xy} & \nu_{xz} \\ \nu_{yx} & \nu_{yy} & \nu_{yz} \\ \nu_{zx} & \nu_{zy} & \nu_{zz} \end{bmatrix}. \quad (4.2)$$

Ως δεύτερο βήμα, μέσω των στοιχείων του  $\mathbf{M}$  κατασκευάζεται το μητρώο  $\mathbf{N}$  ως:

$$\mathbf{N} = \begin{bmatrix} \nu_{xx} + \nu_{yy} + \nu_{zz} & \nu_{yz} - \nu_{zy} & \nu_{zx} - \nu_{xz} & \nu_{xy} - \nu_{yx} \\ \nu_{yz} - \nu_{zy} & \nu_{xx} - \nu_{yy} - \nu_{zz} & \nu_{xy} + \nu_{yx} & \nu_{zx} + \nu_{xz} \\ \nu_{zx} - \nu_{xz} & \nu_{xy} + \nu_{yx} & -\nu_{xx} + \nu_{yy} - \nu_{zz} & \nu_{yz} + \nu_{zy} \\ \nu_{xy} - \nu_{yx} & \nu_{zx} + \nu_{xz} & \nu_{yz} + \nu_{zy} & -\nu_{xx} - \nu_{yy} + \nu_{zz} \end{bmatrix}. \quad (4.3)$$

Στη συνέχεια το βέλτιστο τετραδόνιο  $\hat{\mathbf{q}}$  μπορεί να υπολογιστεί ως το κυρίαρχο ιδιοδύνασμα του μητρώου  $\mathbf{N}$ , δηλαδή το κανονικοποιημένο ιδιοδύνασμα που αντιστοιχεί στη μέγιστη (θετική) ιδιοτιμή του  $\mathbf{N}$ . Απο το οποίο ο βέλτιστος  $\mathbf{R}^*$  μπορεί να εξαχθεί χρησιμοποιώντας τον τύπο του Olinde Rodrigues<sup>4</sup>

$$\mathbf{R}^* = \begin{bmatrix} 1 - 2q_y^2 - 2q_z^2 & 2q_x q_y & 2q_x q_z + 2q_y q_\theta \\ 2q_x q_y + 2q_z q_\theta & 1 - 2q_x^2 - 2q_z^2 & 1 - 2q_y q_z - 2q_x q_\theta \\ 2q_x q_z - 2q_y q_\theta & 2q_y q_z + 2q_x q_\theta & 1 - 2q_x^2 - 2q_y^2 \end{bmatrix}. \quad (4.4)$$

Τέλος έχοντας υπολογίσει τον  $\mathbf{R}^*$  το διάνυσμα ολίσθησης  $\mathbf{t}^*$  μπορεί να υπολογιστεί πολύ εύκολα ως η διαφορά του κεντροειδούς των σημείων του  $\mathbf{W}$  και του μετασχηματισμένου κεντροειδούς των σημείων του  $\mathbf{X}$  ή πιο φορμαλιστικά απο την ακόλουθη σχέση του μητρώου  $\mathbf{X}\mathbf{W}^T$

$$\mathbf{t}^* = \frac{1}{n} \sum_{i=1}^n \mathbf{w}_i - \frac{1}{n} \mathbf{R}^* \sum_{i=1}^n \mathbf{x}_i \quad (4.5)$$

<sup>2</sup>[http://en.wikipedia.org/wiki/Quaternions\\_and\\_spatial\\_rotation](http://en.wikipedia.org/wiki/Quaternions_and_spatial_rotation)

<sup>3</sup>Gimbal lock ονομάζεται η απώλεια ενός βαθμού ελευθερίας κατά την περιστροφή στον τρισδιάστατο χώρο και συμβαίνει όταν οι άξονες περιστροφής δύο εκ των τριών δακτυλίων gimbals γίνουν παράλληλοι.

<sup>4</sup>[http://en.wikipedia.org/wiki/Rodrigues%27\\_rotation\\_formula](http://en.wikipedia.org/wiki/Rodrigues%27_rotation_formula)

### 4.1.3 Επίλυση βασισμένη σε ορθοκανωνικά μητρώα

Σε αυτή την υποενότητα παρουσιάζουμε την λύση που δόθηκε από τους Horn et.al<sup>[5]</sup> και η οποία είναι πανομοιότυπη με την λύση των Arun et.al<sup>[7]</sup>. Να σημειώσουμε ότι και οι δύο δημοσιεύτηκαν την ίδια χρονική περίοδο και βασίζονται στην εύρεση του βέλτιστου ορθογωνίου  $\mathbf{R}$  μέσω της πολικής διάσπασης (**Polar Decomposition**). Η λύση έχει ως εξής. Αρχικά γίνεται 'αποσύνδεση' (**decoupling**) του διανύσματος ολίσθησης  $\mathbf{t}$  προσθαφαιρώντας τους αριθμητικούς μέσους των δύο μητρώων:

$$(4.1) \triangleq \sum_{i=1}^n \|\mathbf{w}_i - \bar{\mathbf{w}} + \bar{\mathbf{w}} - \mathbf{R}\mathbf{x}_i - \mathbf{R}\bar{\mathbf{x}} + \mathbf{R}\bar{\mathbf{x}} - \mathbf{t}\|_2^2 = \sum_{i=1}^n \|(\mathbf{w}_i - \bar{\mathbf{w}}) - \mathbf{R}(\mathbf{x}_i - \bar{\mathbf{x}}) - (\mathbf{t} - \bar{\mathbf{w}} + \mathbf{R}\bar{\mathbf{x}})\|_2^2, \quad (4.6)$$

όπου τα  $\bar{\mathbf{w}}$  και  $\bar{\mathbf{x}}$  ορίζονται ως:

$$\bar{\mathbf{w}} = \frac{1}{n} \sum_{i=1}^n \mathbf{w}_i, \quad \bar{\mathbf{x}} = \frac{1}{n} \sum_{i=1}^n \mathbf{x}_i \quad (4.7)$$

Στη συνέχεια, ορίζοντας τις παρακάτω ποσότητες:

$$\begin{aligned} \bar{\mathbf{w}}_i &= \mathbf{w}_i - \bar{\mathbf{w}}, \\ \bar{\mathbf{x}}_i &= \mathbf{x}_i - \bar{\mathbf{x}}, \\ \bar{\mathbf{t}} &= \mathbf{t} - \bar{\mathbf{w}} + \mathbf{R}\bar{\mathbf{x}} \end{aligned} \quad (4.8)$$

η (4.6) γράφεται ως:

$$\sum_{i=1}^n \|\bar{\mathbf{w}}_i - \mathbf{R}\bar{\mathbf{x}}_i - \bar{\mathbf{t}}\|_2^2, \quad (4.9)$$

η οποία γράφεται ισοδύναμα ως ακολούθως

$$\sum_{i=1}^n \|\bar{\mathbf{w}}_i - \mathbf{R}\bar{\mathbf{x}}_i\|_2^2 - 2\bar{\mathbf{t}}^T \left( \sum_{i=1}^n (\bar{\mathbf{w}}_i - \mathbf{R}\bar{\mathbf{x}}_i) \right) + n \|\bar{\mathbf{t}}\|_2^2. \quad (4.10)$$

Παρατηρούμε ότι η ελαχιστοποίηση της (4.10) ως προς  $\bar{\mathbf{t}}$ , επιτυγχάνεται όταν το  $\bar{\mathbf{t}}$  είναι το μηδενικό διάνυσμα. Από την παρατήρηση αυτή και την (4.8) εύκολα βρίσκουμε:

$$\mathbf{t} = \bar{\mathbf{w}} - \mathbf{R}\bar{\mathbf{x}}. \quad (4.11)$$

Στο επόμενο βήμα πρέπει να ελαχιστοποιήσουμε την συνάρτηση κόστους (4.10) ως προς  $\mathbf{R}$ . Είναι εύκολο να παρατηρήσουμε ότι μόνο ο πρώτος όρος του αθροίσματος εξαρτάται από το  $\mathbf{R}$ , ο οποίος πλέον δεν περιέχει  $\mathbf{t}$ . Αναλύοντας την Ευκλείδεια νόρμα του πρώτου όρου της (4.10) προκύπτει:

$$\sum_{i=1}^n \left[ \|\bar{\mathbf{w}}_i\|_2^2 - 2\bar{\mathbf{w}}_i^T \mathbf{R}\bar{\mathbf{x}}_i + \bar{\mathbf{x}}_i^T \mathbf{R}^T \mathbf{R}\bar{\mathbf{x}}_i \right]. \quad (4.12)$$

Εφόσον γνωρίζουμε ότι το ζητούμενο μητρώο  $\mathbf{R}$  είναι ορθογώνιο:

$$\bar{\mathbf{x}}_i^T \mathbf{R}^T \mathbf{R}\bar{\mathbf{x}}_i = \bar{\mathbf{x}}_i^T \mathbf{I}\bar{\mathbf{x}}_i = \|\bar{\mathbf{x}}_i\|_2^2 \quad \forall i, \quad (4.13)$$



και επομένως η (4.12) απλοποιείται στην:

$$\sum_{i=1}^n \|\bar{\mathbf{w}}_i\|_2^2 - 2 \sum_{i=1}^n (\bar{\mathbf{w}}_i^T \mathbf{R} \bar{\mathbf{x}}_i) + \|\bar{\mathbf{x}}\|_2^2 . \quad (4.14)$$

Παρατηρούμε ότι μόνο ο δεύτερος όρος του αθροίσματος εξαρτάται από το  $\mathbf{R}$  και επιπλέον ότι η ποσότητα (4.14) ελαχιστοποιείται όταν μεγιστοποιηθεί το

$$\Phi(\mathbf{R}) = \sum_{i=1}^n (\bar{\mathbf{w}}_i^T \mathbf{R} \bar{\mathbf{x}}_i) \quad (4.15)$$

το οποίο, χρησιμοποιώντας τις ιδιότητες του ίχνους (trace) έχουμε ότι

$$\Phi(\mathbf{R}) = \text{tr}(\mathbf{W}^T \mathbf{R} \mathbf{X}) = \text{tr}(\mathbf{R} \mathbf{X} \mathbf{W}^T) . \quad (4.16)$$

Χρησιμοποιώντας το λήμμα του μέγιστου ίχνους που δίνεται αναλυτικά στην [Ενότητα 3.2.1] βρίσκουμε ότι ο βέλτιστος  $\mathbf{R}$  μπορεί να υπολογιστεί ως:

$$\mathbf{R} = \mathbf{V} \mathbf{U}^T, \quad (4.17)$$

όπου τα  $\mathbf{V}$  και  $\mathbf{U}$  προκύπτουν από την διάσπαση ιδιαζουσών τιμών του  $\mathbf{X} \mathbf{W}^T$  ως:

$$\mathbf{U} \mathbf{\Sigma} \mathbf{V}^T = \mathbf{X} \mathbf{W}^T . \quad (4.18)$$

Στη συνέχεια το βέλτιστο διάνυσμα ολισθήσεων  $\mathbf{t}$  μπορεί να υπολογιστεί από τη σχέση (4.11) αντικαθιστώντας τον βέλτιστο  $\mathbf{R}$  από την σχέση (4.17):

$$\mathbf{t} = \bar{\mathbf{w}} - \mathbf{V} \mathbf{U}^T \bar{\mathbf{x}} . \quad (4.19)$$

Συμπληρωματικά αναφέρουμε ότι σε ιδιαίζουσες περιπτώσεις δηλαδή περιπτώσεις που τα σημεία είναι συνεπίπεδα, συνευθειακά ή περιέχουν υψηλό ποσοστό θορύβου, η ορίζουσα του μητρώου  $\mathbf{V} \mathbf{U}^T$  μπορεί να προκύψει ίση με  $-1$  το οποίο συνεπάγεται αντικατοπτρισμό (reflection). Σε αυτή την περίπτωση προτείνεται από τους συγγραφείς, ο  $\mathbf{R}$  να υπολογίζεται ως εξής:

$$\mathbf{R} = \mathbf{V} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & |\mathbf{V} \mathbf{U}^T| \end{bmatrix} \mathbf{U}^T . \quad (4.20)$$

#### 4.1.4 Η τροποποιημένη επίλυση του Shinji Umeyama

Η τελευταία μέθοδος επίλυσης του προβλήματος απόλυτου προσανατολισμού που θα παρουσιάσουμε δόθηκε από τον Shinji Umeyama<sup>[6]</sup> το 1991 και βασίζεται στους πολλαπλασιαστές Lagrange. Η πρωταρχική λύση που δίνεται στο<sup>[6]</sup> αφορά την ελαχιστοποίηση της συνάρτησης κόστους (4.1). Εδώ θα παρουσιάσουμε μια τροποποιημένη απόδειξη στην οποία κάθε όρος της (4.1) έχει και ένα βάρος  $b_i$ . Φορμαλιστικά η ποσότητα που θα ελαχιστοποιήσουμε είναι η ακόλουθη:

$$\{\mathbf{R}^*, \mathbf{t}^*\} = \arg \min_{\mathbf{R}, \mathbf{t}} \frac{1}{2} \sum_{i=1}^n b_i \|\mathbf{w}_i - \mathbf{R} \mathbf{x}_i - \mathbf{t}\|_2^2 , \quad (4.21)$$

όπου τα  $\mathbf{W}$  και  $\mathbf{X}$  είναι  $3 \times n$  μητρώα - νέφη όπως και προηγουμένως και  $\mathbf{b}$  είναι ένα  $n \times 1$  διάνυσμα με στοιχεία  $b_i \geq 0$  .

Ως πρώτο βήμα ορίζουμε την συνάρτηση κόστους  $f(\mathbf{R}, \mathbf{t})$  ως το όρισμα ελαχιστοποίησης της (4.21) και ακολούθως την μετατρέπουμε σε μητρική μορφή:

$$\begin{aligned} f(\mathbf{R}, \mathbf{t}) &= \text{tr}\{[\mathbf{W} - \mathbf{R}\mathbf{X} - \mathbf{t}\mathbf{h}^T]\mathbf{B}\mathbf{B}^T[\mathbf{W} - \mathbf{R}\mathbf{X} - \mathbf{t}\mathbf{h}^T]^T\} = \\ &= \text{tr}\{[\mathbf{W}\mathbf{B} - \mathbf{R}\mathbf{X}\mathbf{B} - \mathbf{t}\mathbf{h}^T\mathbf{B}][\mathbf{W}\mathbf{B} - \mathbf{R}\mathbf{X}\mathbf{B} - \mathbf{t}\mathbf{h}^T\mathbf{B}]^T\} = \\ &= \|\mathbf{W}\mathbf{B} - \mathbf{R}\mathbf{X}\mathbf{B} - \mathbf{t}\mathbf{h}^T\|_{\mathbf{F}}^2, \end{aligned} \quad (4.22)$$

όπου το  $\mathbf{B}$  είναι ένα  $n \times n$  διαγώνιο μητρώο της μορφής:

$$\mathbf{B} = \begin{bmatrix} \sqrt{b_1} & 0 & 0 & \cdots & 0 \\ 0 & \sqrt{b_2} & 0 & \cdots & 0 \\ 0 & 0 & \sqrt{b_3} & \cdots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \sqrt{b_n} \end{bmatrix}, \quad (4.23)$$

και το  $\mathbf{h}$  ένα  $n \times 1$  διάνυσμα "κλωνοποίησης" (**replication**):

$$\mathbf{h} = \underbrace{[1 \ 1 \ 1 \ \cdots \ 1 \ 1]^T}_{n \text{ times}}. \quad (4.24)$$

Για την μετατροπή της (4.21) βασιστήκαμε στον ορισμό της στάθμης του Frobenius:

$$\|\mathbf{A}\|_{\mathbf{F}} = \sqrt{\text{tr}(\mathbf{A}\mathbf{A}^T)}.$$

Με βάσει τα  $\mathbf{B}$  και  $\mathbf{h}$  η σχέση (4.22) μπορεί να γραφεί πιο συμπτηγμένα ως:

$$f(\mathbf{R}, \mathbf{t}) = \|\tilde{\mathbf{W}} - \mathbf{R}\tilde{\mathbf{X}} - \tilde{\mathbf{t}}\tilde{\mathbf{h}}^T\|_{\mathbf{F}}^2, \quad (4.25)$$

θέτωντας επιπλέον:

$$\tilde{\mathbf{X}} = \mathbf{X}\mathbf{B},$$

$$\tilde{\mathbf{W}} = \mathbf{W}\mathbf{B},$$

$$\tilde{\mathbf{h}} = \mathbf{B}\mathbf{h}.$$

Υπολογίζοντας τις μερικές παραγώγους της σχέσης (4.25) ως προς  $\mathbf{t}$  και εξισώνοντας με μηδέν προκύπτει ότι

$$\begin{aligned} \frac{\partial f}{\partial \mathbf{t}} = 0 &\Rightarrow 2(\tilde{\mathbf{W}} - \mathbf{R}\tilde{\mathbf{X}} - \tilde{\mathbf{t}}\tilde{\mathbf{h}}^T)(-\tilde{\mathbf{h}}) = 0 \Rightarrow \\ &\mathbf{t}(\tilde{\mathbf{h}}^T\tilde{\mathbf{h}}) = \tilde{\mathbf{W}}\tilde{\mathbf{h}} - \mathbf{R}\tilde{\mathbf{X}}\tilde{\mathbf{h}} \Rightarrow \\ &\mathbf{t} = \frac{\tilde{\mathbf{W}}\tilde{\mathbf{h}}}{\tilde{\mathbf{h}}^T\tilde{\mathbf{h}}} - \frac{\mathbf{R}\tilde{\mathbf{X}}\tilde{\mathbf{h}}}{\tilde{\mathbf{h}}^T\tilde{\mathbf{h}}}. \end{aligned} \quad (4.26)$$

Αντικαθιστώντας το  $\mathbf{t}$  στην σχέση (4.25) έχουμε

$$\|\tilde{\mathbf{W}}(\mathbf{I}_n - \frac{\tilde{\mathbf{h}}\tilde{\mathbf{h}}^T}{\tilde{\mathbf{h}}^T\tilde{\mathbf{h}}}) - \mathbf{R}\tilde{\mathbf{X}}(\mathbf{I}_n - \frac{\tilde{\mathbf{h}}\tilde{\mathbf{h}}^T}{\tilde{\mathbf{h}}^T\tilde{\mathbf{h}}})\|_{\mathbf{F}}^2. \quad (4.27)$$

Θέτοντας επίσης την ποσότητα<sup>5</sup>

$$\mathbf{I}_n - \frac{\tilde{\mathbf{h}}\tilde{\mathbf{h}}^T}{\tilde{\mathbf{h}}^T\tilde{\mathbf{h}}} = \mathbf{K}, \quad (4.28)$$

η ως προς  $\mathbf{R}$  αντικειμενική συνάρτηση γράφεται ως

$$\|(\tilde{\mathbf{W}} - \mathbf{R}\tilde{\mathbf{X}})\mathbf{K}\|_{\mathbf{F}}^2. \quad (4.29)$$

Έχοντας υπολογίσει το διάνυσμα  $\mathbf{t}$  συναρτήσεως του  $\mathbf{R}$  μπορούμε να προχωρήσουμε στην ελαχιστοποίηση της  $f$  ως προς  $\mathbf{R}$ .

Ο υπολογισμός του ακροτάτου της (4.29) πρέπει να γίνει λαμβάνοντας υπόψη μας ότι το  $\mathbf{R}$  πρέπει να είναι μητρώο περιστροφής. Η παραπάνω συνθήκη μας οδηγεί σε δύο επιπλέον περιορισμούς: Την ορθογωνιότητα, δηλαδή ότι  $\mathbf{R}^T\mathbf{R} = \mathbf{R}\mathbf{R}^T = \mathbf{I}_3$  και το γεγονός ότι πρέπει  $|\mathbf{R}| = 1$ . Για να μπορέσουμε να ικανοποιήσουμε αυτούς τους περιορισμούς πρέπει να επεκτείνουμε την συνάρτηση κόστους με χρήση πολλαπλασιαστών Lagrange. Η συνάρτηση κόστους  $\mathcal{F} = \text{obj}(\mathbf{R}, \mathbf{L}, g)$  θα έχει πλέον τη μορφή:

$$\begin{aligned} \mathcal{F} &= \|(\tilde{\mathbf{W}} - \mathbf{R}\tilde{\mathbf{X}})\mathbf{K}\|_{\mathbf{F}}^2 + \text{tr}\{\mathbf{L}(\mathbf{R}^T\mathbf{R} - \mathbf{I}_3)\} + g(|\mathbf{R}| - 1) \Rightarrow \\ \mathcal{F} &= \text{tr}\{\mathbf{K}(\tilde{\mathbf{W}}^T - \tilde{\mathbf{X}}^T\mathbf{R}^T)(\tilde{\mathbf{W}} - \mathbf{R}\tilde{\mathbf{X}})\mathbf{K}\} + \text{tr}\{\mathbf{L}(\mathbf{R}^T\mathbf{R} - \mathbf{I}_3)\} + g(|\mathbf{R}| - 1) \Rightarrow \\ \mathcal{F} &= \text{tr}\{\mathbf{K}\tilde{\mathbf{W}}^T\tilde{\mathbf{W}} - 2\mathbf{R}\tilde{\mathbf{X}}\mathbf{K}\tilde{\mathbf{W}}^T + \mathbf{R}\tilde{\mathbf{X}}\mathbf{K}\tilde{\mathbf{X}}^T\mathbf{R}^T\} + \text{tr}\{\mathbf{R}\mathbf{L}\mathbf{R}^T - \mathbf{L}\} + g(|\mathbf{R}| - 1) \end{aligned} \quad (4.30)$$

όπου  $\mathbf{L}$  είναι ένα συμμετρικό, εκ της συμμετρίας της συνθήκης, μητρώο αποτελούμενο από πολλαπλασιαστές lagrange και  $g \in \mathfrak{R}_+^*$ .

Αναζητούμε το ακρότατο αυτής της συνάρτησης, οπότε από το θεώρημα του Fermat έχουμε ότι στο σημείο του ακροτάτου  $\nabla\mathcal{F} = \vec{0}$ .

$$\frac{\partial\mathcal{F}}{\partial\mathbf{R}} = \mathcal{O}_{3 \times 3} \Rightarrow -2\tilde{\mathbf{W}}\mathbf{K}\tilde{\mathbf{X}}^T + 2\mathbf{R}\tilde{\mathbf{X}}\mathbf{K}\tilde{\mathbf{X}}^T + 2\mathbf{R}\mathbf{L} + g|\mathbf{R}|(\mathbf{R}^{-1})^T = \mathcal{O}_{3 \times 3}, \quad (4.31)$$

$$\frac{\partial\mathcal{F}}{\partial\mathbf{L}} = \mathcal{O}_{3 \times 3} \Rightarrow \mathbf{R}^T\mathbf{R} = \mathbf{I}_{3 \times 3}, \quad (4.32)$$

$$\frac{\partial\mathcal{F}}{\partial g} = \mathcal{O}_{3 \times 3} \Rightarrow |\mathbf{R}| = 1, \quad (4.33)$$

όπου οι μερικές παράγωγοι της  $\mathcal{F}$  υπολογίστηκαν σύμφωνα με τους ακόλουθους κανόνες

$$\frac{\partial \text{tr}(\mathbf{A}\mathbf{X}\mathbf{B})}{\partial\mathbf{X}} = \mathbf{A}^T\mathbf{B}^T, \quad (4.34)$$

$$\frac{\partial \text{tr}(\mathbf{X}\mathbf{B}\mathbf{X}^T)}{\partial\mathbf{X}} = \mathbf{X}\mathbf{B}^T + \mathbf{X}\mathbf{B}, \quad (4.35)$$

$$(4.36)$$

και τον τύπο του Jacobi για τον υπολογισμό της παραγώγου της ορίζουσας:

$$\frac{\partial|\mathbf{X}|}{\partial\mathbf{X}} = |\mathbf{X}| \cdot (\mathbf{X}^{-1})^T. \quad (4.37)$$

<sup>5</sup>Το μητρώο  $\mathbf{K}$  είναι συμμετρικό και ταυτοδύναμο, i.e.  $\mathbf{K}^n = \mathbf{K} = \mathbf{K}^T$ .

Χρησιμοποιώντας το αποτέλεσμα της (4.32) και της (4.33) η εξίσωση (4.31) απλοποιείται στην:

$$\mathbf{R}(\tilde{\mathbf{X}}\mathbf{K}\tilde{\mathbf{W}}^T + \mathbf{L} + \frac{1}{2}g\mathbf{I}) = \tilde{\mathbf{W}}\mathbf{K}\tilde{\mathbf{X}}^T. \quad (4.38)$$

Σε αυτό το σημείο θέτουμε την ποσότητα μέσα στην παρένθεση της εξίσωσης (4.38) ως

$$\mathbf{L}' = \tilde{\mathbf{X}}\mathbf{K}\tilde{\mathbf{W}}^T + \mathbf{L} + \frac{1}{2}g\mathbf{I}. \quad (4.39)$$

Όστε τελικά η εξίσωση (4.38) να μπορεί να γραφεί ως:

$$\mathbf{R}\mathbf{L}' = \tilde{\mathbf{W}}\mathbf{K}\tilde{\mathbf{X}}^T. \quad (4.40)$$

Αν τώρα αναστρέψουμε και τα δύο μέλη της (4.40) καταλήγουμε στο ότι:

$$\mathbf{L}'\mathbf{R}^T = \tilde{\mathbf{X}}\mathbf{K}\tilde{\mathbf{W}}^T, \quad (4.41)$$

όπου το  $\mathbf{L}'$  είναι ένα άθροισμα συμμετρικών μητρώων όποτε είναι και ο ίδιος συμμετρικός. Εάν τώρα πολλαπλασιάσουμε κάθε μέλος της (4.41) απο τα δεξιά, με κάθε μέλος της (4.40) προκύπτει:

$$\mathbf{L}'^2 = \tilde{\mathbf{X}}\mathbf{K}\tilde{\mathbf{W}}^T \tilde{\mathbf{W}}\mathbf{K}\tilde{\mathbf{X}}^T. \quad (4.42)$$

Για το επόμενο βήμα χρειαζόμαστε την διάσπαση ιδιαζουσών τιμών του  $\tilde{\mathbf{W}}\mathbf{K}\tilde{\mathbf{X}}^T$  η οποία υπολογίζεται ως

$$\tilde{\mathbf{W}}\mathbf{K}\tilde{\mathbf{X}}^T = \mathbf{U}\mathbf{D}\mathbf{V}^T, \quad (4.43)$$

όπου  $\mathbf{U}$  και  $\mathbf{V}^T$  είναι τα ορθοκανονικά μητρώα που περιέχουν τα αριστερά και δεξιά ιδιάζοντα διανύσματα του  $\tilde{\mathbf{W}}\mathbf{K}\tilde{\mathbf{X}}^T$ , δηλαδή τα κανονικοποιημένα ιδιοδιανύσματα του  $\tilde{\mathbf{W}}\mathbf{K}\tilde{\mathbf{X}}^T\tilde{\mathbf{X}}\mathbf{K}\tilde{\mathbf{W}}^T$  και του  $\tilde{\mathbf{X}}\mathbf{K}\tilde{\mathbf{W}}^T\tilde{\mathbf{W}}\mathbf{K}\tilde{\mathbf{X}}^T$  αντίστοιχα, και  $\mathbf{D}$  είναι ένα μή αρνητικό διαγώνιο μητρώο που περιέχει της ιδιάζουσες τιμές  $\tilde{\mathbf{W}}\mathbf{K}\tilde{\mathbf{X}}^T$  υπολογισμένες ως τις ρίζες των ιδιοτιμών του  $\tilde{\mathbf{X}}\mathbf{K}\tilde{\mathbf{W}}^T\tilde{\mathbf{W}}\mathbf{K}\tilde{\mathbf{X}}^T$  (ή ισοδύναμα του  $\tilde{\mathbf{W}}\mathbf{K}\tilde{\mathbf{X}}^T\tilde{\mathbf{X}}\mathbf{K}\tilde{\mathbf{W}}^T$ ), με την ίδια σειρά.

Εάν τώρα ξαναγράψουμε την (4.42) χρησιμοποιώντας την (4.43) προκύπτει:

$$\mathbf{L}'^2 = \mathbf{V}\mathbf{D}\mathbf{U}^T\mathbf{U}\mathbf{D}\mathbf{V}^T = \mathbf{V}\mathbf{D}^2\mathbf{V}^T \Rightarrow \mathbf{L}' = \mathbf{V}\mathbf{D}\mathbf{S}\mathbf{V}^T, \quad (4.44)$$

όπου το μητρώο  $\mathbf{S}$  κρύβει ένα πρόσημο το οποίο ίσως να χάθηκε στον τετραγωνισμό

$$\mathbf{S} = \begin{cases} \mathbf{I}_{3 \times 3} & \text{εάν } |\mathbf{U}| \cdot |\mathbf{V}| = 1 \\ \text{diag}(1, 1, -1) & \text{εάν } |\mathbf{U}| \cdot |\mathbf{V}| = -1. \end{cases} \quad (4.45)$$

Αντικαθιστώντας τώρα την (4.44) στην (4.40) και λύνοντας' ως προς  $\mathbf{R}$ , τουλάχιστον με την έννοια των ελαχίστων τετραγώνων στην περίπτωση που το  $\mathbf{L}$  δεν έχει πλήρη τάξη, καταλήγουμε στο ζητούμενο:

$$\begin{aligned} \mathbf{R}\mathbf{L}' &= \tilde{\mathbf{W}}\mathbf{K}\tilde{\mathbf{X}}^T & \Rightarrow \\ \mathbf{R} &= \mathbf{U}\mathbf{D}\mathbf{V}^T\mathbf{L}'^{-1} & \Rightarrow \\ \mathbf{R} &= \mathbf{U}\mathbf{D}\mathbf{V}^T(\mathbf{V}\mathbf{D}\mathbf{S}^{-1}\mathbf{V}^T) & \Rightarrow \\ \mathbf{R} &= \mathbf{U}\mathbf{D}\mathbf{V}^T(\mathbf{V}\mathbf{D}^{-1}\mathbf{S}\mathbf{V}^T) & \Rightarrow \\ \mathbf{R} &= \mathbf{U}\mathbf{S}\mathbf{V}^T. & \end{aligned} \quad (4.46)$$

Να σημειώσουμε εδώ ότι οι  $\mathbf{S}$  και  $\mathbf{D}$  είναι διαγώνια μητρώα και επιπλέον αντιμετατίθονται (commute).

Τέλος αντικαθιστώντας το  $\mathbf{R}$  που υπολογίσαμε στην (4.46), στην (4.26) υπολογίζουμε το βέλτιστο διάνυσμα ολίσθησης:

$$\begin{aligned} \mathbf{t} &= \tilde{\mathbf{W}} \frac{\tilde{\mathbf{h}}}{\tilde{\mathbf{h}}^T \tilde{\mathbf{h}}} - \mathbf{USV}^T \tilde{\mathbf{X}} \frac{\tilde{\mathbf{h}}}{\tilde{\mathbf{h}}^T \tilde{\mathbf{h}}} \quad \Rightarrow \\ \mathbf{t} &= \frac{\tilde{\mathbf{W}} \mathbf{B} \mathbf{h}}{\mathbf{h}^T \mathbf{B} \mathbf{B} \mathbf{h}} - \frac{\mathbf{USV}^T \tilde{\mathbf{X}} \mathbf{B} \mathbf{h}}{\mathbf{h}^T \mathbf{B} \mathbf{B} \mathbf{h}} \quad \Rightarrow \\ \mathbf{t} &= \frac{1}{\sum_{i=1}^n b_i^2} \left( \tilde{\mathbf{W}} - \mathbf{USV}^T \tilde{\mathbf{X}} \right) \mathbf{B} \mathbf{h}. \end{aligned} \quad (4.47)$$

#### 4.1.5 Η συνέχεια Lipschitz και άλλες μετρικές σφάλματος

Έως τώρα παρουσιάσαμε την εύρεση των παραμέτρων του μετασχηματισμού ελαχιστοποιώντας την μετρική ‘αθροίσματος Ευκλείδειων αποστάσεων’, είναι όμως αυτή η μοναδική μετρική μέσω της οποίας μπορούμε να λύσουμε το πρόβλημα; Όπως θα δούμε στην επόμενη ενότητα, όχι μόνο δέν είναι μοναδική, αλλά ίσως σε κάποιες περιπτώσεις να υπάρχουν πολύ καλύτερες. Η αυστηρή ονομασία της μετρικής αυτής είναι (point to point error metric). Αν σκεφτούμε μία περίπτωση δύο νεφών τα οποία περιέχουν μεγάλο πλήθος outliers, δηλαδή μη πραγματικά σημεία που έχουν προέλθει από κάποια μορφή θορύβου και δέν ανήκουν σε κάποια από τις προς αντιστοίχιση επιφάνειες. Τότε η χρήση της μετρικής point to point, χωρίς κάποια τροποποίηση, θα έδινε αποτελέσματα πολύ μακριά από τις πραγματικές τιμές. Μία συνήθης παραλλαγή είναι η μετρική point to plane - point to surface. Η μετρική αυτή δεν λαμβάνει υπόψη της τις αποστάσεις των σημείων αυτές κάθε αυτές αλλά τις αποστάσεις των σημείων του ενός νέφους από την επιφάνεια του άλλου. Η επιφάνεια αυτή συνήθως έχει προέλθει από ανακατασκευή (triangulation) μέσω των σημείων-δειγμάτων και απαρτίζεται από μικρά επίπεδα (συνήθως πολύγωνα ή τρίγωνα), οπότε κατά βάθος η μετρική εξαρτάται από τις αποστάσεις των σημείων από τα κοντινότερα, σε αυτά, επίπεδα. Από την άλλη πλευρά για την ελαχιστοποίηση της point to point μετρικής υπάρχουν κλειστού τύπου λύσεις ενώ για την ελαχιστοποίηση της point to surface πρέπει να χρησιμοποιηθεί κάποια επαναληπτική μέθοδος η οποία ίσως παρουσιάζει προβλήματα σύγκλισης. Εκτός αυτών των δύο βασικών κατηγοριών μετρικών υπάρχουν και άλλες μετρικές που χρησιμοποιούνται ως συναρτήσεις κόστους-οφέλους σε προβλήματα αντιστοίχισης. Συνήθεις είναι οι μετρικές πληροφορίας (information based metrics) και οι μετρικές πιθανοφάνειας (likelihood based metrics). Αξίζει να αναφέρουμε ότι κάποιοι ερευνητές χρησιμοποιούν την συνέχεια Lipschitz για να επιταχύνουν την εύρεση του ακροτάτου της συνάρτησης κόστους. Αυτό το επιτυγχάνουν μετατρέποντας τη συνάρτηση κόστους σε μια άλλη, συνεχή κατά Lipschitz<sup>6</sup> συνάρτηση, για την οποία γνωρίζουν ότι παρουσιάζει ακρότατα στα ίδια σημεία με την αρχική. Το

<sup>6</sup>Μια συνάρτηση  $f(x)$  ονομάζεται συνεχής κατά Lipschitz αν υπάρχει πεπερασμένη σταθερά  $K \geq 0$  τέτοια ώστε  $\frac{df(x)}{dx} \leq K \forall x \in Df$ . Με άλλα λόγια η ιδιότητα αυτή περιορίζει το μέγιστο ρυθμό μεταβολής της συνάρτησης.

πλεονέκτημα που τελικά έχουν είναι να μπορούν να αυξάνουν το βήμα της επαναληπτικής μεθόδου που χρησιμοποιούν για την εύρεση του ακρότατου, γνωρίζοντας ότι εξαιτίας της ‘αργής μεταβολής’ της τροποποιημένης συνάρτησης, το ακρότατο δεν μπορεί να βρισκείται πολύ κοντά στην προηγούμενη εκτίμηση.

Για παράδειγμα η συνάρτηση  $f_1(x) = x^2$  δεν είναι συνεχής κατα Lipshitz διότι η  $f_1^{(1)}(x)$  δεν είναι φραγμένη. ενώ αντίθετα η  $f_2(x) = \exp(-f_1(x)) = \exp(-x^2)$  είναι διότι  $f_2^{(1)}(x) = 2x \exp(-x^2) \leq \sqrt{\frac{2}{e}} \forall x$ . Οι δυο αυτές συναρτήσεις παρουσιάζουν ακρότατα στα ίδια σημεία, οπότε αν θέλαμε να βρούμε το ακρότατο της  $f_1$  μέσω κάποιας επαναληπτικής μεθόδου θα μπορούσαμε να βρούμε το ισοδύναμο ακρότατο της  $f_2$  σε πολύ λιγότερα βήματα.

## 4.2 Άγνωστες αντιστοιχίες

Έως τώρα παρουσιάσαμε τρόπους εύρεσης του γεωμετρικού μετασχηματισμού μεταξύ δύο νεφών υπό την προϋπόθεση ότι έχουμε λύσει το πρόβλημα των αντιστοιχίων σημείων ([correspondence problem](#)). Όμως το βαρύνων υποπρόβλημα της αντιστοίχισης νεφών δεν είναι η εύρεση του μετασχηματισμού αλλά η εύρεση των αντιστοιχίων σημείων. Σε αυτή την υποενότητα παρουσιάζουμε τις κυρίαρχες προσεγγίσεις ([State of the art methods](#)) που έχουν προταθεί για τη λύση του προβλήματος αντιστοίχισης δύο νεφών υπο πραγματικές συνθήκες.

### 4.2.1 Ο αλγόριθμος ICP και οι παραλλαγές του

Ο πρώτος αλγόριθμος που προτάθηκε για τη λύση του πλήρους προβλήματος αντιστοίχισης ([correspondence problem and estimation of geometric transformation](#)) αναπτύχθηκε απο τους Besl and McKay<sup>[9]</sup> το 1992. Ο αλγόριθμος που πρότειναν ονομάζεται [Iterative Closest Point \(ICP\)](#)<sup>7</sup>. Όπως δηλώνει και η ονομασία του, πρόκειται για έναν επαναληπτικό αλγόριθμο, κάθε επανάληψη του οποίου αποτελείται απο δύο βήματα: ένα που αφορά την εύρεση αντίστοιχων σημείων με βάση το κριτήριο ελάχιστης απόστασης ([nearest neighbor criterion](#)) και ένα δεύτερο που αφορά την εκτίμηση του μετασχηματισμού μέσω των αντιστοιχιών που υπολογίστηκαν στο προηγούμενο βήμα. Αναλυτικότερα<sup>8</sup> ας υποθέσουμε δύο, όχι αναγκαστικά ισοπληθή, σύνολα σημείων  $\mathcal{X}$  και  $\mathcal{Y}$  και ας υποθέσουμε επίσης ότι αυτά τα δύο σύνολα συνδέονται με μετασχηματισμό ισομετρίας περιγραφόμενο από ένα μητρώο περιστροφής  $\mathbf{R}$  και ένα διάνυσμα ολίσθησης  $\mathbf{t}$ . Χωρίς απώλεια της γενικότητας υποθέτουμε ότι ο μετασχηματισμός εφαρμόζεται στα σημεία του συνόλου  $\mathcal{X}$ . Ο αλγόριθμος ICP υποθέτει ότι το ένα εκ των δύο συνόλων, έστω το  $\mathcal{X}$ , είναι το μοντέλο και στη συνέχεια αναζητά για κάθε σημείο  $\mathbf{x}_i \in \mathcal{X}$  το κοντινότερο σε αυτό σημείο  $\mathbf{y}_j \in \mathcal{Y}$  με απώτερο σκοπό να ελαχιστοποιήσει την ακόλουθη συνάρτηση κόστους:

$$\mathfrak{E}(\mathbf{R}, \mathbf{t}) = \sum_{i=1}^N \sum_{j=1}^M a_{ji} \|\mathbf{y}_j - (\mathbf{R}\mathbf{x}_i + \mathbf{t})\|_2^2, \quad (4.48)$$

όπου τα  $N$  και  $M$  συμβολίζουν τις πληθικότητες των συνόλων  $\mathcal{X}$  και  $\mathcal{Y}$  και τα  $a_{ji}$  είναι τα βάρη της αντίστοιχης νόρμας, τα οποία στην περίπτωση του ICP είναι δυαδικά. Συγκεκριμένα  $a_{ji} = 1$  εάν το  $\mathbf{y}_j$  είναι το κοντινότερο σημείο στο

<sup>7</sup> Αρχετοί ερευνητές προτιμούν τον όρο Iterative Corresponding Point

$\mathbf{x}_i$  και  $a_{ji} = 0$  σε άλλη περίπτωση. Εξαιτίας της ύπαρξης μηδενικών βαρών η εξίσωση (4.48) απλοποιείται στήν:

$$\mathcal{E}(\mathbf{R}, \mathbf{t}) = \sum_{i=1}^N \| \mathbf{y}_i - (\mathbf{R}\mathbf{x}_i + \mathbf{t}) \|_2^2 . \quad (4.49)$$

Παρατηρούμε σε αυτό το σημείο ότι έχουμε μια απλοποιημένη μορφή φίλτραρίσματος μόλυνσης (**outlier filtering**) υπο την έννοια ότι κάποια σημεία του συνόλου  $\mathcal{Y}$  μπορεί να μείνουν χωρίς αντιστοιχία και τελικά να μην λάβουν μέρος στη συνάρτηση κόστους  $\mathcal{E}$ . Σε κάθε βήμα (**iteration**), αφού έχουμε υπολογίσει τα  $a_{ji}$ , ο βέλτιστος μετασχηματισμός  $\mathbf{R}, \mathbf{t}$  μπορεί να βρεθεί με μία από τις μεθόδους που παρουσιάστηκαν στην [Ενότητα 4.1]. Να σημειώσουμε ότι ερευνητές που πρωτοπρότειναν τον αλγόριθμο χρησιμοποίησαν την μέθοδο των ορθομοναδιαίων τετραδονίων για την περίπτωση 3D σημείων-μορφών (shapes). Σε περίπτωση που ο αλγόριθμος χρησιμοποιείται για την αντιστοίχιση  $mD$  σημείων<sup>8</sup> οποιαδήποτε SVD based μέθοδος μπορεί να χρησιμοποιηθεί. Συνοψίζοντας, ο αλγόριθμος ICP παρατίθεται σε μορφή ψευδοκώδικα για ευκολία υλοποίησης και σύγκρισης με άλλους αλγόριθμους που θα παρουσιαστούν στη συνέχεια.

---

**Algorithm** Iterative-Closest-Point (ICP)

---

**Require:** Point-set matrices  $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_N]$  and  $\mathbf{Y} = [\mathbf{y}_1, \dots, \mathbf{y}_M]$ ,  
 $maxNumIter$ ,  $endeVal$ , Initial transformation parameters:  $\mathbf{R}_0, \mathbf{t}_0$  .

1: Initialize iteration counter:

$$q \leftarrow 1 .$$

2: **repeat**

3:   **for**  $i = 1, \dots, N$  **do**

4:     Find Correspondences using nearest neighbor search:

$$\mathbf{y}_i \leftarrow \arg \min_{\mathbf{y}_j \in \mathcal{Y}} \| \mathbf{y}_j - (\mathbf{R}^{q-1}\mathbf{x}_i + \mathbf{t}^{q-1}) \|_2^2$$

5:   **end for**

6:   Estimate optimum transformation using the newly estimated  $\mathbf{y}_i$ 's:

$$\mathbf{R}^q, \mathbf{t}^q \leftarrow \arg \min_{\mathbf{R}, \mathbf{t}} \sum_{i=1}^N \| \mathbf{y}_i - \mathbf{R}\mathbf{x}_i - \mathbf{t} \|_2^2 .$$

7:   Compute iteration error :

$$err(q) \leftarrow \| \mathbf{R}^q - \mathbf{R}^{q-1} \|_{\mathbf{F}}^2 .$$

8:   increase iteration counter:

$$q \leftarrow q + 1 .$$

9: **until**  $err(q) < endeVal$  ||  $q > maxNumIter$  .

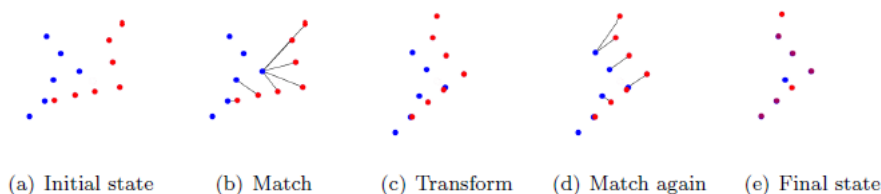
10: **return**  $\mathbf{R}^q, \mathbf{t}^q$  .

---

Επιπλέον στο επόμενο Σχήμα δίνεται το οπτικό αποτέλεσμα των βημάτων του

<sup>8</sup>Η μέθοδος των ορθομοναδιαίων τετραδονίων, εκ κατασκευής, λειτουργεί μόνο για σύνολα τρισδιάστατων σημείων.

αλγορίθμου:



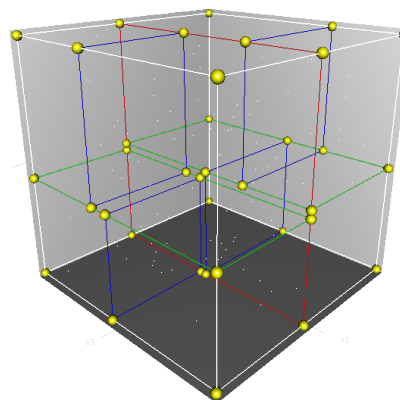
Σχήμα 4.1: Οπτικοποίηση των βημάτων του βασικού ICP, με  $\bullet \in \mathcal{X}$  και  $\bullet \in \mathcal{Y}$ .

