



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΤΜΗΜΑ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ
ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΣΗΜΑΤΩΝ, ΕΛΕΓΧΟΥ ΚΑΙ ΡΟΜΠΟΤΙΚΗΣ

**Εφαρμογές τρισδιάστατης ανακατασκευής με γεωμετρία
πολλών όψεων**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΤΟΥ

**Νίκου Ε.
Μελανίτη-Παρασκευά**

**Επιβλέπων: Πέτρος Μαραγκός
Καθηγητής Ε.Μ.Π.**

Αθήνα, Ιούλιος 2014



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

**Εφαρμογές τρισδιάστατης ανακατασκευής με γεωμετρία
πολλών όψεων**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΤΟΥ

Νίκου Ε. Μελανίτη-Παρασκευά

Επιβλέπων: Πέτρος Μαραγκός
Καθηγητής Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 14η Ιουλίου 2014.

.....
Πέτρος Μαραγκός
Καθηγητής Ε.Μ.Π.

.....
Ηλίας Γλύτσης
Καθηγητής Ε.Μ.Π.

.....
Γεώργιος Παπαβασιλόπουλος
Καθηγητής Ε.Μ.Π.

Αθήνα, Ιούλιος 2014

.....

Νίκος Ε. Μελανίτης-Παρασκευάς

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

© Νίκος Μελανίτης-Παρασκευάς, 2014.

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

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

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

ΠΕΡΙΛΗΨΗ

Η παρούσα διπλωματική εργασία έχει ως αντικείμενο την τρισδιάστατη ανακατασκευή της γεωμετρίας της σκηνής, από δισδιάστατες εικόνες της, με χρήση τεχνικών όρασης υπολογιστών και γεωμετρίας πολλαπλών όψεων. Έμφαση δίνεται σε εξωτερικές αρχιτεκτονικές σκηνές και επιχειρείται η εφαρμογή των μεθόδων σε χώρους πολιτιστικού ενδιαφέροντος. Αρχικά, αναπτύσσεται μέθοδος ανάκτησης των εσωτερικών (χαρακτηριστικά κάμερας) και εξωτερικών (σχετική θέση φωτογραφιών) παραμέτρων των φωτογραφιών. Η μέθοδος χρησιμοποιεί ζεύγη φωτογραφιών και δεν προβαίνει στην υπόθεση ταυτόσημων εσωτερικών παραμέτρων που συνηθίζεται στην βιβλιογραφία. Στη συνέχεια, η μέθοδος ενσωματώνεται σε κατάλληλα σχεδιασμένα συστήματα τρισδιάστατης ανακατασκευής. Τέλος, η προηγούμενη μέθοδος βελτιώνεται με την αξιοποίηση γεωμετρικών ιδιοτήτων του προβλήματος, στο αρχικό ταίριασμα περιγραφών σημείων (χαρακτηριστικών) των εικόνων και στην εκτίμηση των εξωτερικών και εσωτερικών παραμέτρων των φωτογραφιών. Η αξιοποίηση της γεωμετρικής δομής του προβλήματος οδηγεί σε βελτιωμένα αποτελέσματα και πιο αξιόπιστες μεθόδους. Πέρα από τις συνεισφορές της εργασίας που παρουσιάζονται στο δεύτερο μέρος, τα στάδια και οι καθιερωμένοι αλγόριθμοι τρισδιάστατης ανακατασκευής περιγράφονται με λεπτομέρεια στο πρώτο μέρος.

ΛΕΞΕΙΣ ΚΛΕΙΔΙΑ : Όραση υπολογιστών, γεωμετρία πολλαπλών όψεων, τρισδιάστατη ανακατασκευή, μη-επαναληπτικό SfM, L_∞ , βαθμονόμηση με μεταβαλλόμενες εσωτερικές παραμέτρους, γεωμετρική επαλήθευση, στάθμιση περιστροφών

ABSTRACT

The aim of this thesis is to reconstruct scene's 3D geometry, using its 2D images, employing computer vision and multiple view geometry techniques. We focus on outdoor architectural scenes and attempt to reconstruct cultural heritage sites. Initially, we develop a method to recover intrinsic (describing the cameras) and extrinsic (describing relative positions) parameters for each view. The method uses photograph pairs and allows, in contrast with previous approaches, varying internal parameters. Next, we incorporate the previous method to an accordingly designed 3D reconstruction pipeline. Finally, we utilize geometric properties inherent to the reconstruction problem to improve the initial matching between point descriptions (features) as well as the recovery of internal and external camera parameters. That results in improvements in performance and reliability. In addition to our contributions, presented in the second part, we provide a thorough presentation of the various stages and the most influential techniques that constitute a 3D reconstruction pipeline.

KEYWORDS: Computer vision, multiple view geometry, 3D reconstruction, non-incremental SfM, L_∞ , calibration with varying internal parameters, geometric verification, rotation averaging

Με τα λίγα μπορούν να ευτυχούν μόνο όσοι γέρασαν,- σωματικά ή ψυχικά,- επειδή αυτοί, μη έχοντας να περιμένουν τίποτα, είναι ευχαριστημένοι με τα ψίχουλα του δρόμου. Γι'αυτό και οι πικραμένοι άνθρωποι είναι στο βάθος οι πιο αισιόδοξοι: επειδή σ'αυτούς έμεινε ακόμα το προνόμιο ν'αγαπούν, να πιστεύουν στον άνθρωπο και στη ζωή, το προνόμιο να κυνηγούνε χίμαιρες.

Ηλίας Βενέζης, Γαλήνη, Πρόλογος στην Τρίτη Έκδοση

ΠΕΡΙΕΧΟΜΕΝΑ

| | |
|---|-----------|
| Εισαγωγή | 10 |
| Αντικείμενο Διπλωματικής Εργασίας | 10 |
| Διάρθρωση Διπλωματικής Εργασίας | 11 |
| I Βιβλιογραφική Ανασκόπηση | 14 |
| 1. Εισαγωγικές έννοιες | 15 |
| 1.1. Δισδιάστατη γεωμετρία | 15 |
| 1.1.1. Μετασχηματισμοί | 16 |
| 1.2. Τρισδιάστατη γεωμετρία | 17 |
| 1.3. Προβολές από τον τρισδιάστατο στον δισδιάστατο κόσμο. Μοντε- λοποίηση κάμερας. | 18 |
| 1.3.1. Κεντρική Προβολή | 18 |
| 1.3.2. Γενικεύσεις και επεκτάσεις στο μοντέλο της μικροπικής κάμερας | 19 |
| 1.3.3. Εξαγωγή γεωμετρικών χαρακτηριστικών από την Μήτρα της κάμερας | 22 |
| 1.3.4. Μη γραμμικά φαινόμενα. Η ακτινική παραμόρφωση . . | 23 |
| 1.3.5. Ανακεφαλαίωση | 24 |
| 2. Ανάκτηση Παραμέτρων Κάμερας & Αραιή Τρισδιάστατη Ανακατα- σκευή. Συστήματα SfM | 27 |
| 2.1. Μέθοδοι ανακατασκευής τρισδιάστατης γεωμετρίας και προσ- διορισμού παραμέτρων κάμερας | 27 |
| 2.1.1. Βασικές έννοιες. Σφάλμα επαναπροβολής | 28 |
| 2.1.2. Παραγοντοποίηση με μοντέλο προοπτικής προβολής . | 28 |
| 2.1.3. Επιπολική Γεωμετρία. Ο θεμελιώδης πίνακας F. | 30 |
| 2.1.4. Το πρόβλημα του Τριγωνισμού | 32 |
| 2.1.5. DLT και RANSAC:Εφαρμογή στον προσδιορισμό της μή- τρας προβολής | 34 |
| 2.1.6. Η βασική τεχνική του προβλήματος: Bundle Adjustment | 37 |
| 2.2. Παραπομπές στην βιβλιογραφία. Διαθέσιμες υλοποιήσεις . . . | 51 |
| 3. Στερέωση Πολλών Όψεων-MVS | 52 |
| 3.1. Επισκόπηση μεθόδων MVS | 52 |
| 3.1.1. Αναπαράσταση της σκηνής | 52 |
| 3.1.2. Φωτομετρική συνοχή | 53 |
| 3.1.3. Ορατότητα | 53 |
| 3.1.4. Αλγόριθμος ανακατασκευής | 54 |
| 3.1.5. Απαιτήσεις αρχικοποίησης και περιορισμός σχήματος . | 55 |

| | |
|--|------------|
| 3.1.6. Καταλληλότητα μεθόδων για μεγάλα, μη δομημένα, σύνολα φωτογραφιών | 55 |
| 3.2. Μια μέθοδος MVS βασισμένη σε χάρτες βάθους | 56 |
| 3.2.1. Γενική περιγραφή και σύνδεση με προηγούμενες προσεγγίσεις | 56 |
| 3.2.2. Περιγραφή του προτεινόμενου αλγόριθμου | 58 |
| 3.3. Το σύστημα PMVS-Patch Multi View Stereo | 64 |
| 3.3.1. Γενική περιγραφή συστήματος PMVS. Δομές και συναρτήσεις | 64 |
| 3.3.2. Περιγραφή Αλγορίθμου- Ανακατασκευή patch | 66 |
| 3.3.3. Περιγραφή Αλγορίθμου-Πλέγμα (mesh) | 70 |
| 3.3.4. Ενδεικτικά αποτελέσματα | 70 |
| 3.3.5. CMVS: Ένα σύστημα αποσύνθεσης του προβλήματος ανακατασκευής σε μικρότερα υποπροβλήματα | 73 |
| 3.4. Τάσεις & εξελίξεις στον τομέα της στερέωσης πολλών όψεων | 81 |
| 3.4.1. Σύνολα αξιολόγησης | 81 |
| 3.4.2. Παρουσίαση μεθόδων | 81 |
| 3.4.3. Διαθέσιμο Λογισμικό | 86 |
| 4. Ανίχνευση χαρακτηριστικών | 87 |
| 4.1. Η καθιερωμένη προσέγγιση στην ανίχνευση χαρακτηριστικών: Scale Invariant Feature Transform (SIFT) | 88 |
| 4.1.1. Γενικά στοιχεία | 88 |
| 4.1.2. Αναλυτική περιγραφή σταδίων αλγορίθμου SIFT | 89 |
| 4.2. Αξιολόγηση των ανιχνευτών χαρακτηριστικών. Βελτίωση της διακριτικής ικανότητας του SIFT | 94 |
| 4.2.1. Δομή πειραμάτων αξιολόγησης | 95 |
| 4.3. Ο ανιχνευτής SURF | 99 |
| 4.3.1. Χαρακτηριστικά συστήματος ανίχνευσης χαρακτηριστικών SURF | 100 |
| 4.4. Συστηματοποίηση σχεδιασμού ανιχνευτών χαρακτηριστικών | 103 |
| 4.4.1. Γενικά στοιχεία | 103 |
| 4.4.2. Δομικά στοιχεία (blocks) αλγορίθμου | 104 |
| 4.4.3. Αποτελέσματα-Συμπεράσματα | 107 |
| 4.5. DAISY: Ακριβής και αποδοτικός υπολογισμός χαρακτηριστικών. 108 | |
| 4.5.1. Δομή περιγραφέα DAISY | 109 |
| 4.6. Επεκτάσεις και διαθέσιμο λογισμικό | 112 |
| 4.6.1. Επεκτάσεις | 112 |
| 4.6.2. Διαθέσιμο Λογισμικό | 114 |
| II Πειραματικό Μέρος | 116 |
| 5. Μετρική ανακατασκευή & βαθμονόμηση από 2 όψεις | 117 |
| 5.1. Περιγραφή | 117 |
| 5.1.1. Εισαγωγικοί ορισμοί | 117 |
| 5.1.2. Εξισώσεις βαθμονόμησης | 120 |
| 5.1.3. Επίλυση των εξισώσεων και σχέσεις των λύσεων | 126 |
| 5.2. Πρακτικοί προβληματισμοί | 150 |
| 5.2.1. Επίλυση εξισώσεων | 150 |
| 5.2.2. Η περίπτωση του Θορύβου | 153 |

| | |
|--|-----|
| 5.3. Πειράματα | 154 |
| 5.3.1. Πειράματα με συνθετικά δεδομένα | 154 |
| 5.3.2. Εκφυλισμένες διατάξεις. Μια θεωρητική και πειραμα- τική προσέγγιση | 162 |
| 5.3.3. Πειράματα με πραγματικά δεδομένα | 166 |
| 6. Τελική Ανακατασκευή Σκηνης | 179 |
| 6.1. Στόχος Κεφαλαίου | 179 |
| 6.2. Περιγραφή Μεθόδου | 180 |
| 6.2.1. SfM χωρίς την επαναληπτική εφαρμογή BA | 180 |
| 6.2.2. Περιγραφή και ανάλυση της μεθόδου που υλοποιήσαμε | 182 |
| 6.3. Πειράματα | 186 |
| 6.4. Παρουσίαση Αποτελεσμάτων | 190 |
| 6.5. Εφαρμογή της μεθόδου στην Ακρόπολη | 202 |
| 6.6. Προτάσεις για βελτίωση και ανάπτυξη του αλγορίθμου | 207 |
| 7. Βελτιώσεις στην Βασική προσέγγιση | 210 |
| 7.1. Γεωμετρική επαλήθευση των αρχικών αντιστοιχιών | 210 |
| 7.1.1. Γενικά για την Γεωμετρική Επαλήθευση | 210 |
| 7.1.2. Αναγωγή του προβλήματος στο πρόβλημα της μέγιστης κοινής υπακολουθίας-Longest Common Subsequence | 211 |
| 7.1.3. Αποκλίσεις από το αρχικό μοντέλο και βελτιώσεις στην βασική μέθοδο | 214 |
| 7.2. Βελτιώσεις στις εκτιμήσεις των εξωτερικών και εσωτερικών πα- ραμέτρων των όψεων | 222 |
| 7.2.1. Εισαγωγή: κίνητρα και στόχοι | 222 |
| 7.2.2. Σκιαγραφώντας την Γενική Μέθοδο | 222 |
| 7.2.3. Μια σύντομη περιγραφή της στάθμισης περιστροφών με τον αλγόριθμο του Weiszfeld | 223 |
| 7.2.4. Πρώτη υλοποίηση | 224 |
| 7.2.5. Βελτιώσεις στην πρώτη υλοποίηση | 225 |
| 7.3. Αποτελέσματα | 238 |
| 7.3.1. Αξιολόγηση με δείκτες σφάλματος | 238 |
| 7.3.2. Οπτική Αξιολόγηση | 240 |
| 7.4. Επίλογος | 242 |
| 8. Επίλογος | 244 |
| 8.1. Συμβολή της διπλωματικής εργασίας | 244 |
| 8.2. Μελλοντικές κατευθύνσεις | 245 |
| Παραρτήματα | 246 |
| A Παράρτημα: Προσθήκη Θορύβου σε εικόνες για πειράματα με συν- θετικά δεδομένα | 247 |
| B Παράρτημα: Υπολογισμός σφάλματος στην θέση και τον προσανατο- λισμό κάμερας | 249 |
| Παράρτημα: Κατάλογος σχημάτων | 251 |
| Παράρτημα: Κατάλογος πινάκων | 255 |

Αντικείμενο Διπλωματικής Εργασίας

Η παρούσα Διπλωματική Εργασία, ασχολείται με το πρόβλημα της *Ανακατασκευής στην Όραση Υπολογιστών*. Δοσμένου ενός συνόλου φωτογραφιών, η ανακατασκευή της σκηνής στοχεύει στον υπολογισμό ενός τρισδιάστατου μοντέλου της σκηνής. Η αναπαράσταση του μοντέλου αυτού μπορεί να γίνει με διαφορετικούς τρόπους, από την απλούστερη αναπαράσταση ως σύνολο τρισδιάστατων σημείων μέχρι ένα ολοκληρωμένο τρισδιάστατο μοντέλο της επιφάνειας.

Στο πρόβλημα της ανακατασκευής, το είδος των διαθέσιμων εικόνων καθώς και οι επιπλέον πληροφορίες που έχουμε σχετικά με αυτές μπορεί να διαφέρουν αρκετά. Για παράδειγμα ενδέχεται να έχουμε ένα μη-οργανωμένο σύνολο φωτογραφιών, μια ακολουθία φωτογραφιών, ένα video κλπ. Επίσης, άλλοι παράμετροι του προβλήματος, όπως η σχετική επικάλυψη των φωτογραφιών, η σειρά λήψης τους, οι σχετικές θέσεις λήψης του, οι παράμετροι των συσκευών φωτογράφισης κ.α, ενδέχεται να θεωρηθούν γνωστά.

Συμπεραίνουμε έτσι, ότι το πρόβλημα της ανακατασκευής, περιλαμβάνει ένα μεγάλο εύρος διαφορετικών μεταξύ τους προβλημάτων, τα οποία αντιμετωπίζονται με ένα αντίστοιχα μεγάλο εύρος διαφορετικών τεχνικών, συχνά όχι παρόμοιων.

Στην συγκεκριμένη εργασία, θεωρήσαμε ότι έχουμε ένα μη οργανωμένο σύνολο φωτογραφιών. Εκτός από τις ίδιες τις φωτογραφίες, υποθέσαμε ότι δεν δίνεται κάποιας μορφής άλλη πληροφορία. Θεωρήσαμε δηλαδή ότι τόσο οι θέσεις λήψεως των φωτογραφιών, όσο και τα χαρακτηριστικά των καμερών που χρησιμοποιήθηκαν για την λήψη τους (εσωτερικές παράμετροι κάμερας) είναι άγνωστα.

Εστίασαμε την εργασία στην ανάκτηση των πληροφοριών αυτών, ώστε να μπορέσουμε να αξιοποιήσουμε τεχνικές της περιοχής της ανακατασκευής που αναπτύχθηκαν κυρίως στο τέλος της δεκαετίας του 00', και οι οποίες έχουν εξελιχθεί σε ικανοποιητικό βαθμό. Επίσης, προσπαθήσαμε να ακολουθήσουμε σε αυτά τα προβλήματα, όσο το δυνατόν πιο ανεπτυγμένες τεχνικές, ακολουθώντας τις ερευνητικές τάσεις που αναπτύχθηκαν ακόλουθα της δημοσίευσης των αρχικών επιτυχημένων, και πλέον καθιερωμένων, τεχνικών. Στην επόμενη Ενότητα, θα συνοψίσουμε το περιεχόμενο της εργασίας ανά Κεφάλαιο. Αναφέρουμε εδώ κάποιες από τις κύριες κατευθύνσεις και συνεισφορές της Εργασίας

- Αναλυτική παρουσίαση της δομής του προβλήματος, των περιοχών στις

οποίες αυτό χωρίζεται και των διαθέσιμων τεχνικών. Παρουσιάζονται αναλυτικά τα προβλήματα της ανίχνευσης χαρακτηριστικών, του ταιριασματος σημειακών χαρακτηριστικών εικόνων, της επίλυσης των προβλημάτων Self-Calibration, Pairwise Metric Reconstruction, Structure from Motion, Multi-View Stereo και επίσης βασικών προβλημάτων της γεωμετρίας πολλών όψεων.

- Μελετάται το πρόβλημα της γεωμετρικής επαλήθευσης των αντιστοιχιών ανάμεσα στις εικόνες και προτείνεται μέθοδος κατάλληλη για το πρόβλημα της ανακατασκευής αρχιτεκτονικών σκηνών
- Μελετάται το πρόβλημα της βαθμονόμησης και της μετρικής ανακατασκευής ζεύγους καμερών. Προτείνεται κατάλληλη μέθοδος για την επίλυση του προβλήματος, η οποία αναλύεται τόσο θεωρητικά όσο και πειραματικά.
- Η προηγούμενη μέθοδος βελτιώνεται με την εισαγωγή τεχνικής για την στάθμιση των πολυάριθμων εκτιμήσεων για τις εσωτερικές παραμέτρους των καμερών και με την εφαρμογή αλγορίθμου για την στάθμιση των πολυάριθμων εκτιμήσεων για τις σχετικές περιστροφές των καμερών. Επιλέγεται κατάλληλος αλγόριθμος για το υπόλοιπο της διαδικασίας της ανακατασκευής που δεν απαιτεί την γνώση των σχετικών μετατοπίσεων των καμερών (πρόβλημα Structure from Motion with known Rotations)
- Γίνεται ενσωμάτωση των προηγούμενων τεχνικών που αναπτύξαμε σε γνωστές μη ακολουθιακές προσεγγίσεις στο πρόβλημα Structure from Motion που βασίζονται στην επίλυση προβλημάτων βελτιστοποίησης χρησιμοποιώντας το πλαίσιο της L_∞ νόρμας κατάλληλα ορισμένου σφάλματος ανά σημείο δεδομένων. Στην περίπτωση μας, κάθε ταιριασμένο ζεύγος σημείων δύο εικόνων είναι ένα σημείο δεδομένων.
- Το αποτέλεσμα που λαμβάνουμε από την επίλυση του προβλήματος Structure from Motion, μετατρέπεται κατάλληλα ώστε να μπορεί να χρησιμοποιηθεί ως είσοδος σε διαθέσιμες υλοποιήσεις για την τελική ανακατασκευή της τρισδιάστατης επιφάνειας και (ζαναμετετρέπεται) για το texture-mapping της επιφάνειας αυτής.

Διάρθρωση Διπλωματικής Εργασίας

Η διπλωματική εργασία χωρίζεται σε δύο μέρη. Οι *συνεισφορές* της εργασίας περιλαμβάνονται στο *Δεύτερο μέρος* (βλ. Πίνακα 2). Μια γρήγορη επισκόπηση ορισμένων αποτελεσμάτων μας είναι δυνατή στο¹. Στο *Πρώτο Μέρος*, γίνεται μια ενδελεχής ανασκόπηση της σχετικής βιβλιογραφίας, των προβλημάτων που σχετίζονται με την ανακατασκευή και των διαθέσιμων τεχνικών για την επίλυσή τους (βλ. Πίνακα 1). Το μέρος αυτό λειτουργεί συμπληρωματικά, ώστε να είναι διαθέσιμο το σχετικό θεωρητικό υπόβαθρο και να γίνουν σαφείς διαφορετικές προσεγγίσεις που έχουν επιχειρηθεί για την αντιμετώπιση των σχετικών με την ανακατασκευή προβλημάτων.

Στο Κεφάλαιο 1, παρουσιάζουμε βασικές αρχές της προβολικής γεωμετρίας, μοντέλα δισδιάστατων και τρισδιάστατων μετασχηματισμών, την διαδικασία

¹<http://www.youtube.com/channel/UCp5O-gzDEOZ77QpQl6BTQUg/videos>

της προβολής και την μοντελοποίηση μέσω αυτής της διαδικασίας σχηματισμού εικόνων από φωτογραφικές κάμερες. Κάνουμε τέλος και σύντομες αναφορές στις αποκλίσεις από το μοντέλο αυτό.

Στο Κεφάλαιο 2, εισάγουμε τεχνικές και μεθόδους της γεωμετρίας πολλών όψεων (Θεμελιώδης Πίνακας, σχετικά σφάλματα για εφαρμογές γεωμετρίας πολλών όψεων, DLT, RANSAC) καθώς και μικρά προβλήματα που προκύπτουν φυσικά κατά τον χειρισμό του προβλήματος της ανακατασκευής, σε διάφορα σημεία οποιουδήποτε αλγορίθμου (π.χ τριγωνισμός σημείων, resectioning κ.α). Επίσης περιγράφουμε τις μεθόδους και υλοποιήσεις για την επίλυση του προβλήματος Structure From Motion.

Στο Κεφάλαιο 3, περιγράφουμε μεθόδους και υλοποιήσεις για την αναβάθμιση της αναπαράστασης ως αραιό σύννεφο τρισδιάστατων σημείων που έχουμε μετά την επίλυση του προβλήματος SfM, σε αναπαράσταση πυκνού σύννεφου τρισδιάστατων σημείων και τρισδιάστατης επιφάνειας.

Στο Τελευταίο Κεφάλαιο του Πρώτου μέρους (Κεφάλαιο 4), περιγράφουμε μεθόδους ανίχνευσης περιοχών ενδιαφέροντος και περιγραφής των περιοχών αυτών (Region detectors and descriptors).

Στο δεύτερο μέρος, συμπεριλαμβάνονται οι τεχνικές που εισάγαμε και που χρησιμοποιήσαμε για την επίλυση του προβλήματος ανακατασκευής, όπως αυτό περιγράφηκε και στην αμέσως προηγούμενη Ενότητα. Επίσης, συμπεριλαμβάνονται τα αποτελέσματα στα διάφορα πειράματα που διεξήχθησαν για την αξιολόγηση των μεθόδων και για την περιγραφή των ιδιοτήτων τους. Στο Κεφάλαιο 5, εξετάζεται το πρόβλημα της μετρικής ανακατασκευής και βαθμονόμησης 2 όψεων(φωτογραφιών). Αναλύεται η δομή του προβλήματος και προτείνουμε κατάλληλη γραμμική μέθοδο επίλυσης, η οποία όμως οδηγεί σε αμφισημία δύο λύσεων. Αναλύουμε θεωρητικά τις ιδιότητες της τεχνικής επίλυσης (εκφυλισμένες διατάξεις, επίδραση θορύβου) και την σχέση των δύο λύσεων στον χώρο. Στην συνέχεια αναλύουμε πειραματικά ιδιότητες της μεθόδου και σχολιάζουμε την εφαρμογή της και σε πραγματικά δεδομένα για τα οποία γνωρίζουμε τις ορθές τιμές των υπό προσδιορισμό μεγεθών.

Στο Κεφάλαιο 6, αναλύουμε πως μπορεί να συνδυαστούν οι μέθοδοι του προηγούμενου Κεφαλαίου με μη-ακολουθιακές μεθόδους επίλυσης του προβλήματος Structure From Motion στο πλαίσιο της L_∞ νόρμας. Περιγράφουμε την μέθοδο που οδηγεί στην επίλυση του συνολικού προβλήματος και την εφαρμόζουμε τόσο σε σύνολα δεδομένων σχεδιασμένα για το πρόβλημα της Ανακατασκευής, για τα οποία γνωρίζουμε τις ορθές σχετικές θέσεις των καμερών, όσο και σε ακόμα λιγότερο δομημένα σύνολο που δημιουργήσαμε εμείς με στοιχειώδη φωτογραφικό εξοπλισμό. Τα αποτελέσματα της μεθόδου συγκρίνονται και αναλύονται, ενώ σχολιάζεται και η επίδραση διάφορων τροποποιήσεων παραμέτρων και αλγορίθμων της μεθόδου ανακατασκευής.

Στο Κεφάλαιο 7, επιστρέφουμε στις μεθόδους των προηγούμενων Κεφαλαίων με σκοπό την βελτίωση τους ώστε να καταστούν περισσότερο σθεναρές και αξιόπιστες, αλλά και να περιοριστούν τα σφάλματα στην τρισδιάστατη γεωμετρία και τις θέσεις των καμερών. Βελτιώνουμε τις αντιστοιχίες σημείων των φωτογραφιών, εφαρμόζοντας γεωμετρική επαλήθευση σε αυτές. Για την περάτωση της γεωμετρικής επαλήθευσης, προτείνουμε κατάλληλη μέθοδο βασισμένη στο πρόβλημα της Μέγιστης Κοινής Υπακολουθίας. Επίσης βελτιώνουμε τις εκτιμήσεις για τις θέσεις των φωτογραφιών, χρησιμοποιώντας σε κατάλληλα σημεία τους αλγορίθμου επίλυσης μέθοδο στάθμισης περιστροφών βασισμένο στον αλγόριθμο του Weiszfeld. Τέλος, βελτιώνουμε τις εκτιμήσεις για το εστιακό μήκος κάθε κάμερας, με την εισαγωγή κατάλλη-

λης μεθόδου στάθμισης των διάφορων εκτιμήσεων που διαθέτουμε για το εστιακό μήκος κάθε κάμερας.

| Κεφάλαιο | Κύριες πηγές |
|------------|----------------------------------|
| Κεφάλαιο 1 | [90] |
| Κεφάλαιο 2 | [57], [130], [51], [249] |
| Κεφάλαιο 3 | [66], [64], [193], [78] |
| Κεφάλαιο 4 | [133], [149], [18], [243], [225] |

Πίνακας 1: Βασικές βιβλιογραφικές πηγές που παρουσιάζονται στο πρώτο μέρος

| Κεφάλαιο | Κύρια συνεισφορά |
|------------|--|
| Κεφάλαιο 5 | Βαθμονόμηση καμερών με μεταβαλλόμενες εσωτερικές παραμέτρους |
| Κεφάλαιο 6 | Διαμόρφωση πλήρους συστήματος ανακατασκευής |
| Κεφάλαιο 7 | Γεωμετρική επαλήθευση αντιστοιχιών, στάθμιση περιστροφών για εκτίμηση σχετικής θέσης όψης και χρήση δομής προβλήματος για εκτίμηση εσωτερικών παραμέτρων |

Πίνακας 2: Βασικές συνεισφορές Κεφαλαίων δεύτερου μέρους

Μέρος Ι
Βιβλιογραφική Ανασκόπηση

Το Κεφάλαιο βασίζεται στο [90]

1.1 Δισδιάστατη γεωμετρία

Η πιο συνηθισμένη αναπαράσταση σημείων επιπέδου είναι μέσω των καρτεσιανών συντεταγμένων. Έτσι ένα σημείο στο επίπεδο αναπαρίσταται ως $\mathbf{x} = [x \ y]^T$. Για συντομία θα γράφουμε τα διανύσματα ως $\mathbf{x} = (x, y)$.

Συχνά είναι προτιμότερο να χρησιμοποιούμε για τα σημεία του επιπέδου *ομογενείς συντεταγμένες*. Σε αυτές τα σημεία αναπαρίστανται ως $\tilde{\mathbf{x}} = (a, b, c)$. Κάθε σημείο (x, y) παριστάνεται από όλα τα σημεία με ομογενείς συντεταγμένες (a, b, c) για τα οποία ισχύει $x = a/c, y = b/c$. Τα διανύσματα του $\mathbb{R}^3 - (0, 0, 0)$ σχηματίζουν τον προβολικό χώρο P^2 . Τα σημεία του P^2 της μορφής $(a, b, 0)$ λέγονται σημεία στο άπειρο και θα τα αναλύσουμε παρακάτω. Για τα ιδανικά (στο άπειρο) σημεία δεν υπάρχει παράσταση σε μη-ομογενείς συντεταγμένες. Κάθε σημείο του διανυσματικού χώρου \mathbb{R}^2 ορίζει μια κλάση ισοδυναμίας στον P^2 . Τα σημεία στο άπειρο δεν έχουν εικόνα στον \mathbb{R}^2 . Τέλος το σημείο $(0, 0, 0)$ του \mathbb{R}^3 δεν ανήκει στον προβολικό χώρο P^2 .

Ευθείες

Είναι γνωστό ότι μια ευθεία ορίζεται από την σχέση

$$ax + by + c = 0$$

Παρατηρούμε ότι οι ευθείες μπορούν να παρασταθούν στον P^2 ως διανύσματα του χώρου, δηλαδή η προηγούμενη ευθεία παρίσταται ως (a, b, c) . Η κλιμάκωση με σταθερά είναι ξανά αδιάφορη, δηλαδή ισοδύναμη παράσταση είναι κάθε παράσταση της μορφής $\lambda(a, b, c), \lambda \neq 0$. Παρατηρούμε ότι στον προβολικό χώρο ευθείες και σημεία παρίστανται με τον ίδιο τρόπο. Στο εξής θα συμβολίζουμε τις ευθείες ως l . Όπως είναι φανερό και από την παράσταση της ευθείας σε ομογενείς συντεταγμένες, μια ευθεία του επιπέδου έχει 2 βαθμούς ελευθερίας.

Ένα σημείο ανήκει σε μια ευθεία αν και μόνο αν

$$l^T \tilde{\mathbf{x}} = 0$$

Η διατύπωση αυτή είναι απλά μια συμπαγής αναπαράσταση της εξίσωσης ευθείας με την χρήση του εσωτερικού γινομένου διανυσμάτων σε ομογενείς

συντεταγμένες.

Το σημείο τομής δύο ευθειών δίνεται από την σχέση

$$\tilde{\mathbf{x}} = \mathbf{l}_1 \times \mathbf{l}_2$$

Υπενθυμίζουμε ότι το εξωτερικό γινόμενο δύο διανυσμάτων δίνεται από την σχέση

$$\mathbf{x}_1 \times \mathbf{x}_2 = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \end{vmatrix}$$

Από την γνωστή διανυσματική σχέση $\mathbf{x}_1^T(\mathbf{x}_1 \times \mathbf{x}_2) = 0$ έχουμε ότι

$$\mathbf{l}_1^T(\mathbf{l}_1 \times \mathbf{l}_2) = \mathbf{l}_2^T(\mathbf{l}_1 \times \mathbf{l}_2) = 0$$

Άρα το σημείο $(\mathbf{l}_1 \times \mathbf{l}_2)$ ανήκει πράγματι και στις δύο ευθείες και συνεπώς είναι το σημείο τομής τους.

Δύο παράλληλες ευθείες $\mathbf{l} = (a, b, c)$, $\mathbf{l}' = (a, b, d)$ τέμνονται στο σημείο $(b, -a, 0)$. Παρατηρούμε επομένως ότι όλες οι ευθείες τέμνονται και οι παράλληλες ευθείες τέμνονται σε σημεία στο άπειρο. Δεν είναι δύσκολο να επαληθεύσουμε ότι η ευθεία

$$\mathbf{l}_\infty = (0, 0, 1)$$

περιέχει όλα τα σημεία στο άπειρο. Ακόμα μπορούμε να σκεφτούμε την \mathbf{l}_∞ ως το σύνολο των διευθύνσεων των ευθειών του επιπέδου. Δύο σημεία ορίζουν μια ευθεία και η εξίσωση της ευθείας δίνεται από την σχέση

$$l = \tilde{\mathbf{x}}_1 \times \tilde{\mathbf{x}}_2$$

Μπορούμε επίσης να ορίσουμε επιφάνειες δευτέρου βαθμού στο επίπεδο, στις οποίες θα αναφερόμαστε εξ ου με τον όρο κωνικές. Οι επιφάνειες αυτές ορίζονται από σχέσεις της μορφής $\tilde{\mathbf{x}}^T Q \tilde{\mathbf{x}}$, όπου ο Q είναι συμμετρικός πίνακας. Στην γενική περίπτωση οι βαθμοί ελευθερίας είναι 5.

1.1.1 Μετασχηματισμοί

Παρακάτω παρουσιάζουμε μετασχηματισμούς στις δύο διαστάσεις. Αρχίζουμε από τις πιο γενικές μορφές μετασχηματισμών και εξειδικεύουμε. Αυτό σημαίνει ότι οι μετασχηματισμοί που παρουσιάζονται αργότερα είναι ειδικές περιπτώσεις όσων παρουσιάστηκαν πριν από αυτούς.

Προβολικός Μετασχηματισμός ή Ομογραφία

Ο μετασχηματισμός αυτός ορίζεται από τον 3×3 ομαλό πίνακα και την σχέση

$$\tilde{\mathbf{x}}' = H \tilde{\mathbf{x}}$$

Από τον ορισμό των ομογενών συντεταγμένων είναι προφανές ότι κάθε πίνακας λH ορίζει τον ίδιο μετασχηματισμό, οπότε έχουμε 8 βαθμούς ελευθερίας. Στην γενική αυτή περίπτωση μπορούμε να πούμε ότι διατηρούνται οι *ευθείες*, δηλαδή ότι τα ευθύγραμμα τμήματα παραμένουν ευθύγραμμα μετά τον μετασχηματισμό. Είναι η πιο γενική μορφή μετασχηματισμού στις 2 διαστάσεις. Ο πίνακας του μετασχηματισμού μπορεί να γραφεί σε μπλοκ μορφή ως

$$H = \begin{bmatrix} A_{2 \times 2} & \mathbf{t} \\ \mathbf{v}^T & u \end{bmatrix}$$

Παρατηρούμε ότι ιδανικά σημεία απεικονίζονται σε σημεία του επιπέδου και αντίστροφα σημεία του επιπέδου απεικονίζονται στο άπειρο. Από αυτή την παρατήρηση είναι προφανές ότι η παραλληλία ευθειών δεν διατηρείται. Τα μήκη τμημάτων μιας ευθείας επίσης δεν διατηρούνται. Διατηρείται όμως ο λόγος των λόγων αυτών των μηκών.

Αφινικός μετασχηματισμός

Ένας αφινικός μετασχηματισμός έχει την μορφή

$$\tilde{\mathbf{x}}' = H_A \tilde{\mathbf{x}} = \begin{bmatrix} A_{2 \times 2} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix} \tilde{\mathbf{x}}$$

Ο αφινικός μετασχηματισμός είναι γραμμικός μετασχηματισμός των μη ομογενών μεταβλητών και έχει 6 βαθμούς ελευθερίας. Διατηρεί την *παραλληλία*. Αν και το σχήμα παραμορφώνεται στους αφινικούς μετασχηματισμούς, μια σημαντική ιδιότητα τους είναι ότι δρουν ομοιόμορφα στον χώρο. Έτσι η μεταβολή στην επιφάνεια των αντικειμένων είναι ίδια παντού ($\det(A)$) και δεν εξαρτάται από την θέση τους. Όμοια η αλλαγή στον προσανατολισμό μιας ευθείας εξαρτάται μόνο από τον αρχικό προσανατολισμό της και όχι από την θέση της. Στους προβολικούς μετασχηματισμούς η θέση των γεωμετρικών οντοτήτων παίζει καθοριστικό ρόλο στον μετασχηματισμό τους. Άλλο μέγεθος που διατηρείται είναι οι λόγοι από μήκη παράλληλων ευθύγραμμων τμημάτων.

Μετασχηματισμός ομοιότητας

Ο μετασχηματισμός ομοιότητας αποτελείται από μια στροφή με ισοτροπική κλιμάκωση ακολουθούμενη από μια μεταφορά. Ο πίνακας του μετασχηματισμού δίνεται από την σχέση

$$H_s = \begin{bmatrix} sR_{2 \times 2} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix}$$

όπου ο R είναι πίνακας στροφής, δηλαδή $RR^T = I, \det(R) = 1$. Οι βαθμοί ελευθερίας του μετασχηματισμού είναι 4. Ο μετασχηματισμός ομοιότητας διατηρεί γωνίες μεταξύ ευθειών. Επομένως στον μετασχηματισμό ομοιότητας διατηρείται και το σχήμα των αντικειμένων. Επίσης διατηρείται και ο λόγος από μήκη οποιονδήποτε ευθύγραμμων τμημάτων.

Ευκλείδειος μετασχηματισμός

Αποτελείται από περιστροφή και μεταφορά. Έχουμε

$$H_E = \begin{bmatrix} R_{2 \times 2} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix}$$

Πρόκειται για κίνηση στερεού σώματος. Στον μετασχηματισμό αυτό διατηρείται η ευκλείδεια απόσταση, τα μήκη και οι επιφάνειες.

1.2 Τρισδιάστατη γεωμετρία

Η επέκταση των προηγούμενων από τις δύο στις τρεις διαστάσεις είναι άμεση. Τα σημεία του τρισδιάστατου χώρου μπορούν να αναπαρασταθούν με χρήση μη ομογενών συντεταγμένων ως $\mathbf{X} = (X, Y, Z)$. Μπορούμε επίσης να ορίσουμε τον προβολικό χώρο P^3 και να χρησιμοποιήσουμε την ομογενή αναπαράσταση (X, Y, Z, T) . Γενικεύουμε την ευθεία του επιπέδου με το επίπεδο στον

χώρο. Ένα επίπεδο περιγράφεται από την εξίσωση

$$aX + bY + cZ + d = 0$$

Η εξίσωση αυτή μπορεί να παρασταθεί με συμπαγή τρόπο σε ομογενείς συντεταγμένες ως

$$\Pi^T \tilde{\mathbf{X}} = 0$$

όπου $\Pi = (a, b, c, d)$ το κάθετο στο επίπεδο διάνυσμα. Όπως ήταν αναμενόμενο αντίστοιχη έννοια με την ευθεία στο άπειρο έχει το επίπεδο στο άπειρο. Τέλος ορίζουμε αντίστοιχα και δευτεροβάθμιες επιφάνειες με εξισώσεις της μορφής $\tilde{\mathbf{X}}^T Q \tilde{\mathbf{X}}$.

Οι μετασχηματισμοί των τρισδιάστατων συντεταγμένων ορίζονται και αυτοί σε αντιστοιχία με τους δισδιάστατους. Συνοπτική παρουσίασή τους γίνεται στον Πίνακα 1.1

| Μετασχηματισμός | Περιγραφή | Βαθμοί Ελευθερίας | Αναλλοίωτες |
|-----------------|--|-------------------|------------------------------------|
| Προβολικός | $\begin{bmatrix} A_{3 \times 3} & \mathbf{t} \\ \mathbf{v}^T & u \end{bmatrix}$ | 15 | Τομές επιφανειών |
| Αφινικός | $\begin{bmatrix} A_{3 \times 3} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix}$ | 12 | Παραλληλία επιπέδων Λόγοι Όγκων |
| Ομοιότητας | $\begin{bmatrix} sR_{3 \times 3} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix}$ | 7 | Γωνίες |
| Ευκλείδειος | $\begin{bmatrix} R_{3 \times 3} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix}$ | 6 | Όγκοι |

Πίνακας 1.1: Τρισδιάστατοι Μετασχηματισμοί

1.3 Προβολές από τον τρισδιάστατο στον δισδιάστατο κόσμο. Μοντελοποίηση κάμερας.

Σε αυτή την ενότητα θα περιγράψουμε απεικονίσεις από τον τρισδιάστατο χώρο σε δισδιάστατες εικόνες. Οι διατάξεις οι οποίες κάνουν αυτές τις απεικονίσεις θα ονομάζονται *κάμερες*. Για να μοντελοποιήσουμε τις κάμερες χρησιμοποιούμε πίνακες που περιγράφουν τις απεικονίσεις αυτές. Αυτή η αναπαράσταση μας επιτρέπει να υπολογίζουμε εύκολα γεωμετρικές οντότητες όπως το κέντρο προβολής, το επίπεδο της εικόνας κ.α. τα οποία θα συναντήσουμε ακόλουθα.

1.3.1 Κεντρική Προβολή

Ξεκινάμε με το μοντέλο της *κεντρικής προβολής* η οποία γίνεται μέσω της *μικροπικής κάμερας*. Σε αυτό το μοντέλο, ένα σημείο του χώρου απεικονίζεται στο επίπεδο της εικόνας μέσω της ακτίνας που το συνδέει με το *οπτικό*

κέντρο. Η τομή της ακτίνας αυτής με το επίπεδο της εικόνας είναι η εικόνα του σημείου του χώρου. Για να απλοποιήσουμε τις προκύπτουσες σχέσεις, θεωρούμε τρισσορθογώνιο σύστημα συντεταγμένων στον χώρο με αρχή $(0, 0, 0)$ το οπτικό κέντρο C και το επίπεδο της εικόνας κάθετο στον άξονα z . Το κάθετο διάνυσμα στο επίπεδο της εικόνας, εδώ ο άξονας z , ονομάζεται *πρωτεύων άξονας* της προβολής. Ο πρωτεύων άξονας τέμνει το επίπεδο της εικόνας στο σημείο p το οποίο ονομάζεται *πρωτεύον σημείο*. Το οπτικό κέντρο απέχει από το επίπεδο της εικόνας απόσταση f , στην οποία θα αναφερόμαστε ως *εστιακό μήκος*. Για να υπολογίσουμε την εικόνα ενός σημείου του χώρου μπορούμε να χρησιμοποιήσουμε απλές γεωμετρικές σχέσεις και συγκεκριμένα το Θεώρημα του Θαλή (όμοια τρίγωνα). Προκύπτουν άμεσα οι εξής σχέσεις που συνδέουν ένα σημείο του χώρου $\mathbf{X} = (X, Y, Z)$ με την εικόνα του

$$(X, Y, Z) \rightarrow (f^X/z, f^Y/z, f)$$

Αν αγνοήσουμε την 3η συντεταγμένη του σημείου, τότε μπορούμε να συμβολίσουμε την εικόνα του X με $(f^X/z, f^Y/z)$. Η επιλογή αυτή υπονοεί ότι το σύστημα συντεταγμένων που επιλέξαμε στο επίπεδο της εικόνας έχει άξονες x, y ομόροπους με τους αντίστοιχους του συστήματος συντεταγμένων του χώρου και αρχή το πρωτεύον σημείο.

Η εισαγωγή ομογενών συντεταγμένων μας επιτρέπει να παραστήσουμε την προβολή ως γραμμική σχέση. Έχουμε:

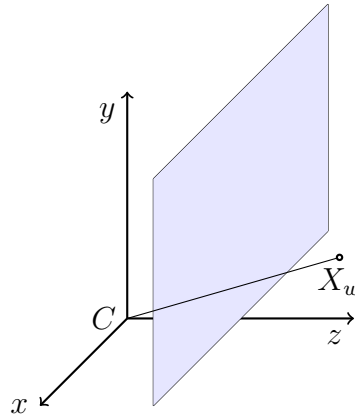
$$\begin{bmatrix} fx \\ fy \\ z \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

Την ανωτέρω σχέση θα την συμβολίζουμε συνήθως

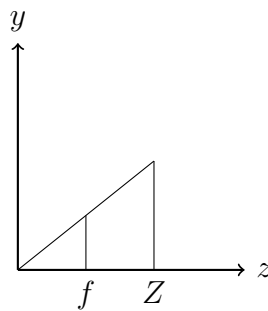
$$\mathbf{x} = P\mathbf{X} = \text{diag}(f, f, 1)[I|\mathbf{0}]\mathbf{X}$$

1.3.2 Γενικεύσεις και επεκτάσεις στο μοντέλο της μικροπικής κάμερας

Το μοντέλο που παρουσιάσαμε πρέπει να γενικευθεί προκειμένου να καταστεί πιο ρεαλιστικό και κατά συνέπεια κατάλληλο για χρήση σε εφαρμογές. Στην προηγούμενη ανάλυση θεωρήσαμε το παγκόσμιο σύστημα συντεταγμένων και το σύστημα συντεταγμένων της κάμερας (πάνω στο επίπεδο της εικόνας) ευθυγραμμισμένα. Αυτή η βολική υπόθεση απλοποίησε σημαντικά τις εξισώσεις μας, πολύ σπάνια όμως πληρείται. Διαφορετικά συστήματα συντεταγμένων εικόνας χρησιμοποιούνται από διάφορες εφαρμογές. Ακόμα, η ταύτιση της αρχής των αξόνων με το οπτικό κέντρο, η οποία εξάλου δεν μπορεί να γίνει αν έχουμε κάμερες σε πάνω από ένα σημεία του χώρου, πρέπει να αναιρεθεί. Οι ευκλείδειοι μετασχηματισμοί μεταξύ των συστημάτων συντεταγμένων που υπονοήσαμε συναποτελούν και τις *γενικεύσεις* του μοντέλου. Ακόμα, το μοντέλο της μικροπικής κάμερας *επεκτείνεται* εύκολα για να συμπεριλάβει φαινόμενα όπως *skew*, *aspect ratio* κ.α.



Σχήμα 1.1: Προοπτική προβολή. Το Οπτικό κέντρο C ταυτίζεται με την αρχή των αξόνων. Ο άξονας z είναι ο κύριος άξονας της κάμερας. Το σημείο τομής $(0, 0, f)$ του κυρίου άξονα με το επίπεδο της εικόνας (μπλε επίπεδο) είναι το πρωτεύον σημείο p . Η απόσταση $(Cp) = f$ είναι το εστιακό μήκος της κάμερας. Το σημείο του χώρου X_w προβάλλεται μέσω της εικονιζόμενης ακτίνας που διέρχεται από το οπτικό κέντρο C στο επίπεδο της εικόνας (μπλε επίπεδο).



Σχήμα 1.2: Χρησιμοποιώντας τα όμοια τρίγωνα, υπολογίζουμε τις συντεταγμένες της προβολής στο επίπεδο της εικόνας

Εσωτερικές παράμετροι. Ο πίνακας βαθμονόμησης

Θεωρούμε ότι το οπτικό κέντρο και η αρχή των συντεταγμένων ταυτίζονται. Θα αναφερόμαστε στις συντεταγμένες αυτές ως *συντεταγμένες κάμερας*, για μεγαλύτερη σαφήνεια. Για να επεκτείνουμε το μοντέλο θεωρούμε ότι έχουμε στην διάθεση μας μετρήσεις σημείων από τον αισθητήρα της κάμερας, δηλαδή μετρήσεις των δισδιάστατων συντεταγμένων του επιπέδου της εικόνας. Θεωρούμε ότι οι εικόνες προέρχονται από ψηφιακές συσκευές, επομένως οι μετρήσεις μας είναι σε pixels. Σκοπός μας είναι να μετατρέψουμε τις μετρήσεις αυτές σε ακτίνες στον χώρο (3 διαστάσεις), με τις εκφράσεις για τον τρισδιάστατο χώρο να γίνονται στις συντεταγμένες της κάμερας. Τα φαινόμενα που θέλουμε να ενσωματώσουμε είναι δύο:

- Την διαμέριση του αισθητήρα σε pixels κατά μήκος και πλάτος της σχηματισμένης εικόνας m_x, m_y .
- Τον προσανατολισμό του συστήματος συντεταγμένων του αισθητήρα σε σχέση με το οπτικό κέντρο, δηλαδή περιστροφή και μετάθεση.

Ενσωματώνουμε τα προηγούμενα στην περιγραφή

$$\mathbf{X}_c = [R_s | \mathbf{c}_s] \begin{bmatrix} m_x & & \\ & m_y & \\ & & 1 \end{bmatrix} \begin{bmatrix} x_s \\ y_s \\ 1 \end{bmatrix} \equiv M \mathbf{x}_s$$

Με \mathbf{X}_c συμβολίζουμε ένα σημείο στον χώρο σε μη ομογενείς συντεταγμένες κάμερας. Η αρχή των συντεταγμένων του συστήματος της εικόνας σε συντεταγμένες κάμερας είναι το διάνυσμα \mathbf{c}_s . Μπορούμε να αντιστρέψουμε την προηγούμενη σχέση και να πάρουμε την σχέση

$$\mathbf{x} = K \mathbf{X}_c$$

Ο πίνακας K ονομάζεται πίνακας *βαθμονόμησης* της κάμερας. Είναι διαστάσεων 3×3 και περιγράφει τις εσωτερικές παραμέτρους της κάμερας.

Εξωτερικές παράμετροι. Ο πίνακας της κάμερας

Για να ενσωματώσουμε τις *εξωτερικές* παραμέτρους, έχουμε ότι ένα σημείο σε συντεταγμένες κάμερας συνδέεται με ένα σημείο σε συντεταγμένες χώρου μέσω ενός μετασχηματισμού $\mathbf{X}_c = [R|\mathbf{T}]\mathbf{X}_w$, όπου \mathbf{X}_w είναι σημείο σε ομογενείς συντεταγμένες χώρου. Έτσι η προηγούμενη εξίσωση γίνεται

$$\mathbf{x} = K[R|\mathbf{T}]\mathbf{X}_w = P\mathbf{X}_w$$

Ο πίνακας P είναι ο πίνακας της κάμερας. Παρατηρούμε στη προηγούμενη σχέση $P = K[R|\mathbf{T}] = KR_1^T[R_1|R_1\mathbf{T}]$, αφού ο πίνακας στροφής R_1 είναι ορθογώνιος και $R_1R_1^T = I$. Επομένως είναι αδύνατον να διαχωρίσουμε τις εσωτερικές και τις εξωτερικές παραμέτρους με μετρήσεις της εικόνας. Για αυτό ο πίνακας της κάμερας P παραγοντοποιείται κατά QR σε γινόμενο πινάκων με τον πίνακα K *άνω τριγωνικό* και τέτοιοι θα τον θεωρούμε. Η πιο απλή μορφή πίνακα βαθμονόμησης που μπορούμε να θεωρήσουμε επομένως είναι της μορφής

$$K = \begin{bmatrix} f & 0 & p_x \\ 0 & f & p_y \\ 0 & 0 & 1 \end{bmatrix}$$

Σε αυτή την περίπτωση θεωρούμε ότι έχουμε ομοιόμορφη κλιμάκωση στις δύο διευθύνσεις της εικόνας και (p_x, p_y) είναι οι συντεταγμένες του πρωτεύοντος σημείου στις συντεταγμένες κάμερας. Σε αυτή την περίπτωση έχουμε 3 βαθμούς ελευθερίας από τον πίνακα βαθμονόμησης και άλλους 6 από τις εξωτερικές παραμέτρους (3 για την μεταφορά και 3 για την περιστροφή). Συνολικά καταλήγουμε με 9 *βαθμούς ελευθερίας*. Αν η κλιμάκωση σε κάθε κατεύθυνση δεν είναι ίδια, όπως για παράδειγμα συμβαίνει αν δεν έχουμε τετράγωνα pixels τότε ο πίνακας K παίρνει την μορφή

$$K = \begin{bmatrix} a_x & 0 & x_o \\ 0 & a_y & y_o \\ 0 & 0 & 1 \end{bmatrix}$$

Τέλος, αν και πολύ σπάνια παρατηρείται, μπορούμε να ενσωματώσουμε στον πίνακα βαθμονόμησης ενδεχόμενη λοξότητα (skew) του συστήματος συντεταγμένων της εικόνας, δηλαδή την μη ορθογωνιότητα των αξόνων. Έχουμε τότε τον πίνακα βαθμονόμησης:

$$K = \begin{bmatrix} a_x & s & x_o \\ 0 & a_y & y_o \\ 0 & 0 & 1 \end{bmatrix}$$

Κλείνουμε την παράγραφο με μερικές παρατηρήσεις. Πρώτον, μπορούμε να συνδέσουμε το σύστημα συντεταγμένων της κάμερας και το σύστημα συντεταγμένων του χώρου με την σχέση $\mathbf{X}_C = \mathbf{R}(\mathbf{X}_W - \mathbf{C})$, όπου όλα τα μεγέθη είναι σε μη ομογενείς συντεταγμένες και \mathbf{C} είναι οι συντεταγμένες του οπτικού κέντρου στο σύστημα συντεταγμένων του χώρου. Έτσι είναι δυνατόν να επαναδιατυπώσουμε τον πίνακα της κάμερας ως $P = K[R] - RC$.

Δεύτερον οποιοσδήποτε πίνακας 3×4 με ομαλό 3×3 αριστερό μπλοκ ορίζει έναν πίνακα κάμερας, αφού μπορεί να παραγοντοποιηθεί κατά QR και να μας δώσει αποδεκτή μορφή πίνακα βαθμονόμησης και πίνακα εξωτερικών παραμέτρων. Προκύπτει επομένως το ερώτημα αν οι 3×4 πίνακες που δεν έχουν ομαλό 3×3 αριστερό μπλοκ και έχουν βαθμό 3 μπορούν να θεωρηθούν πίνακες κάμερας. Η απάντηση είναι καταφατική και οι προκύπτουσες κάμερες έχουν το οπτικό κέντρο στο άπειρο.

1.3.3 Εξαγωγή γεωμετρικών χαρακτηριστικών από την Μήτρα της κάμερας

Όπως εξηγήθηκε, η περιγραφή της κάμερας γίνεται με μια 3×4 μήτρα

$$P = [M|\mathbf{p}]$$

Όταν η μήτρα M είναι ομαλή έχουμε πεπερασμένη κάμερα. Από την μελέτη της μήτρας P είναι δυνατόν να εξαγάγουμε χρήσιμα και ενδιαφέροντα συμπεράσματα.

Το κέντρο προβολής

Είναι διαισθητικά αναμενόμενο το κέντρο προβολής να μην έχει εικόνα στο επίπεδο προβολής. Από την εξίσωση της προβολής σε σύστημα συντεταγμένων κάμερας έχουμε $\mathbf{x} = K\mathbf{X}_o = K\mathbf{0} = \mathbf{0}$. Στην προηγούμενη εξίσωση αντικαταστήσαμε τις συντεταγμένες κάμερας του οπτικού κέντρου $X_o = \mathbf{0}$. Συνεπώς βλέπουμε ότι πράγματι το οπτικό κέντρο δεν αντιστοιχεί σε έγκυρο σημείο της εικόνας. Επεκτείνουμε την προηγούμενη συλλογιστική και αναζητούμε το οπτικό κέντρο ως τον *δεξιό μηδενochώρο της μήτρας κάμερας* P . Η διάσταση του μηδενochώρου της P είναι 1, αφού ο βαθμός της είναι 3. Επειδή η μήτρα M είναι ομαλή έχουμε:

$$\mathbf{0} = [M|\mathbf{p}] [\mathbf{C} \quad 1]^T = M\mathbf{C} + \mathbf{p} \iff \\ \mathbf{C} = -M^{-1}\mathbf{p}$$

Στήλες και γραμμές του P

Οι στήλες του P είναι διανύσματα 3×1 και μπορούν να ερμηνευθούν ως ομογενείς συντεταγμένες σημείων της εικόνας. Εύκολα διαπιστώνεται ότι είναι οι εικόνες των διευθύνσεων (σημείων στο άπειρο) των αξόνων X, Y, Z του συστήματος συντεταγμένων του χώρου. Δηλαδή έχουμε

$$\mathbf{p}_i = P\mathbf{X}_i, i = 1, 2, 3$$

όπου \mathbf{p}_i η i στήλη του P και \mathbf{X}_i το ομογενές διάνυσμα με 1 στην θέση i και 0 αλλού.

Μπορούμε αντίστοιχα να ερμηνεύσουμε τις γραμμές του P ως επίπεδα στον χώρο. Συμβολίζουμε τον πίνακα $P = [P_1^T, P_2^T, P_3^T]$

- Ορίζουμε επίπεδο παράλληλο με το επίπεδο της εικόνας και το οποίο διέρχεται από το οπτικό κέντρο. Ονομάζουμε το επίπεδο αυτό *πρωτεύον επίπεδο*. Για κάθε σημείο του, η ευθεία που διέρχεται από το οπτικό κέντρο είναι παράλληλη με το επίπεδο της εικόνας. Αυτό είναι ξεκάθαρο από την κατασκευή του πρωτεύοντος επιπέδου. Συνεπώς η ακτίνα τέμνει το επίπεδο της εικόνας στο άπειρο. Συνεπώς το πρωτεύον επίπεδο αποτελείται από τα σημεία τα οποία απεικονίζονται στο άπειρο, δηλαδή έχουν εικόνες της μορφής

$$(x, y, 0) \iff \mathbf{P}_3^T \mathbf{X} = 0$$

- Τα επίπεδα $\mathbf{P}_2^T, \mathbf{P}_1^T$ αποτελούνται από τα σημεία που απεικονίζονται στους άξονες x, y . Η ιδιότητα αυτή όμως εξαρτάται από το επιλεγμένο σύστημα συντεταγμένων και δεν είναι τόσο σημαντική.

Τώρα που εντοπίσαμε το πρωτεύον επίπεδο, επάγεται ότι ο *πρωτεύων άξονας* έχει διεύθυνση το κάθετο διάνυσμα του επιπέδου, δηλαδή το διάνυσμα $(p_{31}, p_{32}, p_{33}, 0)$, όπου p_{ij} : το $\langle ij \rangle$ στοιχείο της μήτρας P . Αν προβάλουμε την διεύθυνση αυτή με την μήτρα κάμερας P βρίσκουμε την εικόνα του *πρωτεύοντος σημείου*.

1.3.4 Μη γραμμικά φαινόμενα. Η ακτινική παραμόρφωση

Στην προηγούμενη ανάλυση δεν συμπεριλάβαμε στους υπολογισμούς μας το γεγονός ότι οι προβολές γίνονται μέσω φακών. Οι φακοί προκαλούν μη γραμμικές παραμορφώσεις στην προβολή. Η ακτινική παραμόρφωση γίνεται εμφανής από την στρέβλωση ευθειών της εικόνας. Η παραμόρφωση γίνεται εντονότερη όσο το εστιακό μήκος ελαττώνεται.

Ας συμβολίσουμε με (x, y) τις συντεταγμένες της εικόνας ενός σημείου σε μονάδες οπτικού μήκους, δηλαδή πριν την κλιμάκωση τους κατά f , όπως αυτές προβλέπονται από το *γραμμικό μοντέλο* που αναλύσαμε στις προηγούμενες παραγράφους. Δηλαδή:

$$(x, y, 1) = [I|0]\mathbf{X}_{\text{cam}}$$

Με (x_d, y_d) παριστάνουμε τις συντεταγμένες του σημείου στην εικόνα, μετά την παραμόρφωση που προκαλεί ο φακός. Η σχέση που συνδέει αυτά τα μεγέθη είναι

$$\mathbf{x}_d = L(r)\mathbf{x}$$

Η συνάρτηση $L(r)$ εκφράζει την ακτινική παραμόρφωση. Συνηθίζεται να αναπτύσσουμε την ακτινική παραμόρφωση σε σειρά Taylor και να κρατάμε τους όρους έως r^4 , δηλαδή παίρνουμε $L(r) = 1 + k_1 r^2 + k_2 r^4$. Με r συμβολίζουμε την απόσταση του σημείου της εικόνας από το κέντρο παραμόρφωσης (συχνά αλλά όχι απαραίτητα το πρωτεύον σημείο, $r^2 = x^2 + y^2$). Τα διορθωμένα σημεία της εικόνας προκύπτουν από τις σχέσεις

$$\begin{aligned}x_s &= (x - x_c)L(r) + x_c \\y_s &= (y - y_c)L(r) + y_c \\x_c, y_c &: \text{κέντρο παραμόρφωσης} \\x, y &: \text{σημεία όπως μετρήθηκαν στην εικόνα} \\r^2 &= (x - x_c)^2 + (y - y_c)^2\end{aligned}$$

Μετά την διόρθωση έχουμε επιστρέψει σε ένα γραμμικό σύστημα απεικόνισης, επομένως μπορούμε να χρησιμοποιήσουμε την προηγούμενη ανάλυση και να θεωρήσουμε τους πίνακες της προβολής όπως προηγούμενα.

Για τον υπολογισμό της συνάρτησης L , μπορούμε να την ενσωματώσουμε στην διαδικασία της δημιουργίας εικόνων και να υπολογίσουμε τους συντελεστές του αναπτύγματος της κατά Taylor κατά την διαδικασία προσδιορισμού της μήτρας της κάμερας P . Εναλλακτικά μπορούμε να χρησιμοποιήσουμε διατάξεις όπως αυτή του Tsai [234] και από την απαίτηση οι ευθείες γραμμές να μην στρεβλώνονται με επαναληπτικές διαδικασίες να καταλήξουμε σε κατάλληλες τιμές συντελεστών στο ανάπτυγμα Taylor.

1.3.5 Ανακεφαλαίωση

Κλείνουμε αυτή την ενότητα ανακεφαλαιώνοντας της διαδικασία σχηματισμού εικόνων από κάμερες. Σκοπός είναι να παρουσιάσουμε συγκεντρωτικά την διαδικασία και να την χωρίσουμε σε διακριτά και ξεκάθαρα βήματα. Βασιστήκαμε στην εργασία του [234]. Μπορούμε να περιγράψουμε τον σχηματισμό εικόνων με τα βήματα που φαίνονται στο επόμενο διάγραμμα.

1. Μετασχηματισμός κίνησης *στερεού* σώματος από τις συντεταγμένες του *χώρου* στις συντεταγμένες της *κάμερας*. Η αναπαράσταση στροφή ακολουθούμενη από μετατόπιση καθιστά τον μετασχηματισμό *μοναδικό*.

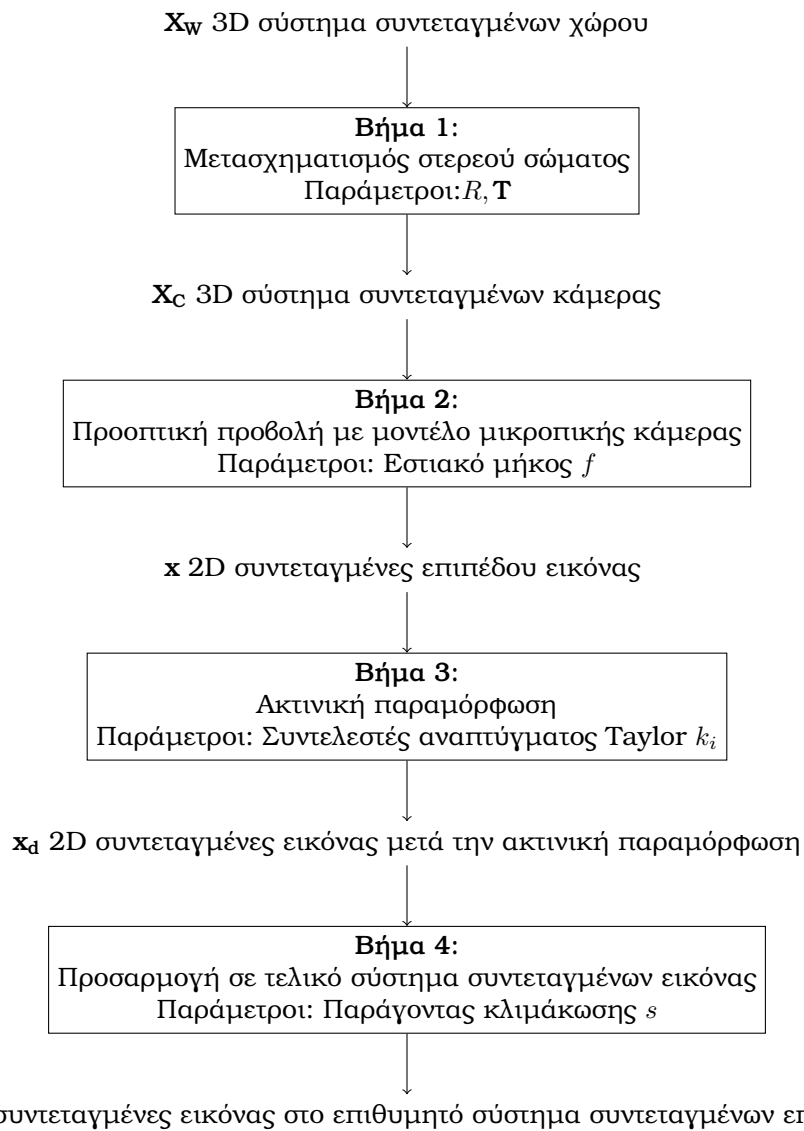
$$\mathbf{X}_C = R\mathbf{X}_W + \mathbf{T}$$

2. Μετασχηματισμός των συντεταγμένων *κάμερας* \mathbf{X}_C σε συντεταγμένες *εικόνας* \mathbf{x} . Οι συντεταγμένες εικόνας σε αυτό το σημείο είναι αποτέλεσμα ιδανικής προβολής μικροπικής κάμερας(προβολή με *γραμμικό* μοντέλο ως προς τις ομογενείς συντεταγμένες).
3. Μετασχηματισμός των *ιδανικών* \mathbf{x} σημείων εικόνας σε *στρεβλωμένα* από την ακτινική παραμόρφωση σημεία \mathbf{x}_d .

$$\begin{aligned}\mathbf{x} &= \mathbf{x}_d(1 + k_1 r^2 + k_2 r^4 \dots) \\r^2 &= x_d^2 + y_d^2\end{aligned}$$

Σύμφωνα με τον [234] κρατάμε μόνο μέχρι το r^2 . Άλλες πηγές, συνιστούν 2 όρους στο ανάπτυγμα. Η προσθήκη περισσότερων όρων είναι περιττή και δημιουργεί θέματα αριθμητικής ευστάθειας.

4. Μετασχηματισμός των συντεταγμένων της εικόνας στο σύστημα συντεταγμένων που χρησιμοποιούμε για τις μετρήσεις της εικόνας.



Σχήμα 1.3: Η διαδικασία σχηματισμού εικόνας χωρισμένη σε βήματα. Ανα-
παραγωγή από [234]

II

ΑΝΑΚΤΗΣΗ ΠΑΡΑΜΕΤΡΩΝ ΚΑΜΕΡΑΣ & ΑΡΑΙΗ ΤΡΙΣΔΙΑΣΤΑΤΗ ΑΝΑΚΑΤΑΣΚΕΥΗ. ΣΥΣΤΗΜΑΤΑ SFM

Το Κεφάλαιο βασίζεται κυρίως στα [57], [130], [51], [249]

2.1 Μέθοδοι ανακατασκευής τρισδιάστατης γεωμετρίας και προσδιορισμού παραμέτρων κάμερας

Το πρόβλημα που θα αντιμετωπίσουμε σε αυτό το κεφάλαιο διατυπώνεται ως εξής. Έχουμε σαν δεδομένα αντιστοιχίες μεταξύ σημείων *δισδιάστατων* εικόνων \mathbf{x}_i^j . Γνωρίζουμε δηλαδή ότι τα σημεία \mathbf{x}_i^j των *όψεων* $i = 1 \dots m$ είναι εικόνες του ίδιου σημείου του τρισδιάστατου χώρου \mathbf{X}_j . Το σημείο \mathbf{X}_j είναι άγνωστο, όπως και οι μήτρες προβολής P^i . Ο προσδιορισμός των άγνωστων αυτών ποσοτήτων είναι το βασικό πρόβλημα που θα αντιμετωπίσουμε. Συνηθίζεται να παραμετροποιούμε κάθε όψη ως ανεξάρτητη *κάμερα*, δηλαδή με ανεξάρτητη μήτρα P^i . Θα συμβολίζουμε γενικά το πλήθος των *αντιστοιχιών* μεταξύ εικόνων ως n και το πλήθος των *καμερών* ως m . Το πρόβλημα αυτό είναι γνωστό στην διεθνή βιβλιογραφία με την ονομασία "*Structure From Motion (SfM)*". Όπως θα διαπιστώσουμε σύντομα, πρόκειται για πρόβλημα βελτιστοποίησης. Αναζητούμε την διαμόρφωση καμερών και τρισδιάστατων σημείων τα οποία σχετιζόμενα μεταξύ τους μέσω των εξισώσεων της προοπτικής προβολής είναι μεγιστικά συνεπή με τις δοσμένες αντιστοιχίες δισδιάστατων σημείων[209]. Η συνέπεια συνήθως εκφράζεται μέσω του *σφάλματος επαναπροβολής*. Το πρόβλημα είναι έντονα μη γραμμικό και έμφαση δίνεται στην σθεναρότητα και ικανότητα κλιμάκωσης (με το μέγεθος της εισόδου) της μεθόδου λύσης. Ξεκινάμε αρχικά με μεθόδους που μπορούν να χρησιμοποιηθούν για να *αρχικοποιήσουμε* τις παραμέτρους κοντά στις βέλτιστες τιμές τους.

2.1.1 Βασικές έννοιες. Σφάλμα επαναπροβολής

Οι μετρήσεις σημείων \mathbf{x} σε μία εικόνα δεν είναι απόλυτα ακριβείς. Αντίθετα, και όπως είναι αναμενόμενο, επηρεάζονται από θόρυβο τον οποίο συνήθως θεωρούμε γκαουσιανής μορφής. Έτσι λύνοντας κάποιο υπερορισμένο πρόβλημα γεωμετρίας πολλαπλών όψεων, δεν μπορούμε να περιμένουμε την ακριβή επαλήθευση, από τα πειραματικά σημεία, των εξισώσεων προβολής. Η λύση που αναζητούμε θέλουμε να ελαχιστοποιεί το *σφάλμα επαναπροβολής* το οποίο ορίζεται:

$$\sum d^2(\hat{P}^i \hat{\mathbf{X}}_j, \mathbf{x}_j^i)$$

Με την ένδειξη $\hat{\cdot}$ συμβολίζουμε εκτιμήσεις μεγεθών που προκύπτουν ως λύσεις του προβλήματος. Ο όρος $\hat{P}^i \hat{\mathbf{X}}_j$ εκφράζει την επαναπροβολή του εκτιμηθέντος σημείου του χώρου από το μοντέλο προβολής P που υποθέσαμε. Η ελαχιστοποίηση του αθροίσματος των τετραγώνων των αποστάσεων των επαναπροβολών από τα σημεία-μετρήσεις εκφράζει επομένως με φυσικό τρόπο το ταίριασμα των μετρήσεων μας με την λύση του προβλήματος που προτείνουμε. Σημειώνεται τέλος ότι το ολικό ελάχιστο του προβλήματος που θεωρήσαμε είναι στην περίπτωση του γκαουσιανού θορύβου και λύση *μέγιστης πιθανοφάνειας*.

2.1.2 Παραγοντοποίηση με μοντέλο προοπτικής προβολής

Ας θεωρήσουμε σημείο της εικόνας $\mathbf{x} = \lambda(x, y, 1)$ και μήτρα προβολής $P = [M | \mathbf{P}_4]$. Υποθέτουμε ότι το σημείο \mathbf{x} είναι η εικόνα του σημείου του χώρου $\mathbf{X} = (X, Y, Z, 1) = (\tilde{\mathbf{X}}, 1)$ δηλαδή ότι ισχύει η σχέση $\mathbf{x} = P\mathbf{X}$. Ακόμα θεωρούμε ότι το κέντρο της προβολής έχει συντεταγμένες $\mathbf{C} = (\tilde{\mathbf{C}}, 1)$. Αν συμβολίσουμε με \mathbf{P}^3 το διάνυσμα της τρίτης γραμμής του πίνακα P έχουμε:

$$\begin{aligned} \lambda &= \mathbf{P}^{3T} \mathbf{X} \\ &= \mathbf{P}^{3T} (\mathbf{X} - \mathbf{C}), \text{ αφού } P\mathbf{C} = \mathbf{0} \\ &= \mathbf{m}^{3T} (\tilde{\mathbf{X}} - \tilde{\mathbf{C}}) \end{aligned}$$

Γράψαμε το $\mathbf{P}^3 = \begin{bmatrix} \mathbf{m}^3 \\ P_4^3 \end{bmatrix}$.

Αν θεωρήσουμε ότι το \mathbf{m}^3 έχει κανονικοποιηθεί ώστε να είναι μοναδιαίο διάνυσμα, παρατηρούμε ότι το λ είναι η απόσταση του σημείου \mathbf{X} από το οπτικό κέντρο \mathbf{C} στην διεύθυνση του οπτικού άξονα. Έτσι θα αποκαλούμε την ποσότητα λ *προοπτικό βάθος* στα επόμενα.

Θεωρούμε ότι *όλα τα σημεία είναι ορατά σε κάθε όψη*. Μπορούμε να γράψουμε τις εξισώσεις προβολής $\lambda_j^i(x_j^i, y_j^i, 1) = P^i \mathbf{X}_j$ σε συμπαγή μορφή ως

$$\begin{bmatrix} \lambda_1^1 \mathbf{x}_1^1 & \lambda_2^1 \mathbf{x}_2^1 & \cdots & \lambda_n^1 \mathbf{x}_n^1 \\ \lambda_1^2 \mathbf{x}_1^2 & \lambda_2^2 \mathbf{x}_2^2 & \cdots & \lambda_n^2 \mathbf{x}_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_1^m \mathbf{x}_1^m & \lambda_2^m \mathbf{x}_2^m & \cdots & \lambda_n^m \mathbf{x}_n^m \end{bmatrix} = \begin{bmatrix} P^1 \\ P^2 \\ \vdots \\ P^m \end{bmatrix} [\mathbf{X}_1 \quad \mathbf{X}_2 \quad \cdots \quad \mathbf{X}_n]$$

Τον πίνακα στο αριστερό σκέλος της προηγούμενης εξίσωσης τον ονομάζουμε *πίνακα μετρήσεων* και τον συμβολίζουμε με W . Τα στοιχεία αυτού του πίνακα είναι οι μετρήσεις πάνω στις εικόνες και τα βάθη, τα οποία θεωρούνται γνωστά. Ο πίνακας αυτός πρέπει να έχει *βαθμό 4*, αφού είναι το γινόμενο δύο

μητρών με 4 στήλες και γραμμές αντίστοιχα. Επειδή οι μετρήσεις μας περιέχουν θόρυβο, η προηγούμενη εξίσωση δεν ικανοποιείται ακριβώς. Για αυτό αναζητούμε μήτρα \hat{W} , πλησιέστερη στην W κατά την νόρμα Frobenius, που έχει βαθμό 4. Η λύση αυτού του προβλήματος μπορεί να βρεθεί από την αποσύνθεση SVD της μήτρας W . Έχουμε $W = UDV^T$ και η επιθυμητή λύση δίδεται ως $\hat{W} = U\hat{D}V^T$, όπου η μήτρα \hat{D} προκύπτει από την D αν θέσουμε όλα τα διαγώνια στοιχεία, με εξαίρεση τα τέσσερα πρώτα, μηδέν. Έχουμε πλέον στην διάθεση μας την μήτρα \hat{W} και μπορούμε να ανακτήσουμε τους αγνώστους του προβλήματος από τις σχέσεις

$$\begin{bmatrix} P^1 \\ P^2 \\ \vdots \\ P^m \end{bmatrix} = U\hat{D} \text{ και } [\mathbf{X}_1 \ \mathbf{X}_2 \ \dots \ \mathbf{X}_n] = V^T$$

Η λύση που πήραμε δεν είναι μοναδική. Αντίθετα μπορούμε να παρεμβάλουμε στο δεξί μέρος της εξίσωσης μετρήσεων οποιονδήποτε προβολικό μετασχηματισμό H και τον αντίστροφο του H^{-1} και να πάρουμε μια αποδεκτή λύση. Έχουμε επομένως *αβεβαιότητα προβολικού μετασχηματισμού* στην λύση μας.

Οι τιμές των λ_j^i αν και θεωρήθηκαν γνωστές, στην πράξη δεν είναι διαθέσιμες. Διαφορετικές στρατηγικές αντιμετώπισης του προβλήματος είναι :

1. Θέτουμε τις τιμές των $\lambda \leftarrow 1$ και υπολογίζουμε έναν αρχικό προοπτικό μετασχηματισμό. Οι τιμές που προέκυψαν μπορούν να χρησιμοποιηθούν σαν βελτιωμένες εκτιμήσεις για τα βάθη και η διαδικασία να επαναληφθεί. Έχουν βρεθεί κανόνες ανανέωσης των τιμών μεταξύ των επαναλήψεων που μπορούν να εγγυηθούν σύγκλιση [164].
2. Με μεθόδους δύο ή/και τριών όψεων υπολογίζουμε αρχικές εκτιμήσεις για τα βάθη.

Για να δείξουμε καλύτερα την διαδικασία που ακολουθούμε και να ξεκαθαρίσουμε τον ρόλο των μεταβλητών λ στην λύση του προβλήματος ξαναγράφουμε την υπό ελαχιστοποίηση ποσότητα

$$\|W - \hat{W}\| = \sum_{ij} \|\lambda_j^i \mathbf{x}_j^i - \hat{\lambda}_j^i \hat{\mathbf{x}}_j^i\|^2 = \sum_{ij} (\lambda_j^i x_j^i - \hat{\lambda}_j^i \hat{x}_j^i)^2 + (\lambda_j^i y_j^i - \hat{\lambda}_j^i \hat{y}_j^i)^2 + (\lambda_j^i - \hat{\lambda}_j^i)^2$$

Παρατηρούμε επομένως ότι αν ικανοποιούνται οι επόμενες δύο συνθήκες, τότε η ποσότητα που ελαχιστοποιεί η λύση μας ταυτίζεται με το σφάλμα επαναπροβολής. Έχουμε:

- Οι εκτιμήσεις των λ_j^i να είναι κοντά στις σωστές τιμές τους.
- Όλες οι τιμές των λ_j^i να είναι περίπου ίσες μεταξύ τους.

Αν και αρχικά οι συνθήκες αυτές φαίνονται δύσκολες να ικανοποιούνται στην γενική περίπτωση, θα δείξουμε ότι τα βάθη μπορούν να κανονικοποιηθούν κατά βούληση. Πράγματι από τις εξισώσεις της προοπτικής προβολής έχουμε:

$$\begin{aligned} \lambda \mathbf{x} &= P\mathbf{X} \iff \\ (\alpha\beta\lambda)\mathbf{x} &= (\alpha P)(\beta\mathbf{X}) \end{aligned}$$

Συνεπώς μπορούμε να κανονικοποιήσουμε το λ σε μονάδα πολλαπλασιάζοντας την κατάλληλη γραμμή του πίνακα μετρήσεων με α και την κατάλληλη στήλη με β . Ένας ευριστικός τρόπος να επιτύχουμε μοναδιαία λ είναι να κανονικοποιήσουμε τις γραμμές του πίνακα μετρήσεων ώστε να έχουν μοναδιαίο μέτρο και ακολούθως να κάνουμε την αντίστοιχη κανονικοποίηση των στηλών του πίνακα μετρήσεων. Κατάλληλος αριθμός επαναλήψεων της μεθόδου οδηγεί στο επιθυμητό αποτέλεσμα.

Όπως είναι συνηθισμένο στις αριθμητικές μεθόδους, οι συντεταγμένες των σημείων της εικόνας πρέπει να κανονικοποιηθούν και αυτές προκειμένου να έχουμε αριθμητική ευστάθεια. Μια αποδεκτή στρατηγική κανονικοποίησης είναι η *ισοτροπική κανονικοποίηση*, στην οποία:

- Με κατάλληλο μετασχηματισμό μετάθεσης, μεταφέρουμε το *κεντροϊδές*, $c^i = \sum_j x_j^i$, των σημείων της εικόνας x_j^i στο $(0, 0, 0)$.
- Κάνουμε κλιμάκωση στα σημεία της εικόνας ώστε το *μέσο μήκος* των διανυσμάτων θέσης των σημείων να είναι $\sqrt{2}$.

Για να καταλάβουμε την σημασία της κανονικοποίησης αυτής πρέπει να αναλύσουμε την διαδικασία που ακολουθούμε για την εύρεση της λύσης. Ο πίνακας W των μετρήσεων μεταβάλλεται στον πίνακα \hat{W} ώστε το άθροισμα των διαφορών των στοιχείων του νέου πίνακα \hat{W} από τον αρχικό W να είναι ελάχιστο. Αν δεν είναι όλα τα στοιχεία του W της ίδιας τάξης μεγέθους, τα μικρότερα στοιχεία τού μπορεί να μεταβληθούν σημαντικά χωρίς να επηρεάσουν εξίσου την προς ελαχιστοποίηση ποσότητα. Η ποσότητα προς ελαχιστοποίηση δηλαδή καθορίζεται ουσιαστικά μόνο από τα μεγάλα στοιχεία του W , κάτι που είναι ανεπιθύμητο.

Οι μέθοδοι παραγοντοποίησης έχουν το πλεονέκτημα να παρέχουν λύση σε κλειστό τύπο η οποία μπορεί να χρησιμοποιηθεί για να αρχικοποιήσει επαναληπτικές διαδικασίες επίλυσης του προβλήματος SfM όπως το Bundle Adjustment. Μπορούν επίσης να χρησιμοποιηθούν εναλλακτικές μεθοδολογίες στην θέση των τεχνικών παραγοντοποίησης, όπως για παράδειγμα με την χρήση ενός κοινού επιπέδου το οποίο φαίνεται από όλες τις κάμερες [186].

2.1.3 Επιπολική Γεωμετρία. Ο θεμελιώδης πίνακας F .

Στην ενότητα αυτή μελετάμε την γεωμετρία δύο όψεων. Θα εξάγουμε τις βασικές σχέσεις που τις συνδέουν και χαρακτηρίζουν τον μετασχηματισμό που τις συνδέει.

Έστω σημείο X του τρισδιάστατου χώρου και δύο όψεις (επίπεδα εικόνας). Συμβολίζουμε \mathbf{x}, \mathbf{x}' τις προβολές του X σε κάθε εικόνα αντίστοιχα. Γενικά συμβολίζουμε ως τονούμενα τα μεγέθη που αναφέρονται στην δεύτερη όψη και άτονα τα μεγέθη που αναφέρονται στην πρώτη.

Παρατηρούμε ότι τα σημεία $X, \mathbf{x}, \mathbf{x}'$ και τα κέντρα των προβολών C, C' ανήκουν σε ένα κοινό επίπεδο. Σε αυτή την παρατήρηση θα στηριχθούμε για να εξάγουμε τις επόμενες ιδιότητες. Συμβολίζουμε το κοινό επίπεδο των $X, \mathbf{x}, \mathbf{x}', C, C'$ ως π .

Ας υποθέσουμε ότι γνωρίζουμε μόνο την θέση του σημείου στην πρώτη όψη, \mathbf{x} . Το σημείο \mathbf{x}' στην δεύτερη όψη πρέπει να βρίσκεται στην τομή του επιπέδου π με το επίπεδο της δεύτερης εικόνας. Συνεπώς η αβεβαιότητα μας για την θέση του \mathbf{x}' περιορίστηκε από ολόκληρο το επίπεδο της εικόνας στην ευθεία που τέμνει το επίπεδο π το επίπεδο της εικόνας. Η ευθεία αυτή λέγεται

επιπολική ευθεία. Η ευθεία που ενώνει τα δύο κέντρα προβολής C, C' τέμνει τα επίπεδα της πρώτης και της δεύτερης εικόνας στα σημεία e, e' αντίστοιχα. Τα σημεία αυτά ονομάζονται *επιπολικά σημεία*.

Για να αναπαραστήσουμε με αλγεβρικό τρόπο την επιπολική γεωμετρία θα εισάγουμε τον θεμελιώδη πίνακα F . Αναζητούμε μια απεικόνιση μεταξύ ενός σημείου \mathbf{x} και της επιπολικής ευθείας l' . Από κάθε σημείο \mathbf{x} της πρώτης εικόνας προβάλλουμε μια ακτίνα η οποία διέρχεται από το κέντρο προβολής C και το σημείο \mathbf{x} . Η προβολή της ακτίνας αυτής στο επίπεδο της δεύτερης εικόνας είναι η επιπολική ευθεία η οποία αντιστοιχεί στο σημείο \mathbf{x} . Για καλύτερη κατανόηση, αρκεί να σκεφτούμε ότι κάθε σημείο της ακτίνας προβάλλεται στο σημείο \mathbf{x} στην πρώτη όψη, και σε κάποιο διαφορετικό σημείο της επιπολικής ευθείας l' στην δεύτερη όψη. Συνεχίζουμε με την σκιαγράφηση δύο εναλλακτικών τρόπων εξαγωγής του θεμελιώδους πίνακα.

Γεωμετρική προσέγγιση: Προσδιορίζουμε τον θεμελιώδη πίνακα με μια διαδικασία που χωρίζεται σε δύο βήματα:

- **Μεταφορά σημείου μέσω επιπέδου** Μεταφέρουμε ένα σημείο \mathbf{x} της πρώτης εικόνας σε ένα επίπεδο του χώρου π_n που δεν διέρχεται από κανένα από τα δύο κέντρα προβολής C, C' . Η ακτίνα που διέρχεται από τα C, \mathbf{x} τέμνει το επίπεδο π_n στο σημείο \mathbf{X} . Το σημείο \mathbf{X} προβάλλεται στην συνέχεια στην δεύτερη εικόνα στο σημείο \mathbf{x}' . Έτσι επάγεται μια δισδιάστατη ομογραφία μεταξύ των δύο όψεων.
- **Κατασκευή επιπολικής ευθείας l'** Η επιπολική ευθεία κατασκευάζεται ως η ευθεία που διέρχεται από το επιπολικό σημείο e' και το σημείο \mathbf{x}' .