### Αδυναμίες του βασικού αλγορίθμου και η ανάγκη τροποποίησης του

Η δυο δεκαετίες που πέρασαν από τότε που πρωτοπροτάθηκε ο ICP, έφεραν στο φως ποικίλα μειονεκτήματα καθώς και υποπροβλήματα που μπορούν να καταστήσουν αδύνατη τη σύγκλιση του στην βέλτιστη λύση. Για την πλειονότητα των αδυναμιών έχουν προταθεί τροποποιήσεις, τις κυριότερες των οποίων αναφέρουμε στη συνέχεια. Οι τροποποιήσεις αυτές μπορούν να ομαδοποιηθούν σε πέντε κατηγορίες, ανάλογα με το σημείο της ροής στο οποίο επεμβαίνουν.

### I . Υπολογιστική πολυπλοκότητα

Με μια πιο προσεκτική ματιά κάποιος θα μπορούσε να παρατηρήσει ότι κυρίαρχο υπολογιστικό κόστος του αλγορίθμου οφείλεται στην εύρεση των κοντινότερων σημείων και την εκτίμηση του βέλτιστου μετασχηματισμού. Εκτίμηση η οποία βασίζεται στην SVD του μητρώου ετεροδιασπορών (**crosscovariance matrix**) των υποσυνόλων των αντίστοιχων σημείων. Η αναζήτηση κοντινότερου ‘γείτονα’ έχει υπολογιστικό κόστος  $\mathcal{O}(MN)$ , το υπολογιστικό κόστος της κατασκευής του μητρώου ετεροδιασπορών, από τα δείγματα, είναι  $\mathcal{O}(N^2)$  και τέλος το υπολογιστικό κόστος της SVD μπορούμε να το θεωρήσουμε αμελητέο στην περίπτωση τρισδιάστατων σημείων. Κάνοντας την υπόθεση ότι  $M \gg N$  καταλήγουμε ότι η πολυπλοκότητα του αλγορίθμου είναι της τάξης του  $\mathcal{O}(MN)$ . Ευλόγως συμπεραίνουμε τελικά ότι το σημείο συμφόρησης είναι η αναζήτηση του κοντινότερου γείτονα. Για την μείωση της πολυπλοκότητας αυτού του βήματος έχει προταθεί<sup>[10]</sup> η χρήση ενός kd-tree για την αποθήκευση και γρήγορη αναζήτηση των σημείων του μοντέλου. Το kd-tree είναι μια δομή αποθήκευσης πολυδιάστατων σημείων η βασική ιδέα της οποίας είναι η διαμέριση του χώρου (που ζούν τα σημεία) σε κελιά. Η κατασκευή του kd-tree βασίζεται στην δυαδική (binary search tree like) διαμέριση, διαδοχικά ως προς κάθε διάσταση. Με χρήση πολύπλοκων αλγορίθμων για την εύρεση μεσαίων τιμών median-finding algorithms) η κατασκευή του kd-tree απαιτεί τουλάχιστον  $\mathcal{O}(N \log N)$  βήματα. Στη συνέχεια η κάθε αναζήτηση ενός κοντινότερου γείτονα κοστίζει  $\mathcal{O}(\log N)$  βήματα



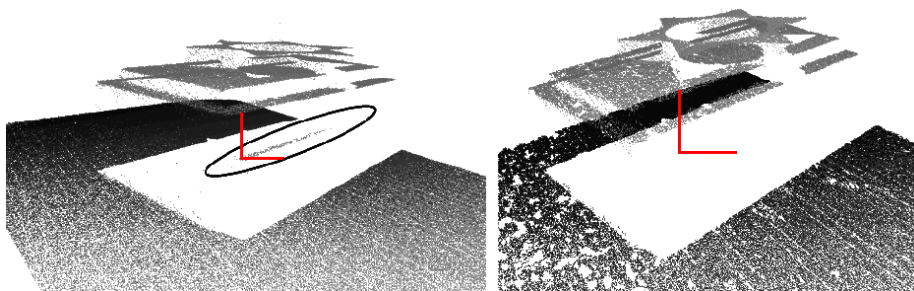
Σχήμα 4.2: Διαμέριση του χώρου με χρήση ενός τρισδιάστατου kd-tree.



στη μέση περίπτωση και  $\mathcal{O}(kN^{1-\frac{1}{k}}) = \mathcal{O}(3\sqrt{N^3})$  βήματα στην χειρότερη περίπτωση (για τρισδιάστατα σημεία). Να σημειώσουμε επίσης ότι χρήση ενός τέτοιου δέντρου συμφέρει μόνο όταν η διάσταση των σημείων είναι πολύ μικρότερη του πλήθους τους. Ένα καλό φράγμα είναι το  $N \gg 2^k$ .

## II . Φιλτράρισμα δεδομένων ειδόδου

Ίσως η πιό εμφανής αδυναμία του ICP είναι ότι εξαναγκάζει "1 – 1" αντιστοιχίες (i.e, it forces injective correspondences). Με άλλα λόγια υποχρεώνει όλα τα σημεία του μόντελου να έχουν αντιστοιχία στο άλλο σύνολο. Ο περιορισμός αυτός μπορεί να δημιουργήσει εσφαλμένες αντιστοιχίες (in the matching step). Για παράδειγμα, σε περιπτώσεις ύπαρξης ισχυρής μόλυνσης των δεδομένων ή περιπτώσεις που η επικάλυψη των επιφανειών είναι περιορισμένη η παραπάνω συνθήκη θα εξαναγκάσει σημεία που υπάρχουν μόνο στο ένα εκ των δύο συνόλων να βρουν εικονικό αντίστοιχο στο άλλο σύνολο. Για την υπέρβαση αυτής της αδυναμίας έχουν αναπτυχθεί ποικίλες τεχνικές. Για παράδειγμα το 2008 προτάθηκε από τους Shi Q. et. al<sup>[11]</sup> η χρήση του αλγορίθμου link clustering για την απομάκρυνση της μόλυνσης από τα δεδομένα εισόδου. Η εφαρμογή την οποία βελτιστοποιούν οι συγγραφείς στο<sup>[11]</sup> είναι ένα σύστημα ελέγχου ακεραιότητας της επιφάνειας μηχανικών μερών σε μονάδα παραγωγής (surface inspection system). Στο οποίο σύστημα λόγω των συσκευών σάρωσης παρατηρήθηκε ότι υπήρχαν δυο είδη μολύνσεων. Λανθασμένες μετρήσεις λόγω σκοτεινών σημείων της επιφάνειας των αντικειμένων, και outliers που εμφανίζονται στα άκρα (boundary) ανεπαρκώς φωτιζόμενων περιοχών<sup>9</sup> οι οποίοι είναι ιδιαίτερα προβληματικοί λόγω του ότι η φωτεινή πηγή είναι συνήθως σε διαφορετικό σημείο από τον σαρωτή της επιφάνειας. Η λειτουργία του αλγορίθμου link clustering διαφέρει από τα κοινά φίλτρα (median and mean filters) στο ότι προσπαθεί να αποφύγει την εξομάλυνση (smoothing) του σημειο-συνόλου. Η αρχή στην οποία βασίζεται είναι η επαναληπτική ομαδοποίηση των σημείων σε συστάδες (clusters) όπου οι ενσωμάτωση δύο συστάδων σε μία μεγαλύτερη επιτρέπεται όταν αυτές βρίσκονται σε απόσταση μικρότερη από ένα κατώφλι. Μετά το πέρας των επαναλήψεων συστάδες οι οποίες βρίσκονται "πολύ" μακριά από τις υπόλοιπες θεωρούνται μόλυνση και αποκόπτονται. Επίσης, το 2009 προτάθηκε από τους Ridene



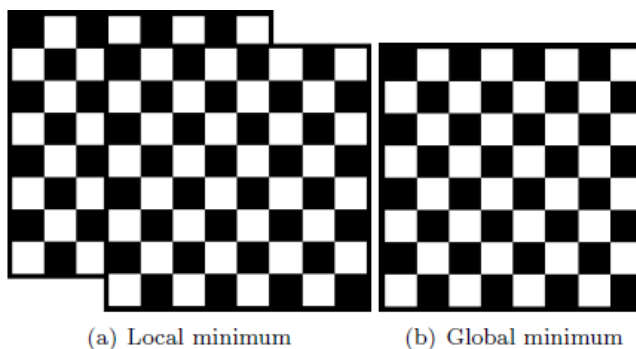
Σχήμα 4.3: link clustering example: Το μολυσμένο νέφος [αριστερά] και το ίδιο νέφος μετά την απομάκρυνση της μόλυνσης (σημεία στο εσωτερικό της έλειψης)

et. al<sup>[12]</sup> η χρήση μίας άλλης μεθόδου για την απομάκρυνση των outliers. Η ιδέα

<sup>9</sup> Παρόμοια μορφή μόλυνσης προέκυψε και στα δικά μας δεδομένα, για την διόρθωση των οποίων χρησιμοποιήθηκε φίλτρο γεωμετρικού μεσαίου (geometric median) βασισμένο στον αλγόριθμο του Endre Weiszfeld.

που πρότειναν έχει δύο υπομέρη. Το πρώτο εφαρμόζεται μία μόνο φορά στη φάση της αρχικοποίησης ενώ το δεύτερο επαναλαμβάνεται μετά από κάθε επανάληψη του ICP. Η συλλογιστική του πρώτου υπομέρους βασίζεται στην αφαίρεση ενός μεγάλου ποσοστού της μόλυνσης με χρήση του αλγορίθμου [RANSAC](#). Ο αλγόριθμος RANSAC (Random Sample Consensus) προτάθηκε από τους Fischler et. al<sup>[13]</sup> το 1981 και είναι ένας επαναληπτικός αλγόριθμος που χρησιμοποιείται για την αφαίρεση των μολυσμένων σημείων από ένα σύνολο δεδομένων βασισμένος σε ένα (συνήθως παραμετρικό) μοντέλο το οποίο περιγράφει τους inliers του συνόλου. Δηλαδή σημεία τα οποία μπορούν να έχουν 'προέλθει' από το μοντέλο. Η αρχή λειτουργίας του βασίζεται στην επαναληπτική επανεκτίμηση των παραμέτρων του μοντέλου μέσω τυχαία επιλεγμένων στοιχείων από το σύνολο, και στη συνέχεια της κατηγοριοποίησης των υπολοίπων στοιχείων-δεδομένων ως inliers ή μόλυνσης βάσει των τρεχόντων παραμέτρων. Οι εναπομένοντες outliers δεν μπορούν να αφαιρεθούν με αυτή την τεχνική διότι πλέον αποτελούν σημεία που στηρίζουν το μοντέλο. Εδώ λαμβάνει χώρα το δεύτερο υπομέρος της μεθόδου των Ridene et. al το οποίο ουσιαστικά είναι μια μέθοδος δυναμικής κατωφλίωσης της μέγιστης ακτίνας αναζήτησης αντίστοιχου σημείου (κοντινότερου γείτονα). Το κατώφλι αυτό αρχικά τίθεται ως η μέγιστη των αναλύσεων των δύο point clouds και στη συνέχεια η τιμή του ανανεώνεται ως το άθροισμα της μέσης τιμής και της τυπικής απόκλισης της τρέχουσας αντιστοίχισης των δύο νεφών. Να σημειώσουμε εδώ ότι οι συγγραφείς του<sup>[12]</sup> χρησιμοποιούν ως σύνολα σημείων αεροφωτογραφίες (2.5D data) όπου η προκύπτουσα τιμή του κατωφλίου είναι βαθμωτή ποσότητα. Στην περίπτωση τρισδιάστατων σημείων η τιμή ανανέωσης θα μπορούσε για παράδειγμα να είναι μια συνάρτηση της φασματικής ακτίνας του μητρώου ετεροδιασπορών των τρεχόντων αντίστοιχων σημείων.

### III. Υψηλή ευαισθησία στην αρχικοποίηση



Σχήμα 4.4: Εγκλωβισμός σε τοπικό ελάχιστο λόγω κακής αρχικοποίησης

Μία ακόμα αδυναμία του βασικού ICP είναι η ευαισθησία στην αρχικοποίηση των παραμέτρων του γεωμετρικού μετασχηματισμού, δηλαδή στο πόσο 'καλά' αντιστοιχισμένα είναι αρχικά τα νέφη σημείων. Η βάση του προβλήματος είναι ότι λόγω της αρχικοποίησης θα προκύψει εύρεση εσφαλμένων αντιστοιχιών η οποία τελικά θα επιφέρει εγκλωβισμό σε τοπικό ελάχιστο. Σε κάποιες περιπτώσεις η απομάκρυνση των outliers είναι αρκετή για να μπορέσει ο αλγόριθμος να συγκλίνει στη λύση, ακόμα και αν δεν έχουμε καλή αρχική εκτίμηση των παραμέτρων. Υπάρχουν όμως περιπτώσεις, όπως αυτή του Σχήματος (4.4), που μπορούν να καταστήσουν αδύνα-

τη τη σύγκλιση ακόμα και χωρίς την ύπαρξη outliers. Συνήθως η αδυναμία αυτή αποφεύγεται με ανθρώπινη επέμβαση και χειροκίνητη αντιστοίχιση των νεφών. Ως αυτοματοποιημένες λύσεις έχουν προταθεί: Αρχικοποίηση με τυχαία  $\mathbf{R}, \mathbf{t}$  και στη συνέχεια χρήση simulated annealing για απεγκλωβισμό από το τοπικό ελάχιστο ή εφαρμογή SIFT (scale invariant feature transform) για αρχικοποίηση αντιστοιχιών<sup>[15]</sup>.

#### IV. Πλεονάζουσα πληροφορία αγνοείται

Κλείνοντας αξίζει να γίνει μια αναφορά στην αξιοποίηση πλεονάζουσας πληροφορίας που μπορεί να έχουν τα δεδομένα, όπως για παράδειγμα πληροφορία χρώματος ή (ανακτασχευασμένης) επιφάνειας. Ο βασικός ICP δεν αξιοποιεί τέτοιου είδους πληροφορία. Έχουν γίνει όμως διάφορες τροποποιήσεις που αξιοποιούν τέτοιου είδους δεδομένα. Για παράδειγμα οι Bae et. al<sup>[16]</sup> εξάγουν πληροφορία επιφάνειας για να βελτιώσουν τις αντιστοιχίες των σημείων σε κάθε επανάληψη, ενώ οι Johnson et. al<sup>[14]</sup> και οι Joung et. al<sup>[15]</sup> χρησιμοποιούν υπάρχουσα πληροφορία χρώματος για τον ίδιο σκοπό. Αναλυτικότερα, το 1997 οι Johnson et. al<sup>[14]</sup> πρότειναν μια βελτίωση της αναζήτησης κοντινότερου γείτονα τροποποιώντας την μετρική Ευκλείδεια απόστασης με την εισαγωγή πληροφορίας χρώματος. Συγκεκριμένα η απόσταση δύο σημείων  $\mathbf{x}, \mathbf{y}$  υπολογίζεται ως

$$d_{color} = \| \mathbf{x} - \mathbf{y} \| + \sqrt{q_1(Y_x - Y_y)^2 + q_2(I_x - I_y)^2 + q_3(Q_x - Q_y)^2}, \quad (4.50)$$

όπου  $\| \mathbf{x} - \mathbf{y} \|$  είναι η Ευκλείδεια απόσταση των σημείων  $x$  και  $y$ , και όπου  $Y_x, I_x, Q_x$  και  $Y_y, I_y, Q_y$  είναι αντίστοιχες χρωματικές τιμές του κάθε σημείου μετατρεμμένες από το σύστημα RGB στο YIQ<sup>10</sup> ως

$$\begin{bmatrix} Y_x \\ I_x \\ Q_x \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.595716 & -0.274453 & -0.321263 \\ 0.211456 & -522591 & 0.311135 \end{bmatrix} \begin{bmatrix} R_x \\ G_x \\ B_x \end{bmatrix}. \quad (4.51)$$

Η επέκταση που έγινε από τους Joung et. al<sup>[15]</sup> στην παραπάνω τροποποίηση είναι ότι πλέον αγνοούν πλήρως σημεία με μεγάλη χρωματική διαφορά.

#### V . Επιτάχυνση σύγκλισης και ο αλγόριθμος IDC

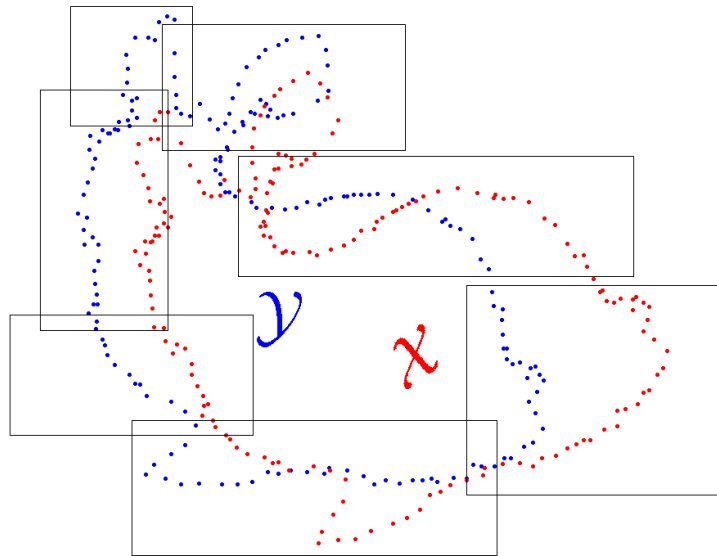
Έχει παρατηρήθει ότι το μητρώο περιστροφής  $\mathbf{R}$  του μετασχηματισμού αργεί να συγκλίνει πολύ περισσότερο από το διάνυσμα ολίσθησης  $\mathbf{t}$ . Για να επιταχύνουν τη σύγκλιση του  $\mathbf{R}$  οι Lu, Milios<sup>[18]</sup> πρότειναν την εξής τροποποίηση: Σε κάθε επανάληψη του ICP θα υπολογίζεται με τον συνήθη τρόπο ένας μετασχηματισμός  $\mathcal{T}_1 = \{\mathbf{R}_1, \mathbf{t}_1\}$  και ακολούθως θα υπολογίζεται και ένας δεύτερος μετασχηματισμός  $\mathcal{T}_2 = \{\mathbf{R}_2, \mathbf{t}_2\}$  όπου οι αντιστοιχίες για τον υπολογισμό του  $\mathcal{T}_2$  θα βρίσκονται συγκρίνοντας τις γωνιακές αποκλίσεις των σημείων  $\mathbf{x}_i, \mathbf{y}_j$  υπο πολικές συντεταγμένες. Πιο αναλυτικά η αναζήτηση κοντινότερων σημείων για κάθε  $\mathbf{x}_i$  θα γίνεται μετατρέποντας τα  $\mathcal{X}$  και  $\mathcal{Y}$  σε πολικές συντεταγμένες και ψάχνοντας το  $\mathbf{y}_j$  με την μικρότερη απόκλιση οριζόντιας και διαμήκους γωνίας (latitudinal and longitudinal angles). Στη συνέχεια ο εφαρμοζόμενος μετασχηματισμός του επόμενου βήματος προτάθηκε να κατασκευάζεται ως  $\mathcal{T} = \{\mathbf{R}_2, \mathbf{t}_1\}$ . Ιστορικά, ο τροποποιημένος αυτός αλγόριθμος ονομάζεται IDC (Iterative Dual Correspondence) με

<sup>10</sup>YIQ is the color-difference space with luma information Y and the color-differences I (orange-blue range) and Q (purple-green range).

την έννοια ότι σε κάθε βήμα υπολογίζει δύο ζεύγη αντιστοιχιών, ένα με βάση τις Ευκλείδειες και ένα με βάση τις γωνιακές αποκλίσεις των σημείων, κάθε ένα εκ των οποίων είναι υπεύθυνο για την εκτίμηση του αντιστοίχου υπομέρους του τελικού μετασχηματισμού (του τρέχοντος βήματος).

#### 4.2.2 Ο αλγόριθμος NDT

Ο αλγόριθμος **NDT (Normal Distributions Transform)** παρουσιάστηκε αρχικά από τους Biber and Straßer<sup>[19]</sup> το 2003, για την αντιστοίχιση δισδιάστατων νεφών point clouds και στη συνέχεια επεκτάθηκε στην τρισδιάστατη περίπτωση, από τον Magnusson<sup>[20]</sup> το 2006. Η καινοτομία του αλγορίθμου είναι η εισαγωγή στοχαστικών στοιχείων (**normal distributions**) στην διαδικασία εύρεσης αντιστοιχιών. Συγκεκριμένα αυτό που προσπάθησε να υπερβεί ο NDT ήταν οι 1 – 1 αντιστοιχίες που επέβαλε ο ICP καθώς και το υψηλό υπολογιστικό κόστος εύρεσης αυτών. Για να το πετύχει αυτό, αντί να χρησιμοποιεί τα σημεία ως έχουν, αναπαριστά το ένα εκ των δύο νεφών, έστω το  $\mathcal{X}$ , ως συνδιασμό από κανονικές κατανομές.



Σχήμα 4.5: Η αρχή λειτουργίας του NDT: Για την αντιστοίχιση των  $\mathcal{X}, \mathcal{Y}$ , αντί των αυτούσιων σημείων του νέφους  $\mathcal{X}$  χρησιμοποιείται μια αναπαράσταση που προέρχεται από τη διαμέριση του χώρου του  $\mathcal{X}$  (μαύρα ορθογώνια) και την εκτίμηση κατανομών από τα περιεχόμενα σημεία.

Για την κατασκευή των οποίων αρχικά διαμερίζει το χώρο που καταλαμβάνει το  $\mathcal{X}$  σε  $\Omega$  ισομεγέθους κύβους (**voxels**) και στη συνέχεια για κάθε voxel  $\omega$ ,  $\omega = 1, \dots, \Omega$  το οποίο περιέχει έναν ικανοποιητικό αριθμό σημείων υπολογίζεται το κεντροειδές

(centroid) και το μητρώο συνδιασπορών (covariance matrix) του ως

$$\mu_\omega = \frac{1}{N_\omega} \sum_{k=1}^{N_\omega} \mathbf{x}_k^\omega, \quad (4.52)$$

$$\mathbf{C}_\omega = \frac{1}{N_\omega - 1} \sum_{k=1}^{N_\omega} (\mathbf{x}_k^\omega - \mu_\omega)(\mathbf{x}_k^\omega - \mu_\omega)^T, \quad (4.53)$$

όπου  $\mathbf{x}_k^\omega, k = 1, \dots, N_\omega$  είναι σημεία του συνόλου  $\mathcal{X}$  τα οποία περιέχονται στον κύβο  $\omega$  και όπου  $N_\omega$  το πλήθος αυτών. Έχοντας υπολογίσει τα  $\mu_\omega, \Sigma_\omega$  για κάθε έναν απο τους κύβους-διαμερίσματα το νέφος  $\mathcal{X}$  μπορεί πλέον να μοντελοποιηθεί ως το σύνολο των κατανομών  $N(\mu_\omega, \Sigma_\omega)$  ορισμένων ως:

$$\mathcal{N}(\mu_\omega, \Sigma_\omega) \sim p(\mathbf{y}) = \frac{1}{2} \exp\left(-\frac{1}{2}(\mathbf{y} - \mu_\omega)^T \Sigma_\omega^{-1}(\mathbf{y} - \mu_\omega)\right). \quad (4.54)$$

Έχοντας ορίσει τις κατανομές  $N(\mu_\omega, \Sigma_\omega)$  μπορούμε να κατασκευάσουμε τη συνάρτηση κόστους  $h$  την οποία τελικά θα πρέπει να ελαχιστοποιήσουμε για να βρούμε τις παραμέτρους του μετασχηματισμού των  $\mathcal{X}, \mathcal{Y}$ . Η συνάρτηση  $h$  ορίζεται ως:

$$h(\vec{\vartheta}) = -\sum_{j=1}^M p(\mathcal{T}(\mathbf{y}_j; \vec{\vartheta})), \quad (4.55)$$

Όπου  $\mathcal{T}(\mathbf{y}_j; \vec{\vartheta})$  είναι τα σημεία  $\mathbf{y}_j$  του συνόλου  $\mathcal{Y}$  αφού τους έχει εφαρμοστεί ο μετασχηματισμός που περιγράφεται απο το διάνυσμα παραμέτρων  $\vec{\vartheta}$  και όπου  $p(\mathcal{T}(\mathbf{y}_j; \vec{\vartheta}))$  είναι η πιθανοφάνεια (likelihood probability) του σημείου  $\mathcal{T}(\mathbf{y}_j; \vec{\vartheta})$  υπολογισμένη βάση της κατανομής του voxel στο οποίο εμπεριέχεται. Επιπλέον στην περίπτωση μας το διάνυσμα  $\vec{\vartheta} = [x, y, z, \theta, t_x, t_y, t_z]^T$  περιέχει τις 7 παραμέτρους του Ευκλείδειου μετασχηματισμού, όπου τα  $(x, y, z, \theta)$  περιγράφουν, αντίστοιχα, τον άξονα και τη γωνία περιστροφής ενώ τα  $(t_x, t_y, t_z)$  ορίζουν το διάνυσμα ολίσθησης. Για την ελαχιστοποίηση της  $h(\vec{\vartheta})$  δέν υπάρχει κλειστού τύπου λύση οπότε θα πρέπει να χρησιμοποιηθεί κάποια επαναληπτική μέθοδος. Οι ερευνητές που πρότειναν τον NDT χρησιμοποίησαν την μέθοδο Newton<sup>11</sup>. Η μέθοδος αυτή δεδομένου ενός αρχικού διανύσματος παραμέτρων  $\vec{\vartheta}$  προσεγγίζει επαναληπτικά το ακρότατο της (4.55), ανανεώνοντας σε κάθε επανάληψη το  $\vec{\vartheta}$  ως  $\vec{\vartheta} \leftarrow \vec{\vartheta} + \Delta\vec{\vartheta}$ , όπου το  $\Delta\vec{\vartheta}$  υπολογίζεται απο τη λύση της

$$\mathfrak{B}\Delta\vec{\vartheta} = -\mathfrak{J}, \quad (4.56)$$

όπου  $\mathfrak{B}$  και  $\mathfrak{J}$  είναι ο Hessian και ο vectorized Jacobian matrix της συνάρτησης  $h$  αντίστοιχα. Θέτωντας, για λόγους αναγνωσιμότητας,  $\tilde{\mathbf{y}}_j = \mathcal{T}(\mathbf{y}_j; \vec{\vartheta}) - \mu_\omega$  τα στοιχεία (διανύσματα-blocks) του  $\mathbf{g}$  και του  $\mathfrak{B}$  μπορούν να υπολογιστούν ως

$$\mathbf{g}_\eta = \frac{\partial h}{\partial \vartheta_\eta} = \sum_{j=1}^M \tilde{\mathbf{y}}_j^T \Sigma_\omega^{-1} \frac{\partial \tilde{\mathbf{y}}_j}{\partial \vartheta_\eta} \exp\left(-\frac{1}{2}\tilde{\mathbf{y}}_j^T \Sigma_\omega^{-1} \tilde{\mathbf{y}}_j\right), \quad (4.57)$$

$$\mathbf{h}_{\eta\kappa} = \frac{\partial^2 h}{\partial \vartheta_\eta \partial \vartheta_\kappa} = \exp\left(-\frac{1}{2}\tilde{\mathbf{y}}_j^T \Sigma_\omega^{-1} \tilde{\mathbf{y}}_j\right) \left( \left( \tilde{\mathbf{y}}_j^T \Sigma_\omega^{-1} \frac{\partial \tilde{\mathbf{y}}_j}{\partial \vartheta_\eta} \right) \left( -\tilde{\mathbf{y}}_j^T \Sigma_\omega^{-1} \frac{\partial \tilde{\mathbf{y}}_j}{\partial \vartheta_\kappa} \right) + \tilde{\mathbf{y}}_j^T \Sigma_\omega^{-1} \frac{\partial^2 \tilde{\mathbf{y}}_j}{\partial \vartheta_\eta \partial \vartheta_\kappa} + \frac{\partial \tilde{\mathbf{y}}_j^T}{\partial \vartheta_\eta} \Sigma_\omega^{-1} \frac{\partial \tilde{\mathbf{y}}_j}{\partial \vartheta_\kappa} \right), \quad (4.58)$$

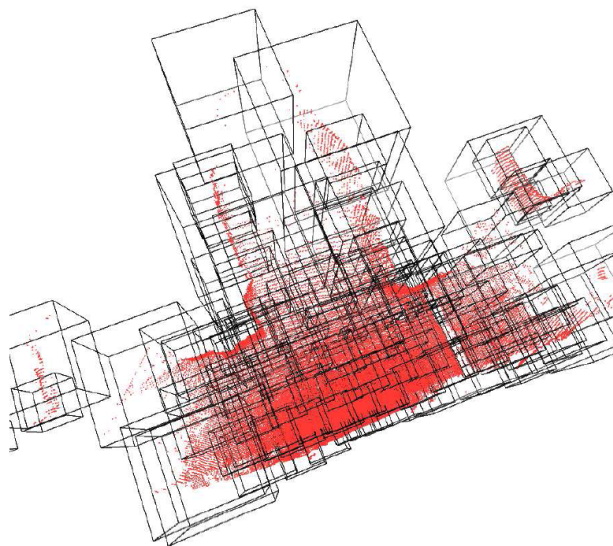
<sup>11</sup>see Newton's method in Optimization

όπου  $\vartheta_\eta, \vartheta_\kappa$  είναι στοιχεία του διανύσματος παραμέτρων  $\vec{\vartheta}$ . Παρατηρούμε ότι οι σχέσεις υπολογισμού των στοιχείων των  $\mathfrak{B}$  και  $\mathfrak{J}$  έχουν την ίδια μορφή για οποιαδήποτε διάσταση σημείων  $\mathbf{y}_j$ . Αυτό που καθορίζει την διάσταση των διανυσμάτων  $\mathbf{h}_{\eta\kappa}$  και  $\mathbf{g}_\eta$  είναι οι μερικές παράγωγοι  $\frac{\partial \tilde{\mathbf{y}}_i}{\partial \vartheta_\eta}, \frac{\partial \tilde{\mathbf{y}}_i}{\partial \vartheta_\kappa}$ . Οι οποίες στην περίπτωση τρισδιάστατων σημείων είναι διανύσματα διαστάσεων  $3 \times 1$ . Έχοντας κάνει αυτή την παρατήρηση μπορούμε να μορφοποιήσουμε τους  $\mathfrak{B}$  και  $\mathfrak{J}$  κατάλληλα ώστε να χρησιμοποιηθούν στη λύση της Σχέσης (4.56). Ο  $\mathfrak{B}$  θα είναι ένα μητρώο διαστάσεων  $21 \times 7$  ενώ το  $\mathfrak{J}$  θα είναι ένα διάνυσμα  $21 \times 1$  κατασκευασμένα αντίστοιχα ως:

$$\mathfrak{B} = \begin{bmatrix} \mathbf{h}_{11} & \dots & \mathbf{h}_{17} \\ \vdots & \ddots & \vdots \\ \mathbf{h}_{71} & \dots & \mathbf{h}_{77} \end{bmatrix}, \quad \mathfrak{J} = \begin{bmatrix} \mathbf{g}_1 \\ \vdots \\ \mathbf{g}_7 \end{bmatrix}. \quad (4.59)$$

### Συμβιβασμοί και παραμετροποιήσεις

Όπως θα παρατήρησε κάποιος, στην κατασκευή του NDT έγιναν διάφοροι συμβιβασμοί. Αρχικά οι συναρτήσεις πιθανοφάνειας που χρησιμοποιούνται στην Σχέση (4.54) δεν είναι πλήρεις κατανομές διότι απουσιάζει από αυτές ο όρος κανονικοποίησης  $1/\sqrt{(2\pi)^3|\Sigma|}$ . Όπως αποδεικνύεται στο <sup>[20]</sup> ο όρος αυτός δεν επηρεάζει την θέση των ακροτάτων της  $h$  και μπορεί να παραληφθεί. Επίσης στις σχέσεις υπολογισμού των  $\mathfrak{B}$  και  $\mathfrak{J}$  ο σταθερός όρος  $\frac{1}{2}$  παραλείπεται αφού απλοποιείται στην Σχέση (4.56). Συμπληρωματικά θα πρέπει να αναφέρουμε τι συμβαίνει στην περίπτωση που κάποια σημεία  $\mathbf{y}_j$  δεν ανήκουν σε κάποιο voxel. Σε μία τέτοια περίπτωση η τυπική ενέργεια είναι να τα αγνοήσουμε ενώ μια εναλλακτική τεχνική είναι να θεωρήσουμε ότι ανήκουν στο κοντινότερο σε αυτά voxel. Η εναλλακτική τεχνική επιφυλλάσει



Σχήμα 4.6: Διαμέριση του χώρου με αδρότητα ανάλογη με την πυκνότητα του νέφους

τον κίνδυνο, outliers που βρίσκονται συνήθως στο boundary του νέφους να επηρεάσουν την αντιστοίχιση. Για την αποφυγή παρεμφερών προβλημάτων όπως επιρροή



απο outliers ή αποδεκατισμού του νέφους εξαιτίας μεγάλου πλήθους αγνοούμενων σημείων. Η λύση που προτείνεται στο<sup>[20]</sup> είναι η επαναληπτική διαμέριση με χρήση όχι υποχρεωτικά ισομεγεθών διαμερισμάτων (voxels). Η επαναληπτική διαμέριση επαναδιαχωρίζει τον χώρο του  $\mathbf{X}$  σε κάθε επανάληψη, ξεκινώντας αρχικά απο μεγάλα παραλληλεπίπεδα και στη συνέχεια μικραίνοντας το μέγεθος και αυξάνοντας το πλήθος τους έτσι ώστε να επιτύχει εκλέπτυνση της αντιστοίχισης. Συνοψίζοντας, ο αλγόριθμος δίνεται στη συνέχεια σε μορφή ψευδοκώδικα. Ο αλγόριθμος

---

**Algorithm** Normal-Distributions-Transform (NDT)

---

**Require:** Point-set matrices  $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_N]$  and  $\mathbf{Y} = [\mathbf{y}_1, \dots, \mathbf{y}_M]$ , initial parameters vector  $\vec{\vartheta}$ , *maxNumIter*, *minimumChange*,  $\Omega$ .

- 1: Voxelize the space occupied by  $\mathbf{X}$  in  $\Omega$  voxels.
  - 2: **for**  $\omega = 1, \dots, \Omega$  **do**
  - 3:   Construct the modified normal distribution of the voxel as:  

$$\mathcal{N}(\mu_\omega, \Sigma_\omega)$$
 using equation (4.54) and (4.52).
  - 4: **end for**
  - 5: Initialize iteration counter:  $q \leftarrow 1$ .
  - 6: **repeat**
  - 7:   **for each**  $\mathbf{y}_j \in \mathbf{Y}$  **do**
  - 8:     Find the voxel  $\omega$  where  $\mathcal{T}(\mathbf{y}_j; \vec{\vartheta})$  belongs
  - 9:     Compute its likelihood  $p(\mathbf{y}_j)$  based on  $\mathcal{N}(\mu_\omega, \Sigma_\omega)$ .
  - 10:     update cost function  $\mathfrak{h}$   

$$\mathfrak{h} \leftarrow \mathfrak{h} - p(\mathbf{y}_j)$$
.
  - 11:     update Hessian matrix  $\mathfrak{B}$  based on (4.58)
  - 12:     update vectorized Jacobian  $\mathfrak{J}$  based on (4.57)
  - 13:   **end for**
  - 14:   Solve  $\mathfrak{B}\Delta\vec{\vartheta} = -\mathfrak{J}$  using Newton's or any other suitable method.
  - 15:   Update current parameters estimation:  

$$\vec{\vartheta} \leftarrow \vec{\vartheta} + \Delta\vec{\vartheta}$$
.
  - 16:   increase iteration counter:  

$$q \leftarrow q + 1$$
.
  - 17: **until**  $\Delta\vec{\vartheta} < \text{minimumChange} \parallel q > \text{maxNumIter}$ .
  - 18: **return**  $\vec{\vartheta}$
- 

NDT αποτελεί μια πρώιμη προσπάθεια εφαρμογής στατιστικών μεθόδων στην αντιστοίχιση νεφών σημείων. Η μετρική  $\mathfrak{h}$ , κάτω απο μερικές συμβάσεις, αποτελεί μια μορφή [Likelihood](#). Πλέον η μοντελοποίηση τους ενός εκ' των δύο νεφών ως σύνολο πιθανοτικών κατανομών αποτελεί μία απο τις πιο συνήθεις μεθόδους μοντελοποίησης προβλημάτων αντιστοίχισης.

### 4.2.3 Ο αλγόριθμος ECM-PR

Στην ενότητα αυτή θα παρουσιάσουμε μία πρόσφατη μέθοδο αντιστοίχισης νεφών η οποία προσπαθεί να επιλύσει το πρόβλημα χρησιμοποιώντας μείγματα Γκαουσιανών κατανομών (Gaussian Mixture Models). Ο αλγόριθμος που θα παρουσιάσουμε στο [κεφάλαιο 6] είναι στην πραγματικότητα, μια επέκταση αυτής της μεθόδου για την συμμετρική αντιστοίχιση πολλαπλών συνόλων. Η μέθοδος Expectation Conditional Maximization for Point Registration (ECM-PR) παρουσιάστηκε από τους Horaud et. al<sup>[1]</sup> το 2011 και αφορά την αντιστοίχιση αρθρωτών σωμάτων (articulated body registration). Οι δυνατές παραμορφώσεις που εξετάζει θεωρούνται ως μια συλλογή μετασχηματισμών άκαμπτου υλικού (rigid body transformations), ένας για κάθε κινητό μέρος του αρθρωτού σώματος. Στην παρούσα διπλωματική υλοποιήσαμε τον μηχανισμό εύρεσης των άκαμπτων υπομετασχηματισμών τον οποίο και παρουσιάζουμε.

Έστω τα δύο προς αντιστοίχιση σύνολα σημείων  $\mathcal{X} = \{\mathbf{x}_i\}_{1 \leq i \leq m}$  και  $\mathcal{Y} = \{\mathbf{y}_i\}_{1 \leq i \leq m}$  οργανωμένα ως μητρώα  $\mathbf{X}$  και  $\mathbf{Y}$  αντίστοιχα και έστω  $\mathbf{R}, \mathbf{t}$  ο ζητούμενος μετασχηματισμός που εφαρμοζόμενος στο  $\mathcal{X}$  δίνει το  $\mathcal{Y}$ . Η επίλυση ξεκινάει υποθέτοντας ότι τα σημεία του συνόλου  $\mathcal{X}$  αναπαριστούν τα κέντρα των συνιστώσων ενός Gaussian Mixture Model (GMM)  $\mathfrak{A}$  [Ενότητα 3.4] ενώ τα σημεία του  $\mathcal{Y}$  αποτελούν παρατηρήσεις (observations) που έχουν προέλθει από το  $\mathfrak{A}$ . Η αντιστοίχιση των  $\mathcal{X}$  και  $\mathcal{Y}$  μοντελοποιείται ως η προσαρμογή (fitting) του βέλτιστου  $\mathfrak{A}$  στα δεδομένα  $\mathcal{Y}$ . Μια τυπική κανονική κατανομή χαρακτηρίζεται από μία μέση τιμή  $\mu$  και ένα μητρώο συνδιασπορών  $\Sigma$ . Επίσης σε ένα τυπικό GMM τα κεντροειδή και τα μητρώα συνδιασπορών θεωρούνται ελεύθερες παράμετροι. Στην περίπτωση μας υποθέτουμε την ύπαρξη ενός GMM του οποίου οι συνιστώσες είναι τα  $n$  μετασχηματισμένα σημεία του  $\mathcal{X}$  τα οποία μπορούν να κινούνται ως ένα συμπαγές σώμα βάση των παραμέτρων  $\mathbf{R}, \mathbf{t}$ . Από την θεώρηση αυτή παρατηρούμε ότι τα κεντροειδή του  $\mathfrak{A}$  δεν αποτελούν πλέον ελεύθερες παραμέτρους αλλά εξαρτώνται έμμεσα από τον λανθάνοντα γεωμετρικό μετασχηματισμό, αφού κατασκευάζονται ως  $\mu_i = \mathbf{R}\mathbf{x}_i + \mathbf{t}$ . Συνοψίζοντας, μπορούμε να συνθέσουμε το σύνολο των αγνώστων παραμέτρων του  $\mathfrak{A}$  ως  $\theta = \{\mathbf{R}, \mathbf{t}, \Sigma_1, \Sigma_2, \dots, \Sigma_n\}$ . Σε αυτό το σημείο χρειαζόμαστε μια μετρική η οποία δεδομένου ενός συνόλου παραμέτρων  $\theta$  θα εκτιμά πόσο καλά περιγράφει το μοντέλο, τις παρατηρήσεις  $\mathcal{Y}$ . Η μετρική αυτή είναι γνωστή ως συνάρτηση πιθανοφάνειας (Likelihood). Η πιθανοφάνεια  $\mathcal{L}$  ενός συνόλου παραμέτρων  $\theta$  ενός στατιστικού μοντέλου δεδομένου ενός ενδεχομένου (outcome)  $x$  ορίζεται ως.  $\mathcal{L}(\theta|x) = \mathcal{P}(x|\theta)$ . Υπό την θεώρηση του  $\mathcal{Y}$  ως συνόλου ενδεχομένων-παρατηρήσεων προερχόμενων από το  $\mathfrak{A}(\theta)$  μπορούμε να ορίσουμε την Παρατηρούμενη λογαριθμική πιθανοφάνεια (Observed data Log Likelihood) του  $\theta$  δεδομένων των  $\mathcal{Y}$  ως.

$$\mathcal{L}(\theta|\mathcal{Y}) = \log(P(\mathcal{Y}; \theta)). \quad (4.60)$$

**ΣΗΜΕΙΩΣΗ:** Το σύμβολο  $|$  στη δεσμευμένη πιθανότητα του δεξιού μέλος της Σχέσης (4.60) αντικαταστάθηκε από το ; διότι το  $\theta$  δεν αποτελεί τυχαία μεταβλητή αλλά σύνολο παραμέτρων. Επίσης πρέπει να αναφέρουμε ότι η τιμή της  $\mathcal{L}$  ως τιμή δεν έχει κάποια πρακτική σημασία. Η χρησιμότητα της είναι στη σύγκριση τιμών για διαφορετικά σύνολα παραμέτρων και ενδεχομένων. Υπό αυτή την έννοια η εφαρμογή του λογαρίθμου δεν επηρεάζει τη σύγκριση αφού πρόκειται για γνησίως αύξουσα συνάρτηση αλλά από την άλλη απλοποιεί τους υπολογισμούς σε περιπτώσεις όπου οι κατανομές του στατιστικού μοντέλου προέρχονται από την εκθετική οικογένεια μέσω της μετατροπής γινομένων σε αθροίσματα.



Είναι εμφανές πλέον ότι εάν τα σύνολα  $\mathcal{X}$  και  $\mathcal{Y}$  ήταν αντιστοιχισμένα ή αντίστοιχα γνωρίζαμε τα βελτιστά  $\mathbf{R}^*$ ,  $\mathbf{t}^*$  τότε σημεία  $\mathbf{y}_j$  του  $\mathcal{Y}$  που υπάρχουν και στο  $\mathcal{X}$  θα είχαν πέσει πάνω στο αντιστοιχο σημείο τους  $\mathbf{x}_i$ . Σε μία τέτοια (υποθετική) κατάσταση η τιμή της παρατηρούμενης πιθανοφάνειας  $\mathcal{L}(\theta^*|\mathcal{Y})$  για το αντίστοιχο σύνολο παραμέτρων  $\theta^* = \{\mathbf{R}^*, \mathbf{t}^*, \Sigma_1^*, \dots, \Sigma_n^*\}$  θα ήταν υψηλότερη σε σχέση με τις τιμές άλλων συνόλων  $\theta$  αφού οι δεσμευμένες πιθανότητες  $p(\mathbf{y}_j; \theta)$  των  $\mathbf{y}_j$  είναι αντιστρόφως ανάλογες της απόστασης τους από το κεντροίδες της κλάσης<sup>12</sup> αντιστοιχίας  $\mathbf{x}_i$ . Αντιστρέφοντας τη συλλογιστική παρατηρούμε ότι η μεγιστοποίηση της παρατηρούμενης πιθανοφάνειας θα έδινε ως αποτέλεσμα τις βέλτιστες παραμέτρους  $\theta^*$  του  $\mathfrak{A}$ .

Η εύρεση του μεγίστου της  $\mathcal{L}$  είναι αδύνατη εξαιτίας της ύπαρξης αγνώστων δεδομένων, συγκεκριμένα τις αντιστοιχίες (correspondences) μεταξύ των  $\mathcal{X}$  και  $\mathcal{Y}$ . Για το λόγο αυτό θα ορίσουμε μία άλλη ποσότητα που ονομάζεται *Αναμενόμενη λογαριθμική πιθανοφάνεια πλήρων δεδομένων διεπομένων από τις παρατηρήσεις* (Expected Complete Data Log Likelihood conditioned by the observations) και την οποία θα συμβολίσουμε με  $\mathcal{E}$ . Έχει αποδειχθεί<sup>[21]</sup> ότι η μεγιστοποίηση της  $\mathcal{E}$  συνεπάγεται μεγιστοποίηση<sup>13</sup> της παρατηρούμενης πιθανοφάνειας  $\mathcal{L}$ . Οι προαναφερθείσες αντιστοιχίες θεωρούνται ελλείποντα στοιχεία (missing data) ή πιο τυπικά ως λανθάνουσες μεταβλητές στα πλαίσια της 'Μέγιστης πιθανοφάνειας' (ML framework). Πιο συγκεκριμένα κάθε σημείο  $\mathbf{y}_j$  μπορεί να κατηγοριοποιηθεί σε μία από τις  $n$  συνιστώσες του  $\mathfrak{A}$  ή σε μια κλάση μολυσμένων δεδομένων  $U$  η οποία χαρακτηρίζεται από μία ομοιόμορφη κατανομή που ορίζεται στη συνέχεια. Το σύνολο των λανθανουσών αυτών μεταβλητών θα το συμβολίζουμε  $\mathcal{Z} = \{Z_j\}_{j=[1, \dots, m]}$  όπου κάθε  $Z_j$  κωδικοποιεί την αντιστοιχία της παρατήρησης (observation)  $\mathbf{y}_j$  στην κλάση  $\mathbf{x}_i$ . Για παράδειγμα ο συμβολισμός  $Z_j = i$  ερμηνεύεται ως: *Η παρατήρηση  $\mathbf{y}_j$  κατηγοριοποιείται στην κλάση  $\mathbf{x}_i$*  ή με απλά λόγια ότι το σημείο  $\mathbf{y}_j$  είναι το μετασχηματισμένο  $\mathbf{x}_i$ . Με σκοπό να διατηρήσουμε συνεκτικότητα, πριν συνεχίσουμε θα πρέπει να ορίσουμε τα βάρη μίξης (mixing weights) των συνιστωσών του  $\mathfrak{A}$ . Οι συντελεστές αυτοί παίζουν το ρόλο της εκ των προτέρων πιθανότητας των κλάσεων  $\mathbf{x}_i$ . Εφόσον δεν έχουμε κάποια πληροφορία για την σημαντικότητα κάποιας κλάσης ανάμεσα στις υπόλοιπες θα ήταν αβιάσιμο να αναθέσουμε διαφορετικές εκ των προτέρων πιθανότητες σε κάποια εξ αυτών. Άρα ορίζουμε δύο σφαίρες, μία αρκετά ευρεία  $\mathcal{V}$  ώστε να χωρά κάθε σημείο με ακτίνα  $R$  και μία πολύ μικρότερη  $u$  με ακτίνα  $r \ll R$  με κέντρο σε ένα σημείο  $\mathbf{x}_i$ . Πιο τυπικά:

$$u = \frac{4\pi r^3}{3}, \quad V = \frac{4\pi R^3}{3}. \quad (4.61)$$

Έχοντας ορίσει τα  $u$  και  $V$  τους συντελεστές μίξης  $p_i = P(Z_j = i)$  ως:

$$p_i = \begin{cases} \frac{u}{V}, & \text{if } 1 \leq i \leq n \\ \frac{V-nu}{V}, & \text{if } i = n+1 \text{ [πιθανότητα κλάσης μολυσμένων δεδ.]} \end{cases} \quad (4.62)$$

Όπως θα δούμε στη συνέχεια χρειαζόμαστε μόνο την ακτίνα των inliers  $r$ .

Έχοντας ορίσει τις μεταβλητές κατηγοριοποίησης  $\mathcal{Z}$  καθώς και τις εκ των προτέρων πιθανότητες  $p_i$  μπορούμε πλέον να κατασκευάσουμε τυπικά την ακόλουθη

<sup>12</sup>Αφού σε μία τέτοια κατάσταση οι διασπορές των κλάσεων  $\mathbf{x}_i$  θα έτειναν στο μηδέν με σκοπό να αποδώσουν μέγιστη πιθανοφάνεια  $p(\mathbf{y}_j|\mathbf{y}_j \text{ corresponds to } \mathbf{x}_i)$  στο αντίστοιχο σημείο-ενδεχόμενο  $\mathbf{y}_j$

<sup>13</sup>Για την σχιαγράφηση της απόδειξης σύγκλισης του αλγορίθμου ECM ο αναγνώστης μπορεί να ανατρέξει στην [Ενότητα 3.4.2] ή στο βιβλίο του C.M.Bishop<sup>[28]</sup>.

συνάρτηση.

$$\mathcal{E}(\theta|\mathcal{Y}, \mathcal{Z}) = \log \mathcal{P}(\mathcal{Y}, \mathcal{Z}; \theta). \quad (4.63)$$

Η μεγιστοποίηση της  $\mathcal{E}$  είναι επίσης αδύνατη αφού δεν γνωρίζουμε την  $\mathcal{Z}$ . Για αυτό το λόγο περιθωριοποιούμε (marginalize) την μεταβλητή  $\mathcal{Z}$  εφαρμόζοντας τον τελεστή αναμενόμενης τιμής<sup>14</sup> και στην συνέχεια προσπαθούμε να μεγιστοποιήσουμε την Αναμενόμενη τιμή της  $\mathcal{E}$ :

$$\mathcal{E}_{\mathcal{Z}}(\theta|\mathcal{Y}, \mathcal{Z}) = \mathbb{E}_{\mathcal{Z}}[\log \mathcal{P}(\mathcal{Y}, \mathcal{Z}; \theta)|\mathcal{Y}]. \quad (4.64)$$

Η μεγιστοποίηση της (4.64) θα γίνει με χρήση μίας παραλλαγής του αλγορίθμου EM (Expectation Maximization) η οποία ονομάζεται ECM (Expectation Conditional Maximization) αλλά πριν εξηγήσουμε την διαδικασία επίλυσης πρέπει να μορφοποιήσουμε την  $\mathcal{E}_{\mathcal{Z}}(\mathcal{Y}, \mathcal{Z}; \theta)$  για την περίπτωση του GMM.

Αναλύοντας το όρισμα του λογαρίθμου της Σχέσης (4.64) και υποθέτοντας ότι τα  $\mathbf{y}_j$  είναι ανεξάρτητα και πανομοιότυπα καταναμημένα (i.i.d) καταλήγουμε ότι:

$$\begin{aligned} \log P(\mathcal{Y}, \mathcal{Z}; \theta) &= \log \prod_{j=1}^m P(\mathbf{y}_j, Z_j; \theta) &= \\ &= \log \prod_{j=1}^m P(\mathbf{y}_j|Z_j; \theta)P(Z_j) &= \\ &= \log \prod_{j=1}^m \prod_{i=1}^{n+1} \left( p_i P(\mathbf{y}_j|Z_j = i; \theta) \right)^{\delta(Z_j - i)}, \end{aligned} \quad (4.65)$$

όπου  $\delta(Z_j - i)$  είναι η διακριτή ακολουθία  $\delta$  ορισμένη ως:

$$\delta(Z_j - i) = \begin{cases} 1, & Z_j = i \\ 0, & \text{αλλιώς.} \end{cases} \quad (4.66)$$

**ΣΗΜΕΙΩΣΗ:** Το δεύτερο γινόμενο ( $\triangleq i$ ) είναι στην ουσία ένας μοναδικός όρος αφού η ακολουθία *kroncker* δίνει μηδενικό εκθέτη σε όλους τους άλλους όρους του γινομένου. Ο σκοπός του δεύτερου γινομένου θα γίνει εμφανής στη συνέχεια αλλά βασικά χρησιμοποιείται για να μετατρέψουμε την  $\mathcal{E}$  σε μία συνάρτηση της μορφής της [Ενότητας 4.1.4] ώστε να έχουμε κλειστού τύπου λύση για τις παραμέτρους του μετασχηματισμού. Εάν τώρα εφαρμόσουμε τον τελεστή  $\mathbb{E}_{\mathcal{Z}}[\cdot]$  στην σχέση (4.65) και χρησιμοποιήσουμε την ιδιότητα του λογαρίθμου έχουμε:

$$\mathcal{E}_{\mathcal{Z}}(\theta|\mathcal{Y}, \mathcal{Z}) = \mathbb{E}_{\mathcal{Z}} \left[ \sum_{j=1}^m \sum_{i=1}^{n+1} \delta(Z_j - i) \underbrace{(\log p_i + \log P(\mathbf{y}_j|Z_j = i; \theta))}_{\mathbf{I}} \middle| \mathcal{Y} \right]. \quad (4.67)$$

Βασίζόμενοι στο ότι το  $\mathbb{E}_{\mathcal{Z}}[\cdot]$  είναι γραμμικός τελεστής και επίσης στο ότι ο όρος  $\mathbf{I}$  είναι ανεξάρτητος<sup>15</sup> του  $\mathcal{Z}$ , η ποσότητα (4.67) απλοποιείται στην:

$$\mathcal{E}_{\mathcal{Z}}(\theta|\mathcal{Y}, \mathcal{Z}) = \sum_{j=1}^m \sum_{i=1}^{n+1} \mathbb{E}_{\mathcal{Z}} [\delta(Z_j - i)|\mathcal{Y}] \left( \log p_i + \log P(\mathbf{y}_j|Z_j = i, \theta) \right). \quad (4.68)$$

<sup>14</sup> Παρατηρούμε ότι ο τελεστής  $\mathbb{E}_{\mathcal{Z}}$  διατρέχει τις  $n+1$  τιμές της μεταβλητής  $\mathcal{Z}$  και επίσης δρά σε δεσμευμένη πιθανότητα (i.e it is a conditional expectation)

<sup>15</sup> Η τιμή της  $Z$  του όρου  $\mathbf{I}$  είναι ήδη επιλεγμένη απο το πρώτο άθροισμα οπότε ο τελεστής αναμενόμενης τιμής την 'βλέπει' ως σταθερά.

Η ποσότητα  $\mathbb{E}_{\mathcal{Z}} [\delta(Z_j - i)|\mathcal{Y}]$  μπορεί να απλοποιηθεί περαιτέρω ως:

$$\begin{aligned}
 \mathbb{E}_{\mathcal{Z}}[\delta[Z_j - i]|\mathcal{Y}] &\doteq \sum_{k=1}^{n+1} \delta [i - k] P(Z_j = k|\mathbf{y}_j) &= \\
 &P(Z_j = i|\mathbf{y}_j) &= \\
 &\frac{P(\mathbf{y}_j|Z_j = i)P(Z_j = i)}{\sum_{k=1}^{n+1} p_k P(\mathbf{y}_j|Z_j = k)} &= \\
 &\frac{\frac{u}{V} \frac{1}{\sqrt{(2\pi)^3 |\Sigma_i|}} \exp\left(-\frac{1}{2} \|\mathbf{y}_j - \mathbf{R}\mathbf{x}_i - \mathbf{t}\|_{\Sigma_i}^2\right)}{\frac{u}{V} \sum_{k=1}^n \left[ \frac{1}{\sqrt{(2\pi)^3 |\Sigma_k|}} \exp\left(-\frac{1}{2} \|\mathbf{y}_j - \mathbf{R}\mathbf{x}_k - \mathbf{t}\|_{\Sigma_k}^2\right) \right] + \frac{V-nu}{V} P(\mathbf{y}_j|Z_j = n+1)} &= \\
 &\frac{\frac{u}{V\sqrt{(2\pi)^3}} \left[ |\Sigma_i|^{-\frac{1}{2}} \exp\left(-\frac{1}{2} \|\mathbf{y}_j - \mathbf{R}\mathbf{x}_i - \mathbf{t}\|_{\Sigma_i}^2\right) \right]}{\frac{u}{V\sqrt{(2\pi)^3}} \left[ \sum_{k=1}^n |\Sigma_k|^{-\frac{1}{2}} \exp\left(-\frac{1}{2} \|\mathbf{y}_j - \mathbf{R}\mathbf{x}_k - \mathbf{t}\|_{\Sigma_k}^2\right) + \frac{V-nu}{V} \frac{1}{V} \frac{V\sqrt{(2\pi)^3}}{u} \right]} &= \\
 &\frac{|\Sigma_i|^{-\frac{1}{2}} \exp\left(-\frac{1}{2} \|\mathbf{y}_j - \mathbf{R}\mathbf{x}_i - \mathbf{t}\|_{\Sigma_i}^2\right)}{\sum_{k=1}^n \left[ |\Sigma_k|^{-\frac{1}{2}} \exp\left(-\frac{1}{2} \|\mathbf{y}_j - \mathbf{R}\mathbf{x}_k - \mathbf{t}\|_{\Sigma_k}^2\right) \right] + \frac{3(2\pi)^{\frac{3}{2}}}{4\pi r^3}} &= \\
 &\frac{|\Sigma_i|^{-\frac{1}{2}} \exp\left(-\frac{1}{2} \|\mathbf{y}_j - \mathbf{R}\mathbf{x}_i - \mathbf{t}\|_{\Sigma_i}^2\right)}{\sum_{k=1}^n \left[ |\Sigma_k|^{-\frac{1}{2}} \exp\left(-\frac{1}{2} \|\mathbf{y}_j - \mathbf{R}\mathbf{x}_k - \mathbf{t}\|_{\Sigma_k}^2\right) \right] + 1.5\sqrt{2\pi}r^{-3}} = a_{ji}. \quad (4.69)
 \end{aligned}$$

Αντικαθιστώντας τις ποσότητες  $a_{ji}$  βάσει της (4.69), καθώς και τις δεσμευμένες πιθανότητες  $P(\mathbf{y}_j|Z_j = i; \theta)$  με τις αντίστοιχες πολυμεταβλητές κανονικές κατανομές τους<sup>16</sup>, η Σχέση (4.68) γράφεται:

$$\begin{aligned}
 \mathcal{E}_{\mathcal{Z}}(\theta|\mathcal{Y}, \mathcal{Z}) &= \\
 &\sum_{j=1}^m \sum_{i=1}^n a_{ji} \left( \log(p_i) + \log\left( (2\pi)^{-\frac{3}{2}} |\Sigma_i|^{-\frac{1}{2}} \exp\left(-\frac{1}{2} \|\mathbf{y}_j - \mathbf{R}\mathbf{x}_i - \mathbf{t}\|_{\Sigma_i}^2\right) \right) \right) = \\
 &\sum_{j=1}^m \sum_{i=1}^n a_{ji} \left( \log(p_i) + \log\left( (2\pi)^{-\frac{3}{2}} - \frac{1}{2} \log |\Sigma_i| - \frac{1}{2} \|\mathbf{y}_j - \mathbf{R}\mathbf{x}_i - \mathbf{t}\|_{\Sigma_i}^2 \right) \right) = \\
 &-\frac{1}{2} \sum_{j=1}^m \sum_{i=1}^n a_{ji} \left( \log |\Sigma_i| + \|\mathbf{y}_j - \mathbf{R}\mathbf{x}_i - \mathbf{t}\|_{\Sigma_i}^2 \right) + \log(p_i) + \frac{3}{2} \log(2\pi). \quad (4.70)
 \end{aligned}$$

Οι δύο τελευταίοι όροι της παρένθεσης είναι ανεξάρτητοι των παραμέτρων  $\theta$  και μπορούν να παραληφθούν στα πλαίσια της μεγιστοποίησης του  $\mathcal{E}(\cdot)$  ως προς  $\theta$ . Εν συνεχεία η ποσότητα  $\mathcal{E}(\cdot)$  μπορεί να γραφεί απλοποιημένα ως:

$$\mathcal{E}_{\mathcal{Z}}(\theta|\mathcal{Y}, \mathcal{Z}) = -\frac{1}{2} \sum_{j=1}^m \sum_{i=1}^n a_{ji} \left[ \log |\Sigma_i| + \|\mathbf{y}_j - \mathbf{R}\mathbf{x}_i - \mathbf{t}\|_{\Sigma_i}^2 \right]. \quad (4.71)$$

<sup>16</sup>Απο την θεωρία των GMM έχουμε ότι  $\mathcal{N}(\mathbf{y}_j|\mathbf{x}_i, \Sigma_i; \theta) \doteq P(\mathbf{y}_j|Z_j = i; \theta)$

Πλέον είμαστε σε θέση να εφαρμόσουμε αλγόριθμο **ECM**. Σε μία τυπική επανάληψη του αλγορίθμου EM οι βελτιστοποιούμενες παράμετροι είναι οι κεντροειδής και τα μητρώα συνδιασπορών της κάθε συνιστώσας του  $\mathfrak{A}$ . Στην περίπτωση μας οι κεντροειδείς εξαρτώνται από τις παραμέτρους αντιστοίχισης  $\mathbf{R}, \mathbf{t}$ , οπότε η μεγιστοποίηση της (4.71) παραμένει δύσκολη. Η ιδέα που προτάθηκε από τους Radu et. al<sup>[1]</sup> είναι να χωρίσουμε τον υπολογισμό σε δύο βήματα. Στο πρώτο βήμα κάθε επανάληψης θα εκτιμούνται οι παράμετροι  $\mathbf{R}, \mathbf{t}$  χρησιμοποιώντας τις συνδιασπορές  $\Sigma_i$  της προηγούμενης επανάληψης ενώ στο δεύτερο οι συνδιασπορές θα επανεκτιμούνται χρησιμοποιώντας την τελευταία εκτίμηση των  $\mathbf{R}, \mathbf{t}$ . Αυτή η μέθοδος ονομάζεται<sup>17</sup> ECM<sup>[22]</sup> και ανήκει στην κατηγορία των γενικευμένων αλγορίθμων EM. Οι αλγόριθμοι της οικογένειας των **Generalized Expectation Maximization** δέν εγγυώνται πλέον ότι θα βρουν την μέγιστη τιμή της πιθανοφάνειας αλλά είναι αποδεδειγμένο ότι την αυξάνουν<sup>[21][23]</sup>. Για να υπολογίσουμε το μέγιστο (ελάχιστο της αντεστραμμένης (negated)) (4.71) χρησιμοποιώντας τον ECM, για το πρώτο βήμα θα θεωρήσουμε ότι οι συνδιασπορές παραμένουν σταθερές και ότι μόνο τα  $\mathbf{R}, \mathbf{t}$  μεταβάλλονται. Με αυτή τη θεώρηση η ελαχιστοποίηση της (4.71) ως προς  $\mathbf{R}, \mathbf{t}$  διατυπώνεται ως:

$$\{\mathbf{R}^*, \mathbf{t}^*\} = \arg \min_{\mathbf{R}, \mathbf{t}} \frac{1}{2} \sum_{j=1}^m \sum_{i=1}^n a_{ji} \| \mathbf{y}_j - \mathbf{R} \mathbf{x}_i - \mathbf{t} \|_{\Sigma_i}^2. \quad (4.72)$$

Η Σχέση (4.72) μπορεί να απλοποιηθεί παίρνοντας τις ακόλουθες ποσότητες:

- $\mathbf{w}_i$  τα οποία παίζουν τον ρόλο εικονικών σημείων του  $\mathcal{Y}$  ώστε το καθένα να αντιστοιχεί σε μία κλάση  $\mathbf{x}_i$ , και ορίζονται ως:

$$\mathbf{w}_i = \frac{\sum_{j=i}^m a_{ji} \mathbf{y}_j}{\sum_{j=i}^m a_{ji}}, \quad (4.73)$$

- και τα  $\lambda_i$  τα οποία παίζουν το ρόλο του ζυγίσματος της κάθε απόστασης στην συνάρτηση κόστους ή πιο αφαιρετικά το ρόλο της ανάθεσης μίας τιμής εμπιστοσύνης (confidence) σε κάθε  $\mathbf{w}_i$  ανάλογη του αθροίσματος των εκ των υστέρων πιθανοτήτων  $a_{ji}$  της κλάσης  $\mathbf{x}_i$  για όλα τα σημεία  $\mathbf{y}_j$ :

$$\lambda_i = \sum_{j=i}^m a_{ji}. \quad (4.74)$$

Αφού αντικαταστήσουμε την (4.73) και την (4.74) στην (4.72) και παραλήψουμε σταθερούς όρους, δηλαδή όρους που δέν επηρεάζουν την ελαχιστοποίηση καταλήγουμε ότι:

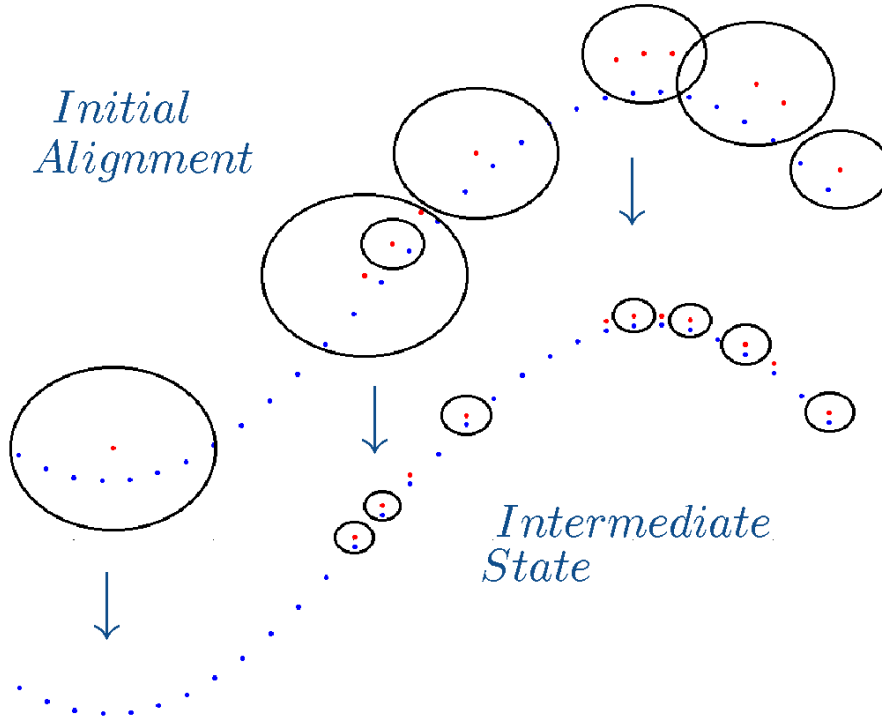
$$\{\mathbf{R}^*, \mathbf{t}^*\} = \arg \min_{\mathbf{R}, \mathbf{t}} \frac{1}{2} \sum_{i=1}^n \lambda_i \| \mathbf{w}_i - \mathbf{R} \mathbf{x}_i - \mathbf{t} \|_{\Sigma_i}^2. \quad (4.75)$$

**ΣΗΜΕΙΩΣΗ:** Στην πραγματικότητα η (4.72) και η (4.75) δέν είναι απολύτως ισοδύναμες διότι έχουν διαφορετικούς σταθερούς όρους. Παρόλα αυτά παρουσιάζουν

<sup>17</sup> Expectation Conditional Maximization με την έννοια ότι μεγιστοποιεί ένα cond.  $\mathbb{E}[\cdot]$ .

ακρότατα στα ίδια σημεία.

Η περίπτωση ισοτροπικών συνδιασπορών



Σχήμα 4.7: Η αρχή του ισοτροπικού αλγορίθμου ECM-PR: ( $\bullet \in \mathcal{X}$ ,  $\bullet \in \mathcal{Y}$ )

**Η ΑΡΧΗ ΤΟΥ ΑΛΓΟΡΙΘΜΟΥ:** Στην ισοτροπική περίπτωση κάθε σημείο  $\bullet$  του συνόλου  $\mathcal{X}$  δημιουργεί μια πολυδιάστατη κανονική κατανομή με μέση τιμή το ίδιο το σημείο και ισοτροπική διασπορά της μορφής  $\Sigma^2 = \sigma^2 \mathbf{I}$  που σχεδιάζεται με μαύρο κυκλο ακτίνας  $\sigma^2$  γύρω από το σημείο. Η κατανομές αυτές αποτελούν τις συνιστώσες του  $\mathfrak{A}$  ενώ τα  $\bullet$  σημεία  $\in \mathcal{Y}$  είναι παρατηρήσεις που έχουν προέλθει από κάποια από τις συνιστώσες-κατανομές του  $\mathfrak{A}$  και στην συνέχεια έχουν υποστεί έναν μετασχηματισμό ισομετρίας,  $\mathbf{R}, \mathbf{t}$  ο οποίος θεωρείται κοινός για όλες. Αρχικά οι διασπορές των κεντροιδών είναι μεγάλες ώστε να αποδώσουν όσο το δυνατόν ομοιόμορφες πιθανοφάνειες στα γειτονικά σημεία. Στη συνέχεια συρρικνώνονται προσπαθώντας να αποδώσουν μεγαλύτερη πιθανοφάνεια στο σημείο που έχει προέλθει από αυτές, ενώ παράλληλα οι παράμετροι του μετασχηματισμού ανανεώνονται ώστε να επιταχύνουν-βελτιώσουν αυτή τη διαδικασία. Μετά το πέρας των επαναλήψεων η διαδικασία έχει καταλήξει σε τοπικό ελάχιστο της παρατηρούμενης πιθανοφάνειας και εμμέσως έχει επιτύχει αντιστοίχιση των δύο συνόλων.

Με σκοπό να αποφύγουμε την χρήση επαναληπτικών μεθόδων για την εύρεση ακροτάτων, προσπαθούμε να καταλήξουμε σε συνάρτηση κόστους η οποία έχει κλειστού τύπου λύση. Για να το πετύχουμε κάνουμε την θεώρηση ότι τα μητρώα συνδιασπορών της κάθε συνιστώσας του  $\mathfrak{A}$  είναι ισοτροπικά, δηλαδή έχουν την μορφή  $\Sigma_i = \sigma_i^2 \mathbf{I}_3$ . Υπό αυτή τη συνθήκη η Mahalanobis απόσταση της Σχέσης (4.75)

μετατρέπεται σε βεβαρυσμένη Ευκλείδεια:

$$\{\mathbf{R}^*, \mathbf{t}^*\} = \arg \min_{\mathbf{R}, \mathbf{t}} \frac{1}{2} \sum_{i=1}^n \frac{\lambda_i}{\sigma_i^2} \|\mathbf{w}_i - \mathbf{R}\mathbf{x}_i - \mathbf{t}\|_2^2. \quad (4.76)$$

Η Σχέση (4.76) είναι ένα γενικευμένο πρόβλημα απόλυτου προσανατολισμού και μπορεί να λυθεί με την μέθοδο της [Ενότητας 4.1.4]. Σε κάθε επανάληψη, αφού βρούμε τα βέλτιστα  $\mathbf{R}^*, \mathbf{t}^*$  μπορούμε να προχωρήσουμε στο δεύτερο βήμα του ECM που είναι η επανεκτίμηση των (συν)διασπορών. Υπολογίζοντας τις μερικές παραγώγους της  $\mathcal{E}_{\mathcal{Z}}(\theta|\mathcal{Y}, \mathcal{Z})$ , σχέση (4.71), για κάθε  $\sigma_i$  εξισώνοντας με μηδέν καταλήγουμε στην σχέση επανεκτίμησης:

$$\begin{aligned} \frac{\partial \mathcal{E}(\theta)}{\partial \sigma_i} = 0 &\Rightarrow \frac{\partial \left( -\frac{1}{2} \sum_{j=1}^m a_{ji} [\log(\sigma_i^2) + \frac{1}{\sigma_i^2} \|\mathbf{y}_j - \mathbf{R}^*\mathbf{x}_i - \mathbf{t}^*\|_2^2] \right)}{\partial \sigma_i} = 0 \Rightarrow \\ &3\sigma_i^2 \sum_{j=1}^m a_{ji} = \sum_{j=1}^m a_{ji} \|\mathbf{y}_j - \mathbf{R}^*\mathbf{x}_i - \mathbf{t}^*\|_2^2 \Rightarrow \\ &\sigma_i^2 = \frac{\sum_{j=1}^m a_{ji} \|\mathbf{y}_j - \mathbf{R}^*\mathbf{x}_i - \mathbf{t}^*\|_2^2}{3 \sum_{j=1}^m a_{ji}} \quad \forall i \in [1, n]. \quad (4.77) \end{aligned}$$

Τα πειράματα αντιστοίχισης ζεύγους νεφών που παρατίθενται στο [κεφάλαιο 7] προέρχονται από την υλοποίηση αλγορίθμου ECM-PR με χρήση ισοτροπικών συνιστωσών και την σύγκριση του με τον ICP. Σε αυτό το σημείο παραθετούμε τον αλγόριθμο σε μορφή ψευδοκώδικα και επιπλέον μερικά σημεία τα οποία θα πρέπει να προσέξει ο αναγνώστης σε μία πιθανή υλοποίηση. Αρχικά ένα συχνά εμφανιζόμενο πρόβλημα είναι κάποιο σημείο  $\mathbf{y}_j$  του συνόλου παρατηρήσεων να βρεθεί στη ίδια ακριβώς θέση με κάποιο κεντροειδές  $\mathbf{x}_i$ . Σε μία τέτοια περίπτωση η Mahalanobis distance  $\|\mathbf{y}_j - \mathbf{R}\mathbf{x}_i - \mathbf{t}\|_{\Sigma_i}^2$  μηδενίζεται με αποτέλεσμα η πιθανοφάνεια του  $\mathbf{y}_j$  ως  $p(\mathbf{y}_j|\mathbf{x}_i, \Sigma_i)$  προς αυτή τη συνιστώσα να μεγιστοποιείται και τελικά ο αλγόριθμος να κλειδώνει σε τοπικό ελάχιστο. Μια διαδικασία αποφυγής αυτού του προβλήματος είναι η διαταραχή (perturbation) των σημείων του  $\mathbf{Y}$  με κάποια μορφή ασθενούς εικονικού θορύβου. Ένα άλλο πρόβλημα που μπορεί να προκύψει είναι η διασπορά  $\sigma_i$  κάποιας συνιστώσας να μηδενιστεί λόγω μεγάλης απόστασης αυτής από τις παρατηρήσεις αρα και μηδενισμό του αριθμητή της Σχέσης (4.77). Μια τέτοια κατάσταση μηδενίζει τις εκ των υστέρων πιθανότητες της συγκεκριμένης κλάσης με αποτέλεσμα να δημιουργεί NaNs κατά τον υπολογισμό των  $\mathbf{w}_i$  της Σχέσης (4.73). Για να αποφύγουμε τέτοιου είδους παθογένειες συνηθίζεται να προστίθεται μια θετική σταθερά  $\epsilon \rightarrow 0$  κατά την ανανέωση του  $\sigma_i$  σε κάθε επανάληψη. Η σταθερά αυτή μπορεί να είναι στατική ή να μεταβάλλεται με το πέρασμα των επαναλήψεων. Παρατηρήσαμε ότι μια σταθερή τιμή της τάξης του  $\epsilon = \mathcal{O}(10^{-5})$  έδινε ικανοποιητικά αποτελέσματα. Η τελευταία παράμετρος που πρέπει να θέσουμε είναι η "ακτίνα των inliers"  $r$  που υπεισέρχεται στον παρανομαστή της σχέσης (4.69). Η ποσότητα αυτή παίζει το ρόλο ζυγίσματος της κλάσης των μολυσμένων δεδομένων, πιο συγκεκριμένα εμφωλεύει το γινόμενο της εκ των προτέρων πιθανότητας  $p(n+1)$  αυτής της κλάσης επί την αντίστοιχη πιθανοφάνεια  $p(\mathbf{y}_j|Z_j = n+1)$  ενός σημείου<sup>18</sup>  $\mathbf{y}_j$ . Για την παράμετρο αυτή χρησιμοποιήσαμε

<sup>18</sup>The law of total probability

**Algorithm** Expectation-Conditional-Maximization-Point-Registration (ECMPR)

**Require:** matrices  $\mathbf{X}$  and  $\mathbf{Y}$ ,  $maxNumIter$ ,  $endeVal$ ,  $r$ , Initial parameter set  $\theta^0 = \{\mathbf{R}^0, \mathbf{t}^0, \sigma_1^0, \sigma_2^0, \dots, \sigma_n^0\}$ .

1: Initialize iteration counter :

$$q \leftarrow 1 .$$

2: **repeat**

3: compute posteriors :

$$a_{ji} = \frac{(\sigma_i^{q-1})^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\|\mathbf{y}_j - \mathbf{R}^{q-1}\mathbf{x}_i - \mathbf{t}^{q-1}\|_{\sigma_i^{q-1}}^2\right)}{\sum_{k=1}^n \left[ (\sigma_k^{q-1})^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\|\mathbf{y}_j - \mathbf{R}^{q-1}\mathbf{x}_i - \mathbf{t}^{q-1}\|_{\sigma_k^{q-1}}^2\right) \right] + 1.5\sqrt{2\pi}r^{-3}} . \quad 4.69\uparrow$$

4: Compute virtual points :

$$\mathbf{w}_i \leftarrow \frac{\sum_{j=1}^m a_{ji}\mathbf{y}_j}{\sum_{j=1}^m a_{ji}} . \quad 4.73\uparrow$$

5: Compute virtual point confidence :

$$\lambda_i \leftarrow \sum_{j=1}^m a_{ji} . \quad 4.74\uparrow$$

6: Estimate optimum :

$$\mathbf{R}^q, \mathbf{t}^q \leftarrow \arg \min_{\mathbf{R}, \mathbf{t}} \frac{1}{2} \sum_{i=1}^n \frac{\lambda_i}{\sigma_i^{q-1}} \|\mathbf{w}_i - \mathbf{R}\mathbf{x}_i - \mathbf{t}\|_2^2 . \quad 4.21\uparrow$$

7: Update covariances :

$$\sigma_i^q \leftarrow \left( \frac{\sum_{j=1}^m a_{ji} \|\mathbf{y}_j - \mathbf{R}^q\mathbf{x}_i - \mathbf{t}^q\|_2^2}{3 \sum_{j=1}^m a_{ji}} \right)^{\frac{1}{2}} . \quad 4.77\uparrow$$

8: Compute iteration error :

$$err(q) \leftarrow \|\mathbf{R}^q - \mathbf{R}^{q-1}\|_{\mathbf{F}}^2 .$$

9: increase iteration counter :

$$q \leftarrow q + 1 .$$

10: **until**  $err(q) < endeVal$  ||  $q > maxNumIter$  .

11: **return**  $\mathbf{R}^q, \mathbf{t}^q$  .

τιμές της τάξης του  $r \approx 10^0$ . Μια άλλη προσέγγιση είναι η δυναμική ανανέωση των εκ των προτέρων πιθανοτήτων των κλάσεων μέσω της θεωρίας των GMM και στη συνέχεια χρήση ενός  $r$  που πλέον θα αναπαριστά μόνο την πιθανοφάνεια της  $n + 1$  "μόλυσμένης" συνιστώσας με αποτέλεσμα να έχει μειωμένη ισχύ. Άλλες παραμετροποιήσεις που έχουν προταθεί για το συγκεκριμένο πρόβλημα, αντί του

outlier component χρησιμοποιούν δυναμικές βεβαρύνσεις των  $a_{ji}$  μέσω βοηθητικών κατανομών οι οποίες εισάγονται στο μοντέλο ως εκ των προτέρων πιθανότητες των κλάσεων και δρουν στους συντελεστές μίξης<sup>19</sup> του GMM. Συνήθως τέτοιες κατανομές επιλέγονται να είναι οι εκ των προτέρων συζυγείς<sup>20</sup> των αντιστοίχων κατανομών πιθανοφάνειας με σκοπό η εκ των υστέρων κατανομή να είναι γνωστή.

### 4.3 Σύντομη επισκόπηση μερικών μεθόδων αντιστοίχισης εύκαμπτων και αρθρωτών σωμάτων

Εως τώρα σκοπός μας ήταν να ανακτήσουμε τον βέλτιστο μετασχηματισμό ισομετρίας μεταξύ δύο συνόλων. Οδηγούμενοι από την ειδική μορφή του ακροτάτου, σε όλες τις μοντελοποιήσεις αναζητήσαμε περιορισμένα ακρότατα τα οποία περιέχονται στο πολύπτυχο μόρφωμα (manifold) των Ευκλείδειων μετασχηματισμών,  $SE(3)$ . Υπάρχουν όμως εφαρμογές στις οποίες έχουμε τη δυνατότητα να επιτρέψουμε επιπλέον βαθμούς ελευθερίας στον λανθάνοντα μετασχηματισμό των προς αντιστοίχιση συνόλων. Παραδείγματος χάριν σε μία περίπτωση αντιστοίχισης εύκαμπτων σωμάτων (nonrigid registration): η ανάκτηση ενός μετασχηματισμού συγγένειας ίσως έδινε καλύτερα αποτελέσματα από τον αντίστοιχο μετασχηματισμό που θα προέκυπτε από επίλυση με υπόθεση Ευκλείδειας παραμόρφωσης. Από μαθηματική σκοπία η μοντελοποίηση του προβλήματος με υπόθεση λανθάνοντα μετασχηματισμού συγγένειας είναι απλούστερη σε σχέση με την περίπτωση ισομετρίας. Φορμαλιστικά: υποθέτοντας ότι η μοντελοποίηση του προβλήματος καταλήγει σε βελτιστοποίηση μίας ποσότητας της μορφής της (4.21):

$$\{\mathbf{A}^*, \mathbf{t}^*\} = \arg \min_{\mathbf{A}, \mathbf{t}} \frac{1}{2} \sum_{i=1}^n b_i \|\mathbf{w}_i - \mathbf{A}\mathbf{x}_i - \mathbf{t}\|_2^2, \quad (4.78)$$

όπου  $\mathbf{A}^*$  είναι το λανθάνον μητρώο μετασχηματισμού συγγένειας και  $\mathbf{t}^*$  το λανθάνον διάλυμα ολίσθησης. Εφαρμόζοντας την μέθοδο που παρουσιάσαμε στην [Ενότητα 4.1.4] έως και τη Σχέση (4.26) η επίλυση είναι πανομοιότυπη με την αναζήτηση Ευκλείδειου μετασχηματισμού. Στην συνέχεια, η εύρεση του  $\mathbf{A}$  απλοποιείται διότι πλέον αναζητούμε ένα ακρότατο της  $\mathcal{F}$  χωρίς καποιο επιπλέον περιορισμό (unconstrained extremum). Συγκεκριμένα η προς ελαχιστοποίηση ποσότητα γράφεται:

$$\begin{aligned} \mathcal{F} &= \|\tilde{\mathbf{W}} - \mathbf{A}\tilde{\mathbf{X}}\mathbf{K}\|_{\mathbb{F}}^2 && \Rightarrow \\ \mathcal{F} &= \text{tr}\{\mathbf{K}(\tilde{\mathbf{W}}^T - \tilde{\mathbf{X}}^T\mathbf{A}^T)(\tilde{\mathbf{W}} - \mathbf{A}\tilde{\mathbf{X}}\mathbf{K})\} && \Rightarrow \\ \mathcal{F} &= \text{tr}\{\mathbf{K}\tilde{\mathbf{W}}^T\tilde{\mathbf{W}} - 2\mathbf{A}\tilde{\mathbf{X}}\mathbf{K}\tilde{\mathbf{W}}^T + \mathbf{A}\tilde{\mathbf{X}}\mathbf{K}\tilde{\mathbf{X}}^T\mathbf{A}^T\} \end{aligned} \quad (4.79)$$

Υπολογίζοντας τις μερικές παραγώγους<sup>21</sup>,  $\frac{\partial \mathcal{F}}{\partial \mathbf{A}}$ , και εξισώνοντας με 0 καταλήγουμε στην σχέση υπολογισμού του βέλτιστου μετασχηματισμού συγγένειας που συνδέει

<sup>19</sup> Στην περίπτωση μας τον ρολο των mixing coefficients παίζουν οι εκ των προτερων πιθανότητες των κλασεων  $P(Z_j = i)$ .

<sup>20</sup> Εκ των προτέρων συζυγείς (conjugate priors) είναι ζεύγη οικογενειών κατανομών οι οποίες όταν βρεθούν σε αντιστροφή Bayes, (Bayes inversion), ως prior και conditional αντίστοιχα, δίνουν posterior distribution της οικογένειας της εκ των προτέρων με δεδομένες παραμέτρους.

<sup>21</sup> Για τον τρόπο υπολογισμού του  $\frac{\partial \mathcal{F}}{\partial \mathbf{A}}$  ο αναγνώστης μπορεί να ανατρέξει στην [ενότητα ;].



τα  $\mathbf{X}$  και  $\mathbf{W}$ :

$$\frac{\partial \mathcal{F}}{\partial \mathbf{A}} = \mathbf{0}_3 \Rightarrow \mathbf{A} \tilde{\mathbf{X}} \mathbf{K} \tilde{\mathbf{X}}^T = \tilde{\mathbf{X}} \mathbf{K} \tilde{\mathbf{W}}^T \Rightarrow \mathbf{A} = \tilde{\mathbf{X}} \mathbf{K} \tilde{\mathbf{W}}^T (\tilde{\mathbf{X}} \mathbf{K} \tilde{\mathbf{X}}^T)^{-1}. \quad (4.80)$$

Στη συνέχεια το διάνυσμα ολίσθησης μπορεί να υπολογιστεί ως

$$\mathbf{t} = \tilde{\mathbf{W}} \frac{\tilde{\mathbf{h}}}{\tilde{\mathbf{h}}^T \tilde{\mathbf{h}}} - \mathbf{A} \tilde{\mathbf{X}} \frac{\tilde{\mathbf{h}}}{\tilde{\mathbf{h}}^T \tilde{\mathbf{h}}}. \quad (4.81)$$

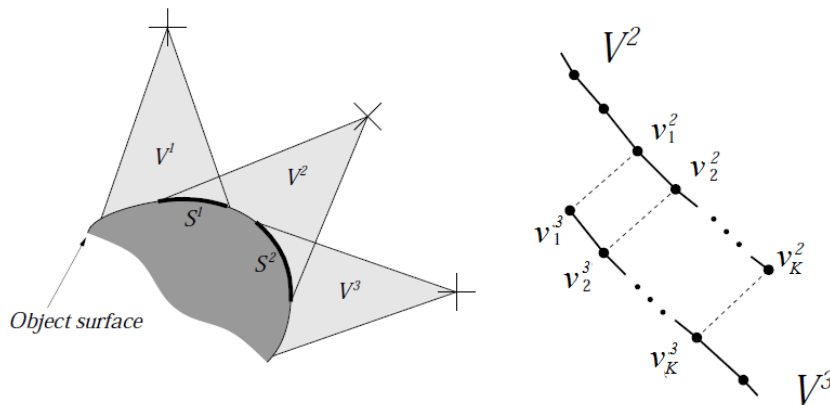
Ιστορικά, το Δεκέμβριο του 2010 οι Myronenko et.al<sup>[27]</sup> παρουσίασαν μια μέθοδο για την αντιστοίχιση ενός ζεύγους τρισδιάστατων νεφών η οποία βασιζόταν σε μία στοχαστική μοντελοποίηση του προβλήματος: πανομοιότυπη με αυτή του αλγορίθμου ECM-PR, [Ενότητα 4.2.3]. Η δημοσίευση περιλάμβανε την επίλυση του προβλήματος με υπόθεση Ευκλείδειου αλλά και **Affine** μετασχηματισμού. Οι διαφοροποιήσεις με την μέθοδο των Horaud et.al<sup>[4]</sup> είναι ότι οι μόνι έκαναν την υπόθεση ενός GMM με ιστροπικές συνιστώσες χαρακτηριζόμενες απο κοινή συνδιασπορά (**global isotropic setting**) ενώ οι δεύτεροι έλυσαν το πρόβλημα θεωρώντας γενικά μητρώα συνδιασπορών. Επίσης οι πρώτοι παρουσίασαν μια μέθοδο αντιστοίχισης 'εύκαμπτων' συνόλων (nonrigid registration) βασισμένη σε εξωτερική παραμετροποίηση των νέφων (**extrinsic parametrization**) μέσω bending surface models, ενώ οι δεύτεροι επεκτάθηκαν σε μια μέθοδο αντιστοίχισης 'αρθρωτών' νεφών. Δε θα σταθούμε στην μέθοδο εύκαμπτης αντιστοίχισης, καθώς απομακρύνεται απο τα πλαίσια της διπλωματικής, αλλά αξίζει να κάνουμε μια συνοπτική παρουσίαση της μεθόδου 'αρθρωτής' αντιστοίχισης μιας και βασίζεται σε πολλαπλές αντιστοιχίσεις άκαμπτου υλικού. Για την αντιστοίχιση 'αρθρωτών' συνόλων χρειαζόμαστε ένα κινηματικό μοντέλο που να περιγράφει τα κινητά μέρη του σώματος καθώς και τους βαθμούς ελευθερίας αυτών. Οι Horaud et.al υποθέτουν ότι τα προς αντιστοίχιση σύνολα περιγράφουν ένα σώμα το οποίο αποτελείται απο το κύριο μέρος-ρίζα (root-part) το οποίο έχει υποστεί έναν Ευκλείδειο μετασχηματισμό χαρακτηριζόμενο απο 7 (ή 6 αν χρησιμοποιηθούν γωνίες Euler) βαθμούς ελευθερίας, και κάποια τμήματα, μηχανικά συνδεδεμένα σε αυτό, τα οποία χαρακτηρίζονται από έναν, δύο ή τρεις περιστροφικούς βαθμούς ελευθερίας, ανάλογα με το πλήθος των αξόνων γύρω απο τους οποίους μπορούν να περιστρέφονται. Θεωρώντας ως δεδομένο σε ποιά κινητό μέρος του σώματος ανήκει κάθε σημείο του συνόλου κεντροειδών  $\mathcal{X}$ , η εύρεση του συνολικού 'αρθρωτού' μετασχηματισμού επιτυγχάνεται ως εξής: Αρχικά, μέσω του ECM-PR, υπολογίζεται ο ολικός Ευκλείδειος μετασχηματισμός που συνδέει τα δύο νέφη. Στη συνέχεια γίνεται κατηγοριοποίηση των σημείων του νέφους παρατηρήσεων  $\mathcal{Y}$  στα τελικά clusters του νέφους κεντροειδών. Τα σημεία που κατηγοριοποιήθηκαν ως inliers απορρίπτονται (discarded) υπο τη θεώρηση ότι ανήκουν στο κύριο μέρος (ρίζα) και ίσως και σε κάποια κινητά υπομέρη τα οποία αντιστοιχίστηκαν επαρκώς μέσω του ολικού μετασχηματισμού. Τα εναπομένοντα σημεία του  $\mathcal{Y}$ , αφού τους εφαρμοστεί ο ολικός μετασχηματισμός που υπολογίστηκε προηγουμένως, υποβάλλονται σε αντιστοίχιση με τα (γνωστά) σημεία του κάθε κινητού υπομέρους του συνόλου  $\mathcal{X}$ , αναδρομικά. Μετά το πέρας της κάθε αντιστοίχισης η διαδικασία κατηγοριοποίησης επαναλαμβάνεται και τα εναπομένοντα σημεία αντιστοιχίζονται με τα υπόλοιπα υπομέρη του  $\mathcal{X}$  διαδοχικά. Η διαδικασία σταματάει όταν επιτευχθεί αντιστοίχιση με όλα τα κινητά μέρη του  $\mathcal{X}$  ή το σύνολο των εναπομένοντων σημείων του  $\mathcal{Y}$  δέν είναι πλέον ικανό να δώσει λύση της τάξης των βαθμών ελευθερίας του τρέχοντος μετασχηματισμού.