Αλγεβρική προσέγγιση: Στην δεύτερη αυτή προσέγγιση, την οποία θα αναλύσουμε καλύτερα, ξεκινάμε με την περιγραφή της ακτίνας που διέρχεται από το κέντρο προβολής C και ένα σημείο της πρώτης εικόνας \mathbf{x} . Τα σημεία της συγκεκριμένης ακτίνας θα ικανοποιούν την εξίσωση προβολής

$$\mathbf{x} = P\mathbf{X} \quad (2.1)$$

Λύνουμε την Εξίσωση 2.1 και λαμβάνουμε την παραμετρική εξίσωση ευθείας

$$\mathbf{X}(\lambda) = P^\dagger \mathbf{x} + \lambda \mathbf{C} \quad (2.2)$$

Με P^\dagger συμβολίζουμε τον ψευδοαντίστροφο του P , δηλαδή πίνακα με την ιδιότητα $PP^\dagger = I$. Από την Εξίσωση 2.2 μπορούμε για διάφορες τιμές του λ να πάρουμε σημεία της ακτίνας. Διαλέγουμε τιμές $\lambda = 0, \lambda = \infty$ και έχουμε

$$\begin{aligned} \mathbf{X}_0 &= P^\dagger \mathbf{x} \\ \mathbf{X}_\infty &= \mathbf{C} \end{aligned}$$

Προβάλλουμε τα δύο αυτά σημεία της ακτίνας μέσω της μήτρας προβολής P' στο επίπεδο της δεύτερης εικόνας και έχουμε

$$\begin{aligned} \mathbf{x}'_0 &= P'P^\dagger \mathbf{x} \\ \mathbf{x}'_\infty &= P'\mathbf{C} \end{aligned}$$

Σημειώνουμε την εμφανή ιδιότητα από την γεωμετρία του σχήματος και τον ορισμό του επιπολικού σημείου

$$\mathbf{x}'_\infty = e'$$

Η επιπολική ευθεία l' θα διέρχεται από τα σημεία $\mathbf{x}'_0, \mathbf{e}'$ επομένως έχουμε

$$l' = (P' \mathbf{C}) \times (P' P^\dagger \mathbf{x}) \iff l' = \mathbf{e}' \times (P' P^\dagger) \mathbf{x} \quad (2.3)$$

Ορίζουμε τον πίνακα της απεικόνισης $\mathbf{e}' \times (P' P^\dagger) \equiv F$ ως τον θεμελιώδη πίνακα και έχουμε

$$l' = F \mathbf{x} \quad (2.4)$$

Τώρα θα συνεχίσουμε με μερικές ιδιότητες του θεμελιώδους πίνακα F . Καταρχάς έχουμε ότι το σημείο της δεύτερης όψης \mathbf{x}' που αντιστοιχεί στο \mathbf{x} της πρώτης όψης θα βρίσκεται πάνω στην επιπολική ευθεία l' , δηλαδή

$$\mathbf{x}'^T l' = 0 \quad (2.5)$$

Αντικαθιστώντας την 2.5 στην 2.4 λαμβάνουμε την σχέση

$$\mathbf{x}'^T F \mathbf{x} = 0 \quad (2.6)$$

Η προηγούμενη Σχέση 2.6 μας επιτρέπει να εκφράσουμε τα στοιχεία της μήτρας F με όρους αποκλειστικά σημείων των δύο όψεων και χωρίς την παρεμβολή των μητρών προβολής P, P' . Αυτό μας δίνει την δυνατότητα να προσδιορίσουμε τον πίνακα F από αντιστοιχίες σημείων των όψεων και στην συνέχεια να ανακτήσουμε τους πίνακες προβολής P, P' . Ο θεμελιώδης πίνακας είναι διαστάσεων 3×3 ομογενής πίνακας, συνεπώς θα μπορούσε να έχει το πολύ 8 βαθμούς ελευθερίας. Όμως, από την Σχέση ορισμού 2.3, μας δείχνει ότι η θεμελιώδης μήτρα είναι το γινόμενο μιας μήτρας βαθμού 2, που προκύπτει από το εξωτερικό γινόμενο, και μιας ομογραφίας που περιγράφεται από μήτρα βαθμού 3. Συνεπώς ο θεμελιώδης πίνακας έχει βαθμό 2, επομένως

$$\det(F) = 0$$

Διαπιστώσαμε έτσι ότι ο θεμελιώδης πίνακας έχει 7 βαθμούς ελευθερίας. Για να προσδιορίσουμε λοιπόν τα στοιχεία του θεμελιώδη πίνακα, αρκούν 7 αντιστοιχίες σημείων εφόσον κάθε αντιστοιχία θα μας δώσει και μία γραμμική εξίσωση στα στοιχεία του πίνακα, όπως είναι ξεκάθαρο από την Σχέση 2.6. Η χρήση μικρότερου αριθμού σημείων είναι σημαντική για την διαδικασία του RANSAC, αφού αντιστοιχεί σε μικρότερη πληθικότητα δείγματος. Η διαδικασία RANSAC θα αναλυθεί σε επόμενες ενότητες.

Επίσης μπορεί να δειχθεί ότι οι μήτρες προβολής που προσδιορίζονται από τον θεμελιώδη πίνακα F δύο όψεων έχουν ακρίβεια ενός προβολικού μετασχηματισμού. Ένας τρόπος να εκφράσουμε τις μήτρες προβολής που δίνονται από τον θεμελιώδη πίνακα F είναι

$$P = [I | \mathbf{0}] P' = [\mathbf{e}' \times F | \mathbf{e}']$$

Τέλος σε περίπτωση που γνωρίζουμε τις εσωτερικές παραμέτρους των καμερών, μπορούμε να εκφράσουμε την επιπολική γεωμετρία με πίνακα 5 βαθμών ελευθερίας, αντίστοιχο του θεμελιώδη πίνακα (Essential Matrix).

2.1.4 Το πρόβλημα του Τριγωνισμού

Το πρόβλημα του τριγωνισμού συνίσταται στην εύρεση των τρισδιάστατων σημείων του χώρου \mathbf{X} , δεδομένων αντιστοιχιών σημείων μεταξύ εικόνων $\mathbf{x} \leftrightarrow \mathbf{x}'$ και των μητρών προβολής P . Οι μετρήσεις των εικόνων θεωρούνται θορυβώδεις, με συχνή την παραδοχή Γκαουσιανού θορύβου.

Ένας απλός τρόπος επίλυσης του προβλήματος τριγωνισμού είναι να επιλέξουμε ως λύση \mathbf{p} , το σημείο του τρισδιάστατου χώρου που βρίσκεται πιο κοντά στις ακτίνες των σημείων των εικόνων \mathbf{x}_j , που του αντιστοιχούν. Με \mathbf{x}_j συμβολίζουμε σημεία εικόνων με πίνακες προβολής $P_j = K_j[R_j|\mathbf{t}_j]$ που αντιστοιχούν στο ίδιο σημείο του χώρου. Υπενθυμίζουμε πως το κέντρο προβολής \mathbf{c}_j της όψης j είναι το μοναδικό σημείο χωρίς εικόνα στην όψη j και μπορεί να βρεθεί με τον ακόλουθο συλλογισμό:

$$\begin{aligned} \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} &= K[R|\mathbf{t}] \begin{bmatrix} \mathbf{c} \\ 1 \end{bmatrix} \iff \\ \mathbf{0} &= KR\mathbf{c} + K\mathbf{t} \iff \\ \mathbf{c} &= -R^{-1}\mathbf{t} \end{aligned} \quad (2.7)$$

Για να βρούμε μια έκφραση των σημείων που ανήκουν στην ακτίνα που διέρχεται από τα $\mathbf{c}_j, \mathbf{x}_j$ γράφουμε:

$$\mathbf{p}_j = \mathbf{c}_j + \hat{\mathbf{x}}_j d_j \quad (2.8)$$

Στην προηγούμενη Εξίσωση 2.8 με $\hat{\mathbf{x}}_j$ συμβολίσαμε το μοναδιαίο διάνυσμα διεύθυνσης της ακτίνας. Δεδομένου ότι όλα τα σημεία της ακτίνας προβάλλονται στην όψη j στο ίδιο σημείο \mathbf{x}_j έχουμε τον ακόλουθο συλλογισμό:

$$P_j \begin{bmatrix} \mathbf{p}_j \\ 1 \end{bmatrix} = A\mathbf{x}_j \quad (2.9)$$

Όπου η A είναι κατάλληλη σταθερά και το αποτέλεσμα είναι άμεση συνέπεια της προβολής στο σημείο της εικόνας \mathbf{x}_j και της παράστασης με ομογενείς συντεταγμένες. Με χρήση της 2.8 για να αντικαταστήσουμε το \mathbf{p}_j και της 2.7 για να γράψουμε την μήτρα προβολής P_j έχουμε

$$\begin{aligned} P_j \mathbf{p}_j &= d_j KR\hat{\mathbf{x}}_j = A\mathbf{x}_j \iff \\ \hat{\mathbf{x}}_j &= N(R_j^{-1}K_j^{-1})\mathbf{x}_j \end{aligned} \quad (2.10)$$

Αναζητούμε το σημείο του χώρου \mathbf{p}^* που ελαχιστοποιεί την απόσταση από τα σημεία της ακτίνας 2.8 δηλαδή ελαχιστοποιεί την ποσότητα

$$\|\mathbf{c}_j + d_j \hat{\mathbf{x}}_j - \mathbf{p}^*\|^2 \quad (2.11)$$

Το σημείο ελάχιστου της 2.11 βρίσκεται

$$d_j^* = \hat{\mathbf{x}}_j(\mathbf{p}^* - \mathbf{c}_j) \quad (2.12)$$

Η διατύπωση του προβλήματος ως σύνθηρες πρόβλημα ελαχίστων τετραγώνων ολοκληρώνεται αθροίζοντας τα σφάλματα 2.11 για όλες τις όψεις j αντικαθιστώντας τα d_j^* από την 2.12 και βελτιστοποιώντας ως προς το σημείο \mathbf{p}^* . Εναλλακτική και καλύτερη προσέγγιση είναι να ελαχιστοποιήσουμε το σφάλμα στις εξισώσεις προβολής

$$\mathbf{x}_j = P\mathbf{X}$$

Έχουμε δηλαδή τις δύο εξισώσεις

$$x_j = \frac{\mathbf{p}^1 \mathbf{X}}{\mathbf{p}^3 \mathbf{X}} \quad (2.13)$$

$$y_j = \frac{\mathbf{p}^2 \mathbf{X}}{\mathbf{p}^3 \mathbf{X}} \quad (2.14)$$

Χρησιμοποιήσαμε τον συμβολισμό

$$P = \begin{bmatrix} \mathbf{p}^1 \\ \mathbf{p}^2 \\ \mathbf{p}^3 \end{bmatrix}$$

Οι εξισώσεις 2.13 μπορούν να μετατραπούν σε γραμμικές ως προς τον άγνωστο \mathbf{x} πολλαπλασιάζοντας με τον κοινό παρανομαστή και να λυθούν με μεθόδους SVD ή και τον αλγόριθμο DLT που θα περιγράψουμε επόμενα.

2.1.5 DLT και RANSAC: Εφαρμογή στον προσδιορισμό της μήτρας προβολής

Ο μετασχηματισμός DLT

Ο μετασχηματισμός DLT (Direct Linear Transform) μπορεί να χρησιμοποιηθεί για την επίλυση διαφόρων ειδών προβλημάτων γεωμετρίας πολλαπλών όψεων, αφού η εφαρμογή του βασίζεται στην μορφή των εξισώσεων προβολής οι οποίες έχουν παρόμοια μορφή για διάφορα είδη προβλημάτων (π.χ. ομογραφίες μεταξύ επιπέδων, προβολές από τον τρισδιάστατο χώρο σε επίπεδα, προσδιορισμός θεμελιώδους μήτρας). Οι άγνωστοι του προβλήματος (π.χ. παράμετροι προβολής, σημεία επιπέδου, σημεία χώρου) μπορούν να μεταβάλλονται χωρίς την παρεμπόδιση της εφαρμογής της τεχνικής. Διαλέγουμε λοιπόν να παρουσιάσουμε την εφαρμογή του αλγορίθμου DLT στο πρόβλημα του προσδιορισμού της μήτρας προβολής P μιας κάμερας δεδομένων αντιστοιχιών σημείων της εικόνας με σημεία του τρισδιάστατου χώρου $\mathbf{x}_j \leftrightarrow \mathbf{X}_j$. Δεν υποθέτουμε ιδιαίτερη μορφή του πίνακα P για απλότητα, η ενσωμάτωση όμως τέτοιων περιορισμών είναι εύκολη.

Ξεκινώντας από τις εξισώσεις προβολής $\mathbf{x}_j = P\mathbf{X}_j$ σημειώνουμε ότι έχουν ακρίβεια μιας πολλαπλασιαστικής σταθεράς, εξαιτίας της χρήσης ομογενών συντεταγμένων που καθιστούν τα σημεία $\mathbf{x}_j, a\mathbf{x}_j$ ισοδύναμα. Για αυτό γράφουμε τις εξισώσεις προβολής ως

$$\mathbf{x}_j \times P\mathbf{X}_j = \mathbf{0} \quad (2.15)$$

Αν συμβολίσουμε τον πίνακα

$$P = \begin{bmatrix} \mathbf{p}^{1T} \\ \mathbf{p}^{2T} \\ \mathbf{p}^{3T} \end{bmatrix}$$

αντικαταστήσουμε στην 2.15 και εκτελέσουμε τις πράξεις καταλήγουμε στην σχέση:

$$\begin{bmatrix} \mathbf{0}^T & -w_j \mathbf{X}_j^T & y_j \mathbf{X}_j^T \\ w_j \mathbf{X}_j^T & \mathbf{0}^T & -x_j \mathbf{X}_j^T \\ -y_j \mathbf{X}_j^T & x_j \mathbf{X}_j^T & \mathbf{0}^T \end{bmatrix} \begin{bmatrix} \mathbf{p}^1 \\ \mathbf{p}^2 \\ \mathbf{p}^3 \end{bmatrix} = \mathbf{0} \quad (2.16)$$

Στην Εξίσωση 2.16 παρατηρούμε:

- Η εξίσωση είναι γραμμική ως προς τα άγνωστα στοιχεία της μήτρας προβολής P .
- Η εξίσωση είναι τετραγωνική ως προς τις συντεταγμένες των γνωστών σημείων $\mathbf{x}_j, \mathbf{X}_j$.

- Η μήτρα έχει βαθμό δύο. Αυτό προκύπτει δι επισκοπήσεως, όσο και από την εξαγωγή της με χρήση του τελεστή εξωτερικού γινομένου.

Επομένως αρκούν μόνο δύο από τις εξισώσεις της Σχέσης 2.16 και μπορούμε να χρησιμοποιούμε στην θέση του το νέο σύστημα:

$$\begin{bmatrix} \mathbf{0}^T & -w_j \mathbf{X}_j^T & y_j \mathbf{X}_j^T \\ w_j \mathbf{X}_j^T & \mathbf{0}^T & -x_j \mathbf{X}_j^T \end{bmatrix} \begin{bmatrix} \mathbf{p}^1 \\ \mathbf{p}^2 \\ \mathbf{p}^3 \end{bmatrix} = \mathbf{0} \quad (2.17)$$

Στην περίπτωση που θεωρούμε τις μετρήσεις μας για τα σημεία ακριβείς, χρειαζόμαστε 11 εξισώσεις για να λύσουμε ως προς τα άγνωστα στοιχεία του P . Επομένως με 5 σημεία και μία συντεταγμένη (x ή y) ενός έκτου σημείου η εξίσωση μπορεί να λυθεί και να προσδιοριστεί πλήρως ο πίνακας P . Είναι όμως σύνηθες, να έχουμε περισσότερες αντιστοιχίες σημείων και να θεωρούμε τις μετρήσεις των σημείων θορυβώδεις. Σε αυτή την περίπτωση θα έχουμε ένα υπερορισμένο σύστημα και οι εξισώσεις δεν θα ικανοποιούνται ακριβώς. Για να λύσουμε τότε το πρόβλημα διατυπώνουμε το πρόβλημα ελαχιστοποίησης

$$\begin{aligned} \min \mathbf{A} \mathbf{p} \\ \text{s.t. } \|\mathbf{p}\| = 1 \end{aligned} \quad (2.18)$$

Στην προηγούμενη Εξίσωση 2.18 χρησιμοποιήσαμε τον συμβολισμό

$$\mathbf{p} = \begin{bmatrix} \mathbf{p}^1 \\ \mathbf{p}^2 \\ \mathbf{p}^3 \end{bmatrix}$$

και ο πίνακας A προκύπτει από την ένωση των μητρών του αριστερού μέλους της 2.17 για όλες τις διαθέσιμες μετρήσεις σημείων, δηλαδή εκφράζει το σύνολο των εξισώσεων του προβλήματος. Το σφάλμα που ελαχιστοποιούμε στην 2.18 είναι γνωστό ως *αλγεβρικό σφάλμα*.

Η λύση του προβλήματος 2.18 έχει απλή μορφή και μπορεί να δειχθεί ότι είναι το ιδιοδιάνυσμα που αντιστοιχεί στην ελάχιστη ιδιοτιμή του πίνακα (αποσύνθεση SVD).

Αναφέρουμε ακόμα ότι θα πρέπει να *κανονικοποιούμε* τα δεδομένα του προβλήματος πριν εφαρμόσουμε την τεχνική του DLT. Προηγούμενες αναφορές στην ισοτροπική κανονικοποίηση καλύπτουν το θέμα και εδώ (Ενότητα 2.1.2)

Αν θέλουμε να ελαχιστοποιήσουμε διαφορετικό σφάλμα, για παράδειγμα το *σφάλμα επαναπροβολής* προτείνεται να χρησιμοποιήσουμε τον μετασχηματισμό DLT και να επιλύσουμε το γραμμικό Σύστημα 2.18. Η λύση που λαμβάνουμε στην συνέχεια μπορεί να χρησιμοποιηθεί για την αρχικοποίηση επαναληπτικής μεθόδου βελτιστοποίησης, όπως για παράδειγμα είναι ο αλγόριθμος LM που περιγράφεται στην Ενότητα 2.1.6.

Η μέθοδος RANSAC

Η μέθοδος RANSAC (**R**andom **S**ampling **C**onsensus) είναι μια γενική μέθοδος που σκοπό έχει την ρωμαλέα εκτίμηση λύσεων και τον διαχωρισμό των δεδομένων σε αναξιόπιστα και αξιόπιστα. Εισήχθη για πρώτη φορά από

τους [57] για να αντιμετωπίσουν το πρόβλημα της εκτίμησης των παραμέτρων της κάμερας (πόζας), δεδομένων αντιστοιχιών μεταξύ σημείων του χώρου και της εικόνας. Η μέθοδος όμως είναι γενική και μπορεί να χρησιμοποιηθεί ανεξάρτητα από την μορφή του προβλήματος.

Ας εξετάσουμε τώρα την μέθοδο RANSAC στα πλαίσια του προβλήματος που μελετάμε σε αυτή την ενότητα. Από τις αντιστοιχίες σημείων που έχουμε κάνει, θεωρούμε ότι μερικές είναι πραγματικές αντιστοιχίες ενώ άλλες είναι λάθος ταιριάσματα. Εφαρμόζοντας την μέθοδο RANSAC θα πάρουμε μια *εκτίμηση της λύσης* καθώς και ένα διαχωρισμό των αντιστοιχιών σε αυτές που ταιριάζουν με αυτήν, επομένως θεωρούνται *πραγματικές*, και σε αυτές που δεν ταιριάζουν με την λύση μας, θεωρούνται *εσφαλμένες* και θα πρέπει να τις αφαιρέσουμε από τα δεδομένα μας.

Για να κάνουμε την σθεναρή αυτή εκτίμηση, διαλέγουμε ένα δείγμα αντιστοιχιών μεγέθους N από τα δεδομένα μας. Το μέγεθος του δείγματος πρέπει να είναι αρκετά μεγάλο ώστε να μπορούμε να επιλύσουμε το σύστημα του προβλήματος. Θυμίζουμε ότι το σύστημα θεωρείται υπερδιορισμένο. Για παράδειγμα για να λύσουμε το πρόβλημα 2.18 χρειαζόμαστε 5.5 σημεία τουλάχιστον (11 εξισώσεις). Είναι σημαντικό να κρατήσουμε το μέγεθος του δείγματος N όσο το δυνατόν μικρότερο. Αυτό γιατί όπως θα δούμε σύντομα, ο αριθμός των επαναλήψεων που απαιτείται για την επίτευξη ενός συγκεκριμένου *βαθμού εμπιστοσύνης* αυξάνεται με αλματώδη ρυθμό συναρτήσει του μεγέθους του δείγματος.

Για κάθε δείγμα που διαλέγουμε *τυχαία*, υπολογίζουμε μια λύση του προβλήματος που προκύπτει αποκλειστικά από τα δεδομένα του δείγματος. Στην συνέχεια χρησιμοποιούμε κάποια μετρική για να υπολογίσουμε τα σημεία που δικαιολογούνται από την προτεινόμενη λύση και τα σημεία που θεωρούμε εσφαλμένες αντιστοιχίες και δεν εξηγούνται ικανοποιητικά από την προταθείσα λύση. Για παράδειγμα, σε προβλήματα γεωμετρίας πολλαπλών όψεων μπορούμε να χρησιμοποιήσουμε το σφάλμα επαναπροβολής για να αξιολογήσουμε το ταίριασμα των δεδομένων και της λύσης.

Αφού επαναλάβουμε την προηγούμενη διαδικασία για ικανοποιητικό, σύμφωνα με στατιστική ανάλυση, αριθμό δειγμάτων επιλέγουμε την λύση του προβλήματος που *μεγιστοποιεί* τον αριθμό των σημείων που μπορεί να *δικαιολογήσει*. Για να υπολογίσουμε ένα κάτω φράγμα των απαιτούμενων επαναλήψεων του αλγορίθμου, υπολογίζουμε την πιθανότητα να υπάρχει τουλάχιστον ένα δείγμα το οποίο να *απαρτίζεται αποκλειστικά* από *έγκυρες* αντιστοιχίες. Η πιθανότητα αυτή υπολογίζεται

$$\begin{aligned}
 1 - P_s &= (1 - p^n)^S, && \text{πιθανότητα αποτυχίας μεθόδου} \\
 \iff S &= \frac{\log(1 - P_s)}{\log(1 - p^n)}, && \text{απαιτούμενος αριθμός επαναλήψεων} \quad (2.19) \\
 P_s &= && \text{πιθανότητα ύπαρξης τουλάχιστον ενός} \\
 &&& \text{δείγματος που απαρτίζεται αποκλειστικά} \\
 &&& \text{από έγκυρες αντιστοιχίες} \\
 p &= && \text{πιθανότητα έγκυρης αντιστοιχίας} \\
 n &= && \text{μέγεθος δείγματος}
 \end{aligned}$$

Επιβεβαιώνεται και από την Εξίσωση 2.19 η ευαισθησία του αριθμού των επαναλήψεων στο μέγεθος του δείγματος.

Σε πολλά προβλήματα γεωμετρίας πολλαπλών όψεων, έχουμε να ελαχιστοποιήσουμε ένα τετραγωνικό κριτήριο σε περιβάλλον γκαουσιανού θορύβου.

Παρά την θεωρητική βελτιστότητα της λύσης ελαχίστων τετραγώνων, η συνάρτηση κόστους επηρεάζεται δραματικά από εσφαλμένες αντιστοιχίες σημείων στα δεδομένα εισόδου, γενικότερα από εσφαλμένες μετρήσεις που δεν σχετίζονται με τα υπόλοιπα δεδομένα. Για την αντιμετώπιση του προβλήματος αυτού, χρησιμοποιούμε σθεναρές μεθόδους ελαχίστων τετραγώνων. Συνηθισμένη είναι επίσης η εφαρμογή μιας μη γραμμικής σθεναρής συνάρτησης ρ στο σφάλμα e και η ελαχιστοποίηση των τετραγώνων $\rho^2(e)$. Όμως συχνά μια τέτοια προσέγγιση δεν είναι αρκετή και απαιτείται η απομάκρυνση των εσφαλμένων αντιστοιχιών από τα δεδομένα μας, με χρήση του αλγορίθμου RANSAC.

Για παραλλαγές και εξελίξεις της βασικής μεθόδου, αναφέρουμε τις εργασίες των [162] που είναι και εφαρμοσμένη στο πρόβλημα του Structure From Motion και την εργασία των [35].

2.1.6 Η βασική τεχνική του προβλήματος: Bundle Adjustment

Προηγούμενα περιγράψαμε τεχνικές οι οποίες μπορούν να χρησιμοποιηθούν είτε για την επίλυση κάποιου είδους (συγκεκριμένος αριθμός όψεων, γνωστές παράμετροι κάμερας κ.α.) SfM είτε για την επίλυση κάποιου παρεμφερούς προβλήματος (διαφορετικοί γνωστοί και άγνωστοι). Όταν αντιμετωπίζουμε το πρόβλημα SfM στην γενική του μορφή, δηλαδή με αυθαίρετο αριθμό όψεων και χωρίς πληροφορίες για την παραμετροποίηση των καμερών, χρησιμοποιούμε τις ανωτέρω τεχνικές *επικουρικά* στην εφαρμογή του B.A. (Bundle Adjustment), όπως θα αποσαφηνίσουμε σύντομα.

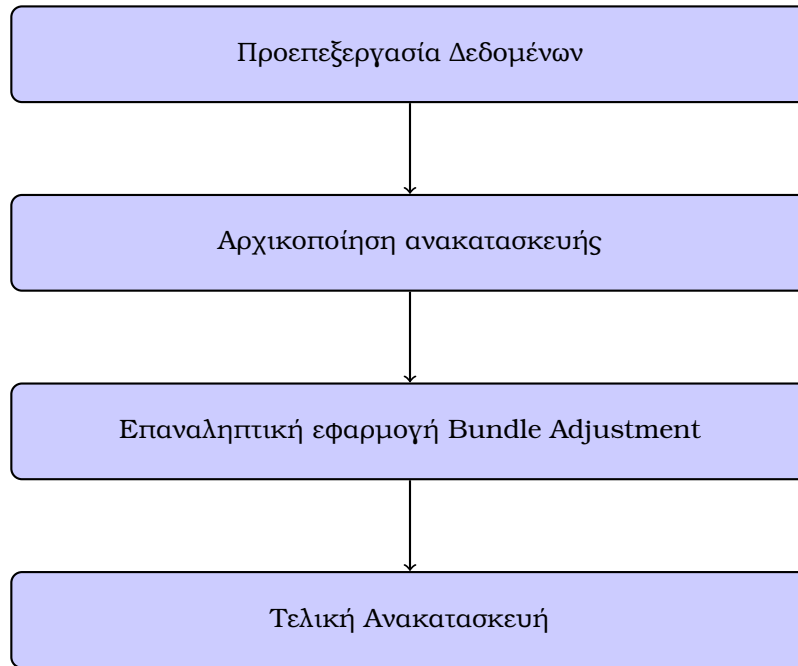
Η τεχνική του B.A. είναι μία μέθοδος που ελαχιστοποιεί το σφάλμα επαναπροβολής του προβλήματος, ή όποια (σθεναρή) αντικειμενική συνάρτηση έχουμε επιλέξει για το πρόβλημα του SfM. Επειδή είναι μια επαναληπτική, αριθμητική μέθοδος είναι σημαντικό να αρχικοποιηθεί καλά. Όπως προαναφέραμε ένας τρόπος αρχικοποίησης είναι η τεχνική της προοπτικής παραγοντοποίησης, υπάρχουν όμως και εναλλακτικές μέθοδοι, οι οποίες μπορεί να βασίζονται σε επιπλέον πληροφορίες που διαθέτουμε για το πρόβλημα [186, 9].

Το όνομα "Bundle Adjustment", αναφέρεται στην προσαρμογή (adjustment) δεσμών φωτός που πηγάζουν από τα τρισδιάστατα σημεία και συγκλίνουν στα οπτικά κέντρα των καμερών, με σκοπό την ελαχιστοποίηση κάποιου κριτηρίου σφάλματος ανακατασκευής.

Η ευελιξία της μεθόδου είναι ένα ακόμα σημαντικό πλεονέκτημα της. Δεν εξαρτάται από την συγκεκριμένη μορφή της συνάρτησης σφάλματος, είναι ανεκτική σε απώλεια δεδομένων (για παράδειγμα αν δεν είναι όλα τα σημεία

| n | p | S |
|----------|----------|----------|
| 3 | 0.5 | 35 |
| 6 | 0.6 | 97 |
| 6 | 0.5 | 292 |

Πίνακας 2.1: Απαιτούμενος αριθμός επαναλήψεων RANSAC για 0.01 πιθανότητα αποτυχίας. Από [212].



Σχήμα 2.1: Συγκεντρωτική μορφή ενός συστήματος Structure from Motion

ορατά από όλες τις κάμερες), κάθε μέτρηση μπορεί ανεξάρτητα από τις υπόλοιπες να χαρακτηρίζεται από κάποια συνάρτηση πυκνότητας πιθανότητας. Ακόμα η μέθοδος μπορεί να επεκταθεί για να ενσωματώσει περιορισμούς, για παράδειγμα στην θέση των σημείων ή στις παραμέτρους των καμερών. Πέρα από την ανάγκη της καλής αρχικοποίησης, η υπολογιστική πολυπλοκότητα είναι ένα άλλο σημαντικό πρόβλημα της μεθόδου. Καταρχάς, ας υπολογίσουμε το μέγεθος του προβλήματος. Έστω ότι έχουμε n σημεία και m όψεις. Τότε έχουμε τουλάχιστον 11 παραμέτρους για κάθε όψη και 3 για κάθε σημείο, συνολικά επομένως τουλάχιστον $11m + 3n$ παραμέτρους. Αν χρησιμοποιήσουμε την γνωστή αριθμητική μέθοδο Levenberg-Marquardt τότε πρέπει να αντιστρέψουμε τετραγωνικούς πίνακες διάστασης $11m + 3n \times 11m + 3n$. Με την προσέγγιση $n \gg m$, η οποία πολύ εύκολα ικανοποιείται, η λύση του προβλήματος ελαχιστοποίησης του σφάλματος επαναπροβολής, το οποίο είναι ένα μη γραμμικό πρόβλημα ελαχίστων τετραγώνων, έχει τάξη πολυπλοκότητας $\mathcal{O}(n^3)$ με $\mathcal{O}(n^4)$. Πρέπει λοιπόν να χρησιμοποιήσουμε μεθόδους που αξιοποιούν την ειδική μορφή του προβλήματος και όχι γενικές αριθμητικές μεθόδους.

Παρουσιάζουμε συνοπτικά στρατηγικές που μπορούν να ακολουθηθούν για την υπολογιστική απλοποίηση του προβλήματος

- Περιορισμός των αγνώστων. Για παράδειγμα μπορεί να μην συμπεριληφθούν όλες οι όψεις ή όλα τα σημεία και στην συνέχεια να προστεθούν στο μοντέλο με τεχνικές resectioning και τριγωνισμού αντίστοιχα. Εναλλακτικά μπορούμε να χωρίσουμε τα δεδομένα εισόδου σε σύνολα, να λύσουμε το πρόβλημα για κάθε σύνολο ξεχωριστά και να τα ενώσουμε στο τέλος. Η λύση που λαμβάνουμε στις περιπτώσεις αυτές είναι υποβέλτιστη.
- Εναλλαγή του τρόπου μείωσης του σφάλματος επαναπροβολής. Εναλλάσσουμε μεταξύ ρύθμισης των παραμέτρων προβολής για τις κάμερες

και της ρύθμισης των τρισδιάστατων σημείων. Η μέθοδος αυτή όμως απαιτεί περισσότερες επαναλήψεις έως την σύγκλιση. Η συνάρτηση που ελαχιστοποιείται όμως είναι ταυτόσημη με την αρχική, οπότε λαμβάνουμε βέλτιστη λύση [233].

- Μέθοδοι που εκμεταλλεύονται την αραιή δομή του προβλήματος. Παραδείγματα είναι η υλοποίηση στο [130], η χρήση αραιών μεθόδων ελαχίστων τετραγώνων (αραιή Παραγοντοποίηση Cholesky [129, 34], μέθοδοι Conjugate Gradient [100, 4]).
- Χωρισμός του προβλήματος σε υποχάρτες (υποπεριοχές) και εφαρμογή του Β.Α. σε κάθε υποχάρτη ξεχωριστά. Οι υποχάρτες διαλέγονται από ελάχιστες τομές σε γράφο που εκφράζει τις αλληλοεξαρτήσεις των παραμέτρων του προβλήματος ώστε να εξασφαλιστεί μέγιστη ανεξαρτησία μεταξύ τους [160].
- Εφαρμογή του Β.Α. στον σκελετικό γράφο του προβλήματος, όπως στο [210].
- Προσπάθειες για λύση του προβλήματος σε μία επανάληψη. Για παράδειγμα μπορούμε να χρησιμοποιήσουμε την L_∞ νόρμα ή να λύσουμε ανεξάρτητα για κάποιες ομάδες παραμέτρων του προβλήματος [79, 87, 141, 202].

Προβλήματα βελτιστοποίησης και αριθμητικές μέθοδοι επίλυσης τους

Η μέθοδος του Bundle Adjustment λύνει ένα πρόβλημα ελαχιστοποίησης και συγκεκριμένα ένα πρόβλημα μη γραμμικών ελαχίστων τετραγώνων. Για αυτό ξεκινάμε από μια γενικότερη περιγραφή και εξειδικεύουμε σταδιακά στην περίπτωση του Bundle Adjustment προβλήματος.

Έστω ότι διαθέτουμε μια συνάρτηση προς ελαχιστοποίηση $c(\mathbf{x})$, όπου \mathbf{x} το διάνυσμα των παραμέτρων του προβλήματος. Αν αναπτύξουμε την c σε δευτεροβάθμιο ανάπτυγμα Taylor γύρω από το σημείο \mathbf{x} έχουμε την σχέση

$$c(\mathbf{x} + \mathbf{dx}) \approx c(\mathbf{x}) + \nabla c(\mathbf{x})^T \mathbf{dx} + 1/2 \mathbf{dx}^T H_c(\mathbf{x}) \mathbf{dx} \quad (2.20)$$

Θυμίζουμε τους ορισμούς της κλίσης

$$\nabla c = \left[\frac{\partial c}{\partial x_1} \quad \cdots \quad \frac{\partial c}{\partial x_n} \right]^T$$

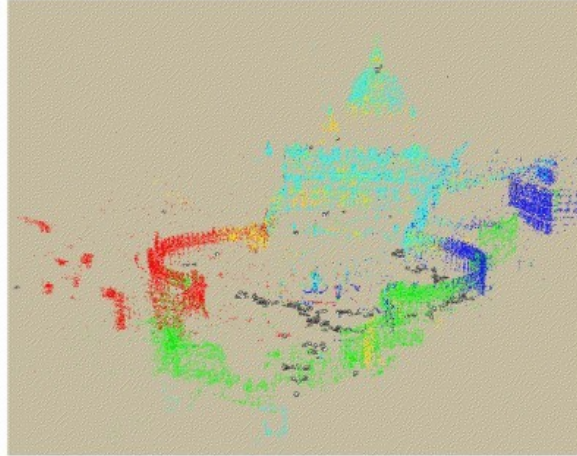
και της μήτρας δευτέρων παραγώγων Hessian

$$\begin{bmatrix} \frac{\partial^2 c}{\partial x_1 \partial x_1} & \cdots & \frac{\partial^2 c}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 c}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 c}{\partial x_n \partial x_n} \end{bmatrix}$$

Βρίσκουμε το ελάχιστο της έκφρασης 2.20 μηδενίζοντας την παράγωγο της ως προς το διάνυσμα \mathbf{dx} και καταλήγουμε στο γραμμικό σύστημα

$$H_c \mathbf{dx} = -\nabla c(\mathbf{x}) \quad (2.21)$$

Η μέθοδος αυτή ανανέωσης του \mathbf{x} από το $\mathbf{x} \leftarrow \mathbf{x} + \mathbf{dx}$ που προκύπτει ως λύση του 2.21 δεν παρέχει εγγυήσεις ότι θα βελτιώσει την τιμή της αντικειμενικής



Σχήμα 2.2: Χωρισμός προβλήματος SfM σε υποχάρτες. Ο διαχωρισμός φαίνεται με τα διαφορετικά χρώματα. Από [160]

συνάρτησης c . Για αυτό ακολουθούμε την εξής στρατηγική: αν ανανέωση του \mathbf{x} δεν μειώνει την τιμή της αντικειμενικής συνάρτησης c , δεν την εφαρμόζουμε και κάνουμε ένα βήμα στην κατεύθυνση της αρνητικής κλίσης $-\nabla c(\mathbf{x})$. Μικρά βήματα στην κατεύθυνση $-\nabla c(\mathbf{x})$ μειώνουν με βεβαιότητα την τιμή της c . Διαθέτουμε δηλαδή δύο στρατηγικές ανανέωσης του \mathbf{x} , με χρήση της λύσης της 2.21 και ακολουθώντας την κατεύθυνση της αρνητικής κλίσης, τις οποίες εναλλάσσουμε κατά την διαδικασία ελαχιστοποίησης της c .

Εξετάζουμε τώρα την περίπτωση στην οποία η προς ελαχιστοποίηση συνάρτηση c , είναι άθροισμα τετραγώνων συναρτήσεων, δηλαδή μπορούμε να γράψουμε

$$c(\mathbf{x}) = \mathbf{f}^T(\mathbf{x})\mathbf{f}(\mathbf{x})$$

Η διανυσματική συνάρτηση \mathbf{f} είναι η συνάρτηση σφάλματος του προβλήματος. Το πρόβλημα ελαχιστοποίησης που σχηματίσαμε τώρα γράφεται ισοδύναμα

$$\min_x \frac{1}{2} \|\mathbf{f}(\mathbf{x})\|^2 \quad (2.22)$$

Αναπτύσσουμε την συνάρτηση $\mathbf{f}(\mathbf{x})$ σε ανάπτυγμα Taylor πρώτου βαθμού και έχουμε $\mathbf{f}(\mathbf{x} + \mathbf{dx}) \approx \mathbf{f}(\mathbf{x}) + J(\mathbf{x})\mathbf{dx}$. Με $J(\mathbf{x})$ συμβολίζουμε την ιακωδιανή της \mathbf{f} στο σημείο \mathbf{x} και υπενθυμίζουμε τον ορισμό της

$$\begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{bmatrix}$$

Αντικαθιστούμε τώρα το ανάπτυγμα Taylor της \mathbf{f} στην c , εξισώνουμε την παράγωγο της με 0 και καταλήγουμε στην εξίσωση ενημέρωσης του \mathbf{x}

$$J^T(\mathbf{x})J(\mathbf{x})\mathbf{dx} = -J(\mathbf{x})\mathbf{f}(\mathbf{x}) \quad (2.23)$$

Η εξίσωση αυτή είναι αντίστοιχη με την 2.21. Παρατηρούμε ότι $\nabla c(\mathbf{x}) = 2J^T(\mathbf{x})\mathbf{f}(\mathbf{x})$ και ότι προσεγγίζουμε την Hessian ως $H_c \approx 2J^T(\mathbf{x})J(\mathbf{x})$. Παρατηρούμε ότι δεν είναι απαραίτητο πλέον να υπολογίσουμε δεύτερες παραγώγους για να σχηματίσουμε τον πίνακα Hessian. Αν στην Εξίσωση 2.23 πολλαπλασιάσουμε την διαγώνιο του αριστερού μέρους με $1 + \lambda$, καταλήγουμε στην μέθοδο Levenberg-Marquardt. Ελέγχουμε την παράμετρο λ με

προσαρμοστικό τρόπο: Αν η Εξίσωση 2.23 πετυχαίνει μείωση στην τιμή της c μειώνουμε την τιμή του λ , σε αντίθετη περίπτωση την αυξάνουμε. Ξαναγράφουμε την Εξίσωση 2.23 μετά την τροποποίηση του αριστερού μέρους της σύμφωνα με τον αλγόριθμο LM ως

$$N(\mathbf{x})d\mathbf{x} = -J(\mathbf{x})\mathbf{f}(\mathbf{x}) \quad (2.24)$$

όπου ο πίνακας N είναι ίδιος με τον $J^T J$ στα εκτός διαγωνίου στοιχεία. Για την διαγώνιο ισχύει $N_{ii} = (J^T J)_{ii}(1 + \lambda)$. Για μεγάλες τιμές του λ ο πίνακας είναι διαγώνιος και ακολουθούμε την κατεύθυνση της αρνητικής κλίσης. Η σύγκλιση είναι αργή, αλλά εγγυημένη. Αυτό θα συμβεί όταν η πρωτοβάθμια προσέγγιση της συνάρτησης λάθους διαφέρει σημαντικά από την πραγματική μορφή της συνάρτησης. Ταυτόχρονα το πλάτος του βήματος ελαττώνεται όταν το λ μεγαλώνει. Αντίθετα, αν το γραμμικό μοντέλο είναι ικανοποιητικό θα κινηθούμε γρήγορα προς το ελάχιστο που δίνεται από το τετραγωνικό βήμα. Επίσης η προσθήκη του όρου απόσβεσης προσθέτει ευστάθεια στην περίπτωση που ο πίνακας $J^T J$ δεν είναι ομαλός, επειδή τον καθιστά αντιστρέψιμο.

Η μέθοδος LM λοιπόν, είναι μια μέθοδος ελαχιστοποίησης αθροίσματος τετραγώνων μη-γραμμικών συναρτήσεων η οποία συνδυάζει τις μεθόδους απότομης κατάβασης (steepest descent) και την μέθοδο Gauss-Newton.

Μια παραλλαγή του αλγορίθμου LM που χρησιμοποιείται στο [51], διαιρεί την προτεινόμενη εξέλιξη $d\mathbf{x}$ του διανύσματος \mathbf{x} με λ σε περίπτωση που αυτή αποτυγχάνει να μειώσει το κόστος c . Δηλαδή σε περίπτωση αποτυχίας εξελίσσουμε το \mathbf{x} με την προσθήκη του $d\mathbf{x} = -1/\lambda(J^T(\mathbf{x})J(\mathbf{x}))^{-1}J^T(\mathbf{x})\mathbf{f}(\mathbf{x})$. Το διάλυμα $d\mathbf{x}$ του προηγούμενου κανόνα είναι το γινόμενο ενός μικρού θετικά ορισμένου πίνακα με την αρνητική κλίση της συνάρτησης κόστους, συνεπώς προκαλεί μείωση του κόστους. Το όφελος από την χρήση αυτής της μεθόδου είναι ότι δεν ξαναλύνουμε το γραμμικό Σύστημα 2.23 όταν αυξάνουμε το λ . Αν η συνάρτηση κόστους c υπολογίζεται εύκολα και η λύση του συστήματος είναι χρονοβόρα, τα οφέλη είναι σημαντικά.

Ο αλγόριθμος LM μπορεί να διατυπωθεί και από διαφορετική οπτική, όπως θα κάνουμε τώρα. Για να λύσουμε το αρχικό πρόβλημα, ακολουθούμε μια γενικευμένη στρατηγική, και λύνουμε μια ακολουθία προσεγγίσεων του αρχικού προβλήματος. Αν αντικαταστήσουμε την προσέγγιση κατά Taylor της f στην 2.22 παίρνουμε την ακολουθία προβλημάτων

$$\min_{dx} 1/2 \|J(\mathbf{x})d\mathbf{x} + \mathbf{f}(\mathbf{x})\|^2 \quad (2.25)$$

Η ελαχιστοποίηση του 2.25 γίνεται μηδενίζοντας της πρώτη παράγωγο του ως προς $d\mathbf{x}$ και οδηγεί στην αντίστοιχη της Εξίσωσης 2.21 για το νέο πρόβλημα. Συνεπώς, όπως αναλύσαμε προηγουμένως, η ακολουθία αυτών των προβλημάτων δεν συγκλίνει πάντοτε σε λύση. Η λύση είναι να ακολουθούμε όποτε είναι απαραίτητο την κατεύθυνση της αρνητικής κλίσης, ακριβώς όπως το παρουσιάσαμε προηγουμένως. Για αυτό ορίζουμε νέα ακολουθία προς επίλυση προβλημάτων που αντιστοιχεί σε αυτή την στρατηγική. Ορίζουμε τα προβλήματα ως

$$\min_{dx} 1/2 \|J(\mathbf{x})d\mathbf{x} + \mathbf{f}(\mathbf{x})\|^2 + \lambda \|D(\mathbf{x})d\mathbf{x}\|^2 \quad (2.26)$$

Ο πίνακας D είναι ένας διαγώνιος πίνακας με στοιχεία μη-αρνητικά. Συνήθως εκλέγεται ώστε να ικανοποιεί την σχέση $D^2 = \text{diag}(J^T J)$.

Η συνάρτηση f που εισήχθη προηγουμένως, είναι στην περίπτωση του Bundle Adjustment η συνάρτηση του σφάλματος επαναπροβολής. Για να βελτιώσουμε την ευστάθεια του συστήματος ανακατασκευής της τρισδιάστατης γεωμετρίας, χρησιμοποιούμε συνήθως μια σθεναρή συνάρτηση σφάλματος που δεν επηρεάζεται πολύ από απομακρυσμένα σημεία (outliers). Για να συμβαίνει αυτό θα πρέπει η συνάρτηση f να κάνει μια μη γραμμική απεικόνιση η οποία θα περιορίζει τα μεγάλα σφάλματα, έτσι ώστε να μην επικρατούν αυτά στην συνάρτηση c . Για παράδειγμα στο [51] οι συγγραφείς χρησιμοποιούν την κατανομή Cauchy για το σφάλμα επαναπροβολής e . Σφάλμα e συνεισφέρει

$$\ln \left(1 + \frac{e^2}{\sigma^2} \right) \quad (2.27)$$

στην συνάρτηση σφάλματος c . Πιο αναλυτικά, αν (x, y) είναι το σφάλμα επαναπροβολής πριν την σθεναροποίηση του, και $e = (x_r, y_r)$ η εικόνα του αρχικού σφάλματος μετά την εφαρμογή της σθεναρής απεικόνισης έχουμε:

$$x_r = \sqrt{\ln \left(1 + \frac{x^2 + y^2}{\sigma^2} \right)} \frac{x}{\sqrt{x^2 + y^2}} \quad (2.28)$$

$$y_r = \sqrt{\ln \left(1 + \frac{x^2 + y^2}{\sigma^2} \right)} \frac{y}{\sqrt{x^2 + y^2}} \quad (2.29)$$

Επομένως όταν υψώσουμε στο τετράγωνο και αθροίσουμε τις συντεταγμένες του σθεναρού σφάλματος $e = (x_r, y_r)$ παίρνουμε την επιθυμητή συνεισφορά στην συνάρτηση κόστους, σύμφωνα με την Εξίσωση 2.27.

Για να λύσουμε το πρόβλημα που διατυπώσαμε, ο περισσότερος χρόνος υπολογισμού αναλώνεται στην λύση του μη γραμμικού προβλήματος ελαχίστων τετραγώνων 2.26. Θα μπορούσαμε να λύσουμε το πρόβλημα αυτό με γενικούς αριθμητικούς αλγόριθμους, όμως κάτι τέτοιο δεν οδηγεί σε αποδοτική αντιμετώπιση του. Για αυτό θα αναλύσουμε την ειδική δομή του προβλήματος B.A., προκειμένου να την εκμεταλλευτούμε κατάλληλα και να καταστεί δυνατή η λύση μεγάλων προβλημάτων SfM.

Πρωτεύουσα και δευτερεύουσα αραιή δομή του προβλήματος Bundle Adjustment.

Αρχικά, πρέπει να αποσαφηνίσουμε τα περιεχόμενα του διανύσματος \mathbf{x} των παραμέτρων του προβλήματος. Οι παράμετροι χωρίζονται σε δύο κατηγορίες

- Παράμετροι για μοντελοποίηση των όψεων/καμερών \mathbf{y}_i . Θεωρούμε ότι οι παράμετροι αυτές είναι p το πλήθος και ότι κάθε \mathbf{y}_i έχει διάσταση c . Λογικές τιμές του c ανάλογα με την πολυπλοκότητα του προς προσδιορισμό μοντέλου προβολής είναι 6 – 9.
- Παράμετροι για μοντελοποίηση των σημείων \mathbf{z}_i . Θεωρούμε ότι οι παράμετροι αυτές είναι q το πλήθος και ότι κάθε \mathbf{z}_i έχει διάσταση s . Λογικές τιμές του s είναι 3 – 4.

Η πρώτη ενδιαφέρουσα παρατήρηση έχει να κάνει με τον τρόπο που διαπλέκονται οι δύο αυτές κατηγορίες παραμέτρων. Οι παράμετροι για τα χαρακτηριστικά της σκηνής \mathbf{z}_i συνδυάζονται με τις παραμέτρους για την περιγραφή των καμερών \mathbf{y}_i για να "προβλέψουν" τις μετρήσεις των εικόνων. Όμως οι μεταβλητές \mathbf{z}_i δεν συνδυάζονται άμεσα με άλλες μεταβλητές \mathbf{z}_i και αντίστοιχα

για τις μεταβλητές \mathbf{y}_i . Δηλαδή οι μεταβλητές \mathbf{y}_i και οι μεταβλητές \mathbf{z}_i είναι ανεξάρτητες μεταξύ τους. Για αυτό ο γράφος του προβλήματος είναι διμερής, γεγονός που πρέπει να εκμεταλλευτούμε κατά την επίλυση του προβλήματος. Σε αυτή την ιδιότητα αναφερόμαστε και ως *πρωτεύουσα αραιή δομή* του προβλήματος Bundle Adjustment. Για να μπορέσουμε να εκμεταλλευτούμε την αραιή δομή σε αυτό το επίπεδο, είναι απαραίτητο να διατάξουμε τις παραμέτρους του προβλήματος είτε με σειρά "κάμερα και μετά σημεία", είτε με την ανάποδη. Πάντως πρέπει να αποφύγουμε την ανάμειξη των δύο αυτών ομάδων παραμέτρων με τυχαίο τρόπο.

Ακόμα, συνήθως η συνάρτηση f έχει τέτοια δομή ώστε κάθε στοιχείο της f_i να είναι συνάρτηση ενός μόνο (y_i, z_i) , δηλαδή να έχουμε $f_i(\mathbf{x}) = f(\mathbf{y}_i, \mathbf{z}_i)$. Επιλέγουμε στα επόμενα να διατάξουμε τις παραμέτρους των όψεων πριν από αυτές των σημείων, δηλαδή $\mathbf{x} = (\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{z}_1, \mathbf{z}_2, \dots)$.

Στην συνέχεια γράφουμε την Ιακωβιανή της f ως $J = [J_C J_P]$ με J_C, J_P τις ιακωβιανές της f ως προς τις παραμέτρους των όψεων και των σημείων αντίστοιχα. Η προσέγγιση της Hessian που χρησιμοποιούμε στην Εξίσωση 2.23 θα είναι τότε

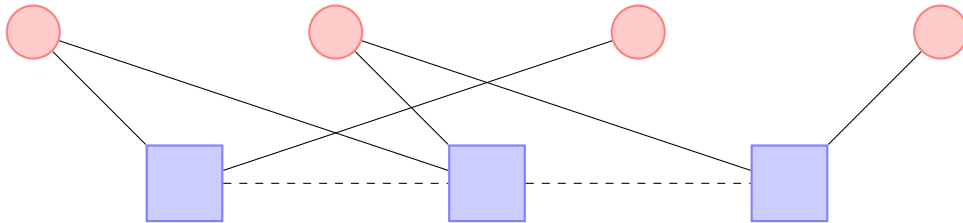
$$H = \begin{bmatrix} J_C^T J_C & J_C^T J_P \\ J_P^T J_C & J_P^T J_P \end{bmatrix} = \begin{bmatrix} H_{CC} & H_{PC}^T \\ H_{PC} & H_{PP} \end{bmatrix} \quad (2.30)$$

Ο πίνακας που χρησιμοποιείται από τον αλγόριθμο LM έχει επαυξημένη διαγώνιο, αυτό όμως δεν επηρεάζει την δομή του όπως την περιγράφουμε παρακάτω

- H_{CC} Μπλοκ διαγώνιος πίνακας διαστάσεων $pc \times pc$. Αποτελείται από p μπλοκς διαστάσεων $c \times c$.
- H_{PP} Μπλοκ διαγώνιος πίνακας διαστάσεων $qs \times qs$. Αποτελείται από q μπλοκς διαστάσεων $s \times s$.
- H_{PC} Γενικής μορφής αραιός πίνακας διαστάσεων $qs \times pc$. Αποτελείται από μπλοκς διαστάσεων $c \times s$ για κάθε παρατήρηση. Ένα μπλοκ είναι μη μηδενικό αν το σημείο που του αντιστοιχεί είναι ορατό από την κάμερα που του αντιστοιχεί. Η δομή του πίνακα αυτού δεν αξιοποιείται στα πλαίσια της πρωτεύουσας αραιής δομής.

Στην συνέχεια, υπό το πρίσμα των διαφωτιστικών αυτών παρατηρήσεων, εξετάζουμε ξανά το πρόβλημα που προέκυψε κατά τον σχηματισμό του αλγορίθμου LM 2.26 και συγκεκριμένα επαναδιατυπώνουμε την Σχέση 2.24 χρησιμοποιώντας την νέα μπλοκ μορφή της Hessian

$$\begin{bmatrix} H_{CC} & H_{PC}^T \\ H_{PC} & H_{PP} \end{bmatrix} \begin{bmatrix} d\mathbf{y} \\ d\mathbf{z} \end{bmatrix} = \begin{bmatrix} \mathbf{v} \\ \mathbf{w} \end{bmatrix}, \text{ με τον ορισμό } \begin{bmatrix} \mathbf{v} \\ \mathbf{w} \end{bmatrix} = J^T(\mathbf{y}, \mathbf{z})\mathbf{f}(\mathbf{y}, \mathbf{z}) \quad (2.31)$$



Σχήμα 2.3: Παράδειγμα γράφου για πρόβλημα SfM. Οι συμπαγείς γραμμές ενώνουν σημεία με κάμερες που βλέπουν αυτά τα σημεία. Οι διακεκομμένες γραμμές ενώνουν κάμερες που βλέπουν ένα κοινό σημείο στον χώρο.

Εφαρμόζουμε στην 2.31 μπλοκ απαλοιφή Gauss. Από την δεύτερη γραμμή του μπλοκ-διαχωρισμένου πίνακα παίρνουμε την σχέση

$$\mathbf{dz} = H_{PP}^{-1}(\mathbf{w} - H_{PC}\mathbf{dy}) \quad (2.32)$$

Σημειώνουμε ότι ο πίνακας H_{PP} είναι μπλοκ διαγώνιος πίνακας απαρτιζόμενος από q μικρά μπλοκς $s \times s$. Συνεπώς αντιστρέφεται πολύ εύκολα σε χρόνο $\mathcal{O}(q)$, ο οποίος είναι αμελητέος για οποιαδήποτε υλοποίηση λύσης στο πρόβλημα. Μπορούμε τώρα να πάρουμε την σχέση για το \mathbf{dz} (2.32) και να την αντικαταστήσουμε στο Σύστημα 2.31, καταλήγοντας σε μία σχέση ως προς το διάνυσμα \mathbf{dy} μόνο, συγκεκριμένα:

$$[H_{CC} - H_{PC}^T H_{PP}^{-1} H_{PC}] \mathbf{dy} = \mathbf{v} - H_{PC}^T H_{PP}^{-1} \mathbf{w} \quad (2.33)$$

Την ποσότητα εντός της αγκύλης $[H_{CC} - H_{PC}^T H_{PP}^{-1} H_{PC}]$ την ονομάζουμε *συμπλήρωμα του Schur* και την συμβολίζουμε S . Ένα άλλο όνομα για τον πίνακα S είναι *ελαττωμένη μήτρα κάμερας*.

Σε αυτό το σημείο μπορούμε να εισάγουμε την *δευτερεύουσα αραιή δομή* του προβλήματος. Η μήτρα S είναι με την σειρά της αραιή. Η μορφή της όμως εξαρτάται από τα ιδιαίτερα χαρακτηριστικά της διαμόρφωσης του προβλήματος και για αυτό είναι δυσκολότερο να προβλεφθεί και να αξιοποιηθεί.

Ας εξηγήσουμε όμως την δομή της S . Όπως είναι ξεκάθαρο και από μια επισκόπηση του ορισμού της, η μήτρα S είναι διαστάσεων $pc \times pc$ και ένα μπλοκ διαστάσεων $c \times c$ δύναται να είναι διάφορο του μηδενός μόνο όταν οι κάμερες που αντιστοιχούν στο μπλοκ αυτό παρατηρούν κάποιο κοινό σημείο. Σε αντίθετη περίπτωση το μπλοκ είναι μηδενικό. Είναι ξεκάθαρο ότι σε προβλήματα με μεγάλο σύνολο δεδομένων φωτογραφιών, η παρατήρηση κοινού σημείου μεταξύ δύο όψεων είναι αρκετά σπάνιο φαινόμενο και επομένως στις περιπτώσεις αυτές ο πίνακας S πρέπει να είναι έντονα αραιός.

Αφού τελειώσαμε με τις παρατηρήσεις στην δομή του προβλήματος, πρέπει να καταδείξουμε και την κατάλληλη ακολουθία επίλυσης των εξισώσεων του προβλήματος. Το κρίσιμο σημείο στις επαναλήψεις του αλγορίθμου LM όπως έχουμε επισημάνει είναι η επίλυση του συστήματος 2.31, όποτε αυτή απαιτείται. Για να λύσουμε το Σύστημα 2.31 εφαρμόζουμε το τέχνασμα του Schur (Εξισώσεις 2.33, 2.32) και επιλύουμε το σύστημα ως προς \mathbf{dy} . Αντικαθιστώντας την ευρεθείσα τιμή του \mathbf{dy} στην 2.32 ανακτάμε και την λύση ως προς \mathbf{dz} .

Τα βήματα που περιγράψαμε προηγουμένως μπορούν να εκτελεστούν με περισσότερους από έναν τρόπους. Ας τα σχολιάσουμε εκτενώς ένα-ένα.

Επίλυση συστήματος 2.31: Αντί να λύσουμε απευθείας αυτό το σύστημα, και επειδή ισχύει $p \ll q$, εφαρμόζουμε το τέχνασμα του Schur, η εφαρμογή του οποίου οδηγεί σε σημαντικά υπολογιστικά οφέλη[29]. Αυτό είναι συνέπεια της πρωτεύουσας αραιής δομής, όπως έχουμε ήδη αναλύσει. Η υπολογιστική πολυπλοκότητα της εφαρμογής απαλοιφής Gauss στο Σύστημα 2.31, δηλαδή του σχηματισμού του συμπληρώματος του Schur στη Εξίσωση 2.33, υπολογίζεται εύκολα από τις μπλοκ διαστάσεις των επιμέρους πινάκων που εμπλέκονται στην έκφραση σε $\mathcal{O}(qp^2)$

Επίλυση συστήματος 2.33: Διάφορες μέθοδοι έχουν προταθεί για την αποτελεσματική λύση του. Η εφαρμογή γενικών αριθμητικών μεθόδων είναι μια επιλογή. Η λύση του συστήματος με παραγοντοποίηση Cholesky [230] κοστίζει χρονικά $\mathcal{O}(p^3)$ και χωρικά $\mathcal{O}(p^2)$. Αποθηκεύει και χειρίζεται τον S σαν πυκνό πίνακα και αδιαφορεί για την δευτερεύουσα αραιή δομή του προβλήματος. Εναλλακτικά μπορούμε να αναδιατάξουμε γραμμές και στήλες του S

$$V = \begin{bmatrix} \sum_{j=1}^3 B_{1j}^T \Sigma_{x_{1j}}^{-1} B_{1j} & & & \\ & \sum_{j=1}^3 B_{2j}^T \Sigma_{x_{2j}}^{-1} B_{2j} & & \\ & & \sum_{j=1}^3 B_{3j}^T \Sigma_{x_{3j}}^{-1} B_{3j} & \\ & & & \sum_{j=1}^3 B_{4j}^T \Sigma_{x_{4j}}^{-1} B_{4j} \end{bmatrix}$$

$$W = \begin{bmatrix} A_{11}^T \Sigma_{x_{11}}^{-1} B_{11} & A_{21}^T \Sigma_{x_{21}}^{-1} B_{21} & A_{31}^T \Sigma_{x_{31}}^{-1} B_{31} & A_{41}^T \Sigma_{x_{41}}^{-1} B_{41} \\ A_{12}^T \Sigma_{x_{12}}^{-1} B_{12} & A_{22}^T \Sigma_{x_{22}}^{-1} B_{22} & A_{32}^T \Sigma_{x_{32}}^{-1} B_{32} & A_{42}^T \Sigma_{x_{42}}^{-1} B_{42} \\ A_{13}^T \Sigma_{x_{13}}^{-1} B_{31} & A_{23}^T \Sigma_{x_{23}}^{-1} B_{23} & A_{33}^T \Sigma_{x_{33}}^{-1} B_{33} & A_{43}^T \Sigma_{x_{43}}^{-1} B_{43} \end{bmatrix}$$

Παρατηρήστε ότι επειδή όλα τα σημεία είναι ορατά σε όλες τις όψεις το συμπλήρωμα του Schur θα προκύψει πυκνός πίνακας. Εφόσον ο πίνακας W είναι πυκνός και με τον συμβολισμό της παραγράφου $S = U - WV^{-1}W^T$, η προηγούμενη παρατήρηση είναι προφανής.

Εφαρμογή: Μια υλοποίηση πραγματικού χρόνου (real-time)

Στην συνέχεια περιγράφουμε τα χαρακτηριστικά σημεία της υλοποίησης του αλγορίθμου Bundle adjustment από τους [51]. Βασικό χαρακτηριστικό της συγκεκριμένης υλοποίησης είναι ότι μπορεί να τρέχει σε πραγματικό χρόνο με την καταγραφή π.χ με βίντεο του χώρου και να παράγει την αναπαραγωγή της γεωμετρίας στα πρότυπα του προβλήματος SfM. Περιγράφουμε αρχικά τις αριθμητικές μεθόδους που χρησιμοποιούνται στην υλοποίηση και στην συνέχεια τον τρόπο που χρησιμοποιούνται από την προτεινόμενη μέθοδο.

Τρόπος επίλυσης συστήματος: Η εξίσωση του αλγορίθμου LM που χρησιμοποιείται σε αυτή την υλοποίηση είναι η 2.23 στην οποία έχουμε επαυξήσει την διαγώνιο της προσέγγισης της Hessian πολλαπλασιάζοντας την με $1 + \lambda$. Κάθε φορά που αποτυγχάνει η ανανέωση του διανύσματος \mathbf{x} να βελτιώσει την τιμή του σφάλματος, αποφεύγουμε να ξαναλύσουμε το Σύστημα 2.23 και διαιρούμε το προταθέν βήμα με λ για να βρούμε την νέα τιμή του βήματος ανανέωσης του \mathbf{x} . Σχηματίζουμε κανονικά το συμπλήρωμα του Schur 2.33 και λύνουμε το σύστημα με τον αλγόριθμο παραγοντοποίησης του Cholesky. Η συνάρτηση σφάλματος που χρησιμοποιούμε είναι η σθεναρή συνάρτηση 2.27.

Μέθοδος υλοποίησης: Η αρχικοποίηση της μεθόδου γίνεται με την χρήση τριών καρέ. Για τους σκοπούς της αρχικοποίησης εφαρμόζουμε διαδικασία RANSAC χρησιμοποιώντας την μέθοδο σχετικού προσανατολισμού 5 σημείων όπως περιγράφεται στο [161]. Στην συνέχεια εφαρμόζουμε Bundle adjustment στην αρχικοποίηση των τριών όψεων.

Κάθε φορά που εντάσσουμε μια νέα όψη στους υπολογισμούς, πρώτα αρχικοποιούμε την θέση της νέας κάμερας. Αυτό γίνεται εφαρμόζοντας διαδικασία RANSAC σε υποθέσεις που παράγονται από resections τριών σημείων.

Στην συνέχεια γίνεται τριγωνισμός των 3D σημείων που είναι ορατά από την νέα όψη, χρησιμοποιώντας την νέα όψη και την πιο παλιά (ως προς την εισαγωγή της στους υπολογισμούς του B.A.) όψη στην οποία αυτά είναι ορατά. Τέλος εκτελούμε τον αλγόριθμο B.A., τις λεπτομέρειες του οποίου περιγράψαμε στους τρόπους επίλυσης του συστήματος. Χρησιμοποιούμε τα τελευταία m καρέ και εκτελούμε το πολύ n επαναλήψεις του B.A. Οι όψεις που δεν είναι στις m πρόσφατες είναι παγωμένες και δεν μεταβάλλονται, συνεισφέρουν όμως στον υπολογισμό του λάθους. Δεν χρειάζεται να εντοπίσουμε outliers, αφού χρησιμοποιούμε σθεναρή συνάρτηση σφάλματος.

Διαγράμματα για την κατανομή του χρόνου υπολογισμού παρουσιάζονται παρακάτω. Αν αυξήσουμε σημαντικά τον αριθμό των ελεύθερων σε μεταβο-

λές όψεων, η επίλυση του συστήματος 2.33 που κλιμακώνεται κυβικά με τον αριθμό των όψεων θα κυριαρχήσει τελικά στον χρόνο υπολογισμού. Η ικανότητα του συστήματος να τρέχει σε πραγματικό χρόνο χάνεται σύμφωνα με τους συγγραφείς στις 50 – 100 ελεύθερες όψεις.

Εφαρμογή:Υλοποιήσεις για μεγάλα σύνολα δεδομένων

Στην συνέχεια παρουσιάζουμε υλοποιήσεις του αλγορίθμου B.A. που μπορούν να εφαρμοστούν σε μεγάλα σύνολα δεδομένων, της τάξης των χιλιάδων φωτογραφιών. Αρχίζουμε με την υλοποίηση του [4] και κάνουμε στην συνέχεια μια σύντομη αναφορά στην βελτίωση της από τον [249].

Τρόπος επίλυσης συστήματος: Το γραμμικό σύστημα εξισώσεων 2.26 το οποίο λύνουμε σε κάθε επανάληψη του αλγορίθμου LM είναι υπολογιστικά απαιτητικό και αποτελεί την στενωπό του προβλήματος. Συνεπώς είναι λογικό να προσπαθήσουμε να αλλάξουμε στοιχεία που αφορούν την επίλυση του προκειμένου να πετύχουμε ικανοποιητικές επιδόσεις. Αρχικά παρατηρούμε μια αδικαιολόγητη εμμονή να λύνουμε το Σύστημα 2.26 ακριβώς. Αδικαιολόγητη, καθώς δεν είναι καν το πρόβλημα που μας ενδιαφέρει. Αποφασίζουμε λοιπόν να λύνουμε το Σύστημα 2.26 μόνο κατά προσέγγιση και όχι ακριβώς. Οι μέθοδοι αυτοί ονομάζονται και περικεκομμένοι/ανακριβείς επιλυτές Newton (Truncated/Inexact Newton solvers).

Οι ανακριβείς επιλυτές Newton απαρτίζονται από δύο μέρη:

- Μια υπολογιστικά φτηνή, επαναληπτική μέθοδο, κατά προσέγγιση επίλυσης γραμμικών συστημάτων. Τυπική επιλογή είναι επιλυτές Conjugate Gradients [163].
- Έναν κανόνα τερματισμού για τον επιλυτή. Ακολουθίες κανόνων τερματισμού που οδηγούν σε συγκλίνουσες ακολουθίες λύσεων παρουσιάζονται στο [245].

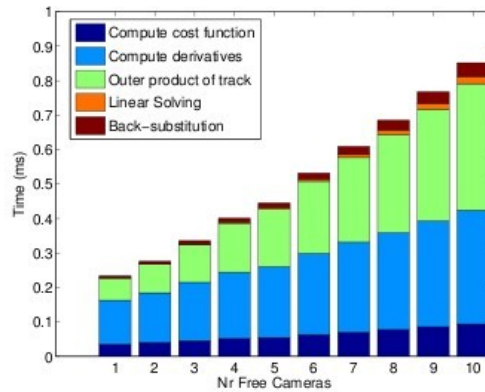
Η απλή εφαρμογή μεθόδων Conjugate Gradients δεν επαρκεί για την αντιμετώπιση του προβλήματος Bundle Adjustment. Αυτό γιατί ο ρυθμός σύγκλισης της μεθόδου εξαρτάται από την κατανομή των ιδιοτιμών της μήτρας N της Εξίσωσης 2.24. Γενικότερα για την επίλυση του συστήματος $Ax = b$ με επιλυτές Conjugate Gradients, ο ρυθμός σύγκλισης καθορίζεται από την κατανομή των ιδιοτιμών της μήτρας A [190].

Για να αντιμετωπίσουμε το πρόβλημα χρησιμοποιούμε την μέθοδο preconditioned Conjugate Gradients. Συμφωνά με αυτή την μέθοδο, αντί να λύσουμε το σύστημα $Ax = b$ λύνουμε το σύστημα $M^{-1}Ax = M^{-1}b$. Ο ρυθμός σύγκλισης της μεθόδου Conjugate Gradients εξαρτάται τώρα από την δομή της μήτρας $M^{-1}A$, η οποία αποσκοπούμε να γίνει ευνοϊκότερη. Προφανώς ο πίνακας M πρέπει να είναι εύκολα αντιστρέψιμος.

Στο πρόβλημα του B.A. έχουμε την επιλογή να λύσουμε με την μέθοδο των preconditioned Conjugate Gradients είτε το Σύστημα 2.31 είτε το Σύστημα 2.33. Επίσης υπάρχουν εναλλακτικές επιλογές για την επιλογή του πίνακα M σε κάθε περίπτωση. Ιδιότητες και προτάσεις για τους πίνακες M παρουσιάζονται στα [4, 50, 137].

Στο [4] παρουσιάζονται διάφορες μέθοδοι preconditioned Conjugate Gradients και συγκρίνονται οι επιδόσεις τους. Παρουσιάζουμε τα αποτελέσματα τους συνοπτικά στους επόμενους πίνακες.

Οι [249] βελτίωσαν την μέθοδο των [4]. Στην υλοποίηση τους δεν αποθη-



Σχήμα 2.4: Η κατανομή του χρόνου υπολογισμού στις επιμέρους διεργασίες. Από το [51]

| Μέθοδος | LM | Schur | Cholesky |
|--------------------------------|------------------------------------|---|-----------------------|
| explicit-direct | Ακριβής | Κατασκευάζεται | Πυκνή παραγοντοποίηση |
| explicit-sparse | Ακριβής | Κατασκευάζεται | Αραιή παραγοντοποίηση |
| normal-jacobi | Preconditioned Conjugate Gradients | Κατασκευάζεται | - |
| explicit-jacobi | Preconditioned Conjugate Gradients | Κατασκευάζεται | - |
| implicit-jacobi, implicit-ssor | Preconditioned Conjugate Gradients | Μόνο διανυσματικά γινόμενα με τον S κατασκευάζονται | - |

Πίνακας 2.2: Παρουσίαση μεθόδων που εξετάζονται στο [4]

| Μέγεθος Προβλήματος | Είδος κατάλληλης μεθόδου |
|---------------------|--|
| Μικρά προβλήματα | Προτείνονται ακριβείς μέθοδοι με πυκνή παραγοντοποίηση Cholesky |
| Μεγάλα προβλήματα | Προτείνονται επαναληπτικές μέθοδοι (με ανακριβείς επιλυτές Newton) |

Πίνακας 2.3: Συμπεράσματα από [4]

κεύουν κανέναν πίνακα στην μνήμη (Schur,Hessian,Jacobian) και πετυχαίνουν υψηλή παραλληλοποίηση κώδικα. Βελτιώνουν τις προηγούμενες επιδόσεις κατά πολλαπλασιαστικό παράγοντα της τάξης του 10 – 30.

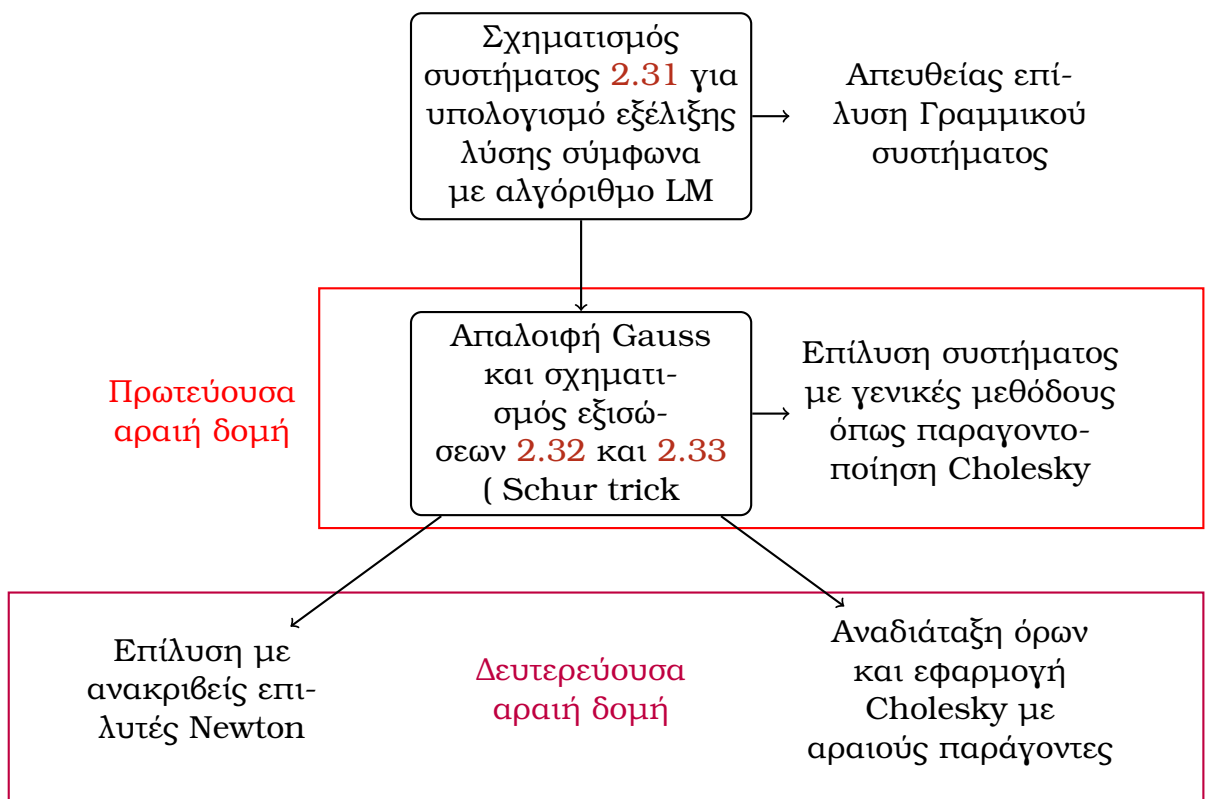
Εφαρμογή:Φωτογραφικός τουρισμός

Στην συνέχεια παρουσιάζουμε το σύστημα ανακατασκευής αραιούς τρισδιάστατης γεωμετρίας από μεγάλα σύνολα φωτογραφιών του διαδικτύων των [210]. Η ανακατασκευή της γεωμετρίας στο συγκεκριμένο σύστημα γίνεται επαναληπτικά. Αντί να εισαχθούν στο πρόβλημα βελτιστοποίησης όλες οι όψεις ταυτόχρονα, εισάγεται μία όψη σε κάθε επανάληψη.

Αρχική επανάληψη: Αρχικά διαλέγουμε τις δύο όψεις με τις περισσότερες αντιστοιχίες χαρακτηριστικών (σημείων). Επίσης λαμβάνεται κατάλληλη μέριμνα προκειμένου να αποφύγουμε εκφυλισμένες περιπτώσεις (όπως να μπορούν οι δύο όψεις να περιγραφούν ικανοποιητικά μέσω ομογραφίας).

Προσθήκη νέων όψεων και σημείων του χώρου: Σε κάθε επανάληψη εισάγουμε την όψη στην οποία είναι ορατά τα περισσότερα ανακατασκευασμένα σημεία του χώρου. Η αρχικοποίηση των εσωτερικών και εξωτερικών παραμέτρων γίνεται με χρήση διαδικασίας RANSAC. Οι εκτιμήσεις των παραμέτρων που χρησιμοποιούνται στον RANSAC δίνονται από υπολογισμό του Direct Linear transform. Για τις εσωτερικές παραμέτρους της κάμερας χρησιμοποιούμε και EXIF δεδομένα, όταν υπάρχουν. Επίσης εισάγουμε στο πρόβλημα βελτιστοποίησης κάποια από τα νέα τρισδιάστατα σημεία που φαίνονται στην νέα όψη. Αυτά πρέπει να φαίνονται σε τουλάχιστον μία όψη που εισήχθη παλαιότερα στο πρόβλημα και η ανακατασκευή τους με τριγωνισμό να δίνει μια καλώς ορισμένη θέση τους στον χώρο.

Πέρα από τον βασικό πυρήνα που περιγράψαμε, περισσότερες λεπτομέρειες μπορούν να αναζητηθούν στο [210].



Σχήμα 2.5: Διάφορες προσεγγίσεις για την περάτωση μιας επανάληψης του Bundle Adjustment.

2.2 Παραπομπές στην βιβλιογραφία. Διαθέσιμες υλοποιήσεις

Για μια ενδελεχή ανάλυση του προβλήματος μη γραμμικών ελαχίστων τετραγώνων σημαντικές παραπομπές είναι [46, 81, 93]. Για τον αλγόριθμο LM παραπέμπουμε τους αναγνώστες στα [163, 154, 24, 136].

Μια δημόσια διαθέσιμη υλοποίηση του αλγορίθμου Bundle adjustment που βασίζεται στην πρωτεύουσα αραιή δομή του προβλήματος και το συμπλήρωμα του Schur περιγράφεται στο [130] και είναι διαθέσιμη από [131]. Ακόμα οι βελτιώσεις του [249] ενσωματώθηκαν στο πρόγραμμα VisualSfM το οποίο διαθέτει πολλές δυνατότητες πέρα από την ανακατασκευή αραιούς 3d γεωμετρίας και υλοποιεί και γραφικό περιβάλλον χειρισμού του. Είναι διαθέσιμο από το [247]. Τέλος πακέτο λογισμικού στο οποίο συμπεριλαμβάνεται η υλοποίηση του SfM που περιγράφεται στο [210] είναι διαθέσιμο με το όνομα Bundler από [207].

Το Κεφάλαιο βασίζεται κυρίως στα [66], [64], [193], [78]

Το πρόβλημα της στερέωσης πολλών όψεων (MVS- Multi-View Stereopsis) συνίσταται στην ανακατασκευή ενός πλήρους τρισδιάστατου μοντέλου από φωτογραφίες που έχουν ληφθεί από διαφορετικές θέσεις στον χώρο[193]. Οι κάμερες θεωρούνται βαθμονομημένες, γνωρίζουμε δηλαδή τις παραμέτρους προβολής. Ξεκινάμε κάνοντας μια επισκόπηση της περιοχής και των διαφορετικών προσεγγίσεων στο πρόβλημα (3.1). Στην συνέχεια παρουσιάζουμε συγκεκριμένες μεθόδους που ενδείκνυται για την χρήση σε μεγάλα, μη-δομημένα, σύνολα φωτογραφιών (3.2,3.3). Τέλος αναφέρουμε πρόσφατες εξελίξεις συστημάτων MVS και παρουσιάζουμε συνοπτικά μεθόδους και αποτελέσματα (3.4).

3.1 Επισκόπηση μεθόδων MVS

Ακολουθώντας την ταξινόμια που προτείνεται στο [193] χωρίζουμε τους αλγόριθμους MVS σε κατηγορίες σύμφωνα με τις ακόλουθες 6 ιδιότητες:

1. Τρόπος αναπαράστασης της σκηνής
2. Μέτρο φωτομετρικής συνοχής (Photoconsistency)
3. Μοντέλο ορατότητας
4. Αλγόριθμος ανακατασκευής
5. Απαιτήσεις αρχικοποίησης
6. Εκ των προτέρων σχήμα (Shape prior)

3.1.1 Αναπαράσταση της σκηνής

Η αναπαράσταση της σκηνής μπορεί να γίνει διάφορους τρόπους. Πρώτον, μπορούμε να χρησιμοποιήσουμε ένα πλέγμα (grid) με ομοιόμορφη δειγματοληψία. Η γεωμετρία αναπαρίσταται σε αυτό το πλέγμα με κάποια διακριτή συνάρτηση πληρότητας (voxels [61, 188, 221, 194, 49, 45, 121, 26, 120, 22, 251, 25, 231, 205, 240]) ή κωδικοποιείται με την απόσταση από την κοντινότερη επιφάνεια (επιπεδοσύνολα [53, 178, 211, 102, 48, 103, 179]).

Οι μέθοδοι αυτοί είναι δημοφιλείς για την απλότητα, την ομοιομορφία τους και την ικανότητα να προσεγγίσουν οποιαδήποτε επιφάνεια. Δεύτερον, μπορούμε να χρησιμοποιήσουμε πολυγωνικά πλέγματα (meshes) και να αναπαραστήσουμε την επιφάνεια ως σύνολο συνδεδεμένων επίπεδων εδρών [62, 185, 258, 97, 92, 253]. Τα πλέγματα μπορούν να αποθηκευτούν αποδοτικά και διευκολύνουν τους υπολογισμούς ορατότητας. Τέλος, η γεωμετρία μπορεί να αναπαρασταθεί με χάρτες βάθους, έναν για κάθε εικόνα του συνόλου δεδομένων. Συστήματα MVS μπορεί να χρησιμοποιούν ένα μόνο τρόπο αναπαράστασης ή να χρησιμοποιούν διαφορετική αναπαράσταση ανάλογα με το στάδιο της ανακατασκευής[47, 73, 261, 116, 108, 219].

3.1.2 Φωτομετρική συνοχή

Βασικό συστατικό κάθε συστήματος MVS είναι η αξιολόγηση της οπτικής συμβατότητας της ανακατασκευής και των δεδομένων (εικόνων) από τις οποίες προέκυψε. Τον σκοπό αυτό εξυπηρετεί το μέτρο φωτομετρικής συνοχής του συστήματος. Η επικρατέστερη αντιμετώπιση του θέματος γίνεται με βάση την συσχέτιση των pixels μεταξύ των εικόνων. Κατηγοριοποιούμε τα μέτρα φωτομετρικής συνοχής σε μέτρα χώρου σκηνής και χώρου εικόνας, σύμφωνα με το [211].

Τα μέτρα *χώρου σκηνής* παίρνουν ένα στοιχείο (σημείο, όγκο κ.α) της ανακατασκευασμένης γεωμετρίας, το προβάλλουν στις εικόνες εισόδου και αξιολογούν την ποσότητα της συμφωνίας των προβολών αυτών. Για την αξιολόγηση, διάφοροι τρόποι έχουν προταθεί στην βιβλιογραφία. Οι [194, 121] χρησιμοποιούν την διασπορά των προβαλλόμενων pixels στις εικόνες εισόδου. Οι [53, 102, 92] συγκρίνουν δύο εικόνες την φορά και χρησιμοποιούν μετρικές που βασίζονται σε παραθυροποίηση όπως αθροίσματα τετραγώνων διαφορών (SSD) ή κανονικοποιημένες ετεροσυσχετίσεις (NCC). Ακόμα οι [120, 205] προτείνουν μέτρα σθεναρά σε μικρές μετατοπίσεις και άλλα φαινόμενα.

Τα μέτρα *χώρου εικόνας* χρησιμοποιούν την ανακατασκευασμένη γεωμετρία για να παραμορφώσουν μια εικόνα από μια δεδομένη γωνία θέασης στον χώρο και να προβλέψουν (μέσω της παραμόρφωσης της εικόνας) μια διαφορετική όψη. Στην συνέχεια συγκρίνουν την πρόβλεψη, δηλαδή την παραμορφωμένη εικόνα, με την μέτρηση, δηλαδή την εικόνα του συνόλου δεδομένων από την συγκεκριμένη γωνία θέασης (σφάλμα πρόβλεψης). Χρήση τέτοιων μετρικών γίνεται στα [179, 220]. Με τέτοιες μεθόδους αποδίδουμε μεγαλύτερο βάρος στα σημεία της σκηνής που έχουν φωτογραφηθεί πιο συχνά, αποτέλεσμα επιθυμητό αν υποθέσουμε ότι η συχνότητα φωτογράφισης σχετίζεται με την σημαντικότητα των αντικειμένων που απεικονίζονται.

Τα προηγούμενα μέτρα φωτομετρικής συνοχής υποθέτουν ότι η ανακλαστικότητα των επιφανειών δεν εξαρτάται από την γωνία θέασης (Lambertian επιφάνειες). Προσπάθειες έχουν γίνει για την μοντελοποίηση πιο σύνθετων συναρτήσεων ανάκλασης (BRDF)[251, 25, 231, 211, 102, 253].

3.1.3 Ορατότητα

Κατά τον υπολογισμό των μέτρων φωτομετρικής συνάφειας, πρέπει να προσδιορίσουμε ποιες όψεις θα χρησιμοποιήσουμε. Θέλουμε δηλαδή να υπολογίσουμε τα κομμάτια της σκηνής που είναι ορατά σε κάθε όψη και να λάβουμε υπόψιν την πιθανή ύπαρξη εμποδίων στις εικόνες που κρύβουν τμήματα της

σκηνής. Οι τεχνικές μπορούν να χωριστούν σε τρεις κατηγορίες

1. Γεωμετρικές
2. Ψευδό-γεωμετρικές
3. Βασισμένες σε outliers

Οι *γεωμετρικές* τεχνικές προσπαθούν να μοντελοποιήσουν την διαδικασία σχηματισμού εικόνων και το σχήμα των επιφανειών της σκηνής προκειμένου να εξάγουν συμπεράσματα ορατότητας. Σε μεθόδους οι οποίες εξελίσσουν επιφάνειες, μια συνηθισμένη προσέγγιση είναι να χρησιμοποιούμε την τρέχουσα εκτίμηση της γεωμετρίας για να υπολογίσουμε σε ποιες όψεις είναι ορατό κάθε σημείο της επιφάνειας [121, 61, 120, 240, 53, 258, 97, 257]. Μπορεί ακόμα να δειχθεί ότι, αν ξεκινήσουμε με αρχική επιφάνεια μια σφαίρα που περικλείει την σκηνή και την "σμιλέψουμε" για να προσαρμοστεί στις σωστές επιφάνειες του χώρου (space carving), οι εκτιμήσεις ορατότητας είναι συντηρητικές [121, 205]. Δηλαδή το σύνολο των όψεων στο οποίο εκτιμούμε ότι ένα σημείου του χώρου είναι ορατό είναι υποσύνολο του συνόλου των όψεων στις οποίες πράγματι είναι ορατό.

Οι *ψευδό-γεωμετρικές* τεχνικές χρησιμοποιούν προσεγγιστικούς γεωμετρικού υπολογισμούς για να εξάγουν σχέσεις ορατότητας. Έχουν προταθεί γενικά διάφορες ευριστικές και προσεγγιστικές τεχνικές προκειμένου να περιορίσουμε τους υπολογισμούς. Για παράδειγμα μπορούμε να περιορίσουμε τους υπολογισμούς φωτομετρικής συνάφειας σε ομάδες (clusters) γειτονικών χωρικά καμερών [158, 92]. Περισσότερες λεπτομέρειες και παραδείγματα μπορούν να βρεθούν στην βιβλιογραφία [121, 61, 25, 240, 122, 201].

Οι τεχνικές που βασίζονται στους *outliers* αποφεύγουν τον γεωμετρικό συμπερασμό και αντιμετωπίζουν τα εμπόδια που προκαλούν μερική απόκρυψη της σκηνής σε κάποιες όψεις, ως outliers. Σε περιπτώσεις που τα σημεία της σκηνής είναι πιο συχνά ορατά από κρυμμένα, απλές τεχνικές απόρριψης outlier έχουν παρουσιάσει ικανοποιητικά αποτελέσματα. Μια συνήθης ευριστική που εφαρμόζεται σε τέτοιες περιπτώσεις είναι να αποφεύγουμε να συγκρίνουμε όψεις που βρίσκονται πολύ μακριά, με σκοπό να αυξήσουμε το ποσοστό των inliers [92, 108, 73, 47].

3.1.4 Αλγόριθμος ανακατασκευής

Θα ταξινομήσουμε τους αλγόριθμους ανακατασκευής σε 4 κατηγορίες. Η *πρώτη* κατηγορία αλγορίθμων βασίζεται στα *voxels*. Στην κατηγορία αυτή αλγορίθμων, υπολογίζουμε πρώτα μια συνάρτηση κόστους σε έναν τρισδιάστατο όγκο και στην συνέχεια εξάγουμε μια επιφάνεια από τον όγκο αυτό. Απλό παράδειγμα είναι ο αλγόριθμος *voxel coloring* [194, 231], ο οποίος σε κάθε πέρασμα του υπολογίζει το κόστος κάθε voxel και ανακατασκευάζει όσα είναι αρκετά φθηνά, σύμφωνα με κάποιο κατώφλι. Άλλοι αλγόριθμοι διαφοροποιούνται στον ορισμό της συνάρτησης κόστους και στον τρόπο εξαγωγής της επιφάνειας. Αναφέρουμε ενδεικτικά αλγορίθμους που ορίζουν ένα Markov Random Field και εξάγουν την βέλτιστη επιφάνεια με αλγορίθμους γράφων [188, 240, 201].

Η *δεύτερη* κατηγορία αλγορίθμων *εξελίσσει* επαναληπτικά *επιφάνειες*, με σκοπό την βελτιστοποίηση μιας αντικειμενικής συνάρτησης. Διαφορετικές τεχνικές βασίζονται σε voxels, επιπεδοσύνολα και πλέγματα επιφανειών. Οι

τεχνικές *space carving* αφαιρούν ασυνεπή voxels από μια αρχική επιφάνεια [61, 121, 49, 120, 22, 205, 257, 191]. Παραλλαγές της μεθόδου προσθέτουν και αφαιρούν voxels [204, 251]. Τεχνικές που βασίζονται σε επιπεδοσύνολα βελτιστοποιούν ένα σύνολο μερικών διαφορικών εξισώσεων που ορίζεται στον όγκο. Συνήθως συρρικνώνουν έναν αρχικό όγκο, μπορούν όμως και να επεκτείνουν τον όγκο σε κάποιες περιοχές. Τέλος, οι τεχνικές που βασίζονται σε πλέγματα, τα εξελίσσουν σύμφωνα με εσωτερικές και εξωτερικές δυνάμεις.

Η *τρίτη* κατηγορία μεθόδων παράγουν *χάρτες βάθους* και επιβάλλουν περιορισμούς συμφωνίας μεταξύ των διαφορετικών χαρτών [77, 215, 219, 116, 261, 73]. Η τελική τρισδιάστατη επιφάνεια προκύπτει από την σύμμιξη των χαρτών.

Τέλος η *τέταρτη* κατηγορία αλγορίθμων εξάγει και ταιριάζει ένα σύνολο από χαρακτηριστικά και στην συνέχεια προσαρμόζει μια επιφάνεια στα χαρακτηριστικά αυτά [54, 138, 155, 222].

3.1.5 Απαιτήσεις αρχικοποίησης και περιορισμός σχήματος

Η επιβολή *περιορισμών σχήματος* μπορεί να βοηθήσουν τον αλγόριθμο ανακατασκευής να προσεγγίζει καλύτερα την υπό ανακατασκευή επιφάνεια, ειδικά σε περιοχές χωρίς έντονη υφή στις οποίες η φωτομετρική συνάφεια δεν επαρκεί. Το εκ των προτέρων σχήμα που θα χρησιμοποιήσουμε ουσιαστικά μεροληπτεί υπέρ συγκεκριμένων επιθυμητών χαρακτηριστικών κατά την διάρκεια της ανακατασκευής. Δεν είναι όμως τόσο σημαντική η χρήση του, επειδή οι περιορισμοί που προκύπτουν από τις πολλαπλές όψεις είναι ισχυροί.