## Κεφάλαιο 5

# ΣΥΝΔΙΑΣΜΕΝΗ ΑΝΤΙΣΤΟΙΧΙΣΗ ΠΟΛΛΑΠΛΩΝ ΝΕΦΩΝ

Ας υποθέσουμε  $M$  μερικώς επικαλυπτόμενα σύνολα-νέφη τρισδιάστατων σημείων  $V_j$ ,  $j = 1, \dots, M$  τα οποία προέρχονται από την δειγματοληψία ενός άτλαντα ο οποίος υποθέτουμε ότι αποτελεί ανακτασκευή μιας τρισδιάστατης σκηνής. Τα σύνολα αυτά από εδώ και στο εξής θα τα ονομάζουμε όψεις ([views](#)).



Σχήμα 5.1: Το πρόβλημα αντιστοίχισης πολλαπλών όψεων: Οι περιοχές  $S$  υποδηλώνουν τις μερικές επικαλύψεις μεταξύ των όψεων. (δεξιά τα  $S$  σε μεγένθυση)

Αυτό που θέλουμε να πετύχουμε είναι η συνένωση-αντιστοίχιση αυτών των όψεων σε ένα κοινό σύστημα αναφοράς το οποίο αποτελεί θεμέλιο λίθο για την ανακτασκευή του άτλαντα. Έως τώρα μελετήσαμε [Ενότητα 4] μεθόδους για την αντιστοίχιση δύο μερικώς επικαλυπτόμενων συνόλων, οι οποίες θα μπορούσαν να χρησιμοποιηθούν και για την αντιστοίχιση περισσότερων των δύο όψεων διαδοχικά. Το μειονέκτημα της ακολουθιακής αντιστοίχισης είναι η συσσώρευση και διάδοση σφαλμάτων εξαιτίας των επιμέρους αντιστοιχίσεων. Σε αυτή την ενότητα προσπαθούμε να κάνουμε μια διαφορετική προσέγγιση του προβλήματος όπου η

αντιστοιχισή θα επιτυγχάνεται ταυτόχρονα για όλες τις όψεις.

## 5.1 Εκ των προτέρων γνωστές αντιστοιχίες

Το πρόβλημα συνδιασμένης αντιστοιχισής δέν είναι τόσο εκτενώς μελετημένο όσο το πρόβλημα αντιστοιχισής ζεύγους. Παρόλα αυτά έχουν προταθεί ποικίλες μέθοδοι για την επίλυση του. Στην [Ενότητα 5.1.1] παρουσιάζουμε μία μέθοδο για την εύρεση των επιμέρους μετασχηματισμών των  $\mathcal{V}^j$  με την θεώρηση ότι οι μεταξύ τους αντιστοιχίες είναι γνωστές. Η μέθοδος αυτή ενσωματώνεται πολύ εύκολα στο πλαίσιο ενός ICP πολλαπλών όψεων [ενοτητα 5.2.1] και θα αποτελέσει το σημείο αναφοράς της απόδοσης της μεθόδου που αναπτύσσουμε στο [κεφάλαιο 6].

### 5.1.1 Απόλυτος προσανατολισμός πολλαπλών όψεων

Η μέθοδος που θα παρουσιάσουμε προτάθηκε από τους Williams et.al<sup>[26]</sup> και είναι μια συμμετρική μέθοδος ταυτόχρονου υπολογισμού των μετασχηματισμών ισομετρίας ενός συνόλου από επικαλυπτόμενες όψεις, όπου σε κάθε όψη εφαρμόζεται και ένας μετασχηματισμός. Η μέθοδος θεωρεί γνωστά τα ζεύγη επικαλυπτόμενων όψεων καθώς και της αντιστοιχίες μεταξύ τους. Επιπρόσθετα, στην ακόλουθη υποενότητα θα παρουσιάσουμε μια επέκταση του ICP η οποία βασίζεται στην επίλυση των Williams et.al και την οποία υλοποιήσαμε για σύγκριση απόδοσης της μεθόδου που παρουσιάζουμε στο [κεφάλαιο 6]. Η καινοτομία της επίλυσης των Williams et.al είναι η συμμετρικότητα, με την έννοια ότι κάθε όψη θεωρείται ότι έχει υποστεί και έναν μετασχηματισμό. Το έμμεσο πλεονέκτημα αυτού του χαρακτηριστικού είναι ότι δέν εξαναγκάζει το σύστημα αναφοράς να συμπίπτει με κάποια από τις όψεις.

#### Μια σύντομη ανασκόπηση

Για την επίλυση του προβλήματος απόλυτου προσανατολισμού πολλαπλών όψεων έχουν προταθεί και άλλες μέθοδοι οι οποίες οστόσο εφαρμόζουν παρόμοιο τρόπο επίλυσης με κάποια από τις αναφερθείσες μεθόδους. Αναφορικά, οι Benjema et.al<sup>[24]</sup> παρουσίασαν μια μέθοδο βασισμένη σε ορθομοναδιαία τετραδόνια, η λύση της οποίας είναι η επαναληπτική εφαρμογή της μεθόδου της [ενοτητας 4.1.2] σε μία συνάρτηση κόστους η οποία εξαρτάται από τα μετασχηματιζόμενα τετραδόνια όλων των όψεων εκτός μίας η οποία θεωρείται ως σύστημα αναφοράς. Σε κάθε βήμα η συνάρτηση ελαχιστοποιείται ως προς ένα εξ' αυτών θεωρώντας τα υπόλοιπα σταθερά. Μια άλλη ενδιαφέρουσα προσέγγιση προτάθηκε από τους Krishnan et.al<sup>[25]</sup>. Σε αυτή την προσέγγιση οι συγγραφείς επιλέγουν αντί της αναζήτησης ενός περιορισμένου (constrained) ακροτάτου να καταφύγουν σε βελτιστοποίηση μιας unconstrained συνάρτησης κόστους εντός του πολύπτυχου μορφώματος των μητρώων περιστροφής (Manifold of Rotations  $SO(3)$ ). Πιο επεξηγηματικά, αντί να αναζητούν το ακρότατο της συνάρτησης κόστους στον  $\mathbb{R}^3$  και στη συνέχεια να εξάγουν από αυτό, το βέλτιστο μητρώο περιστροφής  $\mathbf{R} \in SO(3)$  επιλέγουν να εφαρμόσουν αναζήτηση ακροτάτου της συνάρτησης 'κινούμενοι' μόνο εντός του  $SO(3)$ . Για να το επιτύχουν αυτό εφαρμόζουν μία τροποποιημένη μέθοδο Newton στην οποία το μητρώο ανανέωσης  $\Delta \mathbf{R}$  δέν προστίθεται πλέον στην προηγούμενη εκτίμηση  $\mathbf{R}_q$  αλλά η προηγούμενη εκτίμηση πολλαπλασιάζεται με μία εκθετική

απεικόνιση (exponential map) του μητρώου ανανέωσης  $\mathbf{R}_{q+1} = \mathbf{R}_q \exp(\Delta \mathbf{R})$ , όπου  $\mathbf{R}$  συμβολίζει το block μητρώο που προέρχεται από την συνένωση των μητρώων περιστροφής κάθε όψης<sup>1</sup>.

### Η Επίλυση των J. Williams και M. Bennamoun

Έστω οι  $M$  επικαλυπτόμενες όψεις  $\mathcal{V}_j$ ,  $j = 1, \dots, M$  και έστω ότι μεταξύ τους υπάρχουν  $P$  ζευγάρια<sup>2</sup> τα οποία δηλώνουν ποιές δύο όψεις επικαλύπτονται. Τα ζεύγη αυτά είναι οργανωμένα σε ένα  $P \times 2$  μητρώο  $\mathbf{S}$ , κάθε γραμμή  $p$  του οποίου περιέχει δυο δείκτες  $S(p, 1)$  και  $S(p, 2)$ , οι οποίοι παίρνουν τιμές από  $1, \dots, M$  και δηλώνουν δύο επικαλυπτόμενες όψεις. Επίσης για κάθε όψη  $\mathcal{V}_j$ ,  $j = 1, \dots, M$  υποθέτουμε έναν μετασχηματισμό ισομετρίας  $\mathbf{R}_j, \mathbf{t}_j$  ο οποίος εφαρμοζόμενος στα σημεία  $\mathbf{v}_i^j$  αυτής, ως  $\mathbf{R}_j \mathbf{v}_i^j + \mathbf{t}_j$  τα μεταφέρει στην αρχική τους θέση πάνω στον άτλαντα. Με άλλα λόγια αντιστοιχίζει την όψη αυτή με τις υπόλοιπες  $M - 1$ . Οι  $M$  αυτοί μετασχηματισμοί είναι οργανωμένοι σε ένα  $3 \times 3M$  block μητρώο περιστροφών  $\mathbf{R}$  και ένα  $3M \times 1$  block διάνυσμα  $\mathbf{T}$  ολισθήσεων κατασκευασμένα ως

$$\mathbf{R} = [\mathbf{R}_1 \quad \mathbf{R}_2 \quad \dots \quad \mathbf{R}_M], \quad (5.1)$$

και ως

$$\mathbf{T} = [\mathbf{t}_1^T \quad \mathbf{t}_2^T \quad \dots \quad \mathbf{t}_M^T]^T, \quad (5.2)$$

αντίστοιχα. Επιπλέον υποθέτουμε ότι τα αντίστοιχα σημεία μεταξύ των όψεων κάθε ζεύγους  $p$  είναι δεδομένα σε ένα προς ένα αντιστοιχία, ως στήλες, στα  $3 \times N^p$  μητρώα  $\mathbf{X}^p$  και  $\mathbf{Y}^p$ , όπου το κάθε μητρώο  $\mathbf{X}^p$  και  $\mathbf{Y}^p$  περιέχει ένα υποσύνολο  $N^p$  των σημείων της όψης  $\mathcal{V}^{S(p,1)}$  και  $\mathcal{V}^{S(p,2)}$  αντίστοιχα. Το πλήθος  $N^p$  των αντιστοιχίων σημείων κάθε ζεύγους μπορεί να είναι διαφορετικό από άλλα ζεύγη.

### Μαθηματική τεκμηρίωση της μεθόδου

Η συνάρτηση κόστους  $\Phi$  που θεωρούν οι Williams et. al έχει την γνωστή μορφή άθροισματος τετραγώνων Ευκλειδίων αποστάσεων των αντιστοιχίων σημείων ως προς όλα τα ζεύγη όψεων του  $\mathbf{S}$ . Πιο φορμαλιστικά γράφεται ως

$$\Phi = \sum_{p=1}^P \sum_{i=1}^{N^p} b_i^p \left\| (\mathbf{R}_{S(p,1)} \mathbf{x}_i^p + \mathbf{t}_{S(p,1)}) - (\mathbf{R}_{S(p,2)} \mathbf{y}_i^p + \mathbf{t}_{S(p,2)}) \right\|_2^2, \quad (5.3)$$

όπου  $\mathbf{x}_i^p, \mathbf{y}_i^p$  είναι σημεία στήλες των μητρώων  $\mathbf{X}^p, \mathbf{Y}^p$  αντίστοιχα και  $b_i^p$  είναι βαθμωτά βάρη για κάθε απόσταση κάθε ζεύγους. Αναπτύσσοντας την νορμισμένη ποσότητα μπορούμε να επαναμορφοποιήσουμε την  $\Phi$  ως άθροισμα δύο ποσοτήτων όπου η μία εξαρτάται από τα  $\mathbf{R}, \mathbf{T}$  ενώ η δεύτερη μόνο από το μητρώο  $\mathbf{R}$ .

$$\begin{aligned} \Phi = & \sum_{p=1}^P \sum_{i=1}^{N^p} b_i^p \left( \mathbf{t}_{S(p,1)} - \mathbf{t}_{S(p,2)} \right)^T \left[ 2 \left( \mathbf{R}_{S(p,1)} \mathbf{x}_i^p - \mathbf{R}_{S(p,2)} \mathbf{y}_i^p \right) \right. \\ & \left. + \left( \mathbf{t}_{S(p,1)} - \mathbf{t}_{S(p,2)} \right)^T \right] + \sum_{p=1}^P \sum_{i=1}^{N^p} b_i^p \left\| \mathbf{R}_{S(p,1)} \mathbf{x}_i^p - \mathbf{R}_{S(p,2)} \mathbf{y}_i^p \right\|_2^2. \quad (5.4) \end{aligned}$$

<sup>1</sup> Η περιγραφή της μεθόδου δέν είναι, μαθηματικά, απόλυτα ορθή· σκοπός της παραγράφου είναι να δώσει στον αναγνώστη μια διαισθητική επεξήγηση.

<sup>2</sup> Ζευγάρια που εμφανίζονται περισσότερες από μία φορές δεν επηρεάζουν την επίλυση. Επίσης ζευγάρια που δέν περιέχουν ουσιαστική πληροφορία, για παράδειγμα το  $[j \ j]$ , ότι δηλαδή μία όψη έχει επικάλυψη με τον εαυτό της, δέν επηρεάζουν την επίλυση.

Φορμαλιστικά οι δύο αυτές ποσότητες θα συμβολίζονται ως  $\Phi_{\mathbf{R},\mathbf{T}}$  και  $\Phi_{\mathbf{R}}$ , όπου

$$\Phi_{\mathbf{R}} = \sum_{p=1}^P \sum_{i=1}^{N^p} b_i^p \|\mathbf{R}_{\mathbf{S}(p,1)} \mathbf{x}_i^p - \mathbf{R}_{\mathbf{S}(p,2)} \mathbf{y}_i^p\|_2^2, \quad (5.5)$$

$$\Phi_{\mathbf{R},\mathbf{t}} = \sum_{p=1}^P \sum_{i=1}^{N^p} b_i^p (\mathbf{t}_{\mathbf{S}(p,1)} - \mathbf{t}_{\mathbf{S}(p,2)})^T \left[ 2(\mathbf{R}_{\mathbf{S}(p,1)} \mathbf{x}_i^p - \mathbf{R}_{\mathbf{S}(p,2)} \mathbf{y}_i^p) + (\mathbf{t}_{\mathbf{S}(p,1)} - \mathbf{t}_{\mathbf{S}(p,2)})^T \right]. \quad (5.6)$$

Στη συνέχεια, ορίζοντας ως  $w^p$  το άθροισμα των βαρών  $b_i^p$

$$w^p = \sum_{i=1}^{N^p} b_i^p, \quad (5.7)$$

και ως  $\bar{\mathbf{x}}^p, \bar{\mathbf{y}}^p$  τους σταθμισμένους μέσους των μητρώων  $\mathbf{X}^p, \mathbf{Y}^p$ , αντίστοιχα ως

$$\bar{\mathbf{x}}^p = \frac{\sum_{i=1}^{N^p} b_i^p \mathbf{x}_i^p}{w^p} \quad \text{και} \quad \bar{\mathbf{y}}^p = \frac{\sum_{i=1}^{N^p} b_i^p \mathbf{y}_i^p}{w^p} \quad (5.8)$$

υπολογισμένους για κάθε ζεύγος  $p \in [1, P]$ , η σχέση (5.6) μπορεί να γραφεί απλοποιημένα ως:

$$\begin{aligned} \Phi_{\mathbf{R},\mathbf{T}} &= \sum_{p=1}^P w^p (\mathbf{t}_{\mathbf{S}(p,1)} - \mathbf{t}_{\mathbf{S}(p,2)})^T (\mathbf{t}_{\mathbf{S}(p,1)} - \mathbf{t}_{\mathbf{S}(p,2)}) \\ &\quad + 2 \sum_{p=1}^P w^p (\mathbf{t}_{\mathbf{S}(p,1)} - \mathbf{t}_{\mathbf{S}(p,2)}) (\mathbf{R}_{\mathbf{S}(p,1)} \bar{\mathbf{x}}^p - \mathbf{R}_{\mathbf{S}(p,2)} \bar{\mathbf{y}}^p). \end{aligned} \quad (5.9)$$

### Επίλυση ως προς $\mathbf{T}$

Σε αυτό το σημείο θα πρέπει να ορίσουμε τα μητρώα  $\mathbf{C}^x$  και  $\mathbf{C}^y$  και  $\mathbf{W}$ . Τα μητρώα  $\mathbf{C}^x, \mathbf{C}^y$  έχουν διαστάσεις  $3M \times 3P$  και χρησιμοποιούνται για την επιλογή του αντίστοιχου μετασχηματισμού-block από τα  $\mathbf{R}$  και  $\mathbf{t}$  για κάθε όψη, κάθε ζεύγος του  $\mathbf{S}$ . Η δομή αυτών των μητρώων είναι η εξής. Κάθε  $3 \times 3P$  block-γραμμή των  $\mathbf{C}^x, \mathbf{C}^y$  αποτελείται από  $3 \times 3$  μητρώα μηδενικών  $\mathbf{0}_3$  και ταυτοτικά μητρώα  $\mathbf{I}_3$ . Διαισθητικά, κάθε block-γραμμή των  $\mathbf{C}^x$  και  $\mathbf{C}^y$  περιγράφει τα ζευγη στα οποία συμμετεχει μία όψη. Κατασκευαστικά, στην  $j$ -οστή block-γραμμή του  $\mathbf{C}^x$  τα ταυτοτικά μητρώα  $\mathbf{I}_3$  βρίσκονται στις θέσεις<sup>3</sup> των ζεύγων στις οποίες η  $j$ -οστή όψη παίρνει μέρος ως μητρώο  $\mathbf{X}$  ενώ στην  $j$ -οστή block-γραμμή του  $\mathbf{C}^y$  τα ταυτοτικά μητρώα βρίσκονται στις θέσεις των ζεύγων στα οποία η  $j$ -οστή όψη παίρνει μέρος ως μητρώο  $\mathbf{Y}$ . Για παράδειγμα, για την περίπτωση του (Σχήματος 5.1) το μητρώο  $\mathbf{S}$  θα μπορούσε να ήταν το

$$\mathbf{S} = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \quad (5.10)$$

<sup>3</sup>Ως θέση εννοούμε μια τριάδα στηλών. Για παράδειγμα η θέση του ζεύγους  $p$  είναι οι στήλες  $3(p-1)+1, 3(p-1)+2, 3(p-1)+3$ .

Για το οποίο τα  $\mathbf{C}^x$  και  $\mathbf{C}^y$  θα ήταν τα

$$\mathbf{C}^x = \begin{bmatrix} \mathbf{I}_3 & \mathbf{0}_3 \\ \mathbf{0}_3 & \mathbf{I}_3 \\ \mathbf{0}_3 & \mathbf{0}_3 \end{bmatrix}, \quad \mathbf{C}^y = \begin{bmatrix} \mathbf{0}_3 & \mathbf{0}_3 \\ \mathbf{I}_3 & \mathbf{0}_3 \\ \mathbf{0}_3 & \mathbf{I}_3 \end{bmatrix}. \quad (5.11)$$

Το μητρώο  $\mathbf{W}$  είναι ένα  $3P \times 3P$  διαγώνιο μητρώο κατασκευασμένο ως

$$\bigoplus_{p=1}^P w^p \mathbf{I}_3, \quad (5.12)$$

όπου το σύμβολο  $\bigoplus$  δηλώνει το κατευθύνον άθροισμα (direct sum)<sup>4</sup>. Έχοντας ορίσει αυτές τις ποσότητες η συνάρτηση  $\Phi_{\mathbf{R}, \mathbf{T}}$  μπορεί να γραφεί σε μητρική μορφή ως

$$\Phi_{\mathbf{R}, \mathbf{T}} = \mathbf{T}^T (\mathbf{C}^x - \mathbf{C}^y) \mathbf{W} (\mathbf{C}^x - \mathbf{C}^y)^T + 2\mathbf{T} (\mathbf{C}^x - \mathbf{C}^y) \mathbf{W} \mathbf{z}, \quad (5.13)$$

όπου το διάνυσμα  $\mathbf{z}$  κατασκευάζεται ως

$$\mathbf{z} = \begin{bmatrix} \mathbf{R}_{\mathbf{S}(1,1)} \bar{\mathbf{x}}^1 - \mathbf{R}_{\mathbf{S}(1,2)} \bar{\mathbf{y}}^1 \\ \mathbf{R}_{\mathbf{S}(2,1)} \bar{\mathbf{x}}^2 - \mathbf{R}_{\mathbf{S}(2,2)} \bar{\mathbf{y}}^2 \\ \vdots \\ \mathbf{R}_{\mathbf{S}(P,1)} \bar{\mathbf{x}}^P - \mathbf{R}_{\mathbf{S}(P,2)} \bar{\mathbf{y}}^P \end{bmatrix}. \quad (5.14)$$

Απο την ελαχιστοποίηση της  $\Phi(\mathbf{R}, \mathbf{T})$  (άρα και της  $\Phi$ ) ως προς  $\mathbf{T}$  προκύπτει ότι το βέλτιστο  $\mathbf{T}$  υπολογίζεται, συναρτήσει του  $\mathbf{R}$ , ως

$$\mathbf{T} = - [(\mathbf{C}^x - \mathbf{C}^y) \mathbf{W} (\mathbf{C}^x - \mathbf{C}^y)^T]^\# (\mathbf{C}^x - \mathbf{C}^y) \mathbf{W} \mathbf{z}, \quad (5.15)$$

όπου το σύμβολο  $\#$  δηλώνει τον ψευδοαντίστροφο ενός μητρώου<sup>5</sup> (matrix pseudoinverse) υπο τον ορισμό των Moore-Penrose.

#### Ανάπτυξη της $\Phi_{\mathbf{R}, \mathbf{T}}$

Απο την ελαχιστοποίηση της (5.13) προκύπτει επίσης ότι η ελάχιστη τιμή της  $\Phi_{\mathbf{R}, \mathbf{T}}$  είναι η

$$\Phi_{\mathbf{R}, \mathbf{T}_{min}} = -\mathbf{z} \mathbf{\Gamma} \mathbf{z}, \quad (5.16)$$

όπου το  $\mathbf{\Gamma}$  είναι ένα  $3P \times 3P$  μητρώο ορισμένο ως

$$\mathbf{\Gamma} = \mathbf{W} (\mathbf{C}^x - \mathbf{C}^y)^T [(\mathbf{C}^x - \mathbf{C}^y) \mathbf{W} (\mathbf{C}^x - \mathbf{C}^y)^T]^\# (\mathbf{C}^x - \mathbf{C}^y) \mathbf{W}. \quad (5.17)$$

**ΣΗΜΕΙΩΣΗ:** Το μητρώο  $\mathbf{\Gamma}$  είναι σταθερό ως προς  $\mathbf{R}$  και επίσης έχει block δομή αφού κάθε  $3 \times 3$  ορθογώνιος είναι ένα διαγώνιο μητρώο *Toeplitz*.

Η ποσότητα  $\Phi_{\mathbf{R}, \mathbf{T}_{min}}$ , σχέση (5.16), ως τετραγωνική μορφή (quadratic form) μπορεί να γραφεί σε μορφή ίχνους

$$\Phi_{\mathbf{R}, \mathbf{T}_{min}} = \text{tr}(\mathbf{R} \mathbf{Q}_{\mathbf{R}, \mathbf{T}} \mathbf{R}), \quad (5.18)$$

όπου το  $3M \times 3M$  μητρώο  $\mathbf{Q}_{\mathbf{R}, \mathbf{T}}$  προέρχεται απο την υπέρθεση

$$\mathbf{Q}_{\mathbf{R}, \mathbf{T}} = -\mathbf{Q}_{\mathbf{R}, \mathbf{T}}^{xx} + \mathbf{Q}_{\mathbf{R}, \mathbf{T}}^{xy} + \mathbf{Q}_{\mathbf{R}, \mathbf{T}}^{yx} - \mathbf{Q}_{\mathbf{R}, \mathbf{T}}^{yy}. \quad (5.19)$$

<sup>4</sup> Ονομάζεται επίσης και άθροισμα kronecker.

<sup>5</sup> Αν ο ίδιος μετασχηματισμός εφαρμοστεί σε όλες τις όψεις το μητρώο  $(\mathbf{C}^x - \mathbf{C}^y) \mathbf{W} (\mathbf{C}^x - \mathbf{C}^y)^T$  δεν έχει πλήρη τάξη και για αυτό το λόγο καταφεύγουμε στη χρήση του ψευδοαντίστροφου.

Τα μητρώα  $\mathbf{Q}_{\mathbf{R},\mathbf{T}}^{xx}$ ,  $\mathbf{Q}_{\mathbf{R},\mathbf{T}}^{xy}$ ,  $\mathbf{Q}_{\mathbf{R},\mathbf{T}}^{yx}$ ,  $\mathbf{Q}_{\mathbf{R},\mathbf{T}}^{yy}$ , της οποίας, κατασκευάζονται αντίστοιχα ως

$$\begin{aligned}\mathbf{Q}_{\mathbf{R},\mathbf{T}}^{xx} &= \mathbf{C}^x \mathbf{G}^{xx} \mathbf{C}^{xT}, & \mathbf{Q}_{\mathbf{R},\mathbf{T}}^{xy} &= \mathbf{C}^x \mathbf{G}^{xy} \mathbf{C}^{yT}, \\ \mathbf{Q}_{\mathbf{R},\mathbf{T}}^{yx} &= \mathbf{C}^y \mathbf{G}^{yx} \mathbf{C}^{xT}, & \mathbf{Q}_{\mathbf{R},\mathbf{T}}^{yy} &= \mathbf{C}^y \mathbf{G}^{yy} \mathbf{C}^{yT}.\end{aligned}\quad (5.20)$$

Τα  $\mathbf{G}^{xx}$ ,  $\mathbf{G}^{xy}$ ,  $\mathbf{G}^{yx}$ ,  $\mathbf{G}^{yy}$ , με τη σειρά τους κατασκευάζονται<sup>6</sup> ως

$$\begin{aligned}\mathbf{G}^{xx} &= (\mathfrak{X}_{3P} \mathbf{\Gamma}) \circ \vec{\mathbf{x}}\vec{\mathbf{x}}^T, & \mathbf{G}^{xy} &= (\mathfrak{X}_{3P} \mathbf{\Gamma}) \circ \vec{\mathbf{x}}\vec{\mathbf{y}}^T \\ \mathbf{G}^{yx} &= (\mathfrak{X}_{3P} \mathbf{\Gamma}) \circ \vec{\mathbf{y}}\vec{\mathbf{x}}^T, & \mathbf{G}^{yy} &= (\mathfrak{X}_{3P} \mathbf{\Gamma}) \circ \vec{\mathbf{y}}\vec{\mathbf{y}}^T,\end{aligned}\quad (5.21)$$

όπου το  $\mathfrak{X}_{3P}$  δηλώνει ένα  $3P \times 3P$  block διαγώνιο μητρώο με εναδιαίτους<sup>7</sup> ορμαθούς κατασκευασμένο ως

$$\mathfrak{X}_{3P} = \bigoplus_{p=1}^P \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \quad (5.22)$$

ενώ τα  $\vec{\mathbf{x}}$  και  $\vec{\mathbf{y}}$  είναι  $3P \times 1$  διανύσματα που προέρχονται απο τη συνένωση (concatenation) των μέσων κάθε ζεύγους όπως υπολογίστηκαν απο τη σχέση (5.8):

$$\vec{\mathbf{x}} = \begin{bmatrix} \vec{\mathbf{x}}^1 \\ \vec{\mathbf{x}}^2 \\ \vdots \\ \vec{\mathbf{x}}^P \end{bmatrix}, \quad \vec{\mathbf{y}} = \begin{bmatrix} \vec{\mathbf{y}}^1 \\ \vec{\mathbf{y}}^2 \\ \vdots \\ \vec{\mathbf{y}}^P \end{bmatrix}. \quad (5.23)$$

Τέλος, να σημειώσουμε οτι το σύμβολο  $\circ$  δηλώνει πολλαπλασιασμό Hadamard<sup>8</sup>.

### Ανάπτυξη της $\Phi_{\mathbf{R}}$

Η  $\Phi_{\mathbf{R}}$  αποτελεί επίσης τετραγωνική μορφή και μπορεί να γραφεί ως ίχνος

$$\Phi_{\mathbf{R}} = \text{tr}(\mathbf{R}\mathbf{Q}_{\mathbf{R}}\mathbf{R}^T). \quad (5.24)$$

Το  $3M \times 3M$  μητρώο  $\mathbf{Q}_{\mathbf{R}}$  αποτελεί επίσης υπέρθεση επιμέρους μητρώων. Συγκεκριμένα κατασκευάζεται ως

$$\mathbf{Q}_{\mathbf{R}} = \mathbf{C}^x \mathbf{H}^{xx} \mathbf{C}^{xT} - \mathbf{C}^x \mathbf{H}^{xy} \mathbf{C}^{yT} - \mathbf{C}^y \mathbf{H}^{yx} \mathbf{C}^{xT} + \mathbf{C}^y \mathbf{H}^{yy} \mathbf{C}^{yT}. \quad (5.25)$$

Τα μητρώα  $\mathbf{H}^{xx}$ ,  $\mathbf{H}^{yy}$ ,  $\mathbf{H}^{xy}$ ,  $\mathbf{H}^{yx}$  αποτελούν block διαγώνια μητρώα συνδιασπορών και ετεροδιασπορών. Το κάθε διαγώνιο block καθένός εξ αυτών αφορά ένα ζεύγος όψεων και υπολογίζεται απο το εξωτερικό γινόμενο των μητρώων  $\mathbf{X}^p$  ή  $\mathbf{Y}^p$  αυτού του ζεύγους σταθμισμένο με τα αντίστοιχα βάρη  $b_i^p$  των αποστάσεων του ζεύγους.

<sup>6</sup>Η παρένθεση εισάγεται λόγω της μη προαιρετικότητας του πολλαπλασιασμού Hadamard.

<sup>7</sup>Μητρώο αποτελούμενο απο μονάδες

<sup>8</sup>Πολλαπλασιασμός στοιχείου με στοιχείο.



Πιο φορμαλιστικά τα μητρώα αυτά κατασκευάζονται ως

$$\begin{aligned}\mathbf{H}^{xx} &= \bigoplus_{p=1}^M [\mathbf{X}^p \mathbf{B} \mathbf{X}^{pT}], \\ \mathbf{H}^{xy} &= \bigoplus_{p=1}^M [\mathbf{X}^p \mathbf{B} \mathbf{Y}^{pT}], \\ \mathbf{H}^{yx} &= \bigoplus_{p=1}^M [\mathbf{Y}^p \mathbf{B} \mathbf{X}^{pT}], \\ \mathbf{H}^{yy} &= \bigoplus_{p=1}^M [\mathbf{Y}^p \mathbf{B} \mathbf{Y}^{pT}],\end{aligned}\quad (5.26)$$

όπου  $\mathbf{B}$  είναι ένα διαγώνιο μητρώο που περιέχει τα βάρη  $b_i^p$  των αποστάσεων των αντιστοίχων σημείων κάθε ζεύγους και κατασκευάζεται<sup>9</sup> ως

$$\mathbf{B} = \bigoplus_{i=1}^{N^p} b_i^p \quad (5.27)$$

**Επαναληπτική επίλυση της  $\Phi$  ως προς  $\mathbf{R}$**

Έχοντας ορίσει τα  $\mathbf{Q}_{\mathbf{R},\mathbf{T}}$  και  $\mathbf{Q}_{\mathbf{R}}$  η ελαχιστοποίηση της (5.3) μπορεί να επανατεθεί ως

$$\mathbf{R}^* = \arg \min_{\mathbf{R}} \text{tr}(\mathbf{R} \mathbf{Q} \mathbf{R}^T), \quad (5.28)$$

όπου

$$\mathbf{Q} = \mathbf{Q}_{\mathbf{R}} + \mathbf{Q}_{\mathbf{R},\mathbf{T}}. \quad (5.29)$$

Διαμερίζοντας<sup>10</sup> το  $3M \times 3M$  μητρώο  $\mathbf{Q}$  ως

$$\mathbf{Q} = \begin{bmatrix} \mathbf{Q}_{ul} & \mathbf{Q}_{uj} & \mathbf{Q}_{ur} \\ \mathbf{Q}_{jl} & \mathbf{Q}_{jj} & \mathbf{Q}_{jr} \\ \mathbf{Q}_{dl} & \mathbf{Q}_{dj} & \mathbf{Q}_{dr} \end{bmatrix}, \quad (5.30)$$

και στη συνέχεια το  $\mathbf{R}$  ως

$$\mathbf{R} = [\mathbf{R}_l \quad \mathbf{R}_j \quad \mathbf{R}_r] \quad \forall j \in [1, M], \quad (5.31)$$

μπορούμε να μορφοποιήσουμε την προς ελαχιστοποίηση ποσότητα της σχέσης (5.28) ως

$$\begin{aligned} & \text{tr}(\mathbf{R}_l \mathbf{Q}_{ul} \mathbf{R}_l^T) + \text{tr}(\mathbf{R}_r \mathbf{Q}_{ur} \mathbf{R}_r^T) + 2\text{tr}(\mathbf{R}_r \mathbf{Q}_{dl} \mathbf{R}_l^T) \\ & + \text{tr}(\mathbf{R}_j \mathbf{Q}_{jj} \mathbf{R}_j^T) + 2\text{tr}(\mathbf{R}_j [\mathbf{Q}_{jl} \mathbf{R}_l^T + \mathbf{Q}_{jr} \mathbf{R}_r^T]). \end{aligned} \quad (5.32)$$

Η ταυτόχρονη βελτιστοποίηση ολόκληρου του μητρώου  $\mathbf{R}$  δέν είναι εύκολη, για αυτό το λόγο θα καταφύγουμε σε επαναληπτική επίλυση (**alternation**) όπου σε κάθε βήμα θα βελτιστοποιούμε ένα συγκεκριμένο  $3 \times 3$  block  $\mathbf{R}_j$  του  $\mathbf{R}$  θεωρώντας

<sup>9</sup>Το κατευθύνον άθροισμα με βαθμωτά ορίσματα (συγκεκριμένα τα  $b_i^p$ ) δημιουργεί το διαγώνιο μητρώο αυτών.

<sup>10</sup> Η έννοια των δεικτών των blocks του  $\mathbf{Q}$  είναι: u: up, d: down, l: left, r: right.

τα υπόλοιπα ( $\mathbf{R}_{jr}, \mathbf{R}_{jl}$ ) σταθερά. Με βάση αυτή τη θεώρηση και επιπλέον ότι κάθε  $\mathbf{R}_j$  πρέπει να είναι ορθογώνιο, αφού περιγράφει μητρώο περιστροφής, αγνώνοντας σταθερούς όρους η ποσότητα της (5.32) απλοποιείται στην

$$\text{tr}(\mathbf{R}_j [\mathbf{Q}_{jl}\mathbf{R}_l^T + \mathbf{Q}_{jr}\mathbf{R}_r^T]). \quad (5.33)$$

Θέτοντας ως  $\mathbf{E}_j$

$$\mathbf{E}_j = -[\mathbf{Q}_{jl}\mathbf{R}_l^T + \mathbf{Q}_{jr}\mathbf{R}_r^T]. \quad (5.34)$$

Η ελαχιστοποίηση της (5.33) αναδιατυπώνεται ως μεγιστοποίηση της

$$\Upsilon = \text{tr}(\mathbf{R}_j\mathbf{E}_j). \quad (5.35)$$

Βασιζόμενοι στο λήμμα που παρουσιάσαμε στην [Ενότητα 3.2.1] καταλήγουμε ότι ο  $\mathbf{R}_j$  που μεγιστοποιεί την  $\Upsilon$  ισούται με

$$\mathbf{R}_j = \mathbf{V}_j\mathbf{U}_j^T, \quad (5.36)$$

όπου τα  $\mathbf{V}_j, \mathbf{U}_j$  προέρχονται από την διάσπαση ιδιαζουσών τιμών του  $\mathbf{E}_j$  ως

$$\mathbf{E}_j = \mathbf{U}_j\mathbf{\Lambda}_j\mathbf{V}_j^T. \quad (5.37)$$

Η παραπάνω διαδικασία επαναλαμβάνεται διαδοχικά για κάθε  $\mathbf{R}_j$  και σε κάθε βήμα ανανεώνεται το αντίστοιχο block του  $\mathbf{R}$ . Οι συγγραφείς προτείνουν η παραπάνω διαδικασία να επαναληφθεί έως ότου οι ορμαθοί του  $\mathbf{R}$  συγκλίνουν, δηλαδή σταματήσουν να μεταβάλλονται. *Ός επανάληψη εννοούν επανεκτίμηση όλων των  $M$  blocks, δηλαδή όλων των  $M$  μητρώων περιστροφής.* Γενικά παρατηρήσαμε ότι η σύγκλιση επιτυγχανόταν σε λιγότερες των 10 επαναλήψεων. Δεδομένου του βέλτιστου  $\mathbf{R}$ , το block διάλυσμα ολίσθησης  $\mathbf{T}$  μπορεί να υπολογιστεί βάσει της σχέσης (5.15).

## 5.2 Άγνωστες αντιστοιχίες

Γενικά και στην περίπτωση πολλαπλών όψεων η εύρεση αντίστοιχων σημείων προσεγγίζεται με τρόπους όμοιους με αυτούς που παρουσιάστηκαν στο [κεφάλαιο 4]. Σκοπός αυτής της ενότητας είναι να παρουσιάσει μια γενίκευση του αλγορίθμου ICP για αντιστοίχιση πολλαπλών όψεων καθώς και μία επέκταση του αλγορίθμου ECM-PR [Ενότητα 4.2.3] για την εύρεση αντίστοιχων σημείων η οποία μπορεί να χρησιμοποιηθεί ως στοχαστικός αντικαταστάτης του κριτηρίου κοντινότερου γείτονα [Ενότητα 4.2.1].

### 5.2.1 Ο αλγόριθμος ICP πολλαπλών όψεων

Έχοντας παρουσιάσει την μέθοδο επίλυσης του προβλήματος απόλυτου προσανατολισμού [Ενότητα 5.1.1] μπορούμε να επεκτείνουμε τον αλγόριθμο ICP για την αντιστοίχιση πολλαπλών συνόλων. Αυτό που χρειαζόμαστε είναι μια μέθοδος για την εύρεση επικαλυπτόμενων όψεων και αντίστοιχων σημείων μεταξύ αυτών. Εφόσον σε μια πραγματική κατάσταση δεν μπορούμε να γνωρίζουμε ποιές όψεις επικαλύπτονται είμαστε αναγκασμένοι να υποθέσουμε ότι και κάθε ζεύγος όψεων μπορεί να έχει κοινή περιοχή. Για να εφαρμόσουμε την μέθοδο των Williams

et.al πρέπει να ορίσουμε το μητρώο  $\mathbf{S}$ . Το ελάχιστο (χωρίς πλεονάζοντα ζεύγη) μητρώο  $\mathbf{S}$  κατασκευάζεται από την συλλογή (set)  $\mathcal{C}$  των ανά δύο συνδιασμών των  $M$  όψεων

$$\mathcal{C} = \binom{M}{2} = \frac{M!}{(M-2)! 2!}, \quad (5.38)$$

όπου κάθε συνδυασμός-στοιχείο της συλλογής  $\mathcal{C}$  αποτελεί μία γραμμή του μητρώου  $\mathbf{S}$ . Ακολούθως χρειαζόμαστε μια μέθοδο εύρεσης αντίστοιχων σημείων για τα ζεύγη όψεων του  $\mathbf{S}$ . Για την εύρεση των αντίστοιχων σημείων μπορεί να χρησιμοποιηθεί είτε το κριτήριο κοντινότερου γείτονα (standard ICP) [Ενότητα 4.2.1] είτε η μέθοδος ECM-PR [Ενότητα 4.2.3] με την παρακάτω τροποποίηση:

Τροποποίηση του αλγορίθμου ECM-PR για την εύρεση αντίστοιχων σημείων:

Ο αλγόριθμος ECM-PR με την μορφή που παρουσιάστηκε στην [Ενότητα 4.2.3] επιλύει το πρόβλημα αντιστοίχισης ζεύγους υπολογίζοντας τον μετασχηματισμό  $\mathbf{R}_{pairwise}^p, \mathbf{t}_{pairwise}^p$  ενός ζεύγους όψεων  $\mathcal{V}^{\mathbf{S}(p,1)}, \mathcal{V}^{\mathbf{S}(p,2)}$  βασιζόμενος σε εικονικές αντιστοιχίες, συγκεκριμένα τα  $\mathbf{W}, \mathbf{I}$  μεταξύ των δύο όψεων. Η τεχνική αυτή, στα πλαίσια της στατιστικής κατηγοριοποίησης (classification), ονομάζεται **soft assignment** με την έννοια ότι κάθε πλειάδα (instance) του συνόλου δοκιμών (test set) έχει βαθμό συμμετοχής σε όλες της κλάσεις του συνόλου εκπαίδευσης (training set). Ο τυπικός τρόπος κατηγοριοποίησης ονομάζεται **hard assignment** και βασίζεται στην ανάθεση κάθε πλειάδας σε μία μόνο κλάση. Για την κατασκευή των συνολών  $\mathbf{X}^p, \mathbf{Y}^p$  για κάθε ζεύγος  $p$  του  $\mathbf{S}$  αρχικά εφαρμόζουμε τον αλγόριθμο ECM-PR για την αντιστοίχιση των όψεων  $\mathcal{V}^{\mathbf{S}(p,1)}, \mathcal{V}^{\mathbf{S}(p,2)}$  έως ότου επιτευχθεί σύγκλιση του μετασχηματισμού  $\mathbf{R}_{pairwise}^p, \mathbf{t}_{pairwise}^p$  και στην συνέχεια και είτε εφαρμόζουμε τον μετασχηματισμό στο μητρώο της αντίστοιχης όψης και στη συνέχεια χρησιμοποιούμε το κριτήριο κοντινότερου γείτονα για την εύρεση των  $\mathbf{X}^p, \mathbf{Y}^p$ . Είτε επικαλούμαστε το κριτήριο μέγιστης εκ των υστέρων πιθανότητας (Maximum A Posteriori (MAP)) εφαρμοζόμενο στην τελευταία εκτίμηση των  $a_{ji}$  της Σχέσης (4.69), για την κατηγοριοποίηση (hard assigning) των στοιχείων της μίας όψης στα σημεία-clusters της άλλης. Μια άλλη προσέγγιση είναι η απευθείας χρήση των  $\mathbf{W}$ , σχέση (4.73), ως αντίστοιχων σημείων και των  $\mathbf{I}$ , σχέση (4.74), ως βαρών  $b_i^p$  στην σχέση (5.3). Η προσέγγιση αυτή αποφεύχθηκε εξαιτίας του ότι τα  $\mathbf{W}$  και  $\mathbf{I}$  ήταν επιρρεπή στις διαφορές πληθυκοτήτων των  $\mathcal{V}^{\mathbf{S}(p,1)}$  και  $\mathcal{V}^{\mathbf{S}(p,2)}$ .

**Ο τροποποιημένος αλγόριθμος ICP συνοπτικά**

Έχοντας πλέον μία μέθοδο επίλυσης και για τα δύο κύρια μέρη του αλγορίθμου ICP μπορούμε να παρουσιάσουμε τον γενικευμένο ICP πολλαπλών όψεων σε μορφή ψευδοκώδικα ως σύνοψη της συλλογιστικής μας.

**Algorithm** Generalized-Multiview-Iterative-Closest-Point (GMV-ICP)

**Require:** view matrices  $\mathbf{V}^j \quad \forall j \in [1, M]$ ,  $maxNumIter$ , Initial transformation paramters for each view in the block modules  $\mathbf{R} = [\mathbf{R}_1, \dots, \mathbf{R}_M]$  and  $\mathbf{T} = [\mathbf{t}_1^T, \dots, \mathbf{t}_M^T]^T$ ,  $refinement\_iterations$ , **optional:** matrix of pairs of overlapping views  $\mathbf{S}$ .

- 1: **if**  $\mathbf{S}$  is NOT provided **then**
- 2:   Construct  $\mathbf{S}$  by computing all possible combinations of  $M$  by 2, equation (5.38)
- 3: **end if**
- 4: Initialize iteration counter:
 
$$q \leftarrow 1 .$$
- 5: **repeat**
- 6:   **for** each pair of overlapping views  $[\mathbf{S}(p, 1), \mathbf{S}(p, 2)]$ ,  $p = [1, M]$  **do**
- 7:     Find Correspondences between  $\mathcal{V}^{\mathbf{S}(p,1)}$  and  $\mathcal{V}^{\mathbf{S}(p,2)}$  using modified ECM-PR, nearest neighbor search or other appropriate method:
- 8:     Store the corresponding points between  $\mathcal{V}^{\mathbf{S}(p,1)}$  and  $\mathcal{V}^{\mathbf{S}(p,2)}$  as columns in the matrices  $\mathbf{X}^p$  and  $\mathbf{Y}^p$ , respectively.
- 9:   **end for**
- 10:   Compute  $\mathbf{Q}$  using equation (5.29):
- 11:   **for**  $i = 0; i < refinement\_iterations; i \leftarrow i + 1$  **do**
- 12:     Update all  $M$  blocks of  $\mathbf{R}$  sequentially using the method proposed in [section 5.1.1] or other appropriate method.
- 13:   **end for**
- 14:   Solve for the optimal translation vector  $\mathbf{T}$  based on equation (5.15)
- 15:   increase iteration counter :
 
$$q \leftarrow q + 1 .$$
- 16: **until**  $\mathbf{R}$  is stabilized ||  $q > maxNumIter$  .
- 17: **return**  $\mathbf{R}, \mathbf{T}$  .

## Κεφάλαιο 6

# ΠΡΟΤΕΙΝΟΜΕΝΗ ΤΕΧΝΙΚΗ

Μέχρι αυτό το σημείο η παρούσα διπλωματική αφιερώθηκε στο να εισάγει τον αναγνώστη στο πρόβλημα της αντιστοίχισης μέσα από την παρουσίαση διάφορων μεθόδων επίλυσης αυτού. Η ανάπτυξη των περισσότερων μεθόδων επικεντρώθηκε στην βελτίωση καποιας προϋπάρχουσας. Στην προσπάθεια γενίκευσης της εκάστοτε προϋπάρχουσας η νέα μέθοδος εισήγαγε μειονεκτήματα τα οποία γίνονταν εμφανή στην πορεία. Την τελευταία εικοσαετία έχει γίνει εκτενής έρευνα πάνω στο πρόβλημα αντιστοίχισης η οποία είχε ως αποτέλεσμα την εξάλειψη του μεγαλύτερου πλήθους των μειονεκτημάτων. Η τεχνική που θα παρουσιάσουμε στην παρούσα ενότητα έχει ως σκοπό να εξαλήψει ένα από τα εναπομένοντα μειονεκτήματα των μεθόδων πολλαπλής αντιστοίχισης. Συγκεκριμένα την απαίτηση εκ των προτέρων γνώσης των ζευγών επικαλυπτόμενων νεφών. Η καινοτομία του αλγορίθμου βρίσκεται στην θεώρηση ενός άγνωστου (στοχαστικού) άτλαντα όπου κάθε οπτική αντιστοιχίζεται διαδοχικά με την τρέχουσα εκτίμηση αυτού και στη συνέχεια ο άτλαντας επανυπολογίζεται βάσει των εκτιμηθέντων μετασχηματισμών.

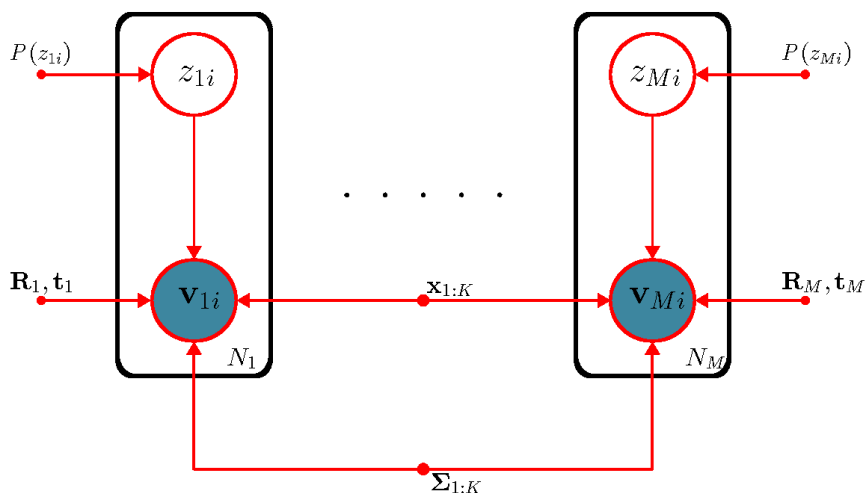
### 6.1 Διατύπωση του προβλήματος

Όπως και προηγουμένως<sup>1</sup> και εδώ υποθέτουμε την ύπαρξη  $M$  επικαλυπτόμενων όψεων  $\mathcal{V}_j$ ,  $j = 1, \dots, M$  οργανωμένων ως μητρώα  $\mathbf{V}_j$ ,  $j = 1, \dots, M$  μεγέθους  $3 \times N_j$  για κάθε ένα από τα οποία αναζητούμε έναν Ευκλείδιο μετασχηματισμό  $\mathbf{R}_j, \mathbf{t}_j$  ο οποίος αντιστοιχίζει την εκάστοτε όψη με όλες τις υπόλοιπες. Τα μετασχηματισμένα σημεία κάθε όψης  $\mathbf{V}_j$ , συγκεκριμένα τα  $\mathbf{R}_j \mathbf{V}_j + \mathbf{t}_j \mathbf{1}_j^{[2]}$  θεωρούνται ως παρατηρήσεις δημιουργούμενες από ένα άγνωστο μίγμα Gaussian κατανομών<sup>3</sup>  $\mathfrak{P}$  οι οποίες έχουν υποστεί έναν Ευκλείδιο μετασχηματισμό  $\mathbf{R}_j, \mathbf{t}_j$ , διαφορετικό για τα σημεία κάθε όψης. Εάν γνωρίζαμε το  $\mathfrak{P}$  τότε θα μπορούσαμε να χρησιμοποιήσουμε την τεχνική της [4.2.3] για να αντιστοιχίσουμε κάθε όψη με τα κέντρα των συνιστωσών του  $\mathfrak{P}$ . Η αρχή της παρούσας μεθόδου είναι η επαναληπτική εκτίμηση

<sup>1</sup>Για μια πιο επεξηγηματική μοντελοποίηση του προβλήματος ο αναγνώστης μπορεί να ανατρέξει στο [κεφάλαιο 5]

<sup>2</sup> $\mathbf{1}_j$  συμβολίζει ένα  $N_j \times 1$  εναδιαίο διάνυσμα στήλη

<sup>3</sup>Gaussian Mixture Models [Ενότητα 3.4.1]



Σχήμα 6.1: Γραφικό μοντέλο του μείγματος  $\mathfrak{P}$  με χρήση plate notation. Οι σκιασμένοι κύκλοι συμβολίζουν παρατηρήσεις ενώ οι διαφανείς συμβολίζουν λανθάνουσες μεταβλητές. Τα βέλη υποδηλώνουν εξάρτηση από παραμέτρους.

των κεντροειδών του  $\mathfrak{P}$  μέσω ταυτόχρονης επανεκτίμησης των παραμέτρων  $\mathbf{R}_j, \mathbf{t}_j$  του μετασχηματισμού κάθε όψης.

## 6.2 Μαθηματική Θεμελίωση του Αλγορίθμου

Ένα τυπικό GMM χαρακτηρίζεται πλήρως από τη μέση τιμή και το μητρώο συνδιασπορών κάθε μίας εκ των συνιστωσών του. Εδώ υποθέτουμε ότι το GMM αποτελείται από  $K$  συνιστώσες χαρακτηριζόμενες από μία μέση τιμή  $\mathbf{x}_k$  και ένα μητρώο συνδιασπορών  $\Sigma_k, k=1, \dots, K$ . Με βάση την επιπλέον θεώρηση ότι οι παρατηρήσεις  $\mathbf{v}_j^i$  εμφωλεύουν τις ντετερμινιστικές παραμέτρους  $\mathbf{R}_j, \mathbf{t}_j$  της όψης από την οποία προέρχονται, μπορούμε να ενσωματώσουμε τα  $\mathbf{R}_j, \mathbf{t}_j$  στο σύνολο αγνώστων παραμέτρων του στοχαστικού μοντέλου  $\mathfrak{P}$  επαυξάνοντας το ως

$$\theta = \{\dots, \mathbf{x}_k, \dots, \Sigma_k, \dots, \mathbf{R}_j, \dots, \mathbf{t}_j, \dots\}. \quad (6.1)$$

Η τεχνική με την οποία θα επιτευχθεί η εκτίμηση των βέλτιστων παραμέτρων  $\theta^*$  αποτελεί μια επέκταση του αλγορίθμου ECM-PR [Ενότητα 4.2.3], πυρήνας του οποίου είναι η τεχνική μεγιστοποίησης δεσμευμένης πιθανοφάνειας (ECM) που παρουσιάστηκε στην [Ενότητα 3.4.2]. Για την κατασκευή της αντικειμενικής συνάρτησης ορίζουμε αρχικά την πιθανοφάνεια παρατηρήσεων  $\mathcal{L}(\theta|\mathfrak{A})$  ως:

$$\mathcal{L}(\theta|\mathfrak{A}) = \mathcal{P}(\mathfrak{A}; \theta) \quad (6.2)$$

όπου  $\mathfrak{A}$  είναι η ένωση<sup>4</sup> των συνόλων παρατηρήσεων όλων των όψεων:

$$\mathfrak{A} = \bigcup_{j=1}^M \hat{\mathbf{V}}_j. \quad (6.3)$$

Στην παραπάνω σχέση το  $\hat{\mathbf{V}}_j$  συμβολίζει το  $j^{\text{οστό}}$  σύνολο μετασχηματισμένων παρατηρήσεων ορισμένο ως:

$$\hat{\mathbf{V}}_j = \mathbf{R}_j \mathbf{V}_j + \mathbf{t}_j. \quad (6.4)$$

Είναι γνωστό<sup>[22][21][28]</sup> ότι η απευθείας μεγιστοποίηση της (6.2) ως προς  $\theta$  είναι αδύνατη λόγω της ύπαρξης αγνώστων δεδομένων, συγκεκριμένα των αναθέσεων κάθε σημείου  $\hat{\mathbf{v}}_j^i \in \mathfrak{A}$  σε μία από τις συνιστώσες του μείγματος  $\mathfrak{F}$ . Στο εξής οι αναθέσεις αυτές θα συμβολίζονται ως  $\mathcal{Z} = \{Z_j^i\}$ ,  $1 < j < M, 1 < i < N_j$  και θα θεωρούνται λανθάνουσες μεταβλητές. Κάθε μεταβλητή  $Z_j^i$  αναθέτει μία παρατήρηση  $\hat{\mathbf{v}}_j^i$  στο κεντροϊδές  $\mathbf{x}_k$  της  $k$ -οστής συνιστώσας,  $k = 1, \dots, K$  ή στην ομάδα των μολυσμένων δεδομένων, δεικτοδοτούμενη ως η κλάση  $K+1$ . Οι Dempster et. al<sup>[21]</sup> πρότειναν την αντικατάσταση<sup>5</sup> της  $\mathcal{L}(\theta|\mathfrak{A})$  με την αναμενόμενη λογαριθμική πιθανοφάνεια πλήρων δεδομένων διεπόμενων από τις παρατηρήσεις, η οποία ορίζεται ως ακολούθως

$$\mathcal{E}(\theta|\mathfrak{A}, \mathcal{Z}) = \mathbb{E}_{\mathcal{Z}}[\log P(\mathfrak{A}, \mathcal{Z}; \theta)|\mathfrak{A}]. \quad (6.5)$$

Ο αλγόριθμος EM<sup>[21]</sup> είναι μια τεχνική επαναληπτικής φύσεως που χρησιμοποιείται για την εκτίμηση παραμέτρων στατιστικών μοντέλων υπο την ύπαρξη αγνώστων δεδομένων και την οποία θα χρησιμοποιήσουμε για την εκτίμηση του  $\theta$ . Ο υπολογισμός του μεγίστου της (6.5) ως προς όλες τις παραμέτρους του συνόλου  $\theta$  θα οδηγήσει σε ένα δυσεπίλυτο μη γραμμικό πρόβλημα. Για το λόγο αυτό η μεγιστοποίηση θα γίνει με τη βοήθεια εναλλαγών (alternation)<sup>6</sup>. Η ιδέα των εναλλαγών είναι σε κάθε επανάληψη η μεγιστοποίηση να γίνεται ως προς ένα υποσύνολο των παραμέτρων, στην περίπτωση μας των μετασχηματιστικών  $\mathbf{R}_j, \mathbf{t}_j \forall j$  και στη συνέχεια δεδομένης αυτής της εκτίμησης να γίνεται μεγιστοποίηση ως προς τις εναπομένουσες παραμέτρους  $\mathbf{x}_k, \Sigma_k$ . Όπως θα δούμε στη συνέχεια θα χρειαστεί ένα επιπλέον βήμα για την ανανέωση των  $\Sigma_k$  εξαιτίας της εξάρτησής τους από τα  $\mathbf{x}_k$ . Πρίν περιγράψουμε την επανάληψη του αλγορίθμου πρέπει να κατασκευάσουμε φORMALISΤΙΚΑ την  $\mathcal{E}(\mathcal{V}, \mathcal{Z}; \theta)$   $\mathcal{E}(\mathcal{V}, \mathcal{Z}; \theta)$  για το μείγμα  $\mathfrak{F}$ .

Αναλύοντας το όρισμα της (6.5) και υποθέτωντας ότι τα  $\mathbf{v}_j^i$  είναι ανεξάρτητα και

<sup>4</sup> Στην αξιωματική θεωρία συνόλων, (axiomatic set theory), η πράξη της ένωσης απορρίπτει διπλότυπα, duplicate elements, στην περίπτωση μας ταυτοτικά σημεία διατηρούνται καθώς υπάρχει περίπτωση να ανήκουν σε διαφορετικές όψεις ή να έχουν ταυτοποιηθεί λόγω της στοχαστικής διαδικασίας που τα δημιουργήσε.

<sup>5</sup> Για μία δικαιολόγηση αυτής της επιλογής ο αναγνώστης μπορεί να ανατρέξει στην [Ενότητα 3.4.2].

<sup>6</sup> Στον αλγόριθμο EM αυτό μεταφράζεται σε πολλαπλά M-steps για διαφορετικά υποσύνολα του  $\theta$ .

όμοια κατανεμημένα (iid) έχουμε ότι:

$$\begin{aligned}
 P(\mathfrak{A}, \mathcal{Z}; \theta) &= \prod_{j=1}^M \prod_{i=1}^{N_j} P(\hat{\mathbf{v}}_j^i, Z_j^i = k; \theta) = \\
 &= \prod_{j=1}^M \prod_{i=1}^{N_j} P(Z_j^i = k) P(\hat{\mathbf{v}}_j^i | Z_j^i = k; \theta) = \\
 &= \prod_{j=1}^M \prod_{i=1}^{N_j} \prod_{k=1}^K \left( P(Z_j^i = k) P(\hat{\mathbf{v}}_j^i | Z_j^i = k; \theta) \right)^{\delta_{Z_j^i, k}}, \quad (6.6)
 \end{aligned}$$

όπου το  $\delta_{Z_j^i, k}^k$  συμβολίζει την διακριτή ακολουθία kronecker ορισμένη ως:

$$\delta_{Z_j^i, k} = \begin{cases} 1, & Z_j^i = k \\ 0, & \text{αλλιώς,} \end{cases} \quad (6.7)$$

και όπου  $P(Z_j^i = k)$  οι εκ των προτέρων πιθανότητες<sup>7</sup> των συνιστωσών του  $\mathfrak{F}$ :

$$P(Z_j^i = k) = \begin{cases} \frac{u}{\mathfrak{U}}, & \text{if } 1 \leq k \leq n \\ \frac{\mathfrak{U} - Ku}{\mathfrak{U}}, & \text{if } k = K + 1 \end{cases} \quad [\text{χλάση μολυσμένων δεδομένων}] \quad (6.8)$$

όπου  $u = 4\pi r^3$  είναι ο όγκος μίας μικρής σφαίρας με ακτίνα  $r$  και κέντρο ένα από τα  $\mathbf{x}_k$  και όπου  $\mathfrak{U}$  είναι η ακτίνα μιας περικλείουσας σφαίρας κατασκευασμένης ώστε  $Ku \ll \mathfrak{U}$ .

Εφαρμόζοντας στην συνέχεια τον τελεστή αναμενόμενης τιμής  $\mathbb{E}_{\mathcal{Z}}[\cdot]$  στην (6.6) η ποσότητα  $\mathcal{E}(\theta | \mathfrak{A}, \mathcal{Z})$  επαναγράφεται ως:

$$\mathbb{E}_{\mathcal{Z}} \left[ \sum_{j=1}^M \sum_{i=1}^{N_j} \sum_{k=1}^K \delta_{Z_j^i, k} \left( \underbrace{\log P(Z_j^i = k) + \log P(\hat{\mathbf{v}}_j^i | Z_j^i = k; \theta)}_I \right) | \mathfrak{A} \right]. \quad (6.9)$$

Δεδομένης της γραμμικότητας του τελεστή  $\mathbb{E}_{\mathcal{Z}}[\cdot]$  και της ανεξαρτησίας του όρου  $\mathbb{E}_{\mathcal{Z}}[\cdot]$  από την  $\mathcal{Z}$  η σχέση (6.9) απλοποιείται στην

$$\sum_{j=1}^M \sum_{i=1}^{N_j} \sum_{k=1}^K \mathbb{E}_{\mathcal{Z}} \left[ \delta_{Z_j^i, k} | \mathfrak{A} \right] \left( \log P(Z_j^i = k) + \log P(\hat{\mathbf{v}}_j^i | Z_j^i = k; \theta) \right). \quad (6.10)$$

Πρίν συνεχίσουμε με την ανάπτυξη του  $\mathbb{E}_{\mathcal{Z}}[\delta_{Z_j^i, k} | \mathfrak{A}]$  απαιτείται να ορίσουμε τις δεσμευμένες πιθανότητες- (*likelihoods*) για κάθε παρατήρηση  $\hat{\mathbf{v}}_j^i \in \mathfrak{A}$  δεδομένης της ανάθεσης της στο cluster  $\mathbf{x}_k$ . Η πιθανότητα αυτή θεωρούμε ότι περιγράφεται από κανονική κατανομή με μέση τιμή  $\mathbf{x}_k$  και μητρώο συνδιασπορών  $\mathbf{\Sigma}_k$ . Φορμαλιστικά

$$P(\hat{\mathbf{v}}_j^i | Z_j^i = k) = \mathcal{N}(\hat{\mathbf{v}}_j^i | \mathbf{x}_k, \mathbf{\Sigma}_k), \quad \forall 1 \leq k \leq K. \quad (6.11)$$

Πλέον είμαστε σε θέση να αναπτύξουμε το  $\mathbb{E}_{\mathcal{Z}}[\delta_{Z_j^i, k} | \mathfrak{A}]$ . Με βάση τον κανόνα του Bayes<sup>8</sup> και τις δεσμευμένες πιθανότητες όπως ορίστηκαν παραπάνω, η ποσότητα

<sup>7</sup>Η πιθανότητες  $P(Z_j^i = k)$  μοντελοποιούνται όπως και στην Σχέση (4.62).

<sup>8</sup>Λέγεται επίσης και κανόνας αντιστροφής (Bayes Inversion Rule).



$\mathbb{E}_{\mathcal{Z}}[\delta_{Z_j^i, k} | \mathfrak{A}]$  γράφεται ως

$$\begin{aligned}
\mathbb{E}_{\mathcal{Z}} \left[ \delta_{Z_j^i, k} | \mathfrak{A} \right] &\doteq \sum_{p=1}^{K+1} \delta_{k,p} P(Z_j^i = p | \hat{\mathbf{v}}_j^i) &= \\
&P(Z_j^i = k | \hat{\mathbf{v}}_j^i) &= \\
&\frac{P(\hat{\mathbf{v}}_j^i | Z_j^i = k) P(Z_j^i = k)}{\sum_{p=1}^{K+1} P(Z_j^i = p) P(\hat{\mathbf{v}}_j^i | Z_j^i = p)} &= \\
&\frac{\frac{u}{\mathfrak{A}} \frac{1}{\sqrt{(2\pi)^3 |\Sigma_k|}} \exp\left(-\frac{1}{2} \|\hat{\mathbf{v}}_j^i - \mathbf{x}_k\|_{\Sigma_k}^2\right)}{\frac{u}{\mathfrak{A}} \sum_{p=1}^K \left[ \frac{1}{\sqrt{(2\pi)^3 |\Sigma_p|}} \exp\left(-\frac{1}{2} \|\hat{\mathbf{v}}_j^i - \mathbf{x}_p\|_{\Sigma_p}^2\right) \right] + \frac{u-nu}{\mathfrak{A}} P(\hat{\mathbf{v}}_j^i | Z_j^i = K+1)} &= \\
&\frac{\frac{u}{\mathfrak{A}} \frac{1}{\sqrt{(2\pi)^3 |\Sigma_k|}} \left[ |\Sigma_k|^{-\frac{1}{2}} \exp\left(-\frac{1}{2} \|\hat{\mathbf{v}}_j^i - \mathbf{x}_k\|_{\Sigma_k}^2\right) \right]}{\frac{u}{\mathfrak{A}} \frac{1}{\sqrt{(2\pi)^3} \left[ \sum_{p=1}^K |\Sigma_p|^{-\frac{1}{2}} \exp\left(-\frac{1}{2} \|\hat{\mathbf{v}}_j^i - \mathbf{x}_p\|_{\Sigma_p}^2\right) + \frac{u-nu}{\mathfrak{A}} \frac{1}{\mathfrak{A}} \frac{u\sqrt{(2\pi)^3}}{u} \right]}} &= \\
&\frac{|\Sigma_k|^{-\frac{1}{2}} \exp\left(-\frac{1}{2} \|\hat{\mathbf{v}}_j^i - \mathbf{x}_k\|_{\Sigma_k}^2\right)}{\sum_{p=1}^K \left[ |\Sigma_p|^{-\frac{1}{2}} \exp\left(-\frac{1}{2} \|\hat{\mathbf{v}}_j^i - \mathbf{x}_p\|_{\Sigma_p}^2\right) \right] + \frac{3(2\pi)^{\frac{3}{2}}}{4\pi r^3}} &= \\
&\frac{|\Sigma_k|^{-\frac{1}{2}} \exp\left(-\frac{1}{2} \|\hat{\mathbf{v}}_j^i - \mathbf{x}_k\|_{\Sigma_k}^2\right)}{\sum_{p=1}^K \left[ |\Sigma_p|^{-\frac{1}{2}} \exp\left(-\frac{1}{2} \|\hat{\mathbf{v}}_j^i - \mathbf{x}_p\|_{\Sigma_p}^2\right) \right] + 1.5\sqrt{2\pi} r^{-3}} = a_{ik}^j. \quad (6.12)
\end{aligned}$$

Αντικαθιστώντας τα  $a_{ik}^j$ , τις δεσμευμένες πιθανότητες  $P(\hat{\mathbf{v}}_j^i | Z_j^i = k)$  και απαλείφοντας σταθερούς όρους, δηλαδή όρους που δεν εξαρτώνται από το σύνολο  $\theta$  η σχέση (6.10) απλοποιείται στην

$$\Phi(\theta) = -\frac{1}{2} \sum_{j=1}^M \sum_{i=1}^{N_j} \sum_{k=1}^K a_{ik}^j \left( \|\hat{\mathbf{v}}_j^i - \mathbf{x}_k\|_{\Sigma_p}^2 + \log |\Sigma_k| \right). \quad (6.13)$$

Αποδεικνύεται<sup>[21][22]</sup> ότι η μεγιστοποίηση της  $\Phi(\theta)$  συνεπάγεται αύξηση της  $\mathcal{L}(\theta | \mathfrak{A})$ .

### 6.2.1 Εκτίμηση του $\theta^*$ στην περίπτωση ισοτροπικών διασπορών

Εάν υποθέσουμε ότι οι συνιστώσες του  $\mathfrak{B}$  χαρακτηρίζονται από ισοτροπικά μητρώα συνδιασπορών της μορφής  $\Sigma_k = \sigma_k^2 \mathbf{I}$  η σχέση (6.13) απλοποιείται<sup>9</sup> παιρεταίρω

$$\Phi(\theta) = \sum_{j=1}^M \sum_{i=1}^{N_j} \sum_{k=1}^K \frac{a_{ik}^j}{2\sigma_k^2} \|\hat{\mathbf{v}}_j^i - \mathbf{x}_k\|_2^2 + 3 \sum_{j=1}^M \sum_{i=1}^{N_j} \sum_{k=1}^K a_{ik}^j \log(\sigma_k) \quad (6.14)$$

<sup>9</sup>Το ανεστραμμένο πρόσημο επαναθέτει την μεγιστοποίηση της (6.13) ως ελαχιστοποίηση της (6.14).

Στο σημείο ολικού ελαχίστου το διάνυσμα κλήσεων (**gradient vector**) πρέπει να ισούται με το μηδενικό ή πιο τυπικά  $\nabla\Phi(\theta^*) = \vec{0}$ . Υπολογίζοντας, αρχικά, τις μερικές παραγώγους της  $\Phi$  ως προς κάθε κεντροειδή  $\mathbf{x}_k$  έχουμε:

$$\begin{aligned} \frac{\partial\Phi(\theta^*)}{\partial\mathbf{x}_k} &= \vec{0} \Rightarrow \\ \frac{1}{2} \sum_{j=1}^M \sum_{i=1}^{N_j} \frac{a_{ik}^j}{\sigma_k^2} \left[ -2\hat{\mathbf{v}}_j^i + 2\mathbf{x}_k \right] &= \vec{0} \Rightarrow \\ \sum_{j=1}^M \sum_{i=1}^{N_j} \frac{a_{ik}^j}{\sigma_k^2} \hat{\mathbf{v}}_j^i &= \sum_{j=1}^M \sum_{i=1}^{N_j} \frac{a_{ik}^j}{\sigma_k^2} \mathbf{x}_k \Rightarrow \\ \mathbf{x}_k^* &= \frac{\sum_{j=1}^M \sum_{i=1}^{N_j} a_{ik}^j \hat{\mathbf{v}}_j^i}{\sum_{j=1}^M \sum_{i=1}^{N_j} a_{ik}^j} \quad \forall 1 \leq k \leq K. \end{aligned} \quad (6.15)$$

**ΣΗΜΕΙΩΣΗ:** Ο υπολογισμός των βέλτιστων  $\mathbf{x}_k$  δεν εξαρτάται από τα  $\sigma_k$  και εξ αυτού η ανανέωση τους προηγείται της επανεκτίμησης των  $\sigma_k$ .

Εν συνεχεία υπολογίζουμε τις μερικές παραγώγους της  $\Phi$  ως προς (κάθε)  $\sigma_k$

$$\begin{aligned} \frac{\partial\Phi(\theta^*)}{\partial\sigma_k} &= \vec{0} \Rightarrow \\ \frac{1}{\sigma_k} \left( \frac{-1}{\sigma_k^2} \sum_{j=1}^M \sum_{i=1}^{N_j} a_{ik}^j \|\hat{\mathbf{v}}_j^i - \mathbf{x}_k\|_2^2 + 3 \sum_{j=1}^M \sum_{i=1}^{N_j} a_{ik}^j \right) &= \vec{0} \Rightarrow \\ \sigma_k^2 &= \frac{\sum_{j=1}^M \sum_{i=1}^{N_j} a_{ik}^j \|\hat{\mathbf{v}}_j^i - \mathbf{x}_k\|_2^2}{3 \sum_{j=1}^M \sum_{i=1}^{N_j} a_{ik}^j} \quad \forall 1 \leq k \leq K. \end{aligned} \quad (6.16)$$

**ΣΗΜΕΙΩΣΗ:** Ο όρος  $\frac{1}{\sigma_k}$  απαλείφεται αφού δεν μας ενδιαφέρει η τετριμμένη λύση άπειρης διασποράς.

Σε αυτό το σημείο αυτό που απομένει είναι η ανανέωση των παραμέτρων  $\mathbf{R}_j, \mathbf{t}_j$  των μετασχηματισμών κάθε όψης. Για την επίτευξη αυτής θα μετατρέψουμε την αντικειμενική συνάρτηση  $\Phi$  ώστε να ενσωματωθεί στο πλαίσιο της της επίλυσης του S.Umeyama η οποία αναπτύχθηκε στην [υποενότητα 4.1.4].

$$\begin{aligned} \mathbf{R}_j^*, \mathbf{t}_j^* &= \arg \min_{\mathbf{R}_j, \mathbf{t}_j} \Phi(\theta) = \\ \arg \min_{\mathbf{R}_j, \mathbf{t}_j} \frac{1}{2} \sum_{i=1}^{N_j} \sum_{k=1}^K \frac{a_{ik}^j}{\sigma_k^2} \|\hat{\mathbf{v}}_j^i - \mathbf{x}_k\|_2^2 &= \\ \arg \min_{\mathbf{R}_j, \mathbf{t}_j} \frac{1}{2} \sum_{i=1}^{N_j} \sum_{k=1}^K \frac{a_{ik}^j}{\sigma_k^2} \|\mathbf{R}_j \mathbf{v}_j^i + \mathbf{t}_j - \mathbf{x}_k\|_2^2 & \end{aligned} \quad (6.17)$$

Αρχικά υπολογίζουμε το βέλτιστο  $\mathbf{t}_j$  ως εξής

$$\begin{aligned} \frac{\partial \Phi(\theta)}{\partial \mathbf{t}_j} = 0 \Rightarrow \\ \sum_{i=1}^{N_j} \sum_{k=1}^K \frac{a_{ik}^j}{\sigma_k^2} (\mathbf{R}_j \mathbf{v}_j^i + \mathbf{t}_j - \mathbf{x}_k) = 0 \Rightarrow \\ \mathbf{t}_j = \frac{\sum_{i=1}^{N_j} \sum_{k=1}^K \frac{a_{ik}^j}{\sigma_k^2} (\mathbf{x}_k - \mathbf{R}_j \mathbf{v}_j^i)}{\sum_{i=1}^{N_j} \sum_{k=1}^K \frac{a_{ik}^j}{\sigma_k^2}} \end{aligned} \quad (6.18)$$

Αντικαθιστώντας το βέλτιστο  $\mathbf{t}_j$  στην Σχέση (6.17) αυτή μετατρέπεται ως εξής

$$\begin{aligned} \sum_{i=1}^{N_j} \sum_{k=1}^K \frac{a_{ik}^j}{\sigma_k^2} \|\mathbf{R}_j \mathbf{v}_j^i + \mathbf{t}_j - \mathbf{x}_k\|_2^2 = \\ \sum_{i=1}^{N_j} \sum_{k=1}^K \frac{a_{ik}^j}{\sigma_k^2} \left\| \mathbf{R}_j \left( \mathbf{v}_j^i - \frac{\sum_{i=1}^{N_j} \sum_{k=1}^K \frac{a_{ik}^j}{\sigma_k^2} \mathbf{v}_j^i}{\sum_{i=1}^{N_j} \sum_{k=1}^K \frac{a_{ik}^j}{\sigma_k^2}} \right) - \left( \mathbf{x}_k - \frac{\sum_{i=1}^{N_j} \sum_{k=1}^K \frac{a_{ik}^j}{\sigma_k^2} \mathbf{x}_k}{\sum_{i=1}^{N_j} \sum_{k=1}^K \frac{a_{ik}^j}{\sigma_k^2}} \right) \right\|_2^2 \end{aligned} \quad (6.19)$$

Θέτωντας ως  $\bar{\mathbf{v}}_j^i$  και  $\bar{\mathbf{x}}_k$  τις ακόλουθες ποσότητες

$$\begin{aligned} \bar{\mathbf{v}}_j^i = \mathbf{v}_j^i - \frac{\sum_{i=1}^{N_j} \sum_{k=1}^K \frac{a_{ik}^j}{\sigma_k^2} \mathbf{v}_j^i}{\sum_{i=1}^{N_j} \sum_{k=1}^K \frac{a_{ik}^j}{\sigma_k^2}}, \\ \bar{\mathbf{x}}_k = \mathbf{x}_k - \frac{\sum_{i=1}^{N_j} \sum_{k=1}^K \frac{a_{ik}^j}{\sigma_k^2} \mathbf{x}_k}{\sum_{i=1}^{N_j} \sum_{k=1}^K \frac{a_{ik}^j}{\sigma_k^2}}, \end{aligned} \quad (6.20)$$

η Σχέση (6.19) συμπύσσεται στην

$$\sum_{i=1}^{N_j} \sum_{k=1}^K \frac{a_{ik}^j}{\sigma_k^2} \|\mathbf{R}_j \bar{\mathbf{v}}_j^i - \bar{\mathbf{x}}_k\|_2^2, \quad (6.21)$$

η ελαχιστοποίηση της οποίας, ως προς  $\mathbf{R}_j$ , συνεπάγεται

$$\begin{aligned} \sum_{i=1}^{N_j} \sum_{k=1}^K \frac{a_{ik}^j}{\sigma_k^2} \text{tr}((\bar{\mathbf{v}}_j^i \mathbf{R}_j^T - \bar{\mathbf{x}}_k^T)(\mathbf{R}_j \bar{\mathbf{v}}_j^i - \bar{\mathbf{x}}_k)) = \\ \sum_{i=1}^{N_j} \sum_{k=1}^K \frac{a_{ik}^j}{\sigma_k^2} \text{tr}(\mathbf{R}_j^T \mathbf{R}_j \bar{\mathbf{v}}_j^i \bar{\mathbf{v}}_j^{iT} - 2\mathbf{R}_j \bar{\mathbf{v}}_j^i \bar{\mathbf{x}}_k^T + \bar{\mathbf{x}}_k \bar{\mathbf{x}}_k^T) = \\ \text{tr}(\mathbf{R}_j^T \mathbf{R}_j \sum_{i=1}^{N_j} \sum_{k=1}^K \frac{a_{ik}^j}{\sigma_k^2} \bar{\mathbf{v}}_j^i \bar{\mathbf{v}}_j^{iT} - 2\mathbf{R}_j \sum_{i=1}^{N_j} \sum_{k=1}^K \frac{a_{ik}^j}{\sigma_k^2} \bar{\mathbf{v}}_j^i \bar{\mathbf{x}}_k^T + \sum_{i=1}^{N_j} \sum_{k=1}^K \frac{a_{ik}^j}{\sigma_k^2} \bar{\mathbf{x}}_k \bar{\mathbf{x}}_k^T). \end{aligned} \quad (6.22)$$

Θέτωνατας ως  $\mathbf{Q}$ ,  $\mathbf{P}$  και  $\mathbf{C}$  τις ακόλουθες ποσότητες

$$\mathbf{Q} = \sum_{i=1}^{N_j} \sum_{k=1}^K \frac{a_{ik}^j}{\sigma_k^2} \bar{\mathbf{v}}_j^i \bar{\mathbf{v}}_j^{iT} \quad (6.23)$$

$$\mathbf{P} = \sum_{i=1}^{N_j} \sum_{k=1}^K \frac{a_{ik}^j}{\sigma_k^2} \bar{\mathbf{v}}_j^i \bar{\mathbf{x}}_k^T \quad (6.24)$$

$$\mathbf{C} = \sum_{i=1}^{N_j} \sum_{k=1}^K \frac{a_{ik}^j}{\sigma_k^2} \bar{\mathbf{x}}_k \bar{\mathbf{x}}_k^T, \quad (6.25)$$

η Σχέση (6.22) επαναγράφεται ως

$$\begin{aligned} & \text{tr}(\mathbf{R}_j^T \mathbf{R}_j \mathbf{Q} - 2\mathbf{R}_j \mathbf{P} + \mathbf{C}) = \\ & \text{tr}(\mathbf{R}_j \mathbf{Q}^{1/2} \mathbf{Q}^{1/2} \mathbf{R}_j^T - 2\mathbf{R}_j \mathbf{Q}^{1/2} \mathbf{Q}^{-1/2} \mathbf{P} + \mathbf{C} + \mathbf{P}^T \mathbf{Q}^{-1} \mathbf{P} - \mathbf{P}^T \mathbf{Q}^{-1} \mathbf{P}) = \\ & \|\mathbf{R}_j \mathbf{Q}^{1/2} - \mathbf{P}^T \mathbf{Q}^{-1/2}\|_{\mathbf{F}}^2 + \text{tr}(\mathbf{C} - \mathbf{P}^T \mathbf{Q}^{-1} \mathbf{P}). \end{aligned} \quad (6.26)$$

Η ελαχιστοποίηση της παραπάνω ποσότητας είναι ισοδύναμη με την ελαχιστοποίηση της

$$\|\mathbf{P}^T \mathbf{Q}^{-1/2} - \mathbf{R}_j \mathbf{Q}^{1/2}\|_{\mathbf{F}}^2, \quad (6.27)$$

η οποία έχει την ίδια μορφή με την Σχέση (4.29) και επιλύεται βάσει της μεθόδου της [Ενότητας 4.1.4].

### 6.2.2 Μία πιθανή υλοποίηση

Εξαιτίας της φύσης του EM, ο αλγόριθμος πρέπει να υλοποιηθεί επαναληπτικά. Στην παρούσα υποενότητα παραθέτουμε μία από τις δυνατές υλοποιήσεις. Να σημειώσουμε ότι η σειρά της ανανέωσης των παραμέτρων δέν είναι μοναδική αλλά θα ήταν φρόνιμο τα κεντροειδή  $\mathbf{x}_k$  να ανανεώνονται πρίν από τις διασπορές  $\sigma_k$  διότι οι μέν δεν εξαρτώνται από τις δε ένω το αντίστροφο δέν ισχύει.

**Algorithm** MULTIVIEW-SYMMETRIC-REGISTRATION

**Require:**  $M$  view matrices  $\mathbf{V}_j \ j=1,\dots,M$ ,  $\maxNumIter$ ,  $r$ ,  $K$   
 Initial parameter set  $\theta^0 = \{\dots, \mathbf{x}_k, \dots, \Sigma_k^0, \dots, \mathbf{R}_j^0, \dots, \mathbf{t}_j^0, \dots\}$ .

1: Initialize iteration counter :  $q \leftarrow 1$  .

2: **repeat**

3: compute posteriors :

$$a_{ik}^j = \frac{(\sigma_k^{q-1})^{-3} \exp\left(-\frac{1}{2} \|\mathbf{R}_j^{q-1} \mathbf{v}_j^i + \mathbf{t}_j^{q-1} - \mathbf{x}_k\|_{\sigma_k^{q-1}}^2\right)}{\sum_{p=1}^K \left[ (\sigma_p^{q-1})^{-3} \exp\left(-\frac{1}{2} \|\mathbf{R}_j^{q-1} \mathbf{v}_j^i + \mathbf{t}_j^{q-1} - \mathbf{x}_p\|_{\sigma_p^{q-1}}^2\right) \right] + 1.5\sqrt{2\pi}r^{-3}} . \quad 6.12\uparrow$$

4: **For**  $j = 1$  to  $M$  **do**

5: Estimate optimal translation:

$$\mathbf{t}_j = \frac{\sum_{i=1}^{N_j} \sum_{k=1}^K \frac{a_{ik}^j}{\sigma_k^2} (\mathbf{x}_k - \mathbf{R}_j \mathbf{v}_j^i)}{\sum_{i=1}^{N_j} \sum_{k=1}^K \frac{a_{ik}^j}{\sigma_k^2}} . \quad 6.18\uparrow$$

6: Estimate optimal Rotation by maximizing the following quantity:

$$\| \mathbf{P}^T \mathbf{Q}^{-1/2} - \mathbf{R}_j \mathbf{Q}^{1/2} \|_{\mathbf{F}}^2 . \quad 6.27\uparrow$$

7: **end for**

8: Update component centers  $\mathbf{x}_k$ :

$$\mathbf{x}_k^q = \frac{\sum_{j=1}^M \sum_{i=1}^{N_j} a_{ik}^j \mathbf{R}_j^q \mathbf{v}_j^i + \mathbf{t}_j^q}{\sum_{j=1}^M \sum_{i=1}^{N_j} a_{ik}^j} \quad 6.15\uparrow$$

9: Update covariances :

$$\sigma_k^q = \left( \frac{\sum_{j=1}^M \sum_{i=1}^{N_j} a_{ik}^j \|\mathbf{R}_j^q \mathbf{v}_j^i + \mathbf{t}_j^q - \mathbf{x}_k^q\|_2^2}{3 \sum_{j=1}^M \sum_{i=1}^{N_j} a_{ik}^j} \right)^{\frac{1}{2}} . \quad 6.16\uparrow$$

10: increase iteration counter :

$$q \leftarrow q + 1 .$$

11: **until** convergence  $\| q > \maxNumIter$  .

12: **return**  $\mathbf{R}_j, \mathbf{t}_j \ \forall j = 1, \dots, M$



## Κεφάλαιο 7

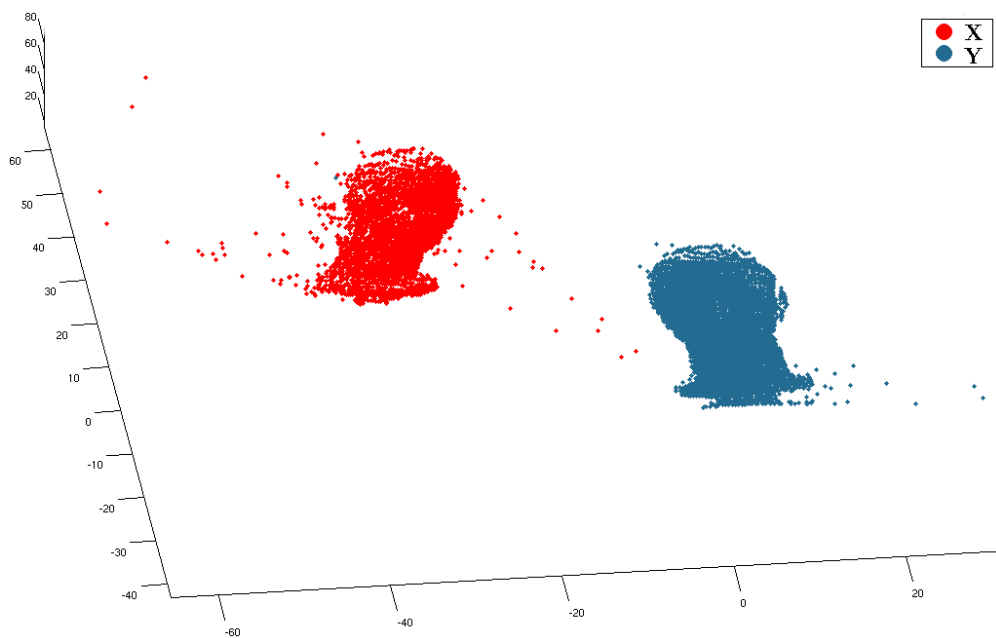
# ΠΕΙΡΑΜΑΤΑ

Οι υλοποιήσεις των πειραμάτων έγιναν σε περιβάλλον MATLAB και συγκεκριμένα στην έκδοση R2012a. Εξαιτίας αλλαγών σε στοιχεία του parser ενδέχεται κάποιοι τελεστές να μην είναι ορισμένοι σε παλαιότερες εκδόσεις. Το κεφάλαιο των πειραμάτων χωρίζεται σε δύο υπόενότητες. Στην πρώτη συγκρίνουμε μεθόδους αντιστοίχισης ζεύγους νεφών πάνω σε πραγματικά και τεχνητά δεδομένα. Συγκεκριμένα θα συγκρίνουμε τους αλγορίθμους ICP [Ενότητα 4.2.1], και ECM-PR [Ενότητα 4.2.3]. Στη δεύτερη ενότητα εστιάζουμε την προσοχή μας σε μεθόδους αντιστοίχισης πολλαπλών νεφών. Συγκεκριμένα θα συγκρίνουμε τον αλγόριθμό μας [κεφάλαιο 6], με μία συμμετρική έκδοση ενός πολλαπλού ICP [Ενότητα 5.2.1], βασισμένη στην επίλυση των Williams et.al [Ενότητα 5.1.1]. Τα πραγματικά δεδομένα προέρχονται από μια TOF-camera [Ενότητα 2.2], συγκεκριμένα μία SwissRanger MESA 4000 τα οποία, προ της αντιστοίχισης, έχουν υποστεί καταφλίωση βάνους για την αποκοπή του υποβάθρου (background). Τα τεχνητά δεδομένα προέρχονται από την ηλεκτρονική βιβλιοθήκη του Stanford, <http://graphics.stanford.edu/data/3Dscanrep/>.