Ανάλογα με το είδος του αλγορίθμου MVS που θα χρησιμοποιήσουμε, οι απαιτήσεις αρχικοποίησης διαφέρουν. Αρχικά οι παράμετροι προβολής και οι σχετικές θέσεις των όψεων θα πρέπει να είναι γνωστά. Συνήθως χρησιμοποιούμε κάποιο σύστημα SfM, όπως αυτά που περιγράψαμε προηγούμενα, για να ανακτήσουμε αυτές τις πληροφορίες. Διάφοροι αλγόριθμοι απαιτούν την οριοθέτηση ενός *όγκου* που *περικλείει* την σκηνή, όπως για παράδειγμα αλγόριθμοι *space carving* και αλγόριθμοι που χρησιμοποιούν επιπεδοσύνολα. Επίσης μπορεί να απαιτείται διαχωρισμός προσκηνίου/παρασκηνίου ή περιορισμός στο εύρος των τιμών βάθους για τα σημεία.

3.1.6 Καταλληλότητα μεθόδων για μεγάλα, μη δομημένα, σύνολα φωτογραφιών

Από τις μεθόδους που αναφέραμε προηγουμένως, δεν είναι όλες κατάλληλες για εφαρμογή σε εξωτερικές σκηνές και μεγάλα σύνολα φωτογραφιών. Μέθοδοι που στηρίζονται στην έννοια του Visually Hull [122], ενδεικτικά αναφέρουμε τις [92, 201, 240] από όσες αναφέρθηκαν προηγούμενα, δεν είναι κατάλληλες για εξωτερικές σκηνές. Ακόμα για λόγους υπολογιστικής πολυπλοκότητας και απαιτήσεων διαθέσιμης μνήμης, μέθοδοι που βασίζονται σε voxels, *space carving*, επιπεδοσύνολα και ογκομετρικούς γράφους πρέπει να απορριφθούν. Τέλος η ύπαρξη εμποδίων που επηρεάζουν την ορατότητα σε κάποιες όψεις, καθιστά μεθόδους που βασίζονται στην θεωρία μεταβολών ακατάλληλες καθώς εγκλωβίζονται σε κακά τοπικά ελάχιστα της συνάρτησης κόστους και δεν είναι δυνατόν να τις αρχικοποιήσουμε αρκετά κοντά στην

επιθυμητή λύση.

Οι μέθοδοι που έχουν αναπτυχθεί για τέτοια σύνολα βασίζονται κυρίως σε χάρτες βάθους και χρήση μικρών προσανατολισμένων επίπεδων επιφανειών (*patches* [69, 124, 83]). Θα παρουσιάσουμε αναλυτικά δύο τέτοιες μεθόδους στις επόμενες ενότητες.

3.2 Μια μέθοδος MVS βασισμένη σε χάρτες βάθους

Στην ενότητα αυτή περιγράφουμε το σύστημα [78] το οποίο είναι μια από τις πρώτες επιτυχημένες προσπάθειες εφαρμογής MVS σε σύνολα φωτογραφιών που προέρχονται το internet. Τα βασικά συστατικά της μεθόδου που την καθιστούν επιτυχημένη είναι ο τρόπος *επιλογής όψεων*, που συνεισφέρει στην εύρεση περισσότερων αντιστοιχιών μεταξύ των εικόνων και στην ενίσχυση της πληρότητας των ανακατασκευασμένων επιφανειών, και η ταυτόχρονη βελτιστοποίηση της *διεύθυνσης* της επιφάνειας.

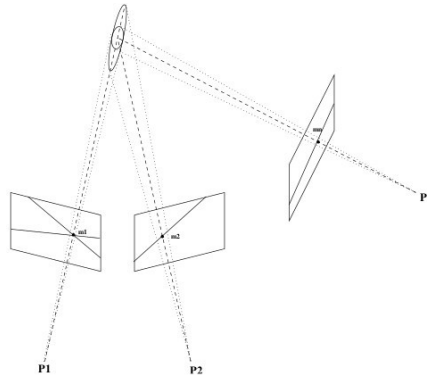
3.2.1 Γενική περιγραφή και σύνδεση με προηγούμενες προσεγγίσεις

Στοιχεία της μεθόδου

Δεδομένου του μεγέθους του συνόλου φωτογραφιών, αναμένουμε ότι μπορούν να βρεθούν φωτογραφίες με παρόμοιες *συνθήκες* λήψης (π.χ εξωτερικός φωτισμός, μετεωρολογικές συνθήκες, ανάλυση) και ικανοποιητική *απόσταση* μεταξύ των θέσεων λήψης των φωτογραφιών. Αν δύο κάμερες βρίσκονται κοντά στον χώρο, ο θόρυβος των μετρήσεων επηρεάζει σημαντικά την εκτιμώμενη θέση στον χώρο ενός σημείου που είναι ορατό και στις δύο κάμερες. Αυτό φαίνεται καλύτερα και στην Εικόνα 3.1. Από την άλλη σημειώνουμε ότι η πολύ μεγάλη απόσταση δύο λήψεων δυσχεραίνει το ταίριασμα σημείων μεταξύ των δύο εικόνων. Για την επιλογή καλών ζευγών εικόνων, θα πρέπει σε επίπεδο εικόνων να ταιριάζουμε *κλίμακα* και *εμφάνιση* και να διασφαλίσουμε ικανοποιητική *απόσταση* των καμερών. Σε επίπεδο pixel θα πρέπει να αντιμετωπίσουμε πιθανά *εμπόδια* (π.χ φύλλα, άνθρωποι) στις *φωτογραφίες*, τον *θόρυβο* και τις τοπικές διαφοροποιήσεις στον *φωτισμό*. Η επιτυχής περάτωση ενός τέτοιου συστήματος είναι η μέθοδος επιλογής όψεων που αναφέραμε και στην εισαγωγή της ενότητας.

Επειδή το σύνολο φωτογραφιών δεν χαρακτηρίζεται από ομοιομορφία, ειδική μέριμνα λαμβάνεται για τον σχεδιασμό ενός σθεναρού συστήματος εντοπισμού αντιστοιχιών μεταξύ εικόνων και δημιουργίας των χαρτών βάθους. Λόγω και της επιτυχίας των μεθόδων SfM σε μη δομημένα σύνολα φωτογραφιών, που περιγράψαμε σε προηγούμενα κεφάλαια, προσεγγίζουμε το πρόβλημα χρησιμοποιώντας το αρχικό αραιό σύνολο αντιστοιχιών που μας παρέχει ένα σύστημα SfM και επεκτείνοντας τις περιοχές αυτές (*region growing*). Οι χάρτες βάθους που παράγονται είναι *συνεχείς*, με σκοπό την επίτευξη ακρίβειας καλύτερης του pixel.

Για να ενισχυθεί η προσαρμοστικότητα σε διαφορές εμφάνισης μεταξύ των όψεων, απορρίπτουμε τις όψεις που δεν βελτιώνουν την ετεροσυσχέτιση ταιριασμένων φωτομετρικών παραθύρων. Επίσης, βελτιστοποιούμε ταυτόχρονα



Σχήμα 3.1: Η συνάρτηση κατανομής του τρισδιάστατου σημείου από θορυβώδεις μετρήσεις εικόνων. Το μέγεθος της έλλειψης εκφράζει την αβεβαιότητα στην θέση του τρισδιάστατου σημείου. Γενικά μεγαλύτερη γωνία μεταξύ των οπτικών ακτίνων των δύο εικόνων δίνει μεγαλύτερη βεβαιότητα στην θέση του τρισδιάστατου σημείου[90]. Μαθηματικά, το πρόβλημα συνίσταται στην εύρεση της ύστερης συνάρτησης κατανομής πιθανότητας $P(\mathbf{X}|\mathbf{x}, \mathbf{x}') \sim P(x', x|X) = P(x'|X)P(x|X)$, όπου υποθέσαμε ομοιόμορφες πρότερες κατανομές και ανεξαρτησία εικόνων. Η εικόνα είναι από [153].

ως προς το *βάθος* και το *κάθετο διάνυσμα της επιφάνειας*.

Σύνδεση με την βιβλιογραφία

Δεδομένης μια όψης αναφοράς, η *εύρεση κατάλληλων όψεων* για την εύρεση κοινών σημείων και τον τριγωνισμό τους, αντιμετωπίζεται κλασσικά με την επιλογή των k πλησιέστερων όψεων [92, 77, 107, 158]. Η επιλογή αυτή υποθέτει *ομοιόμορφη* κατανομή των όψεων στον χώρο. Δεν είναι κατάλληλη όμως για μη δομημένα σύνολα φωτογραφιών, αφού συχνά οι πιο κοντινές όψεις θα είναι σχεδόν ταυτόσημες και δεν θα προσφέρουν αξιόπιστη πληροφορία για τον τριγωνισμό σημείων στον χώρο. Για αυτό και η επίδραση του συστήματος επιλογής όψεων είναι καθοριστική.

Η χρήση τεχνικών εντοπισμού outliers για τον προσδιορισμό εμποδίων που περιορίζουν την ορατότητα στις φωτογραφίες είναι μια κλασσική τεχνική η οποία περιγράφηκε ήδη στην Ενότητα 3.1.3. Επεκτείνεται και χρησιμοποιείται και από το συγκεκριμένο σύστημα. Επιπλέον βιβλιογραφικές αναφορές είναι οι [92, 77].

Μεγάλη δραστηριότητα έχει παρατηρηθεί και γύρω από την δημιουργία *σδεναρών* μεθόδων *ταιριάσματος* σημείων εικόνων. Έχουν αναπτυχθεί μέθοδοι κατάλληλες για σύνολα φωτογραφιών με μεταβολές στον *φωτισμό* [94], με *μη Lambertian* επιφάνειες [102], με μεγάλες αλλαγές στην *εμφάνιση* [112]. Οι συγγραφείς του [78] χρησιμοποιούν απλά την *κανονικοποιημένη ετεροσχέση*. Επίσης επισημαίνουν ότι η χρήση μέτρων αναλλοίωτων σε πολλούς μετασχηματισμούς των εικόνων μειώνει την *διακριτική* ικανότητα τους.

Όσον αφορά την βελτιστοποίηση ως προς το *κάθετο διάνυσμα* στην επιφάνεια, στον τομέα της στερέωσης δύο όψεων η μέθοδος έχει εφαρμοστεί στην εργασία [43]. Μέθοδοι που εφαρμόζουν *επέκταση περιοχών* (region growing) ξεκινώντας από αρχικά σημεία αναπτύσσονται στα [69, 84]. Οι συγγραφείς τους όμως μοντελοποιούν προσεκτικά την ορατότητα για να ελαχιστοποιήσουν την επίδραση των outliers, ενώ στην εργασία που παρουσιάζουμε στηρίζομαστε

σε μεθόδους σθεναρής στατιστικής και στην μέθοδο επιλογής όψεων που αναπτύσσεται.

Σαν βάση του συστήματος χρησιμοποιείται η γνωστή από την φωτογραμμετρία μεθοδολογία ταιριάσματος *MPGC* (Multi Photo Geometrically Constrained Least Squares) [14, 82]. Πιο συγκεκριμένα επεκτείνει τον αλγόριθμο *MVS* του [171] επιβάλλοντας συνέπεια στα κάθετα διανύσματα της επιφάνειας μεταξύ των πολλαπλών όψεων. Τέλος η μέθοδος συνδέεται στενά με τον αλγόριθμο στερέωσης δύο όψεων του [107].

3.2.2 Περιγραφή του προτεινόμενου αλγόριθμου

Προεπεξεργασία και βαθμονόμηση όψεων

Στο πρώτο στάδιο του αλγορίθμου, αφαιρούμε από τις εικόνες την *ακτινική παραμόρφωση*. Οι συγγραφείς του άρθρου χρησιμοποιούν το εμπορικά διαθέσιμο λογισμικό *PTLens* το οποίο χρησιμοποιεί τα δεδομένα από τις ετικέτες *EXIF* των φωτογραφιών. Οι φωτογραφίες οι οποίες δεν μπορούν να διορθωθούν απορρίπτονται. Σημειώνουμε ότι συστήματα όπως το *VisualSFM* [247] και το *Bundler* [207], τα οποία είναι δημόσια διαθέσιμα, μπορούν να χρησιμοποιηθούν και για την αφαίρεση της ακτινικής παραμόρφωσης.

Στην συνέχεια προτείνεται η χρήση κάποιου *σθεναρού συστήματος SfM* όπως τα [30, 210]. Έχουμε τώρα στην διάθεσή μας την εξωτερική και εσωτερική *βαθμονόμηση* κάθε όψης, μια *αραιή αναπαράσταση* της σκηνής και για κάθε ανιχνευθέν *χαρακτηριστικό λίστα* με τις εικόνες στις οποίες ανιχνεύθηκε.

Τέλος *ραδιομετρικοί* μετασχηματισμοί εφαρμόζονται ώστε η παράσταση των εικόνων να γίνεται σε γραμμικό χώρο. Αν δεν γνωρίζουμε την ακριβή καμπύλη απόκρισης του αισθητήρα που κατέγραψε την εικόνα, εφαρμόζουμε τον αντίστροφο μετασχηματισμό *sRGB*.

Καθολικό επίπεδο επιλογής όψεων

Στο δεύτερο στάδιο θέλουμε να εκτιμήσουμε ένα *χάρτη βάθους* για κάθε εικόνα. Κάθε όψη χρησιμοποιείται ως όψη *αναφοράς* ακριβώς μία φορά. Προκειμένου να εντοπίσουμε *καλές* αντιστοιχίες-ταιριάσματα σημείων μεταξύ των εικόνων, εφαρμόζουμε επιλογή εικόνων σε δύο *επίπεδα*. Σε *καθολικό* (global) επίπεδο, καθορίζουμε ένα σύνολο *συνορευουσών* εικόνων για κάθε εικόνα. Σε *τοπικό* (local) επίπεδο, καθορίζουμε για κάθε *pixel*, κατάλληλο υποσύνολο των καθολικώς συνορευουσών εικόνων που αποφέρουν ένα *ευσταδές* στερεοσκοπικό ταιρίασμα. Τις διαδικασίες αυτές θα αναλύσουμε ευθύς αμέσως.

Σε *καθολικό επίπεδο*, αναζητούμε για κάθε όψη αναφοράς R , N γείτονες οι οποίοι είναι καλοί υποψήφιοι για στερεοσκοπικό ταιρίασμα, με κριτήρια το *περιεχόμενο* της σκηνής, την *εμφάνιση* και την *κλίμακα*. Τα χαρακτηριστικά που έχουν ταιριάζει από την εκτέλεση του αλγορίθμου *SfM* είναι μια καλή ένδειξη για την συμβατότητα διαφορετικών όψεων, εμφανίζουν όμως και προβλήματα. Συγκεκριμένα εικόνες με μεγάλο αριθμό ταιριασμένων χαρακτηριστικών είναι σχεδόν ίδιες, επομένως οι θέσεις λήψης τους δεν απέχουν ικανοποιητικά στον χώρο για να δώσουν ακριβή αποτελέσματα έπειτα από τριγωνισμό των κοινών σημείων τους. Επίσης εικόνες με διαφορετική ανάλυση μπορούν να ταιριάζουν πολύ καλά, δεν εξυπηρετούν όμως τους σκοπούς μας. Για αυτό υπολογίζουμε για κάθε εικόνα ένα καθολικό *score* που

αντιπροσωπεύει πόσο καλοί υποψήφιοι είναι για στερεοσκοπικό ταίριασμα με την εικόνα αναφοράς.

Για κάθε εικόνα V στο εσωτερικό μιας γειτονιάς N υπολογίζουμε το σταθμισμένο άθροισμα των κοινών χαρακτηριστικών με την εικόνα αναφοράς σύμφωνα με τον τύπο

$$g_R(V) = \sum_{f \in F_V \cap F_R} W_N(f) W_s(f) \quad (3.1)$$

Με F_X συμβολίζουμε το σύνολο των χαρακτηριστικών που είναι ορατά στην όψη X .

Ο όρος W_N χρησιμοποιείται προκειμένου να αποδώσουμε *χαμηλότερο* score στις όψεις με *γωνίες τριγωνισμού* μικρότερες των *δέκα* μοιρών (10°). Όψεις με *πολύ μεγάλες* γωνίες τριγωνισμού έχουν χαμηλό score από την μικρή πληθικότητα του συνόλου κοινών χαρακτηριστικών. Αναλυτικά ο τύπος υπολογισμού του W_N δίνεται

$$W_N = \prod_A W_\alpha(f, V_i, V_j) \quad (3.2)$$

$$W_\alpha(f, V_i, V_j) = \min((\alpha/\alpha_{max})^2, 1) \quad (3.3)$$

$$A: V_i, V_j \in N \text{ s.t. } f \in F_{V_i} \cap F_{V_j}$$

Με α συμβολίζουμε την γωνία που σχηματίζουν οι ακτίνες που διέρχονται από τα κέντρα προβολής των εικόνων V_i, V_j και τα σημεία των V_i, V_j που αντιστοιχούν στο χαρακτηριστικό f .

Ο όρος W_s έχει σχεδιαστεί για να ενισχύει το score εικόνων με παρόμοια *ανάλυση* με την εικόνα αναφοράς. Στο χαρακτηριστικό f υπολογίζουμε την διάμετρο s_V της σφαίρας με κέντρο το f και διάμετρο, αφού προβληθεί στην εικόνα V , ίση με την απόσταση των pixels στην V . Αντίστοιχα υπολογίζουμε την διάμετρο s_R και υπολογίζουμε τον λόγο $r = s_R(f)/s_V(f)$. Αναλυτικά ο τύπος υπολογισμού του W_S δίνεται

$$W_s = \begin{cases} 2/r, & 2 \leq r \\ 1 & 1 \leq r \leq 2 \\ r & r \leq 1 \end{cases} \quad (3.4)$$

Τώρα για δεδομένο μέγεθος γειτονιάς $|N|$, μπορούμε να βρούμε την *βέλτιστη σύνθεση* της γειτονιάς. Βέλτιστη σύνθεση γειτονιάς θεωρείται αυτή για την οποία το άθροισμα των score των μελών της

$$\sum_{V \in N} g_R(V)$$

είναι μέγιστο. Για λόγους υπολογιστικής πολυπλοκότητας, ακολουθείται υποβέλτιστη *άπληστη* προσέγγιση για τον προσδιορισμό της σύνθεσης της γειτονιάς N . Έτσι για να υπολογίσουμε την γειτονία μεγέθους $\nu + 1$ προσθέτουμε στην τρέχουσα γειτονιά μεγέθους ν την εικόνα V με το μεγαλύτερο score $g_R(V)$ στην τρέχουσα γειτονιά. Η αρχικοποίηση της επαναληπτικής μεθόδου επέκτασης της γειτονιάς γίνεται με την προσθήκη στην κενή γειτονιά της εικόνας αναφοράς R .

Στο τελευταίο βήμα της καθολικής επιλογής όψεων, προσαρμόζουμε την ανάλυση των εικόνων της γειτονιάς, ώστε να περιορίσουμε την διακύμανση της. Αν και ήδη προσπαθήσαμε να επιλέξουμε εικόνες παρόμοιας κλίμακας, το τελευταίο αυτό βήμα είναι απαραίτητο για να εξαλείψουμε το ενδεχόμενο

μη αποδεκτής διακύμανσης κλίμακας. Για αυτό τον λόγο δειγματοληπούμε την εικόνα αναφοράς R ώστε να έχει παρόμοια ανάλυση με την εικόνα της γειτονιάς με την χαμηλότερη ανάλυση V_{min} . Πιο συγκεκριμένα ορίζουμε την συνάρτηση κλίμακας

$$scale_R(V) = \frac{1}{|F_V \cap F_R|} \sum_{f \in F_V \cap F_R} \frac{s_R(f)}{s_V(f)} \quad (3.5)$$

Το V_{min} ορίζεται ως το σημείο ελαχίστου της 3.5. Αν

$$V_{min} \leq t = 0.6$$

κλιμακώνουμε την όψη αναφοράς R ώστε μετά την κλιμάκωση να ισχύει

$$scale_R(V_{min}) = t$$

Στην συνέχεια κλιμακώνουμε όλες τις όψεις με

$$scale_R(V) > 1.2$$

στην κλίμακα της όψης αναφοράς R .

Τοπικό επίπεδο επιλογής όψεων

Στην συνέχεια παρουσιάζουμε το *τοπικό* επίπεδο επιλογής. Για κάθε *τοποθεσία* της εικόνας αναφοράς, επιλέγουμε ένα *υποσύνολο* της γειτονιάς N , $B \subset N$. Αποδεκτή τιμή της πληθικότητας του συνόλου B είναι $|B| = 4$. Η χρησιμοποίηση ενός μικρότερου υποσυνόλου για τον προσδιορισμού του βάθους βελτιώνει και την ταχύτητα υπολογισμού. Δεδομένης μια *τρέχουσας* εκτίμησης του *βάθους* και του *κάθετου* διανύσματος στην υπό εξέταση τοποθεσία, τα στοιχεία του B επιλέγονται ώστε να είναι *φωτομετρικά* συναφή με τις εκτιμήσεις αυτές. Ως μέτρο φωτομετρικής συνάφειας χρησιμοποιούμε την κανονικοποιημένη ετεροσυσχέτιση (NCC), μετρημένη μεταξύ pixels εντός παραθύρου περιμετρικά του υπό εξέταση pixel στις όψεις R και V . Επίσης μας ενδιαφέρει να *αποφύγουμε μικρές γωνίες τριγωνισμού* και να *αποφύγουμε συνεπίπεδα* σημεία. Για να εξυπηρετήσουμε τους σκοπούς αυτούς εισάγουμε μια συνάρτηση αξιολόγησης $l_R(V)$.

Ο *τρόπος επιλογής* της γειτονιάς ενός pixel διαμορφώνεται τελικά ως εξής. Δεδομένης μια αρχικής εκτίμησης του βάθους και του κάθετου διανύσματος, δοκιμάζουμε να προσθέσουμε στο σύνολο B την όψη V με το μεγαλύτερο score $l_R(V)$. Αν η τιμή της NCC για την όψη V ξεπεράσει κάποιο *κατώφλι*, οι συγγραφείς το θέτουν 0.3, προσθέτουμε την όψη V στο B , αλλιώς την απορρίπτουμε. Η διαδικασία επαναλαμβάνεται έως ότου συμπληρωθεί το B ή απορριφθούν όλες οι όψεις. Τονίζουμε ότι για να υπάρχει καλή κατανομή των όψεων στον χώρο, η δομή της συνάρτησης l_R επηρεάζεται από τα στοιχεία του συνόλου B και αποδίδει χαμηλά score σε όψεις που βρίσκονται *κοντά* σε κάποιο στοιχείο του συνόλου B . Κατά την διαδικασία της βελτιστοποίησης, οι τιμές του βάθους και του κάθετου διανύσματος μεταβάλλονται και υπάρχει το ενδεχόμενο κάποιο στοιχείο του συνόλου B να αφαιρεθεί από αυτό. Επαναλαμβάνουμε την ίδια διαδικασία για να προσθέσουμε κάποιο νέο στοιχείο στο B . Ο τερματισμός του αλγορίθμου εξασφαλίζεται, αφού στοιχεία που απορρίφθηκαν *δεν επανεξετάζονται*.

Διαδικασία ανακατασκευής

Τέλος περιγράφουμε την *διαδικασία ανακατασκευής*. Ο αλγόριθμος ανακατασκευής χωρίζεται σε δύο τμήματα: ένα σύστημα επέκτασης περιοχών (*region growing*) και ένα σύστημα *ταιριάσματος*.

Το σύστημα *επέκτασης περιοχών* διατηρεί μια *ουρά προτεραιότητας* Q . Οι εγγραφές της Q είναι οι υποψήφιοι για ταιρίασμα. Για κάθε εγγραφή διατηρούμε την θέση του pixel στην εικόνα αναφοράς R , και αρχικές τιμές για το βάθος και το κάθετο διάνυσμα στο pixel. Η ουρά *αρχικοποιείται* από τα χαρακτηριστικά SfM που είναι ορατά στην όψη αναφοράς R . Τα σημεία αυτά είναι ένα αραιό και σθεναρό σύνολο. Προσθέτουμε ακόμα στην ουρά χαρακτηριστικά που είναι ορατά σε κάθε όψη της γειτονίας N_R , αφού τα προβάλλουμε στην εικόνα R . Κακές αρχικοποιήσεις μπορεί να συμβούν, αλλά το πιο πιθανό είναι ότι θα απορριφθούν στα υπόλοιπα στάδια του αλγορίθμου. Για κάθε εγγραφή της ουράς Q , εκτελούμε διαδικασία στερεοσκοπικού ταιριάσματος προκειμένου να βελτιώσουμε τις τιμές αρχικοποίησης για το βάθος και το κάθετο διάνυσμα.

Η *ιδέα* στην οποία στηρίζεται το σύστημα επέκτασης περιοχών είναι ότι τα *επιτυχώς ταιριασμένα* pixels προσφέρουν καλές τιμές αρχικοποίησης για τους *γείτονες* τους. Επειδή η διαδικασία ταιριάσματος είναι ένα έντονα μη γραμμικό πρόβλημα βελτιστοποίησης, η καλή αρχικοποίηση είναι σημαντική για την ποιότητα της λύσης που θα πετύχουμε. Για αυτό λύνουμε για κάθε pixel πολλές φορές, με διαφορετικές αρχικοποιήσεις κάθε φορά. Στο συγκεκριμένο σύστημα, κάθε φορά που ένα pixel ταιριάζεται επιτυχώς, οι εικόνες στην γειτονία του B εντάσσονται στην ουρά.

Για κάθε *pixel*, η διαδικασία ταιριάσματος δίνει τιμές για το βάθος και το κάθετο διάνυσμα καθώς και μια *τιμή βεβαιότητας* του ταιριάσματος που εκφράζει την αξιοπιστία της λύσης. Η προτεραιότητα στην ουρά Q καθορίζεται από την τιμή της βεβαιότητας. Για να αντικαταστήσουμε τα στοιχεία μιας εγγραφής της ουράς Q από τα πρόσφατα υπολογισμένα, η βεβαιότητα της νέας λύσης θα πρέπει να υπερβαίνει αυτή της προηγούμενης. Η ουρά προτεραιότητας μας επιτρέπει να μην επεκταθούμε σε αναξιόπιστες περιοχές.

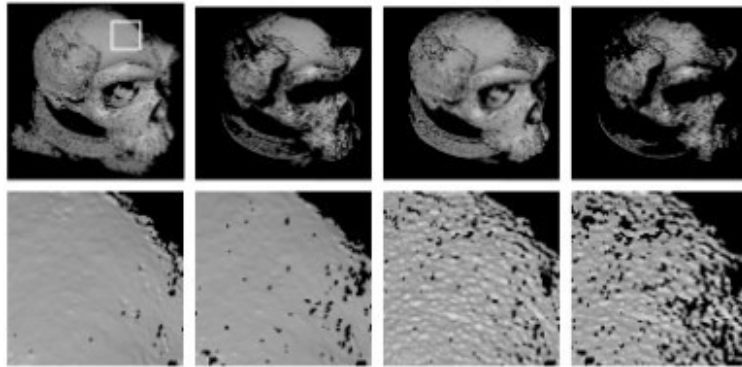
Στην διαδικασία *ταιριάσματος*, ορίζουμε ένα μικρό *παράθυρο* $n \times n$ γύρω από το pixel p της όψης αναφοράς και στοχεύουμε να *βελτιστοποιήσουμε* ως προς το βάθος και τον *προσανατολισμό* της προβολής του παραθύρου στην σκηνή (patch), ώστε να μεγιστοποιηθεί η *φωτομετρική συνάφεια* των προβολών του παραθύρου στην όψη αναφοράς και τις γειτονικές της όψεις. Στην διατύπωση του προβλήματος, η αναπαράσταση του προσανατολισμού της επιφάνειας (προβολής του παραθύρου $n \times n$ στην σκηνή) γίνεται με τρόπο συνεπή ανάμεσα στις διαφορετικές όψεις. Αποφεύγεται δηλαδή η περιγραφή του παραθύρου με παραμέτρους ανά όψη (π.χ. [82]) και εξαλείφονται οι περιττοί βαθμοί ελευθερίας. Στην συνέχεια συνδέουμε την ένταση μεταξύ του παραθύρου στην όψη αναφοράς R και στις γειτονικές όψεις V_R . Στην συνέχεια λύνουμε το υπερρορισμένο σύστημα που προκύπτει με την μέθοδο *MPGC*. Στην επαναληπτική διαδικασία βελτιστοποίησης, ειδική μέριμνα λαμβάνεται για την αποφυγή φαινομένων όπως ταλαντώσεις, αργή σύγκλιση και σύγκλιση σε κακό τοπικό ελάχιστο. Ανάμεσα στις επαναλήψεις της διαδικασίας βελτιστοποίησης, γειτονικές όψεις με χαμηλή NCC αντικαθίστανται από άλλες, καταλληλότερες. Η μέση τιμή της NCC στην γειτονία χρησιμοποιείται στο τέλος για να χαρακτηρίσει την βεβαιότητα της λύσης. Αναφέρουμε πάλι ότι η βεβαιότητα της λύσης καθορίζει πώς θα χρησιμοποιηθεί αυτή για να ανανεώσει τις τιμές των εγγραφών της ουράς Q .

Η *φωτομετρική* μοντελοποίηση γίνεται με όσο το δυνατόν πιο *απλό* τρόπο. Σύνθετα μοντέλα αυξάνουν την πολυπλοκότητα του προβλήματος εισάγοντας επιπλέον παραμέτρους και επηρεάζουν την ευστάθεια του συστήματος. Έτσι τα φαινόμενα ανάκλασης μοντελοποιούνται με έναν συντελεστή c_k , σταθερό σε όλο το παράθυρο στην όψη V_k . Η μοντελοποίηση αυτή περιγράφει πλήρως την περίπτωση επιφάνειας Lambert υπό σταθερό φωτισμό. Αποτυγχάνει όμως σε περιπτώσεις επιφανειών που εμφανίζονται υγρές και στεγνές, που έχουν τραχύ ανάγλυφο και σε περιπτώσεις που ο φωτισμός μεταβάλλεται στο εσωτερικό του παραθύρου (π.χ σκιάσεις).

Για την σύμμιξη των επιμέρους χαρτών βάθους μπορούμε να χρησιμοποιήσουμε πληθώρα τεχνικών, όπως η ανακατασκευή επιφάνειας Poisson (PSR [110]).

Αξιολόγηση και Παρουσίαση ενδεικτικών αποτελεσμάτων

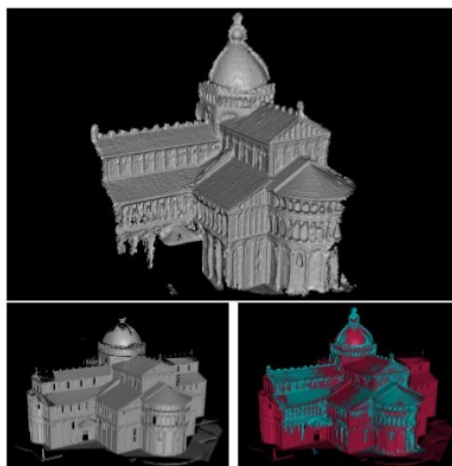
Μερικά αποτελέσματα παρουσιάζονται στις Εικόνες 3.2,3.3,3.4. Για το οικοδόμημα που παρουσιάζεται στην Εικόνα 3.4, σημειώνουμε ότι 90% των ανακατασκευασμένων δειγμάτων βρίσκεται σε απόσταση $0.128m$ από το μοντέλο που έδωσε ο σαρωτής laser (LIDAR).



Σχήμα 3.2: Αποτελέσματα ανακατασκευής. Στην δεύτερη σειρά παρουσιάζεται μεγέθυνση της σηματομενής λεπτομέρειας. Σε στήλες και από τα αριστερά έχουμε: LVS+ON, ON, LVS,Κανένα. Με LVS συμβολίζουμε το σύστημα επιλογής όψεων σε τοπικό επίπεδο και με ON την βελτιστοποίηση ως προς το κάθετο διάνυσμα. Από [78].



Σχήμα 3.3: Αριστερή στήλη: Εικόνες συνόλου δεδομένων. Μεσαία στήλη: Χάρτες βάθους. Δεξιά στήλη: Ανακατασκευασμένο μοντέλο. Από [78]



Σχήμα 3.4: Κορυφή: Ανακατασκευασμένο μοντέλο. Αριστερά: Μοντέλο από σαρωτή laser. Δεξιά: Υπέρθωση των δύο μοντέλων. Από [78]

3.3 Το σύστημα PMVS-Patch Multi View Stereo

Το σύστημα PMVS είναι ένα σύστημα MVS βασισμένο σε patches. Η υλοποίηση του είναι ελεύθερα διαθέσιμη από τον διαδικτυακό τόπο [66]. Ο προσανατολισμός του είναι σε μη δομημένα μεγάλα σύνολα φωτογραφιών εξωτερικών χώρων, έχει σχεδιαστεί δηλαδή στα πλαίσια που μας αφορούν περισσότερο. Συγκρατάλεγγεται στις state-of-the-art τεχνικές αυτή την στιγμή [197].

3.3.1 Γενική περιγραφή συστήματος PMVS. Δομές και συναρτήσεις

Το σύστημα PMVS είναι ένας ευέλικτος αλγόριθμος MVS που παράγει ένα πυκνό σύνολο από μικρά, παραλληλόγραμμα, προσανατολισμένα patches. Τα patches καλύπτουν τις παρατηρούμενες επιφάνειες, εκτός από περιοχές χωρίς σημαντική υφή και περιοχές που είναι κρυμμένες στις εικόνες εισόδου. Η διαδικασία που ακολουθείται για την ανακατασκευή των patches συνοψίζεται σε τρία στάδια

1. Ταίριασμα
2. Επέκταση
3. Διαλογή (φιλτράρισμα)

Στο στάδιο του *ταιριάσματος*, ξεκινάμε βρίσκοντας χαρακτηριστικά στις εικόνες του συνόλου εισόδου. Τα χαρακτηριστικά αυτά εξάγονται με τις μεθόδους της Διαφοράς Γκαουσιανών (DoG-Differences of Gaussians) [140] και των ανιχνευτών Harris [85]. Τα χαρακτηριστικά αυτά ταιριάζονται μεταξύ των εικόνων και παίρνουμε ένα αραιό σύνολο χαρακτηριστικών. Από τα αρχικά αυτά ταιριάσματα, τα επόμενα δύο στάδια επαναλαμβάνονται n φορές ($n = 3$ στο [69]).

Στο στάδιο της *επέκτασης*, χρησιμοποιούμε τις αντιστοιχίες για να παράγουμε ένα πυκνό σύνολο patch. Οι τεχνικές που χρησιμοποιούνται είναι αντίστοιχες με αυτές των [124, 83, 171, 56].

Στο στάδιο της *διαλογής*, εφαρμόζουμε περιορισμούς *ορατότητας* προκειμένου να απορρίψουμε εσφαλμένα ταιριάσματα. Η μέθοδος είναι αντίστοιχη με αυτή που εφαρμόζεται στο [124]. Όμως οι συγγραφείς του [69] αντικαθιστούν την άπληστη προσέγγιση των [124] με εναλλαγή ανάμεσα στις διαδικασίες διαλογής και επέκτασης. Ως αποτέλεσμα, σύνθετες επιφάνειες ανακατασκευάζονται επιτυχώς και γίνεται απόρριψη των outliers.

Στην συνέχεια, περαιτέρω επεξεργασία απαιτείται για την μετατροπή του συνόλου των patch σε μοντέλο *πλέγματος*. Μετά την αρχική κατασκευή του, το πλέγμα επεξεργάζεται με μεθόδους MVS που έχουν σχεδιαστεί για πλέγματα προκειμένου να *εκλεπτινθεί* η αναπαράσταση και να επιβάλλουμε *φωτομετρική* συνάφεια. Όπως και στην μέθοδο που περιγράψαμε στην Ενότητα 3.2, βελτιστοποιούνται και τα κάθετα διανύσματα στην επιφάνεια.

Δομές και συναρτήσεις

Ας αποσαφηνίσουμε καταρχάς την δομή του *patch*. Ένα patch p είναι ουσιαστικά μια τοπική προσέγγιση της επιφάνειας με ένα εφαπτόμενο επίπεδο. Η περιγραφή του p γίνεται με τα μεγέθη:

- Κέντρο $\mathbf{C}(p)$.
- Μοναδιαίο κάθετο διάνυσμα $\mathbf{n}(p)$ με προσανατολισμό προς την κάμερα που παρατηρεί το p .
- Εικόνα αναφοράς $R(p)$ στην οποία το p είναι ορατό.

Το p ορίζεται λοιπόν ως ένα τετράγωνο στον τρισδιάστατο χώρο, προσανατολισμένο έτσι ώστε οι ακμές του να είναι ευθυγραμμισμένες με τον x -άξονα της όψης αναφοράς $R(p)$. Η έκταση του p διαμορφώνεται ώστε το μικρότερο τετράγωνο της εικόνας $R(p)$, που εσωκλείει την προβολή του p στην $R(p)$, και είναι ευθυγραμμισμένο με τον άξονα x της $R(p)$, να είναι διαστάσεων $\mu \times \mu$ pixels (Στο [69] $\mu = 5$ ή 7).

Η συνάρτηση φωτομετρική *διαφωνίας* (το αντίθετο της φωτομετρικής συνάφειας) ορίζεται στο σύστημα ως

$$g(p) = \frac{1}{|V(p) \setminus R(p)|} \sum_{I \in |V(p) \setminus R(p)|} h(p, I, R(p)) \quad (3.6)$$

θα εξηγήσουμε τώρα τους συμβολισμούς της 3.6. Το σύνολο $V(p)$ περιέχει τις εικόνες από τις οποίες το p είναι ορατό. Θα αναλύσουμε αργότερα τον υπολογισμό του. Η συνάρτηση $h(p, I, R(p))$ περιγράφει την φωτομετρική διαφωνία των εικόνων που δέχεται ως ορίσματα. Για να υπολογίσουμε την h ακολουθούμε τα επόμενα βήματα:

- Καλύπτουμε το p με ένα πλέγμα διαστάσεων $\mu \times \mu$.
- Προβάλλουμε το πλέγμα στις εικόνες I_1, I_2 . Δειγματοληπτούμε με διγραμμική παρεμβολή το χρώμα των pixels στις προβολές των σημείων του πλέγματος στις εικόνες $\mathbf{q}(p, I)$.
- Υπολογίζουμε την ετεροσυσχέτιση C των $\mathbf{q}(p, I_1), \mathbf{q}(p, I_2)$. Η φωτομετρική διαφωνία $h(p, I_1, I_2) = 1 - C$.

Η προταθείσα μέθοδος είναι κατάλληλη για *Lambert* επιφάνειες. Άλλα ανακλαστικά φαινόμενα που μπορεί να εμφανίζονται καθώς και η παρουσία εμποδίων στις φωτογραφίες δημιουργούν *προβλήματα*. Για να τα αντιμετωπίσουμε *αγνοούμε* εικόνες με χαμηλή φωτομετρική συνάφεια. Αυτό επιτυγχάνεται με την χρήση ενός κατωφλιού α και τους ακόλουθους ορισμούς:

$$g^*(p) = \frac{1}{|V^*(p) \setminus R(p)|} \sum_{I \in |V^*(p) \setminus R(p)|} h(p, I, R(p)) \quad (3.7)$$

$$V^*(p) : \{I | I \in V(p), h(p, I, R(p)) < \alpha\} \quad (3.8)$$

Αφού ορίσαμε την συνάρτηση φωτομετρικής διαφωνίας g^* και το patch p , Θα σκιαγραφήσουμε τώρα την διαδικασία *βελτιστοποίησης* των patches. Στόχος της διαδικασίας αυτής είναι να καταλήξουμε σε patch με *χαμηλές* τιμές φωτομετρικής *διαφωνίας* g^* . Κάθε patch p ανακατασκευάζεται ξεχωριστά σε *δύο στάδια*

1. *Αρχικοποίηση* των $\mathbf{C}(p), \mathbf{n}(p), V^*(p), R(p)$.
2. *Βελτιστοποίηση* του p σε σχέση με τις *γεωμετρικές* παραμέτρους του $\mathbf{C}(p), \mathbf{n}(p)$.

Αναφέρουμε κάποια στοιχεία για την *παράσταση των γεωμετρικών χαρακτηριστικών* του p . Για το κέντρο $C(p)$, θεωρούμε ότι βρίσκεται σε συγκεκριμένη ακτίνα, ώστε η προβολή του σε κάποια εικόνα (ποια εικόνα θα αποσαφηνιστεί αργότερα) να παραμένει αμετάβλητη. Συνεπώς η περιγραφή διαθέτει *ένα (1) βαθμό ελευθερίας*. Για την διεύθυνση του διανύσματος $\mathbf{n}(p)$, χρησιμοποιούμε γωνίες *Euler* και επομένως έχουμε *τρεις (3) βαθμούς ελευθερίας*. Η βελτιστοποίηση γίνεται με την αριθμητική μέθοδο *conjugate gradients*. Μια καλή υλοποίηση της είναι διαθέσιμη στο [159].

Ένα από τα προβλήματα που πρέπει να αντιμετωπίσουμε είναι η *έλλειψη πληροφορίας σύνδεσης* των patch p . Για αυτό ορίζουμε *κελιά* $C_i(x, y)$ μεγέθους $\beta_1 \times \beta_1$ ($\beta_1 = 2$ στο [69]) σε κάθε εικόνα I_i . Τα κελιά αυτά δημιουργούν ένα πλέγμα που καλύπτει κάθε εικόνα. Κάθε patch p προβάλλεται στο σύνολο των εικόνων $V(p)$ στις οποίες είναι ορατό. Τα κελιά $C_i(x, y)$ θυμούνται το *σύνολο των patch που προβάλλονται σε αυτά*. Το σύνολο των patch που προβάλλονται στο κελί $C_i(x, y)$ της εικόνας I_i θα συμβολίζεται με $Q_i(x, y)$. Αντίστοιχα ορίζουμε το σύνολο $Q_i^*(x, y)$ με χρήση των συνόλων V^* .

Στις επόμενες ενότητες θα εξηγήσουμε αναλυτικά το σύστημα PMVS και τον τρόπο χρήσης των οντοτήτων που ορίσαμε προηγουμένως.

3.3.2 Περιγραφή Αλγορίθμου- Ανακατασκευή patch

Το σύστημα ανακατασκευής patch προσπαθεί να ανακατασκευάσει τουλάχιστον *ένα patch ανά κελί εικόνας* $C_i(x, y)$. Όπως αναφέραμε και στην εισαγωγική Ενότητα 3.3.1, τα τρία στάδια της ανακατασκευής είναι το ταίριασμα των χαρακτηριστικών, η επέκταση και η διαλογή. Το αρχικό ταίριασμα χαρακτηριστικών μας δίνει ένα αραιό σύνολο patch, ενώ τα δύο επόμενα στάδια με την επανάληψη τους n φορές μας δίνουν ένα πυκνό σύνολο patch από το οποίο έχουν αφαιρεθεί εσφαλμένα ταίριασματα. Αναλύουμε τώρα διεξοδικά τα στάδια αυτά.

Ταίριασμα χαρακτηριστικών

Αρχικά *ανιχνεύουμε* χαρακτηριστικά Harris και DoG στις εικόνες. Παραλείπουμε τις λεπτομέρειες (παραμετροποίηση ανιχνευτών) και παραπέμπουμε τον αναγνώστη στο [69] για την περισσότερες πληροφορίες. Για να εξασφαλίσουμε *ομοιόμορφη κάλυψη* των εικόνων από ανιχνευθέντα χαρακτηριστικά, εφαρμόζουμε την ακόλουθη μέθοδο. Καλύπτουμε *κάθε εικόνα* με πλέγματα αποτελούμενα από *κελιά διαστάσεων* $\beta_2 \times \beta_2$, αντίστοιχα με τα κελιά $C_i(x, y)$ που χρησιμοποιούμε στην διαδικασία βελτιστοποίησης και που αναλύσαμε προηγουμένως. Σε *κάθε κελί*, κρατάμε η χαρακτηριστικά. Τα επιλεγμένα χαρακτηριστικά είναι τοπικά μέγιστα σε σχέση με τις αποκρίσεις τους στους τελεστές ανίχνευσης χαρακτηριστικών. Με αυτό τον τρόπο αποσκοπούμε στην επιλογή των πιο "δυνατών" και ξεκάθαρων ανιχνευθέντων χαρακτηριστικών. Στο [69] οι παράμετροι επιλέγονται $\beta_2 = 32, \eta = 4$.

Στην συνέχεια επιχειρείται το *ταίριασμα* των ανιχνευθέντων χαρακτηριστικών μεταξύ των εικόνων. Έστω εικόνα I_i και το οπτικό κέντρο της $\mathbf{O}(I_i)$. Για *κάθε ανιχνευθέν χαρακτηριστικό* f στην I_i βρίσκουμε το σύνολο F_f χαρακτηριστικών f' στις υπόλοιπες εικόνες που ταιριάζουν με το f . Τα χαρακτηριστικά f, f' είναι ίδιου τύπου (DoG ή Harris). Τα χαρακτηριστικά f' επιλέγονται με κριτήριο την *απόσταση τους από την επιπολική γραμμή* που ορίζει το f στην I' . Συγκεκριμένα πρέπει να απέχουν έως δύο (2) pixels από την επιπολική

ευθεία. Αποσαφηνίζουμε ότι συμβολίσαμε με I' την εικόνα στην οποία ανήκει το f' .

Τα ζευγάρια ταιριασμένων χαρακτηριστικών (f, f') χρησιμοποιούνται για να μας δώσουν με τριγωνισμό τα αντίστοιχα σημεία του χώρου $p_{ff'}$. Τα σημεία αυτά είναι υποψήφια για κέντρα patch. Τα διατάσσουμε με σειρά αύξουσας απόστασης από το οπτικό κέντρο $\mathbf{O}(I_i)$, και τα αξιολογούμε σε αυτή την σειρά. Η αξιολόγηση τους οδηγεί είτε στην δημιουργία patch με κέντρο το $p_{ff'}$, είτε στην απόρριψη του σημείου. Ας περιγράψουμε τώρα την διαδικασία της αξιολόγησης για την δημιουργία ενός patch p . Αρχικοποιούμε τις παραμέτρους του σύμφωνα με τις ακόλουθες σχέσεις

$$\mathbf{C}(p) \leftarrow p_{ff'} \quad (3.9)$$

$$\mathbf{n}(p) \leftarrow \frac{\overrightarrow{\mathbf{C}(p)\mathbf{O}(I_i)}}{|\overrightarrow{\mathbf{C}(p)\mathbf{O}(I_i)}|} \quad (3.10)$$

$$R(p) \leftarrow I_i \quad (3.11)$$

Ας υπολογίσουμε τώρα το σύνολο $V(p)$ των εικόνων στις οποίες το p είναι ορατό. Το κριτήριο που καθορίζει την ορατότητα του patch p στην εικόνα I είναι το ακόλουθο: υπολογίζουμε την γωνία δ μεταξύ του κάθετου διανύσματος $\mathbf{n}(p)$ και της διεύθυνσης από το patch p στο οπτικό κέντρο $\mathbf{O}(I)$. Αν $\delta < t$, τότε προσθέτουμε την I στο $V(p)$. Στο [69] χρησιμοποιείται $t = \pi/3 \text{ rads}$. Παρατηρούμε την στήριξη του προτεινόμενου κριτηρίου από την διαίσθηση μας, αναμένουμε συνεπώς αξιοπρεπείς εκτιμήσεις ορατότητας. Εκφράζουμε τώρα το κριτήριο ορατότητας με μαθηματικό συμβολισμό και έχουμε την σχέση

$$V(p) \leftarrow \left\{ I : \mathbf{n}(p) \frac{\overrightarrow{\mathbf{C}(p)\mathbf{O}(I)}}{|\overrightarrow{\mathbf{C}(p)\mathbf{O}(I)}|} < \cos(t) \right\} \quad (3.12)$$

Για να καθορίσουμε και το αντίστοιχο σύνολο $V^*(p)$, χρησιμοποιούμε την Σχέση 3.8.

Εκτελούμε στην συνέχεια για το patch p την διαδικασία βελτιστοποίησης που περιγράψαμε στην Ενότητα 3.3.1. Κατά την διάρκεια της βελτιστοποίησης, περιορίζουμε το $\mathbf{C}(p)$ ώστε να ανήκει σε κατάλληλη ακτίνα που διατηρεί την προβολή του στην εικόνα $R(p)$ αμετάβλητη. Παίρνουμε έτσι μία νέα εκλεπτυσμένη γεωμετρία του patch p την οποία χρησιμοποιούμε για να υπολογίσουμε ξανά τα σύνολα $V(p)$, $V^*(p)$ από τις σχέσεις 3.12, 3.8.

Για να κατασκευάσουμε το υπό εξέταση patch p στηριζόμενα στην πληθικότητα του συνόλου των εικόνων στις οποίες το p είναι αξιόπιστα ορατό. Χρησιμοποιούμε δηλαδή το σύνολο $V^*(p)$ και κατώφλι γ . Το κριτήριο της επιλογής δίνεται με την ακόλουθη σχέση

$$|V^*(p)| \leq \gamma \quad (3.13)$$

Σε περίπτωση επιτυχούς (σύμφωνα με την 3.13) ανακατασκευής του p , το αποθηκεύουμε στα κατάλληλα σύνολα Q, Q^* των αντίστοιχων εικόνων.

Όσον αφορά την τιμή του κατωφλίου α στην Εξίσωση 3.8 χρησιμοποιούμε τιμές 0.6, 0.3 πριν και μετά την βελτιστοποίηση αντίστοιχα ([69]). Για να επιταχύνουμε τους υπολογισμούς, στα κελιά των εικόνων που ανακατασκευάζουμε ένα patch, διαγράφουμε τα υπόλοιπα χαρακτηριστικά και δεν τα χρησιμοποιούμε περαιτέρω.

Επέκταση περιοχών

Στόχος του σταδίου επέκτασης περιοχών είναι η ανακατασκευή τουλάχιστον ενός patch σε κάθε κελί $C_i(x, y)$ κάθε εικόνας I_i . Για κάθε υπάρχον patch p καθορίζουμε ένα σύνολο $C(p)$ από γειτονικά κελιά εικόνων που ικανοποιούν συγκεκριμένα κριτήρια. Στην συνέχεια εφαρμόζουμε την διαδικασία επέκτασης σε κάθε σύνολο $C(p)$. Τα κελιά στα οποία θα επεκταθούμε καθορίζονται από τις γειτονίες $C(p)$ των κατασκευασμένων patch p .

Ας δούμε πως κατασκευάζεται η γειτονιά $C(p)$ για κάποιο patch p . Αρχικοποιούμε την $C(p)$ με την συλλογή των κελιών εικόνων, από όλες τις εικόνες στις οποίες το p είναι ορατό, που ικανοποιούν την σχέση

$$C(p) = \{C_i(x', y') | p \in Q_i(x, y), |x - x'| + |y - y'| = 1\} \quad (3.14)$$

Ας σχολιάσουμε σύντομα την 3.14. Για να χαρακτηρίσουμε τυχόν κελί εικόνας γειτονικό του patch p ζητάμε την ικανοποίηση δύο συνθηκών: μιας συνθήκης ορατότητας και μιας συνθήκης γεωμετρικής εγγύτητας στην αντίστοιχη προβολή του p στην υπό εξέταση εικόνα.

Διαμορφώνουμε στην συνέχεια την γειτονιά $C(p)$ ώστε να αφαιρεθούν από αυτή κελιά στα οποία η επέκταση κρίνεται περιττή. Συγκεκριμένα κελιά αφαιρούνται από την $C(p)$ όταν

1. Στο κελί $C_i(x', y')$ έχει ανακατασκευαστεί patch p' γειτονικό του p .
2. Υπάρχει μια ασυνέχεια στο βάθος.

Σε αυτό το σημείο είναι απαραίτητο να αποσαφηνίσουμε κάποια σημεία. Αρχικά, να ορίσουμε την έννοια της γειτνίασης για τα patch, την οποία χρησιμοποιήσαμε στο 1. Λέμε ότι το patch p' είναι γειτονικό του p αν ισχύει η σχέση

$$|(\mathbf{C}(p) - \mathbf{C}(p')) \mathbf{n}(p)| + |(\mathbf{C}(p) - \mathbf{C}(p')) \mathbf{n}(p')| < 2\rho_1 \quad (3.15)$$

Η Εξίσωση 3.15 εκφράζει την απαίτηση για χωρική εγγύτητα των γειτονικών patch p, p' . Η εγγύτητα των patches ποσοτικοποιείται με την απόσταση των κέντρων τους στις διευθύνσεις των κάθετων διανυσμάτων $\mathbf{n}(p), \mathbf{n}(p')$. Η τιμή του κατωφλιού ρ_1 επιλέγεται ώστε να περιγράφει μετατόπιση β_1 pixels στην εικόνα αναφοράς $R(p)$.

Ακόμα, σχετικά με την συνθήκη που διατυπώθηκε στο 2, σχολιάζουμε ότι θα μπορούσε να παραληφθεί. Η διαδικασία διαλογής που ακολουθεί αφαιρεί τα εσφαλμένα patches. Η προσθήκη της όμως, βελτιώνει την υπολογιστική πολυπλοκότητα του σταδίου επέκτασης και κατ'επέκταση του συστήματος γενικότερα.

Προχωράμε τώρα στην περιγραφή της διαδικασίας επέκτασης, αφού τελειώσαμε με το πεδίο εφαρμογής της διαδικασίας. Για κάθε κελί $C_i(x', y')$ που ανήκει στην γειτονιά $C(p)$ ακολουθούμε την επόμενη διαδικασία. Αρχικοποιούμε τα χαρακτηριστικά του νέου patch p' , $\mathbf{n}(p')$, $R(p')$, $V(p')$, με τις αντίστοιχες τιμές για το patch p . Το κέντρο του patch $\mathbf{C}(p')$ βρίσκεται από την τομή της οπτικής ακτίνας που διέρχεται από το $\mathbf{O}(I_i)$ και το κέντρο του $C_i(x', y')$ με το επίπεδο που περιέχει το patch p .

Για να υπολογίσουμε το σύνολο $V^*(p')$ χρησιμοποιούμε την Εξίσωση 3.8 και θέτουμε κατώφλι $\alpha = 0.6$. Στην συνέχεια εκτελούμε την γνωστή διαδικασία βελτιστοποίησης ως προς τα $\mathbf{n}(p')$, $\mathbf{C}(p')$. Προσοχή απαιτείται στον περιορισμό ως προς την ακτίνα στην οποία κείται το $\mathbf{C}(p')$ στην συγκεκριμένη βελτιστοποίηση. Χρησιμοποιούμε στον περιορισμό την εικόνα I_i , γεγονός που μας

εξασφαλίζει ότι η προβολή του p' στην I_i θα παραμείνει στο κελί $C_i(x, y)$. Αφού ολοκληρωθεί η διαδικασία βελτιστοποίησης, *ανανεώνουμε το $V(p')$ προσθέτοντας* σε αυτό εικόνες στις οποίες το p' θα έπρεπε να είναι ορατό. Η προσθήκη γίνεται χρησιμοποιώντας *χάρτες βάθους*. Βάθος υπολογίζεται για κάθε κελί και όχι για κάθε pixel όπως συνηθίζεται. Στην συνέχεια υπολογίζουμε και το ανανεωμένο $V^*(p')$ κατά τα γνωστά, με χρήση της Εξίσωσης 3.8 και $\alpha = 0.3$.

Η απόφαση να προσθέσουμε νέες εικόνες στο $V(p')$ μετά την βελτιστοποίηση και όχι να αντικαταστήσουμε πλήρως τα περιεχόμενα του $V(p')$ οδηγεί σε *ασυνέπεια*, όσον αφορά την πληροφορία ορατότητας, μεταξύ των *διαφορετικών patch*. Το γεγονός αυτό θα *αξιοποιηθεί* στην διαδικασία *διαλογής*, για αυτό επιλέγεται η συγκεκριμένη προσέγγιση.

Τέλος χρησιμοποιούμε, σε πλήρη αντιστοιχία με την αρχική διαδικασία ανακατασκευής patch, την Σχέση 3.13 και ανανεώνουμε τα κατάλληλα σύνολα Q, Q^* .

Σημειώνουμε μια ακόμα λεπτομέρεια, σχετική με την εκτέλεση του σταδίου. Οι *τιμές κατωφλίου α* της Σχέσης 3.8 πριν και μετά την βελτιστοποίηση, χαλαρώνονται (*αυξάνουν*) σε κάθε επανάληψη του σταδίου κατά 0.2 προκειμένου να αντιμετωπιστούν δύσκολες (χωρίς υφή ή ομογενείς) περιοχές σε επόμενες επαναλήψεις.

Διαλογή (φιλτράρισμα)

Στην διαδικασία διαλογής στοχεύουμε στην αφαίρεση εσφαλμένων patch. Τρία (3) φίλτρα χρησιμοποιούνται για τον παραπάνω σκοπό:

1. Το πρώτο φίλτρο στοχεύει στην επιβολή συνέπειας της πληροφορίας ορατότητας. Έστω patch p και σύνολο $U(p)$ που περιέχει τα patch p' που είναι ασυνεπή σε Σχέση με το p . Εφαρμόζουμε αλγεβρικό κριτήριο (θα παρουσιαστεί ακολούθως) για την απόρριψη του p .
2. Υπολογίζουμε τον αριθμό των εικόνων στις οποίες το p είναι ορατό, σύμφωνα με κριτήριο που βασίζεται στους χάρτες βάθους. Αν είναι ορατό σε λιγότερες από γ εικόνες, απορρίπτουμε το p . Το κριτήριο αποσκοπεί στην επιβολή συνέπειας της πληροφορίας ορατότητας, όμοια με το 1.
3. Συλλέγουμε για το patch p το σύνολο των patch που βρίσκονται στο ίδιο και σε γειτονικά κελιά με το κελί του p . Αν το p είναι γείτονας με λιγότερο του 25% από αυτά τα patch, απορρίπτεται.

Ας αποσαφηνίσουμε τώρα κάποιες από τις έννοιες που εισήχθησαν στην περιγραφή του φίλτρου 1. Ξεκινάμε με το σύνολο $U(p)$. Τα στοιχεία του καθορίζονται ως εξής: $p' \in U(p)$ αν p, p' δεν είναι γείτονες (Σχέση 3.14) αλλά είναι και τα δύο αποθηκευμένα σε κάποιο κελί στο οποίο το p είναι ορατό. Η σχέση που καθορίζει την απόρριψη του p δίνεται

$$|V^*(p)|(1 - g^*(p)) < \sum_{p_i \in U(p)} 1 - g^*(p_i) \quad (3.16)$$

Σε περιπτώσεις που το p έχει ανακατασκευαστεί λάθος, περιμένουμε να έχει μεγάλη φωτομετρική διαφωνία $g^*(p)$ και μικρό πληθικό σύνολο $V^*(p)$ και συνεπώς να απορριφθεί από την Σχέση 3.16.

3.3.3 Περιγραφή Αλγορίθμου-Πλέγμα (mesh)

Μέχρι στιγμής έχουμε ανακατασκευάσει μικρά προσανατολισμένα παραλληλόγραμμα στο χώρο- patches. Αν και ο συγκεκριμένος τρόπος αναπαράστασης επιφανειών παρουσιάζει αύξουσα δημοτικότητα, είναι επιθυμητό να μετατρέψουμε την περιγραφή σε πολυγωνικό πλέγμα.

Κατασκευή πλέγματος

Για την κατασκευή του πλέγματος, μπορούμε να χρησιμοποιήσουμε κάποιο διαθέσιμο λογισμικό που χρησιμοποιεί για την παραγωγή του πλέγματος την μέθοδο Poisson [110]. Απορρίπτουμε εσφαλμένα κομμάτια του πλέγματος, με την διαγραφή τριγώνων με μήκος πλευρών l , $l > 6\bar{l}$ όπου \bar{l} το μέσο μήκος. Αν είναι διαθέσιμη πληροφορία προσκηνίου/παρασκηνίου, μπορούμε να εφαρμόσουμε μεθόδους που βασίζονται στην επαναληπτική παραμόρφωση μιας επιφάνειας και χρησιμοποιούν το visual hull [122]. Μια τέτοια μέθοδος περιγράφεται στο [67].

Εκλέπτυνση αναπαράστασης πλέγματος με μεθόδους MVS

Στο στάδιο αυτό προσπαθούμε να βελτιστοποιήσουμε τις θέσεις των κορυφών του πλέγματος, σε σχέση με ορούς ενέργειας που εκφράζουν την φωτομετρική συμφωνία και την γεωμετρική ομαλότητα της αναπαράστασης. Οι όροι ενέργειας είναι αθροίσματα των αντίστοιχων ενεργειών των κόμβων. Η βελτιστοποίηση γίνεται με την αριθμητική μέθοδο των conjugate gradients. Η βασική ιδέα της μεθόδου μπορεί να αναζητηθεί στο [240]. Οι διαφορές στην προσέγγιση του [69] είναι η βελτιστοποίηση ως προς το βάθος και το κάθετο διάνυσμα στις κορυφές και η χρήση Γκαουσιανής συνάρτησης για τον συνδυασμό των διαφορών υποθέσεων. Για τον ακριβή ορισμό και υπολογισμό των συναρτήσεων ενέργειας παραπέμπουμε τον αναγνώστη στο [69].

3.3.4 Ενδεικτικά αποτελέσματα

Ανακατασκευασμένα μοντέλα αντικειμένων και εξωτερικών σκηνών παρουσιάζονται στις εικόνες 3.6, 3.8. Εικόνες από τα αντίστοιχα σύνολα δεδομένων παρουσιάζονται στις Εικόνες 3.5 και 3.7. Παρατηρούμε σαν *ελάττωμα* της μεθόδου την ύπαρξη *κενών* στο ανακατασκευασμένο μοντέλο, ειδικά σε περιοχές χωρίς έντονη υφή.

Για την *ποιοτική* αξιολόγηση των μεθόδων χρησιμοποιούμε δύο δείκτες

1. Την *ακρίβεια* του μοντέλου. Η ακρίβεια ορίζεται ως η απόσταση d ώστε ένα δεδομένο ποσοστό της ανακατασκευής να βρίσκεται σε απόσταση d από το αληθές μοντέλο.
2. Την *πληρότητα* της ανακατασκευής. Ορίζεται ως το ποσοστό του αληθούς μοντέλου που βρίσκεται εντός δεδομένης απόστασης από το ανακατασκευασμένο μοντέλο.

Οι δείκτες 1,2 ορίζονται στο [193]. Χρησιμοποιώντας αυτά τα κριτήρια, ο αλγόριθμος PMVS παρουσιάζει κορυφαία επίδοση στα σύνολα δεδομένων του [193]. Αντίστοιχα αποτελέσματα έχουμε και για τα σύνολα δεδομένων του [216]. Στο [216] τα σύνολα δεδομένα είναι εξωτερικών σκηνών ενώ, στο [193] είναι αντικείμενα.

| Περιγραφή παραμέτρου | Σύμβολο | Τιμή |
|---|-----------|------------------------------------|
| Επαναλήψεις επέκτασης και διαλογής | n | 3 |
| Έκταση patch | μ | 5 ή 7 pixels |
| Κατώφλι φωτομετρικής συνάφειας | α | Μεταβλητό ανά στάδιο και επανάληψη |
| Μέγεθος κελιών στις εικόνες | β_1 | 2 pixels |
| Μέγεθος κελιών στις εικόνες στην αρχική διαδικασία ταιριάσματος χαρακτηριστικών | β_2 | 32 pixels |
| Αριθμός χαρακτηριστικών ανά κελί | η | 4 |
| Κατώφλι γωνίας για εκτίμηση ορατότητας | t | $\pi/3$ rads |
| Κατώφλι πληθικότητας γειτονίας V^* | γ | 2 ή 3 |

Πίνακας 3.1: Παράμετροι PMVS και τιμές τους. Οι τιμές είναι από [69].

Οι χρόνοι εκτέλεσης της μεθόδου παρουσιάζονται στο 3.2. Παρατηρούμε ότι η κύρια υπολογιστική επιβάρυνση είναι η εκλέπτυνση του πολυγωνικού πλέγματος (mesh). Στις περιπτώσεις που δεν χρησιμοποιήθηκε η μέθοδος Poisson Surface Reconstruction για την αρχικοποίηση του πλέγματος, σημαντική επιβάρυνση εμφανίζεται και σε αυτό το στάδιο. Περισσότερα αποτελέσματα μπορούν να αναζητηθούν στο [69] και στους διαδικτυακούς τόπους [2, 1, 3].



Σχήμα 3.5: Φωτογραφίες από σύνολα δεδομένων. Τα αντίστοιχα σύνολα δεδομένων δίνονται ακολούθως. Από αριστερά προς τα δεξιά: roman, temple, dino, skull, face1, face2. Από [69]

| Σύνολο δεδομένων | Κατασκευή patch | Αρχικοποίηση πλέγματος | Εκτέλεση πλέγματος |
|------------------|-----------------|------------------------|--------------------|
| roman | 20 | 149 | 91 |
| temple | 3 | 89 | 49 |
| dino | 2 | 97 | 69 |
| skull | 6 | 120 | 46 |
| face1 | 9 | 123 | 56 |
| face2 | 4 | 3 | 36 |
| castle | 36 | 3 | 278 |

Πίνακας 3.2: Χρόνος εκτέλεσης σε λεπτά επιμέρους σταδίων PMVS. Οι προσομοιώσεις έγιναν σε επεξεργαστή Dual Xeon 2.66Ghz. Από [69].



Σχήμα 3.6: Ανακατασκευασμένα πολυγωνικά πλέγματα. Τα αντίστοιχα σύνολα δεδομένων δίνονται ακολούθως. Από αριστερά προς τα δεξιά, πρώτη σειρά: roman, temple, dino. Τελευταία σειρά: skull, face1, face2. Από [69]



Σχήμα 3.7: Εικόνες από το σύνολο δεδομένων castle. Από [69]



Σχήμα 3.8: Ανακατασκευασμένο πολυγωνικό πλέγμα για σύνολο δεδομένων castle. Από [69]

3.3.5 CMVS: Ένα σύστημα αποσύνθεσης του προβλήματος ανακατασκευής σε μικρότερα υποπροβλήματα

Η ενότητα αυτή βασίζεται στην εργασία [65]. Η υλοποίηση είναι διαθέσιμη από τον διαδικτυακό τόπο [64].

Η βασική *ιδέα* της ενότητας είναι η *αποσύνθεση* της συλλογής φωτογραφιών που έχουμε ως δεδομένα σε *επικαλυπτόμενα υποσύνολα διαχειρίσιμου μεγέθους*. Το πρόβλημα διατυπώνεται ως πρόβλημα βελτιστοποίησης με περιορισμούς και λύνεται επαναληπτικά. Ο αλγόριθμος συγχώνευσης των λύσεων των υποπροβλημάτων είναι παράλληλος και χρησιμοποιεί βήματα διαλογής (φιλτραρίσματος) προκειμένου να εξαλείψει κακές ανακατασκευές και να επιβάλλει καθολικούς περιορισμούς ορατότητας.

Από τις *προκλήσεις* που πρέπει να αντιμετωπιστούν στον τομέα της τρισδιάστατης ανακατασκευής, η μέθοδος CMVS προσπαθεί να αντιμετωπίσει αυτό της *κλιμάκωσης*. Άλλες προκλήσεις είναι για παράδειγμα η μοντελοποίηση περίπλοκων συναρτήσεων BDRF και η αντιμετώπιση μεταβολών στον φωτισμό.

Για την διαδικασία της *ανακατασκευής* της τρισδιάστατης γεωμετρίας, *αξιόπιστες* μέθοδοι ανίχνευσης χαρακτηριστικών([133]) και ανάκτησης των παραμέτρων προβολής (SfM) έχουν διατυπωθεί. Οι διαδικασίες αυτές μπορούν επιπρόσθετα να εκτελεστούν παράλληλα και αποδοτικά[249, 6, 58]. Ταυτόχρονα, για το τελευταίο στάδιο της ανακατασκευής, το στάδιο εκτέλεσης κάποιου αλγορίθμου MVS, έχουν προταθεί μέθοδοι οι οποίες θεωρούνται ικανοποιητικά αξιόπιστες και στηρίζονται στα ακριβή αποτελέσματα ανίχνευσης χαρακτηριστικών και βαθμονόμησης όψεων[71, 78, 99, 177].

Οι περισσότεροι αλγόριθμοι MVS βασίζονται στην ιδέα της συσχέτισης μετρήσεων από πολλαπλές εικόνες για να εξάγουν πληροφορία για τις τρισδιάστατες επιφάνειες. Παραδείγματα αλγορίθμων που χρησιμοποιούν όλες τις εικόνες ταυτόχρονα είναι τα [69, 92, 180, 240]. Η *προταθείσα μέθοδος* οργανώνει τις εικόνες σε διαχειρίσιμες ομάδες (clusters) και εφαρμόζει σε καθένα ανεξάρτητα κάποιον αλγόριθμο MVS. Έτσι η *ευελιξία* είναι μεγάλη ως προς την *επιλογή* ενός συγκεκριμένου *αλγορίθμου MVS*, αφού πολλοί είναι κατάλληλοι.

Συνοψίζουμε την διαδικασία που ακολουθείται, σκιαγραφώντας τα κύρια σημεία της

- Λύση προβλήματος *ομαδοποίησης* (clustering) των εικόνων εισόδου σε επικαλυπτόμενες ομάδες[16]. Αν δεν χρησιμοποιηθούν επικαλυπτόμενα clusters, η ποιότητα της ανακατασκευής χειροτερεύει στα σύνορα των clusters.
- *Ανακατασκευή* κάθε cluster χωριστά και *συγχώνευση* αποτελεσμάτων.
- Χρήση σθεναρών μεθόδων *διαλογής* (φίλτρων) για τον χειρισμό λαθών ανακατασκευής και της διαφοροποίησης της ποιότητας ανακατασκευής στο κέντρο και στα σύνορα των clusters.

Clustering όψεων

Σε αυτό το στάδιο στοχεύουμε στη δημιουργία *άγνωστου* αριθμού *επικαλυπτόμενων clusters διαχειρίσιμου* μεγέθους. Απαιτείται κάθε σημείο του χώρου που έχουμε στην διάθεση μας από προηγούμενη εκτέλεση κατάλληλου

συστήματος SfM να μπορεί να ανακατασκευαστεί με ακρίβεια από ένα τουλάχιστον cluster.

Με την διατύπωση του προβλήματος πετυχαίνουμε τρία χαρακτηριστικά

1. Απόρριψη πλεοναζόντων εικόνων
2. Μικρά clusters
3. Ελάχιστη απώλεια περιεχομένου

Το **1** καθιστά την αναπαράσταση συμπαγή. Η ιδιότητα της συμπαγείας είναι απαραίτητη για την υπολογιστική αποδοτικότητα της μεθόδου. Όμως βελτιώνει και την ακρίβεια της ανακατασκευής. Αυτό γιατί σε μια μεγάλη συλλογή φωτογραφιών, όπως για παράδειγμα μία που προέρχεται από το internet, υπάρχουν πολλές σχεδόν όμοιες εικόνες και με μικρή διαφορά στις θέσεις λήψεις των φωτογραφιών. Χρήση τέτοιων φωτογραφιών οδηγεί σε φτωχές ανακατασκευές.

Το **3** διατυπώθηκε αρκετά γενικά και θα πρέπει να το εξηγήσουμε. Αναφερόμαστε σε αυτή την ιδιότητα ως κάλυψη. Για την κάλυψη, εξετάζουμε δύο πεδία

1. Τα σημεία του χώρου που προέκυψαν από το SfM
2. Τις εικόνες του συνόλου δεδομένων

Για ένα σημείο P_j του χώρου, λέμε ότι έχει καλυφθεί αν (όπως προαναφέραμε) ανακατασκευάζεται καλά από τουλάχιστον ένα cluster. Για να ποσοτικοποιήσουμε την έννοια της καλής ανακατασκευής, χρησιμοποιούμε το επόμενο αριθμητικό κριτήριο

$$\max_k f(\mathbf{P}_j, C_k \cap V_j) \geq \lambda f(\mathbf{P}_j, V_j) \quad (3.17)$$

Τα σύμβολα και οι συναρτήσεις που χρησιμοποιήσαμε στην Σχέση 3.17 επεξηγούνται ως

$$\begin{aligned} C_k &\leftarrow \text{Σύνολο εικόνων που ανήκουν στο cluster } k \\ \mathbf{P}_j &\leftarrow \text{Σημείο του χώρου από SfM που} \\ V_j &\leftarrow \text{Σύνολο εικόνων στις οποίες το σημείο } P_j \text{ είναι ορατό} \\ f(\mathbf{P}, A) &\leftarrow \text{Ακρίβεια της ανακατασκευής στην τρισδιάστατη} \quad (3.18) \\ &\text{τοποθεσία } \mathbf{P} \text{ από το σύνολο εικόνων } C \end{aligned}$$

Δεν θα αναφερθούμε σε λεπτομέρειες σχετικά με την συνάρτηση 3.18. Η περιγραφή της μπορεί να βρεθεί στο [65].

Η τελική συνθήκη που ενσωματώνουμε στην διατύπωση του προβλήματος για να επιβάλουμε την κάλυψη αφορά τις εικόνες του συνόλου δεδομένων. Συγκεκριμένα, για κάθε εικόνα I_i του συνόλου δεδομένων απαιτούμε ο λόγος κάλυψης (σύμφωνα με την Σχέση 3.17) των σημείων που είναι ορατά στην εικόνα I_i να είναι τουλάχιστον δ . Διατυπώνουμε μαθηματικά το κριτήριο με την σχέση

$$\frac{\text{Καλυμμένα σημεία στην } I_i}{\text{Σημεία στην } I_i} \geq \delta, \forall i \quad (3.19)$$

Η τελική διατύπωση ως πρόβλημα βελτιστοποίησης με περιορισμούς δίνεται

$$\begin{aligned}
 & \text{minimize} && \sum_k |C_k| && \text{(Συμπάγεια)} \\
 & \text{s.t.} && \forall k |C_k| \leq \alpha && \text{(Μέγεθος)} \\
 & && \forall i \frac{\text{Καλυμμένα σημεία στην } I_i}{\text{Σημεία στην } I_i} \geq \delta && \text{(Κάλυψη)}
 \end{aligned} \tag{3.20}$$

Το πρόβλημα 3.20 ικανοποιεί τα σημεία (1,2,3) και ταυτόχρονα ενσωματώνει *ποιοτικά* κριτήρια. Εικόνες χαμηλής ποιότητας έχουν λιγότερα σημεία SfM και συνεπώς η ενσωμάτωσή τους σε Clusters "κοστίζει" περισσότερο.

Επίλυση προβλήματος clustering

Το πρόβλημα που διατυπώσαμε (3.20) διαθέτει την επιθυμητή *εκφραστικότητα*, η επίλυση του όμως *δεν είναι εύκολη*. Για παράδειγμα, δεν είναι σε μορφή που μπορεί να επιλυθεί με ευρύτητα διατυπωμένες τεχνικές όπως ο αλγόριθμος k-means ή η μέθοδος normalised cuts[198].

Η *προσέγγιση* που ακολουθούμε είναι *υποβέλτιστη*. Η υστέρηση αυτή στην συμπάγεια της λύσης, δεν είναι σημαντική για την εφαρμογή του αλγορίθμου. Πριν αναλύσουμε τα βήματα της επίλυσης, εισάγουμε τις *έννοιες* της *γεινίασης εικόνων*, *συνόλων εικόνων* και *σημείων*, για να διευκολυνθούμε στην παρουσίαση. Για δύο εικόνες I_j, I_k θα λέμε ότι γειτονεύουν αν υπάρχει *σημείο SfM ορατό* και στις δύο εικόνες. Αντίστοιχα δύο σύνολα εικόνων γειτονεύουν, αν υπάρχουν *εικόνες* στα σύνολα αυτά που *γειτονεύουν*. Τέλος δύο *σημεία SfM* P_i, P_j γειτονεύουν αν

1. Έχουν παρόμοια σύνολα εικόνων στα οποία είναι ορατά. Συγκεκριμένα απαιτούμε *γεινίαση* των V_i, V_j και
2. Οι προβολές των P_i, P_j σε κάθε εικόνα στο $V_i \cup V_j$ να *απέχουν το πολύ* τ_1 pixels.

Προχωράμε τώρα στην περιγραφή της διαδικασίας επίλυσης του 3.20. Στο *πρώτο* στάδιο, κάνουμε διαλογή των σημείων SfM P_i . Πιο αναλυτικά, *συγχωνεύουμε* σημεία SfM που γειτονεύουν. Το σημείο που προκύπτει από την συγχώνευση δύο σημείων P_i, P_j τοποθετείται χωρικά στην μέση θέση των συγχωνευθέντων σημείων. Το στάδιο αυτό μειώνει τον αριθμό των ανακατασκευασμένων σημείων στον χώρο και ελαφρύνει *υπολογιστικά* τα επόμενα στάδια της διαδικασίας ανακατασκευής. Ταυτόχρονα η πληροφορία ορατότητας των σημείων, κωδικοποιημένη όπως είναι στα σύνολα V_i , γίνεται πιο *αξιόπιστη*. Το σύνολο ορατότητας του σημείου

$$P_M \leftarrow \text{Συγχώνευση } P_i, P_j$$

ορίζεται $V_M = V_i \cup V_j$.

Στο *δεύτερο* στάδιο, αφαιρούμε εικόνες. Μια εικόνα αφαιρείται αν η αφαίρεση της δεν παραβιάζει τον περιορισμό κάλυψης. Οι εικόνες χαμηλής ανάλυσης αφαιρούνται πρώτες.

Στο *τρίτο* στάδιο, *διαιρούμε clusters* τα οποία παραβιάζουν τον σχετικό περιορισμό *μεγέδους*. Η διαίρεση ενός cluster γίνεται με την μέθοδο *normalised cuts* στον γράφο ορατότητας των εικόνων. Οι κόμβοι του γράφου αντιστοιχούν σε εικόνες και οι ακμές μεταξύ των κόμβων εκφράζουν την συνεισφορά

των προσκείμενων κόμβων-εικόνων στην ανακατασκευή MVS. Εικόνες με μεγάλη συνεισφορά συνδέονται με "βαριές" ακμές και είναι λιγότερο πιθανό να κοπούν. Πιο αναλυτικά, το βάρος της ακμής e_{lm} που συνδέει τους κόμβους των εικόνων I_l, I_m στον γράφο ορατότητας δίνεται από την σχέση

$$e_{lm} = \sum_{P_j \in \Theta_{lm}} \frac{f(P_j, \{I_l, I_m\})}{f(P_j, V_j)}$$

$$\Theta_{lm} \leftarrow \{\text{Ορατά σημεία κοινά στις } I_l, I_m\}$$

Στο *τέταρτο* στάδιο, επιβάλλουμε τον περιορισμό *κάλυψης* ο οποίος ενδέχεται να *παραβιάστηκε* από την διαίρεση των clusters στο *τρίτο* στάδιο. Προσθέτουμε για αυτόν το λόγο εικόνες στα clusters προκειμένου να καλύψουμε περισσότερα σημεία SfM. Για την προσθήκη των εικόνων

1. *Κατασκευάζεται* λίστα δράσεων ταξινομημένη ως προς την απόδοσή τους.
2. *Εκτελούνται* δράσεις πάνω από ένα κατώφλι αποδοτικότητας.

Τα βήματα **1** και **2** *επαναλαμβάνονται* μέχρι να ικανοποιηθεί ο περιορισμός κάλυψης. Ας τα δούμε όμως αναλυτικά.

Για να κατασκευάσουμε την λίστα (**1**), εξετάζουμε κάθε ακάλυπτο σημείο SfM. Έστω σημείο P_i . Βρίσκουμε για το P_i το cluster C^i με την μεγαλύτερη κάλυψη για το σημείο, δηλαδή

$$C^i = \arg \max_{C_k} f(P_i, C_k)$$

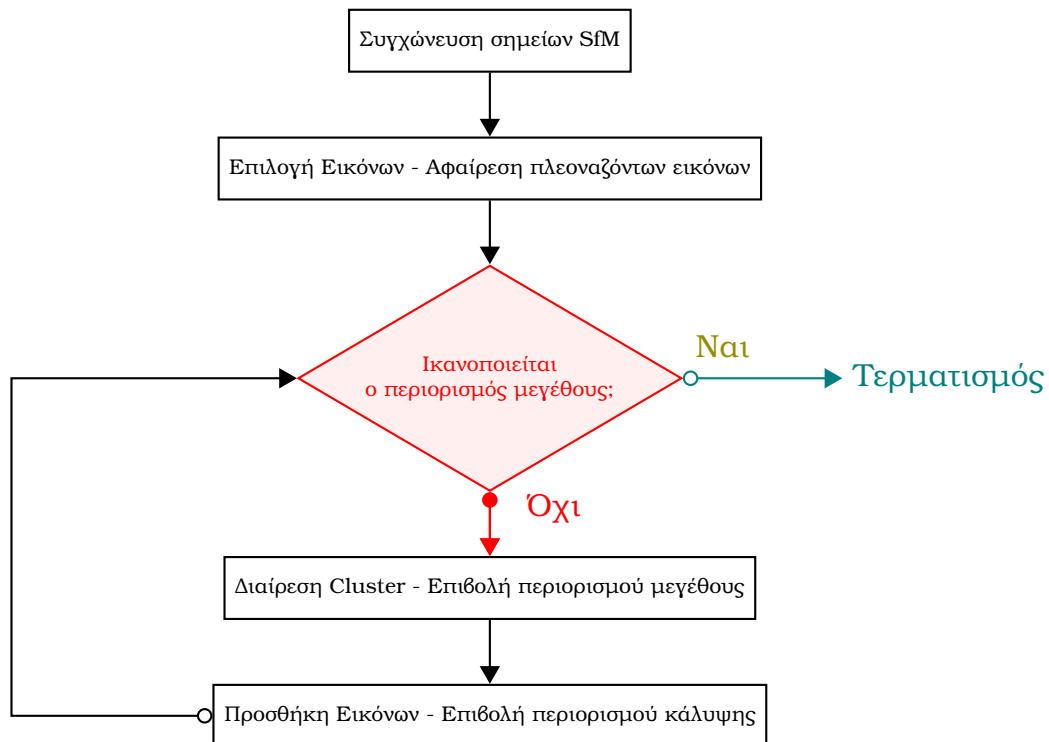
Στην συνέχεια προσθέτουμε δράσεις στην λίστα. Οι *δράσεις* είναι προσθήκες εικόνων I στο cluster C^i . Για κάθε δράση μετράμε την *απόδοση* g η οποία υπολογίζεται ως η διαφορά της ακρίβειας ανακατασκευής πριν και μετά την εκτέλεση της (προσθήκη εικόνας I)

$$g(I, C^i) = f(P_i, C^i \cup \{I\}) - f(P_i, C^i)$$

Τονίζουμε ότι για το σημείο P_i μόνο δράσεις σχετικές με το C^i εξετάζονται, επιλογή υποβέλτιστη αλλά υπολογιστικά αποδοτική. Εξετάζοντας όλα τα ακάλυπτα σημεία, όμοιες δράσεις ενδέχεται να προκύψουν. Στην περίπτωση αυτή συγχωνεύονται και η αποδοτικότητά τους προστίθεται.

Για την εκτέλεση των δράσεων (**2**), θα έπρεπε να υπολογίσουμε εκ νέου την αποδοτικότητα των δράσεων μετά την προσθήκη μιας εικόνας σε cluster (εκτέλεση κάποιας δράσης). Για λόγους υπολογιστικής πολυπλοκότητας, εκτελούμε όλες τις δράσεις με αποδοτικότητα πάνω από 0.7. Οι δράσεις εκτελούνται διαδοχικά και φροντίζουμε να αφαιρούμε από την λίστα αντικρουόμενες δράσεις. Δύο δράσεις A, A' θεωρούνται αντικρουόμενες αν οι εικόνες τους I, I' γειτονεύουν.

Με την ολοκλήρωση του *τέταρτου* σταδίου και τις προσθήκες εικόνων που κάναμε, ενδέχεται ο περιορισμός μεγέθους να έχει παραβιαστεί εκ νέου. Για αυτό το *τρίτο* και το *τέταρτο* στάδιο *επαναλαμβάνονται* μέχρι την ταυτόχρονη ικανοποίηση όλων των περιορισμών του προβλήματος.



Σχήμα 3.9: Διάγραμμα ροής του αλγορίθμου Clustering. Αναπαραγωγή από [65]

Διαλογή σημείων και ένωση λύσεων

Κάθε Cluster ανακατασκευάζεται ξεχωριστά με την χρήση αλγορίθμου MVS. Οι συγγραφείς του [65] χρησιμοποίησαν τον αλγόριθμο PMVS που αναλύσαμε προηγουμένως. Η επιλογή τους όμως δεν είναι δεσμευτική. Για να συγχωνεύσουμε τα ανακατασκευασμένα σημεία, χρησιμοποιούμε δύο φίλτρα. Τα φίλτρα αυτά μπορούν να παραλληλοποιηθούν.

Το πρώτο φίλτρο, προσπαθεί να διαχωρίσει τα σημεία που έχουν ανακατασκευαστεί στα όρια ενός cluster και είναι αναξιόπιστα από όσα ανακατασκευάστηκαν αξιόπιστα στο εσωτερικό του cluster. Τα σημεία στα όρια ενός cluster είναι αραιά, ενώ στο κέντρο του πυκνά. Η επόμενη διαδικασία επαναλαμβάνεται για κάθε cluster C_k . Έστω σημείο P_j ανακατασκευασμένο από το cluster αναφοράς C_k . Συλλέγουμε σημεία, ανακατασκευασμένα από όλα τα clusters, $\{\mathbf{Q}_m\}$ και την αντίστοιχη πληροφορία ορατότητας των σημείων $\{V_m\}$. Τα σημεία $\{\mathbf{Q}_m\}$ ικανοποιούν τις ακόλουθες δύο συνθήκες

1. Οι γωνίες που σχηματίζουν τα κάθετα διανύσματα στα σημεία \mathbf{Q}_m, P_j είναι μικρότερη από 90° .
2. Οι προβολές τους είναι σε απόσταση μικρότερη από τ_2 pixels από την προβολή του P_j , σε κάθε εικόνα στο V_j .

Από τα σημεία $\{\mathbf{Q}_m\}$ υπολογίζουμε το ιστόγραμμα $\{H_l\}$. Το H_l υπολογίζεται αθροίζοντας την ακρίβεια ανακατασκευής όλων των σημείων \mathbf{Q}_m που προέρχονται από το cluster C_l δηλαδή

$$H_l = \sum_{\mathbf{Q}_m \in C_l} f(\mathbf{Q}_m, V_m)$$

Το σημείο P_j απορρίπτεται όταν

$$H_k < 0.5 \max_l H_l$$

Το *δεύτερο* φίλτρο που εφαρμόζουμε επιβάλλει συνέπεια στην πληροφορία *ορατότητας*. Για κάθε σημείο MVS, μετράμε τις φορές c που έρχεται σε *αντίφαση* με ανακατασκευές από διαφορετικά clusters. Αν

$$c \geq 3 \tag{3.21}$$

το σημείο MVS απορρίπτεται.

Έστω cluster αναφοράς C_k και Θ_k το σύνολο των σημείων MVS που έχουν ανακατασκευαστεί από αυτό. Για κάθε εικόνα στο C_k φτιάχνουμε χάρτη βάθους προβάλλοντας τα σημεία του Θ_k στις εικόνες του C_k που είναι ορατά. Επίσης στους χάρτες βάθους αποθηκεύουμε, πέρα από την τιμή βάθους των σημείων, και την ακρίβεια ανακατασκευής.

Για κάθε σημείο MVS P που δεν ανήκει στο cluster αναφοράς μετράμε τον αριθμό των χαρτών βάθους του C_k που έρχονται σε αντίφαση με το P . Ένας χάρτης βάθους έρχεται σε αντίθεση με το P όταν

- Το σημείο και ο χάρτης βάθους διαφέρουν περισσότερο από ένα μικρό περιθώριο.
- Η ακρίβεια ανακατασκευής του P είναι λιγότερο από το μισό της αποθηκευμένης στον χάρτη τιμή ακρίβειας.

Η διαδικασία επαναλαμβάνεται αλλάζοντας το cluster αναφοράς. Για το σημείο P ο αριθμός των χαρτών που έρχονται σε αντίφαση με αυτό προστίθεται συνολικά (μεταξύ των εκτελέσεων με διαφορετικά clusters αναφοράς) και ελέγχεται προκειμένου να αποφασιστεί η διατήρηση ή η απόρριψη του σημείου (Σχέση 3.21).

Με την χρήση αυτού του φίλτρου, επιτυγχάνεται η συνέπεια μεταξύ των clusters. Στο εσωτερικό κάθε cluster η πληροφορία ορατότητας θεωρείται συνεπής. Για παράδειγμα υπενθυμίζουμε το φίλτρο ορατότητας PMVS που σε περίπτωση χρήσης του με το σύστημα CMVS εξασφαλίζει την συνεπή αναπαράσταση στο εσωτερικό κάθε cluster.