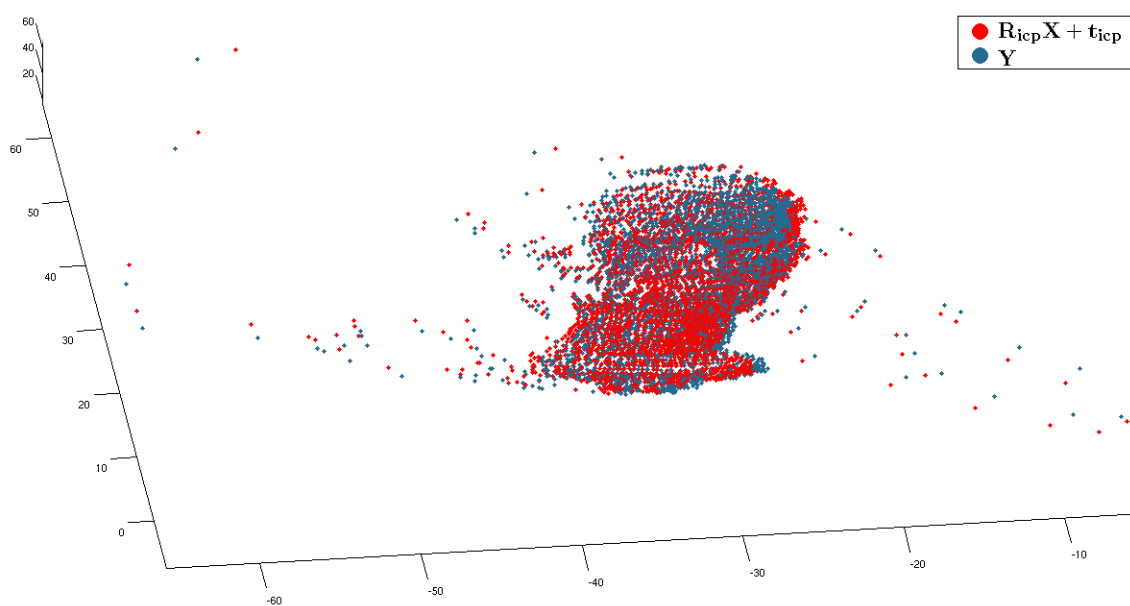
### 7.1 Πειράματα αντιστοίχισης ζεύγους

Να σημειώσουμε εδώ ότι σε όλα τα πειράματα οι αντιστοιχίες θεωρούνται άγνωστες. Αυτό επιβάλλεται έμμεσα από τα πραγματικά δεδομένα, αφού τα νέφη δέν είναι ισοπλήθη (εξαιτίας της καταφλίωσης), ενώ στα τεχνητά δεδομένα αυτό επιτυγχάνεται με υποδειγματοληψία (ώστε να προκύψουν ανισοπλήθη σύνολα). Επίσης όλοι οι αλγόριθμοι αρχικοποιούνται με  $\mathbf{R}_0 = \mathbf{I}_3$  και  $\mathbf{t}_0 = [0 \ 0 \ 0]^T$ .

## 7.1.1 Με χρήση πραγματικών δεδομένων

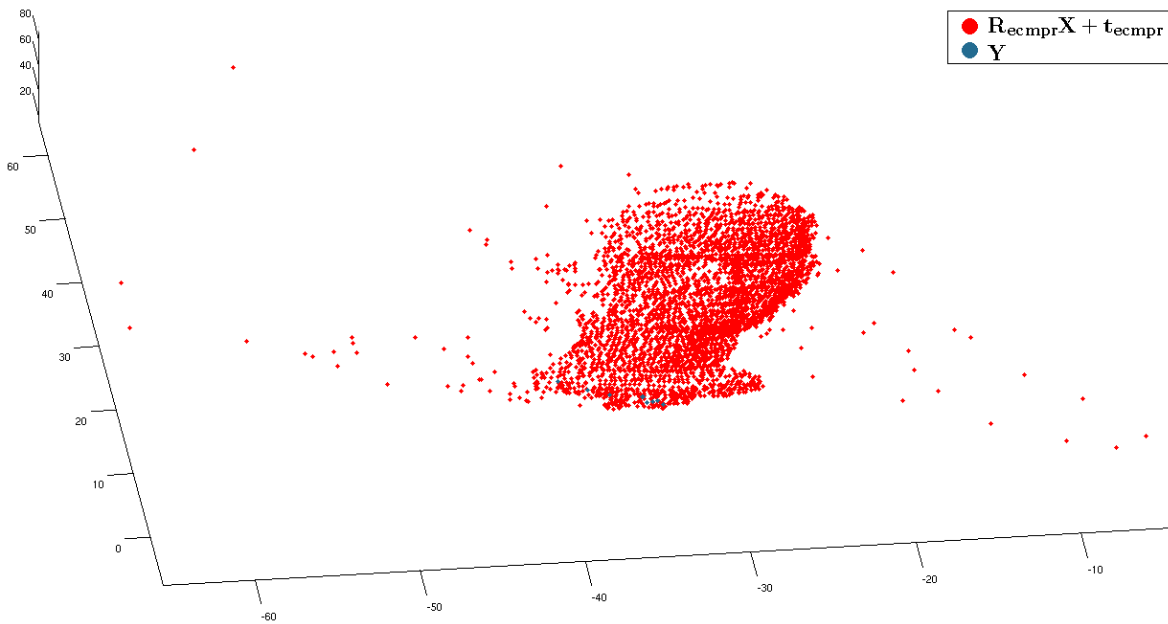


Σχήμα 7.1: Αρχική θέση των δύο προς αντιστοίχιση νεφών

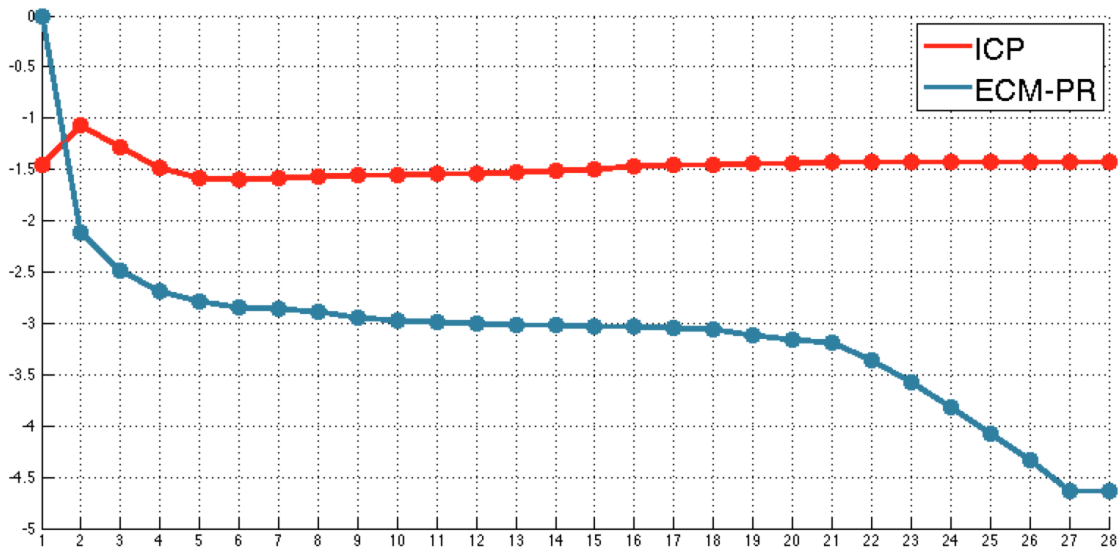


Σχήμα 7.2: Οπτικό αποτέλεσμα της αντιστοίχισης που επιτεύχθηκε μετά από 28 επαναλήψεις με τον αλγόριθμο ICP. Οι αντιστοιχίες αναζητώνται από το μικρότερο πληθυσμό στο μεγαλύτερο χωρίς επαναχρησιμοποίηση των σημείων.



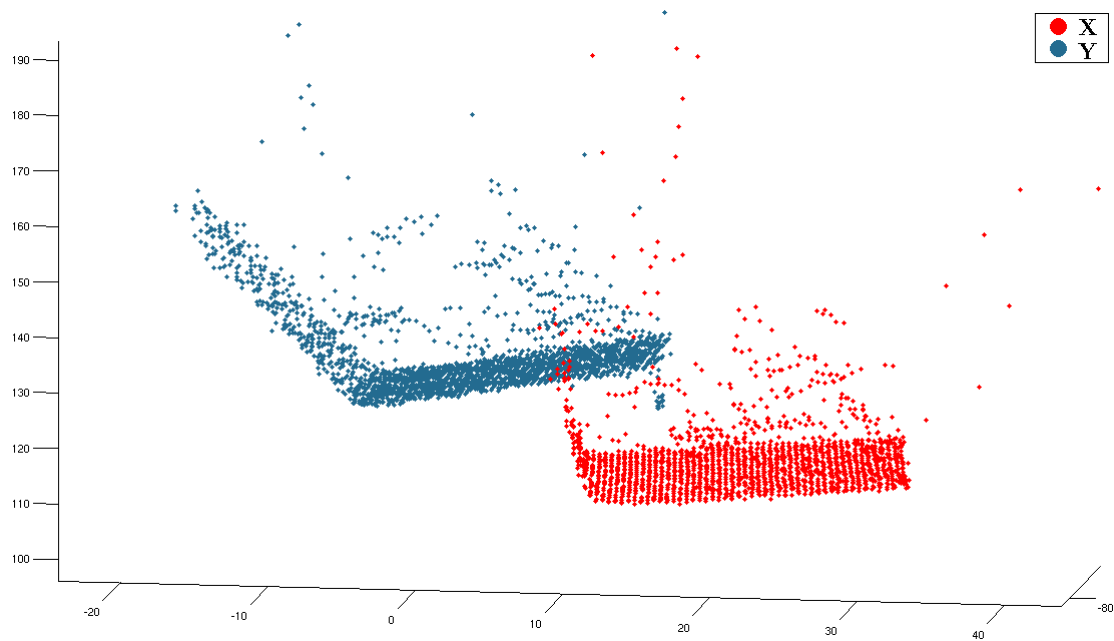


Σχήμα 7.3: Οπτικό αποτέλεσμα της αντιστοίχισης που επιτεύχθηκε μετά από 28 επαναλήψεις<sup>2</sup> με τον αλγόριθμο ECM-PR. Το μόνο στοιχείο που υποδηλώνει την ύπαρξη του  $Y$  είναι τα 8 εναπομένοντα  $\bullet$  σημεία στο κάτω μέρος του αυχένα.

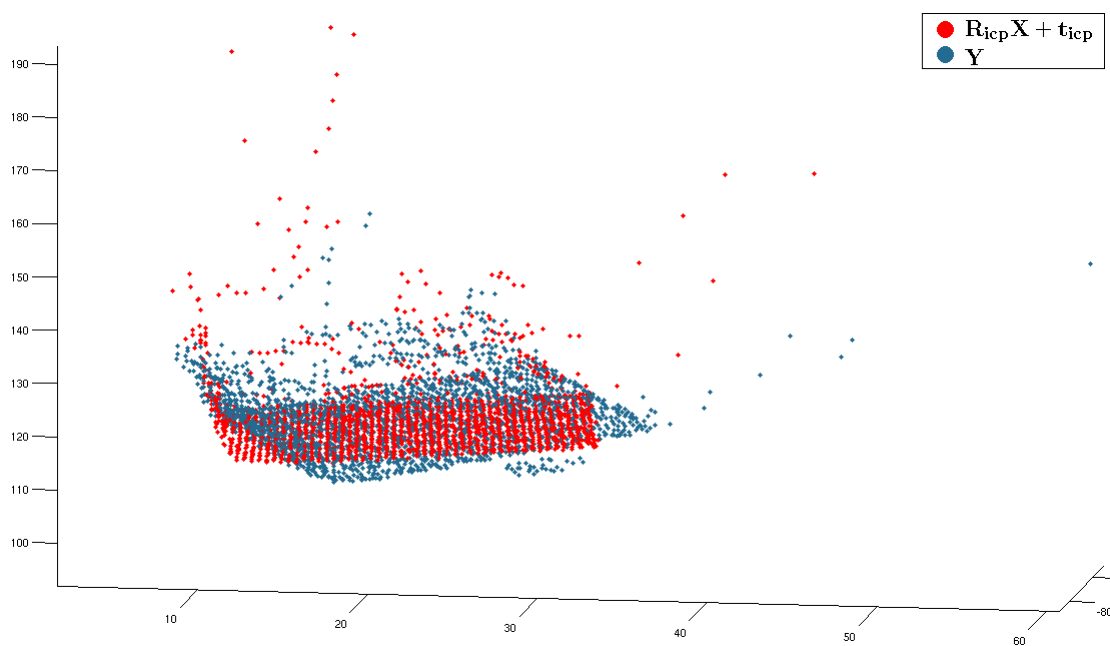


Σχήμα 7.4: Κανονικοποιημένες τιμές του  $\sum_{i=1}^n \|y_i - R\mathbf{x}_i - \mathbf{t}\|_2^2$  μεταξύ αντιστοιχών στοιχείων, σε λογαριθμική κλίμακα για 28 επαναλήψεις.

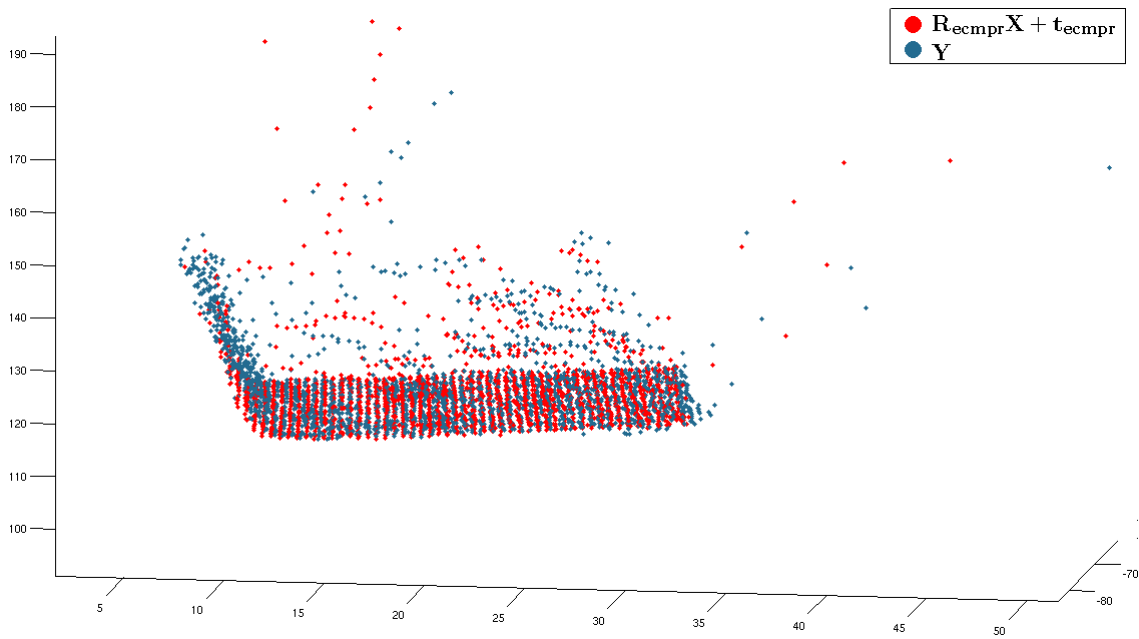
<sup>2</sup>Όπως φαίνεται και στο Σχήμα (34) 28 επαναλήψεις ήταν αρκετές για να σταθεροποιηθούν και οι δύο αλγόριθμοι.



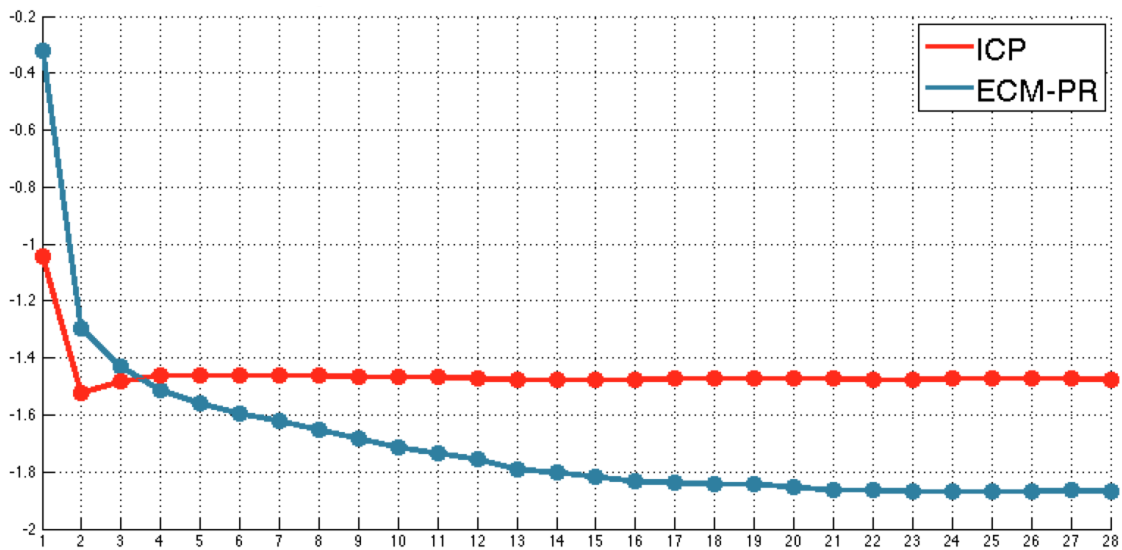
Σχήμα 7.5: Αρχική θέση των δύο προς αντιστοίχιση νεφών τα οποία αναπαριστούν μια πλάγια όψη ενός χαρτοκυτίου.



Σχήμα 7.6: Οπτικό αποτέλεσμα της αντιστοίχισης που επιτεύχθηκε μετά από 28 επαναλήψεις με τον αλγόριθμο ICP.



Σχήμα 7.7: Οπτικό αποτέλεσμα της αντιστοίχισης που επιτεύχθηκε μετά από 28 επαναλήψεις με τον αλγόριθμο ECM-PR.



Σχήμα 7.8: Κανονικοποιημένες τιμές της προς ελαχιστοποίηση μετρικής για κάθε επανάληψη, σε λογαριθμική κλίμακα.

Παρατηρούμε ότι και στα δύο πειράματα ο αλγόριθμος ECM-PR επιτυγχάνει αρκετά καλύτερα οπτικά αποτελέσματα από τον ICP. Συγκρίνοντας τις τιμές της μετρικής σφάλματος για τους δύο αλγόριθμους παρατηρούμε ότι στο πρώτο πείραμα (κεφαλή) ο ICP εκλωβίζεται σε τοπικό ελάχιστο ενώ ο ECM-PR αν και

Ξεκινάει από πιο λανθασμένη εκτίμηση (στις πρώτες επαναλήψεις) καταφέρνει να βελτιώσει το αποτέλεσμα του πριν κλειδώσει<sup>3</sup>. Στο δεύτερο πείραμα παρατηρούμε ότι ενώ και οι δύο αλγόριθμοι έχουν υπολογίσει σχεδόν την ίδια τιμή για την συνάρτηση κόστους Το οπτικό αποτέλεσμα του ICP απέχει πάρα πολύ από το βέλτιστο<sup>4</sup> ενώ αντίθετα ο ECM-PR βρίσκεται πολύ κοντά σε αυτό. Ο λόγος για τον οποίο παρατηρείται αυτή η "ταλάντωση" στη συγκλίση του ECM-PR, στις αρχικές επαναλήψεις, οφείλεται στις αυθαίρετες τιμές που αναθέσαμε, αρχικά, στα μητρώα συνδιασπορών<sup>5</sup>. Οι συνδιασπορές αυτές επηρεάζουν την πρώτη εκτίμηση των  $\mathbf{R}$ ,  $\mathbf{t}$  με αποτέλεσμα αυτή να είναι παραπλανητική.<sup>6</sup>**ΗΜΕΙΩΣΗ:** Η οπτική γωνία των screenshots του πρώτου πειράματος (αντιστοίχισης κεφαλής) έχει Αζιμουθιο (azimuth)  $-7$  και ανύψωση (elevation)  $70$  ενώ η οπτική γωνία των screenshots του δεύτερου πειράματος (χυτίο) έχει αζιμουθιο  $9$  και ανύψωση  $17$ . Οι οπτικές αυτές γωνίες δεν έχουν κάποια πρακτική σημασία, απλώς επιλέχθηκαν, για καλύτερη οπτικοποίηση των αποτελεσμάτων στις 2 διαστάσεις.

### 7.1.2 Με χρήση τεχνητών δεδομένων

Όντας περισσότερο διαισθητικά κατανοητά τα πειράματα αντιστοίχισης πραγματικών δεδομένων παρατέθηκαν πρό των πειραμάτων τεχνητών δεδομένων. Για την αυστηρή σύγκριση δύο αλγορίθμων απαιτείται μια πιο λεπτομερής μελέτη η οποία θα περιέχει ποσοτικοποίηση του σφάλματος απόκλισης. Για την ποσοτικοποίηση αυτών των αποκλίσεων απαιτείται η γνώση του πραγματικού λανθάνοντος μετασχηματισμού των δύο νέφων. Για να έχουμε την γνώση αυτού του μετασχηματισμού τα νέφη των πειραμάτων που διεξάγαγαμε προέρχονταν από ένα νέφος μοντέλο<sup>7</sup> το οποίο μετασχηματίζαμε, εισάγαμε θόρυβο και τέλος δειγματοληπτούσαμε ώστε να προκύψει το προς αντιστοίχιση ζεύγος. Παρατηρώντας τα πραγματικά δεδομένα έγινε αντιληπτό ότι ο θόρυβος μπορεί να μοντελοποιηθεί σε δύο κυρίαρχες μορφές, αυτής των μολυσμένων σημείων (outliers) και αυτής της διασποράς των πραγματικών μετρήσεων (dispersity). Η εισαγωγή διασποράς στα δείγματα επιτεύχθηκε μέσω της συνάρτησης `awgn` της MATLAB η οποία εισάγει ιστροπικό προσθετικό (λευκό) θόρυβο, Gaussian κατανομής και δεδομένου SNR, στα δείγματα ενός συνόλου. Από την άλλη η εισαγωγή outliers επιτεύχθηκε μέσω της συνάρτησης `outlier_constructor`<sup>8</sup> η ιδέα της οποίας είναι η επιλογή ενός πλήθους  $n$  τυχαίων σημείων από το κυρτό περίβλημα (convex hull) του νέφους-μοντέλου κάθε ένα εκ των οποίων χρησιμοποιείται ως κέντρο μίας Gaussian κατανομής η οποία δημιουργεί ένα υποσύνολο των outliers που θα εισαγθούν τελικά στο νέφος. Στα πειράματα που θα ακολουθήσουν το πλήθος  $n$  επιλέγεται ίσο με  $10$  και το συνολικό πλήθος των προς εισαγωγή outliers ισομοιράζεται στις  $n$  συνιστώσες. Η διασπορά κάθε μίας εκ των  $n$  κατανομών επιλέγεται ίση με το  $1/n$  της φασματικής ακτίνας του μητρώου συνδιασπορών `sample covariance matrix` του νέφους-μοντέλου<sup>9</sup>.

<sup>3</sup>Παρατηρούμε ότι περεταίρω επαναλήψεις είναι άσκοπες αφού και οι δύο αλγόριθμοι έχουν κλειδώσει σε στάσιμο σημείο.

<sup>4</sup>Η συνάρτηση κόστους παρουσιάζει δύο ακρότατα με την ίδια σχεδόν τιμή εκ των οποίων μόνο το ένα είναι το οπτικά βέλτιστο.

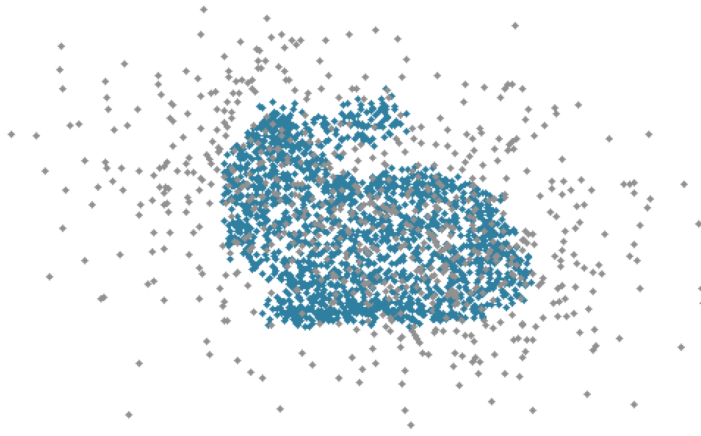
<sup>5</sup>Συγκεκριμένα όλες οι συνιστώσες του  $\mathbf{\Sigma}$  αρχικοποιούνται με διασπορά  $10^3$ .

<sup>6</sup>Μόλις από την δεύτερη κιόλας επανάληψη οι συνδιασπορές αποκτούν τιμές της τάξης του  $10^{-1}$ .

<sup>7</sup> Συγκεκριμένα το stanford bunny.

<sup>8</sup> Αποτελεί ιδιοκατασκευή, ο κώδικας της οποίας παρατίθεται στο [κεφάλαιο ;].

<sup>9</sup>Η φασματική ακτίνα υπολογίζεται πρό της εισαγωγής του θορύβου διασποράς (dispersity noise)



Σχήμα 7.9: Παράδειγμα νέφους (μοντέλου) με προσθετικό θόρυβο (blue dots) και μόλυνση (outliers - gray dots) πλήθους 40% του αρχικού του πλήθους και θορύβου διασποράς με SNR ίσο με 25.

Για την πιο λεπτομερή σύγκριση των αλγορίθμων ICP και ECM-PR διεξαγάγαμε τριών ειδών πειράματα. Ένα στο οποίο έχει γίνει μόνο προσθήκη θορύβου (dispersity noise), ένα μόνο με μολυσμένα σημεία, και ένα με την ύπαρξη και των δύο. Ως λανθάνων μετασχηματισμός, σε όλα τα πειράματα, επιλέχθηκε μία περιστροφή κατά  $\pi/6$  και μια μετάφραση<sup>10</sup> κατά 5.4 προς όλες τις κατευθύνσεις. Ως μετρική ποσοτικοποίησης της απόκλισης επιλέξαμε<sup>11</sup> την Frobenius norm της διαφοράς του πραγματικού μητρώου περιστροφής  $\mathbf{R}_{groundtruth}$  και της εκτίμησης  $\mathbf{R}$  αυτού, από τον εκάστοτε αλγόριθμο, για την τρέχουσα επανάληψη. Τέλος να σημειώσουμε ότι τα πλήθος των πραγματικών σημείων των προς αντιστοίχιση νεφών είναι 1798 για το νέφος-μοντέλο και 1748 για το νέφος το οποίο έχει υποστεί τον μετασχηματισμό και την θορυβοποίηση.

Παρατηρούμε ότι ο αλγόριθμος ECM-PR εξαιτίας της στοχαστικής φύσης του επιτυγχάνει να εκτιμήσει το λανθάνον μητρώο (φθίνουσα γραφική) περιστροφής ακόμα και σε περιπτώσεις ύπαρξης ισχυρού<sup>12</sup> θορύβου. Παρατηρούμε επίσης ότι ο αλγόριθμος ICP στις επιτυχείς περιπτώσεις ασθενούς θορύβου (κιτρινες γραμμές) προσεγγίζει πιο καλά το πραγματικό μητρώο περιστροφής σε αντίθεση με τον ECM-PR που κλειδώνει λίγο νωρίτερα.

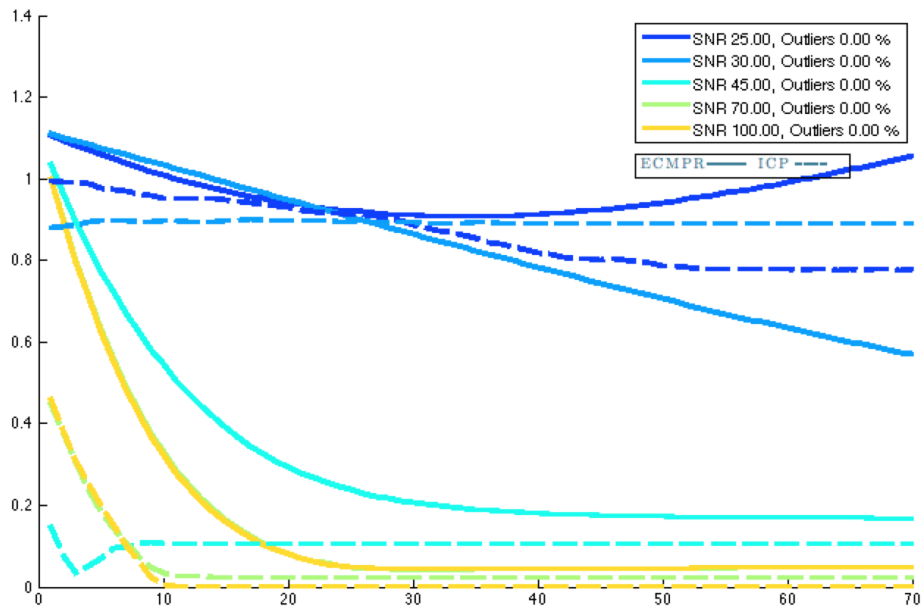
Απο την παραπάνω γραφική γίνεται εμφανές το πόσο υπερτερεί ο αλγόριθμος ECM-PR έναντι του ICP. Ο ECM-PR κατάφερε να συγκλίνει ακόμα και υπο την ύπαρξη πλήθους outliers όσα και τα πραγματικά σημεία ενώ ο ICP συνέκλινε μόνο για την περίπτωση ποσοστού outliers ίσο με το 10% των πραγματικών σημείων.

Η πιο ρεαλιστική περίπτωση είναι αυτή στην οποία έχουμε και τις δύο μορφές θορύβων. Παρατηρούμε και εδώ ότι ο αλγόριθμος ECM-PR καταφέρνει να συγκλίνει σχεδόν σε όλες τις περιπτώσεις έστω και με μειωμένη ταχύτητα σύγκλισης

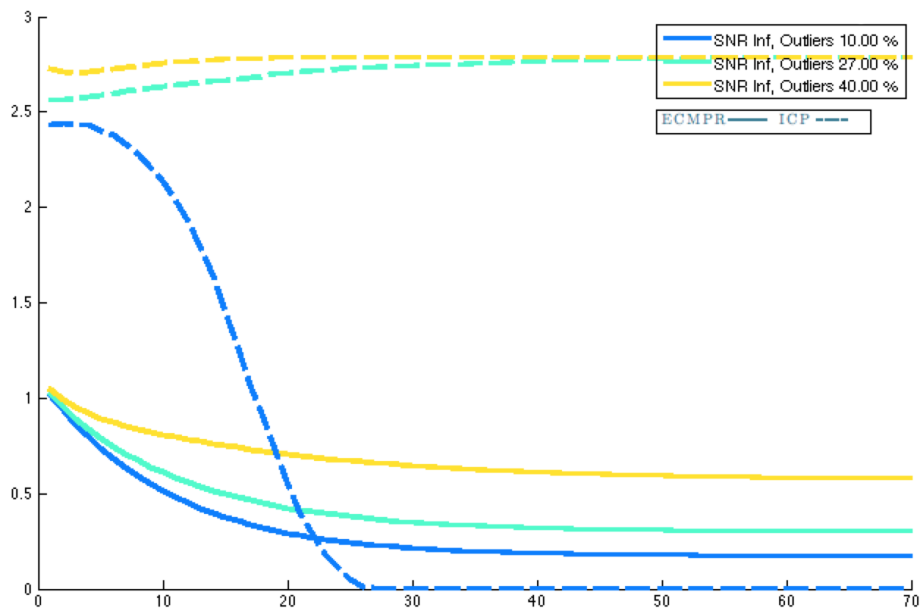
<sup>10</sup>Η μετάφραση επηρεάζει ελάχιστα την αντιστοίχιση (όσο μεγάλη και αν είναι), επίσης να σημειώσουμε ότι εκτιμάται με πολύ μεγάλη ακρίβεια και από τους δύο αλγορίθμους και για αυτό το λόγο δεν λαμβάνεται υπόψη κατά την ποσοτικοποίηση των αποκλίσεων.

<sup>11</sup>Αποφύγαμε να χρησιμοποιήσουμε την  $l_2$  νόρμα εξαιτίας της μεγάλης επιροής της από την κυρίαρχη συνιστώσα του μητρώου διαφορών  $\mathbf{R} - \mathbf{R}_{groundtruth}$ .

<sup>12</sup>Να σημειώσουμε ότι κάτω από τα 25 SNR το νέφος χάνει πλέον την μορφή του (το Σχήμα του λαγού αλιώνεται σε τέτοιο βαθμό που δεν είναι πλέον οπτικά αναγνωρίσιμο).



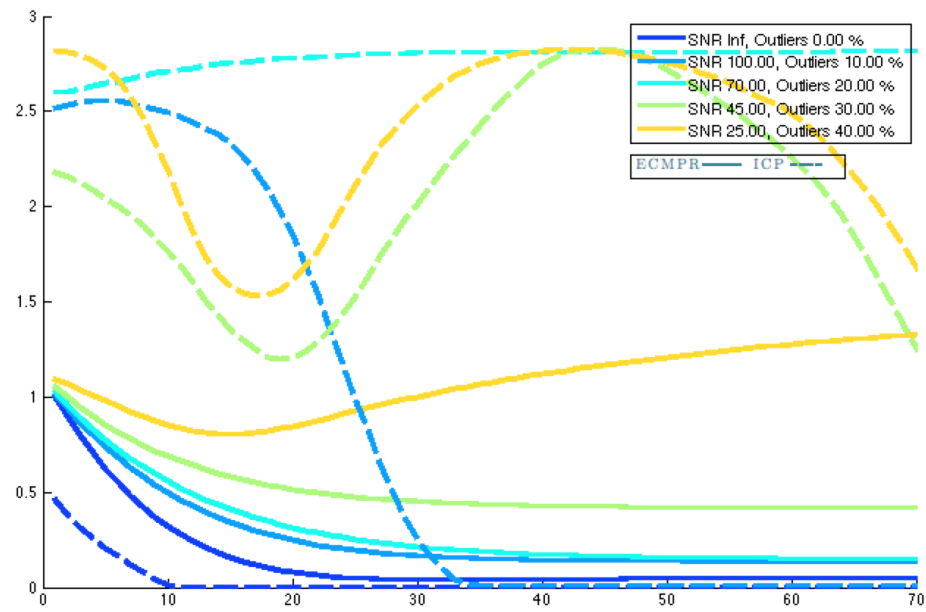
Σχήμα 7.10: Μετρήσεις αποκλίσεων των αλγορίθμων ECM-PR (ενιαία γραμμή) και ICP (διακεκομμένη γραμμή) για 70 επαναλήψεις σε αντιστοίχιση δεδομένων που έχουν υποστεί προσθήκη θορύβου διασποράς.



Σχήμα 7.11: Μετρήσεις αποκλίσεων των αλγορίθμων ECM-PR και ICP για 70 επαναλήψεις σε αντιστοίχιση δεδομένων που έχουν υποστεί προσθήκη outliers.

σε αντίθεση με τον ICP που αποτυγχάνει σε όλες τις περιπτώσεις<sup>13</sup>

<sup>13</sup>Η περίπτωση μηδενικού θορύβου και μηδενικών outliers προφανώς είναι ιδεατή και παρέχεται μόνο για λόγους σύγκρισης (παρέχει την ελάχιστη δυνατή απόκλιση που μπορεί να έχει ένας

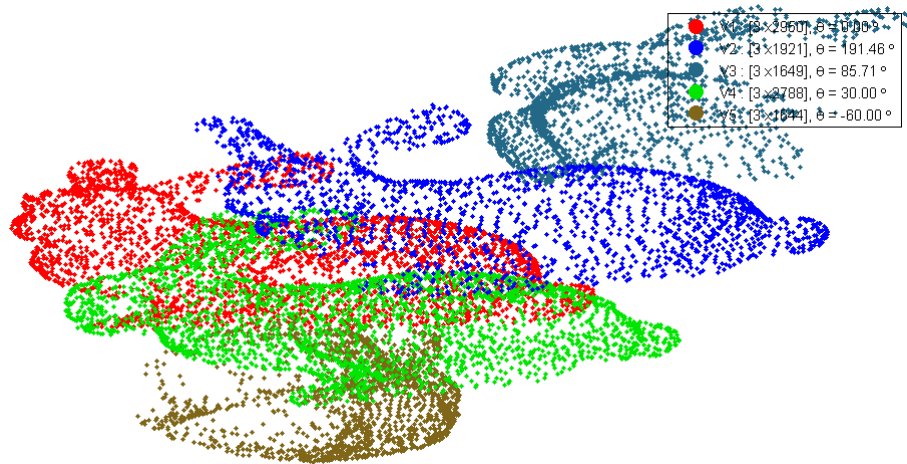


Σχήμα 7.12: Μετρήσεις αποκλίσεων των αλγορίθμων ECM-PR και ICP για 70 επαναλήψεις σε αντιστοίχιση δεδομένων που έχουν υποστεί προσθήκη θορύβου διασποράς και μόλυνσης των δεδομένων.

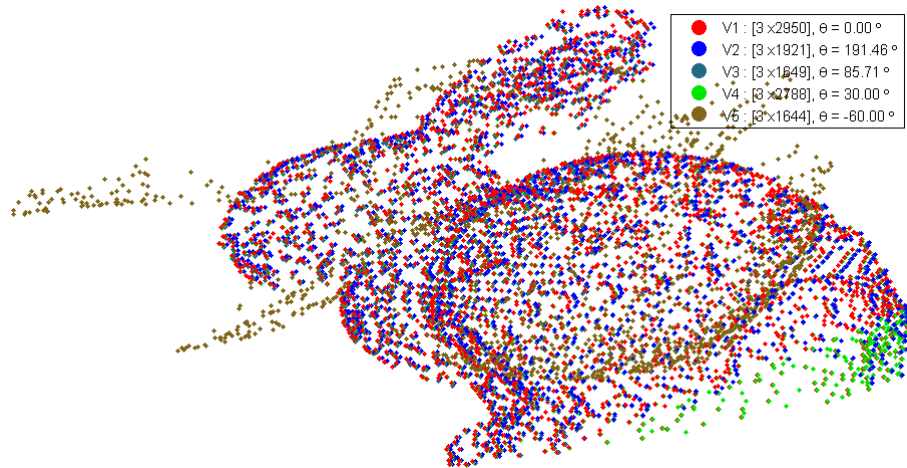
## 7.2 Πειράματα αντιστοίχισης πολλαπλών νεφών

Σύγκριση των αλγορίθμων ICP πολλαπλών όψεων [Ενότητα 5.2.1], και ECM-PR πολλαπλών όψεων [κεφάλαιο 6].

### 7.2.1 Οπτικά αποτελέσματα

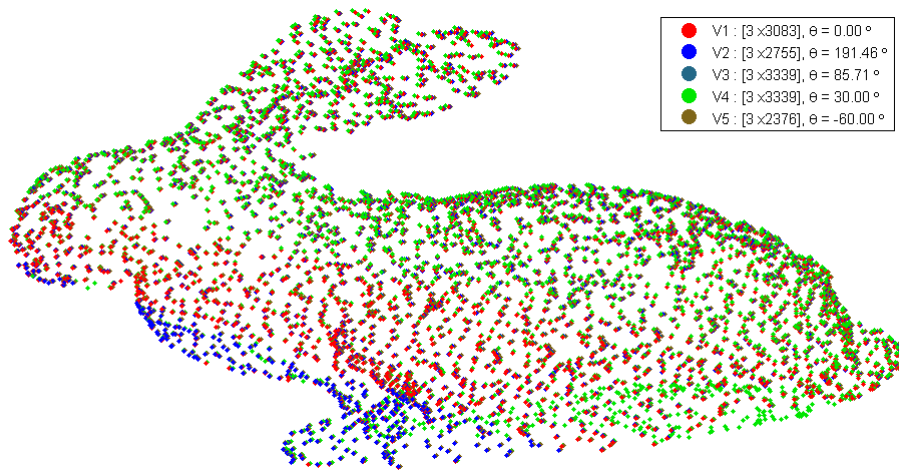


Σχήμα 7.13: Αρχική θέση των προς αντιστοίχιση νεφών. Όλα τα νέφη έχουν προέλθει από μετασχηματισμό και υποδειγματοληψία του πρώτου  $V\{1\}$ .

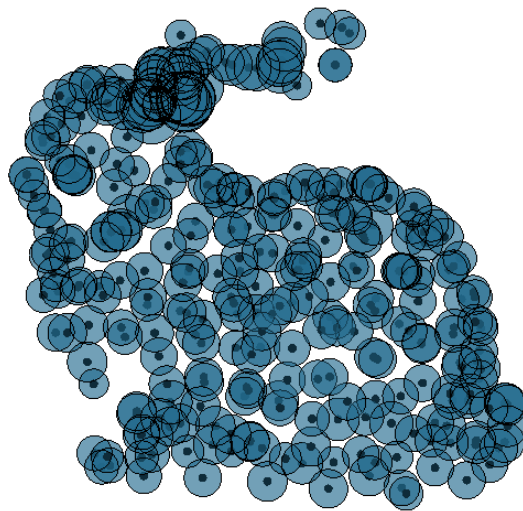


Σχήμα 7.14: Τελική αντιστοίχιση με χρήση του ICP πολλαπλών όψεων μετά από 120 επαναληψεις. Παρατηρούμε ότι αποτυγχάνει να αντιστοιχίσει την γωνία  $-\frac{\pi}{3}$ .





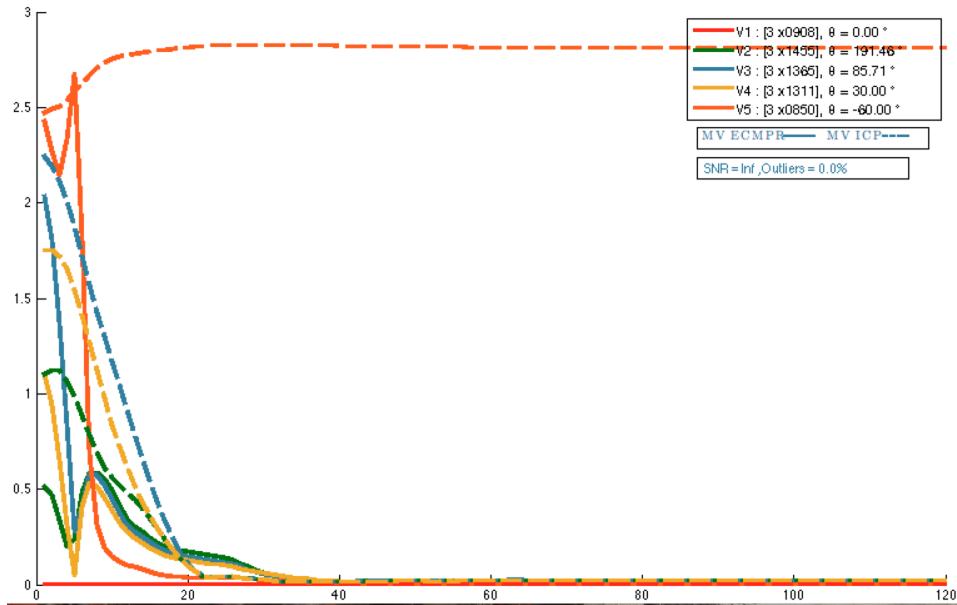
Σχήμα 7.15: Τελική αντιστοίχιση με χρήση του ECM-PR πολλαπλών όψεων μετα απο 120 επαναληψεις.



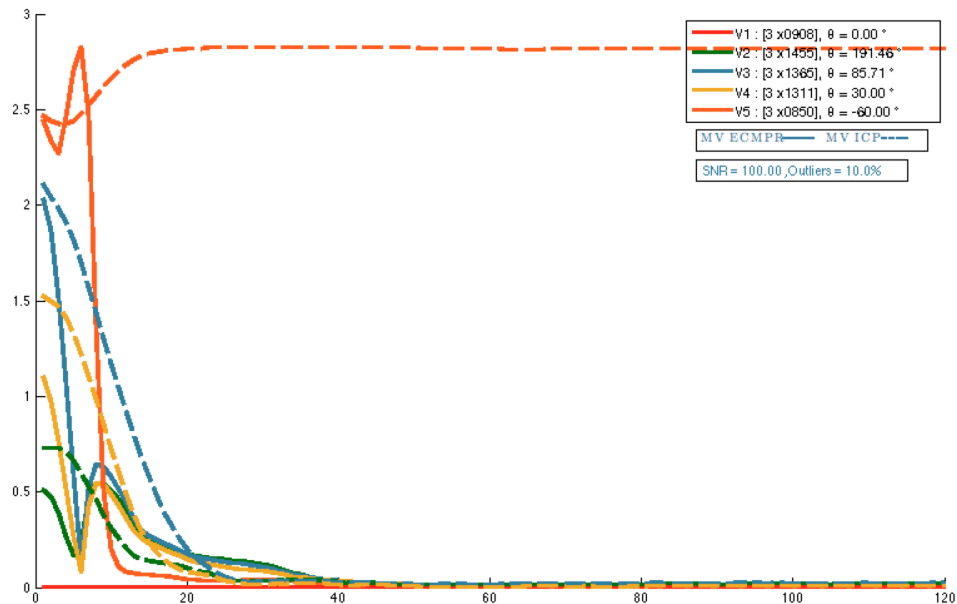
Σχήμα 7.16: Οπτικοποίηση των τελικών 280 κέντρων των κατανομών του ECM-PR πολλαπλών όψεων καθώς και των διασπορών τους. Τα αρχικά κέντρα επιλέχθηκαν τυχαία χωρίς κάποια δομή.