Ενδεικτικά Αποτελέσματα

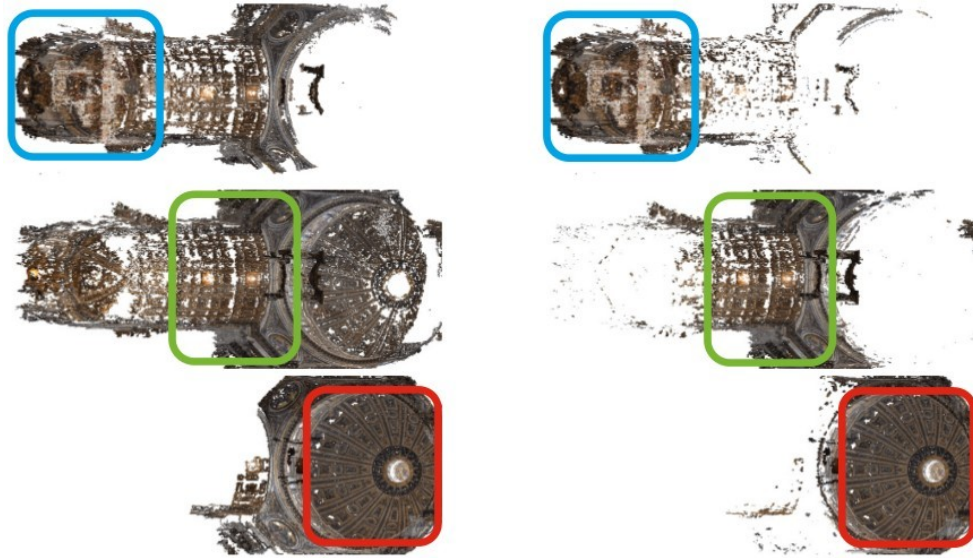
Για να παραστήσουμε τα σημεία MVS, πρέπει να διαλέξουμε κατάλληλο χρώμα για κάθε ένα από αυτά. Το μέσο χρώμα των προβολών ενός σημείου MVS στις εικόνες που είναι ορατό επιλέγεται για να συσχετιστεί με αυτό. Για την οπτικοποίηση των σημείων χρησιμοποιείται το λογισμικό `qsplat` [189].

Η ανάγκη για την χρήση επικαλυπτόμενων clusters παρουσιάζεται με την Εικόνα 3.11. Παρατηρούμε τα κενά στην ανακατασκευή με αμοιβαίως αποκλειόμενα clusters και την διόρθωση τους στην περίπτωση των επικαλυπτόμενων clusters.

Τα ανακατασκευασμένα σημεία των clusters με και χωρίς την χρήση των μεθόδων διαλογής που αναφέραμε παρουσιάζονται στην Εικόνα 3.10. Όπως φαίνεται και στον Πίνακα 3.4 η εφαρμογή των φίλτρων είναι υπολογιστικά απαιτητική και αποτελεί σημαντικό μέρος του συνολικού χρόνου εκτέλεσης. Το γεγονός αυτό οφείλεται κυρίως στο πρώτο φίλτρο που περιγράψαμε.

Ολοκληρώνουμε την περιγραφή του συστήματος CMVS με δύο παρατηρήσεις. Πρώτον, οι συγγραφείς του [65] χρησιμοποίησαν το σύστημα τους και σε

συνδυασμό με το σύστημα MVS των [78] που περιγράψαμε στην Ενότητα 3.2. Η βελτίωση στον χρόνο εκτέλεσης που αναφέρουν είναι σημαντική (υποδιπλασιασμός του), γεγονός που πιστοποιεί την ικανότητα γενικής εφαρμογής του συστήματος CMVS. Δεύτερον, η μείωση της παραμέτρου α που ελέγχει το μέγεθος των clusters οδηγεί στην αύξηση του αριθμού των clusters και αύξηση της επικάλυψής τους. Η μείωση της μέχρις ενός ορίου μειώνει τον παράλληλο χρόνο εκτέλεσης της ανακατασκευής.



Σχήμα 3.10: Επίδραση χρήσης μεθόδων διαλογής για την συγχώνευση των clusters. Τα παραλληλόγραμμα περικλείουν τους πυρήνες των clusters. Αριστερά: πριν την εφαρμογή φίλτρων. Δεξιά: μετά την εφαρμογή φίλτρων. Από [65]



Σχήμα 3.11: Επίδραση χρήσης επικαλυπτόμενων clusters. Για κάθε ζευγάρι, αριστερά έχουμε clusters χωρίς επικάλυψη και δεξιά με επικάλυψη. Από [65]

| Περιγραφή παραμέτρου | Σύμβολο | Τιμή |
|--|-----------|-----------|
| Κατώφλι στον ορισμό της καλής ανακατασκευής 3.17 | λ | 0.7 |
| Κατώφλι στον λόγο κάλυψης εικόνας 3.19 | δ | 0.7 |
| Κατώφλι απόστασης για γειτνίαση σημείων | τ_1 | 64 pixels |
| Κατώφλι απόστασης στην διαλογή σημείων $\{\mathcal{Q}_m\}$ | τ_2 | 8 pixels |
| Μέγεθος clusters 3.20 | α | 150 |

Πίνακας 3.3: Παράμετροι CMVS και τιμές τους. Οι τιμές είναι από [65].

| | |
|--|-----------------|
| Εικόνες | 1275 |
| Μετά την απόρριψη πλεοναζόντων εικόνων | 298 |
| Άθροισμα εικόνων στα clusters | 333 |
| Αριθμός clusters | 4 |
| Σημεία SfM | 515.259 |
| Μετά την συγχώνευση | 14.619 |
| Σημεία MVS | 6.344.656 |
| Μετά την εφαρμογή φίλτρων | 5.107.847 |
| Εκτέλεση clustering | 1.3 min. |
| Εκτέλεση MVS | 305(84) min. |
| Εκτέλεση φίλτρων | 39.9(10.2) min. |

Πίνακας 3.4: Στατιστικά συστήματος CMVS για το σύνολο δεδομένων Basilica. Οι αριθμοί σε παρένθεση είναι χρόνοι παράλληλης εκτέλεσης. Οι προσομοιώσεις έγιναν σε επεξεργαστή Dual Xeon 2.66Ghz. Από [65].

3.4 Τελευταίες τάσεις και εξελίξεις στον τομέα της στερέωσης πολλών όψεων

Στην ενότητα αυτή παρουσιάζουμε σύντομα μεθόδους ανακατασκευής που σε μεγάλο βαθμό στηρίζονται και επεκτείνουν τις ιδέες που αναφέραμε προηγουμένως. Ξεκινάμε παρουσιάζοντας τις μεθόδους αξιολόγησης τους και τα σύνολα αξιολόγησης (benchmarks) που έχουν καθιερωθεί, συνεισφέροντας σημαντικά στην εξέλιξη του τομέα.

3.4.1 Σύνολα αξιολόγησης

Η *πρώτη* προσπάθεια ενιαίας αξιολόγησης των διαθέσιμων αλγορίθμων MVS έγινε με την εργασία των [193]. Για την αξιολόγηση των μεθόδων χρησιμοποιούνται έξι σύνολα φωτογραφιών. Για τα σύνολα αυτά, έχουμε βαθμονομήσει με ακρίβεια τις φωτογραφίες και γνωρίζουμε το αληθές τρισδιάστατο μοντέλο από χρήση laser scanner. Οι μέθοδοι αξιολογούνται βάσει

1. της ακρίβειάς τους.
2. της πληρότητάς τους.

Για την *ακρίβεια* της μεθόδου (1) μετράμε την απόσταση του ανακατασκευασμένου μοντέλου από το αληθές μοντέλο αναφοράς. Εκφράζουμε την ακρίβεια ως απόσταση d από το αληθές μοντέλο, εντός της οποίας βρίσκεται δεδομένο ποσοστό του ανακατασκευασμένου μοντέλου. Τονίζουμε ότι αφού οι αποστάσεις μετρώνται από το ανακατασκευασμένο μοντέλο στο αληθές, ελλείψεις στην ανακατασκευή (τρύπες) δεν αξιολογούνται από την ακρίβεια. Για την *πληρότητα* της μεθόδου (2), χρησιμοποιούμε τις αποστάσεις του αληθούς μοντέλου από το ανακατασκευασμένο. Εκφράζουμε την πληρότητα ως ποσοστό του αληθούς μοντέλου που βρίσκεται σε απόσταση μικρότερη κατωφλίου d από το ανακατασκευασμένο. Ουσιαστικά πρόκειται για την *αντίστροφη* της διαδικασίας του 1.

Αξιολογήσεις μεθόδων MVS και στατιστικά όπως χρόνοι εκτέλεσης, όφσει ευθυγράμμισης κ.α. σύμφωνα με την μέθοδο [193] είναι διαθέσιμα στον διαδικτυακό τόπο [1].

Αν και η συνεισφορά του συνόλου αξιολόγησης [193] ήταν σημαντική, αποτελείται αποκλειστικά από αντικείμενα και δεν αξιολογεί την ικανότητα *ανακατασκευής εξωτερικών σκηνών* των μεθόδων, που είναι για εμάς η πιο ενδιαφέρουσα περίπτωση. Η εργασία των [216] κάλυψε αυτό το κενό. Σε αυτήν, αξιολογούν τα αποτελέσματα της βαθμονόμησης των όψεων και της τρισδιάστατης ανακατασκευής. Χρησιμοποιούν κριτήρια αντίστοιχα των 1, 2 αλλά τα εκφράζουν συναρτήσει της διασποράς των δεδομένων του αληθούς μοντέλου. Το αληθές μοντέλο έχει και πάλι εξαχθεί με χρήση laser scanner. Τα αποτελέσματα τους είναι διαθέσιμα στον διαδικτυακό τόπο [3].

3.4.2 Παρουσίαση μεθόδων

Θα παρουσιάσουμε τώρα τα βασικά σημεία και αποτελέσματα διάφορων μεθόδων



Σχήμα 3.12: Ανακατασκευασμένο μοντέλο με laser scanner. Από [216].

Μέθοδοι με χρήση χαρακτηριστικών DAISY

Τα χαρακτηριστικά DAISY περιγράφονται στο [225]. Μια ενδιαφέρουσα μέθοδος ανακατασκευής η οποία βασίζεται στην χρήση τους παρουσιάζεται στο [227]. Η βασική ιδέα της μεθόδου είναι η χρήση των σθεναρών χαρακτηριστικών DAISY για το αξιόπιστο ταίριασμα χαρακτηριστικών διαφορετικών εικόνων. Η υπολογιστικά απαιτητική διαδικασία βελτιστοποίησης και απόρριψης outlier αποφεύγεται ή απλοποιείται σημαντικά.

Η μέθοδος του [227] ακολουθεί προσέγγιση παρόμοια με αυτή του συστήματος PMVS. Υπολογίζουμε ένα σύννεφο προσανατολισμένων σημείων σε δύο βήματα:

- Ταίριασμα χαρακτηριστικών μεταξύ ζευγαριών εικόνων και τριγωνισμός σημείων στον χώρο.
- Έλεγχος για συνέπεια των σημείων με επαναπροβολή τους σε διαφορετικές εικόνες από το ζευγάρι που χρησιμοποιήθηκε για τον τριγωνισμό τους.

Η μέθοδος παρουσιάζει εντυπωσιακά αποτελέσματα όσον αφορά τον χρόνο εκτέλεσης, ενώ η ακρίβεια της είναι συγκρίσιμη με αυτή του PMVS σε περιπτώσεις εικόνων πολύ υψηλής ανάλυσης (3.5). Ενδεικτικά αναφέρουμε ότι για το σύνολο δεδομένων Herz Jesu του [216] ο χρόνος εκτέλεσης του [227] ήταν 12.2 λεπτά ενώ του PMVS 843 λεπτά (μέγεθος patch 17×17 , στοιχεία από [227]). Τα αποτελέσματα αυτά οφείλονται στον αποδοτικό υπολογισμό των χαρακτηριστικών DAISY και στο γεγονός ότι η αξιοπιστία των χαρακτηριστικών μας επιτρέπει να αποφύγουμε την επίλυση προβλημάτων βελτιστοποίησης για την απόρριψη εσφαλμένων ταιριασμάτων και την εκλεπτύωση της λύσης. Για περισσότερες πληροφορίες και αποτελέσματα, ο αναγνώστης παραπέμπεται στην εργασία [227] και την ιστοσελίδα [226].

Βελτιώνοντας την υπολογιστική πολυπλοκότητα. Η μέθοδος του [197]

Η μέθοδος που προτείνεται στο [197] στηρίζεται στην *συγχώνευση χαρτών βάδους*. Έχει σχεδιαστεί για *μεγάλες εξωτερικές σκηνές* και προσπαθεί να αντιμετωπίσει το πρόβλημα με *ακρίβεια και αποδοτικότητα*.

Χρησιμοποιεί αλγόριθμο βασισμένο σε patch για την παραγωγή των χαρτών

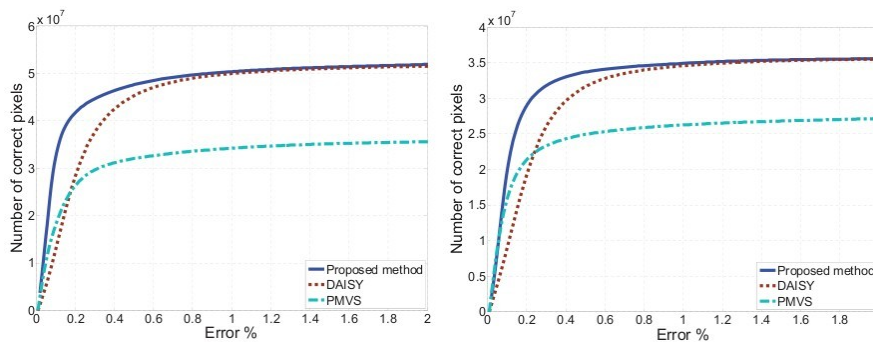
βάθους. Οι χάρτες βάθους στην συνέχεια βελτιώνονται, προκειμένου να επιβληθεί συνέπεια μεταξύ τους. Η μέθοδος είναι *παραλληλοποιήσιμη σε επίπεδο εικόνας*, γιατί κάθε χάρτης βάθους υπολογίζεται ανεξάρτητα. Συγκριτικά με την μέθοδο PMVS που αναλύσαμε στην Ενότητα 3.3 αναφέρεται βελτίωση του χρόνου εκτέλεσης 10 – 20 φορές, ενώ η ακρίβεια είναι αντίστοιχη. Συνοπτικά, τα βήματα της μεθόδου είναι

- Για κάθε εικόνα, επιλογή ζευγαριού (άλλης εικόνας) για στερεοσκοπικό ταίριασμα.
- Υπολογισμός χαρτών βάθους.
- Βελτίωση χαρτών βάθους.
- Συγχώνευση χαρτών βάθους.

Τα αποτελέσματα της μεθόδου παρουσιάζονται στον Πίνακα 3.5 και την Εικόνα 3.13. Παρατηρούμε ότι η προτεινόμενη μέθοδος και η μέθοδος DAISY [227] ανακατασκευάζουν μεγαλύτερο αριθμό σημείων σε σχέση με την μέθοδο PMVS (1.5 φορές περισσότερα σημεία). Η μέθοδος DAISY όμως υστερεί σε σχέση με την προτεινόμενη και την μέθοδο PMVS σε σχέση με την ακρίβεια της ανακατασκευής.

| Μέθοδος | Σωστά pixels | Λάθος pixels | Σωστά \ Λάθος |
|-------------|--------------|--------------|---------------|
| [197] | 4.306.332 | 266.224 | 6.2% |
| DAISY [227] | 4.107.572 | 382.104 | 9.3% |
| PMVS [69] | 2.838.744 | 346.512 | 12.2% |

Πίνακας 3.5: Σύγκριση μεθόδων. Οι τιμές αφορούν το σύνολο δεδομένων του Herz-Jesu-P8 από το [216] . Από [197].



Σχήμα 3.13: Αριθμός σωστά ανακατασκευασμένων σημείων ανά μέθοδο, συναρτήσει του κατωφλιού σφάλματος. Σύνολο δεδομένων: Αριστερά Fountain-P11 , Δεξιά Herz-Jesu-P8 ([216]). Τα διαγράμματα είναι από [197].

Η μέθοδος των [241]

Η συγκεκριμένη μέθοδος είναι κατάλληλη για εφαρμογή σε σύνολα φωτογραφιών *μερικών εκατοντάδων φωτογραφιών*. Η χρήση της σε σύνολα δεδομένων χιλιάδων φωτογραφιών δεν συστήνεται από τους συγγραφείς του [241].

Η αρχική ανακατασκευασμένη επιφάνεια προκύπτει με την επίλυση προβλήματος *ελάχιστης s-t τομής*, διαδικασία που απορρίπτει τους outliers. Στην συνέχεια το πλέγμα βελτιώνεται με χρήση *μεταβολικών τεχνικών*. Συνοπτικά η διαδικασία που ακολουθείται έχει ως εξής

- Υπολογισμός ψευδο-πυκνού σύννεφου σημείων με χρήση καθιερωμένων τεχνικών MVS.
- Εξαγωγή πολυγωνικού πλέγματος (mesh) λαμβάνοντας υπόψιν την ορατότητα. Το πλέγμα κατασκευάζεται από πρόβλημα ελάχιστης s-t (πηγής-καταβόθρας) τομής για να προσαρμόσουμε μια επιφάνεια στον Delaunay τριγωνισμό των σημείων.
- Μεταβολική εκλέπτυνση του πλέγματος για βελτιστοποίηση της φωτομετρικής συνάφειας.

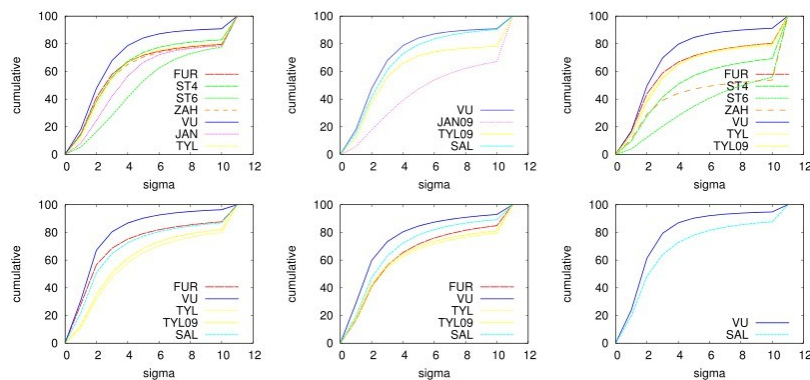
Ενδεικτικά αποτελέσματα της μεθόδου παρουσιάζονται στην εικόνα 3.14

Μερικές ακόμα μέθοδοι

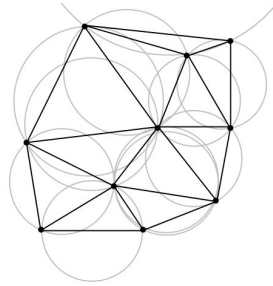
Στην προσέγγιση των [95], χρησιμοποιούνται *χάρτες βάθους* και επιτρέπουμε πολλές τιμές βάθους ανά ρixel σε κάθε χάρτη. Οι *οριστικές αποφάσεις αναβάλλονται όσο το δυνατόν περισσότερο*. Η τελική απόφαση λαμβάνεται με μεθόδους αντίστοιχες του RANSAC. Η βασική ιδέα της μεθόδου είναι ότι όσο αυξάνονται τα δεδομένα, οι λύσεις που λαμβάνουμε στο πρόβλημα χρησιμοποιώντας τους *inliers* θα συγκεντρώνονται γύρω από την σωστή λύση, την οποία μπορούμε επομένως να εντοπίσουμε.

Στο [254], οι συγγραφείς εντάσσουν το πρόβλημα της ανακατασκευής σε *Bayesian* πλαίσιο. Χαρακτηριστικές όψεις επιλέγονται και ανακατασκευάζονται χρησιμοποιώντας το σύστημα PMVS. Στην συνέχεια η ανακατασκευή βελτιώνεται επαναληπτικά, με την ένταξη νέων όψεων. Η βελτίωση της ανακατασκευής γίνεται μέσω της βελτιστοποίησης ενεργειακών όρων φωτομετρικής συνάφειας και γεωμετρικής ομαλότητας. Η επαναληπτική βελτίωση υπάρχοντος τρισδιάστατου μοντέλου γίνεται στα πλαίσια του *bayesian* συμπερασμού (*inference*) και αποτελεί την κύρια συνεισφορά της μεθόδου.

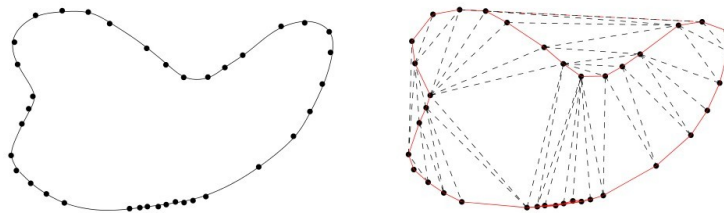
Στο [15] προτείνεται μέθοδος κατάλληλη για εφαρμογή σε *πολύ μεγάλα* σύνολα φωτογραφιών με *μεγάλες μεταβολές στην κλίμακα* των εικόνων. Η ανακατασκευή των επιφανειών γίνεται πάντα στην καλύτερη διαθέσιμη ανάλυση.



Σχήμα 3.14: Σωρευτικό σφάλμα για διάφορες μεθόδους και σύνολα δεδομένων. Με FUR συμβολίζεται η μέθοδος PMVS και με VU η μέθοδος [241]. Από [241].



Σχήμα 3.15: Τριγωνισμός Delaunay ενός συνόλου σημείων στο επίπεδο. Οι περιγεγραμμένοι κύκλοι των τριγώνων που σχηματίζουν τα σημεία έχουν σχεδιαστεί. Από την Wikipedia.



Σχήμα 3.16: Μη ομοιόμορφη δειγματοληψία μιας καμπύλης (Αριστερά) και Delaunay τριγωνισμός τους (Δεξιά). Γειτονικά σημεία σε όλες τις διευθύνσεις ανιχνεύονται. Από [75].

Η μέθοδος στηρίζεται σε χάρτες βάθους και βελτιώνει την μέθοδο επιλογής εικόνων του [78] που αναλύσαμε στην Ενότητα 3.2. Επίσης η υλοποίηση του αλγορίθμου γίνεται αποδοτικά σε GPU, βελτιώνοντας τις απαιτήσεις μνήμης και χρόνου σε σχέση με άλλες μεθόδους.

Αξιόπιστη βαθμονόμηση καμερών με χρήση συστήματος MVS

Στο [68] παρουσιάζεται η χρήση του συστήματος PMVS για να προσδιοριστούν με ακρίβεια οι παράμετροι προβολής των εικόνων του συνόλου δεδομένων. Η ιδέα είναι να εκμεταλλευτούμε πληροφορία που παρέχεται από την αδρή ανακατασκευή των επιφανειών για να καθορίσουμε αντιστοιχίες μεταξύ των εικόνων.

Για την αρχική βαθμονόμηση των όψεων, χρησιμοποιούμε κάποιο σύστημα SfM. Οι συγγραφείς του [68] χρησιμοποιούν το σύστημα Bundler [207]. Στην συνέχεια χρησιμοποιούμε το σύστημα PMVS για να πάρουμε ένα μεγαλύτερο σύνολο σημείων στον χώρο και κατ'επέκταση αντιστοιχιών σημείων μεταξύ εικόνων. Δειγματοληπτούμε τυχαία και ομοιόμορφα στο χώρο τα σημεία MVS και βελτιστοποιούμε την θέση τους στις εικόνες με μεθόδους conjugate gradient και κριτήριο βελτιστοποίησης την κανονικοποιημένη ετεροσυσχέτιση (NCC). Απορρίπτουμε τα χαρακτηριστικά που μετακινήθηκαν περισσότερο από ένα όριο κατά την διαδικασία της βελτιστοποίησης, και επαναλαμβάνουμε την διαδικασία Bundle Adjustment ([207]) χρησιμοποιώντας ως αντιστοιχίες τα σημεία των εικόνων που βελτιστοποιήθηκαν. Έτσι λαμβάνουμε μια βελτιωμένη εκτίμηση των παραμέτρων των όψεων. Η διαδικασία αυτή επαναλαμβάνεται 4 φορές. Σημειώνουμε ότι πριν χρησιμοποιήσουμε

το σύστημα PMVS, κάνουμε υποδειγματοληψία στις εικόνες του συνόλου δεδομένων προκειμένου να προσαρμόσουμε την ανάλυση των εικόνων στο εκτιμώμενο σφάλμα επαναπροβολής. Η προσαρμογή της ανάλυσης των εικόνων κάνει το σφάλμα στην θέση σημείων αρκετά μικρό ώστε το σύστημα PMVS να είναι σθεναρό σε λάθη στις παραμέτρους των καμερών.

Η μέθοδος πετυχαίνει την μείωση του σφάλματος επαναπροβολής, δηλαδή την βελτίωση της ακρίβειας προσδιορισμού των παραμέτρων προβολής. Το τρισδιάστατο μοντέλο που ανακατασκευάζεται χρησιμοποιώντας τις βελτιωμένες παραμέτρους για τις κάμερες, έχει βελτιωμένη ακρίβεια. Η εμφάνιση του όμως, διαφέρει σε λεπτομέρειες μόνο και η βελτίωση δεν γίνεται εύκολα αντιληπτή. Το σύστημα βαθμονόμησης είναι ελεύθερα διαθέσιμο στο [63].

3.4.3 Διαθέσιμο Λογισμικό

Μια online υπηρεσία τρισδιάστατης ανακατασκευής από μη-δομημένα σύνολα φωτογραφιών είναι διαθέσιμη από την ιστοσελίδα [224]. Διαθέσιμα συστήματα τρισδιάστατης ανακατασκευής, για διάφορα λειτουργικά συστήματα και με εύχρηστο γραφικό περιβάλλον μπορούν να αναζητηθούν στις ιστοσελίδες [247, 40].

Κώδικας για τα συστήματα PMVS και CMVS (Ενότητες 3.3 και 3.3.5) είναι διαθέσιμος στις ιστοσελίδες [66, 64], αντίστοιχα. Εναλλακτική υλοποίηση συστήματος mvs (CMPMVS) μπορεί να βρεθεί στην ιστοσελίδα [98]. Τέλος η αξιολόγηση των μεθόδων mvs των [193] μπορεί να βρεθεί στην ιστοσελίδα [1]. Για τις σκηνές εξωτερικών χώρων, το σύνολο δεδομένων του [216], είναι διαθέσιμο στην ιστοσελίδα [3].

IV

ΑΝΙΧΝΕΥΣΗ ΧΑΡΑΚΤΗΡΙΣΤΙΚΩΝ

Το Κεφάλαιο βασίζεται κυρίως στα [133], [149], [18], [243], [225]

Σε αυτό το κεφάλαιο θα ασχοληθούμε με μεθόδους *ανίχνευσης χαρακτηριστικών* σε εικόνες. Έχοντας εξάγει χαρακτηριστικά από τις εικόνες μπορούμε στην συνέχεια να συγκρίνουμε τις περιγραφές των χαρακτηριστικών αυτών και να θεσπίσουμε αντιστοιχίες μεταξύ εικόνων. Οι αντιστοιχίες αυτές χρησιμοποιούνται στην συνέχεια για την ανακατασκευή της τρισδιάστατης γεωμετρίας του χώρου που αποτυπώνεται στις εικόνες. Διαπιστώνουμε λοιπόν την *άμεση σύνδεση* της ανίχνευσης χαρακτηριστικών με το πρόβλημα της *τρισδιάστατης ανακατασκευής*.

Στους αλγορίθμους που αναλύουμε στην συνέχεια, μπορούμε να διακρίνουμε *δύο* στάδια

1. Το στάδιο ανίχνευσης *περιοχών* ενδιαφέροντος. Οι περιοχές αυτές αναπτύσσονται γύρω από σημειακά χαρακτηριστικά που εντοπίζουμε.
2. Το στάδιο *εξαγωγής περιγραφέα* για κάθε περιοχή που ανιχνεύθηκε.

Αξιολόγηση των ανιχνευτών περιοχών(1) γίνεται στο [149]. Θα αναφερθούμε σύντομα στο θέμα στην Ενότητα 4.2. Όμως, *έμφαση* δίνεται στους *περιγραφείς* (2). Θα αναλύσουμε διάφορες εναλλακτικές προσεγγίσεις. Στο [148] γίνεται μια συγκριτική αξιολόγηση περιγραφέων. Λόγω της παλαιότητας της δημοσίευσης, δεν συμπεριλαμβάνονται οι περισσότερες από τις μεθόδους που θα περιγράψουμε. Η *μέθοδος αξιολόγησης* των περιγραφέων του [148] παραμένει όμως σημαντική.

Εναλλακτικά, αντί να ακολουθήσουμε την διαδικασία που περιγράψαμε με τα σημεία 1,2, μπορούμε να κατασκευάσουμε περιγραφέα σε κάθε pixel μιας εικόνας. Καταλήγουμε έτσι σε ένα *πυκνό* σύνολο χαρακτηριστικών, αντίθετα με την προηγούμενη προσέγγιση που παράγει αραιό σύνολο χαρακτηριστικών. *Παράδειγμα* τέτοιας μεθόδου είναι η [225] που θα παρουσιάσουμε αναλυτικά στην Ενότητα 4.5.

Τι *είναι* όμως ένας περιγραφέας και τι *χαρακτηριστικά* επιδιώκουμε να έχει; Τα βασικά αυτά ερωτήματα πρέπει να αποσαφηνιστούν πριν επεκταθούμε περισσότερο. Ο περιγραφέας είναι ένα σύντομο διάνυσμα, αμετάβλητο σε συνήθεις μετασχηματισμούς όπως κλιμάκωση, αφινικοί μετασχηματισμοί, αλλαγές φωτισμού κ.α. Έτσι οι περιγραφείς διαφορετικών όψεων του ίδιου σημείου θα ταυτίζονται (ιδανικά, στην πραγματικότητα περιμένουμε να είναι παρόμοιοι), επιτρέποντας το ταίριασμα των σημείων και την δημιουργία

αντιστοιχίας μεταξύ των διαφορετικών εικόνων.

Στην παρουσίαση που ακολουθεί, χρησιμοποιούμε *χρονολογική σειρά* για την παρουσίαση των εξελίξεων των μεθόδων χαρακτηριστικών. Στην Ενότητα 4.1, παρουσιάζουμε την καθιερωμένη πλέον μέθοδο SIFT. Η μέθοδος αυτή στηρίζεται σε ιστογράμματα, είναι ακριβής και επηρέασε καθοριστικά την μετέπειτα ερεύνα. Ουσιαστικά αποτελεί την βάση για τα επόμενα συστήματα ανίχνευσης χαρακτηριστικών.

Μια πληρέστερη κάλυψη της περιοχής γίνεται στην Ενότητα 4.2, η οποία βασίζεται στο [148]. Στην ίδια Ενότητα παρουσιάζεται και μια βελτίωση των χαρακτηριστικών SIFT (GLOH-Gradient Location Orientation Histogram), και συστηματοποιείται η αξιολόγηση των ανιχνευτών χαρακτηριστικών.

Στη Ενότητα 4.3 παρουσιάζουμε τα χαρακτηριστικά SURF (Speeded Up Robust Features). Στην Ενότητα 4.4 παρουσιάζουμε μια προσπάθεια συστηματοποίησης της διαδικασίας σχεδιασμού περιγραφέων. Εισάγονται ακόμα οι ιδέες που αξιοποιούνται στην Ενότητα 4.5.

Τέλος, στην Ενότητα 4.6 συμπληρώνουμε την ανάλυση που προηγήθηκε και αναφέρουμε χρήσιμες διαθέσιμες υλοποιήσεις σχετικές με την ύλη που κάλυψαμε.

4.1 Η καθιερωμένη προσέγγιση στην ανίχνευση χαρακτηριστικών: Scale Invariant Feature Transform (SIFT)

Περιγράφουμε στην συνέχεια τον ανιχνευτή χαρακτηριστικών SIFT. Η παρουσίαση βασίζεται στο [133].

4.1.1 Γενικά στοιχεία

Τα χαρακτηριστικά SIFT εντάσσονται στην κατηγορία χαρακτηριστικών που βασίζονται σε *ιστογράμματα* και αποτελούν μια *καθιερωμένη* και *αποτελεσματική* μέθοδο εξαγωγής χαρακτηριστικών. Στα περίπου δέκα χρόνια που μεσολάβησαν από την αρχική δημοσίευσή τους, αποτέλεσαν την *βάση* για πολλά συστήματα εξαγωγής χαρακτηριστικών και επιχειρήθηκε η *τροποποίηση* τους με διάφορους τρόπους. Η επιτυχία του αλγορίθμου SIFT αντικατοπτρίζεται και από την υλοποίησή τους σε υλικό (FPGA [252]) καθώς και την υλοποίηση του αλγορίθμου σε GPU [248].

Τα χαρακτηριστικά SIFT είναι *αμετάβλητα* σε *κλιμάκωση* και *περιστροφή*. Έχει δειχθεί ακόμα η σθεναρότητα των χαρακτηριστικών σε *αφινικούς* μετασχηματισμούς, αλλαγή του *σημείου θέασης* στον χώρο, προσθήκη *θορύβου* και αλλαγές στον *φωτισμό*.

Μεγάλος αριθμός χαρακτηριστικών εντοπίζονται σε κάθε εικόνα. Τα χαρακτηριστικά έχουν έντονο *διακριτικό* χαρακτήρα, γεγονός που επιτρέπει κάθε χαρακτηριστικό να ταυριαστεί με ακρίβεια με όμοιά του σε μεγάλα σύνολα χαρακτηριστικών, τα οποία μπορεί να προέρχονται από μεγάλο αριθμό διαφορετικών εικόνων. Η εξαγωγή χαρακτηριστικών και το ταίριασμα τους μπορεί να γίνει με αποτελεσματικό τρόπο.

Οι συγγραφείς του [133] προτείνουν ακόμα την χρήση των χαρακτηριστικών

SIFT σε *συνδυασμό* με άλλα είδη χαρακτηριστικών. Επιγραμματικά αναφέρουμε

- Χαρακτηριστικά που χρησιμοποιούν τα *όρια* των περιοχών. Τα χαρακτηριστικά αυτά επηρεάζονται λιγότερο από το παρασκήνιο κοντά στα όρια των περιοχών [143].
- Χαρακτηριστικά που χρησιμοποιούν τις *τοπικές ακμές* και *αγνοούν* γειτονικές ασυσχέτιστες ακμές. Τα χαρακτηριστικά αυτά δίνουν ευσταθή περιγραφή στα όρια στενών περιοχών [150].
- Χαρακτηριστικά που χρησιμοποιούν την *φάση* των τοπικών χωρικών συχνοτήτων. Εμφανίζουν μεγαλύτερη σθεναρότητα σε μεταβολές του φωτισμού [31].
- Χαρακτηριστικά που αξιοποιούν πληροφορία *χρώματος, κίνησης, βάθους, προσκηνίου-παρασκηίου κ.α.*

Η διαδικασία που ακολουθείται για την εξαγωγή των χαρακτηριστικών μπορεί να περιγραφεί με μια *ακολουθία επιμέρους σταδίων*. Μόνο οι περιοχές που αξιολογούνται θετικά σε ένα στάδιο επεξεργάζονται από το επόμενο του. Έτσι επιτυγχάνεται μεγαλύτερη *αποδοτικότητα* του αλγορίθμου. Συνοπτικά τα στάδια είναι

1. Εντοπισμός ακρότατων στον χώρο κλίμακας (*scale-space*). Η αναζήτηση γίνεται σε όλες τις τοποθεσίες και κλίμακες της εικόνας. Η αποτελεσματική της υλοποίηση βασίζεται στην χρήση Διαφορών Γκαουσιανών (DoG). Τα σημεία που εντοπίζονται είναι αμετάβλητα σε κλίμακα και προσανατολισμό.
2. Σε κάθε υποψήφια τοποθεσία, προσδιορίζεται η κλίμακα και η ακριβής θέση του σημείου ενδιαφέροντος. Επιλέγονται μόνο ευσταθή σημεία.
3. Σε κάθε σημείο ενδιαφέροντος αναθέτουμε προσανατολισμό, βάσει της τοπικής κλίσης.
4. Εξαγωγή περιγραφέα. Υπολογίζουμε τον προσανατολισμό της κλίσης σε περιοχή που εγκλείει το σημείο ενδιαφέροντος. Το μέγεθος της περιοχής εξαρτάται από την κλίμακα που αναθέσαμε στο σημείο ενδιαφέροντος. Στην συνέχεια οι μετρήσεις αυτές μετασχηματίζονται σε περιγραφή που εξασφαλίζει σθεναρότητα σε γεωμετρικές παραμορφώσεις και αλλαγές φωτισμού.

Στην συνέχεια εξετάζουμε αναλυτικότερα τα στάδια του αλγορίθμου SIFT.

4.1.2 Αναλυτική περιγραφή σταδίων αλγορίθμου SIFT

Εντοπισμός ακρότατων στον χώρο κλίμακας

Για τον ορισμό του χώρου κλίμακας της εικόνας χρησιμοποιούμε Γκαουσιανό πυρήνα. Η συνάρτηση χώρου κλίμακας της εικόνας ορίζεται από την συνέλιξη της εικόνας με Γκαουσιανή μεταβλητής κλίμακας

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y)$$

Υπενθυμίζουμε τον ορισμό της ιστροπικής Γκαουσιανής

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-x^2 + y^2/2\sigma^2}$$

Για να εντοπίσουμε τα υποψήφια σημεία ενδιαφέροντος χρησιμοποιούμε την συνάρτηση Διαφοράς Γκαουσιανών (DoG). Συγκεκριμένα εντοπίζουμε τα ακρότατα της συνέλιξης της DoG με την εικόνα, η οποία υπολογίζεται ανάμεσα σε δύο διαδοχικές κλίμακες:

$$\begin{aligned} D(x, y, \sigma) &= (G(x, y, \kappa\sigma) - G(x, y, \sigma)) * I(x, y) \\ &= G(x, y, \kappa\sigma) * I(x, y) - G(x, y, \sigma) * I(x, y) \\ &= L(x, y, \kappa\sigma) - L(x, y, \sigma) \end{aligned} \quad (4.1)$$

Οι λόγοι που χρησιμοποιούμε την Εξίσωση 4.1 είναι πολλοί. Αρχικά μπορεί να υπολογισθεί εύκολα με απλές αφαιρέσεις των εικόνων στις διάφορες κλίμακες. Επίσης είναι μια καλή προσέγγιση της συνάρτησης

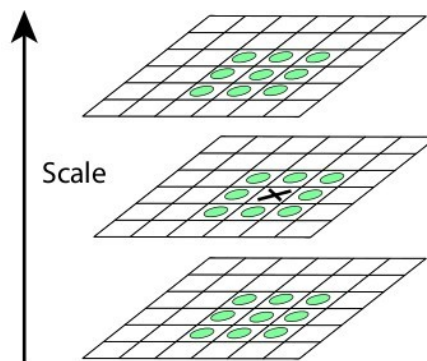
$$\sigma^2 \nabla^2 G$$

η οποία δίνει τα πιο ευσταθή χαρακτηριστικά συγκριτικά με άλλες συναρτήσεις της εικόνας (π.χ Hessian, Harris, gradient).

Για την ανίχνευση των *ακρότατων* της 4.1, ακολουθούμε την ακόλουθη διαδικασία. Έστω δείγμα εικόνας σε κάποια κλίμακα. *Συγκρίνουμε* το δείγμα με τους 8 *γείτονές* του στην ίδια εικόνα, καθώς και με τους 9 γείτονες σε κάθε μία από τις εικόνες στην αμέσως μεγαλύτερη και στην αμέσως μικρότερη κλίμακα. Το *δείγμα επιλέγεται* αν είναι μεγαλύτερο από όλους τους γείτονες ή μικρότερο από όλους. Η διαδικασία παρουσιάζεται στην Εικόνα 4.1.

Είναι σημαντικό να προσδιορίσουμε τον *ρυθμό δειγματοληψίας*, τόσο στο πεδίο της εικόνας (χωρικά) όσο και στο πεδίο της κλίμακας. Η επιλογή αυτή καθορίζει ένα ισοζύγιο μεταξύ αποδοτικότητας των υπολογισμών και πληρότητας στην ανίχνευση ακρότατων. Ο προσδιορισμός των συχνοτήτων δειγματοληψίας γίνεται στο [133] *πειραματικά*, με σκοπό να επιλεγθούν ρυθμοί δειγματοληψίας που δίνουν τα πιο αξιόπιστα αποτελέσματα.

Για την δειγματοληψία στο *πεδίο της κλίμακας*, παρατηρήθηκε μέγιστη επαναληπτικότητα σε δειγματοληψία 3 κλιμάκων ανά οκτάβα. Η *επαναληπτικότητα* ορίζεται ως το ποσοστό χαρακτηριστικών που ανιχνεύονται στην ίδια



Σχήμα 4.1: Εντοπισμός ακρότατων Διαφορών Γκαουσιανών. Εικονίζονται 3 εικόνες σε διαδοχικές κλίμακες. Το υπό εξέταση σημείο συμβολίζεται με Χ και συγκρίνεται με τους "γείτονές" του οι οποίοι συμβολίζονται με Ο. Από [133].

θέση και κλίμακα στην αρχική και την μετασχηματισμένη εικόνα. Ο δείκτης της επαναληπτικότητας χρησιμοποιείται για την αξιολόγηση και σύγκριση ανιχνευτών περιοχών.

Συμπερασματικά, η επαναληπτικότητα εκφράζει αν ένας ανιχνευτής βρίσκει αξιόπιστα τα ίδια σημεία ενδιαφέροντος, υπό διαφορετικές συνθήκες θέασης. Η επαναληπτικότητα δεν αυξάνεται μονότονα με την αύξηση του ρυθμού δειγματοληψίας στην κλίμακα. Αυτό οφείλεται στην ανίχνευση περισσότερων τοπικών ελαχίστων της 4.1, τα οποία δεν είναι ευσταθή σε μικρές διαταραχές. Η παρατηρούμενη εξέλιξη της επαναληπτικότητας συναρτησει του ρυθμού δειγματοληψίας στην κλίμακα παρουσιάζεται στην Εικόνα 4.2. Παρατηρούμε ότι ο απόλυτος αριθμός των χαρακτηριστικών που ανιχνεύονται και ταιριάζονται σωστά ανάμεσα σε δύο εικόνες, αυξάνεται μονότονα με αύξηση της δειγματοληψίας, παρότι η επαναληπτικότητα μειώνεται.

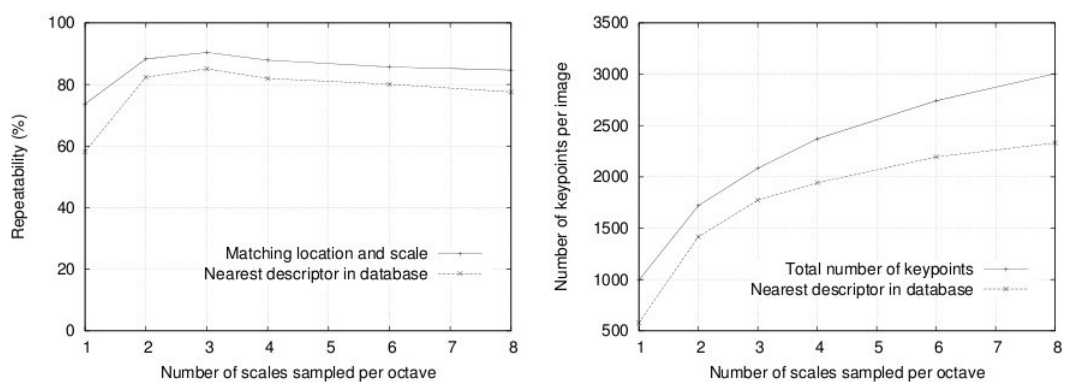
Συνοψίζοντας, η ανίχνευση μεγαλύτερου αριθμού τοπικών ακρότατων της 4.1 συνεπάγεται αύξηση του υπολογιστικού φόρτου. Για αυτό, επιλέγεται η ανίχνευση των πιο *ευσταδών* ακρότατων της.

Στο (χωρικό) πεδίο της εικόνας, καθορίζουμε την κλίμακα σ της Γκαουσιανής που χρησιμοποιείται για την εξομάλυνση της εικόνας πριν κατασκευάσουμε την αναπαράσταση κλίμακας για κάθε οκτάδα. Αν και η επαναληπτικότητα βελτιώνεται με αύξηση της εξομάλυνσης, περιορισμοί που αφορούν την αποδοτική υλοποίηση του αλγορίθμου οδηγούν στην τελική επιλογή του σ ($\sigma = 1.6$ στο [133]).

Η εξομάλυνση της εικόνας πριν την ανίχνευση των ακρότατων με Γκαουσιανές συνιστά βαθυπερατό φιλτράρισμα της εικόνας. Για αυτό τον λόγο επεκτείνουμε την εικόνα με γραμμική παρεμβολή πριν την ανάλυση σε κλίμακες. Ο διπλασιασμός των διαστάσεων της εικόνας οδηγεί σε βελτίωση της απόδοσης έως 4 φορές.

Ακριβής εντοπισμός χαρακτηριστικών και ανάθεση προσανατολισμού

Αφού εντοπίσαμε τα ακρότατα της 4.1 με τον τρόπο που περιγράψαμε προηγουμένως και που απεικονίζεται στην Εικόνα 4.1, προσδιορίζουμε με μεγαλύτερη *ακρίβεια* την θέση του ακρότατου στον χώρο (επίπεδο της εικόνας) και στην κλίμακα. *Απορρίπτουμε* επίσης ακρότατα με χαμηλή αντίθεση



Σχήμα 4.2: Τον οριζόντιο άξονα απεικονίζονται οι κλίμακες που δειγματοληπτούμε ανα οκτάδα. Αριστερά: στον κατακόρυφο άξονα απεικονίζεται η επαναληψιμότητα. Δεξιά: στον κατακόρυφο άξονα απεικονίζεται ο αριθμός των περιοχών που ανιχνεύονται. Από [133].

(*contrast*) και ακρότατα που εντοπίζονται σε *ασταθείς ακμές*.

Για τον *ακριβή εντοπισμό* των σημείων ενδιαφέροντος, χρησιμοποιούμε το δευτεροβάθμιο ανάπτυγμα *Taylor* της 4.1. Εντοπίζουμε το σημείο ακρότατου $\hat{\mathbf{x}}$ του αναπτύγματος και το χρησιμοποιούμε στην θέση της προηγούμενης εκτίμησης μας για την θέση του σημείου.

Η απόλυτη τιμή της συνάρτησης 4.1 $\|D(\hat{\mathbf{x}})\|$, συγκρίνεται με κατάλληλο κατώφλι. Σε περίπτωση που το κατώφλι δεν ξεπερνιέται, το σημείο απορρίπτεται για λόγους χαμηλού *contrast*.

Η Εξίσωση 4.1 αποκρίνεται ισχυρά και στις ακμές. Προκειμένου να απορρίψουμε ασταθείς ακμές, ελέγχουμε τις ιδιοτιμές της Hessian μήτρας της 4.1 στο σημείο του υποψήφιου χαρακτηριστικού. Οι ιδιοτιμές δεν υπολογίζονται ξεχωριστά, εφόσον ενδιαφέρει μόνο ο λόγος τους. Την συγκεκριμένη ιδέα δανειζόμαστε από το [85]. Οι ακμές που θέλουμε να απορρίψουμε χαρακτηρίζονται από μια μεγάλη και μια σημαντικά μικρότερη ιδιοτιμή.

Τέλος αναφέρουμε ότι στα προηγούμενα, ο αποδοτικός υπολογισμός των παραγώγων γίνεται μέσω της προσέγγισης τους από διαφορές γειτονικών δειγμάτων.

Σε κάθε σημείο ενδιαφέροντος, αναθέτουμε *συνεπή προσανατολισμό* βάσει τοπικών ιδιοτήτων της εικόνας. Αυτό μας επιτρέπει να χρησιμοποιήσουμε μέτρα ιδιοτήτων της εικόνας που δεν είναι περιστροφικά αμετάβλητα. Σχηματίζουμε ένα *ιστόγραμμα προσανατολισμών*, στο οποίο χωρίζουμε 36 κατηγορίες για να καλύψουμε τους διαφορετικούς προσανατολισμούς. Σε κάθε κατηγορία δηλαδή, αντιστοιχούν 10° . Οι *προσανατολισμοί* μετρώνται με την *διεύθυνση της κλίσης*, σε σημεία στην *περιοχή του σημείου ενδιαφέροντος*. Η *έκταση* της περιοχής γύρω από το σημείο ενδιαφέροντος, καθορίζεται από την *κλίμακα* στην οποία ανιχνεύθηκε το σημείο. Η κλίση κάθε σημείου της περιοχής υπολογισμού του ιστογράμματος, συνεισφέρει στην αντίστοιχη κατηγορία, *σταδμισμένη* από το *μέτρο* της και από *Γκαουσιανή* τοποθετημένη στο σημείο ενδιαφέροντος.

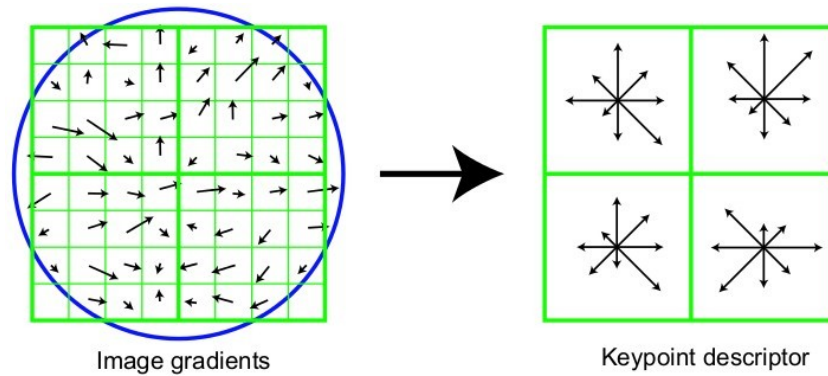
Οι κορυφές του ιστογράμματος, χαρακτηρίζουν τον κυρίαρχο προσανατολισμό στην περιοχή του σημείου ενδιαφέροντος. Περιγραφείς με τους αντίστοιχους προσανατολισμούς δημιουργούνται για όλες τις κατηγορίες του ιστογράμματος που βρίσκονται εντός του 80% του μεγίστου. Τονίζουμε δηλαδή ότι *περισσότερα από ένα χαρακτηριστικά (περιγραφείς) ενδέχεται να δημιουργηθούν από ένα σημείο ενδιαφέροντος*. Η απόφαση αυτή βελτιώνει σημαντικά την ευστάθεια του αλγορίθμου.

Σε κάθε κορυφή του ιστογράμματος που "γεννά" έναν περιγραφέα, ο προσανατολισμός του επιλέγεται ως εξής. Προσαρμόζουμε μια παραβολή στην συγκεκριμένη κορυφή του ιστογράμματος και τις δύο γειτονικές κατηγορίες, ώστε να προσδιορίσουμε με μεγαλύτερη ακρίβεια την κορυφή του ιστογράμματος.

Δομή Περιγραφέα

Για να κατασκευάσουμε τον περιγραφέα ενός σημείου ενδιαφέροντος, *αρχικά* υπολογίζουμε το *μέτρο* και τον *προσανατολισμό* της κλίσης σε *περιοχή* γύρω από το σημείο. Οι κλίσεις στην συνέχεια χρησιμοποιούνται για την κατασκευή *ιστογραμμάτων* προσανατολισμού της κλίσης σε υποπεριοχές του σημείου. Τα ιστογράμματα αυτά θα χρησιμοποιηθούν τελικά για την περιγραφή του σημείου. Η διαδικασία παρουσιάζεται στην Εικόνα 4.3.

Για τον *υπολογισμό των κλίσεων*, εφαρμόζουμε Γκαουσιανή συνέλιξη στην



Σχήμα 4.3: Περιγραφέας SIFT. Αρχικά υπολογίζουμε το μέτρο και και την διεύθυνση της κλίσης σε δείγματα γύρω από το σημείο ενδιαφέροντος (αριστερά). Το σημείο ενδιαφέροντος βρίσκεται στο κέντρο του πλέγματος. Η στάθμιση των κλίσεων με Γκαουσιανή συμβολίζεται με τον μπλε κύκλο. Ο συμψηφισμός των κλίσεων σε ιστογράμματα, για την τελική διαμόρφωση του περιγραφέα, εικονίζεται στην δεξιά εικόνα. Στο συγκεκριμένο παράδειγμα, έχουμε 4 υποπεριοχές (και συνεπώς 4 ιστογράμματα). Από [133].

εικόνα. Το μέγεθος της Γκαουσιανής προσδιορίζεται από την κλίμακα στην οποία ανιχνεύθηκε το σημείο ενδιαφέροντος.

Για να επιτύχουμε περιγραφή *αμετάβλητη σε περιστροφή*, περιστρέφουμε το σύστημα συντεταγμένων σύμφωνα με τον προσανατολισμό του σημείου ενδιαφέροντος, όπως αυτός προσδιορίστηκε προηγουμένως.

Το μέτρο κάθε κλίσης που υπολογίζουμε, *σταθμίζεται από Γκαουσιανή* τοποθετημένη στο κέντρο του παραθύρου υπολογισμού (στο σημείο ενδιαφέροντος), με μέγεθος ανάλογο του μεγέθους του παραθύρου υπολογισμού του περιγραφέα. Η συγκεκριμένη στάθμιση γίνεται για να *αποφύγουμε σημαντικές αλλαγές* του περιγραφέα από *μικρές μετατοπίσεις* της θέσης του σημείου ενδιαφέροντος και για να δώσουμε *μικρότερη έμφαση* σε τιμές της κλίσης σε τοποθεσίες *απομακρυσμένες από το σημείο ενδιαφέροντος*.

Η χρήση *ιστογραμμάτων σε υποπεριοχές*, επιτρέπει την ολίσθηση των τιμών των κλίσεων στην εικόνα χωρίς σημαντικές μεταβολές στην περιγραφή του σημείου. Επίσης, για να αποφύγουμε απότομες αλλαγές στην περιγραφή από ολίσθησεις κλίσεων σε γειτονικές κατηγορίες προσανατολισμού στα ιστογράμματα ή μεταξύ υποπεριοχών υπολογισμού των ιστογραμμάτων, χρησιμοποιούμε *κατάλληλη στάθμιση* των συνεισφορών των κλίσεων στα ιστογράμματα και τις κατηγορίες τους.

Στην προτεινόμενη από τους συγγραφείς του [133] υλοποίηση του περιγραφέα, η περιοχή του σημείου ενδιαφέροντος χωρίζεται σε ένα 4×4 πλέγμα υποπεριοχών. Κάθε ιστογράμματος προσανατολισμού δημιουργείται με 8 κατηγορίες. Συνεπώς η διάσταση του περιγραφέα προκύπτει

$$4 \times 4 \text{ Ιστογράμματα} \times 8 \frac{\text{Προσανατολισμοί}}{\text{Ιστογράμματα}} = 128$$

Για να έχουμε περιγραφή *αμετάβλητη σε αφινικές μεταβολές του φωτισμού*, κανονικοποιούμε το διάνυσμα της περιγραφής σε μοναδιαίο. Ακόμα, *σθεναρότητα σε μη γραμμικές μεταβολές* του φωτισμού επιτυγχάνεται περιορίζοντας την επίδραση των πολύ μεγάλων τιμών της κλίσης στην περιγραφή. Για αυτό τον λόγο θέτουμε κατάλληλο κατώφλι στο μέτρο της κλίσης. Έτσι το μέτρο των πολύ μεγάλων κλίσεων δεν είναι πλέον σημαντικό και δίνεται έμφαση

στον προσανατολισμό.

Η διάσταση του περιγραφέα εξαρτάται όπως παρατηρήσαμε από δύο παραμέτρους

1. Το πλήθος των υποπεριοχών n .
2. Τον αριθμό o των προσανατολισμών σε κάθε ιστόγραμμα.

Το τελικό μέγεθος της περιγραφής προκύπτει $o \cdot n$.

Η αύξηση του μεγέθους του περιγραφέα προσδίδει στην περιγραφή *μεγαλύτερη διακριτική ικανότητα* και καθιστά περισσότερο *χρονοβόρο του υπολογισμού του*. Όμως, αυξάνει και την *ευαισθησία* του σε παραμορφώσεις, γεγονός ιδιαίτερα αρνητικό για την επίδοση του.

Συμπεράσματα

Εισάγαμε σε αυτή την ενότητα τα χαρακτηριστικά SIFT, βασικό εκπρόσωπο μιας μεγάλης κατηγορίας αλγορίθμων ανίχνευσης χαρακτηριστικών που στηρίζονται σε ιστογράμματα και χρησιμοποιούν αναπαραστάσεις υψηλής διάστασης.

Στην επόμενη ενότητα θα παρουσιάσουμε εκτενέστερα το πλαίσιο στο οποίο εντάσσονται τα χαρακτηριστικά SIFT και θα καλύψουμε πληρέστερα την περιοχή της εξαγωγής χαρακτηριστικών.

4.2 Αξιολόγηση των ανιχνευτών χαρακτηριστικών. Βελτίωση της διακριτικής ικανότητας του SIFT

Σε αυτή την Ενότητα θα ερευνήσουμε το πεδίο των περιγραφέων. Παρουσιάζουμε τις διάφορες *προσεγγίσεις* για την περιγραφή ενός σημείου ενδιαφέροντος και εισάγουμε κατάλληλο πλαίσιο *αξιολόγησης* και *κατάταξης* τους. Παρουσιάζουμε ακόμα τον περιγραφέα *GLOH* (Gradient Location and Orientation Histogram), μετεξέλιξη του SIFT (Ενότητα 4.1).

Θα προσπαθήσουμε να αποσαφηνίσουμε σημαντικά *ερωτήματα* που αφορούν την *ανίχνευση* χαρακτηριστικών και το *ταίριασμα* τους σε διαφορετικές εικόνες. Ανάμεσα στους περιγραφείς, θα ξεκαθαρίσουμε ποιοί είναι οι *κατάλληλοι* καθώς και τον τρόπο που η επίδοσή τους επηρεάζεται από τον *ανιχνευτή* περιοχών που χρησιμοποιείται. Επικεντρώναστε σε περιγραφείς για *γκρίζες* εικόνες.

Τεχνικές περιγραφής περιοχών

Για την περιγραφή τοπικών περιοχών εικόνων, πληθώρα τεχνικών έχει χρησιμοποιηθεί. Μπορούμε να κατηγοριοποιήσουμε τους αντίστοιχους περιγραφείς ως εξής:

1. *Ευθύς* υπολογισμός από εντάσεις pixels.
2. Περιγραφείς που βασίζονται σε *κατανομές* ιστογραμμάτων.
3. Περιγραφείς από τεχνικές που χρησιμοποιούν *χωρικές συχνότητες*.
4. *Διαφορικοί* περιγραφείς.

5. Περιγραφείς που βασίζονται στις ροπές.

Η πιο απλή προσέγγιση (1) είναι να χρησιμοποιήσουμε κατευθείαν τις εντάσεις (φωτεινότητας) των pixels. Ο περιγραφέας μιας περιοχής είναι τότε απλά διάνυσμα των pixels της περιοχής. Η σύγκριση περιοχών μπορεί να γίνει μετρώντας την *ετεροσυσχέτιση* των περιγραφών τους. Η *διάσταση* της περιγραφής είναι ιδιαίτερα μεγάλη. Για να αντιμετωπιστούν τα επακόλουθα προβλήματα πολυπλοκότητας χρησιμοποιούνται τεχνικές όπως η ανάλυση σε πρωτεύουσες συνιστώσες (PCA, π.χ Κεφ.6 [223]) και η υποδειγματοληψία.

Οι περιγραφείς που χρησιμοποιούν *κατανομές* (2), χρησιμοποιούν ιστογράμματα για να αναπαραστήσουν διαφορά χαρακτηριστικά σχήματος ή εμφάνισης. Για παράδειγμα, τα χαρακτηριστικά μπορεί να σχετίζονται με τιμές φωτεινότητας των pixels [255], με τις σχετικές θέσεις εντός της γειτονιάς του σημείου ενδιαφέροντος [123], με την κατανομή της κλίσης (Ενότητα 4.1) κ.α.

Πολλές τεχνικές περιγράφουν το *συχνοτικό περιεχόμενο* της εικόνας (3). Ο μετασχηματισμός Fourier χρησιμοποιείται για την αποσύνθεση της εικόνας σε συναρτήσεις βάσεις. Η προσέγγιση αυτή όμως δεν κάνει άμεσες τις χωρικές σχέσεις των σημείων και οι συναρτήσεις βάσεις είναι άπειρες. Τα προβλήματα αυτά μπορούν να ξεπεραστούν με χρήση του μετασχηματισμού Gabor , όμως απαιτείται μεγάλος αριθμός φίλτρων.

Οι *διαφορικοί περιγραφείς* (4), προσεγγίζουν την γειτονιά ενός σημείου με παραγώγους μέχρι μια συγκεκριμένη τάξη [113]. Διάφορες προσεγγίσεις χρησιμοποιούν κατευθυνόμενα (steerable) φίλτρα [60]. Για την ευσταθή εκτίμηση των παραγώγων χρησιμοποιούνται συνελίξεις με εξομαλυντικούς Γκαουσιανούς πυρήνες. Τέλος άλλοι συγγραφείς χρησιμοποιούν φίλτρα από την οικογένεια των μιγαδικών εκθετικών [17, 192].

Τέλος, άλλες τεχνικές χρησιμοποιούν *ροπές* [237]. Οι ροπή της έντασης της εικόνας τάξης $p + q$ και βαθμού α ορίζεται με την σχέση

$$M_{pq}^{\alpha} = \iint_{\Omega} x^p y^q [I(x, y)]^{\alpha} dx dy$$

Οι ροπές χαρακτηρίζουν το σχήμα και την ένταση στην περιοχή Ω , είναι όμως ευαίσθητες σε παραμορφώσεις σχήματος και φωτομετρικές μεταβολές. Υπολογίζονται όμως εύκολα και η διάσταση της περιοχής είναι μικρή, επιτρέποντας στην περίπτωση έγχρωμων εικόνων τον υπολογισμό τους σε κάθε κανάλι χρώματος ξεχωριστά.

4.2.1 Δομή πειραμάτων αξιολόγησης

Εντοπισμός και περιγραφή περιοχών

Για τον εντοπισμό των περιοχών για τις οποίες θα κατασκευαστούν περιγραφείς, μπορούμε αντιστοίχως να χρησιμοποιήσουμε διαφορετικές τεχνικές. Όπως προαναφέραμε, αξιολόγηση και σύγκριση των τεχνικών αυτών μπορεί να βρεθεί στο [149]. Επιγραμματικά και για λόγους πληρότητας, αναφέρουμε κάποιες προσεγγίσεις στην ανίχνευση περιοχών

- Εντοπισμός μεγίστων της κανονικοποιημένης Λαπλασιανής στον χώρο κλίμακας [126]. Όπως είδαμε στην Ενότητα 4.1, στους περιγραφείς SIFT προσεγγίζεται η Λαπλασιανή με διαφορές Γκαουσιανών. Το μέτρο αυτό

είναι αμετάβλητο σε κλιμάκωση. Στο [128] εξελίσσεται σε αφινικά αμετάβλητο με χρήση ροπών δεύτερης τάξης.

- Παραλλαγές του ανιχνευτή Harris [145, 146].
- "Maximally stable extremal regions" [142].
- Μέτρηση εντροπίας σε ιστογράμματα εντάσεων pixel και εύρεση τοπικών μεγίστων στον χώρο αφινικών μετασχηματισμών [104].

Στην αξιολόγηση [148] χρησιμοποιήθηκαν οι επόμενοι ανιχνευτές περιοχών

- Harris [85]. Αμετάβλητος σε περιστροφές. Το μέγεθος της ανιχνευθείσας περιοχής είναι σταθερό.
- Harris-Laplace [145]. Αμετάβλητος σε περιστροφές και κλιμάκωση. Τα σημεία ανιχνεύονται από την (προσαρμοσμένη) συνάρτηση Harris και εντοπίζονται στον χώρο κλίμακας με τον τελεστή LoG (Laplacian of Gaussian).
- Hessian-Laplace [133, 147]. Αμετάβλητος σε περιστροφή και κλιμάκωση. Τα σημεία εντοπίζονται στον χώρο από τα ακρότατα της οριζουσας της Hessian και στην κλίμακα με τον τελεστή LoG.
- Harris-Affine [147]. Αφινικά αμετάβλητος. Παρόμοιος με τον Harris-Laplace.
- Hessian-Affine [147]. Αντίστοιχα με τον ανιχνευτή Harris-Affine. Παρόμοιος με τον Hessian-Laplace.

Περισσότερες λεπτομέρειες για τους ανιχνευτές μπορούν να αναζητηθούν στις αντίστοιχες παραπομπές.

Οι περιοχές που επιστρέφονται από τους ανιχνευτές περιοχών *κανονικοποιούνται*, προκειμένου να βελτιώσουμε την *διακριτικότητα* της περιγραφής τους (που θα ακολουθήσει) και να επιτύχουμε οι περιοχές να είναι *αμετάβλητες* σε συγκεκριμένους μετασχηματισμούς. Για να είναι οι περιοχές περιστροφικά αμετάβλητες, περιστρέφουμε τις περιοχές σύμφωνα με τον επικρατέστερο προσανατολισμό κλίσης, παρόμοια με τον περιγραφέα SIFT. Αντίστοιχα κανονικοποιούμε τις περιοχές για αφινικούς μετασχηματισμούς φωτεινότητας.

Οι περιγραφείς περιοχών που συγκρίνονται στο [148] είναι οι εξής

1. SIFT
2. GLOH
3. PCA-SIFT [111]
4. Spin Images [123]
5. Κατευθυνόμενα φίλτρα (Steerable filters) [60]
6. Διαφορικές αναλλοίωτες (Differential Invariants) [113]
7. Μιγαδικά φίλτρα [192]
8. Αναλλοίωτες ροπών (Moments Invariants) [237]

9. Ετεροσυσχέτιση

10. Shape Context [20]

Ο περιγραφέας *GLOH*, ο οποίος και εισάγεται από τους συγγραφείς του [148], είναι μια επέκταση του περιγραφέα *SIFT*. Χρησιμοποιεί *log-polar* πλέγμα στην περιοχή του σημείου ενδιαφέροντος (Εικόνα 4.4) και μειώνει την διάσταση του περιγραφέα χρησιμοποιώντας *PCA*. Για αυτό και ο υπολογισμός ενός περιγραφέα *GLOH* είναι υπολογιστικά πιο απαιτητικός από τον υπολογισμό περιγραφέα *SIFT*.

Οι περιγραφείς *Shape Context*, βασίζονται επίσης σε ιστογράμματα και είναι παρόμοιοι με τον περιγραφέα *SIFT*. Βασίζονται όμως σε ακμές. Τα ιστογράμματα αφορούν την θέση και τον προσανατολισμό των ακμών. Η διάσταση του περιγραφέα είναι 36.

Για την κατασκευή των περιγραφέων *PCA-SIFT*, υπολογίζουμε την κλίση στις διευθύνσεις x, y της εικόνας και μειώνουμε την διάσταση της περιγραφής σε 36 με *PCA*. Η μικρή διάσταση του περιγραφέα επιτρέπει γρήγορα ταίριασμα, η εφαρμογή *PCA* όμως καθιστά τον υπολογισμό του πιο χρονοβόρο.

Για τους περιγραφείς *Spin Image*, δημιουργούμε ιστογράμματα εντάσεων *pixel* και θέσεων. Χωρίζουμε την περιοχή υπολογισμού του περιγραφέα σε 5 περιοχές και κβαντίζουμε τις εντάσεις των *pixels* σε 10 επίπεδα, δίνοντας στην περιγραφή διάσταση 50.

Στην περίπτωση των *μυγαδικών φίλτρων* χρησιμοποιούμε 15 φίλτρα. Για τις ροπές, υπολογίζουμε τις ροπές της κλίσης της εικόνας στην διεύθυνση x, y μέχρι 2η τάξη και βαθμό.

Στα *κατευθυνόμενα φίλτρα* και τις *διαφορικές αναλλοιώτες*, χρησιμοποιούμε παραγώγους τις οποίες υπολογίζουμε από συνελίξεις με Γκαουσιανές. Οι παράγωγοι υπολογίζονται μέχρι την τέταρτη τάξη (διάσταση περιγραφέα 14) και οι διαφορικές αναλλοιώτες μέχρι την 3 τάξη (διάσταση περιγραφής 8).

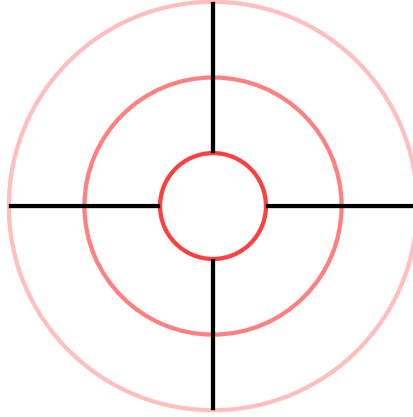
Για να μετρήσουμε την *ομοιότητα των περιγραφέων*, χρησιμοποιούμε κάποιο μέτρο απόστασης. Για τα κατευθυνόμενα φίλτρα, τις διαφορικές αναλλοιώτες, τις αναλλοιώτες ροπών και τα μυγαδικά φίλτρα χρησιμοποιούμε την απόσταση *mahalanobis*. Για τους περιγραφείς που βασίζονται σε ιστογράμματα, χρησιμοποιούμε ευκλείδεια απόσταση.

Αξιολόγηση μεθόδων

Τα αποτελέσματα της αξιολόγησης είναι διαθέσιμα στην διαδικτυακή διεύθυνση [13].

Για την αξιολόγηση των περιγραφέων χρησιμοποιήθηκαν διαφορετικοί γεωμετρικοί και φωτομετρικοί μετασχηματισμοί. Πιο συγκεκριμένα

- Περιστροφή
- Κλιμάκωση
- Αλλαγή θέσης θέασης (viewpoint)
- Θόλωση
- jpeg συμπίεση
- Αλλαγή φωτεινότητας



Σχήμα 4.4: Δομή log-polar grid με διακριτές τις περιοχές συμψηφισμού των τιμών των κλίσεων για την δημιουργία ιστογραμμάτων.

Χρησιμοποιήθηκαν επίσης δομημένες σκηνές με ομογενείς περιοχές και ευδιάκριτα σύνορα αλλά και σκηνές με επαναλαμβανόμενα textures διάφορων μορφών.

Το κριτήριο αξιολόγησης που χρησιμοποιείται είναι αντίστοιχο με αυτό του [111] και βασίζεται στις σωστές και τις λανθασμένες αντιστοιχίες μεταξύ δύο εικόνων. Δύο περιοχές A, B ταιριάζονται αν η απόσταση d μεταξύ των αντίστοιχων περιγραφών D_A, D_B είναι μικρότερη από κατώφλι t . Για κάθε περιγραφέα στην μία εικόνα, ελέγχουμε αν ταιριάζει με κάθε περιγραφέα της άλλης εικόνας. Η αξιολόγηση γίνεται μέσω δύο ανεξάρτητων μεγεθών, της ακρίβειας (precision) και των εύστοχων ταιριασμάτων (recall). Τα μεγέθη αυτά είναι ανεξάρτητα και ορίζονται με τις ακόλουθες σχέσεις

$$1 - \text{precision} = \frac{\# \text{ εσφαλμένα ταιριάσματα}}{\# \text{ εσφαλμένα ταιριάσματα} + \# \text{ σωστά ταιριάσματα}}$$

$$\text{recall} = \frac{\# \text{ σωστά ταιριάσματα}}{\# \text{ ομόλογες περιοχές}}$$

Τα σωστά ταιριάσματα και οι ομόλογες περιοχές καθορίζονται από το σφάλμα επικάλυψης [146]. Το σφάλμα επικάλυψης ϵ_S μετρά πόσο καλά ταιριάζουν δύο περιοχές A, B δεδομένης της ομογραφίας H που συνδέει τις δύο εικόνες και ορίζεται από τον λόγο της τομής και της ένωσης των δύο περιοχών σύμφωνα με τον τύπο

$$\epsilon_S = 1 - \frac{A \cap (H^T B H)}{A \cup (H^T B H)}$$

Θεωρούμε δύο περιοχές ομόλογες, αν το σφάλμα επικάλυψης τους είναι μικρότερο από 50%

Αποτελέσματα

Οι μέθοδοι που δοκιμάζονται για το ταιριασμα των περιγραφών είναι οι εξής

1. Κατώφλι απόστασης. Δύο περιγραφείς ταιριάζονται αν η απόσταση τους είναι μικρότερη από κατώφλι.
2. Κοντινότερος γείτονας. Κάθε περιγραφέας ταιριάζεται με τον κοντινότερο.

3. *Κοντινότερος γείτονας* με χρήση *λόγου αποστάσεων*. Όμοια με το 2, χρησιμοποιούμε όμως και κατώφλι στον λόγο των αποστάσεων του κοντινότερου και του δεύτερου κοντινότερου γείτονα προκειμένου να εγκρίνουμε ένα ταίριασμα.

Το 3 στηρίζεται στην υπόθεση ότι αν το ταίριασμα είναι εσφαλμένο, θα υπάρχουν και άλλοι περιγραφείς σε κοντινή απόσταση. Η υπόθεση αυτή στηρίζεται στην μεγάλη διάσταση του χώρου περιγραφής. Τα 2,3 βελτιώνουν την *ακρίβεια* των περιγραφέων. Η *σχετική κατάταξη* των περιγραφέων όμως *δεν μεταβάλλεται* από την στρατηγική ταιριάσματος που θα εφαρμοστεί.

Όμοια διαπιστώνουμε ότι οι *ανιχνευτές περιοχών* επηρεάζουν τα αποτελέσματα, η *σχετική κατάταξη* των περιγραφέων όμως παραμένει ουσιαστικά αμετάβλητη. Αναφέρουμε όμως ότι οι ανιχνευτές που στηρίζονται στην *Hessian* παρουσιάζουν συστηματικά ελαφρώς καλύτερα αποτελέσματα.

Όσον αφορά το αναλλοίωτο του περιγραφέα σε *μεταβολές φωτισμού*, τα αποτελέσματα είναι σημαντικά βελτιωμένα αν επιτυγχάνεται από την κανονικοποίηση της ανιχνευθίσας περιοχής παρά από την δομή του περιγραφέα.

Για την επιλογή *διάστασης* των περιγραφέων που βασίζονται σε ιστογράμματα, παρατηρήθηκε ότι η επίδοση βελτιώνεται με αύξηση της διάστασης μέχρι ενός ορίου και στην συνέχεια φθίνει με περαιτέρω αύξηση της διάστασης του περιγραφέα.

Τέλος, σημειώνουμε ότι οι *καλύτερες επιδόσεις* σημειώθηκαν συνολικά από τον περιγραφέα *GLOH*, ελαφρώς βελτιωμένες σε σχέση με τις επιδόσεις του σε *SIFT*. Τα αποτελέσματα αυτά δείχνουν την σθεναρότητα και την διακριτική ικανότητα που διαθέτουν οι περιγραφείς που βασίζονται στην σχεδίαση του *SIFT*.

4.3 Βελτίωση του χρόνου υπολογισμού των χαρακτηριστικών. Ο ανιχνευτής SURF

Αν και οι περιγραφείς *GLOH* βελτίωσαν ελαφρώς τις επιδόσεις του *SIFT*, αύξησαν περαιτέρω την πολυπλοκότητα του υπολογισμού των περιγραφέων. Οι συγγραφείς του [18], χρησιμοποιώντας και την μέθοδο αξιολόγησης που περιγράψαμε στην Ενότητα 4.2, σχεδίασαν τον περιγραφέα *SURF* (*Speeded up Robust Features*). Οι *επιδόσεις* των περιγραφέων *SURF* είναι *συγκρίσιμες* σε επίπεδο επαναληψιμότητας, σθεναρότητας και διακριτικότητας σε σχέση με περιγραφείς όπως οι *SIFT* και *GLOH*. Ο *υπολογισμός* τους όμως είναι σημαντικά *ταχύτερος*.

Η *ταχύτητα των υπολογισμών* επιτυγχάνεται με την χρήση ολοκληρωτικών εικόνων (*integral images*) για τον υπολογισμό των συνελίξεων και την κατάλληλη (προκειμένου να μην ζημιωθεί η επίδοση) *απλοποίηση* των περιγραφέων, που βασίζονται σε ιστογράμματα, και των ανιχνευτών, που βασίζονται στην *Hessian*.

Η σθεναρότητα σε μετασχηματισμούς και η πολυπλοκότητα υπολογισμού των περιγραφέων σταθμίζονται με την επιλογή περιγραφέων αμετάβλητων σε κλιμάκωση και περιστροφή. Ακόμα, σύμφωνα με το [133], ο επιπρόσθετος υπολογιστικός φόρτος για την κατασκευή πλήρως αφινικών αναλλοίωτων περιγραφέων δεν δικαιολογείται, αφού πλήττεται η σθεναρότητα της περιγραφής.

Περιγράφοντας συνοπτικά την προσέγγιση που ακολουθήθηκε στους περι-

γραφείς SURF έχουμε:

- Χρήση *ολοκληρωτικών εικόνων* για αποδοτικούς υπολογισμούς.
- Χρήση *ανιχνευτή περιοχών Fast-Hessian*. Στηρίζεται στον ανιχνευτή Hessian-Laplace, ο οποίος παρουσιάζει την καλύτερη επίδοση όπως αναφέραμε και στην Ενότητα 4.2. Προτείνεται κατάλληλο σχήμα προσέγγισης που επιταχύνει τους υπολογισμούς με μικρή επίπτωση στην επίδοση.
- Χρήση *περιγραφέα* που στηρίζεται στα κυματίδια (wavelets) Haar και έχει διάσταση 64, επιτρέποντας γρήγορο υπολογισμό και ταίριασμα.

Οι *ολοκληρωτικές εικόνες* (integral images), επιτρέπουν τον *γρήγορο υπολογισμό συνελίξεων* με συναρτήσεις που έχουν δομή κιβωτίου (box-type). Η εγγραφή $I_{\Sigma}(x, y)$ της ολοκληρωτικής εικόνας I_{Σ} στην θέση (x, y) σχηματίζεται από το άθροισμα όλων των pixels της αρχικής εικόνας I σε μια παραλληλόγραμμη περιοχή που ορίζεται από την αρχή των αξόνων και το σημείο (x, y) . Με μαθηματικό συμβολισμό έχουμε:

$$I_{\Sigma}(x, y) = \sum_{i=0}^{i \leq x} \sum_{j=0}^{j \leq y} I(x, y)$$

Από μια ολοκληρωτική εικόνα, 4 προσθέσεις αρκούν για τον υπολογισμό αθροισμάτων εντός τυχούσης κατακόρυφης, παραλληλόγραμμης περιοχής. Η ανάλυση της εικόνας με *κυματίδια Haar* είναι μια πολυκλιμακωτή ανάλυση με χρήση δενδροειδούς συστοιχίας φίλτρων. Η ακολουθία της εισόδου διαχωρίζεται αρχικά σε δύο εκδοχές : μια αδρή βαθυπερατή εκδοχή και μια λεπτομερή υπερπερατή εκδοχή. Στην συνέχεια η λεπτομερής εκδοχή χωρίζεται σε δύο ακόμα εκδοχές κ.ο.κ. Αναλυτικότερη παρουσίαση των κυματιδίων και των χρήσεων τους στα πλαίσια της πολυκλιμακωτής ανάλυσης γίνεται στο Κεφάλαιο 6 του [223].

Στην συνέχεια αναλύουμε λεπτομερέστερα επιμέρους στοιχεία της σχεδίασης των περιγραφέων SURF.

4.3.1 Χαρακτηριστικά συστήματος ανίχνευσης χαρακτηριστικών SURF

Ο ανιχνευτής περιοχών Fast Hessian

Ο ανιχνευτής περιοχών στηρίζεται στην *Hessian*. Δεν χρησιμοποιούμε διαφορετικό μέτρο για τον προσδιορισμό της θέσης και της κλίμακας των σημείων ενδιαφέροντος (όπως γίνεται στο [145]), αλλά στηριζόμαστε στην *οριζούσα* και για τα δύο. Υπενθυμίζουμε ότι η *Hessian* δίνεται από τον τύπο

$$H(\mathbf{x}, \sigma) = \begin{bmatrix} L_{xx}(\mathbf{x}, \sigma) & L_{xy}(\mathbf{x}, \sigma) \\ L_{xy}(\mathbf{x}, \sigma) & L_{yy}(\mathbf{x}, \sigma) \end{bmatrix}$$

όπου $L_{ij}(\mathbf{x}, \sigma)$ συμβολίσαμε την συνέλιξη της εικόνας I στο σημείο \mathbf{x} με την παράγωγο του Γκαουσιανού πυρήνα σε κλίμακα σ , $\frac{\partial^2 G(\sigma)}{\partial i \partial j}$. Παρά τα γεγονός ότι οι Γκαουσιανές είναι βέλτιστες για ανάλυση στον χώρο κλίμακας ([125]), η *χρήση τους αμφισβητείται* για δύο λόγους

- Σφάλματα *διακριτοποίησης* και *αποκοπής* των φθινουσών "ουρών" της κατανομής.
- Σφάλματα (*aliasing*) λόγω υποδειγματοληψίας.

Έτσι οι *παράγωγοι δεύτερης τάξης* των Γκαουσιανών *προσεγγίζονται* από φίλτρα με δομή *κιβωτίου*, όπως παρουσιάζονται στην Εικόνα 4.5.

Για να αυξήσουμε την *ακρίβεια* της προσέγγισης μας, χρησιμοποιούμε κατάλληλα βάρη στον υπολογισμό της ορίζουσας της Hessian σύμφωνα με την σχέση

$$\text{Det}(H_a) = D_{xx}D_{yy} - (0.9D_{xy})^2 \quad (4.2)$$

Στην Σχέση 4.2 με D_{ij} συμβολίζουμε την προσέγγιση του L_{ij} με την αντικατάσταση των παραγώγων των Γκαουσιανών με φίλτρα δομής κιβωτίου.

Επίσης κανονικοποιούμε τις αποκρίσεις των φίλτρων σε σχέση με το μέγεθος της μάσκας υπολογισμού για να εξασφαλίσουμε σταθερή Frobenius νόρμα, ώστε να είναι κανονικοποιημένα στην κλίμακα [127].

Επειδή χρησιμοποιούμε ολοκληρωτικές εικόνες και φίλτρα με δομή κιβωτίου, ο υπολογισμός των *συνελίξεων* γίνεται σε *σταθερό* χρόνο ανεξάρτητα από το μέγεθος του φίλτρου (και κατά συνέπεια το αντίστοιχο μέγεθος Γκαουσιανής). Συνεπώς δεν υπάρχει λόγος για ακολουθιακό υπολογισμό των *συνελίξεων* σε διαφορετικές κλίμακες και είναι δυνατόν να υπολογίσουμε παράλληλα τις αποκρίσεις.

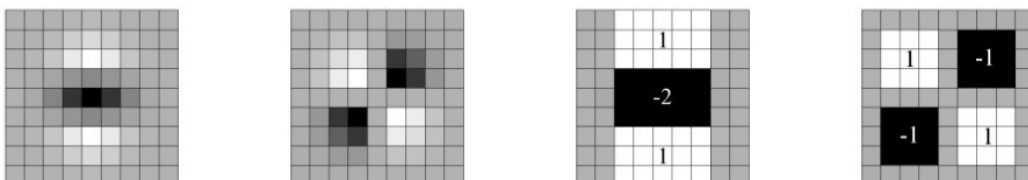
Για τον εντοπισμό των σημείων ενδιαφέροντος, βρίσκουμε τα τοπικά μέγιστα της 4.2 ως προς το διάνυσμα μεταβλητών (x, y, σ) .

Ο περιγραφέας SURF

Ο υπολογισμός του περιγραφέα SURF γίνεται σε δύο βήματα

1. Καθορισμός ενός *προσανατολισμού* που μπορεί να αναπαραχθεί, από κυκλική περιοχή γύρω από το σημείο ενδιαφέροντος.
2. Κατασκευή μιας τετράγωνης περιοχής γύρω από το σημείο ενδιαφέροντος και *εξαγωγή του περιγραφέα SURF* από αυτή.

Ξεκινάμε από το (1), την *ανάθεση προσανατολισμού*. Χρησιμοποιούμε τον μετασχηματισμό κυματιδίων Haar και υπολογίζουμε τις αποκρίσεις στις διευθύνσεις x, y εντός της κυκλικής περιοχής υπολογισμού του προσανατολισμού. Οι αποκρίσεις σταθμίζονται με κατάλληλη Γκαουσιανή που τοποθετείται στο σημείο ενδιαφέροντος. Οι αποκρίσεις εντός παραθύρου συμψηφίζονται, προκειμένου να βρούμε το παράθυρο με την μεγαλύτερη απόκριση,



Σχήμα 4.5: Παράγωγοι δεύτερης τάξης Γκαουσιανών, μετά την διακριτοποίηση και την αποκοπή των ουρών της κατανομής, και προσέγγιση τους από φίλτρα με κυτιοειδή δομή. Από αριστερά προς τα δεξιά: παράγωγος στην διεύθυνση x , παράγωγος στην διεύθυνση xy , προσέγγιση παραγώγου στην διεύθυνση x , προσέγγιση παραγώγου στην διεύθυνση xy . Από [18].

δηλαδή το *μεγαλύτερο κατά μέτρο διάνυσμα*. Η παράμετρος που καθορίζει την *γωνία που καλύπτει* κάθε παράθυρο πρέπει να επιλεγεί προσεκτικά. Μικρές τιμές γωνίας κάλυψης αυξάνουν την επίδραση μεγάλων αποκρίσεων σε μεμονωμένα σημεία. Μεγάλες τιμές γωνίας κάλυψης οδηγούν σε προσανατολισμούς που δεν είναι αξιόπιστοι.

Για την τελική *εξαγωγή του περιγραφέα* (2), κατασκευάζουμε τετράγωνη περιοχή με κέντρο το σημείο ενδιαφέροντος και προσανατολισμό αυτό που βρήκαμε στο 2. Το *μέγεθος* της περιοχής εξαρτάται από την κλίμακα στην οποία ανιχνεύθηκε το σημείο ενδιαφέροντος. Η τετράγωνη περιοχή *χωρίζεται* στην συνέχεια σε πλέγμα 4×4 *υποπεριοχών*. Κάθε υποπεριοχή *δειγματοληπτείται* σε πλέγμα 5×5 και υπολογίζονται για κάθε δείγμα απλά *χαρακτηριστικά* που βασίζονται στα *κυματίδια Haar*. Οι αποκρίσεις σταθμίζονται πάλι από Γκαουσιανή με κέντρο το σημείο ενδιαφέροντος.

Συμβολίζουμε με d_x, d_y τις αποκρίσεις Haar στις διευθύνσεις x, y αντίστοιχα, όπως αυτές υπολογίστηκαν από το 1. Σε κάθε υποπεριοχή συμψηφίζουμε τις αποκρίσεις και δημιουργούμε (ανά υποπεριοχή) το ακόλουθο διάνυσμα χαρακτηριστικών

$$\mathbf{v} = \left(\sum d_x, \sum d_y, \sum |d_x|, \sum |d_y| \right)$$

Ο περιγραφέας SURF προκύπτει από την συνένωση των διανυσμάτων \mathbf{v} όλων των υποπεριοχών και έχει διάσταση

$$4 \times 4 \text{ Υποπεριοχές} \times \frac{4 \text{ Χαρακτηριστικά}}{\text{Υποπεριοχή}} = 64$$

Σημειώνουμε ότι τα χαρακτηριστικά $\sum d_x, \sum d_y$ μας δίνουν πληροφορία για την πολικότητα των αλλαγών της έντασης. Επίσης το διάνυσμα των χαρακτηριστικών SURF (μήκους 64) *κανονικοποιείται σε μοναδιαίο*, προκειμένου να είναι αμετάβλητο σε κλιμάκωση της φωτεινότητας. Το αναλλοίωτο σε όφσει στην φωτεινότητα είναι χαρακτηριστικό των αποκρίσεων κυματιδίων. Πετυχαίνουμε λοιπόν την κατασκευή περιγραφέων αναλλοίωτων σε αφινικούς μετασχηματισμούς φωτεινότητας.

Αν μεταβάλουμε τον *αριθμό του υποπεριοχών* παρατηρούμε

- με *μείωση* των υποπεριοχών, χειρότερη επίδοση και γρηγορότερο ταίριασμα
- με *αύξηση* των υποπεριοχών, μείωση της σθεναρότητας της περιγραφής

Οι καλύτερες επιδόσεις παρατηρήθηκαν με αύξηση της διάστασης του περιγραφέα σε 128. Αυτό έγινε χρησιμοποιώντας 4×4 πλέγμα υποπεριοχών και χρησιμοποιώντας διαφορετικό διάνυσμα χαρακτηριστικών \mathbf{v} . Συγκεκριμένα υπολογίζουμε ξεχωριστά τα αθροίσματα $\sum d_x, \sum |d_x|$ για σημεία με $d_y > 0$ και για σημεία $d_y < 0$. Αντίστοιχα χωρίζουμε τους υπολογισμούς για τα $\sum d_y, \sum |d_y|$ ανάλογα με το πρόσημο του d_x . Οι συγγραφείς όμως προτείνουν την χρήση διανύσματος 64 χαρακτηριστικών όπως αναλύσαμε προηγουμένως, αφού οι επιδόσεις είναι ελαφρά χειρότερες από την χρήση του μεγαλύτερου περιγραφέα (διάστασης 128) και ο χρόνος ταιριάσματος χαρακτηριστικών μικρότερος.

Τέλος για την *επιτάχυνση της διαδικασίας ταιριάσματος*, χρησιμοποιούμε το *πρόσημο της Λαπλασιανής* (ίχνος Hessian). Η Λαπλασιανή ξεχωρίζει Φωτεινές περιοχές σε σκούρο παρασκήνιο από σκούρες περιοχές σε φωτεινό παρασκήνιο. Μόνο περιγραφείς με ίδιο πρόσημο Λαπλασιανής χρειάζεται να συγκριθούν.

Συμπεράσματα

Τα αποτελέσματα της αξιολόγησης του ανιχνευτή και του περιγραφέα μπορούν να βρεθούν στην διαδικτυακή διεύθυνση [13].

Όσον αφορά τον ανιχνευτή περιοχών *Fast Hessian*, παρατηρούμε ότι είναι τρεις φορές ταχύτερος από τον ανιχνευτή περιοχών του [133] που χρησιμοποιεί διαφορές Γκαουσιανών και πέντε φορές ταχύτερος από τον πλήρη ανιχνευτή *Hessian-Laplace* (Ενότητα 4.2). Η επαναληψιμότητα των ανιχνευτών είναι συγκρίσιμη.

Όσον αφορά την αξιολόγηση του περιγραφέα *SURF*, παρατηρούμε συστηματική και σημαντική βελτίωση σε σχέση με τις επιδόσεις των περιγραφέων *GLOH*, *SIFT* και *PCA-SIFT*. Χρησιμοποιώντας τα μέτρα *recall*, *precision* της Ενότητας 4.2, ο περιγραφέας *SURF* υπερέχει (σε ορισμένες περιπτώσεις) έως και 10% στο ποσοστό εύστοχων αποκρίσεων (*recall*) για ίδια τιμή ακρίβειας (*precision*). Παραδείγματα χρόνου υπολογισμού των περιγραφέων παρουσιάζονται στον Πίνακα 4.1.

| | SURF | SURF-128 | SIFT |
|----------------------|-------------|-----------------|-------------|
| Χρόνος σε ms. | 354 | 391 | 1036 |

Πίνακας 4.1: Χρόνοι εκτέλεσης. Χρησιμοποιήθηκε υπολογιστής *Pentium IV 3Ghz*. Οι παράμετροι των μεθόδων μεταβλήθηκαν ώστε να εντοπιστεί ίσος αριθμός σημείων ενδιαφέροντος. Από [18].

4.4 Συστηματοποίηση της διαδικασίας σχεδιασμού ανιχνευτών χαρακτηριστικών.

Η Ενότητα βασίζεται στο [243].

4.4.1 Γενικά στοιχεία

Στις προηγούμενες Ενότητες του Κεφαλαίου, παρουσιάσαμε διαφορετικές προσεγγίσεις σχεδίασης περιγραφέων περιοχών εικόνων. Για την *σχεδίαση ενός περιγραφέα*, χωρίζουμε την διαδικασία σε δύο στάδια

1. την πρόταση συγκεκριμένης *αλγοριθμικής δομής* για το σύστημα ανίχνευσης χαρακτηριστικών.
2. τον προσδιορισμό των *παραμέτρων* του αλγορίθμου 1.

Παρατηρώντας και όσα περιγράψαμε στις προηγούμενες Ενότητες, διαπιστώνουμε ότι η διαδικασία πρότασης(1) ενός *αλγορίθμου* περιγραφέα αντιμετωπίζεται ως *ad-hoc πρόβλημα*. Για παράδειγμα η βασική δομή του αλγορίθμου *SIFT* που περιγράψαμε στην Ενότητα 4.1 διατηρείται στους περιγραφείς *GLOH* και *SURF* (Ενότητες 4.2, 4.3 αντίστοιχα), μεταβάλλονται όμως στοιχεία του με μη συστηματικό τρόπο.

Αντίστοιχα, η *ρύθμιση των παραμέτρων* του αλγορίθμου(2) γίνεται "με το χέρι", διαδικασία που ιδανικά θα θέλαμε να είναι αυτοματοποιημένη. Για

παράδειγμα στο [133] η ρύθμιση των παραμέτρων γίνεται με την επισκόπηση πολυάριθμων γραφικών παραστάσεων, διαδικασία που δεν μπορεί να γενικευθεί σε μεγαλύτερη κλίμακα υπό εξέταση αλγορίθμων.

Στην συνέχεια περιγράφουμε μια *προσπάθεια συστηματοποίησης* της σχεδίασης αλγορίθμων εντοπισμού και περιγραφής χαρακτηριστικών εικόνων και αυτοματοποίησης της ρύθμισης των παραμέτρων του. Έτσι επιχειρείται μια συστηματική εξερεύνηση του χώρου των αντίστοιχων αλγορίθμων. Οι αλγόριθμοι χωρίζονται σε δομικά στοιχεία (μπλοκς) και οι παράμετροι προσαρμόζονται χρησιμοποιώντας σύνολο εικόνων από το πρόβλημα της τρισδιάστατης ανακατασκευής. Οι περιγραφείς που περιγράψαμε στις προηγούμενες Ενότητες εντάσσονται στο πλαίσιο που θα εισάγουμε.

Συνοψίζοντας, οι κύριες *συνεισφορές* είναι:

- *Σύνολο δεδομένων* για τον έλεγχο και την βελτιστοποίηση περιγραφών, με ακριβή δεδομένα επαλήθευσης. Το σύνολο αυτό επεκτείνει προσεγγίσεις όπως της Ενότητας 4.2, που περιορίζονται σε επίπεδους προβολικούς μετασχηματισμούς (ομογραφίες), σε τρισδιάστατους μετασχηματισμούς.
- Μια γενική και συστηματική *σχεδιαστική προσέγγιση*. Ο αλγόριθμος χωρίζεται σε επιμέρους δομικές ενότητες (modules) τα οποία συνδυάζονται για τον σχηματισμό ενός αλγορίθμου ανίχνευσης και περιγραφής χαρακτηριστικών.
- Χρήση (αυτόματης) *μάθησης* για τον προσδιορισμό των παραμέτρων.

Στην συνέχεια αναλύουμε την προτεινόμενη μέθοδο σχεδίασης και εξετάζουμε τα αποτελέσματά της.

4.4.2 Δομικά στοιχεία (blocks) αλγορίθμου

Για να κατασκευάσουμε έναν περιγραφέα, ακολουθούμε την *γενική αλγοριθμική δομή* που παρουσιάζεται στην Εικόνα 4.6. Κάθε δομική μονάδα μπορεί να υλοποιηθεί με *πολλούς διαφορετικούς υποψήφιους αλγορίθμους*. Μερικές προτάσεις υποψήφιων αλγορίθμων θα παρουσιάσουμε στις επόμενες υποενότητες. Επίσης ορισμένοι υποψήφιοι αλγόριθμοι έχουν *ελεύθερες παραμέτρους* τις οποίες επιθυμούμε να θέσουμε σε βέλτιστες τιμές.

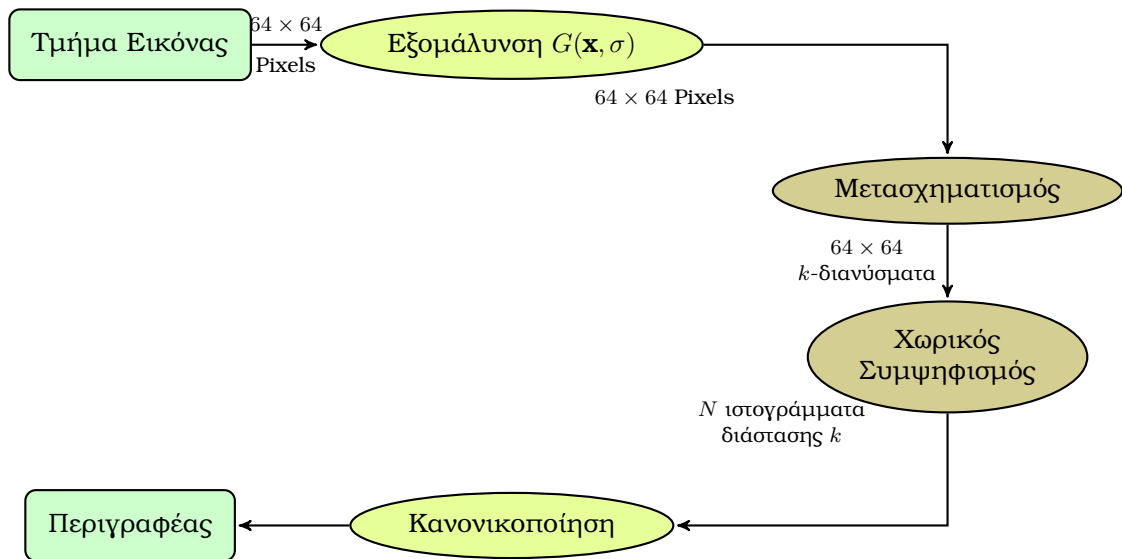
Προκειμένου να συστηματοποιήσουμε τους αλγορίθμους κατασκευής περιγραφών, θεωρούμε προδιαγραφές για την είσοδο και την έξοδο του αλγορίθμου. Σύμφωνα με αυτές έχουμε

- *Είσοδος*: Περιοχές της εικόνας μεγέθους 64×64 pixels.
- *Έξοδος*: Διάνυσμα (περιγραφέας) $D = kN$ αριθμών.

Συνεχίζουμε με την περιγραφή των δομικών μονάδων και την συνοπτική παρουσίαση υποψήφιων αλγορίθμων.

Προεπεξεργασία

Χρησιμοποιούμε Γκαουσιανό πυρήνα με τυπική απόκλιση σ_S για την *εξομάλυνση* της εικόνας πριν την επεξεργασία της από τον αλγόριθμο εξαγωγής περιγραφέα.



Σχήμα 4.6: Στάδια αλγορίθμου κατασκευής περιγραφέα. Αναπαραγωγή από [243]

Μετασχηματισμός (T)

Το στάδιο του μετασχηματισμού απεικονίζει την εξομαλυμένη εικόνα σε ένα πλέγμα αποτελούμενο από δείγματα. Κάθε δείγμα περιγράφεται από ένα διάνυσμα k θετικών αριθμών. Είναι αποδεκτός μεγάλος αριθμός γραμμικών και μη γραμμικών σταδίων μετασχηματισμού. Οι αλγόριθμοι που προτείνονται στο [243] είναι

1. Υπολογίζουμε το μέτρο m και τον προσανατολισμό θ της κλίσης σε κάθε δείγμα. Στην συνέχεια κβαντίζουμε την διεύθυνση σε k προσανατολισμούς. Το μέτρο της κλίσης κατανέμεται γραμμικά στα δύο γειτονικά στοιχεία θ_i, θ_{i+1} του k -διανύσματος προσανατολισμών κλίσης που περιβάλλουν τον προσανατολισμό του δείγματος θ , δηλαδή ικανοποιούν

$$\theta_i \leq \theta \leq \theta_{i+1}$$

Η κατανομή του m στα στοιχεία του διανύσματος γίνεται ανάλογα με την εγγύτητα του προσανατολισμού θ στον αντίστοιχο προσανατολισμό θ_i . Η προσέγγιση αυτή χρησιμοποιείται στους αλγόριθμους *SIFT* και *GLOH* που αναλύσαμε προηγουμένως.

2. Χρησιμοποιούμε το $k = 4$ διάνυσμα

$$\{|\nabla_x| - \nabla_x, |\nabla_x| + \nabla_x, |\nabla_y| - \nabla_y, |\nabla_y| + \nabla_x\}$$

Μπορούμε να χρησιμοποιήσουμε και διάνυσμα διάστασης $k = 8$ υπολογίζοντας την κλίση σε προσανατολισμούς που απέχουν 45° .

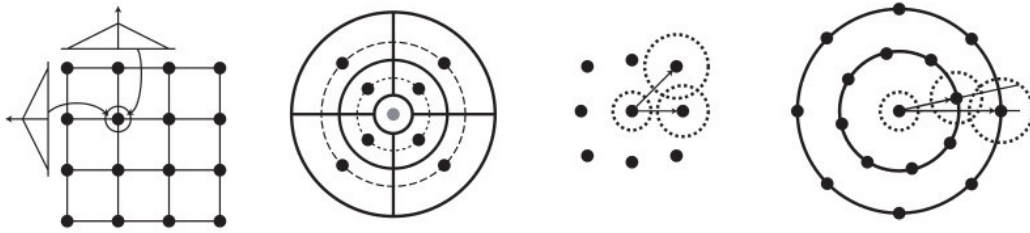
3. Χρήση κατευθυνόμενων (*steerable*) φίλτρων σε k προσανατολισμούς. Τα στοιχεία αυτής της υλοποίησης μπορούν να αναζητηθούν στο [60]. Οι αποκρίσεις των φίλτρων μπορούν να χρησιμοποιηθούν και σε συνδυασμό με το (2) για να δώσουν διάνυσμα διάστασης $4k$. Δοκιμάστηκαν φίλτρα με δευτεροβάθμιες και τεταρτοβάθμιες παραγώγους. Οι τεταρτοβάθμιες παράγωγοι εμφανίζουν μεγαλύτερη διακριτική ικανότητα για τους προσανατολισμούς.

4. Χρήση δύο ιστροπικών Διαφορών Γκαουσιανών (DoG) σε διαφορετικές κλίμακες. Στην συνέχεια συνδυάζουμε τις δύο εξόδους όπως στο (2) για να πάρουμε διάνυσμα $k = 4$ χαρακτηριστικών. Σε κάθε DoG επιλέγουμε σταθερή διαφορά κλίμακας σ_1/σ_2 ανάμεσα στις Γκαουσιανές και στο πρώτο ζευγάρι χρησιμοποιούμε την Γκαουσιανή της προεπεξεργασίας. Συνεπώς μόνο μία Γκαουσιανή για το δεύτερο ζευγάρι DoG πρέπει να προσδιοριστεί.
5. Υπολογίζουμε τον 4×4 μετασχηματισμό *κυμαιδίων Haar* και παίρνουμε τους 15 (χωρίς την DC συνιστώσα) συντελεστές. Πριν την εφαρμογή του μετασχηματισμού κανονικοποιούμε την είσοδο του αλγορίθμου ώστε να έχει μηδενική μέση τιμή και μοναδιαία τυπική απόκλιση.
6. Χρησιμοποιούμε *k-means clustering* ώστε να εντάξουμε την είσοδο του αλγορίθμου σε κάποιο cluster. Χρησιμοποιούμε $k = 16$ clusters. Για την τεχνική του k-means clustering, παραπέμπουμε τον αναγνώστη στο [223].
7. Κανονικοποιούμε την είσοδο του αλγορίθμου ώστε να πετύχουμε αναλλοίωτη περιγραφή σε αφινικούς μετασχηματισμούς της έντασης. Στην συνέχεια κβαντίζουμε την ένταση σε k κατηγορίες και κατανέμουμε την ένταση κάθε δείγματος σε κατηγορίες αντίστοιχα με το (1). Το στάδιο αυτό είναι εμπνευσμένο από τους περιγραφείς *spin image* [123].

Χωρικός συμψηφισμός (S)

Σε αυτό το στάδιο, *συμψηφίζουμε* χωρικά τα k -διανύσματα που σχηματίσαμε στο προηγούμενο στάδιο και καταλήγουμε σε N k -διανύσματα. Ο τελικός *περιγραφέας* D προκύπτει από την *συνένωση* των N διανυσμάτων. Για το N χρησιμοποιούνται κατά περίπτωση οι τιμές $N \in \{4, 9, 16, 17, 25\}$. Παραδείγματα σχημάτων χωρικού συμψηφισμού παρουσιάζονται στην Εικόνα 4.7. Οι αλγόριθμοι που προτείνονται στο [243] είναι

1. Χρήση *τετράγωνου πλέγματος* κέντρων συμψηφισμού. Τα διανύσματα του προηγούμενου σταδίου αθροίζονται σταθμισμένα με κατάλληλα βάρη, σύμφωνα με τις αποστάσεις τους από τα κέντρα συμψηφισμού. Το στάδιο αυτό χρησιμοποιείται στον αλγόριθμο SIFT.
2. *Πολική διάταξη* των περιοχών άθροισης. Τα διανύσματα (παρόμοια με το 1) αθροίζονται σταθμισμένα με διγραμμική παρεμβολή σε πολικές συντεταγμένες, ώστε να συνεισφέρουν σε περισσότερες από μία περιοχές. Το διάνυσμα κάθε περιοχής κανονικοποιείται από την επιφάνεια της περιοχής. Το στάδιο αυτό χρησιμοποιείται στον αλγόριθμο GLOH.
3. *Χρήση Γκαουσιανών για το άθροισμα των διανυσμάτων* σε κάθε κέντρο συμψηφισμού. Το μέγεθος του πλέγματος θεωρείται γνωστό, η θέση των κέντρων συμψηφισμού και τα μεγέθη των Γκαουσιανών όμως είναι παράμετροι προς μάθηση.
4. *Όμοια με το (3)* αλλά αναγκάζουμε τα δείγματα να έχουν *πολική διάταξη*.



Σχήμα 4.7: Παρουσίαση μεθόδων χωρικού συμψηφισμού. Από αριστερά προς τα δεξιά και σύμφωνα με την προηγούμενη αρίθμηση των μεθόδων: 1, 2, 3, 4. Από [243].

Τελική κανονικοποίηση (N)

Στο τελευταίο αυτό στάδιο, *κανονικοποιούμε* τον περιγραφέα για να απαλείψουμε την εξάρτηση της περιγραφής από την αντίθεση (*contrast*) της εικόνας. Ακολουθούμε την διαδικασία του αλγορίθμου SIFT, συγκεκριμένα

- *Κανονικοποιούμε* τον περιγραφέα σε μοναδιαίο διάνυσμα.
- *Ψαλιδίζουμε* τα στοιχεία του περιγραφέα που υπερβαίνουν κατώφλι t .
- *Κανονικοποιούμε* ξανά σε μοναδιαίο διάνυσμα.

Η τιμή του κατώφλιου υποβλήθηκε σε μάθηση.

4.4.3 Αποτελέσματα-Συμπεράσματα

Για την *αξιολόγηση* των αλγορίθμων, φτιάχνουμε τα ιστογράμματα της *ευκλείδειας* απόστασης των περιγραφέων για τις περιπτώσεις ταιριάσματος και μη-ταιριάσματος. Όσο *μικρότερη επικάλυψη* έχουν αυτά τα ιστογράμματα, τόσο *καλύτερη* περιγραφή έχουμε κατασκευάσει. Για την αυτοματοποιημένη αξιολόγηση της επικάλυψης, χρησιμοποιούμε τις καμπύλες ROC.

Ας παρουσιάσουμε σύντομα και *θεωρητικά* τις καμπύλες ROC. Έστω πρόβλημα κατηγοριοποίησης σε δύο κλάσεις A, B . Κάνουμε την γραφική παράσταση των σωστών κατηγοριοποιήσεων στην κλάση B (y -άξονας) και των λάθος κατηγοριοποιήσεων (x -άξονας) στην κλάση A . Μεταβάλλοντας το κατώφλι απόφασης t για τον χαρακτηρισμό ενός διανύσματος χαρακτηρισικών στην κάθε κατηγορία (στην συγκεκριμένη αξιολόγηση το κατώφλι αφορά την ευκλείδεια απόσταση) σχηματίζουμε την καμπύλη ROC. Αν οι κατανομές των κατηγοριών A, B είναι απόλυτα επικαλυπτόμενες προκύπτει η ευθεία $y = x$. Μεγιστοποιώντας το εμβαδό υπό την καμπύλη παίρνουμε τον καλύτερο ταξινομητή. Μια ενδελεχής ανάλυση των καμπύλων ROC γίνεται στο Κεφάλαιο 5 του [223].

Η *μεγιστοποίηση του εμβαδού υπό την καμπύλη ROC* χρησιμοποιήθηκε για την αυτόματη μάθηση των παραμέτρων. Οι παράμετροι προς μάθηση παρουσιάζονται στον Πίνακα 4.2. Επίσης στα T-blocks (Μετασχηματισμού) μεταβάλαμε το μέγεθος k των περιγραφέων και στα S-blocks (Χωρικού συμψηφισμού) το πλήθος N των περιοχών. Οι παράμετροι αυτοί όμως δεν υποβλήθηκαν σε μάθηση.

Όσον αφορά τα *αποτελέσματα*, σημειώνουμε τρεις παρατηρήσεις

- Για τα T-blocks, η αύξηση της διάστασης k γενικά βελτιώνει την ακρίβεια των περιγραφέων. Καλά αποτελέσματα παρουσιάζονται με χρήση

| Δομική μονάδα | Περιγραφή Παραμέτρου |
|---------------|---|
| T4 | Μέγεθος κεντρικής Γκαουσιανής δεύτερου ζεύγους DoG |
| S2 | Ακτίνες των περιοχών ανά επίπεδο. |
| S3 | Μέγεθος Γκαουσιανών & Θέσεις των δειγμάτων του πλέγματος |
| S4 | Ακτίνες για κάθε επίπεδο-δακτύλιο & Μέγεθος Γκαουσιανού πυρήνα που χρησιμοποιείται στα δείγματα κάθε δακτυλίου & Γωνίες για την χωρική τοποθέτηση των δειγμάτων του μεσαίου δακτυλίου |
| N | Κατώφλι ψαλιδισμού t |

Πίνακας 4.2: Παράμετροι που ρυθμίζονται με αυτόματη μάθηση. Συμβολισμός: τα γράμματα συμβολίζουν την ενότητα στην οποία ανήκει η δομική μονάδα. Η αρίθμηση ταυτίζεται με αυτή της αντίστοιχης ενότητας.

των μπλοκς T1 (προσανατολισμός κλίσης), T2(κλίση σε διεύθυνση x, y), T4(Διαφορές Γκαουσιανών). Εκπληξη αποτελεί η επίδοση των κατευθυνόμενων φίλτρων όταν οι έξοδοι των ορθογώνιων ζευγαριών αποθηκεύονται χωριστά ανάλογα με το πρόσημο τους (διατήρηση *πληροφορίας φάσης*). Η διάσταση της περιγραφής είναι όμως μεγάλη.

- Για τα S-blocks, διαπιστώθηκε η υπεροχή των log-polar διατάξεων
- Το κατώφλι ψαλιδισμού της τελικής κανονικοποίησης επηρεάζει σημαντικά την επίδοση των περιγραφών.

Ενδεικτικά αναφέρουμε ότι με *log-polar* διάταξη περιοχών και χαρακτηριστικά από *κατευθυνόμενα φίλτρα*, παρατηρήθηκε μείωση των εσφαλμένων ταιριασμάτων στο $1/3$ των λαθών του SIFT.

4.5 DAISY: Ακριβής και αποδοτικός υπολογισμός χαρακτηριστικών.

Στην Ενότητα αυτή θα αναφερθούμε στους περιγραφείς DAISY [225]. Την χρήση των περιγραφών DAISY σχολιάζουμε στην Ενότητα 3.4.2. Ο σχεδιασμός των περιγραφών DAISY στηρίζεται στις παρατηρήσεις που παρουσιάσαμε στην Ενότητα 4.4. Οι περιγραφείς DAISY είναι σχεδιασμένοι για πυκνό (σε κάθε pixel) υπολογισμό και χρησιμοποιήθηκαν, κατά την εισαγωγή τους, στο πρόβλημα της *τρισδιάστατης ανακατασκευής*. Παραδείγματα πυκνής χρήσης των περιγραφών SIFT και GLOH δίνονται στα [133, 148].

Οι περιγραφείς DAISY προκύπτουν από κατάλληλη *διαμόρφωση* των περιγραφών SIFT/GLOH προκειμένου να είναι *αποδοτικός* ο υπολογισμός τους σε κάθε pixel. Για αυτό πρέπει να υπολογίσουμε τα *ισογράμματα* μόνο *μία φορά* σε κάθε περιοχή και όχι να επαναλαμβάνεται ο υπολογισμός τους σε κάθε pixel. Η προηγούμενη απαίτηση ικανοποιείται με την αλλαγή του

τρόπου υπολογισμού των ιστογραμμάτων. Σε περιγραφείς όπως οι SIFT και GLOH, τα ιστογράμματα προκύπτουν από σταθμισμένα αθροίσματα (με τα σταθμά να εξαρτώνται από το πλάτος της κλίσης σε κάθε σημείο) των προσανατολισμών της κλίσης στην γειτονιά κάθε σημείου. Αυτά τα αθροίσματα αντικαθίστανται από *συνελίξεις* με Γκαουσιανές σε κατάλληλες διευθύνσεις. Επίσης, σημαντική επιρροή στην σχεδίαση του περιγραφέα DAISY ήταν η έννοια του γεωμετρικού θολώματος (*geometric blur*), όπως παρουσιάζεται στο [21]. Σύμφωνα με αυτή την σχεδιαστική επιταγή, προτείνεται η χρήση μικρότερων πυρήνων κοντά στο κέντρο υπολογισμού του περιγραφέα και μεγαλύτερων μακριά από αυτόν.

Η διάταξη των *κέντρων συμψηφισμού* (Ενότητα 4.4) που χρησιμοποιείται παρουσιάζεται στην Εικόνα 4.8. Επίσης στο Πίνακα 4.3 παρουσιάζονται οι παράμετροι του περιγραφέα.

Στον τομέα των *επιδόσεων*, αναφέρουμε ότι διατηρείται η *σθεναρότητα* των περιγραφέων SIFT και GLOH και επιτυγχάνονται *συγκρίσιμα αποτελέσματα*.

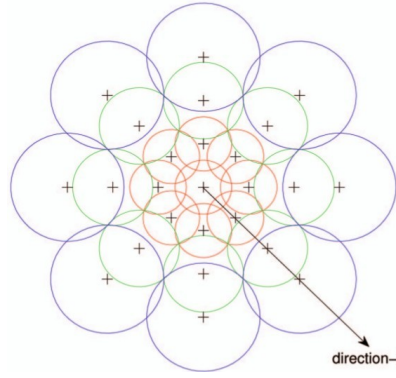
| Παράμετρος | Σύμβολο | Περιγραφή και τυπικές τιμές |
|---------------------------|---------|--|
| Ακτίνα | R | Απόσταση από το κεντρικό pixel στο πιο απομακρυσμένο σημείο του πλέγματος (15) |
| Ακτινική κβάντιση | Q | Πλήθος επιπέδων με διαφορετικό Σ πυρήνα συνέλιξης (3) |
| Γωνιακή κβάντιση | T | Ιστογράμματα σε κάθε επίπεδο (8) |
| Κβάντιση ιστογραμμάτων | H | Κατηγορίες (προσανατολισμοί) σε κάθε ιστόγραμμα (8) |
| Αριθμός σημείων πλέγματος | S | Ιστογράμματα στον περιγραφέα ($Q \cdot T + 1$) |
| Μέγεθος περιγραφέα | D_S | Συνολικό μέγεθος περιγραφέα ($S \cdot H$) |

Πίνακας 4.3: Παράμετροι περιγραφέα Daisy. Από [225].

4.5.1 Δομή περιγραφέα DAISY

Ας θεωρήσουμε μια δεδομένη *εικόνα εισόδου* I στην οποία θέλουμε να υπολογίσουμε τους περιγραφείς DAISY. Υπολογίζουμε αρχικά H *χάρτες προσανατολισμού*

$$G_i, i = 1, \dots, H$$



Σχήμα 4.8: Δομή περιγραφέα Daisy. Οι κύκλοι συμβολίζουν Γκαουσιανές, με την ακτίνα των κύκλων να είναι ανάλογη της κλίμακας της αντίστοιχης Γκαουσιανής. Από [225].

Εξηγούμε τον προηγούμενο συμβολισμό

$$G_o = \left(\frac{\partial I}{\partial o} \right)^+$$

$$(\alpha)^+ = \max(\alpha, 0)$$

Κάθε χάρτης προσανατολισμού G_i συνελίσσεται με Γκαουσιανούς πυρήνες G^Σ σε διάφορες κλίμακες Σ

$$G_i^\Sigma = G^\Sigma * G_o$$

$$= G^\Sigma * \left(\frac{\partial I}{\partial i} \right)^+$$

Υπενθυμίζουμε ότι οι συνελίξεις σε μεγάλες κλίμακες μπορούν να υπολογιστούν από διαδοχικές συνελίξεις με μικρότερους πυρήνες με αποδοτικό υπολογιστικά τρόπο.

Παρατηρούμε επίσης ότι οι τιμές του χάρτη G_i^Σ στην τοποθεσία (x, y) προκύπτουν από το σταθμισμένο (μέσω της συνέλιξης με τον Γκαουσιανό πυρήνα G^Σ) άθροισμα της κλίσης της εικόνας στον προσανατολισμό i , G_i . Έτσι είναι ευθύς η *αντιστοιχία των χαρτών G_i^Σ* με τις τιμές των *ιστογραμμάτων των περιγραφέων SIFT/GLOH*. Κάθε χάρτης G_i^Σ περιέχει παρόμοιες τιμές με την κατηγορία (bin) του ιστογράμματος που αντιστοιχεί στην κλίση στον προσανατολισμό i .

Συμβολίζουμε με $\mathbf{h}_\Sigma(u, v)$ το διάνυσμα που αποτελείται από τις τιμές των χαρτών G_i^Σ , $i = 1, \dots, H$ στην θέση (u, v) της εικόνας I , δηλαδή

$$\mathbf{h}_\Sigma(u, v) = [G_1^\Sigma(u, v) \ \dots \ G_H^\Sigma(u, v)]^T$$

Κανονικοποιούμε το \mathbf{h}_Σ σε διάνυσμα μοναδιαίου μέτρου $\hat{\mathbf{h}}_\Sigma$. Επιλέγουμε να κανονικοποιήσουμε σε αυτό το σημείο ώστε να μειώσουμε την επιρροή εμποδίων που ενδεχομένως να περιορίζουν την ορατότητα σε κάποιες φωτογραφίες.

Συμβολίζουμε με Q το πλήθος των διαφορετικών *κυκλικών επιπέδων* και

έχουμε τον περιγραφέα DAISY $D(u, v)$ στην τοποθεσία (u, v)

$$D(u, v) = \begin{aligned} & [\hat{\mathbf{h}}_{\Sigma_1}^T(u, v), \\ & \hat{\mathbf{h}}_{\Sigma_1}^T(\mathbf{l}_1(u, v, R_1)), \dots, \hat{\mathbf{h}}_{\Sigma_1}^T(\mathbf{l}_T(u, v, R_1)), \\ & \hat{\mathbf{h}}_{\Sigma_2}^T(\mathbf{l}_1(u, v, R_2)), \dots, \hat{\mathbf{h}}_{\Sigma_2}^T(\mathbf{l}_T(u, v, R_2)), \\ & \dots \\ & \hat{\mathbf{h}}_{\Sigma_Q}^T(\mathbf{l}_1(u, v, R_Q)), \dots, \hat{\mathbf{h}}_{\Sigma_Q}^T(\mathbf{l}_T(u, v, R_Q))] \end{aligned} \quad (4.3)$$

Παρατηρούμε στην Σχέση 4.3 την μεταβολή της κλίμακας του Γκαουσιανού πυρήνα μεταξύ των περιοχών R_i , δηλαδή την εφαρμογή του γεωμετρικού θολώματος που αναφέραμε προηγουμένως. Επίσης ο συμβολισμός

$$\mathbf{l}_i(u, v, R)$$

δηλώνει την τοποθεσία που βρίσκεται σε απόσταση R από το (u, v) και σε προσανατολισμό i .

Όπως αναφέραμε και νωρίτερα στην ενότητα, η επιλογή *κυκλικής διάταξης δειγματοληψίας* στην γειτονία του σημείου υπολογισμού του περιγραφέα DAISY δίνει καλύτερα αποτελέσματα από άλλες διατάξεις. Επίσης η κυκλική διάταξη της γειτονίας σε συνδυασμό με την χρήση *ισοτροπικών Γκαουσιανών* κάνουν την περιγραφή *σθεναρή σε περιστροφή*. Τέλος η χρήση *επικαλυπτόμενων περιοχών* αυξάνει την *σθεναρότητα* του περιγραφέα, όμως η υπερβολική αύξηση της επικάλυψης οδηγεί σε απώλεια διακριτικής ικανότητας.

Πολυπλοκότητα Υπολογισμού περιγραφέα DAISY

Θα αναφερθούμε τώρα στην *υπολογιστική πολυπλοκότητα* του αλγορίθμου. Αναφέρουμε για συντομία κατευθείαν τα αποτελέσματα, αναλυτικά οι υπολογισμοί μπορούν να αναζητηθούν στο [225]. Για τον υπολογισμό του περιγραφέα DAISY απαιτούνται:

- $2H \times Q + 1$ μονοδιάστατες συνελίξεις
- $P(Q + T + 1)$ δειγματοληψίες
- $2P \times H$ πολλαπλασιασμοί
- $P \times H$ προσθέσεις

Αντίστοιχα για τον υπολογισμό του περιγραφέα SIFT απαιτούνται:

- W_s Δειγματοληψίες
- $5W_s$ Πολλαπλασιασμοί
- $2W_s$ Αθροίσματα

όπου W_s είναι ο αριθμός δειγμάτων στην γειτονία του σημείου που υπολογίζουμε τον περιγραφέα. Χρησιμοποιώντας τις παραμέτρους που προτείνονται για τους περιγραφείς SIFT στο [133] και για τους περιγραφείς DAISY στο [225], μπορούμε να αντικαταστήσουμε τις παραμέτρους στις προηγούμενες σχέσεις και να οδηγηθούμε στα στοιχεία που παρουσιάζονται στον Πίνακα 4.4. Στον Πίνακα 4.5 παρουσιάζουμε *χρόνους υπολογισμού* για διάφορα μεγέθη εικόνων.

| | DAISY | SIFT |
|-----------------|-------|------|
| Πολλαπλασιασμοί | 122 | 1280 |
| Προσθέσεις | 119 | 512 |
| Δειγματοληψίες | 25 | 256 |

Πίνακας 4.4: Πλήθος πράξεων (ανά κατηγορία) για τον υπολογισμό των περιγραφών DAISY και SIFT . Από [225].

| Μέγεθος Εικόνας | DAISY | SIFT |
|-----------------|-------|------|
| 800 × 600 | 3.8 | 252 |
| 1024 × 768 | 6.5 | 432 |
| 1280 × 900 | 9.8 | 651 |

Πίνακας 4.5: Χρόνοι (σε second) για τον υπολογισμό των περιγραφών DAISY και SIFT. Από [225].

Ερμηνεία και ρύθμιση παραμέτρων περιγραφέα DAISY

Σχολιάζουμε τέλος τις *παραμέτρους* του περιγραφέα Daisy. Ο *περισσότερος χρόνος* δαπανάται στον υπολογισμό των *συνελίξεων*. Επομένως, προκειμένου να βελτιώσουμε τον χρόνο υπολογισμού του περιγραφέα DAISY προτείνεται η μείωση του αριθμού των προσανατολισμών H και των επιπέδων της γειτονιάς Q .

Επειδή οι περιγραφείς DAISY δημιουργήθηκαν με σκοπό την εφαρμογή τους στο στερεοσκοπικό ταίριασμα μεταξύ φωτογραφιών, οι παράμετροι του συστήματος συνίσταται να *μεταβάλλονται* ανάλογα με τα *χαρακτηριστικά των φωτογραφιών*. Τα χαρακτηριστικά που μας ενδιαφέρουν είναι η *υφή*, η *πολυπλοκότητα της σκηνής* και η *απόσταση των δύο όψεων* που χρησιμοποιούμε στο ταίριασμα περιγραφών.

Για τις παραμέτρους Q, T, H μπορούμε να καταλήξουμε ευκολότερα στις τιμές τους. Οι παράμετροι Q, T δεν επηρεάζονται από την απόσταση των όψεων στον χώρο. Αντίθετα, θα πρέπει να χρησιμοποιούμε μεγαλύτερο αριθμό κυκλικών δακτυλίων (Q) στην περίπτωση που οι όψεις είναι απομακρυσμένες στον χώρο. Ο ρόλος του R δεν είναι τόσο ξεκάθαρος. Πάντως παρατηρήθηκε ότι μεγάλες τιμές της παραμέτρου είναι ευνοϊκές σε σκηνές που δεν έχουν έντονη υφή.

4.6 Επεκτάσεις και διαθέσιμο λογισμικό

4.6.1 Επεκτάσεις

Στην Ενότητα αυτή παρουσιάζουμε συγκεκριμένους *άξονες* στους οποίους δομείται η *ανάπτυξη* και *βελτίωση* συστημάτων ανίχνευσης και περιγραφής χαρακτηριστικών σε εικόνες. Επιχειρούμε να *συνδέσουμε* την πορεία αυτή με τις διάφορες *προσεγγίσεις* που αναλύσαμε προηγουμένως.

Ξεκινάμε με την έννοια του *γεωμετρικού δολώματος* [21]. Στο πρόβλημα της

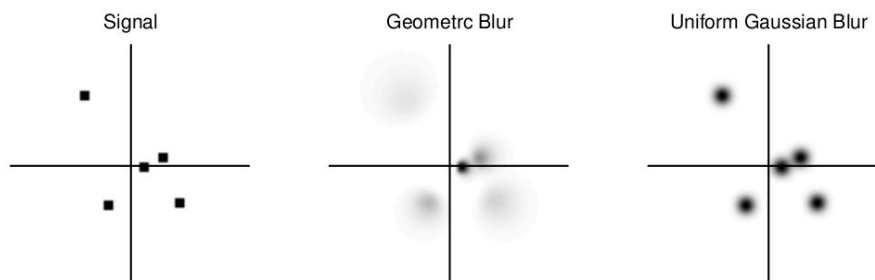
εύρεσης αντιστοιχιών μεταξύ εικόνων που συνδέονται με αφινικούς μετασχηματισμούς, συνηθίζεται να εξομαλύνουμε με κάποιον συνελκτικό πυρήνα τις εικόνες και να ταιριάζουμε τα χαρακτηριστικά σε περισσότερο "τραχύ" επίπεδο. Η χρήση Γκαουσιανού πυρήνα για την εξομάλυνση, έχει σαν αποτέλεσμα όλα τα σημεία γύρω από το σημείο ενδιαφέροντος να γίνονται, ομοιόμορφα, χωρικά αβέβαια κατά ένα παράγοντα σ . Ο συγκεκριμένος τρόπος εξομάλυνσης δεν έχει σχεδιαστεί υπό το πρίσμα της αντιστοίχισης σημείων μεταξύ εικόνων που υπόκεινται σε *γεωμετρικούς μετασχηματισμούς*. Το γεωμετρικό θόλωμα επιχειρεί την *εξομάλυνση* των εικόνων με τρόπο *βέλτιστο για αυτές τις περιπτώσεις*. Προτείνεται η συνέλιξη με Γκαουσιανούς πυρήνες με μεταβλητό σ , ανάλογα με την απόσταση των σημείων από το σημείο ενδιαφέροντος. Μια οπτικοποίηση του γεωμετρικού θολώματος με Γκαουσιανούς πυρήνες, δίνεται στην Εικόνα 4.9.

Μια εφαρμογή της ιδέας του Γεωμετρικού Θολώματος είδαμε στην Ενότητα 4.5. Θυμίζουμε την μεταβολή της τυπικής απόκλισης των Γκαουσιανών πυρήνων, όσο κινούμαστε προς την περιφέρεια του περιγραφέα.

Στο [152], οι συγγραφείς λύνουν το πρόβλημα της εύρεσης του *σωστού συνελκτικού πυρήνα*, για κοινά μοντέλα *κίνησης*. Στα αποτελέσματά τους αναφέρουν σημαντική βελτίωση των επιδόσεων. Συγκεκριμένα, προτείνουν την *εξομάλυνση* της αντικειμενικής *συνάρτησης* του προβλήματος ευθυγράμμισης εικόνων (*image alignment*) που εξετάζουν, αντί για την εξομάλυνση των εικόνων.

Στις εργασίες [114, 115], που προέρχονται από την ίδια ομάδα ερευνητών, εξετάζεται το θέμα της ανάλυσης σε *κλίμακες*. Συγκεκριμένα, η *επιλογή κλίμακας* για τα χαρακτηριστικά είναι συχνά αναξιόπιστη και (πάντα) περιοριστική, οπότε επιχειρείται να *παρακαμφθεί* με κατάλληλο τρόπο. Στο [114], κατασκευάζονται περιγραφείς κατάλληλοι για πυκνό υπολογισμό (σε κάθε pixel της εικόνας), αναλλοίωτοι σε περιστροφή και κλιμάκωση και χωρίς να υπολογιστεί η κλίμακα του κάθε σημείου. Σημαντικό στοιχείο της προσέγγισης είναι ο μετασχηματισμός της εικόνας σε πεδίο στο οποίο η κλιμάκωση της εικόνας μετασχηματίζεται σε μετατόπιση.

Μια πρόσφατη και ενδιαφέρουσα εξέλιξη του περιγραφέα *SURF* (Ενότητα 4.3), παρουσιάζεται στην εργασία [7]. Οι περιγραφείς *Gauge-SURF*, στηρίζονται σε περιγραφή των σημείων (και παραγωγή της φωτεινότητας της εικόνας) σε *ορθογώνιο σύστημα συντεταγμένων* του επιπέδου που ορίζεται από την *τοπική διεύθυνση της κλίσης* και το κάθετο σε αυτή διάνυσμα. Οι περιγραφείς *Gauge-SURF* είναι όμοιοι με τους *SURF*, αντικαθιστώντας τις συνηθισμένες παραγώγους (στις διευθύνσεις x, y) με τις *Gauge-παραγώγους*. Τέλος



Σχήμα 4.9: Σύγκριση Γεωμετρικής και ομοιόμορφης εξομάλυνσης. Αριστερά: Αρχική Εικόνα. Κέντρο: Γεωμετρική εξομάλυνση. Δεξιά: Εξομάλυνση με Γκαουσιανό πυρήνα. Από [21].

η αξιολόγηση των περιγραφών Gauge-SURF γίνεται και σε σύστημα SfM (Ενότητα 2.1).

Στο [242], οι συγγραφείς χρησιμοποιούν την *μεθοδολογία σχεδιασμού* που περιγράψαμε στην Ενότητα 4.4. Χρησιμοποιούν το σχήμα *χωρικού συμψηφισμού των περιγραφών Daisy* (Ενότητα 4.5) και *εξετάζουν* την χρήση χαρακτηριστικών από τον υπολογισμό της *κλίσης* και *κατευθυνόμενων φίλτρων* (Μπλοκ Μετασχηματισμού T). Επίσης εξετάζονται οι επιπτώσεις της *κανονικοποίησης*. Έμφαση δίνεται στην *συμπύεση της περιγραφής*, η οποία επιτυγχάνεται με τον υποβιβασμό της διάστασης με PCA και την μείωση του δυναμικού εύρους των χαρακτηριστικών. Συγκριτικά με τους περιγραφείς SIFT (Ενότητα 4.1), οι συγγραφείς αναφέρουν υποδιπλασιασμό του ρυθμού σφάλματος με μέγεθος περιγραφέα το $1/10$ του SIFT.

Στο μεγαλύτερο τμήμα του Κεφαλαίου, ασχοληθήκαμε με περιγραφείς που στηρίζονται σε ιστογράμματα της διεύθυνσης της κλίσης σε μια περιοχή γύρω από το σημείο ενδιαφέροντος. Στο [23], οι συγγραφείς προτείνουν την χρήση *4 νέων χαρακτηριστικών*, σε συνδυασμό με τους συνήθεις περιγραφείς που στηρίζονται σε ιστογράμματα διευθύνσεων της κλίσης. Τα χαρακτηριστικά αυτά είναι εμπνευσμένα από τις αρχές της *Θεωρίας Gestalt*.

Μια σημαντική *σχεδιαστική επιλογή* που αφορά τους περιγραφείς, είναι η *ρύθμιση του ισοζυγίου αναλλοίωτου-διακριτικής ικανότητας*. Όσο αυξάνει η ικανότητα ενός περιγραφέα να παραμένει αναλλοίωτος σε διάφορους μετασχηματισμούς, μειώνεται η διακριτική του ικανότητα και αντιστρόφως. Στο [238], οι συγγραφείς υποστηρίζουν ότι ανάλογα με το σύνολο δεδομένων που διαθέτουμε, θα πρέπει να προσαρμόσουμε και το συγκεκριμένο ισοζύγιο. Δηλαδή δεν αναζητείται ο καθολικός βέλτιστος περιγραφέας, αλλά ο βέλτιστος περιγραφέας δεδομένου του προβλήματος που θέλουμε να τον χρησιμοποιήσουμε. Η βέλτιστη συνάρτηση πυρήνα για το πρόβλημα βρίσκεται σαν ένας συνδυασμός πυρήνων-βάσης, με την επίλυση ενός προβλήματος κυρτής βελτιστοποίησης με μοναδικό ολικό ακρότατο.

4.6.2 Διαθέσιμο Λογισμικό

Σε αυτή την Ενότητα αναφέρουμε διαθέσιμες υλοποιήσεις ανιχνευτών χαρακτηριστικών και μεθόδων αξιολόγησής τους. Επίσης αναφέρουμε σύνολα δεδομένων για την αξιολόγηση των ανιχνευτών χαρακτηριστικών.

Η αρχική υλοποίηση του ανιχνευτή χαρακτηριστικών SIFT (Ενότητα 4.1) είναι διαθέσιμη από τον διαδικτυακό τόπο [132]. Η υλοποίηση διατίθεται σε μορφή εκτελέσιμου, χωρίς τον κώδικα. Μια Διαθέσιμη υλοποίηση ανοιχτού κώδικα παρέχεται στην βιβλιοθήκη VLFeat (πρώην SIFT++). Η βιβλιοθήκη VLFeat είναι διαθέσιμη από την ιστοσελίδα [239]. Επίσης μια υλοποίηση ανοιχτού κώδικα σε GPU, η SIFTGPU είναι διαθέσιμη από την ιστοσελίδα [246].

Η αρχική υλοποίηση του ανιχνευτή χαρακτηριστικών SURF (Ενότητα 4.3) είναι διαθέσιμη από την ιστοσελίδα [41]. Μια διαδεδομένη υλοποίηση ανοιχτού κώδικα είναι η open SURF, διαθέσιμη από την ιστοσελίδα [52]. Επίσης υλοποίηση των χαρακτηριστικών SURF συμπεριλαμβάνεται στην βιβλιοθήκη ανοιχτού κώδικα Dlib [39].

Ο κώδικας για την υλοποίηση του συστήματος ανίχνευσης χαρακτηριστικών DAISY (Ενότητα 4.5), είναι διαθέσιμος από την ιστοσελίδα [226].

Για την αξιολόγηση των περιγραφών και των ανιχνευτών περιοχών, η μέθοδος αξιολόγησης που παρουσιάσαμε στην Ενότητα 4.2 είναι διαθέσιμη στην

ιστοσελίδα [13]. Τέλος κώδικας για την εφαρμογή εξομάλυνσης για την ευθυγράμμιση εικόνων [152], που περιγράψαμε προηγουμένως, είναι διαθέσιμος στην ιστοσελίδα [151].

Μέρος ΙΙ
Πειραματικό Μέρος

ΜΕΤΡΙΚΗ ΑΝΑΚΑΤΑΣΚΕΥΗ ΚΑΙ ΒΑΘΜΟΝΟΜΗΣΗ ΚΑΜΕΡΑΣ ΜΕ ΜΕΘΟΔΟΥΣ 2 ΟΨΕΩΝ

5.1 Περιγραφή

Τα περιεχόμενα των επόμενων Ενοτήτων, είναι αποτέλεσμα ανεξάρτητης εργασίας. Στις περιπτώσεις που δεν ισχύει αυτό, όπως για παράδειγμα στην Ενότητα 5.1.1, αναφέρονται οι σχετικές πηγές. Όπως έγινε αντιληπτό, αντίστοιχες προσεγγίσεις μπορούν να βρεθούν στην βιβλιογραφία [176, 195], αν και η παρουσίαση που γίνεται είναι σύντομη. Αναφέρουμε ακόμα μερικές διαφορές σε του [176] σε σχέση με την προσέγγιση μας, στις οποίες θα χρησιμοποιήσουμε και κάποιο από τον συμβολισμό που θα εισάγουμε αργότερα. Έχουμε λοιπόν για το [176]:

- Ο περιορισμός για το rank της περιγραφής της δυϊκής απόλυτης κωνικής Q_∞^* δεν επιβάλλεται. Οι συγγραφείς χρησιμοποιούν SVD αποσύνθεση για την επιβολή του σωστού rank (3).
- Σε περίπτωση μη επιβολής του rank, λαμβάνεται μια παραμετρική οικογένεια λύσεων. Σε περίπτωση επιβολής του, 4 λύσεις (περίπτωση 2 όψεων).

5.1.1 Εισαγωγικοί ορισμοί

Πριν περάσουμε στην περιγραφή του προβλήματος, αναφέρουμε σύντομα κάποιες βασικές έννοιες. Πληρέστερη ανάλυση και περιγραφή των εννοιών αυτών μπορεί να αναζητηθεί στο [90].

Θεωρούμε *μετρικό* σύστημα αναφοράς, δηλαδή σύστημα συντεταγμένων στο οποίο το επίπεδο στο άπειρο π_∞ βρίσκεται στην κανονική του θέση $(0 \ 0 \ 0 \ 1)^T$

Ορισμός 1. *Ορίζουμε την απόλυτη κωνική Ω_∞ , ως την κωνική στο επίπεδο στο άπειρο, που περιγράφεται από τις εξισώσεις:*

$$\begin{aligned} X_1^2 + X_2^2 + X_3^2 &= 0 \\ X_4 &= 0 \end{aligned}$$

Η Ω_∞ είναι μια γεωμετρική παράσταση των 5 βαθμών ελευθερίας που απαιτούνται για τον προσδιορισμό των μετρικών ιδιοτήτων του χώρου δοσμένης μιας αφινικής ανακατασκευής. Υπενθυμίζουμε ότι η μετάβαση από μια προβολική ανακατασκευή σε αφινική συνίσταται ουσιαστικά στον προσδιορισμό του π_∞ .

Παρατηρούμε ακόμα, ότι κανένα έγκυρο σημείο του προβολικού χώρου P^3 δεν ανήκει στην Ω_∞ . Μόνο σημεία με αμιγώς φανταστικές συντεταγμένες ικανοποιούν τις εξισώσεις ορισμού της.

Η Ω_∞ είναι μια κωνική η οποία ορίζεται ως γεωμετρικός τόπος σημείων. Τα επίπεδα του χώρου που είναι εφαπτόμενα στην Ω_∞ , ορίζουν την λεγόμενη δυϊκή επιφάνεια της Ω_∞ , την οποία θα συμβολίζουμε Q_∞^* .

Η Q_∞^* είναι μια τετραγωνική επιφάνεια του χώρου και έχει βολική αλγεβρική παράσταση. Συγκεκριμένα περιγράφουμε την Q_∞^* με την 4×4 μήτρα

$$Q_\infty^* = \begin{bmatrix} \mathbf{I}_{3 \times 3} & \mathbf{0}_3 \\ \mathbf{0}_3^T & 0 \end{bmatrix} \quad (5.1)$$

Τονίζουμε και πάλι ότι η εξίσωση αυτή ισχύει σε μετρικό σύστημα αναφοράς. Ένα επίπεδο π ανήκει στην Q_∞^* αν και μόνο αν ικανοποιεί την εξίσωση

$$\pi^T Q_\infty^* \pi = 0$$

Μπορεί να δειχθεί ότι τα επίπεδα που ανήκουν στην Q_∞^* είναι ακριβώς αυτά που είναι εφαπτόμενα στην Ω_∞ . Συνεπώς η Q_∞^* είναι πράγματι δυϊκή της Ω_∞^* .

Στην συνέχεια θα μας απασχολήσουν δύο θέματα:

1. Πως μετασχηματίζονται οι οντότητες που μόλις ορίσαμε, όταν ο χώρος μετασχηματίζεται προβολικά.
2. Πως προβάλλονται (εικονίζονται) οι οντότητες που ορίσαμε, στο επίπεδο της εικόνας.

Για την διερεύνηση των θεμάτων αυτών, παραθέτουμε τα επόμενα αποτελέσματα, ορισμένα χωρίς απόδειξη

Αποτέλεσμα 1. Όταν δύο επίπεδα (χώρος P^2) συνδέονται με ομογραφία H δηλαδή

$$\mathbf{x}' = H\mathbf{x}$$

ισχύει η ακόλουθη σχέση μεταξύ των κωνικών των επιπέδων

$$C' = H^{-T} C H$$

Αποτέλεσμα 2. Σε μετρικό σύστημα αναφοράς, το επίπεδο της εικόνας και το επίπεδο στο άπειρο συνδέονται με την ομογραφία

$$H_\infty = K R$$

Απόδειξη. Σε μετρικό σύστημα αναφοράς, τα σημεία του επιπέδου στο άπειρο έχουν την παράσταση

$$\mathbf{X}^T = (X_1 \ X_2 \ X_3 \ 0) = (\mathbf{x}_\infty^T \ 0)$$

Συμβολίζοντας τον πίνακα προβολής της κάμερας ως

$$P = [K R \ \mathbf{a}]$$

Έχουμε $x = P\mathbf{X} = KR\mathbf{x}_\infty$

□

Χρησιμοποιώντας τις προηγούμενες σχέσεις, μπορούμε τώρα να προσδιορίσουμε την εξίσωση της εικόνας της Ω_∞ , η οποία θυμίζουμε ότι είναι κωνική στο επίπεδο στο άπειρο.

Αποτέλεσμα 3. Η εικόνα της Ω_∞ δίνεται από την σχέση

$$\omega_\infty = K^{-T} K^{-1}$$

Απόδειξη. Από τον Ορισμό 1 της Ω_∞ προκύπτει άμεσα ότι είναι η κωνική $C = I$ στο π_∞ . Με χρήση αυτής της μορφής της C στα Αποτελέσματα 1 και 2 προκύπτει το ζητούμενο. \square

Για να βρούμε την εικόνα της Q_∞^* , παραθέτουμε πρώτα την επόμενη σχέση

Αποτέλεσμα 4. Η προβολή της δυϊκής τετραγωνικής επιφάνειας Q^* από τον πίνακα προβολής P στο επίπεδο της εικόνας είναι η δυϊκή κωνική

$$C^* = P Q^* P^T$$

Με απλή εφαρμογή του προηγούμενου θεωρήματος στην εξίσωση της Q_∞^* (Σχέση 5.1) έχουμε το επόμενο αποτέλεσμα

Αποτέλεσμα 5. Η Q_∞^* προβάλλεται στο επίπεδο της εικόνας στην δυϊκή κωνική

$$\omega_\infty^* = K K^T$$

Στην συνέχεια, προσδιορίζουμε την επίδραση που έχει ο μετασχηματισμός του χώρου με μια τρισδιάστατη 4×4 ομογραφία H στις δυϊκές τετραγωνικές επιφάνειες

Αποτέλεσμα 6. Έστω μετασχηματισμός του χώρου H :

$$\mathbf{X}' = H\mathbf{X}$$

Τότε, οι δυϊκές τετραγωνικές επιφάνειες μετασχηματίζονται σύμφωνα με την σχέση

$$Q^{*'} = H Q^* H^T$$

Επίσης, για ισχύει το ακόλουθο αποτέλεσμα για τις μήτρες προβολής P

Αποτέλεσμα 7. Ο μετασχηματισμός του χώρου με ομογραφία H , δηλαδή $\mathbf{X}' = H\mathbf{X}$, έχει ως αποτέλεσμα τον μετασχηματισμό των μητρών προβολής των καμερών σύμφωνα με την σχέση

$$P' = P H^{-1}$$

Ακόμα, το ακόλουθο αποτέλεσμα αποδεικνύεται εύκολα

Αποτέλεσμα 8. Έστω μετασχηματισμός του χώρου όμοια με το Αποτέλεσμα 7. Τότε, τα επίπεδα μετασχηματίζονται σύμφωνα με την σχέση

$$\pi' = H^{-T} \pi$$

Τα προηγούμενα αποτελέσματα μας επιτρέπουν να διατυπώσουμε τις βασικές εξισώσεις για την βαθμονόμηση των καμερών και τον προσδιορισμό του επιπέδου στο άπειρο π_∞ της προβολικής ανακατασκευής.

5.1.2 Εξισώσεις βαθμονόμησης

Δομή προβλήματος

Στην ενότητα αυτή θα περιγράψουμε την δομή του προβλήματος που θα αντιμετωπίσουμε.

Στο πρόβλημα θα θεωρήσουμε ότι έχουμε 2 κάμερες. Εφόσον δεν μας αφορά η ένταξη των δύο αυτών όψεων σε κάποιο άλλο (χωρικό) σύστημα αναφοράς, αλλά μόνο η σχετική θέση των δύο καμερών, τοποθετούμε το σύστημα συντεταγμένων του χώρου στην θέση της πρώτης κάμερας και το προσανατολίζουμε σύμφωνα με τον προσανατολισμό της πρώτης κάμερας.

Κατά συνέπεια, οι πίνακες προβολής των δύο καμερών θα δίνονται από τις σχέσεις

$$\begin{aligned} P_1 &= [K_1 \quad \mathbf{0}] \\ P_2 &= [K_2 R \quad \mathbf{t}] \end{aligned}$$

Επίσης θεωρούμε ότι οι πίνακες βαθμονόμησης K_i των καμερών την ιδιαίτερα απλή μορφή

$$K = \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Παρατηρούμε ότι το εστιακό μήκος f είναι η μοναδική παράμετρος του πίνακα K . Η επιλογή αυτής της απλοποιημένης μορφής του πίνακα K είναι δικαιολογημένη επαρκώς για την περίπτωση των σύγχρονων ψηφιακών φωτογραφικών μηχανών και εμφανίζεται συχνά στην βιβλιογραφία.

Διατύπωση εξισώσεων

Εστω ότι έχουμε στην διάθεση μας μια προβολική ανακατασκευή των καμερών, δηλαδή πίνακες καμερών P_{P1}, P_{P2} αντίστοιχα οι οποίοι διαφέρουν από τους πραγματικούς κατά ένα προβολικό μετασχηματισμό του τρισδιάστατου χώρου, δηλαδή ισχύει (βλ.Αποτέλεσμα 7)

$$\begin{aligned} P_{M1} &= P_{P1}H \\ P_{M2} &= P_{P2}H \end{aligned} \tag{5.2}$$

Στις προηγούμενες σχέσεις συμβολίζουμε με P_{Mi} την μήτρα προβολής της κάμερας i στο μετρικό πλαίσιο αναφοράς και με P_{Pi} την μήτρα προβολής της κάμερας i στο προβολικό πλαίσιο αναφοράς.

Χρησιμοποιώντας το Αποτέλεσμα 4 την Σχέση 5.1 και την Σχέση 5.2, έχουμε την επόμενη σχέση για τη εικόνα της Q_∞^* στην εικόνα 2

$$\begin{bmatrix} f_2^2 & 0 & 0 \\ 0 & f_2^2 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \omega_2^* = P_{P2}H Q_\infty^* H^T P_{P2}^T \tag{5.3}$$

Για να μπορέσουμε να κάνουμε χρήσιμες τις Σχέσεις 5.3, θα προσδιορίσουμε πρώτα την μορφή της ομογραφίας H .

Θεωρούμε ότι η προβολική ανακατασκευή είναι στην λεγόμενη κανονική μορφή, δηλαδή η μήτρα προβολής της κάμερας 1 έχει την μορφή

$$P_{P1} = [I \quad \mathbf{0}]$$

Από τις Σχέσεις 5.2 έχουμε :

$$\begin{aligned} P_{M1} &= [K_1 \quad \mathbf{0}] \\ P_{P1} &= [I \quad \mathbf{0}] \\ \Leftrightarrow H &= \begin{bmatrix} K_1 & \mathbf{0} \\ \mathbf{v}^T & \sigma \end{bmatrix} \end{aligned} \quad (5.4)$$

όπου το διάνυσμα \mathbf{v}^T και η σταθερά σ δεν έχουν ακόμα προσδιοριστεί.

Αν αγνοήσουμε την κλίμακα της ανακατασκευής, μπορούμε να θέσουμε την σταθερά $\sigma = 1$.

Οι απαραίτητες συνθήκες για τον πλήρη προσδιορισμό του H μπορούν να ληφθούν από την θέση του επιπέδου στο άπειρο π_∞ . Από την Σχέση 5.2 και το Αποτέλεσμα 7 έχουμε

$$\mathbf{X}_P = H\mathbf{X}_M \quad (5.5)$$

Χρησιμοποιώντας τώρα την προηγούμενη Σχέση 5.5 και το Αποτέλεσμα 8 έχουμε

$$\pi_{\infty, P} = H^{-T} \begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix} = \begin{pmatrix} -K_1^{-T}\mathbf{v} \\ 1 \end{pmatrix}$$

Στην συνέχεια θα συμβολίζουμε, για ευκολία

$$\pi_{\infty, P} = \begin{pmatrix} \mathbf{p} \\ 1 \end{pmatrix}$$

Συνεπώς, καταλήξαμε στην ακόλουθη μορφή για την ομογραφία H

$$\begin{bmatrix} K_1 & \mathbf{0} \\ -\mathbf{p}^T K_1 & 1 \end{bmatrix} \quad (5.6)$$

Συμπερασματικά, η Σχέση 5.6 μας δίνει μια περιγραφή της ομογραφίας που συνδέει την προβολική με την μετρική ανακατασκευή μέσω 4 παραμέτρων, συγκεκριμένα την θέση του (σωστού) επιπέδου στο άπειρο $\pi_{\infty, P}$ και του εστιακού μήκους της κάμερας $1, f_1$. Οι σχέσεις αυτές δεν πρέπει να μας ξαφνιάζουν, αφού η μετρική ανακατασκευή συνεπάγεται και την ανάκτηση των αφνικών ιδιοτήτων του χώρου, γεγονός που προϋποθέτει τον ορθό προσδιορισμό του επιπέδου στο άπειρο.

Αντικαθιστούμε τώρα την περιγραφή της H (από την Σχέση 5.6) στις αρχικές εξισώσεις μας (5.3) και χρησιμοποιούμε την μορφή της Q_∞^* για να πάρουμε

$$\omega_2^* = P_{P2} \begin{bmatrix} K_1 K_1^T & -K_1 K_1^T \mathbf{p} \\ -\mathbf{p}^T K_1 K_1^T & \mathbf{p}^T K_1 K_1^T \mathbf{p} \end{bmatrix} P_{P2}^T \quad (5.7)$$

Από την επίλυση των (μη-γραμμικών ως προς τους 5 αγνώστους) Εξισώσεων 5.7, προσδιορίζουμε την μετρική ανακατασκευή της γεωμετρίας του χώρου.

Κατ'αρχάς, παρατηρούμε ότι (από τον ορισμό της) η ω_∞^* έχει συμμετρική περιγραφή, και ότι η κλιμάκωση της περιγραφής κατά μία σταθερά δεν επηρεάζει τις ιδιότητες της (είναι ομογενής ποσότητα). Συνεπώς από τα 9 στοιχεία της μήτρας που περιγράφει την ω_2^* , μπορούμε να έχουμε έως 5 ανεξάρτητες εξισώσεις, όσοι δηλαδή και οι άγνωστοι μας.

Λόγω της συμμετρίας του Συστήματος 5.3, μπορούμε να σχηματίσουμε 6

διαφορετικές εξισώσεις. Αυτές είναι:

$$\omega_2^*(1, 1) = f_2^2 = f_1^2 p_1^2 p_{14}^2 + f_1^2 p_{14}^2 p_2^2 - 2 f_1^2 p_1 p_{11} p_{14} - 2 f_1^2 p_{12} p_{14} p_2 + f_1^2 p_{11}^2 + f_1^2 p_{12}^2 + p_{14}^2 p_3^2 - 2 p_{13} p_{14} p_3 + p_{13}^2$$

$$\omega_2^*(1, 2) = 0 = p_{14} p_{24} p_3^2 - (p_{13} p_{24} + p_{14} p_{23}) p_3 - (f_1^2 p_{14} p_2 - f_1^2 p_{12}) p_{22} - (f_1^2 p_1 p_{14} - f_1^2 p_{11}) p_{21} + (f_1^2 p_1^2 p_{14} + f_1^2 p_{14} p_2^2 - f_1^2 p_1 p_{11} - f_1^2 p_{12} p_2) p_{24} + p_{13} p_{23}$$

$$\omega_2^*(1, 3) = 0 = -(p_{14} p_3 - p_{13}) p_{33} - (f_1^2 p_{14} p_2 - f_1^2 p_{12}) p_{32} - (f_1^2 p_1 p_{14} - f_1^2 p_{11}) p_{31} + (f_1^2 p_1^2 p_{14} + f_1^2 p_{14} p_2^2 - f_1^2 p_1 p_{11} - f_1^2 p_{12} p_2 + p_{14} p_3^2 - p_{13} p_3) p_{34}$$

$$\omega_2^*(2, 2) = f_2^2 = f_1^2 p_{21}^2 + f_1^2 p_{22}^2 + p_{24}^2 p_3^2 + (f_1^2 p_1^2 + f_1^2 p_2^2) p_{24}^2 - 2 p_{23} p_{24} p_3 - 2 (f_1^2 p_1 p_{21} + f_1^2 p_2 p_{22}) p_{24} + p_{23}^2$$

$$\omega_2^*(2, 3) = 0 = -(p_{24} p_3 - p_{23}) p_{33} - (f_1^2 p_2 p_{24} - f_1^2 p_{22}) p_{32} - (f_1^2 p_1 p_{24} - f_1^2 p_{21}) p_{31} - (f_1^2 p_1 p_{21} + f_1^2 p_2 p_{22} - p_{24} p_3^2 - (f_1^2 p_1^2 + f_1^2 p_2^2) p_{24} + p_{23} p_3) p_{34}$$

$$\omega_2^*(3, 3) = 1 = f_1^2 p_{31}^2 + f_1^2 p_{32}^2 + (f_1^2 p_1^2 + f_1^2 p_2^2 + p_3^2) p_{34}^2 - 2 (f_1^2 p_1 p_{31} + f_1^2 p_2 p_{32} + p_3 p_{33}) p_{34} + p_{33}^2$$

Στις προηγούμενες εξισώσεις, αντικαταστήσαμε τον πίνακα P_{P_2} με

$$\begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix}$$

Παρατηρώντας τις εξισώσεις αυτές (καλύτερα από την γραφή τους στην μορφή 5.7), διαπιστώνουμε ότι μπορούμε να ομαδοποιήσουμε τους αγνώστους στα εξής συμπλέγματα

$$\mathbf{x}_0 = \begin{pmatrix} f_1^2 \\ f_2^2 \\ f_1^2 p_1^2 + f_1^2 p_2^2 + p_3^2 \\ p_3 \\ f_1^2 p_1 \\ f_1^2 p_2 \end{pmatrix} \quad (5.8)$$

Αναδιατάσσοντας το σύστημα των εξισώσεων (και εκτελώντας και μια ασήμαντη αναδιάταξη της σειράς τους) παίρνουμε το γραμμικό σύστημα ως προς το διάνυσμα των "αγνώστων" \mathbf{x}_0 :

$$\begin{pmatrix} p_{21}^2 + p_{22}^2 & -1 & p_{24}^2 & -2 p_{23} p_{24} & -2 p_{21} p_{24} & -2 p_{22} p_{24} & -p_{23}^2 \\ p_{21} p_{31} + p_{22} p_{32} & 0 & p_{24} p_{34} & -p_{23} p_{34} - p_{24} p_{33} & -p_{21} p_{34} - p_{24} p_{31} & -p_{22} p_{34} - p_{24} p_{32} & -p_{23} p_{33} \\ p_{31}^2 + p_{32}^2 & 0 & p_{34}^2 & -2 p_{33} p_{34} & -2 p_{31} p_{34} & -2 p_{32} p_{34} & -p_{33}^2 + 1 \\ p_{11} + p_{12}^2 & -1 & p_{14}^2 & -2 p_{13} p_{14} & -2 p_{11} p_{14} & -2 p_{12} p_{14} & -p_{13}^2 \\ p_{11} p_{21} + p_{12} p_{22} & 0 & p_{14} p_{24} & -p_{13} p_{24} - p_{14} p_{23} & -p_{11} p_{24} - p_{14} p_{21} & -p_{12} p_{24} - p_{14} p_{22} & -p_{13} p_{23} \\ p_{11} p_{31} + p_{12} p_{32} & 0 & p_{14} p_{34} & -p_{13} p_{34} - p_{14} p_{33} & -p_{11} p_{34} - p_{14} p_{31} & -p_{12} p_{34} - p_{14} p_{32} & -p_{13} p_{33} \end{pmatrix} \quad (5.9)$$

Όπου στην Εξίσωση 5.9 έχουμε δώσει τον επαυξημένο πίνακα

$$[A|\mathbf{b}]$$

του γραμμικού συστήματος

$$A\mathbf{x}_0 = \mathbf{b} \quad (5.10)$$

Υπενθυμίζουμε ότι από τις 6 εξισώσεις του Συστήματος 5.9, μόνο οι 5 μπορεί να είναι γραμμικώς ανεξάρτητες. Όπως ήδη αναφέραμε, οι βαθμοί ελευθερίας της μήτρας ω_2^* είναι 5 και συνεπώς οι εξισώσεις ήταν εξαρχής εξαρτημένες, γεγονός που φυσικά δεν μπορεί να μεταβληθεί με αναδιάταξη των όρων των εξισώσεων. Το γεγονός αυτό μπορεί να ελεγχθεί και με τον υπολογισμό της ορίζουσας του πίνακα A που δίδεται, που προκύπτει ταυτοτικά 0. Εφόσον όμως δεν ισχύει το ίδιο για όλες τις 6×6 υποορίζουσες του επαυξημένου πίνακα, το σύστημα μπορεί να είναι μη-επιλύσιμο. Θα επιστρέψουμε στο θέμα αυτό και στην επίδραση της κλιμάκωσης, σε επόμενη Ενότητα (Ενότητα 5.2.1).

Συνεπώς, παρουσιάσαμε ένα γραμμικό σύστημα 5 (στην ευνοϊκότερη περίπτωση) γραμμικώς ανεξάρτητων εξισώσεων σε 6 μη-γραμμικά συμπλέγματα των 5 αγνώστων μας. Παρότι ένα γραμμικό σύστημα 5×6 στην γενική περίπτωση δεν έχει μοναδική λύση, το γεγονός ότι έχουμε δευτεροβάθμια πολυώνυμα των 5 υπό προσδιορισμό παραμέτρων, δίνει τις επιπλέον απαραίτητες σχέσεις για να περιορίσουμε τον αριθμό των λύσεων, όπως θα δούμε σύντομα.

Θετικά στοιχεία της διατύπωσης του προβλήματος

Στο σημείο αυτό θα τονίσουμε κάποιες θετικές ιδιότητες της λύσης του προβλήματος της μετρικής αναβάθμισης της ανακατασκευής, μέσω των εξισώσεων που προτείνουμε.

Άλλες μέθοδοι που έχουν προταθεί κατά το παρελθόν, θέτουν πιο χαλαρούς περιορισμούς και δεν λαμβάνουν υπόψιν ιδιότητες, όπως για παράδειγμα ότι η Q_∞^* είναι εκφυλισμένη τετραγωνική επιφάνεια (ο πίνακας περιγραφής της έχει $\text{rank } 3$). Αν και η εξαγωγή των ιδιοτήτων αυτών είναι σχετικά τετριμμένη, πιστεύουμε ότι αξίζει να τις συγκεντρώσουμε. Στην συνέχεια αναλύουμε κάποιες από αυτές τις ιδιότητες.

Αρχικά, οι λύσεις του Συστήματος 5.9, οι οποίες ικανοποιούν επιπλέον και τις μη-γραμμικές σχέσεις μεταξύ των συντεταγμένων του \mathbf{x}_0 , είναι και λύσεις του αρχικού Συστήματος 5.7. Κάνοντας χρήση του Αποτελέσματος 6 παρατηρούμε ότι η εξίσωση

$$HQ_\infty^* H^T$$

δίνει την δυϊκή τετραγωνική επιφάνεια της απόλυτης κωνικής στο προβολικό σύστημα αναφοράς. Η σχέση αυτή, δίνεται σε πιο βολική μορφή μετά την αντικατάσταση του H στην Εξίσωση 5.7. Δια απλής επισκοπήσεως, διαπιστώνουμε ότι το $(\pi_{\infty, P}^T \ 1)^T$ είναι πράγματι το μηδενοδιάνυσμα της μήτρας αυτής, όπως απαιτείται.

Ακόμα, η μήτρα $HQ_\infty^* H^T$ έχει $\text{rank } 3$, όπως εύκολα διαπιστώνεται από τον μη εκφυλισμένο μηδενοχώρο (στον οποίο ανήκει το $\pi_{\infty, P}$) και την 3×3 μη μηδενική υποορίζουσα του $H(1 : 3, 1 : 3) = K_1 K_1^T$. Το ίδιο αποτέλεσμα μπορεί να ληφθεί και χρησιμοποιώντας το γνωστό από την γραμμική άλγεβρα λήμμα

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix} = |AD - BC| \text{ όταν } CD = DC$$

Επίσης, ένα κοινό πρόβλημα που παρουσιάζεται, αφορά τον περιορισμό της θετικά ημιορισμένης ω_i^* . Όταν η συνθήκη αυτή δεν ικανοποιείται, δεν είναι

δυνατόν να γίνει αποσύνθεση Cholesky της μήτρας $\omega_i^* = K_i K_i^T$ για να ανακτηθεί η μήτρα K . Με τον τρόπο που διατυπώσαμε το πρόβλημα, το ελάττωμα αυτό δεν εμφανίζεται. Παρόμοια, η διαγώνια μορφή των μητρών K_i επιβάλλεται.

Γραφή των εξισώσεων με γινόμενα μητρών

Στην ενότητα αυτή θα γράψουμε τις εξισώσεις του Συστήματος 5.9 ως γινόμενα μητρών, γεγονός που βοηθάει στην καλύτερη εποπτεία των εξισώσεων αυτών και εξυπηρετεί σε κάποια σημεία τον αλγεβρικό χειρισμό των εξισώσεων. Από την γραφή αυτή παραλείπουμε την στήλη του πίνακα που αντιστοιχεί στην μεταβλητή f_2 , γιατί είναι σταθερή (ανεξάρτητη της μήτρας P) και δυσχεραίνει την διατύπωση.

Αρχικά παρατηρούμε την εξής σχέση, μεταξύ των στοιχείων της ω_2^* και των στοιχείων της μήτρας P_{P_2}

$$\omega_2^*(i, j) = P_i^T H Q_\infty^* H^T P_j \quad (5.11)$$

Στην προηγούμενη Σχέση (5.11), χρησιμοποιήσαμε τον συμβολισμό

P_i : Το διάνυσμα στήλη που προκύπτει από την i γραμμή του πίνακα P_{P_2}

Κάνοντας μερικές ακόμα παρατηρήσεις, θα γράψουμε ως γινόμενο μητρών τον πίνακα

$$A_r = A(1 : 5, \{1, 3, 4, 5, 6\}) \quad \text{οι πρώτες 5 γραμμές του πίνακα } A, \\ \text{χωρίς την σταθερή δεύτερη στήλη}$$

όπου ο πίνακας A δίνεται στην Σχέση 5.10.

Από την μορφή του πίνακα $H Q_\infty^* H^T$, όπως αυτή είναι φανερά στην Εξίσωση 5.7, καταλήγουμε στην ακόλουθη γραφή

$$A_r = \text{diag}(P_2^T, P_2^T, P_3^T, P_1^T, P_1^T) \cdot \text{diag}(Q^{1-5}, Q^{1-5}, Q^{1-5}, Q^{1-5}, Q^{1-5}) \cdot \begin{bmatrix} D_2 \\ D_3 \\ D_3 \\ D_1 \\ D_2 \end{bmatrix} = C Q D \quad (5.12)$$

όπου

$$Q^{1-5} = [Q_1 \quad Q_2 \quad Q_3 \quad Q_4 \quad Q_5]$$

$$Q_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$Q_2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$Q_3 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \end{bmatrix}$$

$$Q_4 = \begin{bmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{bmatrix}$$

$$Q_5 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}$$

$$D_i = \text{diag}(P_i, P_i, P_i, P_i, P_i)$$

Η μορφή αυτή των εξισώσεων, μας παρέχει ένα συστηματικό τρόπο να εξάγουμε συνθήκες για το rank του πίνακα A_r .

Συγκεκριμένα, λαμβάνοντας υπόψιν ότι

- Ο πίνακας P έχει πάντα $\text{rank} 3$
- Την γνωστή από την γραμμική άλγεβρα ανισότητα

$$\text{rank}(AB) \leq \max(\text{rank}(A), \text{rank}(B))$$

- την μορφή των Q_i και των C, D

μπορούμε να εξάγουμε συμπεράσματα για τα $\text{rank}(CQ)$, $\text{rank}(QD)$ και κατά συνέπεια για το $\text{rank}(CQD) = \text{rank}(A_r)$.

Αρχικά, εξετάζουμε το γινόμενο AC . Η μήτρα που προκύπτει από αυτόν τον πολλαπλασιασμό, έχει για στήλες 5-διανύσματα που προκύπτουν ως γραμμικοί συνδυασμοί (με σταθμά τα στοιχεία των στηλών του πίνακα Q) των στηλών της μήτρας A .

Δύο βολικές ιδιότητες των μητρών C, Q , συγκεκριμένα

- Οι στήλες του C είναι πολλαπλάσια των διανυσμάτων e_i της ορθοκανονικής βάσης I_5
- Κανένα Q_i δεν έχει δύο μη-μηδενικά στοιχεία σε κάποια στήλη του

μας επιτρέπουν την εξαγωγή απλών (ως παράσταση των στοιχείων της μήτρας P) ικανών και αναγκαίων συνθηκών που εξασφαλίζουν ότι το γινόμενο

CQ έχει $\text{rank} 5$. Οι συνθήκες όμως αυτές ικανοποιούνται ταυτοτικά, από την συνθήκη $\text{rank}(P) = 3$, και συνεπώς δεν προσφέρουν επιπλέον πληροφορία για το πρόβλημα.

Αντίστοιχα (με ένα μη αυστηρό τρόπο δυϊκά), μπορούμε να εξετάσουμε το γινόμενο QD σαν γραμμικό συνδυασμό των γραμμών του D με σταθμά τις εγγραφές των γραμμών του Q . Ουσιαστικά η ανάλυση είναι ίδια με πριν, εφόσον νοητικά κάνουμε την αντιστοιχία γραμμές \leftrightarrow στήλες. Καταλήγουμε έτσι στο επόμενο Αποτέλεσμα

Αποτέλεσμα 9. Όταν η σχετική δέση των καμερών 1,2 είναι καθαρή περιστροφή (μηδενική μετατόπιση), ο πίνακας A_r χάνει rank

Το Αποτέλεσμα 9, είναι αναμενόμενο, αφού αφινικές ιδιότητες του χώρου (δηλαδή το επίπεδο στο άπειρο), δεν μπορούν να ανακτηθούν στην περίπτωση της καθαρά περιστροφικής κίνησης. Όμως, το εστιακό βάθος μπορεί να ανακτηθεί.

Δυστυχώς, η ανάλυση δεν μπορεί να επεκταθεί στο γινόμενο $(CQ)D$ ή $C(QD)$, γιατί η ιδιαίτερη δομή των παραγόντων του γινομένου χάνεται και δεν δύναται να εξαχθούν απλές και άμεσες συνθήκες που συνδέουν το rank του γινομένου με τα στοιχεία της μήτρας P .

Συνεπώς, τονίζουμε ότι η ανάλυση μας δεν ήταν εξαντλητική, όσον αφορά το γινόμενο $A_r = CQD$. Ενδεχομένως, υφίστανται και άλλες εκφυλισμένες διατάξεις.

5.1.3 Επίλυση των εξισώσεων και σχέσεις των λύσεων

Λύνουμε το Σύστημα 5.10, από το οποίο έχουμε κρατήσει 5 μόνο από τις εξισώσεις του. Θα το συμβολίζουμε

$$A_5 \mathbf{x}_0 = \mathbf{b}_5 \quad (5.13)$$

Με την προϋπόθεση ότι το Σύστημα 5.10 δεν είναι αδύνατο, δεν έχει σημασία ποιές εξισώσεις θα κρατήσουμε, αφού όπως προαναφέραμε ο πίνακας A έχει, στην γενική περίπτωση, $\text{rank} 5$. Στα επόμενα υποθέτουμε ότι $\text{rank}(A) = 5$ και ότι το Σύστημα 5.10 είναι επιλύσιμο.

Εφαρμόζουμε απαλοιφή Gauss στο Σύστημα 5.13 και το φέρνουμε στην μορφή

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & b_1 \\ 0 & 1 & 0 & 0 & 0 & 0 & b_2 \\ 0 & 0 & 1 & 0 & 0 & 0 & b_3 \\ 0 & 0 & 0 & 1 & 0 & c & b_4 \\ 0 & 0 & 0 & 0 & 1 & d & b_5 \end{bmatrix} \quad (5.14)$$

όπου στην Σχέση 5.14 δώσαμε τον επαυξημένο πίνακα που προκύπτει μετά την εφαρμογή απαλοιφής Gauss. Για την εξαγωγή της μορφής αυτής, θα κάνουμε μια σύντομη απόδειξη.

Καταρχάς, για τα στοιχεία της μήτρας του επαυξημένου Συστήματος 5.14 κάνουμε την ακόλουθη κατανομή σε 3 κατηγορίες:

1. Τα στοιχεία που γράψαμε με την συνήθη (μαύρη) γραμματοσειρά, προκύπτουν όπως δίνονται, για κάθε 5×7 μήτρα στην οποία έχουμε εφαρμόσει απαλοιφή Gauss. Είναι δηλαδή γενικό αποτέλεσμα, και δεν σχετίζεται με την ιδιαίτερη δομή του προβλήματος.

2. Τα στοιχεία που γράψαμε με πράσινη γραμματοσειρά, έχουν τις δοσμένες τιμές λόγω της ιδιαίτερης δομής των εξισώσεων που χρησιμοποιούμε. Θα αποδείξουμε αρχικά αυτό το αποτέλεσμα.
3. Τέλος, το στοιχείο που γράψαμε με την κυανή γραμματοσειρά, έχει την δοσμένη τιμή όταν ο πίνακας προβολής P είναι μέλος του κανονικού ζευγαριού καμερών του προβλήματος, δηλαδή

$$\{[I \ 0], [[\mathbf{a}]_x^F \ \mathbf{a}]\} \quad (5.15)$$

όπου F ο Θεμελιώδης πίνακας που αντιστοιχεί στο πρόβλημα και \mathbf{a} το αριστερό μηδενοδιάνυσμα του F , δηλαδή

$$F^T \mathbf{a} = \mathbf{0}$$

Σε αυτή την περίπτωση, το 3×3 αριστερό μπλοκ του πίνακα P έχει rank 2, επομένως τα διανύσματα γραμμές του είναι γραμμικώς εξαρτημένα. Η τελευταία αυτή ιδιότητα, είναι απαραίτητη για την απόδειξη του συγκεκριμένου αποτελέσματος.

Συμβολίζουμε επίσης με

$$\mathbf{P}_i^\gamma$$

το διάνυσμα-γραμμή που παίρνουμε από την i γραμμή του πίνακα $[\mathbf{a}]_x^F$. Για το Σύστημα 5.13, θα υποθέσουμε ότι έχουμε κρατήσει τις Εξισώσεις 1, 2, 6, 4, 5 του Συστήματος 5.10 (Συμβολίζουμε A_{pr}). Κρατήσαμε δηλαδή τις 5 πρώτες Εξισώσεις του συστήματος, αντικαθιστώντας την 3η εξίσωση με την 6η. Η κίνηση αυτή θα δικαιολογηθεί επαρκώς στις επόμενες Ενότητες της ανάλυσης. Επίσης, η περίπτωση αυτή επαρκεί για τις ανάγκες μας, γεγονός στο οποίο θα στηριχθούμε προκειμένου να μην επεκταθούμε περισσότερο. Επίσης, αποκλειστικά για λόγους απλοποίησης του συμβολισμού και χωρίς άλλη επίπτωση, αναδιατάσσουμε τους όρους του διανύσματος αγνώστων (Σχέση 5.8). Συγκεκριμένα γράφουμε το διάνυσμα των αγνώστων ως

$$\mathbf{x} = \begin{pmatrix} f_1^2 \\ f_2^2 \\ f_1^2 p_1^2 + f_1^2 p_2^2 + p_3^2 \\ f_1^2 p_1 \\ f_1^2 p_2 \\ p_3 \end{pmatrix}$$

και κατά συνέπεια αντιμεταθέτουμε τις στήλες

$$4 \leftrightarrow 6$$

της μήτρας του συστήματος (Εξίσωση 5.10).

Αρχικά, παρατηρούμε την δομή της μήτρας A_{pr} και με την εισαγωγή κατάλληλου, βολικού, συμβολισμού, την γράφουμε ως

$$\begin{bmatrix} \Gamma_1 & \phi_1^1 \mathbf{P}_2^\gamma & \psi_1 \\ \Gamma_2 & \phi_2^1 \mathbf{P}_2^\gamma + \phi_2^2 \mathbf{P}_3^\gamma & \psi_2 \\ \Gamma_3 & \phi_3^1 \mathbf{P}_1^\gamma + \phi_3^2 \mathbf{P}_3^\gamma & \psi_3 \\ \Gamma_4 & \phi_4^1 \mathbf{P}_1^\gamma & \psi_4 \\ \Gamma_5 & \phi_5^1 \mathbf{P}_1^\gamma + \phi_5^2 \mathbf{P}_2^\gamma & \psi_5 \end{bmatrix} \quad (5.16)$$

Στην προηγούμενη μήτρα, έχουμε:

1. \mathbf{r}_i : 1×3 διανύσματα. Δεν παρουσιάζουν κάποιες ειδικές σχέσεις μεταξύ τους και για αυτό δεν θα μας απασχολήσουν στην απόδειξη.
2. Κατάλληλες σταθερές ϕ_j^i . Στην πραγματικότητα είναι στοιχεία της τέταρτης στήλης του P , το γεγονός αυτό όμως δεν επηρεάζει την εξαγωγή των αποτελεσμάτων.
3. Τα στοιχεία της τελευταίας στήλης του συστήματος A_{pr} , τα συμβολίζουμε με τις σταθερές ψ_i , καθώς ξανά δεν παρουσιάζουν ιδιαίτερη δομή.

Το σκεπτικό της απόδειξης που δίνουμε, είναι να μηδενίσουμε τα στοιχεία του πίνακα A_{pr} που βρίσκονται στις γραμμές 1 – 3 και στήλες 4 – 6. Ασχολούμαστε δηλαδή με το τμήμα της μήτρας A_{pr} του οποίου την δομή καταστήσαμε σαφή (Σχέση 5.16), και συγκεκριμένα τις τρεις πρώτες γραμμές του τμήματος αυτού. Αν στην συνέχεια εφαρμόσουμε απαλοιφή Gauss στον πίνακα που προκύπτει, θα λάβουμε την επιθυμητή μορφή που δίνεται στην Σχέση 5.14.

Προχωράμε λοιπόν, στον μηδενισμό των δύο πρώτων γραμμών, ισοδύναμα δηλαδή στην Απόδειξη που αφορά τα πράσινα στοιχεία. Συμβολίζοντας με \mathbf{e}_i την i γραμμή του πίνακα A_{pr} , εκτελώντας γραμμοπράξεις καταλήγουμε στο ζητούμενο. Για παράδειγμα, μια πιθανή ακολουθία γραμμοπράξεων είναι:

$$\begin{aligned} \mathbf{e}_5 &\leftarrow \mathbf{e}_5 + \alpha_1 \mathbf{e}_4 \\ \mathbf{e}_1 &\leftarrow \mathbf{e}_1 + \alpha_2 \mathbf{e}_5 \\ \mathbf{e}_2 &\leftarrow \mathbf{e}_2 + \alpha_3 \mathbf{e}_5 \\ \mathbf{e}_3 &\leftarrow \mathbf{e}_3 + \alpha_4 \mathbf{e}_4 \\ \mathbf{e}_2 &\leftarrow \mathbf{e}_2 + \alpha_5 \mathbf{e}_3 \end{aligned}$$

η οποία οδηγεί στην ακόλουθη μορφή για τις δύο πρώτες γραμμές της μήτρας A_{pr}

$$\begin{bmatrix} \mathbf{r}'_1 & \mathbf{0} & \psi'_1 \\ \mathbf{r}'_2 & \mathbf{0} & \psi'_2 \end{bmatrix}$$

Τώρα με την επιπλέον υπόθεση για τον πίνακα P , μπορούμε να γράψουμε

$$\mathbf{P}_3^\gamma = \mathbf{d}_1 \mathbf{P}_1^\gamma + \mathbf{d}_2 \mathbf{P}_2^\gamma$$

και αντίστοιχα, με κατάλληλες γραμμοπράξεις να καταλήξουμε στην ακόλουθη μορφή για τις τρεις πρώτες γραμμές του πίνακα A_{pr} :

$$\begin{bmatrix} \mathbf{r}'_1 & \mathbf{0} & \psi'_1 \\ \mathbf{r}'_2 & \mathbf{0} & \psi'_2 \\ \mathbf{r}'_3 & \mathbf{0} & \psi'_3 \end{bmatrix}$$

Στην συνέχεια, εφαρμογή απαλοιφής Gauss στην ισοδύναμη μορφή της μήτρας A_{pr} στην οποία καταλήξαμε, δίνει το σύστημα της Σχέσης 5.14.