### 7.2.2 Ποσοτικοποίηση των αποκλίσεων

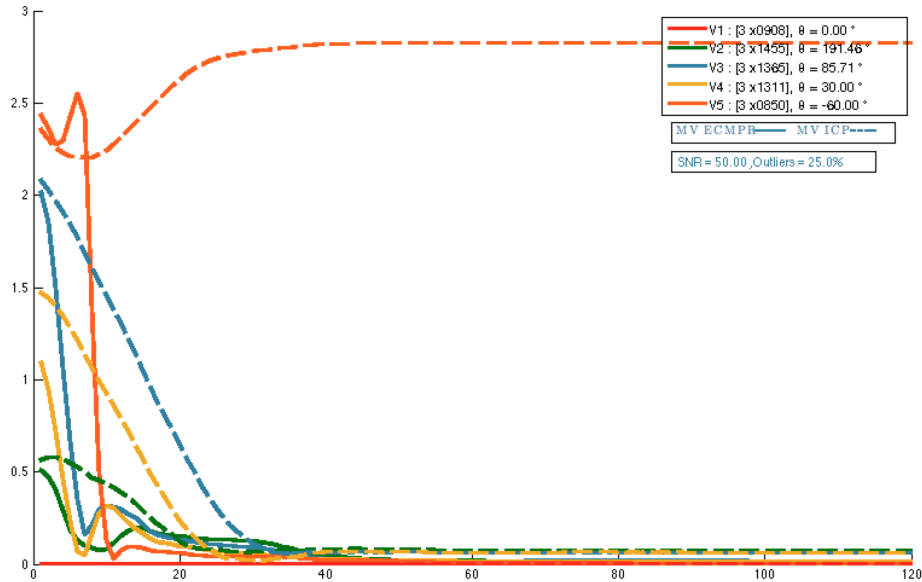
Μετρήσεις αποκλίσεων, κάθε όψης, για διάφορες τιμές (ισχύος) προσθετικού θορύβου και ποσοστού μόλυνσεως.



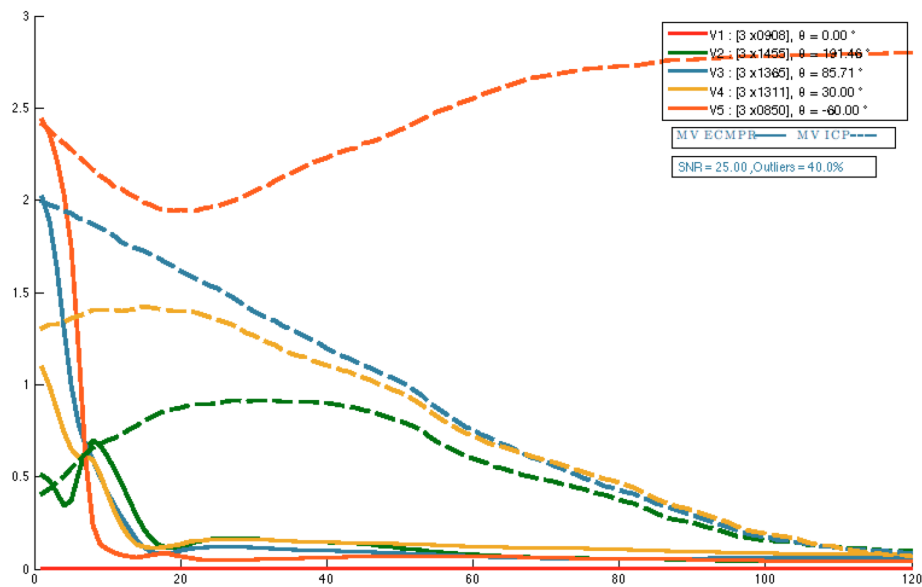
Σχήμα 7.17: Σύγκλιση των μητρών περιστροφής κάθε όψης, επαναπροβεβλημένων στην  $V\{1\}$ , για την ιδεατή περίπτωση μηδενικού θορύβου.



Σχήμα 7.18: Σύγκλιση των αλγορίθμων παρουσία θορύβου.



Σχήμα 7.19: Η αναπήδηση του ECM-PR πιθανώς οφείλεται στο ότι τα κέντρα του GMM δεν έχουν ισορροπήσει ακόμα.



Σχήμα 7.20: Η γραφική αυτή παρατίθεται ως οριακή κατάσταση στην οποία ο θόρυβος αρχίζει να υπερνικά το σήμα (Σχήμα (7.9)).



## Παράρτημα Α΄

# ΥΛΟΠΟΙΗΣΕΙΣ

Στις μεθόδους που παρουσιάσαμε, ως επι το πλείστον, απαιτείται υπολογισμός Ευκλείδειων αποστάσεων. Λόγο της φύσης της MATLAB οι elementwise-Hadamard πολλαπλασιασμοί, λόγω μειωμένης τοπικότητας, καθυστερούν την υνολική εκτέλεση. Για την επιτάχυνση των υλοποιήσεων, ο υπολογισμός των (τετραγώνων των) Ευκλείδειων αποστάσεων έγινε στην γλώσσα C με τη βοήθεια του περιβάλλοντος MEX. Σε περίπτωση αδυναμίας χρήσης της mex function `mex_sqe.c` που ακολουθεί ο αναγνώστης μπορεί να χρησιμοποιήσει την MATLAB υλοποίηση αυτής `mex_sqe.m` που παρατίθεται αμέσως μετά. **ΣΗΜΕΙΩΣΗ** : Για την μεταγλωττισή της `mex_sqe.c` απαιτείται το αρχείο `common.c`<sup>1</sup>.

### `mex_sqe.c`

```
1 /* =====*
2  * sqdist2new.c
3  * Squared Euclidean distances
4  *
5  * c. Georgios Evangelidis
6  *
7  * This function computes the squared Euclidean distances
8  * between two matrices Y, TX, of sizes [3 x m], [3 x n]
9  * respectively. It outputs a [m x n] matrix D whose j,i
10 * element is equal with ||Y(:,j)-TX(:,i)||^2_2 .
11 *
12 * mex_sqe(Y,TX) =
13 *
14 * for i=1:n
15 *     for j=1:m
16 *         D(j,i) = sum((Y(:,j)-X(:,i)).^2);
17 *     end
18 * end
19 *
20 * Where sz(X) = 3xn & sz(Y) = 3xm
21 * =====*
```

<sup>1</sup> <http://www.mathworks.com/matlabcentral/fileexchange/26825-utilities-for-mex-files/content/example/sqdist2.m>

```

22  */
23  #define _PACKAGENAME_ "math:sqdist2"
24  #include "common.c" /* common.c by Levente Hunyadi */
25
26  mxArray* distmatmat(const mxArray* mat1, const mxArray*
    mat2) {
27
28      const double* pmat1 = mxGetPr(mat1);
29      const double* pmat2 = mxGetPr(mat2);
30
31      int m1 = mxGetM(mat1);
32      int n1 = mxGetN(mat1);
33      int m2 = mxGetM(mat2);
34      int n2 = mxGetN(mat2);
35
36      mxArray* res;
37      double* pr;
38      int i,j,k;
39
40      res = mxCreateDoubleMatrix(n1,n2, mxREAL);
41      pr = mxGetPr(res);
42
43      for (i = 0; i < n1; i++) {
44  for (j=0; j<n2; j++) {
45          double *p1 = pmat1 + 3*i;
46          double *p2 = pmat2 + 3*j;
47          double r = 0.0;
48          for (k = 0; k < 3; k++) {
49              double d = *p1 - *p2;
50              r+= d*d;
51              p2++;
52              p1++;
53          }
54          pr[i+j*n1]=r;
55          /* *(pr++) = r;*/
56      }
57  }
58      return res;
59  }
60  /**
61  * Entry point. [ inputs should be checked by the user ]
62  */
63  void mexFunction(int nlhs, mxArray* plhs[], int nrhs,
    const mxArray* prhs[]) {
64  plhs[0] = distmatmat(prhs[0], prhs[1]);
65  }

```

---

Υλοποίηση της mex\_sqe.c σε περιβάλλον MATLAB.

**mex\_sqe\_m.m**

---

```

1 % Replacement for mex_sqe.c for non mex equipped machines.
2 function D = mex_sqe_m(Y,TX)
3
4     m = size(Y,2);
5     n = size(TX,2);
6
7     D(m,n) = 0; % ~2000 times faster than zeros
8
9     for i=1:n
10         D(:,i) = sum((Y-repmat(TX(:,i),1,m)).^2);
11     end

```

---

## A'.1 Υλοποιήσεις αντιστοίχισης ζεύγους

Υλοποίηση του αλγορίθμου ECM-PR, [Ενότητα 4.2.3].

### **ecmpr.m**

---

```

1 function [R,t,iter,vr,degenSign,ssed,cverr] = ...
2     ecmpr(X,Y,varargin)
3 %ECMPR Expectation Conditional Maximization – Point Regi
4 % stration: Computes the Optimum Rigid Transformati
5 % on Parameters (R,t) in order to align and registe
6 % r the 3D point clouds X and Y.
7 %
8 % MEX optimized [DEPENDENCIES: mex_sqe.c]
9 %
10 % [R,t] = ecmpr(X,Y) returns the optimum transformation p
11 % arameters for the registration of point clouds X and Y
12 % in the sense that Y = RX+t.
13 %
14 % X,Y must be matrices of 3Dimensional points organized a
15 % s columns [3 x ..] respectively.
16 %
17 % OPTIONAL INPUT arguments can be given as string-value p
18 % airs
19 %
20 % [R,t] = ecmpr(X,Y,...,'param1',val1..,'param2',val2,..)
21 %
22 %
23 %
24 % Input parameters include :
25 %
26 % 'theta' Initial assumption for the a
27 % ngle between X and Y ,if not
28 % provided the algorithm assum
29 % estheta = 0 as a common angl

```

```

30 %                                     e for the 3 axes xyz.
31 %
32 %
33 % 'R'                                 Is a [3x3] Rotation Matrix u
34 %                                     sed as an initial assumption
35 %                                     of the pose between X and Y
36 %                                     If 'R' is provided 'theta' v
37 %                                     alue is ignored.
38 %
39 %
40 % 't'                                 a [3x1] vector used as initi
41 %                                     al estimation of the transla
42 %                                     tion between the two point c
43 %                                     louds. Default is the zero
44 %                                     vector.
45 %
46 %
47 % 'mode'                               Is a string value specifying
48 %                                     the parameterization of the Ga
49 %                                     ussian Mixture Model used fo
50 %                                     r ECM Can be 'global' or 'lo
51 %                                     cal'.
52 %
53 %                                     'global' :
54 %
55 %                                     A common isotropic covarianc
56 %                                     e matrix  $S^2 \times I$  is assume
57 %                                     d for all gaussian density c
58 %                                     omponents.
59 %
60 %                                     'local' :
61 %
62 %                                     Each mixture component is ch
63 %                                     aracterized by its own isotr
64 %                                     opic covariance matrix
65 %                                      $S(i)^2 \times I$ 
66 %
67 %                                     Default => 'local'
68 %
69 %
70 % 'S'                                 Is a scalar or Vector of ini
71 %                                     tial standard deviations for
72 %                                     each mixture component. S is
73 %                                     scalar when 'mode' = 'global
74 %                                     ' or a vector of length(X) w
75 %                                     hen 'mode' = 'local', if us
76 %                                     er provides a scalar when mo
77 %                                     de = 'local' that same scala
78 %                                     r is assigned to all compone
79 %                                     nts.

```



```

80 %
81 %
82 %
83 %
84 %
85 %
86 %
87 %
88 % 'epsilon'
89 %
90 %
91 %
92 %
93 %
94 %
95 %
96 %
97 %
98 %
99 %
100 %
101 %
102 %
103 %
104 % 'maxNumIter'
105 %
106 %
107 %
108 % 'endeVal'
109 %
110 %
111 %
112 %
113 %
114 %
115 %
116 %
117 %
118 %
119 % 'radius'
120 %
121 %
122 %
123 %
124 %
125 %
126 %
127 %
128 %
129 %

```

Default is 1000 for all components. WARNING : must be large to avoid local maxima. It is the reason of the oscillations appearing at initial iterations.

A scalar value added at the diagonal elements of each covariance Matrix after the iteration's update to artificially flatten them in order to avoid singularities created when the exponent of a gaussian density tends to zero or equivalently a transformed point from Y collapses upon a component centroid from X.

Default is 0.7

Specifies the maximum number of ECM iterations Default =7

Scalar which specifies a termination condition. The algorithm stops when the frobenius norm of the difference between the current and former estimation of the Rotation matrix (R) becomes less than this threshold, specifically  $\|R_{(i-1)} - R_i\|_{\text{fro}}$ . Default =  $1e-7$

Specifies the radius of the sphere which characterizes the uniform distribution of the prior probabilities for each point  $X(:,i)$  to be an inlier. Default =  $1e-40$ . WARNING : r values must be infinitesimally small but not zero! Protects SVD from NaN & Inf's.

```

130 %
131 %   'groundTruth'           Ground Truth rotation matrix
132 %                           for measuring convergence
133 %                           [see output argument cverr]
134 %
135 %
136 % OUTPUT ARGUMENTS :
137 %
138 %
139 %   'R'                       The optimal rotation Matrix
140 %                           which better aligns X and Y
141 %                           in the sense that the sum of
142 %                           squared distances between th
143 %                           e points are minimized.
144 %
145 %   't'                       The optimal translation vect
146 %                           or between X and Y.
147 %
148 %
149 %   'iter'                     Is the number of iterations
150 %                           executed until the terminati
151 %                           on, it will be less or equal
152 %                           to maxNumIter depending on
153 %                           if endeVal was reached earli
154 %                           er.
155 %
156 %   'vr'                       Is the Variation of Rotation
157 %                           s, i.e. how much does R chan
158 %                           ge between two consecutive i
159 %                           terations computed as the
160 %                           frobenius norm of the diffe
161 %                           rence between Rq-1 and Rq:
162 %                           ||Rq-R{q-1}||_F.
163 %
164 %
165 %   'degenSign'               A binary value if it is 1 me
166 %                           ans the algorithm fell into
167 %                           a degenerate case, i.e,
168 %                           |U||V| = -1 So a Reflection
169 %                           provided a better result.
170 %
171 %
172 %   'ssd'                     IF REQUESTED: Is a vector co
173 %                           ntaining the value of the mi
174 %                           nimized metric, i.e the sum
175 %                           of squared Euclidean distanc
176 %                           es between each Yj and Xi.
177 %
178 %
179 %   'cverr'                   If 'groundTruth' provided is

```

```

180 %                                     a vector containing the conv
181 %                                     ergence differences between
182 %                                     the ground truth rotation
183 %                                     and the current estimation
184 %                                     of the rotation matrix, in
185 %                                     the frobenious sense.
186 %
187 %
188 % $ Last revisited : Wed, 29 August, 23:41 PM $
189
190 %% ===== PARSE INPUTS ===== %%
191 [rx ,n] = size(X);
192 [ry ,m] = size(Y);
193
194 if( rx ~= 3 || ry ~= 3)
195     error('Input arrays must be of size [3 x ..]');
196 end
197
198 %% ===== inputParser Struct ===== %%
199 p = inputParser;
200 addOptional(p, 'theta', 0, @(x) isscalar(x) && ~ischar(x));
201 addOptional(p, 'S', 1e3, @isvector);
202 addOptional(p, 't', zeros(3,1), @(x) isvector(x) && ...
203     numel(x)==3);
204 addOptional(p, 'R', NaN(3), @(x) size(x,1)==3 && ...
205     size(x,2)==3);
206 addOptional(p, 'epsilon', 1e-7, @isscalar);
207 addOptional(p, 'maxNumIter', 40, @(x) (round(x)==x) & x>0);
208 addOptional(p, 'endeVal', 1e-7, @(x) isreal(x) && ...
209     isscalar(x) && ~ischar(x));
210 addOptional(p, 'radius', .7, @(x) gt(x,0));
211 addOptional(p, 'mode', 'local', @(x) strcmp(x, 'local') || ...
212     strcmp(x, 'global'));
213 addOptional(p, 'groundTruth', NaN);
214 parse(p, varargin{:});
215 %% ===== %%
216
217
218 %% ===== rename struct values ===== %%
219 theta = p.Results.theta;
220 mode = p.Results.mode;
221 S = p.Results.S(:);
222 t = p.Results.t(:);
223 R = p.Results.R;
224 epsilon = p.Results.epsilon;
225 maxNumIter = p.Results.maxNumIter;
226 endeVal = p.Results.endeVal;
227 r = p.Results.radius;
228 Rgt = p.Results.groundTruth;
229 %% ===== %%

```

```

230
231 if nargout > 6 && nnz(isnan(Rgt))>0
232     error(strcat('Ground Truth Rotation was',...
233         'not provided or contained NaN.'));
234 end
235
236 %===== Input Validation =====%
237 if (strcmp(mode,'local') && numel(S)==1)
238     % means user did not provide S or in R^1
239     S = repmat(S,n,1);
240 end
241
242 if (strcmp(mode,'local') && numel(S)~=n) || ...
243     (strcmp(mode,'global') && numel(S)~=1)
244
245     error('incompatible size for S');
246 end
247
248 %===== create initial R from theta =====%
249 if sum(isnan(R)) % user didnt provide R
250 Rx = [1 0 0;0 cos(theta) -sin(theta);...
251     0 sin(theta) cos(theta)];
252 Ry = [cos(theta) 0 -sin(theta);0 1 0;...
253     sin(theta) 0 cos(theta)];
254 Rz = [cos(theta) -sin(theta) 0;...
255     sin(theta) cos(theta) 0;0 0 1];
256 R = Rx*Ry*Rz;
257 end
258
259 %===== Initialize ECM iteration =====%
260 degenSign = 0;
261 % if 1 it will mean that a degenerate case occurred
262 iter = 1;
263 % Frobenius of R_new - R_old
264 vr(maxNumIter,1)=0;
265 % IF user wants them: preallocate Sum_of_Squared_Mahalans
266 if nargout > 5
267     ssed(maxNumIter,1) = 0;
268 end
269 % to be sure it will pass the condition
270 vr(1) = Inf;
271
272 cverr = zeros(maxNumIter,1); % preallocate gt errors
273
274 while( iter <= maxNumIter && vr(iter) > endeVal)
275
276 %% ===== POSTERIORS ===== %%
277 U3D = 2.121320343559643 * r.^-3; % 1.5 \sqrt{2} r^-3
278
279 % compute Mahalan-distance of each Yi from every Tx_i

```

```

280
281 TX = R*X+repmat(t,1,n);
282
283     a = mex_sqe(Y,TX); % The L2 differences using C
284
285 % (12) ecmpr [a_ji]
286 if strcmp(mode,'local')
287     a = repmat(S'.^-3,m,1).*...
288         exp(-.5*repmat(S'.^-2,m,1).*a);
289     a = a./repmat((sum(a,2)+U3D),1,n);
290 else % [global isotropic]
291     a = exp(-.5/S^2*a); % exponentiate once for locality
292     a = a./repmat((sum(a,2)+U3D*S^3),1,n);
293 end
294
295 % Virtual points with respect to li [l~rowVec]
296 l = sum(a); % (22) ecmpr
297 W = repmat(1./l,3,1).*(Y*a); % (23) ecmpr
298
299 %===== UMEYAMA ===== %%
300 p = sqrt(n*l(:)/2)./S; % {h^ = B*h}
301 Xcond = X.*repmat(p',3,1); % {X^ = X*B}
302 W = W.*repmat(p',3,1); % {W^ = W*B}
303 gamma = p'*p; % {denum = h^T*B*B*h}
304
305 aa =sum(l);
306
307 %===== Estimate new R,t ===== %
308 % - Sinji umeyama
309
310 % [U,D,V] = svd(WKX')
311 % R = USV' where S = diag([1 1 ... 1])
312 % or diag([ 1 ... -1]) if reflection |UV|<0
313
314 % Avoid the exterior product and
315 % the explicit construction of K & B
316 %-----
317 SIGMA = W*Xcond' - ((W*p)*(Xcond*p)')/gamma;
318 %-----
319 % SIGMA = W*K*Xcond' =>
320 % SIGMA = W(I-Bh'h'^T*B/gamma)Xcond^T =>
321 % SIGMA = WXcond^T - W*pp^T*Xcond^T/gamma =>
322 % SIGMA = WXcond^T - ((W*p)*(Xcond*p)^T)/gamma
323
324 [U,~,V] = svd(SIGMA);
325
326 % check if reflection gives a better match
327 Sreflection = eye(3);
328     if det(U*V') < 0
329         degenSign = 1;

```

```

330         Sreflection(end) = -1;
331     end
332
333     Rnew = U*Sreflection*V';
334
335     t = (W-Rnew*Xcond)*p/gamma;
336
337     vr(iter+1) = norm(Rnew-R, 'fro');
338     R = Rnew; % update R
339
340     cverr(iter) = norm(R-Rgt, 'fro');
341
342     %% ===== REESTIMATE S ===== %%
343
344     % Tx = R_{q+1}X+t_{q+1}
345     TX = R*X+repmat(t,1,n);
346
347     if strcmp(mode, 'local')
348         % local isotropic
349         % S(i) = Sum{a_{ji} ||Yj-TXi||} / 3*l(i) + epsilon
350
351     for i=1:n
352         S(i) = mex_sqe(Y,TX(:,i))*a(:,i);
353     end
354     S = sqrt(S./(3*l(:))+epsilon);
355     else % global isotropic
356         % S = Sum{Sum{a_{ji} ||Yj-TXi||}} / 3aa + epsilon
357         S = sqrt( ( sum(Y.^2)*sum(a,2) ...
358                 -2*(a(:)'*reshape(Y'*TX,m*n,1)) ...
359                 +sum(TX.^2)*sum(a) ) / (3*aa) + epsilon);
360     end
361
362     % iteration error
363     if nargout > 5
364         [~,map] = max(a,[],2); %MAP criterion
365         ssd(iter) = sum(sum((Y-TX(:,map)).^2));
366     end
367     iter = iter+1;
368     end % ECM iteration
369     %t = -R*mx+my+t;
370
371     %% ===== DISCARD UNUSED ELEMENTS ===== %%
372     iter = iter-1; % for the user ...
373
374     % Discard unused preallocated entries
375     if nargout > 5
376         ssd = log10(ssd/sum(ssd(1:iter)));
377     end

```

---

Υλοποίηση του αλγορίθμου ICP, [Ενότητα 4.2.1].

**icp.m**


---

```

1 function [R,t,iter ,vr ,ssed ,cverr] = icp(X,Y,varargin)
2 %ICP Iterative Closest Point
3 % Computes the Rigid Transformation Parameters (R,t) in
4 % order to align and register the 3D point clouds X and Y
5 %
6 % MEX optimized [ DEPENDENCIES: mex_sqe.c ]
7 %
8 % [R,t] = icp(X,Y) returns the optimum transformation pa-
9 % rameters for the registration of point clouds X and Y
10 % in the sense that  $Y = RX+t$ 
11 %
12 % X,Y must be matrices of 3Dimensional points organized
13 % as columns [3 x ..]
14 %
15 % OPTIONAL INPUT arguments provided as string-value pairs
16 %
17 % [R,t] = icp(X,Y,... , 'param1',val1... , 'param2',val2 ,...)
18 %
19 %
20 % Input parameters include :
21 %
22 % 'theta' Initial assumption for the a
23 % ngle between X and Y ,if not
24 % provided the algorithm assu-
25 % mes theta theta=0 as a comm-
26 % on angle for the 3 axes xyz.
27 %
28 %
29 % 'R' Is a [3x3] Rotation Matrix
30 % used as an initial assumpti-
31 % on of the pose between X and
32 % Y.If 'R' is provided 'theta'
33 % value is ignored.
34 %
35 %
36 % 't' a [3x1] vector used as init-
37 % ial estimation of the trans-
38 % lation between the two point
39 % clouds.
40 % Default is the zero vector.
41 %
42 %
43 % 'maxNumIter' Specifies the maximum number
44 % of ICP iterations.
45 % Default = 7.
46 %
47 %

```

```

48 % 'endeVal'
49 %
50 %
51 %
52 %
53 %
54 %
55 %
56 %
57 %
58 %
59 % 'groundTruth'
60 %
61 %
62 %
63 %
64 %
65 %
66 %
67 %
68 % OUTPUT ARGUMENTS :
69 %
70 %
71 % 'R'
72 %
73 %
74 %
75 %
76 %
77 % 't'
78 %
79 %
80 %
81 % 'iter'
82 %
83 %
84 %
85 %
86 %
87 %
88 % 'vr'
89 %
90 %
91 %
92 %
93 %
94 %
95 %
96 %
97 % 'ssed'

```

Scalar which specifies a termination condition. The algorithm stops when the frobenius norm of the difference between the current and former estimation of the Rotation matrix (R) becomes less than the following threshold  $\|R_{(i-1)} - R_i\|_{\text{fro}}$ . Default =  $1e-7$ .

Is an optional argument for the user to provide Ground-Truth rotation matrix between X and Y to obtain convergence differences at each iteration [if `cverr` outargument is requested].

OUTPUT ARGUMENTS :

'R' The optimal rotation Matrix which better aligns X and Y in the sense that the sum of squared distances between the points are minimized.

't' The optimal translation vector between X and Y.

'iter' Is the number of iterations executed until the termination, it will be less or equal to `maxNumIter` depending on if `endeVal` was reached earlier.

'vr' Is the Variation of Rotations, i.e. how much does R change between two consecutive iterations computed as the frobenius norm of the difference between  $R_{q-1}$  and  $R_q$ :  $\|R_q - R_{q-1}\|_{\text{F}}$

'ssed' IF REQUESTED: Is a vector co



```

98 %                               ntaining the value of the mi
99 %                               nimized metric, i.e the sum
100 %                              of squared Euclidean distanc
101 %                              es between each Yj and Xi.
102 %
103 %
104 %   'cverr'                       If requested and 'GroundTru-
105 %                               th' is provided it measures
106 %                               the convergence error betwe-
107 %                               en the current estimation
108 %                               and the ground truth for
109 %                               each iteration as the froben
110 %                               ius norm of the differences.
111 %
112 %
113 %
114 %   $ Last revisited : Wed, 23 October, 8:01 PM $
115
116 [rx,n] = size(X);
117 [ry,m] = size(Y);
118
119 if( rx ~= 3 || ry ~= 3)
120     error('Input arrays must be of size [3 x ..]');
121 end
122
123 %===== parse inputs =====%
124 p = inputParser;
125 addOptional(p, 'theta', 0, @(x) ...
126     isscalar(x) && ~ischar(x));
127 addOptional(p, 't', zeros(3,1), @(x) ...
128     isvector(x) && numel(x)==3);
129 addOptional(p, 'R', NaN(3), @(x) ...
130     size(x,1)==3 && size(x,2)==3);
131 addOptional(p, 'maxNumIter', 40, ...
132     @(x) (round(x)==x) & x>0);
133 addOptional(p, 'endeVal', 1e-7, @(x) ...
134     isreal(x)&&isscalar(x)&&~ischar(x));
135 addOptional(p, 'groundTruth', NaN);
136 parse(p, varargin{:});
137 %=====
138
139
140 %===== rename struct values =====%
141 theta = p.Results.theta;
142 t = p.Results.t(:);
143 R = p.Results.R;
144 maxNumIter = p.Results.maxNumIter;
145 endeVal = p.Results.endeVal;
146 Rgt = p.Results.groundTruth;
147 %=====

```

```

148
149 %—— create initial R from theta
150 if sum(isnan(R)) % user didnt provide R
151 R = [1 0 0;0 cos(theta) -sin(theta);0 sin(theta) ...
152      cos(theta)]*[cos(theta) 0 -sin(theta); ...
153      0 1 0;sin(theta) 0 cos(theta)]*[cos(theta) ...
154      -sin(theta) 0; sin(theta) cos(theta) 0;0 0 1];
155 end
156
157 % preallocation ~2000 times faster than zeros()
158 if nargout > 4
159     ssed(maxNumIter,1) = 0;
160 end
161
162 if m <= n % the smallest to the largest
163     search_direction = 1; % search Y into X
164     cpoint_ind = zeros(m,1); % index of corresponding pts
165     K = eye(m)-ones(m)/m; % umeyama on 3xm crosscovariance
166 else
167     search_direction = 2; % search X into Y
168     cpoint_ind = zeros(n,1);
169     K = eye(n)-ones(n)/n; % on 3xn this time
170 end
171
172 iter = 1;
173 vr = zeros(maxNumIter,1); % variations of rotations
174 vr(1) = Inf; % in order to pass the initial condition
175
176 if nargout > 5 && nnz(isnan(Rgt))>0
177     error(strcat('Ground Truth Rotation was not', ...
178                'provided or contained NaN.'));
179 end
180
181 cverr = zeros(maxNumIter,1);
182
183 while( iter <= maxNumIter && vr(iter) > endeVal)
184
185     % Xtemp is used only in NN Search and so is altered
186     Xtemp = R*X+repmat(t,1,n);
187     % because some values may have grow Indefinitely
188     Ytemp = Y;
189
190 % NN Search For Corresponding points using old parameters
191 if search_direction == 1
192     % search each Y into X [Y \equiv X(:,cpoint_ind)]
193     for tt = 1:m
194         [~,cpoint_ind(tt)] = min(mex_sqe(Xtemp,Y(:,tt)));
195         % surjective correspondences are assumed, i.e a
196         % point from the first set can be assigned to
197         % only one point in the latter set.

```

```

198     Xtemp(:, cpoint_ind(tt)) = [Inf; Inf; Inf];
199     end
200     % The Matrices of Containing the Corresponding points
201     Xtemp = X(:, cpoint_ind);
202     Ytemp = Y;
203 else % direction == 2
204     % search each X into Y [X \equiv Y(:, cpoint_ind)]
205     for tt = 1:n
206         [~, cpoint_ind(tt)] = min(mex_sqe(Xtemp(:, tt), ...
207             Ytemp));
208         % "1-1" correspondences are forced
209         Ytemp(:, cpoint_ind(tt)) = [Inf; Inf; Inf];
210     end
211     % The Equicardinal Corresponding Matrices
212     Xtemp = X;
213     Ytemp = Y(:, cpoint_ind);
214 end
215
216 % L2 Estimation of transformation parameters (R_i, t_i)
217 [U, ~, V] = svd(Ytemp*K*Xtemp');
218 Rnew = U*diag([1, 1, det(U*V')])*V';
219 t = mean(Ytemp, 2) - Rnew*mean(Xtemp, 2);
220
221 % The Variation of Rotations
222 vr(iter+1) = norm(Rnew-R, 'fro');
223 R = Rnew;
224
225 % Current iteration convergence error computation
226 cverr(iter) = norm(R-Rgt, 'fro');
227
228 if nargout > 4
229     % ssd(iter) = sum(sum(mex_sqe(Y, R*X+repmat(t, 1, n))));
230     ssd(iter) = sum(sum((Ytemp-R*Xtemp ...
231         -repmat(t, 1, size(Xtemp, 2)).^2));
232 end
233
234 iter = iter+1;
235
236 end % iteration
237
238 % the first is virtual (i.e., Infinity)
239 vr = vr(2:end);
240
241 % The first value is discarded as virtual (it is used
242 % for the condition of iteration 1.
243 if nargout > 4
244     ssd = log10(ssd/sum(ssd));
245 end

```

---

Συνάρτηση εξομοίωσης θορύβου κάμερας χρόνου πτήσης.

**outlier\_constructor.m**

---

```

1 function [Y,num_of_outliers] = outlier_constructor(Y,...
2     SNR,percent)
3 %OUTLIER_CONSTRUCTOR
4 % This function is used to simulate the imperfections of
5 % a TOF camera on an acquisition of a 3D point cloud.
6 %
7 % Y = outlier_constructor(Y,SNR,percent) takes as input a
8 % [3 x m] matrix Y representing a 3D point cloud and ret-
9 % urns another matrix Y of size [3 x (m+num_of_outliers)]
10 % representing also a point cloud, the points of which
11 % are composed by the points of the input p.cloud disper-
12 % sed by additive white gaussian noise of power SNR and
13 % ceil(n*percent) extra 3D-points playing the role of out-
14 % liers. The outliers are more heavily present at the
15 % boundaries and the dark regions of the scene. Therefore
16 % to simulate them we compute the convex hull of Y and
17 % then we randomly select k=10 (unique) points on it
18 % which are used subsequently as mean values (centroids)
19 % of some multivariate normal distributions. Each of the-
20 % se densities is then used to generate ceil(n*percent/k)
21 % outliers according to a m.v. normal centered at that k
22 % with variance equal with the spectral radius of the
23 % sample covariance matrix of the input point cloud Y.
24 %
25 % [Y,num_of_outliers] = outlier_constructor(...) returns
26 % also the total number of outliers added to the input Y.
27 %
28 %
29 % $ Last revisited : Thu, 24 October, 4:29 PM $
30
31 % completion of input arguments
32 if nargin < 3 % percent not provided
33     percent = 0;
34 end
35 if isempty(SNR) % disparsity noise is abscent
36     SNR = Inf;
37 end
38
39 % gross estimation of boundary points before adding noise
40 K = convhull(Y(1,:),Y(2,:),Y(3,:));
41 K = unique(K(:)); % boundary points unique indices on Y
42
43 % spectral radius of sample covariance is computed before
44 % adding disparsity noise for correlation reasons.
45 pcv = sqrt(max(eig(cov(Y'))));
46

```

```

47 % add disparsity noise on the model points (input Y)
48 Y = awgn(Y,SNR, 'measured');
49
50 % add outliers:
51 % if the input cloud is too small then it want have enou-
52 % gh points to support k unique multivariate normals,
53 % thus k is set equal to the cardinality of the hull
54 if numel(K) < 10
55     k = numel(K);
56 else
57     k = 10; % how many distributions in the mixture
58 end
59 m = size(Y,2);
60
61 p_at_center = ceil(percent/k*m); % points at each m.v.n.
62
63 num_of_outliers = p_at_center*k; % total count of
64     outliers
65
66 % generate outliers and concatenate them to the partial Y
67 for i = 1:k
68     Y = [Y pcv*randn(3,p_at_center)+repmat(Y(:,K(i)),1,...
69         p_at_center)];
69 end

```

---

Υλοποίηση των πειραμάτων της [Ενότητας 7.1.1].

#### **pairwise\_on\_real.m**

---

```

1 %PAIRWISE_ON_REAL demo for comparisson of ICP and ECM-PR
2 % algorithms on the pairwise registration of real data.
3 % The point clouds were acquired with the MESA 4000 TOF-
4 % camera.
5 %
6 %
7 % $ Last revisited : Tue, 19 November, 6:18 PM $
8
9 clear
10 close all
11
12 % IF not compiled
13 % mex mex_sqe.c
14
15 % Load the point clouds, each dataset simply contains
16 % a [3 x ..] matrix of coordinates representing a p.cloud
17 load dataset1 % head           [contains the X matrix]
18 load dataset2 %                 [contains the Y matrix]
19 % load dataset3 % box           [X matrix]
20 % load dataset4 %                 [Y matrix]
21

```

```

22 ecyan = [35 107 144]/255;
23
24 % optimal viewing angle for aligning dataset3 + dataset4
25 % opt_view = [9 17];
26 % optimal viewing angle      -||-      dataset1 + dataset2
27 opt_view = [-7 70];
28
29 %===== VISUALIZE INITIAL STATE =====%
30 close all
31 figure;
32 hold on
33 title('INITIAL POSE','fontsize',27,'fontangle','it',...
34       'fontweight','bold','color',[0 0 0]);
35
36 plot(NaN,'o','markersize',15,'markeredgecolor',...
37       'r','markerfacecolor','r');
38 plot(NaN,'o','markersize',15,'markeredgecolor',...
39       ecyan,'markerfacecolor',ecyan);
40 h_legend = legend('\mathbf{X}$','\mathbf{Y}$');
41 set(h_legend,'fontsize',11,'color','none',...
42     'interpreter','latex','fontsize',21);
43
44 plot3(X(1,:),X(2,:),X(3,:),'.','markeredgecolor',ecyan);
45 plot3(Y(1,:),Y(2,:),Y(3,:),'.','markeredgecolor','r');
46
47 view(opt_view); % optimal viewing angle
48
49 set(gca,'Position',get(gca,'OuterPosition') - ...
50     get(gca,'TightInset') * [-1 0 1 0; 0 -1 0 1; ...
51     0 0 1 0; 0 0 0 1]);
52
53 % 16/9 Aspect Ratio
54 a = get(0,'screensize');
55 set(gcf,'position',floor([a(3) a(4) 3*a(3) ...
56     3*a(4)]/5),'color',[1 1 1]);
57
58 axis('tight')
59 hold off
60 pause(.5);
61
62 % Estimate Isometry
63
64 maxNumIter = 28; % 40 for 3+4
65
66 [R_icp,t_icp,~,~,ssed_icp] = ...
67     icp(X,Y,'maxNumIter',maxNumIter,'endeVal',-1);
68 [R_ecmpr,t_ecmpr,~,~,~,ssed_ecmpr] = ...
69     ecmpr(X,Y,'maxNumIter',maxNumIter,'endeVal',-1);
70
71 %===== PLOT ICP FINAL ALIGNMENT =====%

```

```

72 % Apply final transformation for Visualization
73 TX_icip = R_icip*X+repmat(t_icip,1,length(X));
74
75 figure;
76 hold on
77 title('ICP FINAL ALIGNMENT','fontsize',27,...
78       'fontangle','it','fontweight','bold','color',[0 0 0]);
79
80 plot(NaN,'o','markersize',15,'markeredgecolor',...
81       'r','markerfacecolor','r');
82 plot(NaN,'o','markersize',15,'markeredgecolor',...
83       'cyan','markerfacecolor','cyan');
84 h_legend = legend(strcat('\mathbf{R}_{icip}',...
85       '\mathbf{X}+\mathbf{t}_{icip}','\mathbf{Y}$'));
86 set(h_legend,'fontsize',11,'color','none',...
87       'interpreter','latex','fontsize',21);
88
89 plot3(TX_icip(1,:),TX_icip(2,:),TX_icip(3,:),...
90       '.','markeredgecolor','cyan');
91 plot3(Y(1,:),Y(2,:),Y(3,:),...
92       '.','markeredgecolor',[1 0 0]);
93
94 view(opt_view);
95 set(gca,'Position',get(gca,'OuterPosition') - ...
96       get(gca,'TightInset') * [-1 0 1 0; 0 -1 0 1; ...
97       0 0 1 0; 0 0 0 1]);
98
99 % 16/9 Aspect Ratio
100 a = get(0,'screensize');
101 set(gcf,'position',floor([a(3) a(4) 3*a(3) 3*a(4)]/5),...
102       'color',[1 1 1]);
103
104 axis('tight')
105 hold off
106
107 %===== PLOT ECM-PR FINAL ALIGNMENT =====%
108 TX_ecmpr = R_ecmpr*X+repmat(t_ecmpr,1,length(X));
109
110 figure;
111 hold on
112 title('ECMPR FINAL ALIGNMENT','fontsize',27,...
113       'fontangle','it','fontweight','bold','color',[0 0 0]);
114
115 plot(NaN,'o','markersize',15,'markeredgecolor','r',...
116       'markerfacecolor','r');
117 plot(NaN,'o','markersize',15,'markeredgecolor','cyan',...
118       'markerfacecolor','cyan');
119 h_legend = legend(strcat('\mathbf{R}_{ecmpr}',...
120       '\mathbf{X}+\mathbf{t}_{ecmpr}','\mathbf{Y}$'));
121 set(h_legend,'fontsize',11,'color','none',...

```

```

122     'interpreter', 'latex', 'fontsize', 21);
123
124 plot3(TX_ecmpr(1,:), TX_ecmpr(2,:), TX_ecmpr(3,:), ...
125     '.', 'markeredgecolor', cyan);
126 plot3(Y(1,:), Y(2,:), Y(3,:), '.', ...
127     'markeredgecolor', [1 0 0]);
128
129 view(opt_view);
130 set(gca, 'Position', get(gca, 'OuterPosition') - ...
131     get(gca, 'TightInset') * [-1 0 1 0; 0 -1 0 1; ...
132     0 0 1 0; 0 0 0 1]);
133
134 % 16:9 Aspect Ratio
135 a = get(0, 'screensize');
136 set(gcf, 'position', floor([a(3) a(4) 3*a(3) ...
137     3*a(4)]/5), 'color', [1 1 1]);
138
139 axis('tight')
140 hold off
141 %===== Sum Of Squared Euclidean =====%
142 figure();
143 grid on
144 hold on
145
146 % dummy, cheat the legend entries linewidth
147 plot(NaN, 'r', 'linewidth', 2.7);
148 plot(NaN, cyan, 'linewidth', 2.7, 'color', cyan);
149 h_errors = legend('ICP', 'ECM-PR'); % legend order
150 set(h_errors, 'fontsize', 21);
151
152 plot(1:maxNumIter, ssd_icp, 'color', 'r', 'linewidth', 2.7)
153 plot(1:maxNumIter, ssd_ecmpr, 'color', cyan, ...
154     'linewidth', 2.7);
155 plot(1:maxNumIter, ssd_ecmpr, 'marker', 'o', ...
156     'color', cyan, 'markersize', 10, ...
157     'markerfacecolor', cyan);
158 plot(1:maxNumIter, ssd_icp, 'marker', 'o', ...
159     'color', [1 0 0], 'markersize', 10, ...
160     'markerfacecolor', [1 0 0])
161
162 % title('iteration gain : ||R^{new}-R^{old}||_F', ...
163 %     'fontweight', 'bold', 'fontsize', 15, 'color', cyan);
164 title(strcat('$\sum \limits_{i=1}^n', ...
165     '\parallel \mathbf{y}_i - \mathbf{R} \mathbf{x}_i', ...
166     '- \mathbf{t} \parallel_2^2$'), 'interpreter', ...
167     'latex', 'fontsize', 40)
168
169 xlim([1 maxNumIter])
170 set(gca, 'XTickLabel', num2str((1:maxNumIter)'), ...
171     'XTick', 1:maxNumIter)

```



```

172 set(gca, 'Position', get(gca, 'OuterPosition') - ...
173 get(gca, 'TightInset') * [-1 0 1 0; 0 -1 0 1; ...
174 0 0 1 0; 0 0 0 1]);
175
176 a = get(0, 'screensize');
177 set(gcf, 'position', floor([a(3) a(4) 3*a(3) 3*a(4)]/5), ...
178     'color', [1 1 1]);
179 hold off

```

---

Υλοποίηση των πειραμάτων της [Ενότητας 7.1.2].