Για να γίνουν εμφανείς οι σχέσεις των μεταβλητών, ξαναγράφουμε το Σύστημα 5.14

$$f_1^2 = b_1 \quad (5.17)$$

$$f_2^2 = b_2 \quad (5.18)$$

$$f_1^2 p_1^2 + f_1^2 p_2^2 + p_3^2 = b_3 \quad (5.19)$$

$$p_3 + c f_1^2 p_2 = b_4 \quad (5.20)$$

$$f_1^2 p_1 + d f_1^2 p_1 = b_5 \quad (5.21)$$

Ένας τρόπος επίλυσης είναι, να λύσουμε τις Σχέσεις 5.21,5.20 ως προς p_1 και p_3 αντίστοιχα, και να αντικαταστήσουμε τις τιμές αυτές στην Σχέση 5.19. Με αυτόν τον τρόπο προκύπτει μια τετραγωνική εξίσωση στο p_2 με γνωστούς συντελεστές, αφού το f_1^2 μπορεί να αντικατασταθεί από την Σχέση 5.17. Λαμβάνουμε έτσι δύο λύσεις

$$\begin{aligned} \mathbf{x}_0^1 &= (b_1 \quad b_2 \quad b_3 \quad p_3 \quad f_1^2 p_1 \quad f_1^2 p_2)^T \\ \mathbf{x}_0^2 &= (b_1 \quad b_2 \quad b_3 \quad p_3' \quad f_1^2 p_1' \quad f_1^2 p_2')^T \end{aligned} \quad (5.22)$$

Σημειώνουμε ότι τα f_1, f_2 προσδιορίζονται μοναδικά, αφού ορίζονται ως θετικές ποσότητες.

Γεωμετρική ερμηνεία των λύσεων

Για να ερμηνεύσουμε γεωμετρικά τις λύσεις, παρατηρούμε αρχικά από τις Σχέσεις 5.22,5.20 ότι

$$f_1^2 p_1^2 + f_2^2 p_2^2 + p_3^2 = f_1^2 p_1'^2 + f_2^2 p_2'^2 + p_3'^2 = b_3 \quad (5.23)$$

Ερμηνεύουμε τις Εξισώσεις 5.19,5.20,5.21,5.23 ως περιορισμούς στον τρισδιάστατο χώρο, από τους οποίους προκύπτει η λύση των εξισώσεων μας. Έτσι έχουμε

- Αν θεωρήσουμε το τρισδιάστατο διάνυσμα αγνώστων $(f_1 p_1 \quad f_1 p_2 \quad p_3)^T = \vec{v}$ η Εξίσωση 5.23 γράφεται

$$\|\vec{v}\| = r^2$$

και παριστά σφαίρα με κέντρο την αρχή των αξόνων.

- Αν θεωρήσουμε τρισδιάστατο διάνυσμα αγνώστων $(p_1 \quad p_2 \quad p_3)^T \equiv \vec{m}$, η Εξίσωση 5.23 παριστά ελλειψοειδές με δύο ίσους άξονες, αφού οι συντελεστές των δευτεροβάθμιων όρων είναι θετικοί και ίσοι για τα p_1, p_2 .

Η θεώρηση του διανύσματος \vec{v} δίνει απλούστερη επιφάνεια, το διάνυσμα \vec{m} όμως είναι στενότερα συνδεδεμένο με την γεωμετρία του προβλήματος, γιατί $\pi_{\infty, P} = (\mathbf{p}^T \quad 1)^T$, δηλαδή το διάνυσμα \vec{m} έχει την ίδια διεύθυνση με το κάθετο διάνυσμα του π_{∞} στο προβολικό σύστημα συντεταγμένων.

Οι Εξισώσεις 5.20,5.21 μπορούν να ερμηνευθούν ως εξισώσεις επιπέδων. Συνεπώς οι λύσεις (\vec{v} και \vec{m} αντίστοιχα) παριστάνουν την τομή μιας δευτεροβάθμιας επιφάνειας τοποθετημένη στην αρχή των αξόνων (σφαίρας και ελλειψοειδούς αντίστοιχα) και δύο επιπέδων που δεν διέρχονται από την αρχή των αξόνων.

Αν θεωρήσουμε την παράσταση με το διάνυσμα \vec{v} , μπορούμε να δώσουμε και άλλη γεωμετρική ερμηνεία στις Εξισώσεις 5.20,5.21. Χρησιμοποιώντας τις σχέσεις

$$\begin{aligned} \vec{a} \cdot \vec{b} &= \|\vec{a}\| \cdot \|\vec{b}\| \cos \theta \\ \|\vec{v}\| &= \|\vec{v}'\| \end{aligned}$$

γίνεται εμφανές ότι οι λύσεις \vec{v}, \vec{v}' πρέπει να σχηματίζουν ίδια οξεία γωνία με τα διανύσματα που προκύπτουν από τους συντελεστές των Εξισώσεων 5.20,5.21.

Οι παραπάνω ερμηνείες, πέρα από μια γεωμετρική αίσθηση του προβλήματος, μας παρέχουν και σημαντικές κατευθύνσεις. Συγκεκριμένα, υποδεικνύουν ότι δεν υπάρχει μια απλή γεωμετρική σχέση μεταξύ των διανυσμάτων \vec{v}, \vec{m} που να ισχύει ανεξάρτητα από τις παραμέτρους του προβλήματος, δηλαδή ανεξάρτητα από τα στοιχεία του πίνακα P_{P_2} και να μας επιτρέπει να επιλέξουμε μία από τις δύο λύσεις, βάσει γεωμετρικών επιχειρημάτων. Συνεπώς, γεωμετρικές σχέσεις θα πρέπει να αναζητηθούν σε οντότητες παράγωγες της Λύσης 5.22 και όχι κατευθείαν σε αυτήν.

Σχέσεις των ανακτώμενων ανακατασκευών. Επιλογή σωστής λύσης.

Στην Ενότητα αυτή, θα αναλύσουμε την σχετική θέση των ανακατασκευών που προκύπτουν από τις δύο λύσεις 5.22 και θα παρουσιάσουμε έναν τρόπο να αρθεί η αμφισημία, καταλήγοντας σε μοναδική λύση. Ξεκινάμε διατυπώνοντας το προς απόδειξη Θεώρημα

Θεώρημα 1. Έστω

$$\{P_{m1}^1, P_{m2}^1\}, \{P_{m1}^2, P_{m2}^2\}$$

οι "μετρικές" ανακατασκευές που προκύπτουν από τις λύσεις 5.22. Τότε οι κάμερες P_{m2}^1, P_{m2}^2 έχουν τοποθετηθεί σε θέσεις κατοπτρικές ως προς την αρχή των αξόνων (θέση της $P_{m1}^{1,2}$) δηλαδή:

Τα κέντρα προβολής C_{m2}^1, C_{m2}^2 ικανοποιούν την σχέση

$$C_{m2}^1 = -C_{m2}^2 \quad (5.24)$$

Επιπλέον, ισχύουν και οι επόμενες ιδιότητες

Θεώρημα 2. Στο παγκόσμιο σύστημα συντεταγμένων, στο οποίο η κάμερα 1 βρίσκεται στην αρχή των αξόνων και είναι προσανατολισμένη στον άξονα z , θεωρούμε τις διευθύνσεις προσανατολισμού $\mathbf{v}_{m2}^1, \mathbf{v}_{m2}^2$ των καμερών P_{m2}^1, P_{m2}^2 και τα διανύσματα θέσης C_{m2}^1, C_{m2}^2 των αντίστοιχων κέντρων προβολής. Τότε τα C_{m2}^1, C_{m2}^2 είναι οι διχοτόμοι των 2 γωνιών που σχηματίζουν τα διανύσματα $\mathbf{v}_{m2}^1, \mathbf{v}_{m2}^2$ στο επίπεδο που ορίζουν. Έχουμε δηλαδή

$$\angle C_{m2}^1, \mathbf{v}_{m2}^1 = \angle C_{m2}^1, \mathbf{v}_{m2}^2 \quad (5.25)$$

$$\angle C_{m2}^2, \mathbf{v}_{m2}^1 = \angle C_{m2}^2, \mathbf{v}_{m2}^2 \quad (5.26)$$

Εύκολα, χρησιμοποιώντας τα προηγούμενα Θεωρήματα 1,2, προκύπτει και η επόμενη σχέση

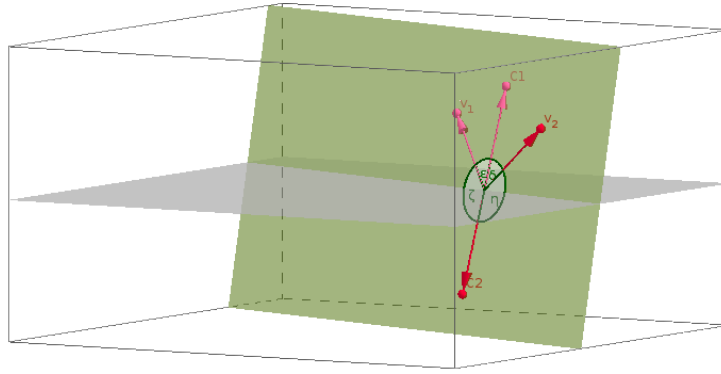
$$\angle C_{m2}^1, \mathbf{v}_{m2}^j + \angle C_{m2}^2, \mathbf{v}_{m2}^j = 180^\circ \quad (5.27)$$

Χρησιμοποιώντας τα Θεωρήματα 1,2, μπορούμε να διαχωρίσουμε την σωστή από την λάθος λύση απαιτώντας τα σημεία του χώρου που είναι ορατά από μια κάμερα, να βρίσκονται μπροστά από αυτή.

Πόρισμα 1. Η σωστή από τις λύσεις 5.22 μπορεί να επιλεγεί, από τον περιορισμό τα σημεία που είναι ορατά από την κάμερα 2, P_{m2} , να είναι μπροστά από αυτήν.

Ξεκινάμε την απόδειξη των Θεωρημάτων 1,2 παραθέτοντας κάποια γνωστά Αποτελέσματα.

Αποτέλεσμα 10. Έστω πίνακας προβολής P . Το σχετικό κέντρο προβολής C_P δεν έχει εικόνα, δηλαδή προβάλλεται στο σημείο O . Ισοδύναμα, το κέντρο προβολής $C_P = (C^T \ 1)^T$ είναι το δεξι μηδενοδιάνυσμα του πίνακα P .



Σχήμα 5.1: Παρουσιάζουμε σχηματικά τα Θεωρήματα 1,2. Στο γράφημα, απεικονίζεται το κέντρο προβολής (C^1, C^2) και η κατεύθυνση προσανατολισμού (v_1, v_2), για κάθε μία από τις λύσεις 5.22. Χρησιμοποιούμε το ροζ χρώμα για την απεικόνιση των στοιχείων της λύσεως 1 και το κόκκινο χρώμα αντίστοιχα για την λύση 2. Απεικονίζουμε ακόμα το κοινό επίπεδο των C_1, C_2, v_1, v_2 καθώς και τις γωνίες

$$\begin{aligned} \angle v_1, C_1 &\equiv \epsilon = \delta \equiv \angle v_2, C_1 \\ \angle v_1, C_2 &\equiv \zeta = \eta \equiv \angle v_2, C_2 \end{aligned}$$

Για την παραγωγή του γραφήματος, αρχικοποιήσαμε τις κάμερες 1, 2 σε τυχαία διάταξη στον χώρο, επιλύσαμε το σύστημα που έχουμε διατυπώσει και απεικονίσαμε τα δεδομένα των λύσεων 5.22 που προέκυψαν.

Απόρροια του προηγούμενου Αποτελέσματος, είναι και το επόμενο:

Αποτέλεσμα 11. Έστω πίνακας προβολής P . Ο P αποσυντίθεται ως

$$P = [KR \quad -KRC]$$

όπου C η μη-ομογενής παράσταση του κέντρου προβολής και K, R ως συνήθως.

Τα προηγούμενα Αποτελέσματα 10 και 11 μας επιτρέπουν να προσδιορίσουμε την θέση C της κάμερας. Τα συμπληρώνουμε με το επόμενο που μας δίνει τον προσανατολισμό της.

Αποτέλεσμα 12. Έστω πίνακας προβολής

$$P = [M \quad \mathbf{p}]$$

Συμβολίζουμε με το διάνυσμα \mathbf{m}_3^T την τρίτη γραμμή του πίνακα M .

Το διάνυσμα

$$\mathbf{v} = \det M \mathbf{m}_3$$

έχει την διεύθυνση του πρωτεύοντος άξονα του P και κατεύθυνση προς τον χώρο μπροστά από την κάμερα P .

Συμπληρώνουμε τα προηγούμενα, με δύο λήμματα που χαρακτηρίζουν τους πίνακες των μετρικών ανακατασκευών $P_{m_2}^1, P_{m_2}^2$ που λαμβάνουμε από τις λύσεις 5.22.

Λήμμα 1. Έστω

$$\begin{aligned} P_{m_2}^1 &= [K_2^1 R^1 \quad \mathbf{a}^1] \\ P_{m_2}^2 &= [K_2^2 R^2 \quad \mathbf{a}^2] \end{aligned}$$

οι πίνακες προβολής για την κάμερα 2 που λαμβάνουμε από τις λύσεις 5.22. Τότε ισχύει

$$\mathbf{a}^1 = \mathbf{a}^2 = \mathbf{a}$$

Απόδειξη. Λαμβάνοντας υπόψιν:

1. Την μορφή της ομογραφίας 5.6
2. Την σχέση $P_{m_2}^i = P_{P_2} H^i$ (Σχέση 5.2), όπου H^i η ομογραφία που λαμβάνουμε αντικαθιστώντας την αντίστοιχη Λύση 5.22

Το λήμμα προκύπτει άμεσα □

Λήμμα 2. Έστω $P_{m_2}^1, P_{m_2}^2$ όπως στο Λήμμα 1. Ισχύει:

$$K_2^1 R_2^1 - K_2^2 R_2^2 = \mathbf{a} \mathbf{n}^T \quad (5.28)$$

με \mathbf{n} κατάλληλο διάνυσμα.

Απόδειξη. Η απόδειξη είναι παρόμοια με αυτήν του Λήμματος 1, χρησιμοποιώντας τις ίδιες σχέσεις, το Λήμμα 1 και την παρατήρηση ότι οι ομογραφίες H^i που αναφέρουμε στην απόδειξη του Λήμματος 1 διαφέρουν μόνο στα διανύσματα \mathbf{v} (Βλ.5.6). □

Ακόμα, σημειώνουμε για λόγους πληρότητας, ότι το Λήμμα 2 είναι συνέπεια ενός γενικότερου αποτελέσματος και δεν εξαρτάται από την μορφή των λύσεων 5.22. Σκιαγραφούμε την παραγωγή αυτού του αποτελέσματος χωρίς να μπούμε σε λεπτομέρειες. Οι παρατηρήσεις που χρειαζόμαστε είναι:

1. Οι ανακατασκευές $P_{m_2}^i$ έχουν κοινή τέταρτη στήλη $\mathbf{a}^i = \mathbf{a}$. Αυτό εξασφαλίζεται από την επιλογή κοινής κλίμακας στις ομογραφίες H^i , $H^i(4, 4) = H^j(4, 4)$.
2. Είναι φανερό ότι ο θεμελιώδης πίνακας F της προβολικής ανακατασκευής (σε κανονική μορφή) αντιστοιχεί και στις μετρικές ανακατασκευές $P_m^{1,2}$. Δείχνεται εύκολα από την απαίτηση η ποσότητα $P_{P_1}^T F P_{P_2}$ να είναι αντισυμμετρική και πολλαπλασιάζοντας την με τις αντίστροφες ομογραφίες των H^i .
3. Η υπόλοιπη διαδικασία είναι παρόμοια με την απόδειξη του Θεωρήματος 9.10 του [90].

Στην συνέχεια, παράγουμε τα επιχειρήματα που οδηγούν στην Απόδειξη του Θεωρήματος 1

Λήμμα 3. Για τις ανακατασκευές $P_{m_2}^1 = [P_1 \quad \mathbf{a}]$, $P_{m_2}^2 = [P_2 \quad \mathbf{a}]$ ισχύει

$$\det P_1 = \pm \det P_2 \quad (5.29)$$

Απόδειξη. Οι μήτρες $P_{m_2}^1, P_{m_2}^2$ είναι λύσεις του Συστήματος 5.7. Συνεπώς, θα πρέπει να δίνουν την ίδια εικόνα ω^* της Q_∞^* . Έτσι, χρησιμοποιώντας και την Εξίσωση 5.1 έχουμε:

$$\begin{aligned}\omega_1^* &= P_{m_2}^1 Q_\infty^* P_{m_2}^{1T} \\ &= P_1 P_1^T \\ &= \omega_2^* \\ &= P_2 P_2^T\end{aligned}$$

Χρησιμοποιούμε τώρα δύο σχέσεις γνωστές από την γραμμική άλγεβρα

- $\det(A \cdot B) = \det A \cdot \det B$
- $\det A = \det A^T$

και έχουμε

$$\begin{aligned}\det P_1 P_1^T &= \det P_2 P_2^T \Rightarrow \\ \det P_1^2 &= \det P_2^2 \Rightarrow \\ \det P_1 &= \pm \det P_2\end{aligned}$$

□

Από την προηγούμενη Απόδειξη, είναι φανερό και το επόμενο Λήμμα

Λήμμα 4. Για τις ανακατασκευές $P_{m_2}^1, P_{m_2}^2$ του Λήμματος 3, έχουμε

$$K_{m_2}^1 = K_{m_2}^2$$

Απόδειξη. Από την ισότητα των ω_1^*, ω_2^* και την διαγώνια μορφή των μητρών εσωτερικής βαθμονόμησης $K_{m_2}^1, K_{m_2}^2$ προκύπτει το ζητούμενο. □

Στην συνέχεια, θα εκλεπτύνουμε το Λήμμα 3, ώστε να αρθεί η αμφισημία στο πρόσημο. Πριν προχωρήσουμε, ανακεφαλαιώνουμε τον συμβολισμό μας

1. $P_{m_2}^1, P_{m_2}^2$ είναι οι 2 λύσεις για την κάμερα 2 που λαμβάνουμε από την επίλυση του Συστήματος 5.13.
2. Σε συμφωνία με τα Λήμματα 1,3,4 έχουμε

$$P_{m_2}^i = [P_i \quad \mathbf{a}] = [K_2 R_2^i \quad \mathbf{a}]$$

3. Σε συμφωνία με το Λήμμα 2

$$P_1 - P_2 = \mathbf{a} \mathbf{n}^T$$

4. Συμβολίζουμε τα αντίστοιχα κέντρα προβολής ως $\mathbf{C}_{m_2}^1, \mathbf{C}_{m_2}^2$.

Λήμμα 5. Για τις ανακατασκευές $P_{m_2}^1, P_{m_2}^2$ ισχύει

$$\mathbf{C}_{m_1}^1 = \mathbf{C}_{m_2}^2 \equiv \mathbf{C} \iff \mathbf{n}^T \mathbf{C} = 0 \quad (5.30)$$

Απόδειξη. Έχουμε

$$\begin{aligned} P_{m_2}^1 \begin{pmatrix} \mathbf{C} \\ 1 \end{pmatrix} &= \mathbf{0} \\ \iff \mathbf{a} &= -K_2 R_2^1 \mathbf{C} \end{aligned} \quad (5.31)$$

Αντίστοιχα, από την μήτρα $P_{m_2}^2$ έχουμε

$$\begin{aligned} P_{m_2}^2 \begin{pmatrix} \mathbf{C} \\ 1 \end{pmatrix} &= \mathbf{0} \\ \iff K_2 R_2^1 \mathbf{C} + \mathbf{a} \mathbf{n}^T \mathbf{C} + \mathbf{a} &= \mathbf{0} \\ \iff K_2 R_2^1 \mathbf{C} + \mathbf{a} \mathbf{n}^T \mathbf{C} - K_2 R_2^1 \mathbf{C} &= \mathbf{0}, \text{ από την 5.31} \\ \iff \mathbf{a} \mathbf{n}^T \mathbf{C} &= \mathbf{0} \\ \iff \begin{pmatrix} \alpha_1 \mathbf{n}^T \\ \alpha_2 \mathbf{n}^T \\ \alpha_3 \mathbf{n}^T \end{pmatrix} \mathbf{C} &= \mathbf{0} \\ \iff \mathbf{n}^T \mathbf{C} &= 0 \end{aligned} \quad (5.32)$$

□

Λήμμα 6. Για τις ανακατασκευές $P_{m_2}^1, P_{m_2}^2$ ισχύει

$$\mathbf{C}_{m_1}^1 = -\mathbf{C}_{m_2}^2 \equiv \mathbf{C} \iff \mathbf{n}^T \mathbf{C} = 2 \quad (5.33)$$

Απόδειξη. Οπώς και στην απόδειξη του Λήμματος 5 που προηγήθηκε

$$\begin{aligned} P_{m_2}^1 \begin{pmatrix} \mathbf{C} \\ 1 \end{pmatrix} &= \mathbf{0} \\ \iff \mathbf{a} &= -K_2 R_2^1 \mathbf{C} \end{aligned} \quad (5.34)$$

Αντίστοιχα, από την μήτρα $P_{m_2}^2$ έχουμε

$$\begin{aligned} P_{m_2}^2 \begin{pmatrix} -\mathbf{C} \\ 1 \end{pmatrix} &= \mathbf{0} \\ \iff -K_2 R_2^1 \mathbf{C} - \mathbf{a} \mathbf{n}^T \mathbf{C} + \mathbf{a} &= \mathbf{0} \\ \iff \mathbf{a} - \mathbf{a} \mathbf{n}^T \mathbf{C} + \mathbf{a} &= \mathbf{0}, \text{ από την 5.34} \\ \iff \mathbf{a}(2 - \mathbf{n}^T \mathbf{C}) &= \mathbf{0} \\ \iff \mathbf{n}^T \mathbf{C} = 2, \text{ με την προϋπόθεση } \mathbf{a} \neq \mathbf{0} \end{aligned} \quad (5.35)$$

□

Συμπληρώνουμε καθένα από τα προηγούμενα Λήμματα 5,6, με μια ισοδύναμη του σχέση. Για την Απόδειξη τους, χρησιμοποιούμε το επόμενο, γνωστό από την γραμμική άλγεβρα, Αποτέλεσμα

Αποτέλεσμα 13. Για κάθε τετραγωνική, αντιστρέψιμη μήτρα X , διάνυσμα στήλη \mathbf{c} και διάνυσμα γραμμή \mathbf{r} ισχύει

$$\det(X + \mathbf{c}\mathbf{r}) = \det X \cdot \det(1 + \mathbf{r}X^{-1}\mathbf{c}) \quad (5.36)$$

Λήμμα 7. Για τις ανακατασκευές $P_{m_2}^1, P_{m_2}^2$ ισχύει

$$\begin{aligned} \det P^1 &= \det P^2 \iff \\ \mathbf{n}^T \mathbf{C}_1 &= 0 \end{aligned} \quad (5.37)$$

Απόδειξη. Χρησιμοποιούμε την Σχέση 5.36. Η σχέση αυτή έχει εφαρμογή, γιατί

1. Ο πίνακας P^1 είναι πλήρους rank (3) για κάθε πίνακα προβολής, με εξαίρεση την περίπτωση της κάμερας στο άπειρο που δεν μας αφορά εδώ.
2. Ο πίνακας P^2 εκφράζεται συναρτήσει του P^1 και των \mathbf{n}, \mathbf{a} , επομένως μπορεί να εφαρμοστεί η Σχέση 13 στον υπολογισμό της $\det P^2$.

Εφαρμόζουμε όσα περιγράψαμε και έχουμε:

$$\begin{aligned}
 \det P^2 &= \det P^1 \\
 \iff 1 + \mathbf{n}^T R_2^{1T} K_2^{-1} \mathbf{a} &= 1 \\
 \iff \mathbf{n}^T R_2^{1T} K_2^{-1} \mathbf{a} &= 0 \\
 \iff -\mathbf{n}^T R_2^{1T} K_2^{-1} K_2 R_2^1 \mathbf{C}_1 &= 0, \text{ από Σχέση 5.31 } \mathbf{a} = -K_2 R_2^1 \mathbf{C}_1 \\
 \iff \mathbf{n}^T \mathbf{C}_1 &= 0, \text{ από } RR^T = I \text{ για πίνακες στροφής}
 \end{aligned} \tag{5.38}$$

□

Λήμμα 8. Για τις ανακατασκευές P_{m2}^1, P_{m2}^2 ισχύει

$$\begin{aligned}
 \det P^1 &= -\det P^2 \iff \\
 \mathbf{n}^T \mathbf{C}_1 &= 2
 \end{aligned} \tag{5.39}$$

Απόδειξη. Όμοια με την Απόδειξη του Λήμματος 7 έχουμε:

$$\begin{aligned}
 \det P^2 &= -\det P^1 \\
 \iff 1 + \mathbf{n}^T R_2^{1T} K_2^{-1} \mathbf{a} &= -1 \\
 \iff \mathbf{n}^T R_2^{1T} K_2^{-1} \mathbf{a} &= -2 \\
 \iff -\mathbf{n}^T R_2^{1T} K_2^{-1} K_2 R_2^1 \mathbf{C}_1 &= -2, \text{ από Σχέση 5.31 } \mathbf{a} = -K_2 R_2^1 \mathbf{C}_1 \\
 \iff \mathbf{n}^T \mathbf{C}_1 &= 2, \text{ από } RR^T = I \text{ για πίνακες στροφής}
 \end{aligned} \tag{5.40}$$

□

Στην συνέχεια, θα αποδείξουμε ότι η περίπτωση των ομόσημων διακρινουσών πρέπει να απορριφθεί, με απαγωγή σε Άτοπο. Τα επιχειρήματα που θα το καταδείξουν, ξεδιπλώνονται ακολούθως.

Πριν διατυπώσουμε το επόμενο λήμμα, παρουσιάζουμε τον συμβολισμό μας:

1. Σε συμφωνία με τις προηγούμενες Ενότητες, συμβολίζουμε τον πίνακα της κάμερας 2 στο προβολικό σύστημα συντεταγμένων με

$$P_{P2}$$

2. Θεωρούμε ότι ο P_{P2} προέρχεται από το κανονικό ζευγάρι καμερών

$$\{ [I \ \mathbf{0}], [[\mathbf{a}]_x F \ \mathbf{a}] \} \tag{5.41}$$

όπου F ο Θεμελιώδης πίνακας που αντιστοιχεί στο πρόβλημα και \mathbf{a} το αριστερό μηδενοδιάνυσμα του F , δηλαδή

$$F^T \mathbf{a} = \mathbf{0}$$

Η επιλογή του συγκεκριμένου ζεύγους καμερών είναι συνηθισμένη στην βιβλιογραφία.

3. Με $[\mathbf{a}]_x$ συμβολίζουμε τον αντισυμμετρικό πίνακα που ορίζεται για τον υπολογισμό του εξωτερικού γινομένου μέσω γινομένου μητρών, δηλαδή

$$[\mathbf{a}]_x \mathbf{v} \equiv \mathbf{a} \times \mathbf{v}$$

4. Συμβολίζουμε \mathbf{e} το δεξί μηδενοδιάνυσμα του F , δηλαδή

$$F \mathbf{e} = \mathbf{0}$$

Λήμμα 9. Έστω

$$P_{P2} = [A \ \mathbf{a}] = [[\mathbf{a}]_x F \ \mathbf{a}]$$

ο πίνακας της κάμερας 2 στο προβολικό πλαίσιο αναφοράς και \mathbf{p}, \mathbf{p}' οι λύσεις για την θέση του π_∞ που λαμβάνουμε από τις Σχέσεις 5.22, δηλαδή

$$\begin{aligned} \mathbf{p}^T &= (p_1 \ p_2 \ p_3) \\ \mathbf{p}'^T &= (p'_1 \ p'_2 \ p'_3) \end{aligned}$$

Τότε ισχύει

$$\mathbf{p} - \mathbf{p}' = \psi \mathbf{e}_f \quad (5.42)$$

όπου

$$\mathbf{e}_f = \begin{pmatrix} e_1/f_1^2 \\ e_2/f_1^2 \\ e_3 \end{pmatrix}$$

f_1 : Το εστιακό μήκος της κάμερας 1

Πριν προχωρήσουμε στην απόδειξη, υπενθυμίζουμε ότι στο πλαίσιο στο οποίο θέσαμε το πρόβλημα, οι πίνακες εσωτερικής βαθμονόμησης K_1, K_2 των καμερών 1, 2 θεωρούνται διαγώνιοι με μοναδική παράμετρο τα αντίστοιχα εστιακά μήκη f_1, f_2 .

Απόδειξη. Χρησιμοποιούμε την Σχέση 5.7, και υπενθυμίζουμε ότι οι λύσεις 5.22 έχουν κοινά f_1 . Έχουμε

$$\begin{aligned} \omega_1^* = \omega_2^* &\iff \\ P_{P2} \begin{bmatrix} K_1 K_1^T & -K_1 K_1^T \mathbf{p} \\ -\mathbf{p}^T K_1 K_1^T & \mathbf{p}^T K_1 K_1^T \mathbf{p} \end{bmatrix} P_{P2}^T &= \\ P_{P2} \begin{bmatrix} K_1 K_1^T & -K_1 K_1^T \mathbf{p}' \\ -\mathbf{p}'^T K_1 K_1^T & \mathbf{p}'^T K_1 K_1^T \mathbf{p}' \end{bmatrix} P_{P2}^T &\iff \\ AK_1 K_1^T A^T - AK_1 K_1^T \mathbf{p} \mathbf{a}^T - \mathbf{a} \mathbf{p}^T K_1 K_1^T A^T + \mathbf{a} \mathbf{p}^T K_1 K_1^T \mathbf{p} \mathbf{a}^T &= \\ AK_1 K_1^T A^T - AK_1 K_1^T \mathbf{p}' \mathbf{a}^T - \mathbf{a} \mathbf{p}'^T K_1 K_1^T A^T + \mathbf{a} \mathbf{p}'^T K_1 K_1^T \mathbf{p}' \mathbf{a}^T & \end{aligned}$$

Παρατηρούμε ότι (Βλ. Λύσεις 5.22 ,Σχέση 5.23)

$$(f_1^2 p_1^2 + f_1^2 p_2^2 + p_3^2) \mathbf{a} \mathbf{a}^T = \mathbf{a} \mathbf{p}^T K_1 K_1^T \mathbf{p} \mathbf{a}^T = \mathbf{a} \mathbf{p}'^T K_1 K_1^T \mathbf{p}' \mathbf{a}^T$$

Προφανώς το ίδιο ισχύει και για τους όρους

$$AK_1 K_1^T A^T$$

που είναι κοινοί στις δύο προς εξίσωση ποσότητες.
Συνεχίζοντας τους προηγούμενους υπολογισμούς και αφού απαλείψουμε τις προηγούμενες ίσες ποσότητες παίρνουμε

$$AK_1K_1^T((\mathbf{p} - \mathbf{p}')\mathbf{a}^T) + (\mathbf{a}(\mathbf{p}^T - \mathbf{p}'^T))K_1K_1^T A^T = 0 \iff Q + Q^T = 0 \quad (5.43)$$

όπου κάναμε τον ορισμό

$$Q \equiv AK_1K_1^T((\mathbf{p} - \mathbf{p}')\mathbf{a}^T)$$

Η Εξίσωση 5.43 μας δίνει ότι η μήτρα Q είναι αντισυμμετρική. Όπως είναι γνωστό, αντισυμμετρικές μήτρες έχουν μηδενική διαγώνιο. Από αυτή την ιδιότητα, δηλαδή τον μηδενισμό των στοιχείων της διαγωνίου, εξάγουμε τις επόμενες σχέσεις

$$\begin{aligned} Q &= AK_1K_1^T((\mathbf{p} - \mathbf{p}')\mathbf{a}^T) \\ &= AK_1K_1^T \begin{pmatrix} (p_1 - p'_1)\mathbf{a}^T \\ (p_2 - p'_2)\mathbf{a}^T \\ (p_3 - p'_3)\mathbf{a}^T \end{pmatrix} \\ &= A \begin{pmatrix} f_1^2(p_1 - p'_1)\mathbf{a}^T \\ f_1^2(p_2 - p'_2)\mathbf{a}^T \\ (p_3 - p'_3)\mathbf{a}^T \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{A}^{1T} \\ \mathbf{A}^{2T} \\ \mathbf{A}^{3T} \end{pmatrix} (\Delta_f \alpha_1 \quad \Delta_f \alpha_2 \quad \Delta_f \alpha_3) \end{aligned} \quad (5.44)$$

όπου στην Σχέση 5.44 ορίσαμε

$$\Delta_f \equiv \begin{pmatrix} f_1^2(p_1 - p'_1) \\ f_1^2(p_2 - p'_2) \\ (p_3 - p'_3) \end{pmatrix}$$

Εξετάζοντας τώρα τα διαγώνια στοιχεία της μήτρας Q και χρησιμοποιώντας την γραφή της Εξίσωσης 5.44 έχουμε τις επόμενες σχέσεις

$$\mathbf{A}^{1T} \Delta_f = 0 \quad (5.45)$$

$$\mathbf{A}^{2T} \Delta_f = 0 \quad (5.46)$$

$$\mathbf{A}^{3T} \Delta_f = 0 \quad (5.47)$$

Οι τρεις Σχέσεις 5.45, 5.46, 5.47 μαζί με την υπόθεση

$$P_{P2} = [A \quad \mathbf{a}] = [[\mathbf{a}]_x F \quad \mathbf{a}]$$

μας δίνουν

$$[\mathbf{a}]_x F \Delta_f = \mathbf{0} \quad (5.48)$$

Από την Εξίσωση 5.48 και από το γεγονός ότι ο πίνακας $[\mathbf{a}]_x F$ έχει για μηδενοδιάνυσμα το \mathbf{e} έχουμε

$$\Delta_f = \psi \mathbf{e} \quad (5.49)$$

όπου ψ κατάλληλη σταθερά.

Συνεπώς, από την Σχέση 5.49 και με απλή αναδιάταξη των όρων παίρνουμε

$$\begin{pmatrix} p_1 - p'_1 \\ p_2 - p'_2 \\ p_3 - p'_3 \end{pmatrix} = \psi \mathbf{e}_f$$

□

Χρησιμοποιώντας τώρα το προηγούμενο Λήμμα 9, και κάνοντας ξανά τις ίδιες υποθέσεις, έχουμε το επόμενο

Λήμμα 10. Έστω ότι ισχύουν οι ίδιες υποθέσεις και ακολουθούμε τον ίδιο συμβολισμό με το Λήμμα 9 και με τα προηγούμενα Λήμματα. Τότε ισχύει

$$\det P_1 = -\det P_2$$

Απόδειξη. Θα αποδείξουμε το ζητούμενο υποθέτοντας

$$\det P_1 = \det P_2$$

και καταλήγοντας σε άτοπο.

Σύμφωνα με το Λήμμα 13 και συγκεκριμένα την Σχέση 5.36, ισοδύναμα απαιτούμε

$$\mathbf{n}^T \mathbf{C}_1 = 0 \quad (5.50)$$

Στην ανωτέρω Σχέση 5.50, πρέπει να αποσαφηνίσουμε το, μέχρι στιγμής απροσδιόριστο, διάνυσμα \mathbf{n} . Για τον σκοπό αυτό:

1. Υπενθυμίζουμε την σχέση ορισμού του \mathbf{n} (Σχέση 5.28)
2. Την σχέση των P_{P_2}, P_{M_2}, H (Σχέσεις 5.2, 5.6)
3. Τους προηγούμενους ορισμούς, συμβολισμούς και υποθέσεις μας

Έχουμε λοιπόν

$$\begin{aligned} P^1 &= AK_1 - \mathbf{a}\mathbf{p}^T K_1 \\ P^2 &= AK_1 - \mathbf{a}\mathbf{p}'^T K_1 \\ \iff P^2 &= P^1 + \mathbf{a}(\mathbf{p} - \mathbf{p}')^T K_1 \equiv P^1 + \mathbf{a}\mathbf{n}^T \end{aligned}$$

Συνεπώς, η Σχέση 5.50 δίνει

$$(\mathbf{p} - \mathbf{p}')^T \mathbf{K}_1 \mathbf{C}_1 = 0 \quad (5.51)$$

Στην συνέχεια, θα εξάγουμε κάποιες βοηθητικές Σχέσεις

$$\begin{aligned} P_{M_2}^1 \begin{pmatrix} \mathbf{C}_1^1 \\ 1 \end{pmatrix} &= 0 \iff \\ P_{P_2} H^1 \begin{pmatrix} \mathbf{C}_1^1 \\ 1 \end{pmatrix} &= 0 \iff \\ P_{P_2} \begin{pmatrix} K_1 \mathbf{C}_1^1 \\ -\mathbf{p}^T K_1 \mathbf{C}_1^1 + 1 \end{pmatrix} &= 0 \end{aligned}$$

Στο σημείο αυτό, χρησιμοποιούμε την υπόθεση ότι ο P_{P2} προέρχεται από το κανονικό ζευγάρι καμερών, συγκεκριμένα ότι έχει την μορφή της Σχέσης 5.41. Συνεπώς, το μηδενοδιάνυσμα του θα έχει την μορφή

$$\begin{pmatrix} \mathbf{e} \\ 0 \end{pmatrix}$$

Έχουμε έτσι:

$$P_{P2} \begin{pmatrix} K_1 \mathbf{C}^1 \\ -\mathbf{p}^T K_1 \mathbf{C}^1 + 1 \end{pmatrix} = 0 \iff \begin{aligned} K_1 \mathbf{C}^1 &= \psi \mathbf{e}, \text{ όπου } \psi \text{ κατάλληλη σταθερά} & (5.52) \\ -\mathbf{p}^T K_1 \mathbf{C}^1 + 1 &= 0 & (5.53) \end{aligned}$$

Χρησιμοποιούμε τώρα το Λήμμα 9, συγκεκριμένα την Σχέση 5.42, και τις Σχέσεις 5.51, 5.52 έχουμε:

$$\begin{aligned} \mathbf{e}_F^T \mathbf{e} &= 0 \iff \\ e_1^2/f_1^2 + e_2^2/f_1^2 + e_3^2 &= 0 \end{aligned} \quad (5.54)$$

Το άτοπο της Εξίσωσης 5.54 στο πεδίο των πραγματικών αριθμών ($\mathbf{e} \neq \mathbf{0}$), είναι προφανές. Συνεπώς από το Λήμμα 3, αποδείχθηκε ότι

$$\det P^1 = -\det P^2$$

□

Συμπληρώνουμε την προηγούμενη Απόδειξη με μερικές παρατηρήσεις:

1. Οι σχέσεις των Λημμάτων 5,6 μπορούν να προκύψουν και από τις Σχέσεις

$$\begin{aligned} P_{m2}^i \mathbf{C} &= P_{P2} H^i \mathbf{C} \\ H^i \mathbf{C} &= \mathbf{C}_P = (\mu \mathbf{e} \quad 0)^T \end{aligned}$$

Η προσέγγιση αυτή είναι περισσότερο γεωμετρική, σε σχέση με τον αλγεβρικό τρόπο που χρησιμοποιήθηκε στην προηγούμενη ανάλυση.

2. Η χρήση της σχέσης

$$\begin{aligned} K_2 R^2 &= K_2 R^1 + \mathbf{a} \mathbf{n}^T \Rightarrow \\ K_2 R^2 \mathbf{C}^2 &= \pm (K_2 R^1 + \mathbf{a} \mathbf{n}^T) \mathbf{C}^1 \end{aligned}$$

δεν παράγει χρήσιμα αποτελέσματα. Δίνει όμως συνθήκες που μπορούν να ερμηνευθούν γεωμετρικά, με την προϋπόθεση ότι γνωρίζουμε την σχέση των $\mathbf{C}^1, \mathbf{C}^2$. Συγκεκριμένα

$$\mathbf{C}^1 = \mathbf{C}^2 \equiv \mathbf{C} \Rightarrow \mathbf{C} \text{ άξονας περιστροφής } R^{1T} R^2$$

Συγκεντρώνοντας τώρα τα προηγούμενα αποδειχθέντα συμπεράσματα, βρίσκoμαστε σε θέση να αποδείξουμε το Θεώρημα 1.

Απόδειξη Θεωρήματος 1. Από το Λήμμα 10 έχουμε

$$\det P^1 = -\det P^2$$

Χρησιμοποιώντας τα Λήμματα 6,8 έχουμε ισοδύναμα

$$\mathbf{C}_{m2}^1 = -\mathbf{C}_{m2}^2$$

□

Στην συνέχεια αποδεικνύουμε το Θεώρημα 2. Συγκεντρώνουμε για μία ακόμη φορά τον συμβολισμό μας

1. Συμβολίζουμε με $\mathbf{v}_{m2}^1, \mathbf{v}_{m2}^2$ τα διανύσματα που μας δίνουν την κατεύθυνση προσανατολισμού των καμερών P_{m2}^1, P_{m2}^2 αντίστοιχα, σε συμφωνία με το Θεώρημα 2.
2. Θεωρούμε ότι για την ανακατασκευή P_{m2}^1 ισχύουν

$$\begin{aligned} \det P^1 &> 0 \\ \mathbf{C} &\equiv \mathbf{C}_{m2}^1 \end{aligned}$$

Απόδειξη Θεωρήματος 2. Με χρήση προηγούμενων Αποτελεσμάτων (Αποτελέσματα 11,12, Λήμμα 1, Θεώρημα 1) έχουμε για τον πίνακα P_{m2}^1 :

$$\begin{aligned} K_2 R^1 \mathbf{C} &= -\mathbf{a} \iff \\ \begin{pmatrix} f_2 \mathbf{R}_1^T \\ f_2 \mathbf{R}_2^T \\ \mathbf{R}_3^T \end{pmatrix} \mathbf{C} &= -\mathbf{a} \iff \\ \mathbf{R}_3^T \mathbf{C} &= -a_3 \end{aligned} \tag{5.55}$$

Επειδή $\det P^1 = \det K_2 R_1 > 0$ έχουμε

$$\mathbf{v}_{2m}^1 = \mathbf{R}_3$$

Συνεπώς, από την Σχέση 5.55, έχουμε

$$\mathbf{v}_{2m}^1{}^T \mathbf{C} = \|\mathbf{v}_{2m}^1\| \|\mathbf{C}\| \cos \angle \mathbf{C}, \mathbf{v}_{2m}^1 = -a_3 \tag{5.56}$$

Στην Σχέση 5.56, έχουμε

$$\|\mathbf{R}_3^T\| = 1, \text{ αφού ο πίνακας } R^1 \text{ είναι πίνακας στροφής}$$

Το διάνυσμα \mathbf{C} , μπορεί πάντοτε να κλιμακωθεί κατάλληλα ώστε να είναι μοναδιαίο. Συγκεκριμένα, η συνθήκη που χρειάζεται να ικανοποιήσουμε είναι

$$\|K_2 \mathbf{a}\| = 1$$

αφού η στροφή δεν επηρεάζει το μέτρο ενός διανύσματος. Σύμφωνα με τα προηγούμενα, η Σχέση 5.56 δίνει

$$\cos \angle \mathbf{C}, \mathbf{v}_{2m}^1 = -a_3$$

Με τελείως αντίστοιχο τρόπο, κάνοντας χρήση των ιδιοτήτων

$$\begin{aligned} \mathbf{C}_{m2}^2 &= -\mathbf{C} \\ \det P^2 &< 0 \\ \mathbf{a}^1 &= \mathbf{a}^2 \end{aligned}$$

καταλήγουμε

$$\cos \angle \mathbf{C}, \mathbf{v}_{m2}^2 = -a_3$$

και τελείως αντίστοιχα αποδεικνύονται και οι υπόλοιπες σχέσεις που συμπληρώνουν το Θεώρημα, δηλαδή

$$\begin{aligned} \cos \angle \mathbf{C}^2, \mathbf{v}_{m2}^1 &= a_3 \\ \cos \angle \mathbf{C}^2, \mathbf{v}_{m2}^2 &= a_3 \\ \cos \angle \mathbf{C}, \mathbf{v}_{m2}^1 + \cos \angle \mathbf{C}^2, \mathbf{v}_{m2}^1 &= 180^\circ \\ \cos \angle \mathbf{C}, \mathbf{v}_{m2}^2 + \cos \angle \mathbf{C}^2, \mathbf{v}_{m2}^2 &= 180^\circ \end{aligned}$$

Για να ολοκληρωθεί η Απόδειξη, μένει να δείξουμε ότι τα $\mathbf{v}_{m2}^{1,2}, \mathbf{C}_{m2}^{1,2}$ ανήκουν σε κοινό επίπεδο. Η απόδειξη αυτή θα γίνει κατασκευαστικά, στην ανάλυση που ακολουθεί, και για αυτό παραλείπεται τώρα. \square

Στην συνέχεια, θα αναλύσουμε τα δεδομένα του προβλήματος όπως αυτά διαμορφώνονται με τα προηγούμενα Αποτελέσματα και θα σχολιάσουμε την ευκολία διαχωρισμού των λύσεων.

Αρχικά, υποστηρίζουμε ότι όσο μεγαλύτερη είναι η γωνία που σχηματίζουν οι κατευθύνσεις των οπτικών αξόνων των καμερών, τόσο πιο εύκολα θα διαχωρίζονται. Για να υπολογίσουμε την γωνία αυτή, χρησιμοποιούμε το Θεώρημα 2. Έχουμε:

$$\angle \mathbf{v}_{m2}^1, \mathbf{v}_{m2}^2 = 2 \angle \mathbf{v}_{m2}^1, \mathbf{C}^1 \equiv 2\phi \quad (5.57)$$

Στην προηγούμενη Σχέση 5.57, Θεωρούμε ότι υπολογίζουμε την μικρότερη γωνία των $\mathbf{v}_{m2}^{1,2}$, και ότι συνεπώς το \mathbf{C}^1 είναι διχοτόμος αυτής της γωνίας. Έχουμε ακόμα:

$$\cos \phi = \mathbf{v}_{m2}^1 \cdot \mathbf{C}^1 = -a_3, \text{ από Απόδειξη Θεωρήματος 2} \quad (5.58)$$

Ερμηνεύοντας την Σχέση 5.58, έχουμε ότι όσο πιο έντονη είναι η κίνηση κατά την διεύθυνση του οπτικού άξονα της κάμερας 2 (a_3), τόσο μικρότερη θα είναι η γωνία ϕ και συνεπώς δυσκολότερος ο διαχωρισμός των λύσεων. Υπενθυμίζουμε ότι επειδή το διάνυσμα \mathbf{a} είναι κανονικοποιημένο, το a_3 μας δίνει την ένταση της κίνηση κατά την συγκεκριμένη διεύθυνση συγκριτικά με τις υπόλοιπες δύο, ορθογώνιες της. Συγκεντρώνουμε τα προηγούμενα σε ένα πόρισμα:

Πόρισμα 2. Η γωνία 2ϕ που σχηματίζουν οι κατευθύνσεις προσανατολισμού των λύσεων $P_{m2}^{1,2}$, εξαρτάται από την κίνηση κατά την διεύθυνση του οπτικού άξονα της κάμερας 2 και μειώνεται όσο η κίνηση σε αυτή την διεύθυνση γίνεται εντονότερη.

Τέλος, κάνουμε ένα σχόλιο σχετικά με τον υπολογισμό της γωνίας ϕ . Μπορούμε να αποφύγουμε την ανάγκη κανονικοποίησης των μητρών προβολής $P_{m2}^{1,2}$, αν υπολογίσουμε την γωνία 2ϕ . Τα διανύσματα $\mathbf{v}_{m2}^{1,2}$ είναι "αυτόματα"

κανονικοποιημένα, ως γραμμές μητρών στροφής, και συνεπώς το εσωτερικό γινόμενο τους δίνει ευθέως το $\cos 2\phi$.

Πέρα από την γωνία των κατευθύνσεων προσανατολισμού των καμερών, ένα άλλο μέγεθος που πρέπει να εξετάσουμε για την επιλογή της πιθανότερης λύσης, είναι η γωνία που σχηματίζουν οι οπτικοί άξονες των καμερών 1, 2. "Εξωτικές" γωνίες καταδεικνύουν λάθος λύση, αφού στην πράξη δεν παρατηρούνται σε κάμερες που παρατηρούν κοινό τμήμα της Σκηνης. Για τον σκοπό αυτό, θα δώσουμε μια τριγωνομετρική σχέση μεταξύ των ϕ, θ_1, θ_2 , όπου με $\theta_{1,2}$ συμβολίζουμε τις γωνίες των οπτικών αξόνων των καμερών για τις λύσεις 1, 2 αντίστοιχα.

Ξεκινάμε διατυπώνοντας την σχέση αυτή, την οποία και αποδεικνύουμε ακολούθως. Η Απόδειξη αυτή καλύπτει και τις εκκρεμότητες στην Απόδειξη του Θεωρήματος 2.

Αποτέλεσμα 14. Έστω $P_{m2}^{1,2}$ οι λύσεις για την κάμερα 2 όπως αυτές προκύπτουν από τις Σχέσεις 5.22. Τότε, ισχύει η σχέση

$$\cos \theta_2 = \cos 2\phi \cos \theta_1 + A \sin 2\phi \quad (5.59)$$

όπου στην προηγούμενη Σχέση 5.59

$$\begin{aligned} \theta_i &= \angle \mathbf{v}_{m2}^i, \mathbf{v}_{m1} \\ 2\phi &= \angle \mathbf{v}_{m2}^1, \mathbf{v}_{m2}^2, \text{ μετρούμενη συμβατά με τον κανόνα του δεξιού χεριού} \\ A &= R_{23}^1 \cos \theta_x + R_{13}^1 \sin \theta_x \\ \tan \theta_x &= a_1/a_2 \end{aligned}$$

Η Σχέση 5.59 γράφεται ισοδύναμα και στην μορφή:

$$\cos \theta_2 = \cos 2\phi \cos \theta_1 + |\sin \theta_1| \cos \psi \sin 2\phi \quad (5.60)$$

$$\text{με } \psi = \angle \left(\begin{pmatrix} R_{13}^1 \\ R_{23}^1 \end{pmatrix}, \begin{pmatrix} \sin \theta_x \\ \cos \theta_x \end{pmatrix} \right)$$

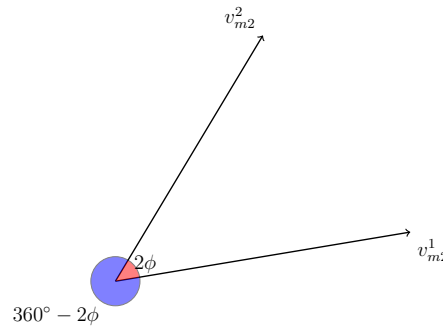
Απόδειξη. Κάνουμε αρχικά κάποιες υποθέσεις, προκειμένου να απλοποιήσουμε τον συμβολισμό μας και χωρίς να βλάψουμε την γενικότητα της Απόδειξης. Διευκρινίζουμε λοιπόν τα ακόλουθα:

1. Θεωρούμε ότι, από τις δύο ανακατασκευές $P_{m2}^{1,2}$ που έχουν προκύψει ως λύσεις των Εξισώσεων 5.22, η σωστή λύση είναι η P_{m2}^1 ενώ η P_{m2}^2 δίνει εσφαλμένη θέση της κάμερας 2 στον χώρο.
2. Θεωρούμε ότι

$$\text{sign}(\det P^1) > 0$$

η υπόθεση αυτή απλοποιεί τον συμβολισμό μας, καθώς μας επιτρέπει να απαλείψουμε από την κατεύθυνση προσανατολισμού της κάμερας και τον μετασχηματισμό που θα χρησιμοποιήσουμε το πρόσημο της προηγούμενης οριζουσας. Όμως, δεν επηρεάζει ουσιαστικά με οποιαδήποτε τρόπο την Απόδειξη.

3. Ιδιαίτερη προσοχή πρέπει να δοθεί στον τρόπο που μετράμε τις γωνίες των διανυσμάτων, ώστε οι εξισώσεις μας να είναι συμβατές μεταξύ τους. Για αυτό, τονίζουμε ότι δεν μπορούμε να θεωρούμε πλέον, όπως προηγούμενα, ότι για παράδειγμα η γωνία ϕ είναι οξεία, αλλά η τιμή τους προκύπτει από γνωστές συμβάσεις.



Σχήμα 5.2: Ανάλογα με το ποιο διάνυσμα θα περιστρέψουμε για να το ευθυγραμμίσουμε με το άλλο, ο πίνακας στροφής θα εκτελεί την μικρότερη (κόκκινη) η την μεγαλύτερη (μπλε) στροφή.

Για την Απόδειξη του ζητήματος, εφαρμόζουμε στον χώρο τρισδιάστατους μετασχηματισμούς ομοιότητας. Οι μετασχηματισμοί ομοιότητας, όπως είναι γνωστό, διατηρούν τις γωνίες και συνεπώς δεν επηρεάζουν την ορθότητα των συλλογισμών μας. Σκοπός μας είναι να μετασχηματίσουμε τον χώρο ώστε το θεωρούμενο σύστημα συντεταγμένων του χώρου να έχει βολικό προσανατολισμό, απλοποιώντας την σχέση των υπό εξέταση οντοτήτων.

Η διαδικασία που ακολουθούμε, μπορεί να οπτικοποιηθεί ως τοποθέτηση μιας "εικονικής" κάμερας και προσανατολισμός της, ώστε τα $\mathbf{v}_{m2}^{1,2}, \mathbf{C}^{1,2}$ να βρίσκονται σε βολικά επίπεδα της κάμερας αυτής (αναφερόμαστε στο πρωτεύων επίπεδο και τα $x-, y-$ επίπεδα της κάμερας) διευκολύνοντας έτσι και την διαίσθηση μας να αναπτυχθεί.

Εφαρμόζουμε λοιπόν στον χώρο, τις περιστροφές

$$R_x R_{perm} R^1 \quad (5.61)$$

όπου

R^1 :ο πίνακας στροφής του P_{m2}^1

R_{perm} :Περιστροφικός μετασχηματισμός που αντιμεταθέτει τα x_1, x_3 ενός διανύσματος $(x_1 \ x_2 \ x_3)^T$

R_x :κατάλληλος περιστροφικός μετασχηματισμός που τοποθετεί το \mathbf{C}^1 στο κατάλληλο επίπεδο

Μετά την εφαρμογή του R^1 , από την ορθογωνιότητα του πίνακα στροφής και το Αποτέλεσμα 12, η κατεύθυνση προσανατολισμού της κάμερας 2 γίνεται

$$R^1 \mathbf{v}_{m2}^1 = (0 \ 0 \ 1)^T \equiv \mathbf{v}_{m2}^{1n}$$

Εφαρμόζουμε στην συνέχεια τον μετασχηματισμό R_{perm} , η χρήση του οποίου εξυπηρετεί καθαρά λόγους οπτικοποίησης των αποτελεσμάτων

$$R_{perm} = R_y(90^\circ) = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{bmatrix}$$

Η περιστροφή γύρω από τον άξονα x ολοκληρώνει τους μετασχηματισμούς, περιστρέφοντας τους άξονες συντεταγμένων ώστε το \mathbf{C}^1 να βρίσκεται στο επίπεδο z . Ταυτόχρονα, δεν αλλοιώνει την μορφή του διανύσματος $\mathbf{v}_{\mathbf{m}2}^1$ στο υπό διαμόρφωση σύστημα συντεταγμένων.

Αρχίζουμε διατυπώνοντας την μορφή του R_x

$$R_x = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \theta_x & -\sin \theta_x \\ 0 & \sin \theta_x & \cos \theta_x \end{bmatrix}$$

Το \mathbf{C}^1 μετασχηματίζεται

$$\begin{aligned} K_2 R^1 \mathbf{C}^1 &= -\mathbf{a} \iff \\ R^1 \mathbf{C}^1 &= -K_2^{-1} \mathbf{a} \iff \\ R_{perm} R^1 \mathbf{C}^1 &= -R_{perm} (-K_2^{-1}) \mathbf{a} = \begin{pmatrix} -a_3 \\ -f_2 a_2 \\ f_2 a_1 \end{pmatrix} \end{aligned}$$

με χρήση του R_x έχουμε:

$$\begin{aligned} R_x R_{perm} R^1 \mathbf{C}^1 &= R_x \begin{pmatrix} -a_3 \\ -f_2 a_2 \\ f_2 a_1 \end{pmatrix} \\ &= \begin{pmatrix} x_1 \\ x_2 \\ -f_2 a_2 \sin \theta_x + f_2 a_1 \cos \theta_x \end{pmatrix} \end{aligned}$$

Προκειμένου να ικανοποιήσουμε την επιθυμητή συνθήκη ($x_3 = 0$)

$$\begin{aligned} -f_2 a_2 \sin \theta_x + f_2 a_1 \cos \theta_x &= 0 \iff \\ \tan \theta_x &= a_1/a_2 \end{aligned}$$

Τώρα, πρέπει να δείξουμε ότι η προηγούμενη επιλογή του R_x τοποθετεί στο z -επίπεδο της εικονικής κάμερας και το $\mathbf{v}_{\mathbf{m}2}^2$. Με τον τρόπο αυτό ολοκληρώνουμε παράλληλα και την Απόδειξη του Θεωρήματος 2. Έχουμε έτσι το επόμενο Λήμμα:

Λήμμα 11. *Θεωρούμε τον συμβολισμό και τις υποθέσεις του Αποτελέσματος 14 και του Λήμματος 2. Τότε ισχύει:*

$$R_x R_{perm} R^1 \mathbf{v}_{\mathbf{m}2}^2 = (x_1 \quad x_2 \quad 0)^T \quad (5.62)$$

Απόδειξη Λήμματος 11. Έστω

$$\mathbf{m}_j^i : j\text{-γραμμή πίνακα } R^i$$

Έχουμε

$$\begin{aligned} R_{perm} R^1 \mathbf{v}_{\mathbf{m}2}^2 &= R_{perm} R^1 (-\mathbf{m}_3^{2T}) \\ &= \left(\mathbf{m}_3^1 (-\mathbf{m}_3^{2T}) \quad \mathbf{m}_2^1 (-\mathbf{m}_3^{2T}) \quad -\mathbf{m}_1^1 (-\mathbf{m}_3^{2T}) \right)^T \\ &= (x_1 \quad x_2 \quad x_3)^T \end{aligned}$$

όπου στις προηγούμενες σχέσεις χρησιμοποιήσαμε ότι $\det P^2 < 0$. Αρκεί τώρα να δείξουμε ότι

$$\begin{aligned} R_x(x_1 \ x_2 \ x_3)^T &= (x'_1 \ x'_2 \ 0) \iff \\ \sin \theta_x \mathbf{m}_2^1 \mathbf{m}_3^{2T} &= \cos \theta_x \mathbf{m}_1^1 \mathbf{m}_3^{2T} \iff \\ a_1(\mathbf{m}_2^2 + f_2^{-1} a_2 \mathbf{n}^T) \mathbf{m}_3^{2T} &= a_2(\mathbf{m}_1^2 + f_2^{-1} a_1 \mathbf{n}^T) \mathbf{m}_3^{2T} \iff \\ a_1 f_2^{-1} a_2 \mathbf{n}^T \mathbf{m}_3^{2T} &= a_2 f_2^{-1} a_1 \mathbf{n}^T \mathbf{m}_3^{2T}, \text{ που ισχύει} \end{aligned}$$

με τις τελευταίες δύο συνεπαγωγές να προκύπτουν από το Λήμμα 2, την διαγώνια μορφή του K_2 , την ισότητα των \mathbf{a}_i , και την ορθογωνιότητα των διανυσμάτων γραμμών ενός πίνακα στροφής. \square

Χρησιμοποιώντας τώρα το Λήμμα που μόλις αποδείξαμε, είναι προφανές ότι τα $\mathbf{v}_{\mathbf{m}2}^1, \mathbf{v}_{\mathbf{m}2}^2, \mathbf{C}^1, \mathbf{C}^2$ ανήκουν στο επίπεδο $z = 0$ του νέου, μετασχηματισμένου Συστήματος συντεταγμένων, δηλαδή στο πρωτεύον επίπεδο της "εικονικής" κάμερας. Στο επίπεδο αυτό, τα διανύσματα $\mathbf{v}_{\mathbf{m}2}^1, \mathbf{v}_{\mathbf{m}2}^2$ σχηματίζουν γωνία 2ϕ . Έχουμε λοιπόν:

$$\begin{aligned} \mathbf{v}_{\mathbf{m}2}^{2n} &= R_z(2\phi) \mathbf{v}_{\mathbf{m}2}^{1n} \iff & (5.63) \\ \mathbf{v}_{\mathbf{m}2}^{2n} &= \begin{bmatrix} \cos 2\phi & -\sin 2\phi & 0 \\ \sin 2\phi & \cos 2\phi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \iff \\ \mathbf{v}_{\mathbf{m}2}^{2n} &= \begin{pmatrix} \cos 2\phi \\ \sin 2\phi \\ 0 \end{pmatrix} \end{aligned}$$

όπου συμβολίσαμε με $\mathbf{v}_{\mathbf{m}2}^{1n}$ τα διανύσματα στο "νέο" -μετασχηματισμένο- σύστημα συντεταγμένων και χρησιμοποιήσαμε την ιδιότητα ότι στο νέο σύστημα συντεταγμένων έχουμε ευθυγραμμίσει τον άξονα x με το διάνυσμα $\mathbf{v}_{\mathbf{m}2}^1$. Για να εισάγουμε τώρα τις γωνίες θ_1, θ_2 στις εξισώσεις μας, έχουμε για την παράσταση του $\mathbf{z} \equiv (0 \ 0 \ 1)^T$, στο νέο σύστημα συντεταγμένων:

$$\mathbf{z}^n = R_x R_{perm} R^1 \mathbf{z} = \begin{pmatrix} R_{33}^1 \\ R_{23}^1 \cos \theta_x + R_{13}^1 \sin \theta_x \\ z_3 \end{pmatrix}$$

Υπολογίζουμε τώρα την γωνία

$$\angle \mathbf{v}_{\mathbf{m}2}^2, \mathbf{z} = \theta_2$$

Την γωνία θ_2 την υπολογίζουμε στο νέο σύστημα συντεταγμένων:

$$\mathbf{v}_{\mathbf{m}2}^{2nT} \cdot \mathbf{z}^n = 1 \cdot 1 \cdot \cos \theta_2 = \cos 2\phi R_{33}^1 + (R_{23}^1 \cos \theta_x + R_{13}^1 \sin \theta_x) \sin 2\phi \quad (5.64)$$

Στην Σχέση 5.64, τα διανύσματα έχουν μοναδιαίο μέτρο, αφού οι μετασχηματισμοί στροφής δεν αλλάζουν το μέτρο ενός διανύσματος και για το $\mathbf{v}_{\mathbf{m}2}^2$ ισχύουν όσα επικαλεστήκαμε για παράδειγμα στην Απόδειξη του Θεωρήματος 2.

Για την γωνία θ_1 , χρησιμοποιούμε το αρχικό σύστημα συντεταγμένων και, αντίστοιχα, έχουμε:

$$\mathbf{v}_{\mathbf{m}2}^{1T} \cdot \mathbf{z} = 1 \cdot 1 \cdot \cos \theta_1 = R_{33} \quad (5.65)$$

Αντικαθιστώντας τώρα την Σχέση 5.65 στην Σχέση 5.64, προκύπτει η ζητούμενη Εξίσωση 5.59.

Ολοκληρώνουμε την Απόδειξη, με την παραγωγή της ισοδύναμης μορφής 5.60. Από την Σχέση 5.65 και επειδή ο πίνακας R^1 ως πίνακας στροφής έχει διανύσματα στήλες μοναδιαίου μέτρου, έχουμε για την τρίτη στήλη του R^1 :

$$\begin{aligned} R_{13}^{1\ 2} + R_{23}^{1\ 2} + R_{33}^{1\ 2} &= 1^2 \iff \\ R_{13}^{1\ 2} + R_{23}^{1\ 2} &= 1 - (\cos \theta_1)^2 \iff \\ \left\| \begin{pmatrix} R_{13}^1 \\ R_{23}^1 \end{pmatrix} \right\| &= |\sin \theta_1| \end{aligned} \quad (5.66)$$

Από την Σχέση 5.59 και την Σχέση 5.66 προκύπτει η Σχέση 5.60. \square

Θα κάνουμε τώρα κάποιες επισημάνσεις, τις περισσότερες από τις οποίες ήδη αναφέραμε, προσπαθώντας να τις επεξηγήσουμε και να τις συγκεντρώσουμε:

- Η γωνία 2ϕ δεν είναι η μικρότερη γωνία των διανυσμάτων $\mathbf{v}_{m2}^1, \mathbf{v}_{m2}^2$. Αυτό προκύπτει από την Σχέση 5.63, που επιβάλλει την μέτρηση των γωνιών κατά την ανθρωπολογιακή φορά (κοιτώντας το επίπεδο με το κάθετο διάνυσμα προσανατολισμένο προς τον αναγνώστη).
- η κατασκευή της απόδειξης, προχωράει σε τρία στάδια
 1. Ο πίνακας R^1 ευθυγραμμίζει τον άξονα z με το διάνυσμα \mathbf{v}_{m2}^1 .
 2. Ο πίνακας R_{perm} εναλλάσσει απλώς τους άξονες x, z .
 3. Ο R_x στρέφει τους άξονες z, y , και κατά συνέπεια τα επίπεδα που ορίζουν, ώστε να βρεθεί το \mathbf{C}^1 στη σωστή θέση στον χώρο, ειδικότερα να ταυτιστεί το επίπεδο $z = 0$ με το επίπεδο που ορίζουν τα $\mathbf{v}_{m2}^{1,2}, \mathbf{C}^1$.

Στην συνέχεια, αναλύουμε τις Σχέσεις 5.59, 5.60. Για να είναι ο όρος A της Εξίσωσης 5.59 σημαντικός, παρατηρούμε από την Εξίσωση 5.60 ότι πρέπει:

- Οι οπτικοί άξονες της πραγματικής κάμερας 2 και της κάμερας 1 (θ_1) να σχηματίζουν μεγάλη γωνία, ώστε ο όρος $\sin \theta_1$ να μην είναι μικρός.
- Η προβολή της διεύθυνσης της μετατόπισης μεταξύ των καμερών, στο επίπεδο της εικόνας 2, να σχηματίζει μικρή γωνία με την προβολή του οπτικού άξονα της πρώτης κάμερας, στο επίπεδο της εικόνας 2.

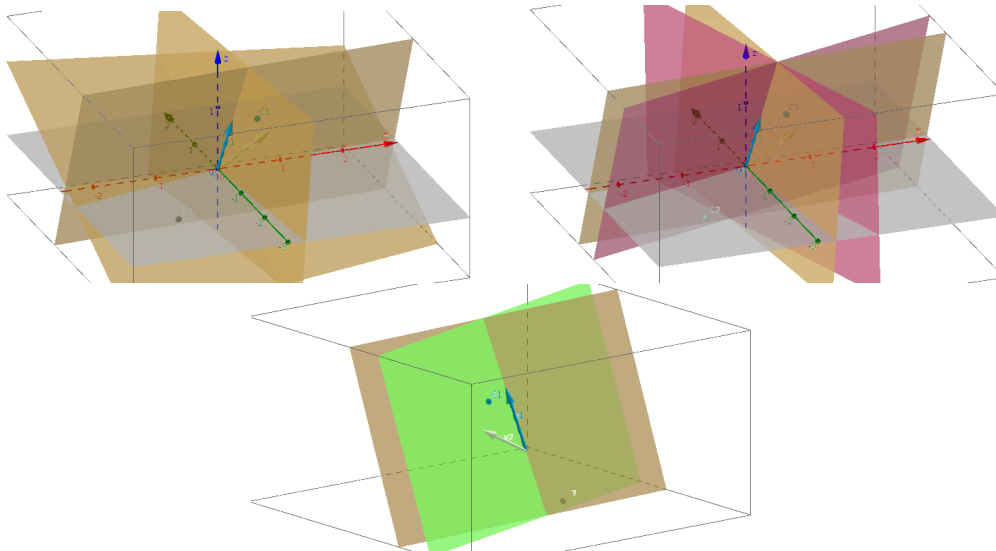
Δηλαδή, ο όρος A είναι σημαντικός αν έχουμε μεγάλη γωνία μεταξύ των οπτικών αξόνων (των δύο καμερών) και η διεύθυνση μετατόπισης είναι "κοντά" στον οπτικό άξονα της κάμερας 1.

Επίσης, από τις Εξισώσεις 5.64, 5.65, παρατηρούμε ότι μπορούμε να επιλέξουμε την γωνία θ_1 στο διάστημα $[0, \pi]$, αφού αυτό καλύπτει το σύνολο τιμών του $\cos \theta_1$. Αυτό μας επιτρέπει να κάνουμε στη Εξίσωση 5.60 την αντικατάσταση:

$$|\sin \theta_1| \leftrightarrow \sin \theta_1$$

αφού το ημίτονο διατηρεί πρόσημο στο διάστημα $[0, \pi]$.

Ανάλογα, εφόσον η θ_2 εμφανίζεται στις εξισώσεις μόνο μέσω του συνημιτόνου

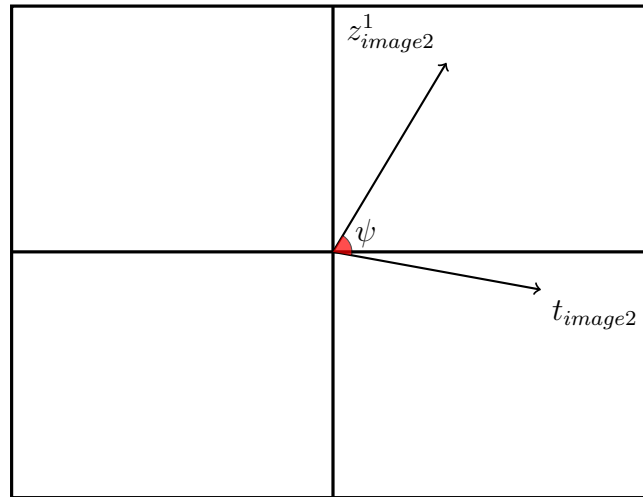


Σχήμα 5.3: Παρουσιάζουμε την κατασκευή της Απόδειξης του Αποτελέσματος 14. Στο γράφημα, απεικονίζεται το κέντρο προβολής (C^1, C^2) και η κατεύθυνση προσανατολισμού (v_1, v_2), για κάθε μία από τις λύσεις 5.22. Χρησιμοποιούμε το μπλε χρώμα για την απεικόνιση των στοιχείων της λύσεως 1 και το λευκό χρώμα αντίστοιχα για την λύση 2. Επίσης, ορατοί είναι και οι άξονες του αρχικού συστήματος συντεταγμένων. Για την παραγωγή του γραφήματος, αρχικοποιήσαμε τις κάμερες 1, 2 σε τυχαία διάταξη στον χώρο, επιλύσαμε το σύστημα που έχουμε διατυπώσει και απεικονίσαμε τα δεδομένα των λύσεων 5.22 που προέκυψαν. Έχουμε:

Κορυφή Αριστερά: Τα τρία επίπεδα του συστήματος συντεταγμένων, όπως αυτά προκύπτουν με την εφαρμογή του περιστροφικού μετασχηματισμού R_1 . Σημειώνουμε ότι παραλείπουμε την εφαρμογή του R_{perm}

Κορυφή Δεξιά: Η προηγούμενη (καφέ), και η προκύπτουσα, θέση (κόκκινη) των επιπέδων που επηρεάζονται από τον περιστροφικό μετασχηματισμό R_x . Μόνο δύο επίπεδα απεικονίζονται, αφού η περιστροφή γίνεται ως προς έναν από τους 3 άξονες του συστήματος συντεταγμένων, με αποτέλεσμα το αντίστοιχο επίπεδο του συστήματος συντεταγμένων να μην μεταβάλλεται.

Κάτω: Λεπτομέρεια της προηγούμενης Εικόνας (Κέντρο). Βλέπουμε πώς η περιστροφή R_x μετακινεί ένα από τα επίπεδα του συστήματος συντεταγμένων (καφέ) σε κατάλληλη θέση (πράσινο), ώστε αυτό να ταυτιστεί με το επίπεδο στο οποίο ανήκουν τα C_1, C_2, v_1, v_2 .



Σχήμα 5.4: Στο γράφημα, παρουσιάζουμε το επίπεδο της εικόνας 2. Σε αυτό το επίπεδο, έχουμε προβάλλει τον οπτικό άξονα της κάμερας 1 (z_{image2}^1). Το διάνυσμα t_{image2} , είναι μοναδιαίο διάνυσμα, ομόρροπο με την προβολή του διανύσματος μετατόπισης t μεταξύ των καμερών 1, 2 στο επίπεδο της εικόνας 2. Το εσωτερικό γινόμενο των διανυσμάτων αυτών δίνει τον όρο A της Εξίσωσης 5.59.

της, μπορούμε να θεωρήσουμε ότι βρίσκεται στο διάστημα $[0, \pi]$. Στην συνέχεια, κάνουμε τον ορισμό

$$\Psi \equiv \cos \psi \quad (5.67)$$

Η Εξίσωση 5.60 δίνει με την προηγούμενη αντικατάσταση και τον Ορισμό 5.67:

$$\cos \theta_2 = \cos 2\phi \cos \theta_1 + \Psi \sin \theta_1 \sin 2\phi \quad (5.68)$$

Η προηγούμενη Σχέση 5.68 μας επιτρέπει να απαλείψουμε τον παράγοντα Ψ , ο οποίος έχει μεν φυσική ερμηνεία, δεν σχετίζεται όμως το ίδιο άμεσα με τα μεγέθη ϕ, θ_1, θ_2 .

Αρχικά, παρατηρούμε ότι, από τον ορισμό του

$$\Psi \in [-1, 1]$$

Επίσης, παρατηρούμε ότι η Εξίσωση 5.68 είναι γραμμική ως προς Ψ . Συνεπώς, η Εξίσωση 5.68 μεταβάλλεται επίσης μονότονα ως προς Ψ , με την ακριβή μορφή της μονοτονίας να εξαρτάται από την τιμή της 2ϕ και να είναι προς το παρόν αδιάφορη.

Επίσης, αναφέρουμε ότι η συνάρτηση τόξου συνημιτόνου \cos^{-1} , είναι γνησίως φθίνουσα συνάρτηση. Από τις προηγούμενες παρατηρήσεις, μπορούμε να βρούμε το σύνολο τιμών της γωνίας θ_2 , για δεδομένα $\theta_1, 2\phi$ και Ψ μεταβλητό στο σύνολο τιμών του.

Συγκεκριμένα, χρησιμοποιούμε την ιδιότητα κάθε συνεχούς συνάρτησης να παίρνει όλες τις τιμές ανάμεσα από δύο γνωστές τιμές της, και την μονοτονία των συναρτήσεων για να βρούμε τις δύο ακραίες (ελάχιστη και μέγιστη) τιμές της συνάρτησης \cos^{-1} με όρισμα που δίνεται από την γραμμική συνάρτηση του Ψ που διατυπώσαμε στην Σχέση 5.68.

Σύμφωνα με όσα αναφέραμε και με χρήση των γνωστών τριγωνομετρικών ταυτοτήτων αθροίσματος γωνιών, έχουμε:

- Για $\Psi = 1$, η Εξίσωση 5.68 δίνει:

$$\begin{aligned}\cos \theta_2 &= \cos 2\phi \cos \theta_1 + \sin \theta_1 \sin 2\phi \\ &= \cos (\theta_1 - 2\phi)\end{aligned}\quad (5.69)$$

- Για $\Psi = -1$, η Εξίσωση 5.68 δίνει:

$$\begin{aligned}\cos \theta_2 &= \cos 2\phi \cos \theta_1 - \sin \theta_1 \sin 2\phi \\ &= \cos (\theta_1 + 2\phi)\end{aligned}\quad (5.70)$$

Μπορούμε τώρα να βρούμε το σύνολο τιμών της διαφοράς

$$\theta_2 - \theta_1$$

Από τις Σχέσεις 5.69, 5.70 και την προηγούμενη συζήτηση, έχουμε:

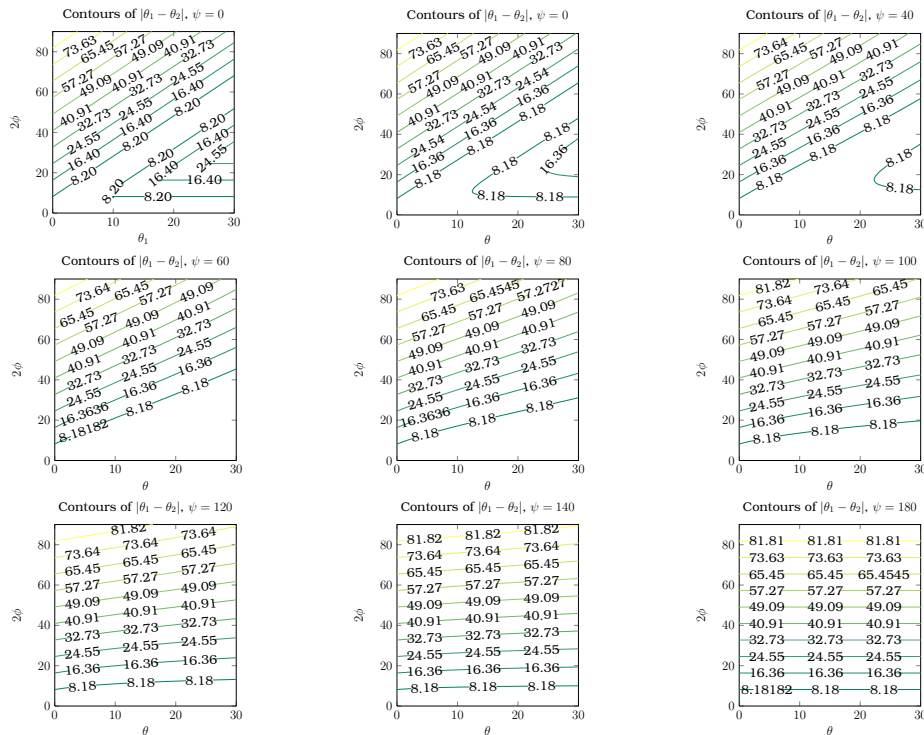
$$\theta_2 - \theta_1 \in [-2\phi, 2\phi] \quad (5.71)$$

Βρισκόμαστε τώρα σε Θέση να διατυπώσουμε το επόμενο Πόρισμα, το οποίο συνδέει την διαφορά των γωνιών θ_1, θ_2 με την γωνία 2ϕ . Μας επιτρέπει δηλαδή να συνδέσουμε τον διαχωρισμό των κατευθύνσεων προσανατολισμού των λύσεων 5.22 με το πόσο συμβατές είναι με την κάμερα 1 (γωνίες θ_1, θ_2).

Πόρισμα 3. Η διαφορά των γωνιών θ_1, θ_2 που σχηματίζουν οι οπτικοί άξονες των δύο λύσεων 5.22 με τον οπτικό άξονα της κάμερας 1, φράσσεται από την γωνία που σχηματίζουν οι κατευθύνσεις προσανατολισμού των δύο λύσεων. Συγκεκριμένα ισχύει η σχέση

$$|\theta_2 - \theta_1| \leq 2\phi \quad (5.72)$$

Το Πόρισμα που μόλις διατυπώσαμε, δεν δίνει ουσιαστικό περιορισμό αν η γωνία 2ϕ είναι η εξωτερική (μεγάλη) γωνία που σχηματίζουν τα διανύσματα $\mathbf{v}_{m2}^1, \mathbf{v}_{m2}^2$. Σε αυτή την περίπτωση, μπορούμε να αντιστρέψουμε τον ρόλο των ανακατασκευών P_{m2}^1, P_{m2}^2 και να καταλήξουμε σε ένα ουσιαστικό φράγμα μέσω της Εξίσωσης 5.72.



Σχήμα 5.5: Ισοσταθμικές της $|\theta_2 - \theta_1|$ για διάφορες τιμές της γωνίας ψ και ενδεικτικό εύρος γωνιών $\theta_1, 2\phi$.

5.2 Πρακτικοί προβληματισμοί

5.2.1 Επίλυση εξισώσεων

Επιστρέφουμε τώρα στο Σύστημα 5.10, για να εξετάσουμε πως μπορούμε να το επιλύσουμε στην πράξη. Για τον σκοπό αυτό, κάνουμε τώρα μια πιο προσεκτική ανάλυση των σχέσεων που οδηγούν στην ομογραφία H που συνδέει την προβολική P_{P1}, P_{P2} με την πραγματική, ευκλείδεια ανακατασκευή P_{GT1}, P_{GT2} .

Για να απλοποιήσουμε τους υπολογισμούς μας, εφαρμόζουμε την τρισδιάστατη ομογραφία H_k στην ευκλείδεια ανακατασκευή, ώστε να την φέρουμε σε κανονική μορφή, δηλαδή $P'_{GT1} = [I \ 0]$. Η ομογραφία αυτή έχει την απλή μορφή

$$\begin{bmatrix} K_1^{-1} & \mathbf{0} \\ \mathbf{0}^T & 1 \end{bmatrix}$$

Για την συνέχεια, χρειαζόμαστε δύο Αποτελέσματα (Αποτέλεσμα 9.8 και Λήμμα 9.11 από [90]), τα οποία θα παραθέσουμε χωρίς Απόδειξη, για συνοπτικότητα.

Αποτέλεσμα 15. Αν συμβολίσουμε με H προβολικό μετασχηματισμό του τρισδιάστατου χώρου (Ομογραφία), τότε οι Θεμελιώδεις πίνακες που αντιστοιχούν στα ζευγάρια $\{P, P'\}, \{PH, P'H\}$ είναι ίδιοι.

Αποτέλεσμα 16. Έστω ότι η rank2 μήτρα F μπορεί να γραφεί με δύο διαφο-

ρευτικούς τρόπους

$$F = [\mathbf{a}]_x A$$

$$F = [\hat{\mathbf{a}}]_x \hat{A}$$

Τότε ισχύουν:

$$\hat{\mathbf{a}} = \kappa \mathbf{a}$$

$$\hat{A} = \kappa^{-1}(A + \mathbf{a}\mathbf{v}^T)$$

Πριν συνεχίσουμε, συμβολίζουμε:

$$\bullet \{P'_{GT1}, P'_{GT2}\} \equiv \{[I \ \mathbf{0}], [A \ \mathbf{a}]\}$$

$$\bullet \{P_{P1}, P_{P2}\} \equiv \{[I \ \mathbf{0}], [\hat{A} \ \hat{\mathbf{a}}]\}$$

Για να βρούμε την μορφή του H , έχουμε: Αφού οι δύο ανακατασκευές συνδέονται με προβολικό μετασχηματισμό H , θα έχουν σύμφωνα με τα προηγούμενα κοινό θεμελιώδη πίνακα F . Σύμφωνα με τον προηγούμενο συμβολισμό μας, γράφουμε:

$$F_P = [[\mathbf{a}]_x A]$$

$$F_{GT} = [[\hat{\mathbf{a}}]_x \hat{A}]$$

Από το Αποτέλεσμα 15, έχουμε ότι οι Θεμελιώδεις πίνακες πρέπει να είναι κοινοί, επειδή είναι όμως ομογενείς ποσότητες, οι θεμελιώδεις πίνακες θα συνδέονται με μια πολλαπλασιαστική σταθερά, δηλαδή:

$$F_P = \epsilon F_{GT}$$

Χρησιμοποιούμε τώρα το Αποτέλεσμα 16 και έχουμε

$$\hat{\mathbf{a}} = \kappa \mathbf{a}$$

$$\hat{A} = \epsilon^{-1} \kappa^{-1}(A + \mathbf{a}\mathbf{v}^T)$$

Μπορεί τώρα να επιβεβαιωθεί, ότι η ομογραφία που συνδέει τις δύο ανακατασκευές, έχει την μορφή

$$H' \equiv \begin{bmatrix} \kappa^{-1} \epsilon^{-1} I & \mathbf{0} \\ \kappa^{-1} \epsilon^{-1} \mathbf{v}^T & \kappa \end{bmatrix}$$

Πράγματι,

$$\begin{aligned} \kappa^{-1} \epsilon^{-1} P'_{GT1} &= P_{P1} H' \\ P'_{GT2} &= P_{P1} H' \end{aligned} \quad (5.73)$$

συνθέτοντας τώρα την ομογραφία H' με την ομογραφία H_k^{-1} , παίρνουμε την μορφή της ομογραφίας που συνδέει τις δύο αρχικές ανακατασκευές, συγκεκριμένα

$$\begin{bmatrix} \kappa^{-1} \epsilon^{-1} K_1 & \mathbf{0} \\ \kappa^{-1} \epsilon^{-1} \mathbf{v}^T K_1 & \kappa \end{bmatrix}$$

Παρατηρούμε τώρα, ότι δεν ενδιαφερόμαστε να ανακτήσουμε την πραγματική κλίμακα της ανακατασκευής, επομένως μπορούμε να απλοποιήσουμε την ομογραφία, καταλήγοντας στην τελική της μορφή

$$\begin{bmatrix} \kappa^{-1} \epsilon^{-1} K_1 & \mathbf{0} \\ \kappa^{-1} \epsilon^{-1} \mathbf{v}^T K_1 & 1 \end{bmatrix} \quad (5.74)$$

Σημειώνουμε επίσης, ότι αν απαλείψουμε τους όρους $\kappa^{-1}\epsilon^{-1}$ από την Σχέση 5.74, παίρνουμε την ομογραφία της Σχέσης 5.6. Χρησιμοποιώντας αυτή την ομογραφία, καταλήγουμε στους επόμενους πίνακες για την μετρική ανακατασκευή

$$P_{M1} = [K_1 \quad \mathbf{0}]$$

$$P_{M2} = [\mu K_2 R_2 \quad \mathbf{a}]$$

όπου μ κατάλληλη σταθερά. Αν και προφανές, τονίζουμε ξανά ότι το διάνυσμα \mathbf{a} έχει την σωστή διεύθυνση, όχι όμως και το σωστό μέτρο σε σχέση με την "ground truth" ανακατασκευή.

Αντικαθιστώντας τώρα την ομογραφία 5.74 στις εξισώσεις υπολογισμού της ω_2^* (Σχέση 5.3), εξάγουμε τις "διορθωμένες" μορφές των Εξισώσεων 5.10. Παραθέτουμε ακόμα την μορφή των εξισώσεων αυτών εδώ, για καλύτερη σύνδεση με την ανάλυση που επιχειρούμε

$$\omega_2^* = P_{P2} \begin{bmatrix} (\kappa\epsilon)^{-2} K_1 K_1^T & -(\kappa\epsilon)^{-2} K_1 K_1^T \mathbf{p} \\ -(\kappa\epsilon)^{-2} \mathbf{p}^T K_1 K_1^T & (\kappa\epsilon)^{-2} \mathbf{p}^T K_1 K_1^T \mathbf{p} \end{bmatrix} P_{P2}^T$$

Πρέπει συμπληρωματικά να τονιστεί, ότι όπως είναι σαφές από τις Σχέσεις 5.73, η εικόνα ω_2^* βρίσκεται στην κανονική της θέση, εφόσον η μήτρα $K_2 R_2$ προσδιορίζεται ακριβώς και όχι πολλαπλασιασμένη με κάποια σταθερά.

Ας κάνουμε τώρα σαφές το πλαίσιο της συζήτησης που θα ακολουθήσει. Οι όροι των εξισώσεων, αναφέρονται στην ομαδοποίηση όπως αυτή εμφανίζεται στο Σύστημα 5.10. Υπενθυμίζουμε ακόμα ότι κάθε εξίσωση αυτού του συστήματος έχει προκύψει από τα στοιχεία της ω_2^* . Έχουμε:

- Οι ομογενείς εξισώσεις, δεν επηρεάζονται από την πολλαπλασιαστική σταθερά $(\kappa\epsilon)^{-1}$. Αυτό γιατί η πολλαπλασιαστική σταθερά πολλαπλασιάζει όλους τους όρους της εξίσωσης, και συνεπώς μπορεί να απαλειφθεί. Οι εξισώσεις αυτές αντιστοιχούν στα στοιχεία εκτός διαγωνίου της ω_2^* .
- Στις δύο εξισώσεις στις οποίες εμφανίζεται ο όρος f_2^2 , θα πρέπει να απορροφήσουμε την πολλαπλασιαστική σταθερά στον ορισμό της μεταβλητής της εξίσωσης, προκειμένου να μπορούμε πάλι να απαλείψουμε την σταθερά-έστω c - που τις πολλαπλασιάζει. Για να αποσαφηνίσουμε τα λεγόμενα μας, κάνουμε την ακόλουθη αντικατάσταση αγνώστων:

$$f_2^2 \rightarrow c^{-1} f_2^2 \quad (5.75)$$

Με αυτή την αντικατάσταση, μπορούμε "απαλείφοντας" την κοινή σταθερά c από τις εξισώσεις, να καταλήξουμε στην μορφή τους ακριβώς όπως αυτή εμφανίζεται στο Σύστημα 5.10. Βέβαια, από την λύση των εξισώσεων αυτών δεν θα ανακτήσουμε το f_2^2 , αλλά την κατά μια πολλαπλασιαστική σταθερά προσέγγιση του $c^{-1} f_2^2$

- Η εξίσωση που προκύπτει από το στοιχείο (3, 3) της ω_2^* , δεν μπορεί να αξιοποιηθεί. Αυτό γιατί, αν και όλοι οι όροι της εξίσωσης πολλαπλασιάζονται με την σταθερά, προκύπτει το ακόλουθο πρόβλημα με τον "σταθερό όρο" της εξίσωσης αυτής

$$-p_{33}^2 + 1 \rightarrow -cp_{33}^2 + 1$$

Η εξίσωση αυτή μας αναγκάζει να εισάγουμε έναν άγνωστο ακόμα (την πολλαπλασιαστική σταθερά), και επομένως δεν είναι αξιοποιήσιμη χωρίς να αλλάξουμε τον αριθμό των αγνώστων.

- Η τελευταία αυτή παρατήρηση μας επιτρέπει να ξεπεράσουμε τους προβληματισμούς που θέσαμε νωρίτερα, συγκεκριμένα αν είναι το Σύστημα 5.10 των 6 εξισώσεων συμβιβαστό, αφού τελικά μόνο 5 από αυτές μπορούν να αξιοποιηθούν.

Συμπερασματικά, προτείνουμε τον ακόλουθο τρόπο επίλυσης

1. Λύνουμε το σύστημα που προκύπτει από τις Εξισώσεις 5.10, με τους συντελεστές όπως ακριβώς εμφανίζονται. Τονίζουμε ότι κρατάμε 5 από τις 6 Εξισώσεις του 5.10, απορρίπτοντας την ακατάλληλη εξίσωση που προκύπτει από το στοιχείο $\omega_2^*(3, 3)$.
2. Με αυτό τον τρόπο η σωστή λύση ανακτάται, με εξαίρεση το f_2 που έχει προσδιοριστεί με ακρίβεια μιας πολλαπλασιαστικής σταθεράς.
3. Διάφοροι τρόποι μπορούν να χρησιμοποιηθούν για τον πλήρη προσδιορισμό του σωστού f_2 . Η λύση που προτείνουμε είναι η ακόλουθη. Αντιστρέφουμε τον ρόλο των καμερών 1, 2, δηλαδή τοποθετούμε την κάμερα 2 στην αρχή των αξόνων. Σχηματίζοντας και επιλύοντας για αυτή την αντεστραμμένη διάταξη καμερών το Σύστημα 5.10, μπορούμε αντίστοιχα να ανακτήσουμε την σωστή τιμή του f_2 . Με αυτόν τον τρόπο προσδιορίζουμε τόσο το σύμπλεγμα $κε$, όσο και το f_2 .
4. Σημειώνουμε, ότι η προηγούμενη αντιστροφή του ρόλου των καμερών 1, 2, γίνεται ιδιαίτερα εύκολα, με την απλή αντιστροφή του πίνακα F που αντιστοιχεί στο αρχικό ζευγάρι.

Καταληκτικά, επιλύουμε τις (σωστές 5) Εξισώσεις του Συστήματος 5.13 και επίσης προσδιορίζουμε την ορθή τιμή του f_2 . Ανάλογα με την Ομογραφία που θα χρησιμοποιήσουμε (Σχέση 5.6, 5.74), κάποιος από τους δύο πίνακες της μετρικής ανακατασκευής P_{m1}, P_{m2} , θα προκύπτει με το αριστερό 3×3 μπλοκ (δηλαδή το $K_i R_i$) πολλαπλασιασμένο κατά μία σταθερά. Το γεγονός αυτό δεν επηρεάζει την ορθότητα της αναπαράστασης.

5.2.2 Η περίπτωση του Θορύβου

Μέχρι τώρα, δεν εξετάσαμε καθόλου πως επηρεάζει ο Θόρυβος την επίλυση του Συστήματος 5.13. Σε μια προσπάθεια να μοντελοποιήσουμε την επίδραση του Θορύβου, γράφουμε:

$$P_{P2} = P_{P2}^{GT} + P_{P2}^n \quad (5.76)$$

Αναλύουμε δηλαδή την έκφραση της μήτρας προβολής 2 που έχουμε εκτιμήσει, σε άθροισμα δύο όρων, ενός πραγματικού όρου P_{P2}^{GT} και ενός όρου θορύβου P_{P2}^n , που θεωρήσαμε έως τώρα μηδενικό.

Θεωρούμε όπως συνήθως, ότι η προβολική αναπαράσταση είναι στην "κανονική" μορφή, συνεπώς η κάμερα 1 δεν επηρεάζεται από τον θόρυβο αφού τίθεται στη τιμή $[I \ 0]$.

Αν αντικαταστήσουμε την Θορυβώδη μορφή της μήτρας P_{P2} στο Σύστημα 5.10, βλέπουμε ότι λαμβάνουμε αντίστοιχα το ιδανικό σύστημα προσαυξημένο κατά

μια διαταραχή, η οποία προκύπτει από την $P_{P_2}^n$.

Από την επίλυση του θορυβώδους συστήματος επομένως, λαμβάνουμε μια θορυβώδη εκτίμηση των 5 υπό προσδιορισμό παραμέτρων. Κατά κάποιο τρόπο, βλέπουμε ότι προσαρμοζόμαστε στον θόρυβο, αφού το σύστημα δεν είναι υπερορισμένο. Όμως ο θόρυβος δεν εφαρμόζεται απευθείας στο Σύστημα 5.10, αλλά στα σημεία των υπό εξέταση εικόνων. Οπότε ο τρόπος που επηρεάζει ο θόρυβος τις προς επίλυση εξισώσεις δεν είναι ξεκάθαρος.

Παρατηρώντας την μορφή του συστήματος που προκύπτει μετά την αντικατάσταση του 5.76 στο 5.10, βλέπουμε ότι υπάρχουν κάποιοι κοινοί όροι θορύβου στις εξισώσεις. Διάφοροι τρόποι δοκιμάστηκαν προκειμένου να αξιοποιηθεί το γεγονός αυτό, με φτωχά όμως αποτελέσματα.

Για να ελαττώσουμε την επίδραση του θορύβου στις εκτιμήσεις των αγνώστων παραμέτρων, είναι σημαντικό να κατανοήσουμε τον τρόπο που επηρεάζει τις Εξισώσεις. Οι θορυβώδεις μετρήσεις δημιουργούν σφάλμα στην εκτίμηση του Θεμελιώδους πίνακα F και η χρήση του εσφαλμένου F τελικά μεταφέρει κάποιο σφάλμα στην μήτρα P_{P_2} . Μια επιτυχημένη ανάλυση, θα μας επέτρεπε να εκτιμήσουμε, δεδομένης της P_{P_2} , την βέλτιστη σύμφωνα με κάποιο κριτήριο αποθορυβοποίησή της, P'_{P_2} . Επιπλέον, χρήσιμα συμπεράσματα θα μπορούσαν να προκύψουν από τον προσδιορισμό της συνάρτησης κατανομής του θορύβου στα στοιχεία της μήτρας P_{P_2} , δεδομένης της κατανομής του θορύβου των μετρήσεων. Τέλος, η χρήση πολλών εκτιμήσεων της παραμέτρου του εστιακού μήκους f κάθε κάμερας, από διαφορετικά ζευγάρια καμερών στα οποία συμμετέχει, μπορεί να βελτιώσει την ακρίβεια των εκτιμήσεων μας και να βοηθήσει στο διαχωρισμό των αξιόπιστων και των αναξιόπιστων ανακατασκευών.

5.3 Πειράματα

5.3.1 Πειράματα με συνθετικά δεδομένα

Δομή συνθετικής σκηνής

Στα πειράματα που εκτελέσαμε με συνθετικά δεδομένα, χρησιμοποιήσαμε την ακόλουθη μεθοδολογία για να τοποθετήσουμε τις κάμερες στον χώρο. Τοποθετήσαμε την μία κάμερα (Κάμερα 1) στην αρχή των αξόνων και προσανατολισμένη κατά τον άξονα z . Για την δεύτερη κάμερα, δειγματοληπούμε ομοιόμορφα επιφάνεια μοναδιαίας σφαίρας, με κέντρο την θέση της κάμερας 1. Έτσι η σχετική θέση των καμερών 1 και 2 επιλέγεται ομοιόμορφα στον χώρο, ανάμεσα σε όλες τις πιθανές σχετικές θέσεις. Η επιλογή της μοναδιαίας ακτίνας της σφαίρας, καθορίζει την παράμετρο της κλίμακας της ανακατασκευής.

Για τον προσανατολισμό της Δεύτερης κάμερας, δημιουργούμε ένα τυχαίο διάνυσμα, και το χρησιμοποιούμε ως άξονα περιστροφής. Οι συντεταγμένες του διανύσματος αυτού είναι ομοιόμορφα κατανεμημένες μεταβλητές στο διάστημα $[0, 1]$. Αντίστοιχα επιλέγουμε ομοιόμορφα και την γωνία στροφής, γύρω από τον προηγούμενο άξονα. Θέτουμε τον περιορισμό οι οπτικοί άξονες των δύο καμερών να σχηματίζουν γωνία μικρότερη των 30° , προκειμένου να δημιουργηθούν πιο ρεαλιστικές σκηνές. Η επιλογή αυτή υποστηρίζεται από την βιβλιογραφία, για παράδειγμα [174]. Τα αποτελέσματα των πειραμάτων δεν φάνηκε πάντως να επηρεάζονται από τον περιορισμό αυτό.

Δείκτες Αξιολόγησης

Για να αξιολογήσουμε τα αποτελέσματα των πειραμάτων, τόσο σε συνθετικά όσο και σε πραγματικά δεδομένα, χρησιμοποιήσαμε τα παρακάτω σφάλματα

1. Για την ποσοτικοποίηση του σφάλματος στην εκτίμηση του εστιακού βάθους, ορίζουμε την ποσότητα

$$\Delta f \equiv \left| \frac{f}{f_{gt}} - 1 \right|$$

όπου f συμβολίζουμε την ευρεθείσα τιμή και f_{gt} την σωστή τιμή του εστιακού μήκους.

2. Για την ποσοτικοποίηση του σφάλματος στην εκτίμηση της θέσης του π_∞ , χρησιμοποιούμε τις επόμενες 2 ποσότητες

- Την γωνία που σχηματίζει το διάνυσμα μετατόπισης \mathbf{t} και το σωστό διάνυσμα μετατόπισης \mathbf{t}_{gt}

$$\Delta t \equiv \angle \mathbf{t}, \mathbf{t}_{gt}$$

- Την γωνία που σχηματίζει η σύνθεση της ευρεθείσας περιστροφής R με την σωστή περιστροφή R_{gt}

$$\Delta R \equiv \angle RR_{gt}^{-1}$$

Παρατηρούμε ότι για την θέση του π_∞ , προσδιορίζουμε έμμεσα το σφάλμα. Αυτό γίνεται γιατί από τον τρόπο που διεξάγουμε τα πειράματα, δεν γνωρίζουμε την "σωστή" θέση του π_∞ στην προβολική ανακατασκευή.

Οι προηγούμενοι δείκτες, είναι συνηθισμένοι στην βιβλιογραφία. Αναφέρουμε ενδεικτικά την χρήση τους στα [74, 119, 32].

Προσθήκη Θορύβου

Όπως είναι φυσικό, τα πειράματα χωρίς θόρυβο στις μετρήσεις επιστρέφουν τις σωστές τιμές των υπό προσδιορισμό παραμέτρων. Για τον λόγο αυτό, δεν μπορούμε να αξιολογήσουμε τις προτεινόμενες μεθόδους, αν δεν προσθέσουμε Θόρυβο στις μετρήσεις.

Λάβαμε ιδιαίτερη μέριμνα, ώστε ο θόρυβος να προστεθεί στο μέγεθος που επηρεάζει στην πραγματικότητα, δηλαδή στις συντεταγμένες των σημείων στην εικόνα. Η προσθήκη θορύβου στους πίνακες προβολής, στις υπό προσδιορισμό παραμέτρους, στα σημεία του χώρου, στον θεμελιώδη πίνακα κ.λ.π. δεν μοντελοποιεί σωστά την διαδικασία σχηματισμού εικόνων, όποτε θα πρέπει να αποφευχθεί.

Για τον σκοπό αυτό, προσθέτουμε στις εικόνες Θόρυβο που ακολουθεί Γκαουσιανή κατανομή. Ο Θόρυβος προστίθεται στα σημεία της εικόνας, με τυπική απόκλιση σχετική με το μέγεθος της εικόνας. Όσον αφορά την κατανομή που επιλέξαμε για τον θόρυβο στις μετρήσεις, ακολουθήσαμε την πιο διαδεδομένη επιλογή. Υπάρχουν όμως αναφορές που διαφωνούν με την χρήση Γκαουσιανού Θορύβου ([87]).

Συγκεκριμένα, δοσομένης της συνθετικής σκηνής, την δημιουργία της οποίας περιγράψαμε προηγουμένως, ακολουθούμε τα εξής βήματα:

1. Δημιουργία στον χώρο τρισδιάστατων σημείων. Τα σημεία δημιουργούνται με χρήση της ομοιόμορφης κατανομής για τις συντεταγμένες τους. Από αυτά επιλέγονται όσα είναι ορατά και από τις δύο κάμερες.
2. Τα σημεία προβάλλονται στις εικόνες, με τους σωστούς πίνακες προβολής που διαθέτουμε.
3. Στα προηγούμενα σημεία των εικόνων, προσθέτουμε Γκαουσιανό Θόρυβο.
4. Από τα ταιριασμένα ζεύγη σημείων στις δύο εικόνες, υπολογίζουμε τον Θεμελιώδη πίνακα.
5. Τέλος, υπολογίζουμε από τον Θεμελιώδη πίνακα το κανονικό ζευγάρι καμερών.

Για τον υπολογισμό του Θεμελιώδους πίνακα, χρησιμοποιούμε την διαθέσιμη υλοποίηση από το computer vision toolbox του matlab. Η υλοποίηση αυτή χρησιμοποιεί την μέθοδο Least Median of Squares (LMedS). Όμως, στην πραγματικότητα είναι μια βασική υλοποίηση, η οποία δεν επιτρέπει τον καθορισμό ικανοποιητικού αριθμού παραμέτρων της μεθόδου. Επιλέχθηκε όμως για τους εξής λόγους:

- Το γεγονός ότι δεν υπάρχουν outliers στα σημεία που έχουμε αντιστοιχήσει στις δύο εικόνες, διευκολύνει τον ακριβή προσδιορισμό του Θεμελιώδους πίνακα
- Δεν θέλουμε να επηρεαστούν τα αποτελέσματα μας από την χρήση κάποιου ιδιαίτερα εξελιγμένου αλγορίθμου προσδιορισμού του Θεμελιώδους πίνακα

Τέλος, για τον θόρυβο σημειώνουμε τα εξής. Διαλέξαμε να προσθέσουμε τον Θόρυβο ποσοστιαία στις συντεταγμένες. Αυτό έχει σαν συνέπεια σημεία με μεγάλες "τιμές" τεταγμένης ή τετμημένης να επηρεάζονται περισσότερο από σημεία με μικρές τιμές. Αυτό δεν είναι επιθυμητό, από τον τρόπο δημιουργίας όμως των σημείων στον χώρο (ομοιόμορφη κατανομή), θεωρούμε ότι εξομαλύνεται μερικώς. Από την άλλη, η προσθήκη Θορύβου με τυπική απόκλιση δοσμένο αριθμό pixels, είναι προβληματική ή και ανεφάρμοστη σε "μικρές" εικόνες ή στο κανονικοποιημένο εύρος εστιακών μηκών που χρησιμοποιούμε. Οπότε θεωρούμε ότι είναι αποδεκτός συμβιβασμός, αν και, όπως παρατηρούμε, σε μεγάλα επίπεδα Θορύβου το μέγεθος της εικόνας μεταβάλλει τα σχετικά αποτελέσματα της μεθόδου.

Παρατηρούμε επίσης, ότι στην βιβλιογραφία ([33, 74, 119, 32]) τα πειραματικά αποτελέσματα αξιολογούνται για επίπεδα θορύβου έως 1 pixel ή 1%, οπότε δίνουμε ιδιαίτερο βάρος σε αυτά τα μεγέθη.

Πειράματα και Αποτελέσματα

Αρχικά, μεταβάλλουμε το επίπεδο του θορύβου, και παρουσιάζουμε τα αποτελέσματα της μεθόδου. Εφαρμόζουμε την μέθοδο που αναπτύξαμε σε

- κανονικοποιημένες διαστάσεις εικόνας, με διαγώνιο περίπου 1. Στην περίπτωση αυτή δειγματοληπτούμε το εστιακό μήκος f στο διάστημα $[0.5, 1.7]$. Έτσι έχουμε μεγάλη διακύμανση στο εύρος τιμών του εστιακού μήκους.

- διαστάσεις εικόνας 700×500 .

Σημειώνουμε ότι ελέγχουμε τα σημεία του χώρου να προβάλλονται μέσα στα όρια της εικόνας.

Η εξέλιξη του σφάλματος σε σχέση με το επίπεδο του θορύβου, παρουσιάζεται στα Διαγράμματα 5.6. Επίσης, στα Διαγράμματα 5.7, παρουσιάζουμε την διασπορά του σφάλματος και στο Διάγραμμα 5.8 την εξέλιξη του ποσοστού επιτυχίας της μεθόδου. Η μέθοδος αποτυγχάνει, όταν οι τιμές που προκύπτουν για τα f_1^2, f_2^2 είναι αρνητικές.

Στα προηγούμενα αποτελέσματα, παρατηρούμε ότι για τιμές θορύβου $< 1\%$, το σφάλμα παραμένει σε ικανοποιητικά χαμηλά επίπεδα. Ακόμα, παρατηρούμε ότι το εύρος τιμών του f δεν επηρεάζει σημαντικά τα αποτελέσματα, ειδικά σε λογικά επίπεδα θορύβου (Σύγκριση κανονικοποιημένου εύρους τιμών f και τιμών για μέγεθος εικόνας 700×500). Η κατανομή του σφάλματος, για επίπεδο θορύβου 1% παρουσιάζεται στα Ιστογράμματα 5.9. Στον Πίνακα 5.1 παρουσιάζουμε το ποσοστό των δειγμάτων με σφάλμα μικρότερο ενός ορίου, και το σφάλμα που φράσσει το 90% των δειγμάτων.

π Στα προηγούμενα αποτελέσματα, παρατηρούμε ότι το σφάλμα του εστιακού μήκους f συγκεντρώνεται καλά κοντά στο μηδενικό σφάλμα, με λίγα δείγματα να παρουσιάζουν μεγάλες τιμές σφάλματος. Όπως φαίνεται και στον Πίνακα 5.1, σημαντικό ποσοστό των πειραμάτων, παρουσιάζει σφάλμα μικρότερο του 10% . Για το σφάλμα στην εκτίμηση της περιστροφής και της μετατόπισης, τα αποτελέσματα δεν είναι τόσο ικανοποιητικά, αφού παρατηρούμε μεγαλύτερη διασπορά του σφάλματος.

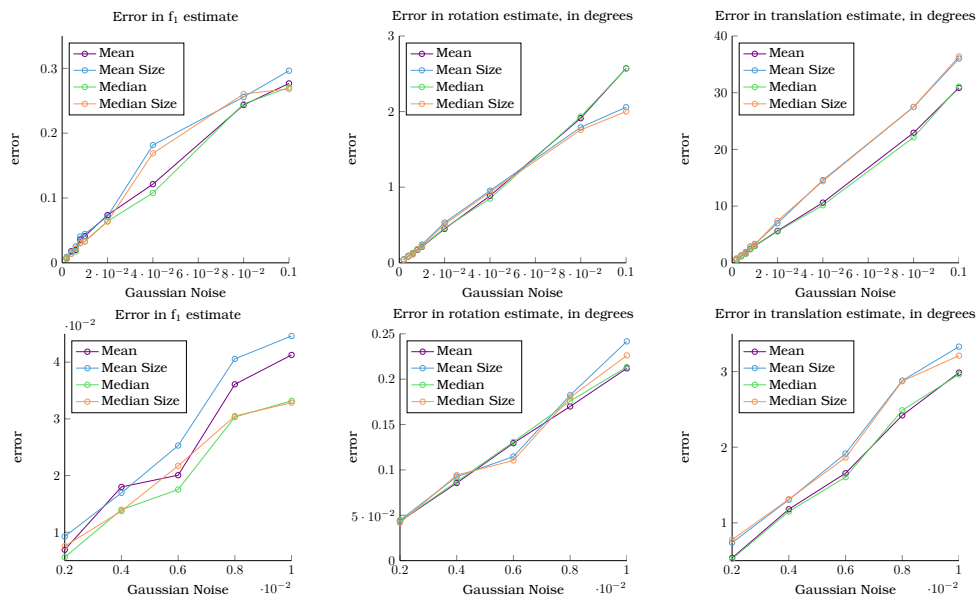
Για να μειώσουμε το σφάλμα των εκτιμήσεων μας, χρησιμοποιούμε την ακόλουθη μέθοδο. Θεωρούμε ότι η κάμερα 1 "βλέπει" κοινά σημεία με n άλλες κάμερες. Λαμβάνουμε έτσι n τιμές για το εστιακό μήκος f_1 , τις οποίες μπορούμε να χρησιμοποιήσουμε για να βελτιώσουμε την εκτίμηση του f_1 . Εικάζουμε ότι οι λύσεις που λαμβάνουμε για το f_1 θα συγκεντρώνονται γύρω από την σωστή τιμή του, ενώ λύσεις που επηρεάζονται από τον θόρυβο θα παρουσιάζονται ως outliers.

Δεδομένων των n εκτιμήσεων f_1^i για το f_1 , προσδιορίζουμε την τιμή του ως εξής:

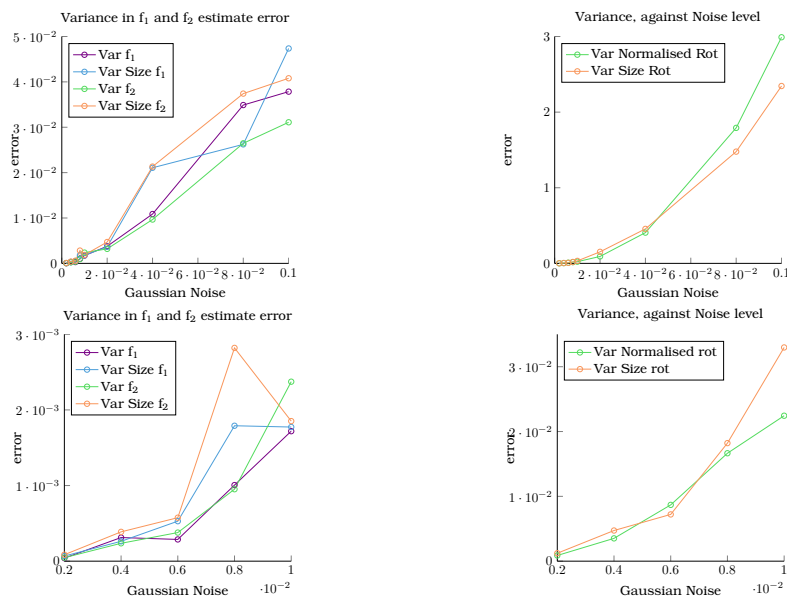
1. Υπολογίζουμε τον μέσο \bar{f}_1 των εκτιμήσεων μας f_1^i . Ο μέσος αν και επηρεάζεται από outliers, προσεγγίζει καλά την πραγματική τιμή του f_1 , όταν το σφάλμα κατανέμεται συμμετρικά γύρω από αυτήν.
2. Επιλέγουμε τις τιμές f_1^i που βρίσκονται εντός συγκεκριμένης ανοχής T

| | | |
|---|-----|-----|
| Ποσοστό αποτελεσμάτων με σφάλμα $\leq 10\%$ | 82% | 80% |
| Σφάλμα Δf_1 που φράσσει το 90% των αποτελεσμάτων | 19% | 17% |

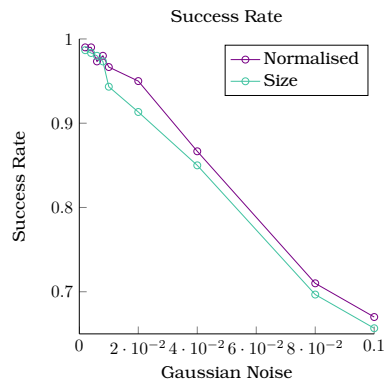
Πίνακας 5.1: Αποτελέσματα για επίπεδο Θορύβου 1% . Πρώτη στήλη: Κανονικοποιημένες διαστάσεις εικόνας, Δεύτερη στήλη: Διαστάσεις εικόνας 700×500



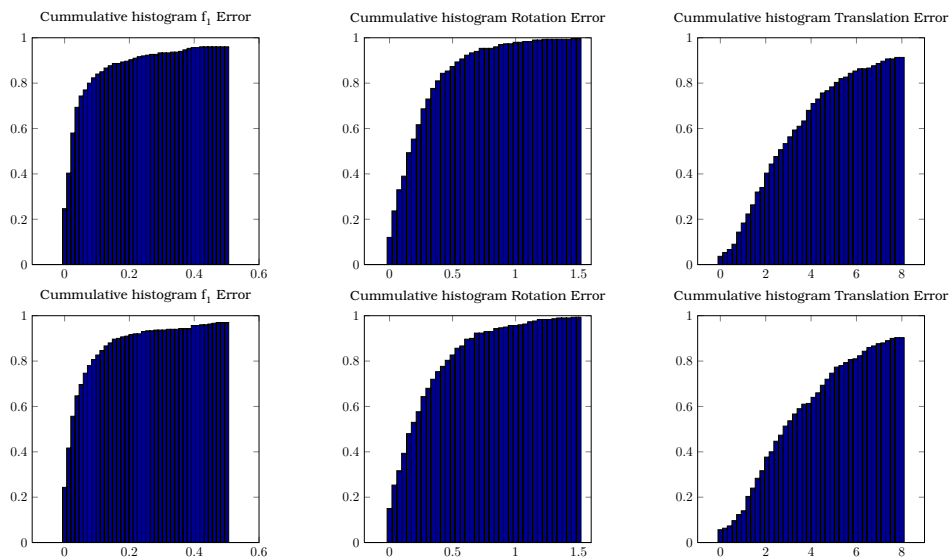
Σχήμα 5.6: Εξέλιξη του σφάλματος στις εκτιμήσεις των εστιακών μηκών, της περιστροφής και της μετατόπισης για διαφορετικά επίπεδα Θορύβου. Ο Θόρυβος περιγράφεται με την τυπική του απόκλιση ως ποσοστό των συντεταγμένων των σημείων της εικόνας.



Σχήμα 5.7: Εξέλιξη της διασποράς του σφάλματος στις εκτιμήσεις των εστιακών μηκών και της περιστροφής για διαφορετικά επίπεδα Θορύβου. Ο Θόρυβος περιγράφεται με την τυπική του απόκλιση ως ποσοστό των συντεταγμένων των σημείων της εικόνας.



Σχήμα 5.8: Εξέλιξη του ποσοστού επιτυχίας της μεθόδου, για διαφορετικά επίπεδα θορύβου. Ο Θόρυβος περιγράφεται με την τυπική του απόκλιση ως ποσοστό των συντεταγμένων των σημείων της εικόνας.



Σχήμα 5.9: Συσσωρευτικά ιστογράμματα για την κατανομή του σφάλματος. Το επίπεδο του Θορύβου είναι 1%. Στην κορυφή παρουσιάζουμε αποτελέσματα με κανονικοποιημένο μέγεθος εικόνας και στην δεύτερη σειρά με μέγεθος εικόνας 500×700

από την μέση τιμή \bar{f}_1 .

3. Υπολογίζουμε ξανά την μέση τιμή, των δειγμάτων που επιλέχθηκαν στο προηγούμενο βήμα. Αυτή είναι η εκτίμηση του f_1

Όσον αφορά την επιλογή του κατώφλιου T που αναφέραμε προηγουμένως, δοκιμάσαμε διάφορες λογικές τιμές. Υπολογίσαμε το σφάλμα μόνο στις περιπτώσεις που υπήρχαν εκτιμήσεις αρκετά κοντά (σύμφωνα με το κατώφλι T) στην μέση τιμή f_1 , και υποθέσαμε ότι σε αντίθετη περίπτωση η μέθοδος αποτυγχάνει λόγω της μεγάλης διασποράς των εκτιμήσεων. Το σφάλμα Δf_1 παρέμεινε σχεδόν σταθερό για τις διάφορες τιμές κατώφλιου. Συνεπώς η διασπορά γύρω από την μέση τιμή, δεν είναι σημαντική παράμετρος για τον προσδιορισμό του σφάλματος στην εκτίμηση.

Για την μέθοδο που μόλις περιγράψαμε, η εξέλιξη του σφάλματος σε σχέση με το επίπεδο του θορύβου, παρουσιάζεται στα Διαγράμματα 5.10. Επίσης, στα Διαγράμματα 5.11, παρουσιάζουμε την διασπορά του σφάλματος.

Στα προηγούμενα αποτελέσματα, παρατηρούμε ότι η διασπορά του σφάλματος μειώνεται με αύξηση του αριθμού των καμερών (συνεπώς και διαθέσιμων εκτιμήσεων του σφάλματος). Έτσι, η μέθοδος γίνεται πιο σταθερή και αξιόπιστη όσο περισσότερα "ζευγάρια" μπορούμε να φτιάξουμε για κάθε κάμερα του συνόλου δεδομένων. Επίσης παρατηρούμε ότι η αύξηση του αριθμού των καμερών βελτιώνει το σφάλμα στις εκτιμήσεις των εστιακών μηκών f_1, f_2 .

Στα διαγράμματα 5.10, παρουσιάσαμε το σφάλμα στην εκτίμηση της περιστροφής, της μετατόπισης και του εστιακού μήκους της δεύτερης κάμερας f_2 , όταν διαθέτουμε περισσότερες από μια εκτιμήσεις για το f_1 . Το σφάλμα αυτό υπολογίστηκε ως το σφάλμα στην ανακατασκευή που έχει την μικρότερη απόκλιση από την εκτιμηθείσα τιμή του f_1 . Σκοπός μας ήταν να δούμε αν οι ανακατασκευές που δίνουν μικρό σφάλμα στην εκτίμηση του f_1 , μπορούν να θεωρηθούν περισσότερο αξιόπιστες και για τα υπόλοιπα προς προσδιορισμό μεγέθη. Όσον αφορά το f_2 , παρατηρούμε ότι πράγματι υπάρχει τέτοια σύνδεση. Για την περιστροφή και την μετατόπιση, αν και τα αποτελέσματα είναι βελτιωμένα, η βελτίωση τους είναι συγκριτικά συγκρατημένη.

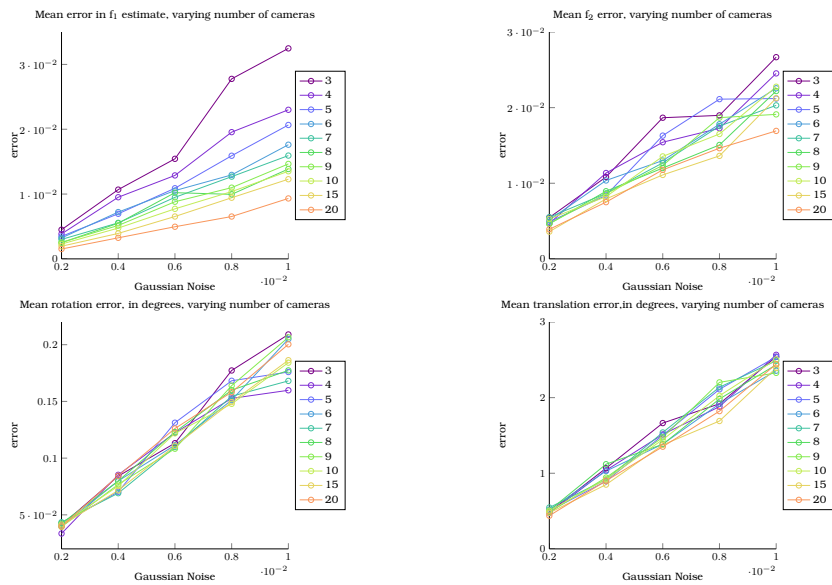
Στα Διαγράμματα 5.12, συγκρίνουμε το σφάλμα στην βασική μέθοδο και στην μέθοδο με το προηγούμενο σύστημα επιλογής. Όμοια με προηγουμένως, παρατηρούμε σημαντική βελτίωση στις εκτιμήσεις των εστιακών μηκών f_1, f_2 και μικρότερη στις εκτιμήσεις της περιστροφής και της μετατόπισης.

Σύγκριση με μεθόδους που επιλύουν τις εξισώσεις Kruppa

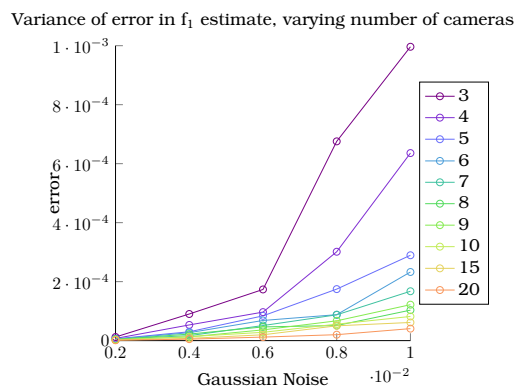
Για τις εκτιμήσεις των εστιακών μηκών f_1, f_2 , συγκρίναμε τα αποτελέσματα της μεθόδου μας με τα αποτελέσματα της μεθόδου που βασίζεται στις εξισώσεις Kruppa. Αυτή είναι και η καθιερωμένη μέθοδος για εκτιμήσεις διαφορετικών εστιακών μηκών f_1, f_2 από 2 εικόνες. Σημειώνουμε τα ακόλουθα:

Στην περίπτωση που δεν έχουμε Θόρυβο στις μετρήσεις, η προτεινόμενη μέθοδος και η μέθοδος Kruppa δίνουν σωστά τις τιμές των f_1, f_2 . Τα σφάλματα είναι της τάξης της αριθμητικής ακρίβειας του υπολογιστή.

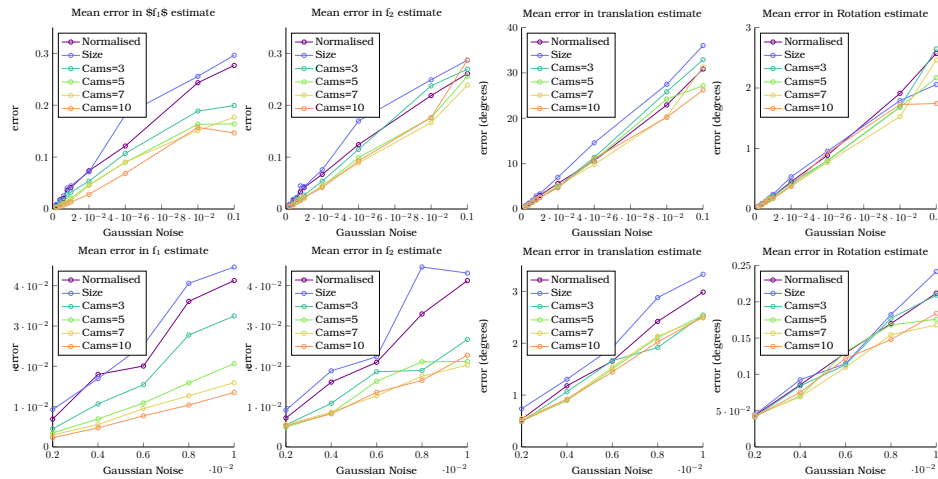
Ελέγξαμε ακόμα την συμπεριφορά της μεθόδου, για επίπεδα θορύβου στο εύρος 0.2% έως 10%. Συγκεντρωτικά τα αποτελέσματα παρουσιάζονται στον Πίνακα 5.1. Σημειώνουμε:



Σχήμα 5.10: Εξέλιξη του σφάλματος στις εκτιμήσεις των εστιακών μηκών, της περιστροφής και της μετατόπισης για διαφορετικά επίπεδα Θορύβου. Ο Θόρυβος περιγράφεται με την τυπική του απόκλιση ως ποσοστό των συνεταγμένων των σημείων της εικόνας.



Σχήμα 5.11: Εξέλιξη της διαφοράς του σφάλματος στο εστιακό μήκος Δf_1 , για διαφορετικά επίπεδα Θορύβου. Ο Θόρυβος περιγράφεται με την τυπική του απόκλιση ως ποσοστό των συνεταγμένων των σημείων της εικόνας.



Σχήμα 5.12: Σύγκριση των αποτελεσμάτων της βασικής μεθόδου, και της μεθόδου με τις πολλαπλές εκτιμήσεις του f_1 . Ο αριθμός των καμέρων δίνεται στο υπόμνημα κάθε γραφήματος. Ο Θόρυβος περιγράφεται με την τυπική του απόκλιση ως ποσοστό των συντεταγμένων των σημείων της εικόνας.

- Όταν η μέθοδος μας επιτυγχάνει, επιτυγχάνει και η μέθοδος Kruppa. Το αντίστροφο δεν ισχύει.
- Όταν και οι δύο μέθοδοι επιτυγχάνουν, τα αποτελέσματα τους ταυτίζονται.
- Όταν η μέθοδος μας αποτυγχάνει, η μέθοδος Kruppa δίνει, σε ποσοστό μεγαλύτερο από το 95% των περιπτώσεων, τελείως εσφαλμένες εκτιμήσεις.

Επομένως, παρατηρούμε ότι οι δύο μέθοδοι είναι ισοδύναμες. Το γεγονός ότι η μέθοδος μας αποτυγχάνει όταν η μέθοδος Kruppa δίνει τελείως εσφαλμένη εκτίμηση των f_1, f_2 , μπορεί να θεωρηθεί ως πλεονέκτημα, αφού αποφεύγουμε έτσι την είσοδο ακραίων τιμών (outliers) στις εκτιμήσεις των f_1, f_2 .

5.3.2 Εκφυλισμένες διατάξεις. Μια θεωρητική και πειραματική προσέγγιση

Στην Ενότητα αυτή, συνεχίζουμε την ανάλυση για τον εντοπισμό και την καταγραφή των εκφυλισμένων διατάξεων. Βασιζόμαστε στην διαθέσιμη βιβλιογραφία για το ζήτημα, εξάγουμε συμπεράσματα και τα επαληθεύουμε πειραματικά. Κύριες αναφορές της Ενότητας είναι [105, 217, 218, 135, 27]. Ξεκινάμε παρουσιάζοντας τα αποτελέσματα του [105], όσα αφορούν άμεσα το πρόβλημα μας. Η ανάλυση που ακολουθεί αφορά εκφυλισμένες διατάξεις που επηρεάζουν το πρόβλημα, ανεξάρτητα από την μέθοδο επίλυσης του. Η μέθοδος επίλυσης που θα επιλέξουμε μπορεί να διαθέτει επιπλέον εκφυλισμένες διατάξεις.

Σχετικά με τον προσδιορισμό της παραμέτρου του εστιακού μήκους f_i , έχουμε την επόμενη πρόταση

Πρόταση 1 (Πρόταση 5.2 [105]). Δοσμένου του επιπέδου στο άπειρο π_∞ , και όλων των εσωτερικών παραμέτρων (μήτρα K) πλην των εστιακών μηκών f_i , η διάταξη είναι εκφυλισμένη αν και μόνο αν υπάρχει το πολύ μία διεύθυνση προσανατολισμού.

| Περιγραφή | Τιμή |
|--|--------------|
| Αποτυχίες προτεινόμενης μεθόδου | 1798 |
| Η προτεινόμενη μέθοδος αποτυγχάνει και η μέθοδος Kruppa δίνει σφάλμα $\leq 10\%$ | 68 |
| Οι δύο μέθοδοι δίνουν ίδιες λύσεις | 18202 |
| Μέσος όρος σφάλματος μεθόδου Kruppa, στις περιπτώσεις που η προτεινόμενη μέθοδος αποτυγχάνει | $\geq 150\%$ |

Πίνακας 5.2: Αποτελέσματα σύγκρισης της προτεινόμενης μεθόδου και της μεθόδου που βασίζεται στις Εξισώσεις Kruppa. Τα στοιχεία προκύπτουν από πειράματα με 2000 δοκιμές ανά επίπεδο θορύβου. Ο θόρυβος ρυθμίστηκε στα επίπεδα (τοις εκατό):0.2, 0.4, 0.6, 0.8, 1, 2, 4, 8, 10

Αν εξετάσουμε δύο όψεις, η παραπάνω πρόταση αντιστοιχεί στην περίπτωση της μηδενικής περιστροφής. Σημειώνουμε ότι, σε αυτή την περίπτωση όλες οι πιθανές Ω_∞ βρίσκονται στο π_∞ . Συνεπώς το π_∞ μπορεί να προσδιοριστεί. Αντίστοιχα, παραθέτουμε τώρα μία πρόταση σχετικά με την δυνατότητα ανάκτησης των αφινικών ιδιοτήτων του χώρου, ισοδύναμα τον προσδιορισμό του π_∞

Πρόταση 2 (Πρόταση 5.3 [105] και Ενότητα 5 [217]). *Δοσμένων των εσωτερικών παραμέτρων των καμερών, εκτός από τα εστιακά μήκη f_i , η σχετική θέση των καμερών είναι κρίσιμη διάταξη για την αφινική ανακατασκευή αν και μόνο αν*

1. Οι κάμερες έχουν συγγραμικά οπτικά κέντρα.
2. Τα κέντρα των καμερών είναι τοποθετημένα σε κωνική του επιπέδου, με τις διευθύνσεις προσανατολισμού τους εφαπτόμενες στην συγκεκριμένη κωνική(διευθύνσεις προσανατολισμού στην διεύθυνση "κίνησης" πάνω στην κωνική των κέντρων).
3. Η κίνηση (μετατόπιση) γίνεται κατά μήκος του οπτικού άξονα, συνδυασμένη με τυχούσες περιστροφές γύρω από τον οπτικό άξονα.

Στις περιπτώσεις 1,2 προκύπτει αμφισημία στις λύσεις (2 λύσεις). Στην περίπτωση 3 έχουμε μια παραμετρική οικογένεια λύσεων.

Σχολιάζουμε τώρα την προηγούμενη πρόταση. Αρχικά, η συνθήκη 1 της προηγούμενης πρότασης, ισχύει πάντα στην περίπτωση των δύο όψεων. Συνεπώς, η αμφισημία σχετικά με την θέση του π_∞ , που εμφανίζεται στην προτεινόμενη μέθοδο και αναλύσαμε εκτενώς προηγούμενα, δεν μπορεί να ξεπεραστεί από καμία μέθοδο.

Μια διαισθητικά ικανοποιητική εξήγηση της αμφισημίας αυτής, δίνεται στο [217]. Έστω ω_1, ω_2 οι προβολές της απόλυτης κωνικής στις Εικόνες 1, 2 αντίστοιχα. Προβάλλοντας τις κωνικές αυτές στον χώρο, παίρνουμε 2 κώνους. Οι κώνοι αυτοί στην γενική περίπτωση έχουν 2 τομές, εξ' ου και η αμφισημία.

Ας εξετάσουμε τώρα μερικές ειδικές περιπτώσεις των προηγούμενων προτάσεων. Εξετάσαμε ήδη την περίπτωση μηδενικής περιστροφής, οπότε εξετάζουμε τώρα την περίπτωση μηδενικής μετατόπισης (καθαρή περιστροφή). Όπως περιγράφεται στο [217], παρ'ότι ο προσδιορισμός αφινικών ιδιοτήτων είναι αδύνατος, τα εστιακά μήκη μπορούν να βρεθούν. Συγκεκριμένα, τα εστιακά μήκη προσδιορίζονται όταν οι οπτικοί άξονες των δύο όψεων δεν ταυτίζονται (η περιστροφή δεν είναι γύρω από τον οπτικό άξονα).

Τέλος, αναφέρουμε ένα ακόμα παράδειγμα. Αν οι σχετικές θέσεις των 2 καμερών περιγράφονται με επίπεδη κίνηση, δηλαδή οι οπτικοί άξονες και τα οπτικά κέντρα ανήκουν σε κοινό επίπεδο, τότε, δεν είναι δυνατός ο προσδιορισμός του π_∞ και των f_1, f_2 [28]. Η προσθήκη τρίτης όψης αίρει την κρισιμότητα της διάταξης.

Στην συνέχεια, αναφέρουμε διατάξεις, οι οποίες είναι κρίσιμες για συγκεκριμένες μεθόδους επίλυσης του προβλήματος της βαθμονόμησης των καμερών (προσδιορισμός της μήτρας εσωτερικής βαθμονόμησης K). Οι διατάξεις αυτές αφορούν συγκεκριμένες μεθόδους, επομένως δεν είναι τόσο γενικές όσο οι προηγούμενες. Για ορισμένες από αυτές τις διατάξεις, επαληθεύσαμε πειραματικά την συμπεριφορά της μεθόδου που προτείνουμε και της μεθόδου που βασίζεται στις εξισώσεις Kruppa, όταν οι 2 κάμερες τοποθετηθούν στον χώρο σύμφωνα με αυτή την διάταξη.

Στο [27], εξετάζεται η περίπτωση δύο εικόνων με προσδιορισμένες τις παραμέτρους εσωτερικής βαθμονόμησης, εκτός από τα εστιακά μήκη f_1, f_2 . Οι συγγραφείς εξαγουν το επόμενο σχετικό συμπέρασμα

Πρόταση 3 (Από [27]). *Ο θεμελιώδης πίνακας F δεν αποσυντίθεται μοναδικά ως $K_1^{-T}BK_2^{-1}$ αν και μόνο αν*

1. *Οι οπτικοί άξονες των δύο καμερών και η ευθεία που ενώνει τα οπτικά τους κέντρα (baseline) είναι συνεπίπεδα*
2. *Ένα οπτικός άξονας, η baseline και το κάθετο διάνυσμα στον δεύτερο οπτικό άξονα και την baseline είναι συνεπίπεδα*

Παρατηρούμε ότι ήδη έχουμε σχολιάσει την περίπτωση 1.

Πειραματικά, ελέγξαμε την συμπεριφορά της προτεινόμενης μεθόδου και των εξισώσεων Kruppa στις διατάξεις 1,2. Παρατηρούμε:

- Το γραμμικό σύστημα που λύνουμε χάνει rank και ο προσδιορισμός των αγνώστων παραμέτρων δεν μπορεί να γίνει μοναδικά.
- Η μέθοδος που βασίζεται στις εξισώσεις Kruppa δίνει εσφαλμένες τιμές εστιακών μηκών f_1, f_2 .

Στο άρθρο [218], οι συγγραφείς παρουσιάζουν επιπλέον κρίσιμες διατάξεις για την μέθοδο που βασίζεται στις εξισώσεις Kruppa. Συγκεκριμένα, δείχνουν ότι η επόμενη διάταξη καμερών είναι εκφυλισμένη

- Τα οπτικά κέντρα των καμερών βρίσκονται σε σφαίρα.
- Οι κάμερες είναι προσανατολισμένες προς το κέντρο της σφαίρας, δηλαδή οι οπτικοί τους άξονες είναι ακτίνες της σφαίρας και η κατεύθυνση τους είναι προς το κέντρο αυτής

Πειραματικά, επαληθεύσαμε ότι η συγκεκριμένη διάταξη είναι κρίσιμη και για την προτεινόμενη μέθοδο, αφού το προς επίλυση γραμμικό σύστημα χάνει rank.

Από την ανάλυση που παρατίθεται στο [218] (συγκεκριμένα στην Ενότητα 8), διαπιστώνουμε ότι οι μέθοδοι που χρησιμοποιούν δύο μόνο όψεις, είναι πιο ευάλωτες σε εκφυλισμένες διατάξεις από μεθόδους με περισσότερες όψεις. Αυτό γιατί, οι εκφυλισμοί είναι συνέπεια ύπαρξης κωνικών του χώρου με ίδια εικόνα σε όλες τις όψεις. Συνεπώς, όσο αυξάνουμε τις όψεις, μειώνονται οι εκφυλισμένες διατάξεις. Στην περίπτωση των εξισώσεων Kruppa, η συγκεκριμένη κωνική του χώρου μπορεί να διαφέρει για κάθε ζευγάρι εικόνων, με την προϋπόθεση ότι η εικόνα της είναι ίδια σε όλες τις εικόνες. Όμως, αφού χρησιμοποιούμε ένα μόνο ζευγάρι εικόνων, αντιμετωπίζουμε το ίδιο πρόβλημα.

Τέλος, στο άρθρο [135], εξετάζονται οι εξισώσεις Kruppa, σε ένα πιο γενικό πλαίσιο, χωρίς να επιβάλλονται περιορισμοί στις εσωτερικές παραμέτρους των καμερών. Από την ανάλυση που γίνεται, καταλήγουν σε δύο συνθήκες για τον εκφυλισμό των εξισώσεων:

1. Άξονας περιστροφής παράλληλος στην μετατόπιση
2. Άξονας περιστροφής κάθετος στην μετατόπιση

Παρακολουθώντας ξανά την ανάλυση του [135], παρατηρούμε ότι στην περίπτωση που μόνο τα εστιακά μήκη είναι άγνωστα, μόνο οι περιπτώσεις γωνίας στροφής $\phi = 0^\circ$ (μηδενική περιστροφή) και $\phi = 90^\circ$ πρέπει να εξετασθούν. Εξετάσαμε ήδη την περίπτωση μηδενικής περιστροφής, οπότε αρκεί να εξετάσουμε την περίπτωση $\phi = 90^\circ$. Έχουμε:

- Στην περίπτωση **1** και οι δύο μέθοδοι αποτυγχάνουν, αντίστοιχα με προηγουμένως
- Η περίπτωση **2** δεν είναι εκφυλισμένη για καμία μέθοδο

Συμπεριφορά σε σχεδόν εκφυλισμένες διατάξεις

Παρουσιάζουμε τώρα, την συμπεριφορά των μεθόδων σε διατάξεις που είναι "κοντά" στις εκφυλισμένες διατάξεις που περιγράψαμε προηγουμένως. Χρησιμοποιούμε τις επόμενες διατάξεις (από [135, 218])

- Στην περίπτωση των παράλληλων οπτικών αξόνων, μεταβάλλουμε την γωνία ϕ του πίνακα στροφής της κάμερας 2 γύρω από τις 90° [135].
- Στην περίπτωση που οι κάμερες ανήκουν σε σφαίρα και "κοιτάζουν" προς το κέντρο της σφαίρας, μετακινούμε το οπτικό κέντρο της μίας κάμερας εκτός σφαίρας. Για να το κάνουμε αυτό, αλλάζουμε το μέτρο του διανύσματος θέσης του οπτικού κέντρου της κάμερας αυτής [218].
- Στην περίπτωση που οι κάμερες ανήκουν σε σφαίρα και "κοιτάζουν" προς το κέντρο της σφαίρας, περιστρέφουμε κατά μικρή γωνία το διάγραμμα προσανατολισμού της μίας κάμερας [218].

Στα πειράματα, χρησιμοποιήσαμε επίπεδο θορύβου 1%.

Τα αποτελέσματα παρουσιάζονται στα Διαγράμματα 5.13. Παρατηρούμε ότι, και οι δύο μέθοδοι έχουν ποιοτικά αντίστοιχη συμπεριφορά, όσον αφορά την ευαισθησία των διατάξεων. Οι διαφορές στο μέγεθος του σφάλματος Δf , δεν μας απασχόλησαν, αφού σε κάθε περίπτωση αυτό βρίσκεται σε πολύ υψηλά επίπεδα.

Ακόμα, όταν μεταβάλλαμε το μέτρο του διανύσματος θέσης του κέντρου της

κάμερας, οδηγηθήκαμε ξανά σε εκφυλισμένη διάταξη. Επίσης, παρατηρούμε ότι η συμπεριφορά των μεθόδων είναι καλύτερη στην περίπτωση των παράλληλων οπτικών αξόνων (σχεδόν άμεση αποκατάσταση φυσιολογικής συμπεριφοράς της μεθόδου), σε σχέση με την διάταξη των καμερών σε σφαίρα. Παρατηρούμε ότι στις διατάξεις σφαίρας, η συμπεριφορά των μεθόδων διαταράσσεται σημαντικά για μεγάλο εύρος "σχεδόν" κρίσιμων διατάξεων.

5.3.3 Πειράματα με πραγματικά δεδομένα

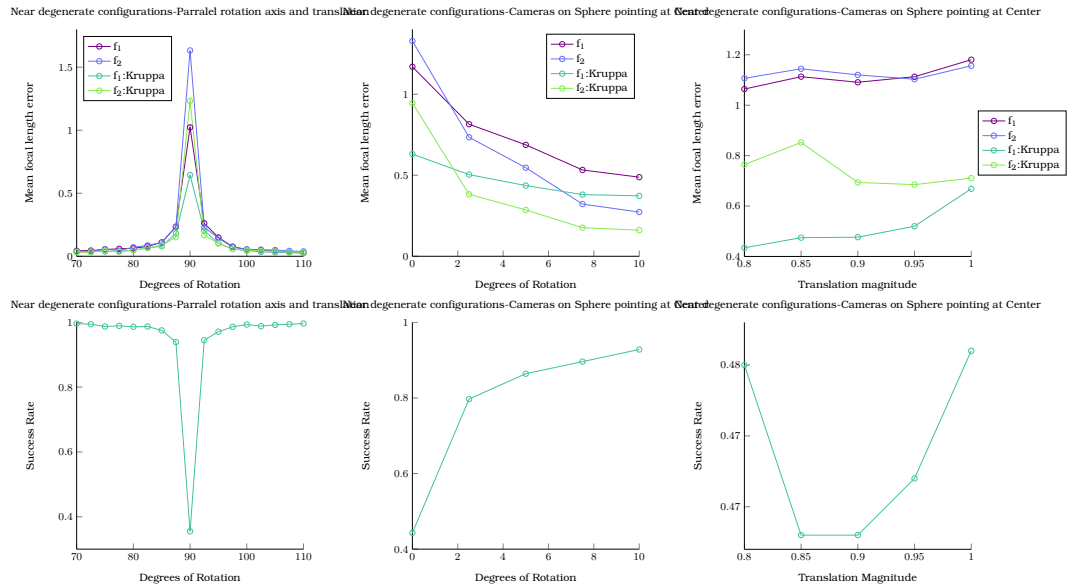
Στην ενότητα αυτή θα παρουσιάσουμε τα αποτελέσματα της εφαρμογής της μεθόδου που προτείνουμε σε σύνολα φωτογραφιών. Χρησιμοποιούμε το σύνολο δεδομένων του [216]. Το συγκεκριμένο σύνολο δεδομένων, αποτελείται από σύνολα φωτογραφιών εξωτερικών σκηνών. Παρέχονται οι μήτρες εσωτερικών και εξωτερικών παραμέτρων κάθε κάμερας, γνωρίζουμε δηλαδή τις σωστές τιμές για τις υπό προσδιορισμό παραμέτρους. Θεωρήσαμε ότι το πρωτεύων σημείο βρίσκεται στο κέντρο της εικόνας, υπόθεση η οποία ικανοποιείται ικανοποιητικά από το σύνολο δεδομένων. Τα διάφορα σύνολα φωτογραφιών, και οι ονομασίες με τις οποίες αναφέρονται στην βιβλιογραφία, βρίσκονται στον Πίνακα 5.3. Σημειώνουμε ότι το συγκεκριμένο σύνολο δεδομένων έχει δημιουργηθεί για αξιολόγηση αλγορίθμων γεωμετρίας πολλών όψεων (*mvs*, *sfm*, κλπ).

Για την επίλυση του προβλήματος ακολουθήσαμε την επόμενη διαδικασία

1. Ανίχνευση χαρακτηριστικών. Ταίριασμα χαρακτηριστικών μεταξύ εικόνων, με το κριτήριο του λόγου των δύο κοντινότερων γειτόνων [133]. Το σύνολο δεδομένων αποτελείται από εικόνες υψηλής ανάλυσης. Κατά συνέπεια μεγάλο πλήθος χαρακτηριστικών ανιχνεύονται σε κάθε εικόνα. Για να επιταχυνθεί το ταίριασμα, χρησιμοποιήσαμε τον προσεγγιστικό αλγόριθμο Approximate Nearest Neighbors (ANN) και την δομή δεδομένων kd-tree. Σε προηγούμενες ενότητες έχουμε αναφερθεί με μεγαλύτερη λεπτομέρεια στα θέματα αυτά (Κεφάλαιο 4.1).
2. Υπολογισμός Θεμελιώδους πίνακα F για κάθε ζευγάρι ταιριασμένων εικόνων. Απορρίπτουμε ζευγάρια με λιγότερες από 50 αντιστοιχίσεις, τιμή που επιλέξαμε αυθαίρετα.
3. Από τον πίνακα F , υπολογίζουμε την κανονική αναπαράσταση των

| Ονομασία συνόλου | Αριθμός φωτογραφιών |
|------------------|---------------------|
| Herz-Jesu-P25 | 25 |
| Herz-Jesu-P8 | 8 |
| Fountain-P11 | 11 |
| entry-P10 | 10 |
| castle-P30 | 30 |
| castle-P19 | 19 |

Πίνακας 5.3: Περιγραφή συνόλων δεδομένων



Σχήμα 5.13: Σφάλμα (πρώτη σειρά) και ποσοστό επιτυχίας (δεύτερη σειρά) για σχεδόν εκφυλισμένες διατάξεις. Πρώτη στήλη: Παράλληλοι οπτικο άξονες. Δεύτερη στήλη: Κάμερες σε σφαίρα και στροφή ενός οπτικού άξονα. Τρίτη στήλη: Κάμερες σε σφαίρα και μετατόπιση του κέντρου της μίας κάμερας. Στην τελευταία περίπτωση, στον άξονα x δίνουμε το μέτρο του διανύσματος θέσης του μετατοπισμένου οπτικού κέντρου. Θυμίζουμε ότι η σφαίρα έχει μοναδιαία ακτίνα.

μητρών προβολής για τις δύο κάμερες P_1, P_2 και επιλύουμε το Σύστημα 5.13.

Για την εκτέλεση των βημάτων 1,2, δοκιμάσαμε διάφορες μεθόδους, τις οποίες θα περιγράψουμε σε επόμενες ενότητες. Επίσης στο τέλος του κεφαλαίου, συγκεντρώνουμε τις διάφορες πηγές λογισμικού που χρησιμοποιήσαμε. Στον Πίνακα 5.4, συγκεντρώνουμε κάποιες παραμέτρους οι οποίες είναι κοινές σε όλες τις μεθόδους, μαζί με βιβλιογραφικές πηγές για τον καθορισμό των τιμών του.

Αρχικά αποτελέσματα

Στην ενότητα αυτή, παρουσιάζουμε κάποια πρώτα αποτελέσματα, χωρίς να βελτιστοποιήσουμε τις μεθόδους που χρησιμοποιούμε και τις παραμέτρους αυτών. Σκοπός είναι να δώσουμε μια εικόνα των επιδόσεων της μεθόδου, στην γενική περίπτωση, χωρίς προσαρμογή στο σύνολο δεδομένων μας (παράμετροι) ή χρήση εξελιγμένων μεθόδων εξαγωγής χαρακτηριστικών και υπολογισμού της μήτρας F .

Για την ανίχνευση και την περιγραφή χαρακτηριστικών, χρησιμοποιούμε τα χαρακτηριστικά SIFT σε περιοχές που ανιχνεύονται με τον τελεστή DoG. Δεν μεταβάλλουμε καθόλου τις παραμέτρους του ανιχνευτή περιοχών που σχετίζονται με την τιμή και την κυρτότητα των ακρότατων που εντοπίζουμε. Θέτουμε την παράμετρο $T_{ratio} = 0.4$. Παρότι η τιμή αυτή είναι αυστηρή σε σχέση με άλλες που έχουν προταθεί στην βιβλιογραφία [133, 210], εντοπίζουμε πολλά χαρακτηριστικά λόγω της μεγάλης ανάλυσης των εικόνων. Προσπαθούμε έτσι να περιορίσουμε τον αριθμό τους κρατώντας μόνο τα καλύτερα από αυτά. Σε μεγαλύτερο βάθος, τα θέματα αυτά αναλύθηκαν στην

| Παράμετρος | Τιμή | Βιβλιογραφική αναφορά |
|---|---------------------------------|-----------------------|
| Ελάχιστες αντιστοιχίες μεταξύ δύο εικόνων T_{ms} | 50 | |
| Λόγος δεύτερου προς πρώτου κοντινότερου γείτονα T_{ratio} | 0.4 – 0.8 | [133, 210] |
| Μέγιστη απόσταση inlier $T_{inliers}$ | 0.6% μέγιστης διάστασης εικόνας | [210] |
| Απόσταση σημείων $d(,)$ | Προσέγγιση Sampson | |

Πίνακας 5.4: Παράμετροι προβλήματος

Ενότητα 4.1.

Μετά το ταιρίασμα των χαρακτηριστικών των εικόνων, έχουμε ταιριασμένα σημεία για κάθε ζεύγος εικόνων. Έστω I_1, I_2 δύο εικόνες. Από την διαδικασία ταιριάσματος, είναι δυνατόν ένα χαρακτηριστικό της I_1 να έχει ταιριαστεί με πολλά χαρακτηριστικά της I_2 . Προφανώς κάποια από αυτά τα ταιριάσματα είναι εσφαλμένα. Για να αντιμετωπίσουμε την κατάσταση αυτή, αφαιρούμε όλα αυτά τα ταιριάσματα (όσα δηλαδή δεν είναι 1 – 1). Η ιδέα αυτή προέρχεται από το [206].

Για τον προσδιορισμό της θεμελιώδους μήτρας F , εφαρμόζουμε μία διαδικασία σε δύο στάδια

1. Βρίσκουμε μια πρώτη εκτίμηση της F , με την μέθοδο LMedS[187]. Επαναλαμβάνουμε ότι χρησιμοποιούμε για μετρική απόστασης την προσέγγιση του Sampson στο σφάλμα επαναπροβολής.
2. Χρησιμοποιούμε την μέθοδο του Torr (M-estimator)[229]. Χρησιμοποιούμε μόνο τις αντιστοιχίες σημείων που είναι inliers στην εκτίμηση της F από το στάδιο 1.

Επειδή έχουμε πολλές εσφαλμένες αντιστοιχίες στις εικόνες (outliers), είναι σημαντικό να χρησιμοποιήσουμε αρχικά μια μέθοδο σθεναρή σε μεγάλο ποσοστό outliers. Για αυτό επιλέξαμε να χρησιμοποιήσουμε την μέθοδο LMedS. Στην συνέχεια, η εκτέλεση της εκτίμησης μας για την F γίνεται με έναν τρόπο γρήγορο, που συχνά εμφανίζεται στην βιβλιογραφία. Η μέθοδος αυτή δεν μπορεί να χρησιμοποιηθεί μόνη της (χωρίς την LMedS), γιατί δεν μπορεί να αντιμετωπίσει πολλούς outliers στα δεδομένα. Ο χρόνος εκτέλεσης της μεθόδου του Torr είναι σημαντικά μικρότερος από μεθόδους που χρησιμοποιούν επαναληπτικούς αλγόριθμους βελτιστοποίησης. Τέλος, απορρίπτουμε εκτιμήσεις που έχουν λιγότερους από 70 inliers.

Για τον υπολογισμό του f_1 από τις διάφορες εκτιμήσεις που διαθέτουμε για αυτό, αφού κάθε εικόνα συμμετέχει σε αυθαίρετο αριθμό ζευγών εικόνων, χρησιμοποιούμε την ίδια μέθοδο που αναλύσαμε στην Ενότητα των πειραμάτων με συνθετικά δεδομένα.

Παρουσιάζουμε στον Πίνακα 5.5 τα αποτελέσματα για το Δf_1 . Τα σύνολα δεδομένων που χρησιμοποιήθηκαν περιγράφονται στην λεζάντα του πίνακα. Παρατηρούμε τα επόμενα:

- Στο Σύνολο 1 έχουμε πολύ καλά αποτελέσματα, ενώ στο Σύνολο 2, εξαιτίας της προσθήκης των "Castle" συνόλων δεδομένων, τα αποτελέσματα γίνονται χειρότερα.
- Στα σύνολα δεδομένων "Castle", η σχετική αποτυχία της μεθόδου μπορεί να οφείλεται στην δυσκολία ταιριάσματος χαρακτηριστικών λόγω επαναλαμβανόμενων δομών και στην μη ευνοϊκή σχετική θέση των φωτογραφικών λήψεων.
- Αν στο Σύνολο 1 αφαιρέσουμε την μία φωτογραφία στην οποία έχουμε πολύ μεγάλο σφάλμα, θεωρώντας ότι η μέθοδος αποτυγχάνει στην περίπτωση αυτή, το μέσο σφάλμα προκύπτει 3%
- Για το Σύνολο 2, συμπληρώνουμε ότι το 31% των εικόνων έχει σφάλμα Δf_1 μεγαλύτερο από 20% και μόλις το 4% πάνω από 40%

Στην συνέχεια, χρησιμοποιήσαμε διαφορετικές μεθόδους εκτίμησης του θεμελιώδους πίνακα (συνοπτικά, τον κανονικοποιημένο αλγόριθμο 8 σημείων του Hartley ως αρχικοποίηση σχήματος βελτιστοποίησης LM στην προσέγγιση Sampson για το σφάλμα) στο σύνολο δεδομένων entry-P10. Το σύνολο αυτό παρουσίαζε από τα μεγαλύτερα σφάλματα στα προηγούμενα πειράματα. Κατάλληλη επιλογή παραμέτρων (ελάχιστος αριθμός inliers στην εκτίμηση του σταδίου 2, κατάλληλη μέθοδος για το στάδιο 1), μπορεί να μειώσει το μέσο σφάλμα σε 11%. Στην συνέχεια θα αναφερθούμε αναλυτικά σε προσπάθειες βελτίωσης του βασικού σχήματος.

Τέλος, για την εκτίμηση του σφάλματος στην μήτρα περιστροφής, ΔR , και στο διάνυσμα μετατόπισης, Δt , χρησιμοποιούμε την ακόλουθη μέθοδο. Για κάθε εικόνα i έστω ότι διαθέτουμε n ανακατασκευές στις οποίες συμμετέχει η εικόνα i . Υπολογίζουμε από αυτές την τιμή του f_i , όπως περιγράφηκε στην ενότητα των πειραμάτων σε συνθετικά δεδομένα. Στην συνέχεια, από τις n ανακατασκευές, απορρίπτουμε όσες δίνουν εκτίμηση f_i^{est} με σφάλμα Δf_i μεγαλύτερο από T_f . Για τον υπολογισμό του σφάλματος Δf_i χρησιμοποιούμε στην θέση του σωστού f_i , την τιμή που υπολογίσαμε από τις n ανακατασκευές. Από αυτές τις "αξιόπιστες" ανακατασκευές, υπολογίζουμε τώρα τα σφάλματα ΔR , Δt που παρουσιάζουμε στον Πίνακα 5.6.

Για τα αποτελέσματα του Πίνακα 5.6 σχολιάζουμε τα ακόλουθα:

- Για τα σύνολα δεδομένων που έχουμε καλύτερες εκτιμήσεις του f_i (τρεις πρώτες στήλες Πίνακα 5.6) έχουμε ελαφρώς καλύτερα αποτελέσματα για τα R, t .

| | Σύνολο 1 | Σύνολο 2 |
|------------------------|----------|----------|
| Μέση τιμή | 5% | 12% |
| Διάμεσος | 1% | 5% |
| $\Delta f_1 \geq 10\%$ | 4(9%) | 47(45%) |
| $\Delta f_1 \geq 30\%$ | 1(2%) | 12(11%) |

Πίνακας 5.5: Σφάλμα στην εκτίμηση του εστιακού μήκους f_1 . Το Σύνολο 1 αποτελείται από τα Σύνολα Fountain-P11, Herz-Jesu-P25, Herz-Jesu-P8. Το Σύνολο 2 αποτελείται από όλα τα σύνολα που παρουσιάζουμε στον Πίνακα 5.3

| | HJ-P8 | HJ-P25 | fountain-Plehntry-P10 | castle-P30 | castle-P19 |
|----------------------|-------|--------|-----------------------|------------|------------|
| Μέση τιμή ΔR | 0.07 | 0.07 | 0.05 | 0.80 | 0.32 |
| Διάμεσος ΔR | 0.04 | 0.06 | 0.05 | 0.35 | 0.17 |
| Μέση τιμή Δt | 15.86 | 0.92 | 0.94 | 8.13 | 5.35 |
| Διάμεσος Δt | 1.34 | 0.65 | 0.57 | 5.69 | 3.27 |

Πίνακας 5.6: Σφάλμα στην εκτίμηση της διεύθυνσης μετατόπισης και της περιστροφής. Οι τιμές είναι σε μοίρες. Χρησιμοποιήθηκε η ίδια βασική μέθοδος εντοπισμού χαρακτηριστικών και υπολογισμού των μητρών F που περιγράψαμε νωρίτερα και που έδωσε τα Αποτελέσματα για το Δf_1 στον Πίνακα 5.5. Χρησιμοποιήσαμε $T_f = 0.1$

- Οι εκτιμήσεις των R, t περιέχουν ακραίες τιμές, γεγονός που φαίνεται από την μεγάλη διαφορά της μέσης τιμής, που είναι ευαίσθητη σε outliers, και της διαμέσου. Η παρατήρηση αυτή έρχεται σε συμφωνία με τις παρατηρήσεις μας στα πειράματα σε συνθετικά δεδομένα, π.χ. Εικόνα 5.9.
- Τα αποτελέσματα είναι χειρότερα, τόσο από άποψη τιμής του σφάλματος όσο και από άποψη ευστάθειας, στην εκτίμηση του διανύσματος μετατόπισης. Επαληθεύονται δηλαδή ξανά οι παρατηρήσεις που κάναμε από τα πειράματα σε συνθετικά δεδομένα.

Υπολογισμός μήτρας F - Σθεναρότητα σε ακραίες τιμές (outliers)

Αρχίζουμε αυτή την Ενότητα, με την περιγραφή των μεθόδων που δοκιμάσαμε για τον υπολογισμό της F . Δίνουμε σύντομα ονόματα στις μεθόδους αυτές, για την εύκολη αναφορά τους.

- **Mat4:** Εφαρμόζουμε την μέθοδο LMedS για την παραγωγή μιας πρώτης εκτίμησης της F και την εκτίμηση των inliers από αυτήν. Χρησιμοποιούμε την απόσταση Sampson. Στην συνέχεια εφαρμόζουμε στους inliers τον M-estimator του Torr.
- **Mat3:** Εφαρμόζουμε την μέθοδο LMedS για την παραγωγή μιας πρώτης εκτίμησης της F και την εκτίμηση των inliers από αυτήν. Χρησιμοποιούμε την απόσταση Sampson. Στην συνέχεια χρησιμοποιούμε τον κανονικοποιημένο αλγόριθμο 8 σημείων (normalised 8 point algorithm [90, σ. 282]) και τους inliers, και εξάγουμε μια νέα εκτίμηση της F . Η εκτίμηση αυτή χρησιμοποιείται ως αρχικοποίηση επαναληπτικής μεθόδου βελτιστοποίησης με τον αλγόριθμο LM. Ως δεδομένα χρησιμοποιούμε τους inliers και για κόστος βελτιστοποίησης την προσέγγιση Sampson του σφάλματος επαναπροβολής. Η μέθοδος αυτή περιγράφεται από τις Εξισώσεις 18, 19 της δημοσίευσης [134]
- **Kov4:** Η Μέθοδος αυτή είναι ίδια με την μέθοδο **Mat 4**. Χρησιμοποιούμε όμως την μέθοδο RANSAC στην θέση της μεθόδου LMedS.

Χρησιμοποιούμε για τις υποθέσεις της μεθόδου RANSAC τον αλγόριθμο 7 σημείων του [90], προκειμένου να μειώσουμε τον αριθμό των επαναλήψεων. Ο αριθμός των επαναλήψεων υπολογίζεται δυναμικά, όπως περιγράψαμε στην Ενότητα 2.1.5.

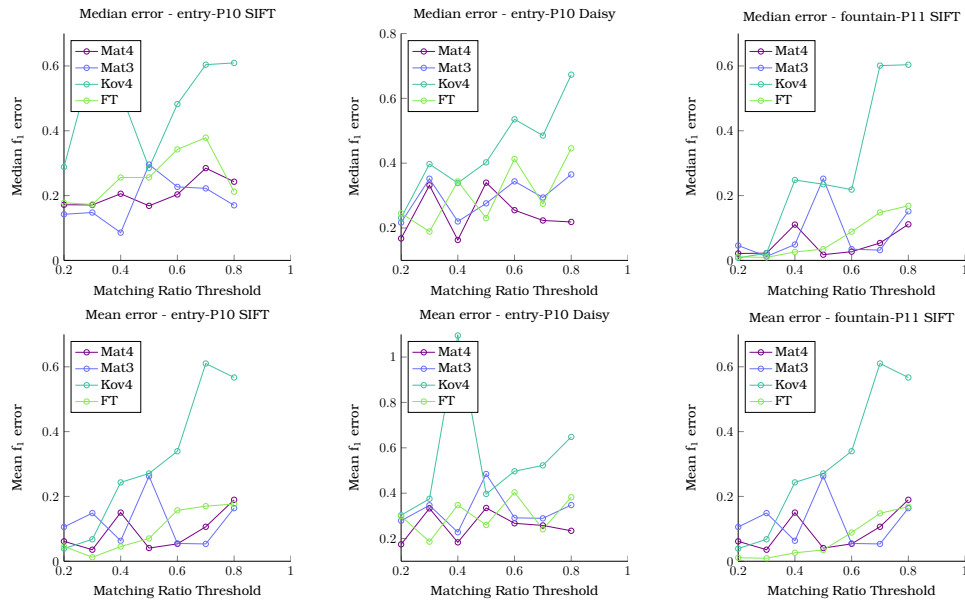
- **Kov3**: Όμοια με την μέθοδο **Mat3**, χρησιμοποιούμε όμως την μέθοδο RANSAC αντί της μεθόδου LMedS. Οι λεπτομέρειες είναι όπως στην μέθοδο **Kov4**.
- **FT**: Χρησιμοποιούμε την μέθοδο LMedS. Θεωρούμε, όπως άλλωστε και προηγουμένως, 50% πιθανότητα να είναι σωστή μια αντιστοιχία σημείων, προκειμένου να υπολογίσουμε τον αριθμό των επαναλήψεων. Για τον υπολογισμό του θεμελιώδους πίνακα, χρησιμοποιούμε ορθογώνια ελάχιστα τετράγωνα (orthogonal least squares). Περισσότερες λεπτομέρειες δίνονται στο [11], περίπτωση 19. Τέλος, όπως και σε όλες τις υπόλοιπες μεθόδους, οι inliers στην λύση υπολογίζονται με την απόσταση Sampson.

Συγκρίνουμε τώρα τα αποτελέσματα που παίρνουμε με χρήση της μεθόδου RANSAC και με χρήση της μεθόδου LMedS, κρατώντας ίδια την υπόλοιπη διαδικασία. Παρατηρούμε ότι, το ποσοστό των outliers είναι αρκετά υψηλό, συνεπώς θα πρέπει να χρησιμοποιήσουμε την μέθοδο LMedS για να αποκτήσουμε αξιόπιστα αποτελέσματα. Η μέθοδος RANSAC, αν και υπολογιστικά πιο αποδοτική, δεν έχει την απαιτούμενη σθεναρότητα. Εποπτικά, βλέπουμε στα Διαγράμματα 5.14, ότι η μέθοδος **Kov4** έχει πάντα χειρότερα αποτελέσματα από την μέθοδο **Mat4**. Επίσης, από το Διάγραμμα 5.14, συγκεκριμένα από τους συνδυασμούς entry-P10 daisy και fountain-P11 SIFT, βλέπουμε ότι για μικρή τιμή του κατωφλίου T_{ratio} , εντοπίζουμε πιο αξιόπιστα τις αντιστοιχίες μεταξύ των εικόνων και σε ορισμένες περιπτώσεις τα αποτελέσματα των μεθόδων που χρησιμοποιούν τον αλγόριθμο RANSAC ταυτίζονται με αυτά των μεθόδων που χρησιμοποιούν την μέθοδο LMedS. Το θέμα της υπολογιστικής πολυπλοκότητας των μεθόδων θα αναλυθεί αργότερα στην υποενότητα "Χρονικοί Περιορισμοί".

Σύγκριση επαναληπτικής βελτιστοποίησης LM και χρήσης M-estimator του Torr[229]

Θα συγκρίνουμε μεθόδους που αφορούν το δεύτερο στάδιο προσδιορισμού της μήτρας F . Όπως βλέπουμε από τα Διαγράμματα 5.14, δεν είναι ξεκάθαρη η σχέση των μεθόδων που χρησιμοποιούν βελτιστοποίηση με τον αλγόριθμο LM του σφάλματος (μέθοδοι **Mat3** και **Kov3**) με τους αντίστοιχους που χρησιμοποιούν τον M-estimator του Torr (μέθοδοι **Mat4** και **Kov4**). Ανάλογα με την τιμή του T_{ratio} , αλλάζει η σχετική κατάταξη των μεθόδων. Παρατηρούμε όμως ότι η απόκλιση των μεθόδων είναι κατά περιπτώσεις μικρή, όπως για παράδειγμα φαίνεται από την διάμεσο του σφάλματος στο fountain-P11 ή το μέσο σφάλμα του entry-P10. Συνεπώς συνίσταται η χρήση της μεθόδου του Torr, αφού γενικά οι επαναληπτικές αριθμητικές μέθοδοι βελτιστοποίησης είναι ακριβείς αλλά και υπολογιστικά απαιτητικές.

Τέλος, συγκρίνοντας όλες τις μεθόδους, παρατηρούμε ότι οι μέθοδοι **Mat4**, **FT** παρουσιάζουν την καλύτερη επίδοση. Η μέθοδος **FT** παρουσιάζει πιο σταθερή συμπεριφορά σε σχέση με το T_{ratio} , συνεπώς περιμένουμε να βρίσκει ευκολότερα γενικότερη εφαρμογή. Επίσης, υπερέχει της μεθόδου **Mat4** στο σύνολο Fountain-P10, στο οποίο έχουμε καλά αποτελέσματα. Δεν μπορούμε



Σχήμα 5.14: Πρώτη σειρά: Διάμεσος σφάλματος Δf_1 . Δεύτερη σειρά: Μέσο σφάλμα. Χρησιμοποιήθηκε η Μέθοδος **Mat4**, με παραμέτρους $T_{inliers} = 30, T_{ratio} = 0.4$

όμως να διαλέξουμε οριστικά υπέρ της μίας ή της άλλης μεθόδου, αφού έχουν συγκρίσιμη επίδοση.

Μέθοδοι περιγραφής χαρακτηριστικών

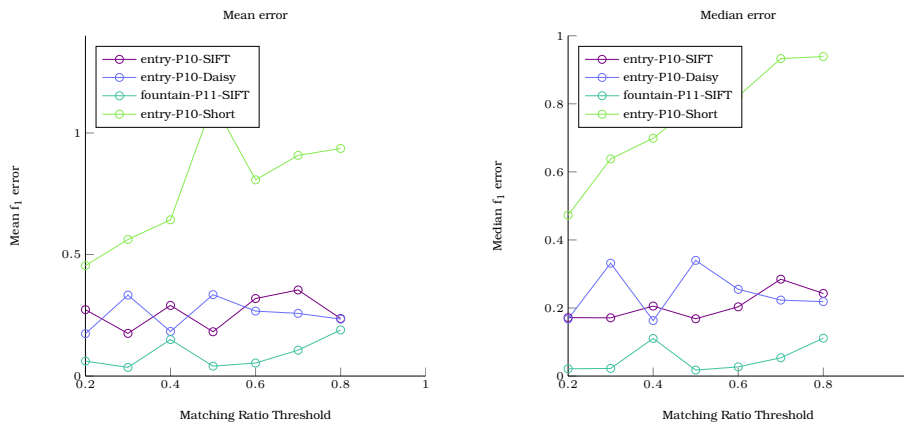
Στην ενότητα αυτή θα συγκρίνουμε διάφορες μεθόδους περιγραφής χαρακτηριστικών. Σε όλες τις περιπτώσεις, χρησιμοποιούμε τον ίδιο ανιχνευτή περιοχών, συγκεκριμένα τον τελεστή DoG των χαρακτηριστικών SIFT. Για λεπτομέρειες, ο αναγνώστης παραπέμπεται στην Ενότητα 4.1. Αναφέρουμε ότι δεν ρυθμίσαμε τις παραμέτρους του ανιχνευτή περιοχών.

Τα χαρακτηριστικά που δοκιμάσαμε στις εικόνες είναι τα ακόλουθα

- Περιγραφείς SIFT. Για τις παραμέτρους χρησιμοποιήσαμε τις καθιερωμένες τιμές. Καταλήξαμε έτσι στον γνωστό περιγραφέα διάστασης 128.
- Περιγραφείς Daisy. Χρησιμοποιήσαμε τις καθιερωμένες τιμές των παραμέτρων, όπως αυτές παρουσιάζονται στον Πίνακα 4.3. Ο περιγραφέας που προκύπτει έχει διάσταση 200.
- Χρήση ενός τροποποιημένου περιγραφέα Daisy. Συγκεκριμένα χρησιμοποιήσαμε για τις παραμέτρους του περιγραφέα τις τιμές που προτείνονται στο [227] προκειμένου να καταλήξουμε σε ένα συμπιεσμένο περιγραφέα διάστασης 36. Οι τιμές αυτές, ακολουθώντας τον συμβολισμό της Ενότητας 4.5, είναι $R = 5, Q = 2, T = 4, H = 4$.

Οι παραπάνω μέθοδοι ανίχνευσης χαρακτηριστικών δεν είναι προφανώς βέλτιστοι. Έχουν όμως το πλεονέκτημα της εύκολης διαθεσιμότητας.

Από τα αποτελέσματα, παρατηρούμε ότι ο μικρός περιγραφέας δεν έχει την απαιτούμενη διακριτική ικανότητα για δύσκολα σύνολα δεδομένων όπως το entry-P10. Ακόμα, η επίδοση των χαρακτηριστικών SIFT και των χαρακτηριστικών Daisy είναι παρόμοια και εξαρτάται από το T_{ratio} . Τα Αποτελέσματα



Σχήμα 5.15: Απόδοση διαφορετικών ανιχνευτών, για διάφορες τιμές του T_{ratio} . Ο θεμελιώδης πίνακας υπολογίζεται σε κάθε περίπτωση με την μέθοδο **Mat4** και με $T_{inliers} = 30$

παρουσιάζονται στα Διαγράμματα 5.15.

Σημειώνουμε τέλος, ότι η προσπάθεια μας να χρησιμοποιήσουμε τους περιγραφείς Daisy παρακινήθηκε από την υψηλή διακριτική ικανότητα της τοπολογίας τους, την καλύτερη ταχύτητα υπολογισμού των περιγραφέντων Daisy σε σύγκριση με τους περιγραφείς SIFT, αλλά και επειδή είναι δημόσια διαθέσιμοι και είναι δυνατή η εύκολη ρύθμιση των παραμέτρων τους.

Μέθοδος απόρριψης λύσεων από το πλήθος των inliers

Μετά το πέρας της εκτίμησης των μητρών F , θέλουμε να απορρίψουμε κάποιες από αυτές, ώστε να κρατήσουμε μόνο όσες θεωρούνται αξιόπιστες. Μέχρι στιγμής, έχουμε απορρίψει μόνο ζευγάρια εικόνων με λιγότερα από T_{ms} αντιστοιχίες. Επίσης, στην προηγούμενη υποενότητα, χρησιμοποιήσαμε ένα κατώφλι $T_{inliers}$ για το πλήθος των inliers στην ανακτηθείσα μήτρα F , ώστε να γίνει αυτή δεκτή.

Προκειμένου να επεκτείνουμε την προηγούμενη λογική, να την εξορθολογίσουμε και να την γενικεύσουμε, εισάγουμε την επόμενη ευριστική διαδικασία απόρριψης μητρών F :

- Το πλήθος των inliers που έχει μια μήτρα F , δείχνει την στήριξη της από τα δεδομένα του προβλήματος. Συνεπώς, είναι λογικό να υποθέσουμε ότι όσο αυξάνονται οι inliers, αυξάνεται και η αξιοπιστία της F . Εισάγουμε έτσι ένα κατώφλι T_l . Έστω μήτρα F_j^i , όπου ο συμβολισμός υποδηλώνει ότι προέρχεται από τις κάμερες i, j , με την κάμερα i τοποθετημένη στην αρχή των αξόνων, δηλαδή

$$\mathbf{x}_j^i F_j^i \mathbf{x}_i = 0$$

Για την *απόρριψη* της μήτρας F_j^i εισάγουμε το κριτήριο

$$\# \text{inliers}_j^i / \# \text{inliers}_{max}^i \leq T_l$$

όπου στον λόγο, στον αριθμητή έχουμε τους inliers της F_j^i και στον παρονομαστή το μέγιστο πλήθος inliers που βρήκαμε στο πρόβλημα για μήτρα F_j^i, j' οποιοδήποτε.

- Αντίστοιχα, υπερβολικός αριθμός inliers μπορεί να σημαίνει ότι έχουμε προσαρμογή στον θόρυβο, δεδομένου ότι οι outliers είναι μεγάλο ποσοστό των αρχικών αντιστοιχιών. Επομένως εισάγουμε αντίστοιχα το κριτήριο *απόρριψης*

$$\# \text{ inliers}_j^i / \# \text{ inliers}_{max}^i \geq T_u$$

Στα Διαγράμματα 5.16, παρουσιάζουμε το μέσο σφάλμα Δf_1 , την διάμεσο του σφάλματος και τον ρυθμό αποτυχίας, για τα σύνολα δεδομένων fountain-P11 και entry-P10. Παρατηρούμε ότι οι καλύτερες τιμές για την παράμετρο T_u είναι το 1, εκτός και αν δεν μας ενδιαφέρει ο ρυθμός αποτυχίας. Αντίστοιχα, το T_l καλύτερα να επιλεγεί στο διάστημα 0.4 – 0.6. Παρατηρούμε τέλος ότι μεταβολές του T_l , επηρεάζουν λιγότερο την διάμεσο και περισσότερο την μέση τιμή.

Χρονικοί Περιορισμοί

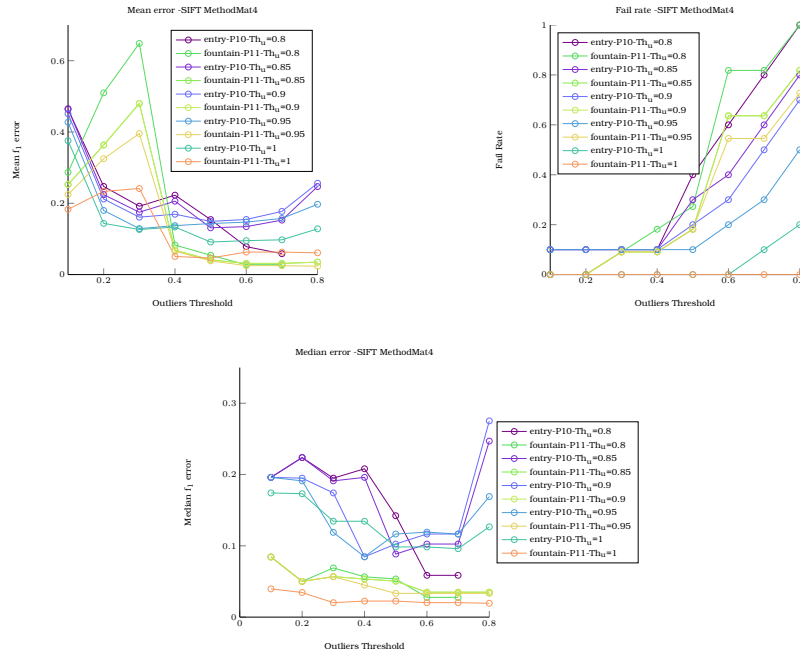
Στην ενότητα αυτή, εξετάζουμε τον χρόνο εκτέλεσης των διάφορων μεθόδων εκτίμησης της μήτρας F . Τα αποτελέσματα για το σύνολο δεδομένων entry-P10 δίνονται στον Πίνακα 5.7. Λεπτομέρειες για τον τρόπο δοκιμής δίνονται στην λεζάντα του πίνακα. Στον Πίνακα 5.7, έχουμε συμπεριλάβει και τις μεθόδους **MAP1,MAP2**. Οι μέθοδοι αυτοί, χρησιμοποιούν τον Αλγόριθμο MAPSAC [228] (Maximum A-posteriori), μια παραλλαγή του RANSAC που δίνει, όπως φανερώνει εξάλλου το όνομα του, την λύση με την μέγιστη A-posteriori (ύστερη) πιθανότητα. Κατά τα άλλα, οι μέθοδοι **MAP1,MAP2**, είναι ίδιες με την μέθοδο **Mat4**. Μεταξύ τους διαφέρουν μόνο στην εκτιμώμενη πιθανότητα λάθους αντιστοιχίας (0.5 και 0.25 αντίστοιχα). Η μέθοδος MAPSAC βρίσκεται ενδιάμεσα από τους αλγορίθμους RANSAC και LMedS, σε σχέση με την υπολογιστική πολυπλοκότητα και με την ανοχή σε outliers. Το γεγονός αυτό μπορούμε να το παρατηρήσουμε από τον Πίνακα 5.7.