#### **pairwise\_on\_artificial.m**

---

```

1  %PAIRWISE_ON_ARTIFICIAL demo for comparisson of the ICP
2  % and ECM-PR algorithms on the pairwise registration of
3  % two partially overlapping unequicardinal 3D-p.clouds
4  % one of which is a corrupted (see outlier_constructor.m)
5  % subsample roto-translated version of the other (model).
6  % As model p.cloud we used the stanford bunny. Stored as
7  % a [3 x ...] matrix in the file "rabbit.mat". The ground
8  % truth rotation on all experiments is pi/6 radians coun-
9  % terclockwise (translation value does not affect the re-
10 % sults so it is ignored at the computation of errors.
11 % Both algorithms executed 70 iterations to be sure that
12 % they reached their stationary point.
13 %
14 %
15 %  $ Last revisited : Wed, 24 October, 6:18 PM $
16
17 clear
18 close all
19
20 % IF not compiled
21 %mex mex_sqe.c
22
23 % construct model p.cloud X and "acquisition" p.cloud Y
24 load rabbit % contains a [3 x ..] matrix I repr. a p.cl
25
26 X = I(:,1:10:end); % Subsample the model point cloud
27
28 % ground truth rotation
29 theta = pi/6;
30 Rx = [1 0 0;0 cos(theta) -sin(theta);...
31     0 sin(theta) cos(theta)];
32 Ry = [cos(theta) 0 -sin(theta);0 1 0;...
33     sin(theta) 0 cos(theta)];
34 Rz = [cos(theta) -sin(theta) 0;...
35     sin(theta) cos(theta) 0;0 0 1];
36 R = Rx*Ry*Rz;
37

```

```

38 % create the "acquisition" (apply Isometry and subsample)
39 Y = R*X+5.4; % add a displacement (translation)
40 Y = Y(:,1:end-50);
41
42 % define iterations (all of them will be run (no endeVal)
43 maxNumIter = 70;
44
45 % some colors for the Visual part
46 ecyan = [35 107 144]/255;
47 secondary_color = [47 109 134]/255;
48
49 % Visual example of the model cloud with some noise disp-
50 % ersity and some outliers added on it
51 % [for visual comprehension of the error data presented
52 % subsequently]
53
54 [Evisual,ext] = outlier_constructor(X,25,.4);
55
56 figure;
57 hold on
58 title('Noise SNR 25%, Outlier percent. 40%',...
59       'fontsize',27,'fontangle','it',...
60       'fontweight','bold','color',[0 0 0]);
61 plot3(Evisual(1,1:end-ext),Evisual(2,1:end-ext),...
62       Evisual(3,1:end-ext),'.','color',ecyan);
63 plot3(Evisual(1,end-ext+1:end),...
64       Evisual(2,end-ext+1:end),Evisual(2,end-ext+1:end),...
65       '.','color',[.5 .5 .5]);
66
67 % Aspect Ratio opt. for 16/9 monitors
68 a = get(0,'screensize');
69 set(gcf,'position',floor([a(3) a(4) 3*a(3)...
70      3*a(4)]/5),'color',[1 1 1]);
71 hold off
72
73 % Define SNR/outlier percentage pairs to measure ground-
74 % truth divergences for the two algorithms. Three packets
75 % nv_combined: presence of both outliers and disparsity n
76 % nv_noise_only: outliers wont be added, disparsity n only
77 % nv_outliers_only: disparsity noise will be absent
78 nv_combined = [Inf 100 70 45 25;
79               0 .1 .2 .3 .4];
80
81 nv_noise_only = [25 30 45 70 100;
82                0 0 0 0 0];
83
84 nv_outliers_only = [Inf Inf Inf
85                   .1 .27 .4];
86
87 % concatenate all experiments in a cell, one figure will

```

```

88 % be created for each element in cell nv_cell.
89 nv_cell = {nv_combined, nv_noise_only, nv_outliers_only};
90
91
92 % Build the required cerror matrices
93
94 % noise types x nv_pairs/type x 2algorithms x iterations
95 cv_err = cell(numel(nv_cell),1); % snr/outliers/ snr+out.
96
97 % -----
98 %   Algorithm 1 is ecmpr / 2 is icp
99 % -----
100
101 % Das Experiments: compute convergence errors for nv_cell
102 for i=1:numel(nv_cell)
103     if (~isempty(nv_cell{i}))
104         fprintf('Measuring set %d\n',i);
105
106         % the block iter x algorithms x snr_values to be stor
107         % ed on each cell element of cv_err.
108         temp_errors = zeros(maxNumIter,2, size(nv_cell{i},2));
109
110         for j=1:size(nv_cell{i},2)
111             % reinstantiate           Y           SNR           outliers
112             Yj = outlier_constructor(Y, nv_cell{i}(1,j), ...
113                 nv_cell{i}(2,j));
114
115             % endeVal -> -1 will force both algorithms to run all
116             % iterations.
117
118             [R_ecmpr, t_ecmpr, ~, ~, ~, ~, temp_errors(:,1,j)] = ...
119                 ecmpr(X, Yj, 'groundTruth', R, ...
120                     'maxNumIter', maxNumIter, 'endeval', -1, 'radius', ...
121                     1e10, 'epsilon', .7);
122
123             [R_icp, t_icp, ~, ~, ~, ~, temp_errors(:,2,j)] = ...
124                 icp(X, Yj, 'groundTruth', R, ...
125                     'maxNumIter', maxNumIter, 'endeVal', -1);
126         end
127
128         % store the block [may not be equicardinal with
129         % the remaining i's]
130         cv_err{i} = temp_errors;
131     end
132 end
133
134 % Visualize convergence errors
135
136 % one plot (figure) for each nv_cell (snr, outliers, ...)
137 for i=1:numel(cv_err)

```

```

138 if (~isempty(nv_cell{i}))
139     figure;
140
141     a = get(0, 'screensize'); % opt. 16/9 AR
142     set(gcf, 'position', floor([a(3) a(4) ...
143         3*a(3) 3*a(4)]/5), 'color', [1 1 1]);
144
145     % The titles, generic title for i > 3 unknow cell el.
146     hold on
147     switch i
148     case 1
149         title('Pairwise on artificial [SNR/OUTLIERS]', ...
150             'fontsize', 27, 'fontangle', 'it', ...
151             'fontweight', 'bold', 'color', [0 0 0]);
152     case 2
153         title('Pairwise on artificial [SNR only]', ...
154             'fontsize', 27, 'fontangle', 'it', ...
155             'fontweight', 'bold', 'color', [0 0 0]);
156     case 3
157         title('Pairwise on artificial [OUTLIER only]', ...
158             'fontsize', 27, 'fontangle', 'it', ...
159             'fontweight', 'bold', 'color', [0 0 0]);
160     otherwise
161         title(strcat('Pairwise on artificial', ...
162             sprintf(' %d [MODIFIED]', i)));
163     end
164
165     % Fetch some colors for the error visualization part.
166     nofpairs = size(cv_err{i}, 3);
167     TTC = jet(10*nofpairs); % expand color differences
168
169     % Construct a readable legend
170     DynamicLegend = cell(nofpairs, 1);
171     for dynlen_entries=1:nofpairs
172         DynamicLegend{dynlen_entries} = ...
173             sprintf('SNR %.2f, Outliers %.2f %%', ...
174                 nv_cell{i}(1, dynlen_entries), ...
175                 nv_cell{i}(2, dynlen_entries)*100);
176     end
177
178     for j=1:nofpairs
179         % Different color for each noise pair, different
180         % linestyle for each algorithm: — ICP / -- ECMPR
181         plot(cv_err{i}(:, 1, j), 'linewidth', ...
182             2.7, 'color', TTC(7*j, :)) % 1 is ECMPR
183     end
184     h_legend = legend(DynamicLegend);
185     set(h_legend, 'fontsize', 11, 'color', 'none');
186     lp = get(h_legend, 'position');
187

```

```

188 % Visualize the linestyles used for each algorithm
189 h = annotation('textbox',[lp(1) lp(2)-.051 ...
190 .19 .035], 'string', strcat('$\mathrm{ECMPR}^{--}', ...
191 '^{\- \- \- \- \- \- \-}^{--} \mathrm{ICP}', ...
192 '^{\- \- \- \- \- \- \-}^{\- \- \- \- \- \- \-}$'), 'interpreter', ...
193 'latex', 'color', secondary_color);
194
195 % The nonlegend plots
196 for j=1:nofpairs
197     plot(cv_err{i}(:,2,j), '—', ...
198         'linewidth', 2.7, 'color', TTC(7*j, :)) % 2 is ICP
199 end
200 hold off
201 end % if no values for an nv_cell elem. noise dont plot
202 end

```

## A'.2 Υλοποιήσεις πολλαπλής αντιστοίχισης

Υλοποίηση του αλγορίθμου ECM-PR πολλαπλών όψεων, [κεφάλαιο 6].

### mv\_ecmpr.m

```

1 function [R,t,X,cv_err] = mv_ecmpr(V,X,varargin)
2 %MVECMMPR Multiple View version of the Expectation Cond-
3 % itional maximization – Point Registration algorithm
4 % Implemented using the method described in section 6.2.1
5 %
6 % [R,t] = mv_ecmpr(V,X) Returns the optimum transformati-
7 % on parameters for the registration of the p.clouds in
8 % cell V in two cells R and t. The registration is achiev
9 % ed by applying the corresponding R{j},t{j} to each V{j}
10 % as R{j} x V{j} +t{j} x ones(1,size(V{j},2)). X contains
11 % the initial centroids of the gaussian mixture model.
12 % Initial variances are all set equal with 1000.
13 %
14 % [R,t,X] = mv_ecmpr(V,X) return also a cell X containing
15 % the final centroids in X{1} and their variances in X{2}
16 %
17 % V must be a cell of M [3 x ...] matrices whose columns
18 % represent 3Dimensional points, each entry of V is treat
19 % ed as an independent view. X must be a [3 x k] matrix.
20 %
21 % OPTIONAL INPUT arguments can be given as string-value p
22 % airs
23 %
24 % [R,t] = mv_icp(V,..., 'param1', val1.., 'param2', val2,..)
25 %
26 %
27 %

```

```

28 % Input parameters include :
29 %
30 % 'Rinit'           Is a cell of [3x3] Rotation
31 %                  matrices used as an initial
32 %                  assumption of the pose betwe
33 %                  en the views V{j} and the
34 %                  common referential. Defaults
35 %                  are the Identity matrices.
36 %
37 %
38 % 't'              Is a cell of [3x1] translati
39 %                  on vectors used as an initi-
40 %                  al assumption of the displac
41 %                  ement of the views V{j}. De-
42 %                  faults are zero vectors.
43 %
44 %
45 % 'maxNumIter'     Specifies the max. number of
46 %                  iterations. Default is 70.
47 %
48 %
49 % 'gt'             Is a cell containing Ground-
50 %                  Truth rotation matrices for
51 %                  each view V{j} assuming that
52 %                  all views are reprojected to
53 %                  V{1}. If 'gt' is provided
54 %                  the optional argument cverr
55 %                  is returned. Cverr contains
56 %                  convergence errors measure-
57 %                  ments for the rotations com-
58 %                  puted at each iteration.
59 %                  Cverr entries are computed
60 %                  as ||Rproj{j} - Rgt{j}||_F^2
61 %                  where Rproj{j} is the pro-
62 %                  jection of V{j} onto V{1}
63 %                  as R{1}'*R{j} and where R is
64 %                  the rotation estimate of the
65 %                  current iteration.
66 %
67 %
68 % 'epsilon'        Is the artificial flatten ad
69 %                  ded to the variances in or-
70 %                  der to avoid zero variances.
71 %                  Default is 1e-5.
72 %
73 %
74 % 'radius'         Is the outlier radius, a wei
75 %                  ghting of the outlier class.
76 %                  The larger the radius, more
77 %                  propable a point to fall in-

```

```

78 %                                     to the outlier class.
79 %                                     Dafult is 10.
80 %
81 %
82 % 'S'                                     Initial covariances for the
83 %                                     centroids of the GMM. Must
84 %                                     be a scalar or k x 1 vector
85 %                                     If it is a scalar all centr-
86 %                                     oids are initialized with
87 %                                     the same variance.
88 %                                     Dafult is 10 for all centers
89 %
90 %
91 % $ Last revisited : Tue, 5 November, 12:57 AM $
92
93 M = numel(V); % number of views i.e. transformations
94 k = size(X,2); % number of cluster centers
95
96 % input parser
97 p = inputParser;
98 addOptional(p, 'maxNumIter', 70, ...
99     @(x) (round(x)==x) & x>0);
100 addOptional(p, 't', NaN, @(x) iscell(x) || isvector(x));
101 addOptional(p, 'Rinit', NaN);
102 addOptional(p, 'epsilon', 1e-5, @isscalar);
103 addOptional(p, 'radius', 1e1, @(x) gt(x,0));
104 addOptional(p, 'gt', {}, @(x) iscell(x));
105 addOptional(p, 'S', 1.5, @isvector);
106 parse(p, varargin{:});
107
108 %rename struct values
109 t = p.Results.t(:)';
110 R = p.Results.Rinit;
111 epsilon = p.Results.epsilon;
112 maxNumIter = p.Results.maxNumIter;
113 r = p.Results.radius;
114 Rgt = p.Results.gt;
115 S = p.Results.S;
116
117 % if user did not provide R, construct the initial R cell
118 if ~iscell(R)
119 R = mat2cell(kron(ones(1,M), eye(3)), 3, 3*ones(M,1));
120 end
121
122 % same for t cell
123 if ~iscell(t)
124 t = mat2cell(zeros(3*M,1), 3*ones(M,1), 1)';
125 end
126
127 % check variances [vectorize if user provided a scalar]

```

```

128 if numel(S) == 1
129     S = repmat(S,k,1);
130 end
131
132 iter = 1;
133 a = cell(M,1);
134 U3D = 2.121320343559643 * r.^-3;    % 1.5 \sqrt{2} r^-3
135
136 % cverr measurement flag, becomes 1 only if gt is provided
137 measure_differences = 0;
138 if nargin > 3
139     if isempty(Rgt)
140         error('gt values not provided');
141     else
142         cverr = zeros(maxNumIter,M);
143         measure_differences = 1;
144     end
145 end
146
147 % zeromean Everything R*(X-imean)+t ==> R*X -R*imean+t
148 imeans = cellfun(@(V) mean(V,2),V,'uniformoutput',false);
149 V = cellfun(@(V,imeans) V-repmat(imeans,1,size(V,2)),...
150     V,imeans,'uniformoutput',false);% the Zero meaned V cell
151
152 while(iter <= maxNumIter)
153     for j=1:M % for all views a.k.a transformations
154
155         m = size(V{j},2);
156
157         % Apply the j-th transform to the centers X
158         TX = R{j}'*(X-repmat(t{j},1,k));
159
160         % posteriors for the current view
161         a{j} = mex_sqe(V{j},TX);
162
163         % logarithmized exponentiation may work better
164         a{j} = repmat(S'.^-3,m,1) .* ...
165             exp(-.5*repmat(S'.^-2,m,1) .* a{j});
166         a{j} = a{j} ./ repmat((sum(a{j},2)+U3D),1,k);
167
168         % Virtual points with respect to li [l~rowVec]
169         l = sum(a{j});
170         W = repmat(1./l,3,1) .* (V{j}.*a{j});
171
172         % UMEYAMA
173         p = sqrt(k*1(:)/2) ./ S;           % {h^ = B*h}
174         Xcond = X.*repmat(p',3,1);        % {X^ = X*B }
175         W = W.*repmat(p',3,1);           % {W^ = W*B }
176         gamma = p'*p;                    % {denum = h^T*B*B*h}
177

```



```

178 % Estimate new R{j} and t{j} using the method of
179 % [section 4.1.4] by inverting the pos. of R and t.
180
181 % [U,D,V] = svd(WKX')
182 % R = USV' where S = diag([1 1 ... 1])
183 % or diag([ 1 ... -1]) if reflection |UV|<0
184 % Avoid the exterior product and
185 % the explicit construction of K & B
186 %-----
187 SIGMA = W*Xcond' - ((W*p)*(Xcond*p)')/gamma;
188 %-----
189 % SIGMA = W*K*Xcond' =>
190 % SIGMA = W(I-Bh'h'^T*B/gamma)Xcond^T =>
191 % SIGMA = WXcond^T - W*pp^T*Xcond^T/gamma =>
192 % SIGMA = WXcond^T - ((W*p)*(Xcond*p)^T)/gamma
193
194 [UU,~,vv] = svd(SIGMA);
195 %diag(Seig)
196
197 % check if reflection gives a better match
198 Sreflection = eye(3);
199 if det(UU*vv') < 0
200     Sreflection(end) = -1;
201 end
202
203 % Umeyama solution of ||V -R_ume*X -t||
204 R_ume = UU*Sreflection*vv';
205 t_ume = (W-R_ume*Xcond)*p/gamma;
206
207 % Symmetrization in ||R{j}V +t{j} - X||
208 R{j} = R_ume';
209 t{j} = -R{j}*t_ume;
210
211 end % end R,t estimation
212
213 % IF REQUESTED
214 if measure_differences
215     % reproject everything onto the first
216     Rproj = cellfun(@(x) R{1}'*x,R, 'uniformoutput',false);
217     cverr(iter,:) = cellfun(@(Rproj,Rgt) ...
218         norm(Rproj-Rgt, 'fro'),Rproj,Rgt);
219 end
220
221 % apply new R,t
222 TV = cellfun(@(x,R,t) R*x+repmat(t,1,size(x,2)),...
223     V,R,t, 'uniformoutput',false);
224
225 % Update cluster centers X_k
226 for v = 1:k
227     tmpx = [0;0;0];

```

```

228         denum = 0;
229     for j=1:M
230         tmpx = tmpx+TV{j}*a{j}(:,v);
231         denum = denum+sum(a{j}(:,v));
232     end
233     X(:,v) = tmpx/denum;
234 end
235
236 % Update covariances (They do not depend on X_k)
237 for v = 1:k
238 % variance numerator and denominator
239     num = 0;
240     denum = 0;
241     for j=1:M
242         num = num+mex_ssqe(TV{j},X(:,v))'*a{j}(:,v);
243         denum = denum+sum(a{j}(:,v));
244     end
245     S(v) = sqrt(num/(3*denum)+epsilon);
246 end
247
248     fprintf('pass %d\n',iter);
249     iter = iter+1;
250
251 end % general iterator
252
253 % Recalculate t for the non zeromeaned V, t' = t-R*v_mean
254 t = cellfun(@(t,R,imeans) t-R*imeans,t,R,imeans,...
255     'uniformoutput',false);
256
257 % Concatenate the final variances to cell X
258 X = {X,S};

```

---

Υλοποίηση του αλγορίθμου ICP πολλαπλών όψεων, [Ενότητα 5.2.1].

**mv\_icp.m**

---

```

1 function [R,t,cverr] = mv_icp(V,varargin)
2 %MVICP Multiple View version of the Iterative Closest
3 % point algorithm. Implemented using typical injective
4 % NN Search and the method for simultaneous estimation of
5 % transformation parameters given by J.Williams and M.Be-
6 % mnamoun in [section 5.1.1] of the current thesis.
7 %
8 % [R,t] = mv_icp(V) Returns the optimum transformation
9 % parameters for the registration of the p.clouds in cell
10 % V in two cells R and t. The registration is achieved by
11 % applying the corresponding R{j},t{j} to each V{j} as
12 % R{j} x V{j} + t{j} x ones(1,size(V{j},2)).
13 %

```

```

14 % V must be a cell of M [3 x ...] matrices whose columns
15 % represent 3Dimensional points, each entry of V is treat
16 % ed as an independent view.
17 %
18 % OPTIONAL INPUT arguments can be given as string-value p
19 % airs
20 %
21 % [R,t] = mv_icp(V,..., 'param1',val1.., 'param2',val2,..)
22 %
23 %
24 %
25 % Input parameters include :
26 %
27 % 'R' Is a cell of [3x3] Rotation
28 % Matrices used as an initial
29 % assumption of the pose betwe
30 % en the views V{j} and the
31 % common referential. Defaults
32 % are the Identity matrices.
33 %
34 %
35 % 't' Is a cell of [3x1] translati
36 % on vectors used as an initi-
37 % al assumption of the displac
38 % ement of the views V{j}. De-
39 % faults are zero vectors.
40 %
41 %
42 % 'maxNumIter' Specifies the max. number of
43 % iterations. Default is 70.
44 %
45 %
46 % 'gt' Is a cell containing Ground-
47 % Truth rotation matrices for
48 % each view V{j} assuming that
49 % all views are reprojected to
50 % V{1}. If 'gt' is provided
51 % the optional argument cverr
52 % is returned. Cverr contains
53 % convergence errors measure-
54 % ments for the rotations com-
55 % puted at each iteration.
56 % Cverr entries are computed
57 % as ||Rproj{j} - Rgt{j}||_F^2
58 % where Rproj{j} is the pro-
59 % jection of V{j} onto V{1}
60 % as R{1}'*R{j} and where R is
61 % the rotation estimate of the
62 % current iteration.
63 %

```

```

64 %
65 % $ Last revisited : Tue, 5 November, 12:18 AM $
66
67 M = numel(V);
68
69 %input parser
70 p = inputParser;
71 addOptional(p, 't', NaN);
72 addOptional(p, 'R', NaN);
73 addOptional(p, 'maxNumIter', 70, ...
74     @(x) (round(x)==x) & x>0);
75 addOptional(p, 'gt', {}, @iscell);
76 parse(p, varargin{:});
77
78 %rename struct values
79 t = p.Results.t(:)'; % The struct columnizes t
80 R = p.Results.R;
81 maxNumIter = p.Results.maxNumIter;
82 Rgt = p.Results.gt;
83
84 % if user did not provide R, construct the initial R cell
85 if ~iscell(R)
86 R = mat2cell(kron(ones(1,M), eye(3)), 3, 3*ones(M,1));
87 end
88
89 % same for t cell
90 if ~iscell(t)
91 t = mat2cell(zeros(3*M,1), 3*ones(M,1), 1)';
92 end
93
94 % cverr measurment flag, becomes 1 only if gt is provided
95 measure_differences = 0;
96 if nargin > 2
97     if isempty(Rgt)
98         error('gt values not provided');
99     else
100         cverr = zeros(maxNumIter, M);
101         measure_differences = 1;
102     end
103 end
104
105 % NOTE: ZeroMeaning of the Input does not alter the res-
106 % ult of MVICP so it was overlooked.
107 for iter=1:maxNumIter
108
109 % Apply current estimation of R,t
110 Vcomparisson = cellfun(@(x,R,t) R*x+repmat(t,1,...
111     size(x,2)), V, R, t, 'uniformoutput', false);
112
113 % NN search

```

```

114
115 % pairs= [ones(M-1,1) (2:M)']; % 1-2,1-3,1-4,...,1-M
116 pairs = combntns(1:M,2); % all possible combinations by 2
117 P = size(pairs,1);
118 Vmod = cell(P,2); % Vleft_pp    Vright_pp
119
120 % For every pair, the view with the smaller cardinality.
121 [cardinals,mc] = min(cellfun(@(x) ...
122     size(x,2),V(pairs)),[],2);
123
124 for pp=1:P
125     % the most/least populated view of this pair
126     small = pairs(pp,mc(pp));
127     big = pairs(pp,3-mc(pp));
128
129     % fetch the Large view and alter it
130     mc3tmp = Vcomparisson{big};
131     mc3ind = zeros(cardinals(pp),1);
132
133     % NN search for each point in most-cardinal (mc) view
134     for i = 1:cardinals(pp)
135         % mc3ind(i) = knnsearch(mc3tmp', ...
136         % Vcomparisson{small}(:,i)'); % old version,same res.
137         [~,mc3ind(i)] = min(mex_sqe(mc3tmp,...
138             Vcomparisson{small}(:,i)));
139         mc3tmp(:,mc3ind(i)) = Inf; % demand injectivity
140     end
141
142     % Modified cell containing sets of corresponding pts
143     % for each pair of views, updated at each iteration.
144     Vmod(pp,[mc(pp) 3-mc(pp) 3]) = ...
145         {V{small} V{big}(:,mc3ind),ones(1,cardinals(pp))};
146 end
147
148 R = cell2mat(R); % recast to cell
149
150 % L2 optimum transformation parameters using J.Williams
151
152 % View selection matrices
153 Ca = zeros(3*M,3*P); % Left views
154 for i=1:P
155     Ca(3*(pairs(i,1)-1)+1:3*pairs(i,1),...
156         3*(i-1)+1:3*i) = eye(3);
157 end
158 Cb = zeros(3*M,3*P); % Right vies
159 for i=1:P
160     Cb(3*(pairs(i,2)-1)+1:3*pairs(i,2),...
161         3*(i-1)+1:3*i) = eye(3);
162 end
163

```

```

164 % The H cells , playing role in the Q_r part of costFun
165 % The weighted covariances XX'
166 Hxx = cellfun(@(x,1) (x.*repmat(1,3,1))*x' ,...
167     Vmod(:,1),Vmod(:,3) , 'uniformoutput' , false);
168 % The weighted covariances WW
169 Hyy = cellfun(@(w,1) (w.*repmat(1,3,1))*w' ,...
170     Vmod(:,2),Vmod(:,3) , 'uniformoutput' , false);
171 % The weighted crosscovariances ( Hxy and Hyx are symm.)
172 Hxy = cellfun(@(x,w,1) (x.*repmat(1,3,1))*w' ,...
173     Vmod(:,1),Vmod(:,2),Vmod(:,3) , 'uniformoutput' , false);
174
175 % Transform them into matrices a.k.a blkdiagonalize
176 Hxx = blkdiag(Hxx{:});
177 Hyy = blkdiag(Hyy{:});
178 Hxy = blkdiag(Hxy{:});
179 Hyx = Hxy';
180
181 % The Gamma Matrix
182 Weights = cellfun(@(1) sum(1)*eye(3),Vmod(:,3) , ...
183     'uniformoutput' , false);
184 % block diagonal of The sum of weights
185 Weights = blkdiag(Weights{:});
186
187 Cdiff = Ca-Cb;
188
189 %PINV(abs(PINV)<eps(3*P*norm(PINV))) = 0;
190 PINV = pinv(Cdiff*Weights*Cdiff');
191
192 %Gamma(abs(Gamma)<eps(3*P*norm(Gamma))) = 0;
193 Gamma = Weights*Cdiff'*PINV*Cdiff*Weights;
194
195 % Fill 3x3 diagonal blocks in order to apply hadamard.
196 Gamma = cell2mat(cellfun(@(x) x+circshift(x,[0 1])+...
197     circshift(x,[0 2]),mat2cell(Gamma,3*ones(P,1) ,...
198     3*ones(P,1)) , 'uniformoutput' , false));
199
200 % The mean vectors of the "Left" and "Right" views.
201 meanVl = cell2mat(cellfun(@(x,1) (x*1')/sum(1) ,...
202     Vmod(:,1),Vmod(:,3) , 'uniformoutput' , false));
203 meanVr = cell2mat(cellfun(@(w,1) (w*1')/sum(1) ,...
204     Vmod(:,2),Vmod(:,3) , 'uniformoutput' , false));
205
206 % G matrices , weighted exteriors of the L,R mean vectors.
207 Gxx = Gamma.*(meanVl*meanVl');
208 Gxy = Gamma.*(meanVl*meanVr');
209 Gyx = Gamma.*(meanVr*meanVl');
210 Gyy = Gamma.*(meanVr*meanVr');
211
212 % The Qr matrix (rotation dependent)
213 Qr = Ca*Hxx*Ca' + Cb*Hyy*Cb' - Ca*Hxy*Cb' - Cb*Hyx*Ca';

```

```

214
215 % The Qrt (Rototranslation dependent)
216 Qrt = -Ca*Gxx*Ca' + Ca*Gxy*Cb' + Cb*Gyx*Ca' -Cb*Gyy*Cb';
217
218 % The Q Matrix (Complete cost function, as maximum trace)
219 Q = Qr+Qrt;
220
221 % Solve iteratively for R = { R{1}, R{2}, ... , R{M} }.
222 for i=1:1e2 % converge?
223     for j=1:M % solve for each Rj
224
225         % The minus objective
226         F = -(Q(3*(j-1)+1:3*j,1:3*(j-1))*R(:,1:3*(j-1))'+...
227             Q(3*(j-1)+1:3*j,3*j+1:end)*R(:,3*j+1:end)');
228
229         % SVD of the partial objective
230         [Uwilliams,~,Vwilliams] = svd(F);
231
232         % The R{i} block of R
233         R(:,3*(j-1)+1:3*j) = Vwilliams*...
234             diag([1 1 det(Uwilliams*Vwilliams)])*Uwilliams';
235     end
236 end
237
238 % The Rotated Left and Right means
239 Rl = mat2cell(R*Ca,3,3*ones(P,1));
240 Rr = mat2cell(R*Cb,3,3*ones(P,1));
241
242 % The R-dependent part of T
243 Z = blkdiag(Rl{:})*meanVl-blkdiag(Rr{:})*meanVr;
244
245 % Solve for T once using the final Rotations, T =
246 % -(PINV x Cdiff x N x (R x Ca x meanVL-R x Cb x meanVR))
247 T = - PINV*Cdiff*Weights*Z;
248
249 % Reshape t and R to cell
250 t = mat2cell(T,3*ones(M,1),1)';
251 R = mat2cell(R,3,3*ones(M,1)); % applied in the beginning
252
253 % IF REQUESTED
254 if measure_differences
255     % reproject everything onto the first, V{1}
256     Rproj = cellfun(@(x) R{1}'*x,R, 'uniformOutput',false);
257     cverr(iter,:) = cellfun(@(Rproj,Rgt) ...
258         norm(Rproj-Rgt, 'fro'),Rproj,Rgt);
259 end
260     disp(['pass ' num2str(iter)])
261 end

```

---

Υλοποίηση των πειραμάτων της [Ενότητας 7.2.1].

### multiview\_visual.m

---

```

1  %MULTIVIEW_VISUAL Visual Comparisson of the Results of
2  % MVECMR and MVICP algorithms on the registration of
3  % M unequicardinal point clouds. The user need only to
4  % specify the latent angles for each View in vector theta
5  % The Constraint of unequal cardinalities is forced using
6  % random subsampling of the model p.cloud. Selection Cri-
7  % terion for k is the number of unique points in the conv
8  % ex hull of the initial pose (concat.) of all M p.clouds.
9  %
10 % $ Last revisited : Sat, 2 November, 3:07 PM $
11
12 close all
13
14 load rabbit % the model-cloud, a [3 x ...] matrix
15 I = I(:,1:7:end); % discard some points
16
17 maxNumIter = 87; % Execution Iteration for both algs
18 theta = [0 pi+.2 pi/2.1 pi/6 -pi/3]; % latent angles
19
20
21 %===== Inline Function For Euclidean Transformation =====%
22 euclidean = @(theta,A,shift) ...
23     [1 0 0;0 cos(theta) -sin(theta);0 sin(theta) ...
24     cos(theta)]*[cos(theta) 0 -sin(theta);0 1 0; ...
25     sin(theta) 0 cos(theta)] *[cos(theta) -sin(theta) ...
26     0;sin(theta) cos(theta) 0; 0 0 1] ...
27     *A+repmat(shift,1,size(A,2));
28 %=====
29
30
31 %===== Construct the Views =====%
32 M = numel(theta); % A view for each rotation, as many as
33 % the angles specified.
34
35 V = cell(1,M);
36
37 for i=1:M
38     % Force Unequal cardinalities ,destroy correspondences
39     endpoint = ceil((rand(1)/2+.45)*size(I,2));
40
41     % Apply Rototranslation and Subsampling on each View
42     V{i} = euclidean(theta(i),I(:,1:endpoint) ,...
43         .3*rand(3,1)); % ~ random initial translation
44 end
45 %=====
46

```



```

47
48 %===== THE DYNAMIC LEGEND =====%
49 DynamicLegend = cell(M,1); % visualization using NaN's
50 for i=1:M
51     DynamicLegend{i} = sprintf(strcat('V%d : [3 x ',...
52         '%.4d], \theta = %.2f \circ'),i,...
53         size(V{i},2),theta(i)*180/pi); % radian to degrees
54 end
55 %===== %
56
57
58 %===== ColorMap & ScreenLimits =====%
59 TTC = [1 0 0;0 0 1;[35 104 135]/255;...
60     [0 .9 0];[.5 .4 .1];[.7 0 0];jet(M-5)]; % colormap
61 a = get(0,'screensize'); % will be used in positioning
62 %===== %
63
64
65 %===== INITIAL ALIGNMENT VISUALIZATION =====%
66 hold on
67 title(...
68     sprintf('Initial Position of The %d Point Clouds',M),...
69     'fontsize',27,'fontangle','it','fontweight','bold',...
70     'color',[0 0 0]);
71
72 % Use NaN's for the projection of the legend [They do not
73 % appear in the plot but they count as legend entriers]
74 for i=1:M
75     plot(NaN,'o','markersize',11,'markeredgecolor',...
76         TTC(i,:), 'markerfacecolor',TTC(i,:));
77 end
78
79 % Place the Legend Here
80 h_legend = legend(DynamicLegend);
81 set(h_legend,'fontsize',11,'color','none');
82
83 % Plot the Real Data with the convenient markersize
84 for i=1:M
85     plot3(V{i}(1,:),V{i}(2,:),V{i}(3:,:),'.',...
86         'markeredgecolor',TTC(i,:));
87 end
88
89 % screen positioning
90 set(gca,'Position',get(gca,'OuterPosition') - ...
91     get(gca,'TightInset') * [-1 0 1 0; 0 -1 0 1; ...
92     0 0 1 0; 0 0 0 1]);
93
94 % 16/9 Aspect Ratio
95 set(gcf,'position',floor([a(3) a(4) 3*a(3) 3*a(4)]/5),...
96     'color',[1 1 1]);

```

```

97
98 axis('tight') % Centralize the View
99 axis('off') % but discard the line/numbers
100 hold off
101 %=====
102
103
104 %===== REGISTRATION USING THE MULTVIEW ICP =====
105 [R_icp, t_icp] = mv_icp(V, 'maxnumiter', maxNumIter);
106
107 TV_icp = cellfun(@(x,R,t) R*x+repmat(t,1, size(x,2)), ...
108     V, R_icp, t_icp, 'uniformoutput', false); % apply the R,t
109 %=====
110
111
112 %===== FINAL ALIGNMENT USING MV_ICP =====
113 figure;
114 hold on
115 title('Final Alignment Using MV\ICP', 'fontsize', 27, ...
116     'fontangle', 'it', 'fontweight', 'bold', 'color', [0 0 0]);
117
118 % Use NaN's for the projection of the legend
119 for i=1:M
120     plot(NaN, 'o', 'markersize', 11, 'markeredgecolor', ...
121         TIC(i,:), 'markerfacecolor', TIC(i,:));
122 end
123
124 % Place the Legend Here
125 h_legend = legend(DynamicLegend);
126 set(h_legend, 'fontsize', 11, 'color', 'none');
127
128 % The Real Data
129 for i=1:M
130     plot3(TV_icp{i}(1,:), TV_icp{i}(2,:), TV_icp{i}(3,:), ...
131         '.', 'markeredgecolor', TIC(i,:));
132 end
133
134 % screen corrections
135 set(gca, 'Position', get(gca, 'OuterPosition') - ...
136     get(gca, 'TightInset') * [-1 0 1 0; 0 -1 0 1; ...
137     0 0 1 0; 0 0 0 1]);
138
139 % 16/9 Aspect Ratio
140 set(gcf, 'position', floor([a(3) a(4) 3*a(3) 3*a(4)]/5), ...
141     'color', [1 1 1]);
142
143 axis('tight') % Centralize the View
144 axis('off') % but discard the line/numbers
145 hold off
146 %=====

```

```

147
148
149 %===== REGISTRATION USING THE MULTIVIEW ECMPR =====%
150 % A criterion for the number of centers
151 k = numel(unique(convhull(cell2mat(V)')));
152 X = randn(3,k); % The initial centers
153
154 [R_ecmpr, t_ecmpr]= mv_ecmpr(V,X, 'maxNumIter',maxNumIter);
155
156 TV_ecmpr = cellfun(@(x,R,t) R*x+repmat(t,1,size(x,2)),...
157     V,R_ecmpr,t_ecmpr, 'uniformoutput',false);
158 %=====
159
160 %===== FINAL ALIGNMENT USING MV_ECMPR =====%
161 figure;
162 hold on
163 title('Final Alignment Using MV\ECMPR','fontsize',27,...
164     'fontangle','it','fontweight','bold','color',[0 0 0]);
165
166
167 % Use NaN's for the projection of the legend
168 for i=1:M
169     plot(NaN,'o','markersize',11,'markeredgecolor',...
170         TTC(i,:), 'markerfacecolor',TTC(i,:));
171 end
172
173 % Place the Legend Here
174 h_legend = legend(DynamicLegend);
175 set(h_legend, 'fontsize',11, 'color', 'none');
176
177 for i=1:M
178     plot3(TV_ecmpr{i}(1,:),TV_ecmpr{i}(2,:),...
179         TV_ecmpr{i}(3,:), '.', 'markeredgecolor',TTC(i,:));
180 end
181
182 % screen corrections
183 set(gca, 'Position', get(gca, 'OuterPosition') - ...
184     get(gca, 'TightInset') * [-1 0 1 0; 0 -1 0 1; ...
185     0 0 1 0; 0 0 0 1]);
186
187 % 16/9 Aspect Ratio
188 set(gcf, 'position', floor([a(3) a(4) 3*a(3) 3*a(4)]/5), ...
189     'color',[1 1 1]);
190
191 axis('tight') % Centralize the View
192 axis('off') % but discard the line/numbers
193 hold off

```

---

Υλοποίηση των πειραμάτων της [Ενότητας 7.2.2].

**multiview\_metricization.m**


---

```

1 %MULTIVIEW_METRICIZATION Quantitative comparisson of the
2 % MVECMPR and MVICP algorithms. The user has to provide
3 % a vector of rotations "theta" and a [2 x ...] matrix of
4 % noise-outlier pairs as columns, named "nvc". This demo
5 % constructs a set of views, one for each entry of theta
6 % and then measures convergence errors of the registrati-
7 % on of these views for both algorithms. The process is
8 % repeated for each column (noise-outlier pair) on nvc.
9 % For the view noising algorithm See OUTLIER_CONSTRUCTOR.
10 %
11 %
12 % $ Last revisited : Mon, 4 November, 10:56 PM $
13
14 close all
15
16 % latent angles, one for each view, random translation
17 theta = [0 pi/3 2*pi/3 pi 4*pi/3 5*pi/3];
18
19 % SNR values (noise) and Outliers percent. combined [nvc]
20 nvc = [Inf 100 50 25;0 .1 .25 .4];
21
22 % Experiment durations, number of iterations (no endeVal)
23 maxNumIter = 250;
24
25 load rabbit % the model-cloud, a [3 x ...] matrix
26 I = I(:,1:10:end); % discard some points
27
28 M = numel(theta); % num of views
29
30 %Construct Cell of GroundTruth Rotations
31 Rgt = arrayfun(@(theta) [1 0 0;0 cos(theta) ...
32     -sin(theta);.0 sin(theta) cos(theta)]*[cos(theta) ...
33     0 -sin(theta);0 1 0;sin(theta) 0 cos(theta)] ...
34     *[cos(theta) -sin(theta) 0;sin(theta) ...
35     cos(theta) 0; 0 0 1],theta, 'uniformoutput', false);
36
37 % The GT is computed by assuming everything is backproj-
38 % ected to the referential of the first cloud, because
39 % the referentials estimated by MVICP & MVECMPR may
40 % differ. [Reprojection to V{1}]
41 Rgt = cellfun(@ctranspose ,Rgt, 'uniformoutput', false);
42
43 % Apply Isometry transformation to a point Cloud - View
44 % using an cellfun and an inline function.
45 euclidean = @(theta ,A, shift) [1 0 0;0 cos(theta) ...
46     -sin(theta);.0 sin(theta) cos(theta)]*[cos(theta) ...
47     0 -sin(theta);0 1 0; sin(theta) 0 cos(theta)] ...

```

```

48     * $[\cos(\theta) \ -\sin(\theta) \ 0; \sin(\theta) \ \cos(\theta) \ \dots$ 
49     0; 0 0 1] *A+repmat(shift,1,size(A,2));
50
51 % Number of views
52 V = cell(1,M);
53
54 % Construct the Views
55 for i=1:M
56     % Force Unequal cardinalities,destroy correspondences
57     endpoint = ceil((rand(1)/2+.45)*size(I,2));
58
59     % Apply Rototranslation and Subsampling on each View
60     V{i} = euclidean(theta(i),I(:,1:endpoint),...
61         .3*rand(3,1)); % ~ random initial translation
62 end
63
64 % Set a ColorMap and get system screensizes
65 a = get(0,'screensize'); % will be used in positioning
66
67 TTC = [ [255 0 0;0 100 0;35 107 144;238 154 0; ...
68     255 69 0]/255 ; jet(M-5)]; % missing colormap entries
69 % are completed using the built-in jet() colormap.
70
71 % initial centroids for MVECMR
72 % k = numel(unique(convhull(cell2mat(V)'))); % optional
73 k = 280;
74 X = randn(3,k); % Their topology doesn't matter
75
76 for i=1:size(nvc,2)
77     fprintf('\nSNR: %.2f, Outliers: %.2f%%\n\n',...
78         nvc(1,i),nvc(2,i)*100);
79
80 % construct a legend with cardinalities and latent angles
81 DynamicLegend = cell(M,1); % visualization using NaN's
82 for j=1:M
83     DynamicLegend{j} = sprintf(strcat('%d : [3 x ',...
84         '%.4d], \theta = %.2f \circ'),j,...
85         size(V{j},2),theta(j)*180/pi); % radian to degrees
86 end
87
88 % summon the outlier constructor for all p.clouds
89 Vmod = cellfun(@(V) outlier_constructor(V,nvc(1,i),...
90     nvc(2,i)),V,'uniformoutput',false); % SNR / outliers
91
92 % execute mv_icp
93 fprintf('----- MultiView ICP\n');
94 [~,~,cverr_icp] = mv_icp(Vmod,...
95     'maxNumIter',maxNumIter,'gt',Rgt);
96
97 % execute mv_ecmpr

```

```

98 fprintf( '----- MultiView ECMPR\n' );
99 [~,~,~,cverr_ecmpr] = mv_ecmpr(Vmod,X,...
100     'maxNumIter',maxNumIter, 'gt',Rgt, 'radius',1e10, 'S'
    ,10);
101
102 % visualize differences (convergence errors)
103 figure;
104 hold on
105
106 % MVECMPR errors (continuous line)
107 for z = 1:M
108 plot( cverr_ecmpr(:,z), 'linewidth',2.7, 'color',TTC(z,:));
109 end
110
111 % set the Dynamic Legend and extract its position for h
112 h_legend = legend(DynamicLegend);
113 set(h_legend, 'fontsize',11, 'color', 'none');
114 lp = get(h_legend, 'position');
115
116 % Textbox for the linestyles used for each algorithm
117 h = annotation( 'textbox', [lp(1)+.185 lp(2)+.04 .242 ...
118     .035], 'string', strcat( '$\mathrm{MV};\text{ECMPR}$', ...
119     '$\{-\{-\{-\{-\{-\{-\{-\}\}\}\}\}\}\}\mathrm{MV};\text{ICP}$', ...
120     '$\{-\{-\{-\{-\{-\{-\{-\}\}\}\}\}\}\}$'), 'interpreter', ...
121     'latex', 'color', [35 104 144]/255);
122
123 % Textbox for current figure's SNR and outlier values
124 l = annotation( 'textbox', [lp(1)+.185 lp(2)-.01 .22 ...
125     .035], 'string', sprintf( strcat( 'SNR = %.2f', ...
126     'Outliers = %.1f%%'), nvc(1,i), nvc(2,i)*100), ...
127     'color', [35 104 144]/255, 'fontweight', 'bold');
128
129 % MVICP errors (dashed line)
130 for view = 1:M
131 plot( cverr_icp(:,view), 'linestyle', '--', ...
132     'linewidth',2.7, 'color',TTC(view,:));
133 end
134
135 % screen corrections
136 set(gca, 'Position', get(gca, 'OuterPosition') - ...
137     get(gca, 'TightInset') * [-1 0 1 0; 0 -1 0 1; ...
138     0 0 1 0; 0 0 0 1]);
139
140 set(gcf, 'position', floor( [a(3) a(4) ...
141     3*a(3) 3*a(4)]/5), 'color', [1 1 1]);
142
143 hold off
144 end

```

---

# Βιβλιογραφία

- [1] Radu Horaud, Florence Forbes, Manuel Yguel, Guillaume Dewaele, and Jian Zhang Rigid and Articulated Point Registration with Expectation Conditional Maximization, *IEEE PAMI 2011*
- [2] Richard Hartley and Andrew Zisserman, Multiple View Geometry in computer vision *Cambridge University Press 2003*
- [3] Gary K.L. Tam, Zhi-Quan Cheng, Yu-Kun Lai, Frank C. Langbein, Yonghuai Liu, David Marshall, Ralph R. Martin, Xian-Fang Sun and Paul L. Rosin, Registration of 3D point clouds and meshes a survey from rigid to non-rigid *IEEE trans. on Visualization and Computer Graphics July 2013*
- [4] Berthold K.P. Horn, Closed-form solution of absolute orientation using unit quaternions, *Journal of the optical society of America, November 1986*
- [5] Berthold K.P. Horn, Hugh M. Hilden, Shahriar Negahdaripour, Closed-form Solution Of Absolute Orientation Using Orthonormal Matrices *Honolulu, Hawaii 1988*
- [6] Shinji Umeyama, Least Squares Estimation of transformation Parameters Between Two Point Patterns *IEEE PAMI 1991*
- [7] K. S. Arun, T. S. Huang and S. D. Blostein, Least-Squares fitting of two 3-D Point Sets *PAMI 1987*
- [8] Felix Klein, Elementary mathematics from an Advanced Standpoint. Macmillan, NewYork, 1939
- [9] P.J.Besl, N.D.McKay, A method for registration of 3-D Shapes, *PAMI 1992*
- [10] Zhang, Zhengyou, Iterative point matching for registration of free-form curves and surfaces, *International Journal of Computer Vision 1994*
- [11] Shi Q., Xi, N., Automated data processing for a rapid 3D Surface Inspection System, Robotics and Automation (CIRA) *IEEE international conference on, May 2008*
- [12] Ridene T., Goullete F., Registration of fixed and mobile terrestrial Laser data sets with DSM, *Computational Intelligence in Robotics Automation (CIRA), IEEE International Symposium on, 2009*
- [13] M. Fischler and R. Bolles. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *1981.*

- [14] Johnson A., K.H. Kang, S.B., Registration and integration of textured 3-d data, Recent Advances in 3-D Digital Imaging and Modeling, *Proceedings of the International Conference on, May 1997*
- [15] Joung, J.H. An, K.H. Kang, J.W. Chung, M.J. Yu, W., 3D environment reconstruction using modified color ICP algorithm by fusion of a camera and a 3D laser range finder, Intelligent Robots and Systems (IROS), *IEEE/RSSJ International Conference on, October 2009*
- [16] K. Bae and D. Lichti. A method for automated registration of unorganised point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing, 2008*
- [17] Lorenz Oswald, Recent Development of the Iterative Closest Point Algorithm An overview of the years 2008 to 2010, *ETH Autumn Term 2010*
- [18] Feng Lu and Evangelos Milios. Robot pose estimation in unknown environments by matching 2D range scans. *Journal of Intelligent and Robotic Systems, March 1997*
- [19] Peter Biber and Wolfgang Straßer. The normal distributions transform: A new approach to laser scan matching. In *Proceedings of the IEEE International Conference on Intelligent Robots and Systems (IROS), 2003.*
- [20] Martin Magnusson, 3D Scan Matching for Mobile Robots with Application to Mine Mapping, *PHD Thesis, University of Orebro 2006*
- [21] A.P.Dempster, N.M.Laird, D.B.Rubin, Maximum likelihood from incomplete data via the EM Algorithm *Journal of the royal statistical society. series B(methodological), Vol 39, No.1, (1977), pp.1-38*
- [22] Xia-Li Meng, Donald B.Rubin, Maximum likelihood estimation via the ECM A general framework, *Biometrika vol 80 ,No. 2,(June 1993)*
- [23] Radford M.Neal, Geoffrey E.Hinton, A view of the EM algorithm that justifies incremental, sparse, and other variants, *Dept. of statistics and computer science, University of Toronto, Toronto, Ontario, Canada 1999*
- [24] Rauf Benjeemaa, Francois Schmitt, A Solution For The Registration Of Multiple 3D point Sets Using Unit Quaternions, *Ecole Nationale Supérieure des Telecommunications, Paris, France, 1997*
- [25] Shankar Krishnan, Pei Yean Lee, John B.Moore, Suresh Venkatasubramanian, Global Registration of Multiple 3D Point Sets via Optimization-on-a-Manifold, *Eurographics Symposium on Geometry Processing (2005)*
- [26] John Williams, Mohammed Bennamoun, Simultaneous Registration of Multiple Corresponding Point Sets, *Computer Vision and Image Understanding (CVIS), 2001*
- [27] Andry Myronenko, Xubo Song, Point Set Registration: Coherent Point Drift, *IEEE PAMI Dec. 2010*
- [28] Christopher M.Bishop, Pattern Recognition and Machine Learning, *Springer-Verlag 2006*