Επίσης, συμπληρώνουμε ότι το σφάλμα των μεθόδων **MAP1,MAP2** είναι ελκυστικά χαμηλό. Με κατάλληλη ρύθμιση των κατωφλιών T_l, T_u , σε συμφωνία με όσα προτείναμε σε προηγούμενη σχετική υποενότητα, έχουμε μέσο σφάλμα $\Delta f_1 = 5\%$ με την μέθοδο **MAP1** (Σύνολο δεδομένων Fountain-P11). Σχολιάζουμε επίσης τον μεγάλο χρόνο εκτέλεσης της μεθόδου **Kov3**. Δύο είναι τα κύρια αίτια:

- Εξαιτίας του μεγάλου ποσοστού outliers, η μέθοδος RANSAC δεν "συγκλίνει". Έτσι εκτελούνται πολλές επαναλήψεις και συχνά εξαντλείται το όριο ανώτατων επαναλήψεων που χρησιμοποιούμε στην υλοποίηση της μεθόδου RANSAC.
- Αντίστοιχα, σημαντικό για τον ρυθμό σύγκλισης των επαναληπτικών αριθμητικών μεθόδων βελτιστοποίησης, είναι η καλή αρχικοποίηση τους.

| Μέθοδος | Kov4 | Mat4 | Mat3 | FT | Kov3 | MAP1 | MAP2 |
|---------|------|------|-------|-------|--------|------|------|
| min. | 3.61 | 3.58 | 34.74 | 42.30 | 193.68 | 1.45 | 1.65 |

Πίνακας 5.7: Αποτελέσματα για εκτίμηση των μητρών F στο σύνολο δεδομένων Fountain-P11. Χρησιμοποιήσαμε κατώφλι $T_{inliers} = 30$ και $T_{ratio} = 0.4$



Σχήμα 5.16: Μέσο σφάλμα Δf_1 , ρυθμός αποτυχίας και διάμεσος του σφάλματος, για μεταβαλλόμενα T_u, T_l . Χρησιμοποιούμε την μέθοδο **Mat4** και χαρακτηριστικά SIFT.

Παρατηρούμε λοιπόν ότι η μέθοδος RANSAC αποτυγχάνει να αρχικοποιήσει σωστά την μέθοδο LM.

Βελτιωμένα Αποτελέσματα

Με την συζήτηση που προηγήθηκε, είμαστε πλέον σε θέση να παρουσιάσουμε βελτιωμένα αποτελέσματα. Παρουσιάζουμε τρεις περιπτώσεις, προκειμένου να συμπεριλάβουμε τις δύο επικρατέστερες μεθόδους υπολογισμού της μήτρας F (**FT** και **Mat4**). Αναλυτικά, οι μέθοδοι επίλυσης είναι οι ακόλουθες:

- **O1:** Ανίχνευση χαρακτηριστικών SIFT. Για το ταίριασμα των χαρακτηριστικών θέτουμε $T_{ratio} = 0.2$. Εκτίμηση Θεμελιώδους μήτρας με την μέθοδο **FT**. Θεωρούμε πιθανότητα σωστού ταιριάσματος 0.5. Για την αποδοχή η απόρριψη των μητρών F , χρησιμοποιούμε κατώφλι $T_l = 0.5$ και $T_u = 1$.
- **O2:** Ανίχνευση χαρακτηριστικών SIFT. Για το ταίριασμα των χαρακτηριστικών θέτουμε $T_{ratio} = 0.2$. Εκτίμηση Θεμελιώδους μήτρας με την μέθοδο **Mat4**. Για την αποδοχή η απόρριψη των μητρών F , χρησιμοποιούμε κατώφλι $T_l = 0.5$ και $T_u = 1$.
- **O3:** Ανίχνευση χαρακτηριστικών SIFT. Για το ταίριασμα των χαρακτηριστικών θέτουμε $T_{ratio} = 0.2$. Εκτίμηση Θεμελιώδους μήτρας με την μέθοδο **Mat4**. Για την αποδοχή η απόρριψη των μητρών F , χρησιμοποιούμε κατώφλι $T_l = 0.6$ και $T_u = 1$.

Τα αποτελέσματα παρουσιάζονται στους Πίνακες 5.8,5.9,5.10.

| | f-P11 | entry-P10 | HJ-P25 | castle-P19 | HJ-P8 | castle-P30 |
|---------------------|-------|-----------|--------|------------|-------|------------|
| O1 Μέση τιμή | 0.02 | 0.13 | 0.02 | 0.08 | 0.01 | 0.11 |
| O1 Διάμεσος | 0.01 | 0.09 | 0.02 | 0.05 | 0.01 | 0.03 |
| O1 Αποτυχίες | 0 | 2 | 0 | 0 | 0 | 0 |
| O2 Μέση τιμή | 0.09 | 0.14 | 0.02 | 0.10 | 0.03 | 0.06 |
| O2 Διάμεσος | 0.09 | 0.14 | 0.02 | 0.04 | 0.03 | 0.01 |
| O2 Αποτυχίες | 0 | 0 | 0 | 0 | 0 | 0 |
| O3 Μέση τιμή | 0.09 | 0.13 | 0.02 | 0.07 | 0.03 | 0.05 |
| O3 Διάμεσος | 0.09 | 0.11 | 0.02 | 0.04 | 0.03 | 0.02 |
| O3 Αποτυχίες | 0 | 0 | 0 | 1 | 0 | 1 |

Πίνακας 5.8: Σφάλμα στην εκτίμηση του εστιακού μήκους Δf . Δίνουμε επίσης και το πλήθος των φωτογραφιών στις οποίες αποτυγχάνει κάθε μέθοδος

| | f-P11 | entry-P10 | HJ-P25 | castle-P19 | HJ-P8 | castle-P30 |
|---------------------|-------|-----------|--------|------------|-------|------------|
| O1 Μέση τιμή | 0.18 | 0.53 | 0.15 | 0.81 | 0.10 | 0.27 |
| O1 Διάμεσος | 0.07 | 0.49 | 0.04 | 0.08 | 0.07 | 0.10 |
| O2 Μέση τιμή | 0.11 | 0.39 | 0.07 | 0.24 | 0.08 | 0.34 |
| O2 Διάμεσος | 0.03 | 0.26 | 0.03 | 0.06 | 0.07 | 0.09 |
| O3 Μέση τιμή | 0.11 | 0.39 | 0.07 | 0.24 | 0.08 | 0.34 |
| O3 Διάμεσος | 0.03 | 0.26 | 0.03 | 0.06 | 0.07 | 0.09 |

Πίνακας 5.9: Σφάλμα στην εκτίμηση της περιστροφής ΔR . Οι τιμές είναι σε μοίρες. Η παράμετρος T_f για την διαλογή των αξιόπιστων μητρών F τέθηκε στην τιμή 0.1.

| | f-P11 | entry-P10 | HJ-P25 | castle-P19 | HJ-P8 | castle-P30 |
|---------------------|-------|-----------|--------|------------|-------|------------|
| O1 Μέση τιμή | 2.35 | 7.35 | 2.47 | 14.94 | 1.94 | 6.77 |
| O1 Διάμεσος | 1.42 | 6.33 | 0.96 | 3.68 | 0.87 | 2.98 |
| O2 Μέση τιμή | 1.66 | 4.25 | 1.48 | 5.84 | 1.67 | 8.18 |
| O2 Διάμεσος | 1.27 | 3.68 | 0.78 | 3.40 | 0.87 | 4.02 |
| O3 Μέση τιμή | 1.66 | 4.25 | 1.48 | 5.84 | 1.67 | 8.18 |
| O3 Διάμεσος | 1.27 | 3.68 | 0.78 | 3.40 | 0.87 | 4.02 |

Πίνακας 5.10: Σφάλμα στην εκτίμηση της διεύθυνσης μετατόπισης Δt . Οι τιμές είναι σε μοίρες. Η παράμετρος T_f για την διαλογή των αξιόπιστων μητρών F τέθηκε στην τιμή 0.1.

Λογισμικό

Στον Πίνακα 5.11 παραθέτουμε χρήσιμο λογισμικό ανοιχτού κώδικα, το οποίο χρησιμοποιήσαμε για την εκτέλεση των πειραμάτων μας. Γίνεται ακόμα σύντομος σχολιασμός και δίνονται οι σχετικές βιβλιογραφικές αναφορές

| Πακέτο Λογισμικού | Αναφορές | Σχόλια |
|---|------------|--|
| MATLAB Functions for Multiple View Geometry | [90, 260] | Υλοποιήσεις αλγορίθμων γεωμετρίας πολλαπλών όψεων |
| Peter Kovese's Functions for Computer Vision and Image Processing | [117] | Υλοποιήσεις αλγορίθμων γεωμετρίας πολλών όψεων |
| Daisy features | [226] | |
| VLfeat | [239] | Συμπεριλαμβάνεται υλοποίηση των χαρακτηριστικών SIFT |
| FLANN | [156, 157] | Βιβλιοθήκη για δημιουργία δομών kd-trees και αναζήτηση σε αυτά με τον αλγόριθμο ANN. Συμπεριλαμβάνεται κατ' αλληλο interface για ρύθμιση των παραμέτρων των αλγορίθμων από ποιοτικούς χαρακτηρισμούς, χωρίς την ανάγκη προσδιορισμού τεχνικών λεπτομερειών [157, σ.22-23]. |
| Epipolar Geometry Toolbox | [181, 139] | Αλγόριθμοι υπολογισμού Θεμελιώδους πίνακα F . Συμπεριλαμβάνονται υλοποιήσεις πολλών αλγορίθμων και συναρτήσεων οπτικοποίησης των αποτελεσμάτων. |
| F-matrix Toolbox | [10, 11] | Υλοποιήσεις πληθώρας αλγορίθμων υπολογισμού μήτρας F . |
| Bill Trigg's functions | [232] | Διάφορες συναρτήσεις γεωμετρίας πολλών όψεων υλοποιημένες από τον Bill Triggs, όπως η μέθοδος βαθμονόμησης με εξισώσεις Kruppa. |

Πίνακας 5.11: Λογισμικό που χρησιμοποιήσαμε στην διεξαγωγή των πειραμάτων

6.1 Στόχος Κεφαλαίου

Σε αυτό το Κεφάλαιο, στοχεύουμε στην σύνθεση ενός συστήματος για την τελική ανακατασκευή της σκηνής. Με τον όρο ανακατασκευή της σκηνής εννοούμε την παράστασή της με τους εξής δύο τρόπους:

- ως αραιό "σύννεφο" σημείων του τρισδιάστατου χώρου. Η παράσταση αυτή είναι το αποτέλεσμα εφαρμογής συστημάτων Structure from Motion. Η ανάλυση τέτοιων συστημάτων έχει γίνει στο Κεφάλαιο 2 και συγκεκριμένα στην Ενότητα 2.1.6, όπου παρουσιάζονται και αποτελέσματα στην μορφή του αραιού σύννεφου σημείων.
- ως ανακατασκευασμένη συνεχή επιφάνεια. Η παράσταση αυτή προκύπτει από την παρεμβολή μιας επιφάνειας σε ένα (πιο πυκνό) σύννεφο σημείων, το οποίο έχει προκύψει συνήθως από σύστημα Multi-view Stereo (MvS). Σχετικές μεθόδους και τα αποτελέσματά τους παρουσιάσαμε στο Κεφάλαιο 3.

Σημειώνουμε επίσης, ότι προκειμένου να είναι εφικτή η ανακατασκευή της σκηνής, προσδιορίζουμε επιπλέον την θέση και τον προσανατολισμό λήψης των φωτογραφιών που χρησιμοποιούμε.

Για την ανακατασκευή της σκηνής, χρησιμοποιούμε:

- Τις εκτιμήσεις για τις εσωτερικές και τις εξωτερικές παραμέτρους των καμερών, που προκύπτουν με την διαδικασία που αναλύσαμε στο Κεφάλαιο 5. Σημειώνουμε ότι διαθέτουμε εκτιμήσεις για τις σχετικές θέσεις ζευγών φωτογραφιών, η σύνθεση των εκτιμήσεων αυτών όμως σε ενιαίο σύστημα συντεταγμένων δεν είναι εύκολη.
- Ένα μη-ακολουθιακό σύστημα SfM για την εκτίμηση της δομής (σημεία του χώρου) και της κίνησης (τοποθέτηση καμερών σε κοινό σύστημα συντεταγμένων).
- Ένα σύστημα στερέωσης πολλών όψεων (MVS) και ανακατασκευής επιφάνειας, προκειμένου να παράγουμε το τελικό μοντέλο.

Στην συνέχεια του Κεφαλαίου, θα περιγράψουμε αρχικά τα προηγούμενα βήματα και θα αναλύσουμε την μέθοδο που εφαρμόσαμε(Ενότητα 6.2). Η

περιγραφή των πειραμάτων που διεξήχθησαν γίνεται στην Ενότητα 6.3 και κάποια Αποτελέσματα παρουσιάζονται στην Ενότητα 6.4. Τέλος, ολοκληρώνουμε το Κεφάλαιο με προτάσεις βελτίωσης της μεθόδου (Ενότητα 6.6).

6.2 Περιγραφή Μεθόδου

Η σημαντικότερη ίσως επιλογή που καθορίζει και τις ιδιαιτερότητες της μεθόδου που εφαρμόσαμε, άφορα την επίλυση του προβλήματος SfM, οπότε και ξεκινάμε από αυτό. Οι μέθοδοι που περιγράψαμε στο Κεφάλαιο 2 (π.χ. [210, 207, 249]), είναι επαναληπτικές μέθοδοι. Ξεκινάνε από ένα ζευγάρι εικόνων με γνωστές εσωτερικές παραμέτρους και στην συνέχεια προσθέτουν επαναληπτικά περισσότερες όψεις και σημεία του χώρου στο πρόβλημα, χρησιμοποιώντας κάθε φορά για την επίλυση του υποπροβλήματος κατάλληλη ρουτίνα bundle adjustment. Η προσέγγιση αυτή έχει μερικά εμφανή μειονεκτήματα, αναφέρουμε ενδεικτικά

- την εξάρτηση του αποτελέσματος από το ζευγάρι εικόνων που αρχικοποιεί το σύστημα
- την επαναλαμβανόμενη επίλυση μεγάλων προβλημάτων βελτιστοποίησης με αριθμητικές μεθόδους
- την ύπαρξη πολλών τοπικών ελαχίστων στην συνάρτηση κόστους

Για αυτούς και άλλους λόγους, επιχειρήθηκε από πολλούς ([44, 168, 256, 170, 106, 87]) μια εναλλακτική προσέγγιση του προβλήματος. Η προσέγγιση αυτή βασίζεται στην επίλυση του προβλήματος SfM, με τις μήτρες περιστροφής των καμερών R_i γνωστές. Για αυτό το πρόβλημα έχουμε:

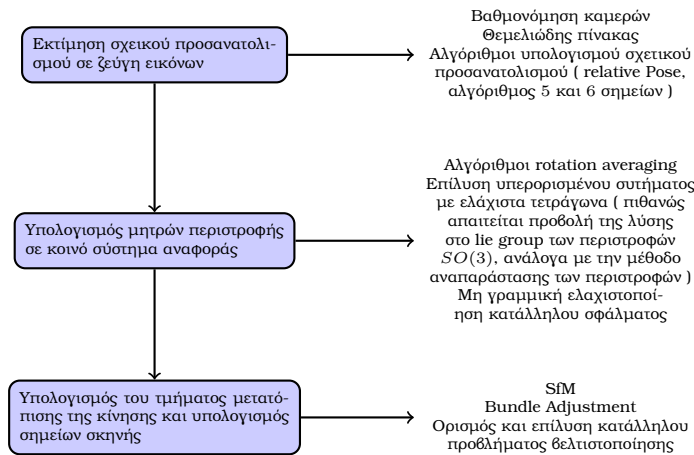
- Η συνάρτηση σφάλματος επαναπροβολής, συγκεκριμένα η συνάρτηση $d^2(\mathbf{x}_i, P\mathbf{X})$ είναι quasiconvex. Συνεπώς, το μέγιστο σφάλμα, ισοδύναμα η L_∞ νόρμα του διανύσματος με στοιχεία το σφάλμα κάθε μέτρησης, είναι επίσης quasiconvex.
- Quasiconvex συναρτήσεις έχουν μοναδικό ελάχιστο.
- Τα προβλήματα αυτά μπορούν να επιλυθούν αποτελεσματικά σαν μια ακολουθία SOCP προβλημάτων ή να προσεγγισθούν από γραμμικά προβλήματα.

Την δεύτερη αυτή προσέγγιση ακολουθήσαμε για την επίλυση του προβλήματος SfM. Στην συνέχεια θα περιγράψουμε τα στάδια ενός τέτοιου συστήματος και στην συνέχεια θα εξειδικεύσουμε περαιτέρω στο τελικό σύστημα που χρησιμοποιήσαμε.

6.2.1 SfM χωρίς την επαναληπτική εφαρμογή ΒΑ

Για τον προσδιορισμό της Δομής (σημείων του χώρου) και της κίνησης (θέση και προσανατολισμός καμερών), χωρίς την επαναληπτική εφαρμογή Β.Α. θα πρέπει να ακολουθήσουμε τα επόμενα βήματα

1. Προσδιορισμός των μητρών βαθμονόμησης K των καμερών.
2. Προσδιορισμός σχετικών περιστροφών R_i^j που περιγράφουν την σχετική θέση των καμερών i, j



Σχήμα 6.1: Οργάνωση του προβλήματος σε στάδια. Σχετικές τεχνικές και αλγόριθμοι.

3. Από τα 1,2 προσδιορίζουμε τις περιστροφές R_i σε κοινό σύστημα συντεταγμένων.
4. Σχηματισμός κατάλληλου προβλήματος βελτιστοποίησης και επίλυση του με τεχνικές (κυρτής) βελτιστοποίησης, για τον προσδιορισμό των θέσεων των καμερών και των σημείων του χώρου (SfM με γνωστές μήτρες περιστροφής).
5. Συνηθίζεται οι εκτιμήσεις των μεγεθών να βελτιώνονται με εκτέλεση B.A.

Παραδείγματα συστημάτων SfM που στηρίζονται στις προηγούμενες ιδέες μπορούν να βρεθούν στην βιβλιογραφία, ενδεικτικά αναφέρουμε [168, 166, 141, 202].

Τα συστήματα αυτά, θα πρέπει να αντιμετωπίσουν τις ακόλουθες προκλήσεις

- Την ευαισθησία της $\text{minimax} (L_\infty)$ διατύπωσης σε outliers. Λανθασμένες αντιστοιχίες σημείων μεταξύ των εικόνων, θα εμφανίσουν μεγάλο σφάλμα επαναπροβολής και επομένως θα καθορίσουν την τιμή της αντικειμενικής συνάρτησης στο πρόβλημα βελτιστοποίησης. Θα πρέπει να αποφύγουμε επομένως την προσαρμογή της λύσης σε αυτές τις μετρήσεις. Επίσης πρόβλημα αποτελούν οι εσφαλμένες επιπολικές γεωμετρίες που προσδιορίζουμε μεταξύ εικόνων που δεν παρατηρούν κοινό μέρος της σκηνής, οι οποίες δυσχεραίνουν την καταχώρηση των φωτογραφιών σε κοινό πλαίσιο αναφοράς.
- Η ταχύτητα της μεθόδου και η κλιμάκωση της με το μέγεθος της εισόδου (αριθμός φωτογραφιών). Μέθοδοι απόρριψης outliers που έχουν προταθεί (π.χ [141, 200]) οδηγούν στην επαναληπτική επίλυση του προβλήματος βελτιστοποίησης και συνεπώς δεν είναι κατάλληλες για μεγάλο αριθμό μετρήσεων. Σε άλλες κατευθύνσεις, έχει επιχειρηθεί η βελτιστοποίηση αλγορίθμων επίλυσης προβλημάτων SOCP και η προσέγγιση του αρχικού προβλήματος με γραμμικά προβλήματα (LP)[196, 5, 169].

6.2.2 Περιγραφή και ανάλυση της μεθόδου που υλοποιήσαμε

Η μέθοδος αυτή είναι βασισμένη στο [168]. Στην συνέχεια περιγράφουμε τα διάφορα στάδια της μεθόδου αυτής και τις τροποποιήσεις που εφαρμόσαμε. Μερικές λεπτομέρειες της αρχικής μεθόδου δεν αναπτύσσονται, όσο αυτές διατηρήθηκαν αναλλοίωτες, καθώς μπορούν να αναζητηθούν στις βιβλιογραφικές αναφορές που δίνονται.

Εκτίμηση μητρών εσωτερικών παραμέτρων K^i των καμερών

Για τις μήτρες εσωτερικών παραμέτρων των καμερών, κάνουμε την παραδοχή ότι περιγράφονται με τις παραμέτρους f (εστιακό μήκος), p_x, p_y (συντεταγμένες πρωτεύοντος σημείου). Την παραδοχή αυτή την έχουμε σχολιάσει σε άλλα σημεία της εργασίας και αναφέρουμε απλά ότι είναι επαρκώς δικαιολογημένη ενώ έχει επικρατήσει στην βιβλιογραφία.

Η πλειονότητα των μεθόδων SfM προχωράει επίσης στην επόμενη παραδοχή για το εστιακό μήκος f

- Το f θεωρείται γνωστό, π.χ. από τα EXIF Tags. Ανάλογα με την μέθοδο (sequential/non-sequential) η παραδοχή αυτή γίνεται για ένα υποσύνολο των φωτογραφιών ή για όλες της φωτογραφίες [207, 168].
- Το f θεωρείται άγνωστο, αλλά κοινό σε όλες τις φωτογραφίες [141]. Η παραδοχή αυτή επιτρέπει την εφαρμογή αξιόπιστων, ρωμαλέων και αποδοτικών μεθόδων για την εύρεση του εστιακού μήκους [214]. Σημειώνουμε ότι είναι έγκυρη παραδοχή στην περίπτωση που το σύνολο των φωτογραφιών έχει ληφθεί από την ίδια κάμερα και χωρίς την χρήση ζουμ.
- Το f προσδιορίζεται με μεθόδους που δεν χρησιμοποιούν μόνο αντιστοιχίες σημείων. Για παράδειγμα μπορούν να χρησιμοποιηθούν μέθοδοι που εντοπίζουν vanishing points (π.χ. [202]). Στην περίπτωση αυτή όμως το περιεχόμενο της σκηνής επηρεάζει την δυνατότητα επιτυχούς προσδιορισμού των f .

Στην περίπτωση μας, επιλέξαμε να μην προχωρήσουμε σε επιπλέον παραδοχές. Έτσι υποθέσαμε άγνωστο και διαφορετικό f για κάθε εικόνα του συνόλου δεδομένων. Ακόμα, χρησιμοποιήσαμε αποκλειστικά αντιστοιχίες σημείων στις εικόνες. Για την επίλυση του προβλήματος, χρησιμοποιήσαμε το σύστημα που αναπτύξαμε και παρουσιάσαμε στο Κεφάλαιο 5. Υπενθυμίζουμε ότι ταυτόχρονα με εκτιμήσεις για τα εστιακά μήκη f^i των καμερών, λαμβάνουμε και εκτιμήσεις για την σχετική θέση $[R \ t]$ των καμερών. Αυτές οι σχετικές περιστροφές R^{ij} μπορούν να χρησιμοποιηθούν στα επόμενα στάδια της ανακατασκευής ή να απορριφθούν και να εφαρμόσουμε διαφορετικές μεθόδους για τον υπολογισμό τους.

Τέλος, για λόγους πληρότητας, υπενθυμίζουμε ότι η θέση του πρωτεύοντος σημείου λαμβάνεται στο κέντρο της εικόνας. Αφού ανακτήσαμε τις εσωτερικές παραμέτρους των καμερών, βρισκόμαστε πλέον στο σημείο αφετηρίας των μεθόδων SfM. Τα επόμενα στάδια βασίζονται σημαντικά στην υλοποίηση του [167].

Προσδιορισμός μητρών σχετικής περιστροφής R^{ij} για ζεύγη φωτογραφιών

Αρχικά, θέλουμε να προσδιορίσουμε για ζεύγη φωτογραφιών τις σχετικές τους θέσεις, συγκεκριμένα την σχετική περιστροφή R^{ij} που συνδέει την θέση λήψης της φωτογραφίας i και της φωτογραφίας j . Οι εκτιμήσεις αυτές μπορούν να ληφθούν μαζί με τις εκτιμήσεις για τα f_i από το σύστημα που περιγράψαμε στο Κεφάλαιο 5. Εναλλακτικά, μπορούμε να εφαρμόσουμε, σε συμφωνία και με την προσέγγιση του [168], τον αλγόριθμο 5 σημείων για την εκτίμηση της σχετικής θέσης βαθμονομημένων καμερών [213], χρησιμοποιώντας τις εκτιμήσεις f_i που διαθέτουμε.

Καταχώρηση μητρών περιστροφής σε κοινό σύστημα συντεταγμένων

Η καταχώρηση των μητρών σχετικής περιστροφής R^{ij} σε κοινό σύστημα (Συμβολισμός: R_i, R_j για τις κάμερες i, j αντίστοιχα) βασίζεται στην ακόλουθη σχέση

$$R_i = R^{ij} R_j \quad (6.1)$$

Για τον σκοπό αυτό, κατασκευάζουμε σταθμισμένο γράφο $G(V, E)$. Οι κόμβοι V αντιστοιχούν στις εικόνες του συνόλου δεδομένων. Για κάθε σχετική περιστροφή που έχουμε υπολογίσει, προσθέτουμε ακμή ij μεταξύ των αντίστοιχων κόμβων του γράφου. Το βάρος της ακμής είναι ο αριθμός των inliers στις αντιστοιχίες των εικόνων i, j , για την εκτίμηση R^{ij} που διαθέτουμε.

Στον γράφο αυτό εκτελούμε μια RANSAC διαδικασία, δειγματοληπώντας spanning trees. Από την Σχέση 6.1, προκύπτει ότι ένα spanning tree επιτρέπει τον μονοσήμαντο καθορισμό των περιστροφών R^i . Η πιθανότητα επιλογής μιας ακμής στο spanning tree είναι ανάλογη του βάρους της ακμής. Για να εκτιμήσουμε την ποιότητα της λύσης, εφαρμόζουμε ξανά την Σχέση 6.1, για τις περιστροφές (ακμές) που δεν συμπεριελήφθησαν στο spanning tree, και τις κατατάσσουμε ανάλογα με το σφάλμα ως inliers ή outliers.

Σημειώνουμε ότι, ενώ σε εικόνες με μικρό baseline (απόσταση μετατόπισης στις σχετικές θέσεις λήψης) η εκτίμηση της διεύθυνσης μετατόπισης είναι δύσκολη, η εκτίμηση της σχετικής περιστροφής γίνεται με ακρίβεια. Συνεπώς η τάση να διαλέγουμε εικόνες με μικρό baseline και κατά συνέπεια πολλές αντιστοιχίες, βελτιώνει την εκτίμηση των περιστροφών.

Σημειώνουμε ότι για την καταχώρηση των σχετικών περιστροφών σε κοινό σύστημα συντεταγμένων, η εφαρμογή αλγορίθμων γράφων είναι αρκετά συνηθισμένη. Παραπέμπουμε τους αναγνώστες στην σειρά δημοσιεύσεων του Govindu [79, 80].

Γενικά για το πρόβλημα, το οποίο είναι γνωστό στην βιβλιογραφία και ως rotation averaging (στάθμιση περιστροφών), παραπέμπουμε στο [89], το οποίο παρέχει μια πλήρη περιγραφή, καλύπτοντας τους διαφορετικούς τρόπους παράστασης των περιστροφών (πίνακες, quaternions), τις μετρικές σφάλματος και τους αλγόριθμους επίλυσης του προβλήματος.

Διατύπωση προβλήματος βελτιστοποίησης και επίλυση του

Μπορούμε τώρα να διατυπώσουμε το πρόβλημα SfM με γνωστές περιστροφές και να προχωρήσουμε στην επίλυση του. Χρησιμοποιούμε τον ακόλουθο

συμβολισμό

\mathbf{x}_i : Σημείο εικόνας (μέτρηση)

\mathbf{X} : Σημείο του χώρου

Για το τετράγωνο του σφάλματος επαναπροβολής, $d^2(\mathbf{x}_i, \mathbf{X})$, θεωρώντας πίνακα προβολής $[R \ t]$ έχουμε

$$d^2(\mathbf{x}_i, \mathbf{X}) = \left\| \left(x_i^1 - \frac{R_1 \mathbf{X} + t^1}{R_3 \mathbf{X} + t^3}, x_i^2 - \frac{R_2 \mathbf{X} + t^2}{R_3 \mathbf{X} + t^3} \right) \right\|^2 \quad (6.2)$$

Τα σημεία που είναι ορατά σε μία εικόνα, πρέπει να βρίσκονται μπροστά από αυτήν (Cheirality). Συνεπώς στο πρόβλημα ισχύουν και οι (γραμμικοί) περιορισμοί

$$R_3 \mathbf{X} + t^3 > 0 \quad (6.3)$$

Κάνουμε μια μικρή παρένθεση, για να ξεκαθαρίσουμε λίγο και την Θεωρία που σκιαγραφήσαμε στην αρχή της Ενότητας 6.2. Οι Σχέσεις 6.2, 6.3, μαζί με το Λήμμα 12 που παραθέτουμε αμέσως μετά, δείχνουν την quasiconvexity του τετραγώνου του σφάλματος επαναπροβολής και συνεπώς την quasiconvexity της L_∞ νόρμας του τετραγώνου των σφαλμάτων.

Λήμμα 12 (Λήμμα 6 [170]). *Αν η συνάρτηση $w_i : S \mapsto \mathbb{R}$ είναι κυρτή και η συνάρτηση $v_i : S \mapsto \mathbb{R}$ είναι κοίλη τότε η συνάρτηση $R_i(x) = \frac{w_i(x)}{v_i(x)}$ είναι pseudoconvex (έννοια ισχυρότερη της quasiconvex) στο $S = \{x | v_i(x) > 0\}$*

Για να διατυπώσουμε τώρα βολικά το πρόβλημα, θεωρούμε ότι το μέγιστο σφάλμα είναι μικρότερο μιας σταθεράς γ και έχουμε να λύσουμε ένα SOCP (συγκεκριμένα να δούμε να είναι εφικτό) με περιορισμούς τις σχέσεις

$$\left\| \left((x_i^1 R_3 - R_1) \mathbf{X} + x_i^1 t^3 - t^1, (x_i^2 R_3 - R_2) \mathbf{X} + x_i^2 t^3 - t^2 \right) \right\| \leq \gamma (R_3 \mathbf{X} + t^3) \quad (6.4)$$

Ο προσδιορισμός του ελάχιστου γ πραγματοποιείται με την επίλυση σειράς SOCP και είναι η βασική προσέγγιση στο SfM με την L_∞ νόρμα [106, 199, 88].

Αντί να λύσουμε αυτό το πρόβλημα, με στόχο τον περιορισμό της επίδρασης των outliers, προσθέτουμε slack variables (μία για κάθε περιορισμό-μέτρηση) και έχουμε στην θέση των 6.4

$$\left\| \left((x_i^1 R_3 - R_1) \mathbf{X} + x_i^1 t^3 - t^1, (x_i^2 R_3 - R_2) \mathbf{X} + x_i^2 t^3 - t^2 \right) \right\| \leq \gamma (R_3 \mathbf{X} + t^3) + s_i \quad (6.5)$$

θα θέλαμε να βελτιστοποιήσουμε ως προς τον αριθμό των μη-μηδενικών slack variables (την L_0 νόρμα του \mathbf{s}), όμως το πρόβλημα αυτό δεν είναι επιλύσιμο εύκολα. Για αυτό ελαχιστοποιούμε την L_1 νόρμα του \mathbf{s} δηλαδή έχουμε το κριτήριο βελτιστοποίησης

$$\min \sum_i s_i$$

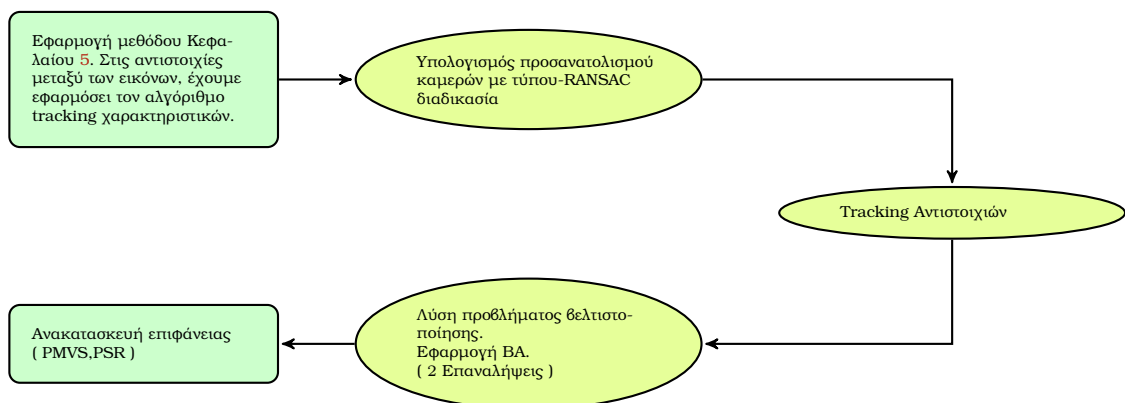
Σημειώνουμε ότι, ενώ έχουμε το όφελος της σθεναρότητας σε outliers και της επίλυσης ενός (και όχι μιας σειράς) προβλημάτων βελτιστοποίησης, σε σχέση με την βασική προσέγγιση, αυτό έρχεται με το κόστος του εκ των προτέρων προσδιορισμού του γ , δηλαδή της "ακτίνας" των inliers.

Μια αρκετά παρόμοια παρουσίαση, γίνεται στο [44], μαζί με την απόδειξη ότι η λύση είναι MAP εκτιμητής για συγκεκριμένες πρότερες κατανομές στις μετρήσεις και στους outliers. Η μόνη διαφορά είναι η θέση των slack variables είναι στον χώρο των "αντικειμένων", δηλαδή μέσα στην νόρμα της εξίσωσης 6.4. Οι διαφορές των μεθόδων αυτών εξηγούνται (γεωμετρικά) στο [256].

Ανακεφαλαίωση συστήματος και επιπλέον χαρακτηριστικά

Συνοπτικά, τα βήματα που ακολουθήσαμε είναι

1. Ανίχνευση χαρακτηριστικών SIFT και ταίριασμα των χαρακτηριστικών.
2. Οι αντιστοιχίες μεταξύ των εικόνων εμπλουτίζονται ενώ απορρίπτονται εσφαλμένες αντιστοιχίες, με χρήση αλγόριθμου γράφων. Περισσότερες λεπτομέρειες δίνονται στο [168]. Αναφέρουμε συνοπτικά ότι ο αλγόριθμος παρουσιάζει ομοιότητες με τον αλγόριθμο του Prim για την εύρεση του MST. Οι κόμβοι του γράφου αντιστοιχούν στα σημεία των εικόνων (χαρακτηριστικά), οι ακμές σε ταιριάσματα χαρακτηριστικών και τα βάρη τους στον συνολικό αριθμό αντιστοιχιών των σχετικών εικόνων.
3. Στην συνέχεια προσδιορίζουμε τις εσωτερικές παραμέτρους των εικόνων με την μέθοδο του Κεφαλαίου 5. Ταυτόχρονα προσδιορίζονται και οι μήτρες σχετικών περιστροφών R^{ij} . Για τον προσδιορισμό των σωστών μητρών (υπενθυμίζουμε ότι βρίσκουμε δύο λύσεις σε κάθε ζεύγος) χρησιμοποιούμε τις συνθήκες Cheirality.
4. Εναλλακτικά, προσδιορίζουμε τις μήτρες σχετικών περιστροφών με τον αλγόριθμο των 5 σημείων [213], ακολουθούμενο από BA.
5. Προσδιορίζουμε τις μήτρες περιστροφών των καμερών σε κοινό σύστημα αναφοράς, με χρήση του αλγορίθμου δειγματοληψίας σε spanning trees που περιγράψαμε νωρίτερα.
6. Εφαρμόζουμε ξανά τον αλγόριθμο του βήματος 2. Αυτήν την φορά οι αντιστοιχίες που χρησιμοποιούνται (κόμβοι του γράφου) και αντίστοιχα τα βάρη των ακμών, προκύπτουν από τους inliers στις σχετικές γεωμετρίες που έχουμε εκτιμήσει.
7. Λύνουμε το πρόβλημα βελτιστοποίησης και βελτιώνουμε την λύση με την εφαρμογή BA. Στην συνέχεια, χρησιμοποιώντας τις νέες εκτιμήσεις των περιστροφών, επαναλαμβάνουμε το βήμα ξανά.
8. Τέλος αφαιρούμε σημεία του χώρου, για τα οποία υπάρχει μεγάλη αβεβαιότητα στο βάθος τους (σε σχέση με κάποια από τις κάμερες στις



Σχήμα 6.2: Τα στάδια της μεθόδου που ακολουθήσαμε, από την καταχώριση των περιστροφών σε κοινό πλαίσιο έως την εκτίμηση της τελικής επιφάνειας

οποίες είναι ορατό). Το κριτήριο βασίζεται στην δεύτερη παράγωγο του σφάλματος επαναπροβολής, στην διεύθυνση του κέντρου της κάμερας.

9. Χρησιμοποιούμε τις θέσεις των καμερών που βρήκαμε ως δεδομένα εισόδου στο σύστημα στερέωσης πολλών όψεων PMVS [66, 69].
10. Το πυκνό σύννεφο σημείων που μας δίνει η μέθοδος PMVS χρησιμοποιείται ως αρχικοποίηση για την παρεμβολή συνεχούς επιφάνειας με την μέθοδο Poisson Surface Reconstruction [110, 109].

6.3 Πειράματα

Σε αυτή την ενότητα των πειραμάτων, χρησιμοποιήσαμε το σύνολο δεδομένων HJ-P8 που παρουσιάσαμε στο Κεφάλαιο 5. Με στόχο την καλύτερη γενίκευση των αποτελεσμάτων, δεν μεταβάλλαμε το κατώφλι T_{ratio} για το ταίριασμα των χαρακτηριστικών, αλλά το αφήσαμε σε μια τυπική τιμή (0.5).

Διαλογή αξιόπιστων ανακατασκευών 2 όψεων και εκτίμηση f

Αρχικά, θα σχολιάσουμε την επιλογή της σωστής από τις δύο λύσεις για την μήτρα $[R \ t]$ που λαμβάνουμε ως λύση, όταν ακολουθούμε την προσέγγιση του Κεφαλαίου 5.

Χρησιμοποιήσαμε την συνθήκη ότι τα σημεία του χώρου που είναι ορατά από μια κάμερα, πρέπει να βρίσκονται μπροστά από αυτήν (Cheirality). Επιλέξαμε την λύση για την οποία τα περισσότερα σημεία ικανοποιούν αυτή την συνθήκη. Από την ανάλυση που κάναμε στο Κεφάλαιο 5, περιμένουμε οι λύσεις να διαχωρίζονται στις περισσότερες των περιπτώσεων εύκολα. Το γεγονός αυτό επαληθεύθηκε στην πράξη. Συγκεκριμένα:

- Στις περισσότερες των περιπτώσεων, μόνο μία από τις δύο λύσεις αντιστοιχεί σε κάμερα που έχει σημεία "μπροστά της".
- Στις περιπτώσεις που δεν ισχύει η προηγούμενη περίπτωση, παρατηρήσαμε μεγάλα σφάλματα στις εκτιμήσεις των εστιακών μηκών f_1, f_2 . Σε αυτές τις περιπτώσεις, εσφαλμένες αντιστοιχίες σημείων έχουν επικρατήσει στην εκτίμηση του μοντέλου, με αποτέλεσμα αυτό να μην ανταποκρίνεται στην σκηνή.

Συνεπώς, παρατηρήσαμε ότι η συγκεκριμένη συνθήκη, διαχωρίζει έντονα τις δύο λύσεις. Ακόμα, διαπιστώσαμε ότι είναι ένας αποτελεσματικός τρόπος για τον εντοπισμό εσφαλμένων γεωμετριών και ότι η εκτίμηση του εστιακού μήκους αποτυγχάνει, όταν οι "εξωτερικές" παράμετροι $[R \ t]$ δεν προσδιορίζονται σωστά.

Υπενθυμίζουμε σε αυτό το σημείο, ότι τα σύνολα δεδομένων που χρησιμοποιήσαμε απαρτίζονται από γεωμετρικές σκηνές με σημαντικές επαναλαμβανόμενες δομές και συμμετρία, ιδιότητες που, όπως είναι προφανές, δυσχεραίνουν σημαντικά το ταίριασμα των χαρακτηριστικών.

Επίσης, σχετικά με την χρήση του κριτηρίου Cheirality, παρατηρούμε ότι δίνει καλύτερα αποτελέσματα σε σχέση με τις ευριστικές μεθόδους που χρησιμοποιήθηκαν στο Κεφάλαιο 5. Το αποτέλεσμα αυτό είναι σημαντικό, λόγω της προφανούς γεωμετρικής σημασίας της συνθήκης αυτής αλλά και της απλότητας της.

Εφαρμογή Bundle Adjustment στις εκτιμήσεις των σχετικών θέσεων 2 όψεων

Επειδή η υλοποίηση του [168] εφαρμόζει στις εκτιμήσεις των μητρών σχετικής περιστροφής R^{ij} που προκύπτουν από τον αλγόριθμο 5 σημείων Bundle Adjustment, δοκιμάσαμε να ακολουθήσουμε και εμείς την αντίστοιχη στρατηγική. Σημειώνουμε ότι τα εστιακά μήκη f_i θέτονται σταθερά στην βελτιστοποίηση με BA.

Παρατηρήσαμε ότι η εφαρμογή Bundle Adjustment βελτιώνει λίγο τα Αποτελέσματα (τις εκτιμήσεις για τα R, t), αρκεί να έχουμε κάνει καλή εκτίμηση του μοντέλου (υπολογιστικά απαιτητικές μέθοδοι υπολογισμού F-matrix, μη γραμμική μέθοδος τριγωνισμού των σημείων). Δικαιολογούμε αυτή την παρατήρηση ως προσαρμογή στις εσφαλμένες αντιστοιχίες, οι οποίες δικαιολογούνται από μια πρόχειρη εκτίμηση του μοντέλου.

Μέθοδος υπολογισμού Θεμελιώδους μήτρας F και ποιότητα αποτελεσμάτων

Για τον υπολογισμό της μήτρας F , μπορούμε να χρησιμοποιήσουμε κλιμακούμενης υπολογιστικής πολυπλοκότητας μεθόδους. Το σημείο αυτό αναπτύχθηκε στο Κεφάλαιο 5, ενώ πιο συγκεκριμένη αναφορά κάνουμε και αργότερα στην Ενότητα. Παρατηρήσαμε ότι η ακρίβεια στον υπολογισμό των μητρών F μεταβάλλεται σύμφωνα με την ακρίβεια στον υπολογισμό των θέσεων των καμερών, αποτέλεσμα λογικό και αναμενόμενο. Όμως, η χρήση διαφορετικών μεθόδων διαλογής των αξιόπιστων ανακατασκευών (για παράδειγμα $\mathbf{h}_3, \mathbf{h}_4$ που ορίζουμε παρακάτω), ενδέχεται να υποδεικνύει και την χρήση συγκεκριμένης μεθόδου υπολογισμού του πίνακα F .

Ακρίβεια υπολογισμού σχετικής θέσης καμερών με την μέθοδο του Κεφαλαίου 5

Αν εφαρμόσουμε την μέθοδο που αναπτύξαμε στο Κεφάλαιο 5, λαμβάνουμε τόσο εκτιμήσεις για τα εστιακά μήκη f_i όσο και για την σχετική θέση ζευγών καμερών R^{ij} . Δοκιμάσαμε να χρησιμοποιήσουμε τις εκτιμήσεις αυτές ή να χρησιμοποιήσουμε μόνο τις εκτιμήσεις f_i και να επανεκτιμήσουμε τον σχετικό προσανατολισμό με τον αλγόριθμο των 5 σημείων.

Παρατηρήσαμε ότι οι δύο αυτές προσεγγίσεις, δίνουν κοντινά μεταξύ τους αποτελέσματα, όταν εφαρμόζουμε Bundle Adjustment στις εκτιμήσεις των R^{ij} ή τουλάχιστον ακολουθούμε μια περισσότερο ακριβή μέθοδο προσδιορισμού της μήτρας F . Επίσης, οι μέθοδοι βρίσκονται πιο κοντά όσον αφορά το σφάλμα ΔR , ενώ αποκλίνουν περισσότερο σε σχέση με το σφάλμα Δt , εξαιτίας και της συσσώρευσης των προηγούμενων σφαλμάτων που προηγούνται του υπολογισμού των t .

Περιγραφή πειραμάτων που διεξήχθησαν και παραμέτρων που εξετάστηκαν

Συγκεντρωτικά, εξετάσαμε τις παρακάτω παραμέτρους της μεθόδου επίλυσης του προβλήματος

- Την μέθοδο προσδιορισμού του f_i από τις διάφορες διαθέσιμες εκτιμήσεις και διαλογής των αξιόπιστων από τις γεωμετρίες που υπολογίσαμε

(σχετική θέση 2 όψεων). Συγκεκριμένα συμβολίζουμε:

- \mathbf{h}_1 : Τρόπος διαλογής όπως στο Κεφάλαιο 5.
 - \mathbf{h}_2 : Απόρριψη όλων των ανακατασκευών για τις οποίες αποτυγχάνει το κριτήριο Cheiraity πριν την εκτίμηση της τιμής του f_i και εφαρμογή του τρόπου διαλογής \mathbf{h}_1 .
 - \mathbf{h}_3 : Απόρριψη όλων των ανακατασκευών για τις οποίες αποτυγχάνει το κριτήριο Cheiraity πριν την εκτίμηση της τιμής του f_i .
 - \mathbf{h}_4 : Απόρριψη όλων των ανακατασκευών για τις οποίες αποτυγχάνει το κριτήριο Cheiraity πριν την εκτίμηση της τιμής του f_i . Μετά την εκτίμηση του f_i , απορρίπτουμε επιπλέον τις ανακατασκευές με σφάλμα Δf μεγαλύτερο από 0.1. Το κριτήριο αυτό οδηγεί στην απόρριψη ανακατασκευών που είναι αμφίβολες, με αποτέλεσμα να ανακατασκευάζονται λιγότερες κάμερες και να μειώνεται το σφάλμα. Σημειώνουμε ότι διαφορετικές τιμές του κατωφλίου απόρριψης μπορεί να οδηγήσουν σε καλύτερη ρύθμιση του ισοζυγίου απόρριψη καμερών-αύξηση σφάλματος.
- Τρόπος Υπολογισμού Θεμελιώδους Πίνακα F . Χρησιμοποιήσαμε τις μεθόδους **FT** (Συμβ. **F**) και **MAT44** (Συμβ. **M**) του Κεφαλαίου 5.
 - Εφαρμογή Bundle Adjustment για την βελτίωση των εκτιμήσεων μας R, \mathbf{t} των σχετικών θέσεων 2 εικόνων (Συμβ. **BA**).
 - Για την ανανέωση των έγκυρων αντιστοιχιών μεταξύ ζευγαριών εικόνων, αφού εκτιμήσουμε κάποια σχετική γεωμετρία $[R \ \mathbf{t}]$, κάνουμε τριγωνοποίηση των σημείων:
 - Γραμμικά, με χρήση της μεθόδου DLT (Συμβ. **L**).
 - Μη-γραμμικά, με βελτιστοποίηση του σφάλματος επαναπροβολής (Συμβ. **N**).

Απορρίπτουμε στην συνέχεια τις αντιστοιχίες του ζεύγους εικόνων, για τις οποίες το σημείο του χώρου

- Βρίσκεται πίσω από κάποια από τις δύο κάμερες (Συμβ. \mathbf{v}_1).
- Βρίσκεται πίσω από κάποια κάμερα ή έχει μεγάλο σφάλμα επαναπροβολής (Συμβ. \mathbf{v}_2).

Στον Πίνακα 6.1, παρουσιάζουμε τα αποτελέσματα που λάβαμε χρησιμοποιώντας τις εκτιμήσεις για τα f_i από τις μεθόδους **O1**, **O2**, **O3** του Κεφαλαίου 5 (δίνονται στον Πίνακα 5.8). Στα αποτελέσματα, παρατηρούμε ότι οι εκτιμήσεις των \mathbf{t} δεν είναι τόσο επιτυχείς, όταν οι εικόνες του συνόλου έχουν μικρό baseline. Επαληθεύουμε δηλαδή, ένα γνωστό συμπέρασμα. Τέλος, στον Πίνακα 6.2 παρουσιάζουμε αποτελέσματα που λάβαμε από την χρήση των εκτιμήσεων f_i, R^{ij} του συστήματος του Κεφαλαίου 5.

| | castle-P19 | entry-P10 | castle-P30 | HJ-P8 | HJ-P25 | f-P11 |
|----------------------|------------|-----------|------------|-------|--------|-------|
| Μέση τιμή ΔR | 2.46 | 3.16 | 1.11 | 0.39 | 0.44 | 0.47 |
| Μέση τιμή Δt | 9.12 | 12.79 | 5.55 | 1.22 | 1.47 | 2.85 |

Πίνακας 6.1: Σφάλμα (σε μοίρες) στις εκτιμήσεις της θέσης και του προσανατολισμού των καμερών. Σημειώνουμε ότι τα σφάλματα αναφέρονται στις τελικές περιγραφές $[R \ t]$ που ανήκουν σε κοινό σύστημα συντεταγμένων. Τα σύνολα δεδομένων περιγράφονται στο Κεφάλαιο 5

| | castle-P19 | entry-P10 | castle-P30 | HJ-P8 | HJ-P25 | f-P11 |
|--------------------------------|-----------------------|------------------|----------------------|------------------|------------------|----------------------|
| Μέθοδος | h_3, FT, BA, L, v_2 | h_4, M, L, v_1 | h_3, M, BA, N, v_2 | h_3, M, L, v_2 | h_3, F, L, v_2 | h_4, M, BA, L, v_2 |
| Μέση τιμή ΔR | 1.80 | 1.06 | 6.39 | 0.66 | 1.25 | 0.50 |
| Μέση τιμή Δt | 10.69 | 6.50 | 18.73 | 2.25 | 1.61 | 1.36 |
| Μέση τιμή Δf | 0.09 | 0.06 | 0.18 | 0.03 | 0.05 | 0.02 |
| Κάμερες που ανακατασκευάζονται | 12 | 9 | 28 | 8 | 25 | 11 |

Πίνακας 6.2: Σφάλμα (σε μοίρες) στις εκτιμήσεις της θέσης και του προσανατολισμού των καμερών και σφάλμα Δf . Σημειώνουμε ότι τα σφάλματα αναφέρονται στις τελικές περιγραφές $[R \ t]$ που ανήκουν σε κοινό σύστημα συντεταγμένων. Τα σύνολα δεδομένων περιγράφονται στο Κεφάλαιο 5

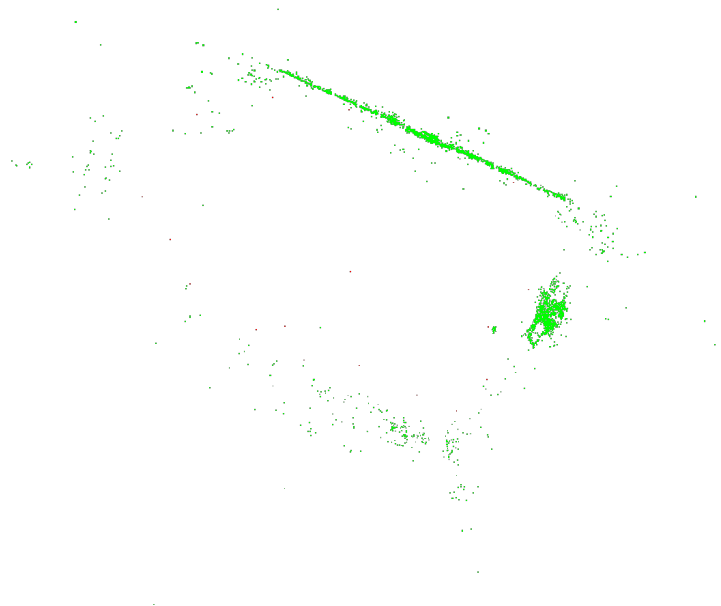
6.4 Παρουσίαση Αποτελεσμάτων

Στην ενότητα αυτή παρουσιάζουμε (οπτικά) Αποτελέσματα για τα διάφορα σύνολα φωτογραφιών που χρησιμοποιήθηκαν στο Κεφάλαιο 5. Τα Αποτελέσματα που πείραμε εφαρμόζοντας τις μεθόδους που εισάγαμε είναι διαθέσιμα στο youtube [144].

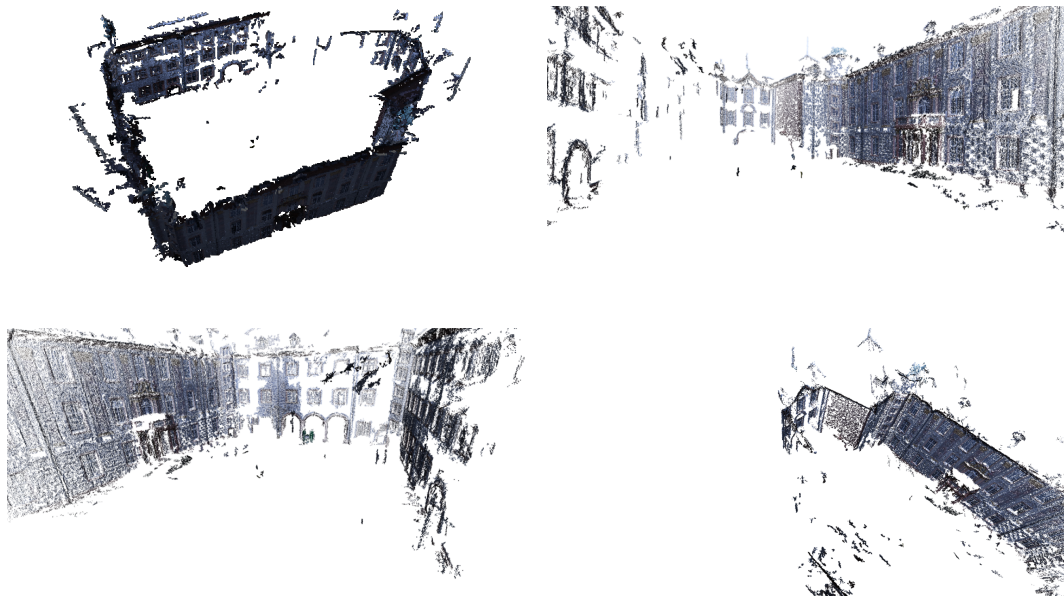
Σύνολο δεδομένων castle-P19



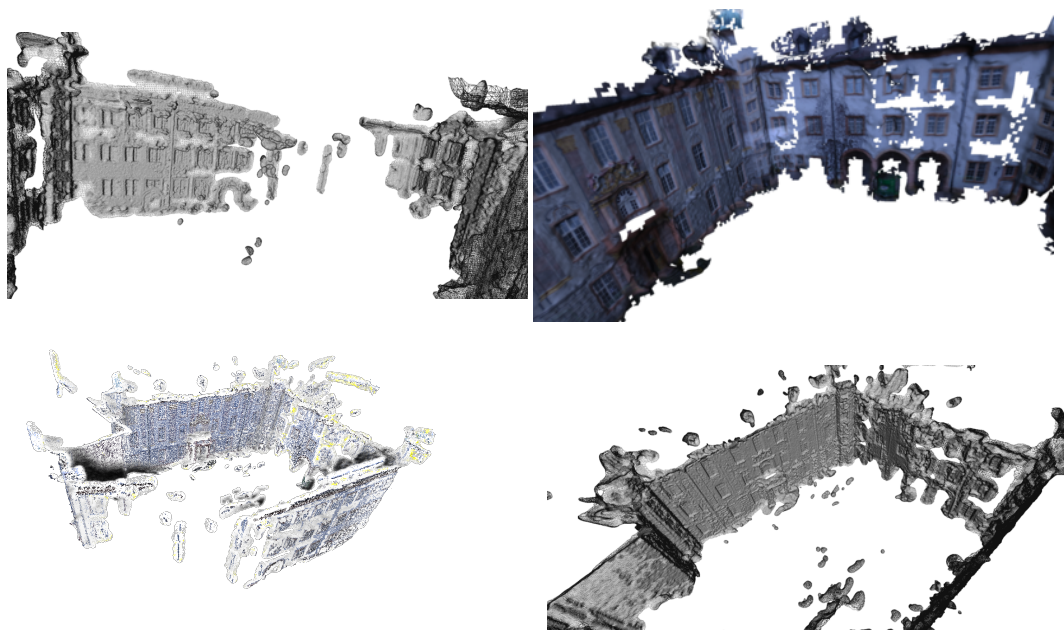
Σχήμα 6.3: Εικόνες συνόλου δεδομένων



Σχήμα 6.4: Άραιο σύννεφο σημείων-έξοδος SfM. Τα κόκκινα σημεία αντιστοιχούν στις θέσεις των καμερών



Σχήμα 6.5: Πυκνό σύννεφο σημείων, έξοδος MvS

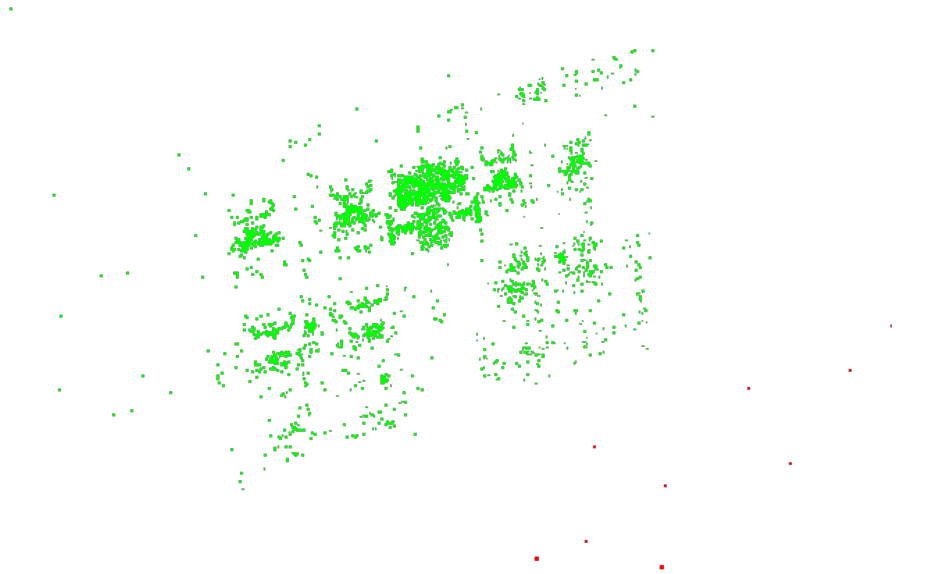


Σχήμα 6.6: Τελική ανακατασκευή της επιφάνειας, όπως προέκυψε από τον αλγόριθμο PSR

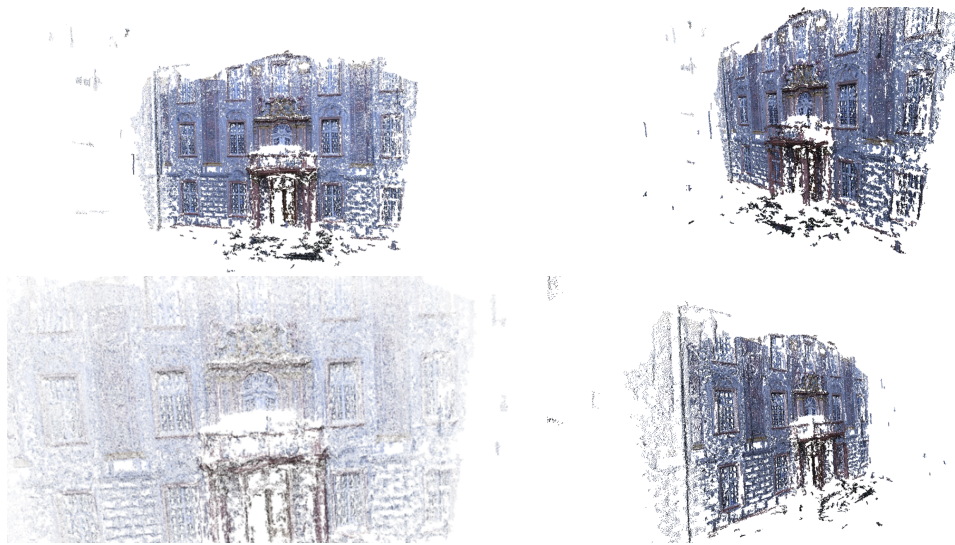
Σύνολο δεδομένων entry-P10



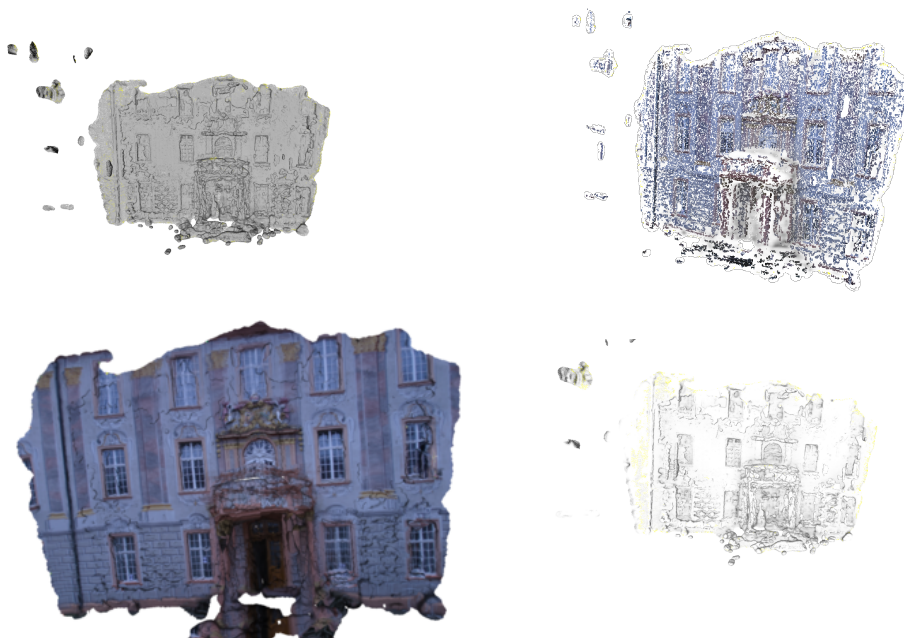
Σχήμα 6.7: Εικόνες συνόλου δεδομένων



Σχήμα 6.8: Άραιο σύννεφο σημείων-έξοδος SfM. Τα κόκκινα σημεία αντιστοιχούν στις θέσεις των καμερών



Σχήμα 6.9: Πυκνό σύννεφο σημείων, έξοδος MvS

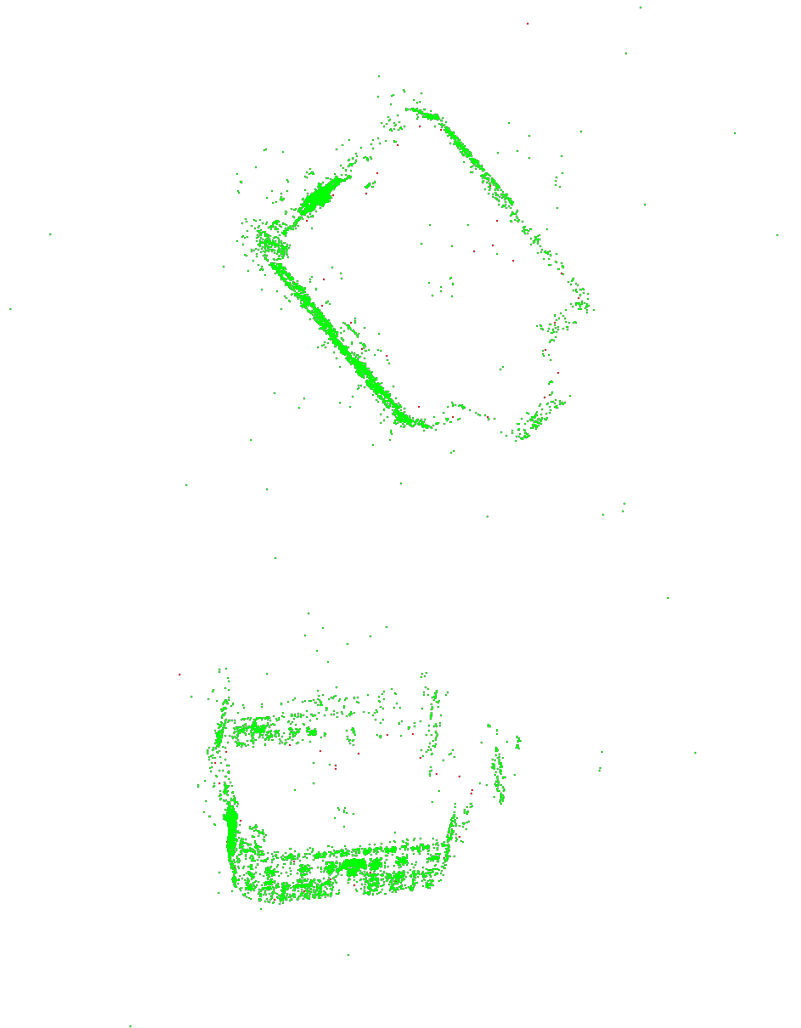


Σχήμα 6.10: Τελική ανακατασκευή της επιφάνειας, όπως προέκυψε από τον αλγόριθμο PSR

Σύνολο δεδομένων castle-P30



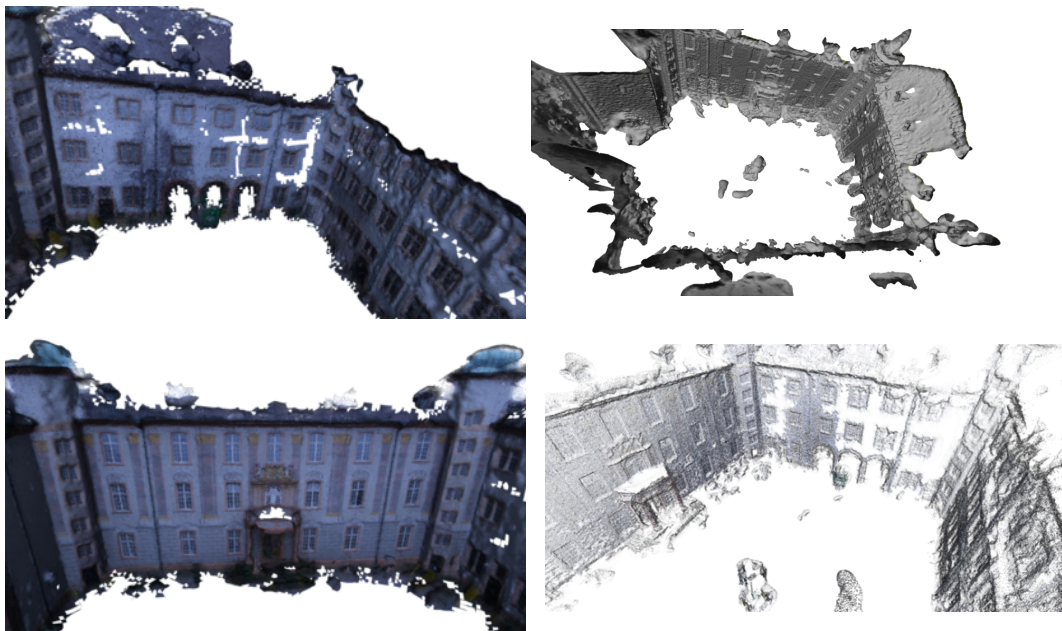
Σχήμα 6.11: Εικόνες συνόλου δεδομένων



Σχήμα 6.12: Άραιο σύννεφο σημείων-έξοδος SfM. Τα κόκκινα σημεία αντι-στοιχούν στις θέσεις των καμερών



Σχήμα 6.13: Πυκνό σύννεφο σημείων, έξοδος ΜνS

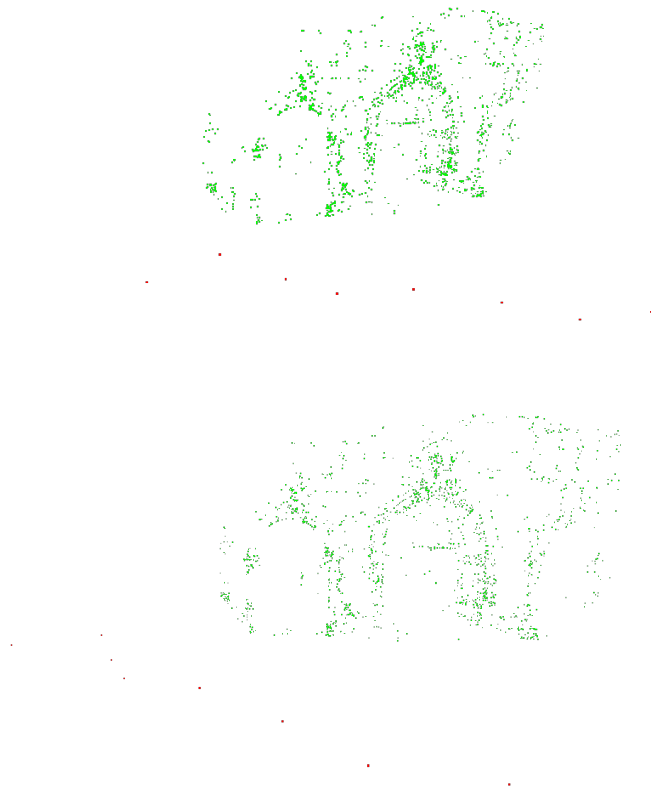


Σχήμα 6.14: Τελική ανακατασκευή της επιφάνειας, όπως προέκυψε από τον αλγόριθμο PSR

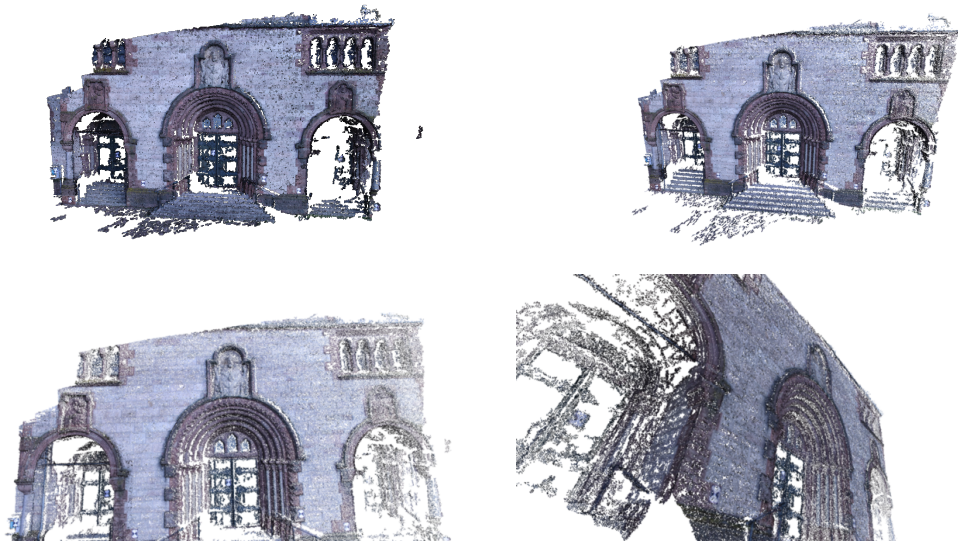
Σύνολο δεδομένων HJ-P8



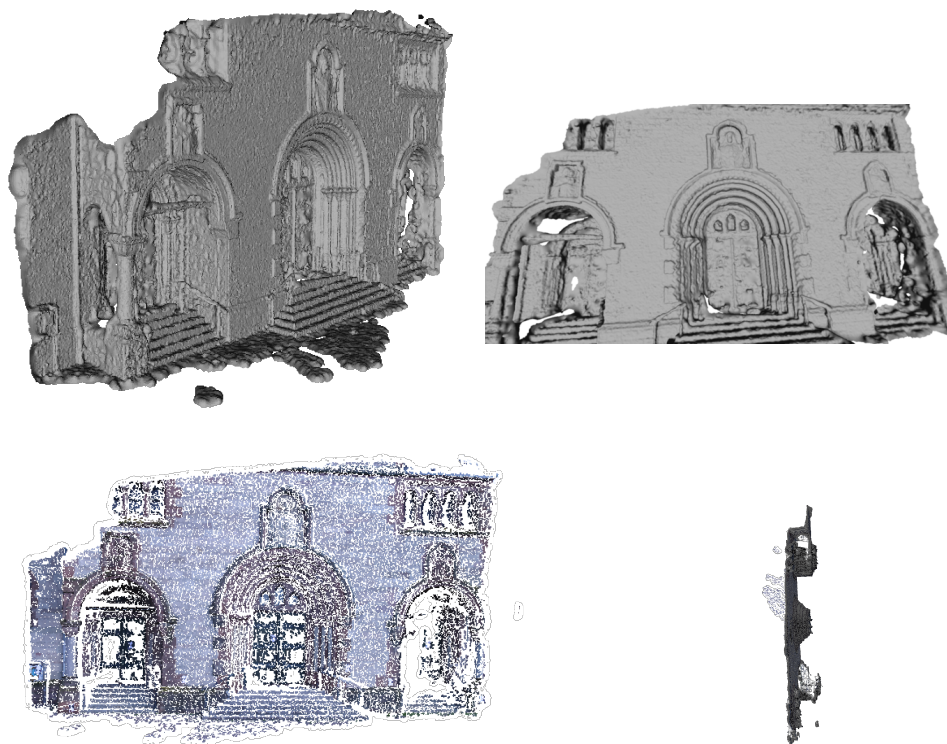
Σχήμα 6.15: Εικόνες συνόλου δεδομένων



Σχήμα 6.16: Άραιο σύννεφο σημείων-έξοδος SfM. Τα κόκκινα σημεία αντιστοιχούν στις θέσεις των καμερών



Σχήμα 6.17: Πυκνό σύννεφο σημείων, έξοδος ΜνS

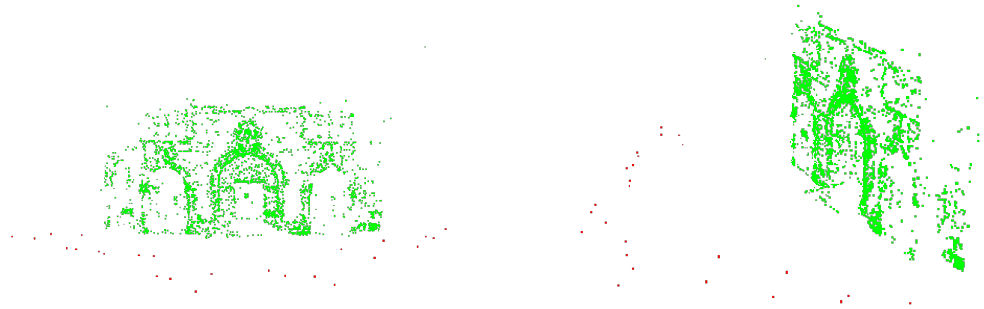


Σχήμα 6.18: Τελική ανακατασκευή της επιφάνειας, όπως προέκυψε από τον αλγόριθμο PSR

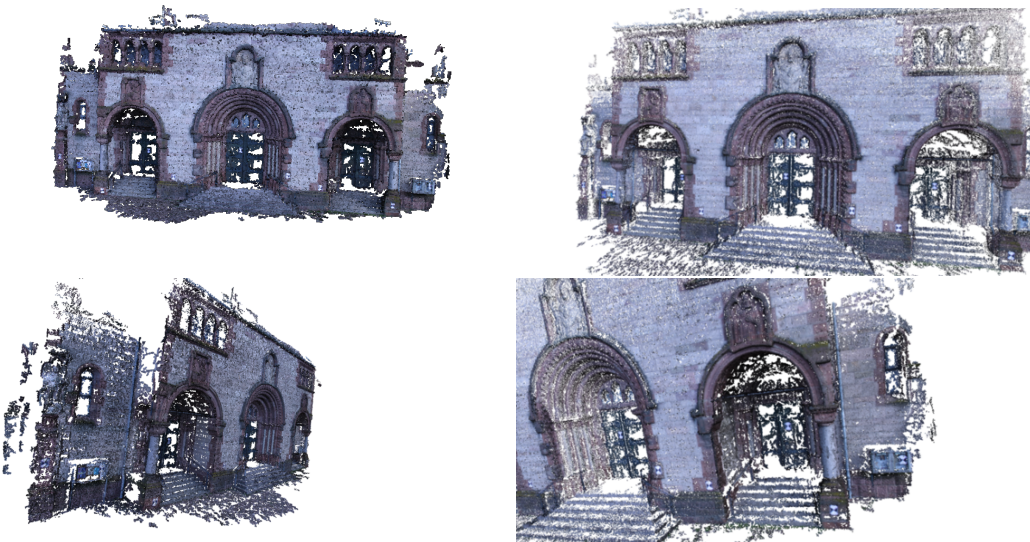
Σύνολο δεδομένων HJ-P25



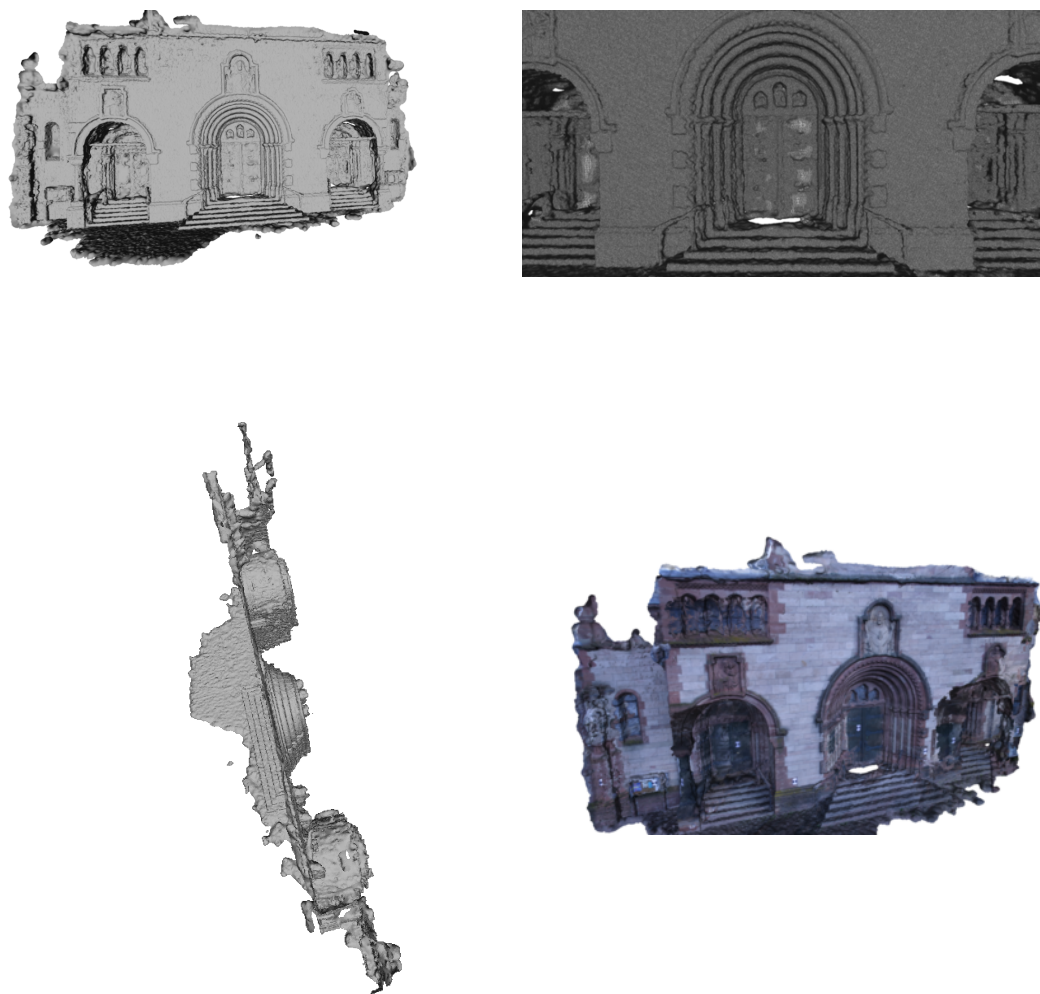
Σχήμα 6.19: Εικόνες συνόλου δεδομένων



Σχήμα 6.20: Άραιο σύννεφο σημείων-έξοδος SfM. Τα κόκκινα σημεία αντιστοιχούν στις θέσεις των καμερών



Σχήμα 6.21: Πυκνό σύννεφο σημείων, έξοδος MvS

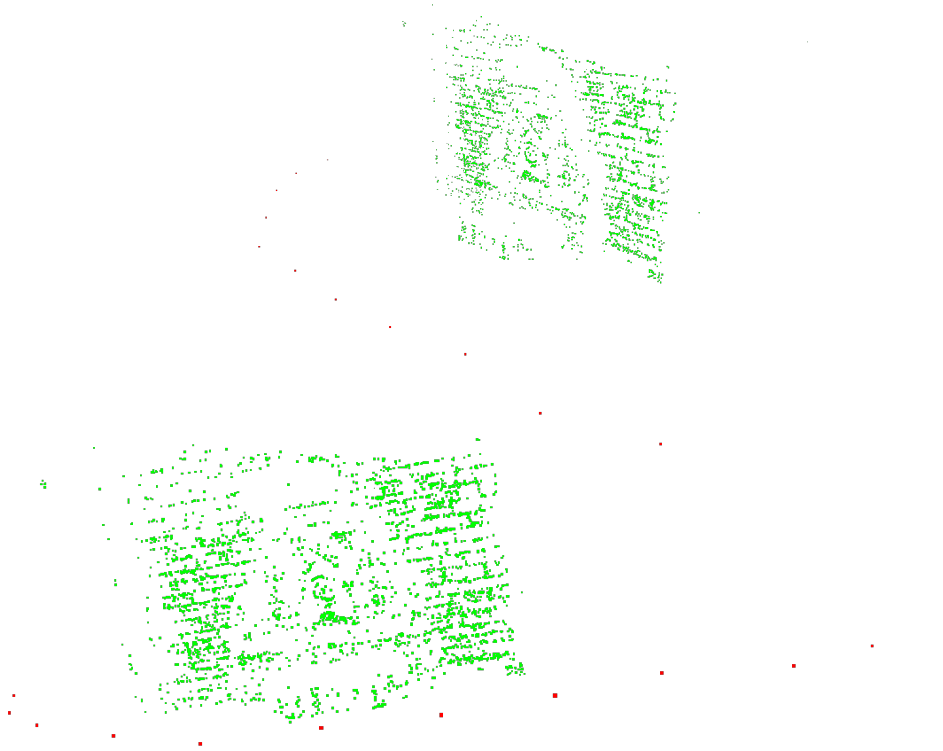


Σχήμα 6.22: Τελική ανακατασκευή της επιφάνειας, όπως προέκυψε από τον αλγόριθμο PSR

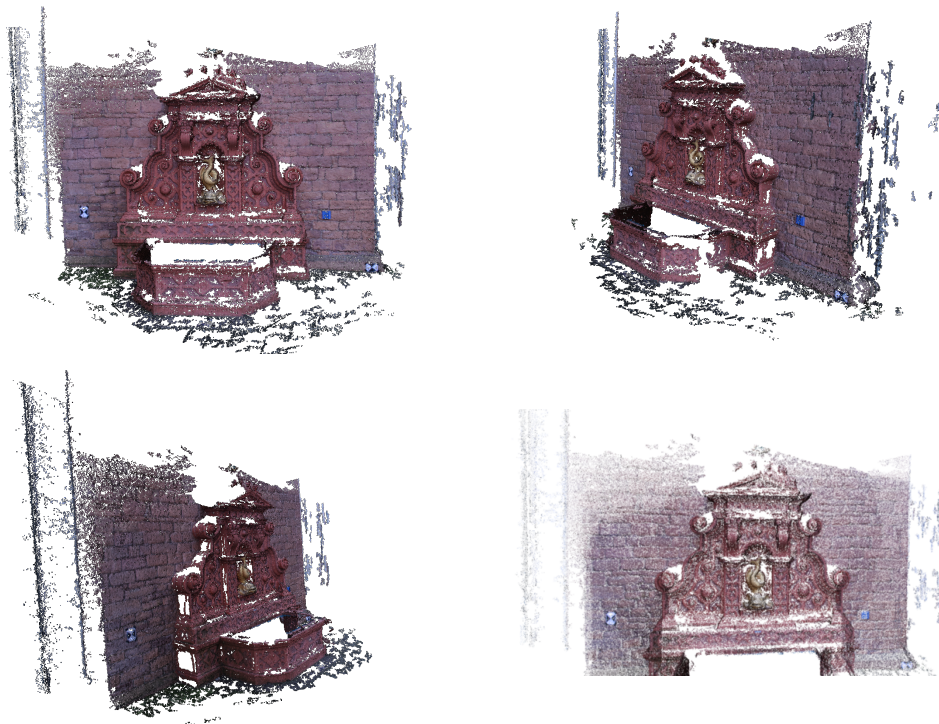
Σύνολο δεδομένων f-P11



Σχήμα 6.23: Εικόνες συνόλου δεδομένων



Σχήμα 6.24: Άραιο σύννεφο σημείων-έξοδος SfM. Τα κόκκινα σημεία αντιστοιχούν στις θέσεις των καμερών



Σχήμα 6.25: Πυκνό σύννεφο σημείων, έξοδος MvS



Σχήμα 6.26: Τελική ανακατασκευή της επιφάνειας, όπως προέκυψε από τον αλγόριθμο PSR

6.5 Εφαρμογή της μεθόδου στην Ακρόπολη

Στην ενότητα αυτή παρουσιάζουμε (οπτικά) ανακατασκευές μνημείων της Ακρόπολης των Αθηνών. Οι φωτογραφίες που χρησιμοποιήθηκαν έχουν ληφθεί με χαμηλού κόστους φωτογραφική μηχανή και δεν χρησιμοποιήθηκε ειδικός εξοπλισμός. Περιορισμοί στην πρόσβαση καθώς και εργασίες συντήρησης και αναστήλωσης, παρακωλύουν την επαρκή (φωτογραφική) κάλυψη των μνημείων, με αποτέλεσμα τα ανακατασκευασμένα μοντέλα να μην είναι πλήρη. Τα Αποτελέσματα είναι επίσης διαθέσιμα στο youtube [144].

Καρυάτιδες



Σχήμα 6.27: Αποτελέσματα ανακατασκευής (ως πύκνο σύννεφο σημείων και ως textured επιφάνεια) για τις Καρυάτιδες.

Ερέχθειο



Σχήμα 6.28: Αποτελέσματα ανακατασκευής (ως πύκνο σύννεφο σημείων και ως textured επιφάνεια) για το Ερέχθειο.

Παρθενώνας



Σχήμα 6.29: Αποτελέσματα ανακατασκευής (ως πύκνο σύννεφο σημείων και ως textured επιφάνεια) για τον Παρθενώνα.

Άλλες τοποθεσίες



Σχήμα 6.30: Μερική ανακατασκευή μνημείου του Λυσικράτη



Σχήμα 6.31: Μερική ανακατασκευή μνημείου Φιλοπάππου

6.6 Προτάσεις για βελτίωση και ανάπτυξη του αλγορίθμου

Ανίχνευση χαρακτηριστικών και ταίριασμα εικόνων

Στα πειράματα που πραγματοποιήσαμε σε αυτό το Κεφάλαιο, χρησιμοποιήσαμε χαρακτηριστικά SIFT σε περιοχές DoG. Η επιλογή αυτή είναι από τις πιο συνηθισμένες στην βιβλιογραφία. Για να βελτιώσουμε τα αρχικά χαρακτηριστικά που ανιχνεύονται στις εικόνες, αλλά και να βελτιώσουμε το εύρος εφαρμογής της μεθόδου, μπορούμε:

- Να χρησιμοποιήσουμε διαφορετικά είδη (περισσότερα από ένα) περιοχών, με σκοπό την αξιόπιστη ανίχνευση περιοχών ενδιαφέροντος σε μεγάλο εύρος σκηνών. Για παράδειγμα είναι γνωστό ότι οι περιοχές MSER έχουν καλύτερη επίδοση σε δομημένες σκηνές, ενώ οι περιοχές Hessian-Affine σε περιοχές με έντονη υφή.
- Αν και η χρήση βελτιστοποιημένων χαρακτηριστικών(π.χ.[243]) θα βελτιώσει τα αποτελέσματα, θεωρούμε περισσότερο ενδιαφέρουσα την χρήση ανιχνευτών που ξεφεύγουν από το μοντέλο "histograms of oriented gradients" του SIFT. Μια ενδιαφέρουσα, πρόσφατη εξέλιξη, είναι τα χαρακτηριστικά LIOP (Local Intensity Order Pattern [259]). Πιστεύουμε ότι η εφαρμογή τους στο πρόβλημα, μπορεί να δώσει ενδιαφέροντα αποτελέσματα.

Το ταίριασμα των εικόνων και στην συνέχεια το ταίριασμα των χαρακτηριστικών, είναι μια σημαντική στενωπός του προβλήματος και δυσχεραίνει καθοριστικά την δυνατότητα εφαρμογής της μεθόδου σε μεγάλα σύνολα φωτογραφιών. Για να επιταχύνουμε την διαδικασία, μπορούμε να ακολουθήσουμε κάποιες από τις επόμενες κατευθύνσεις

- Να χρησιμοποιήσουμε το μοντέλο Bag-of-Visual-Words και τεχνικές αντίστοιχες της μεθόδου tf-idf (γνωστές από την ανάκτηση κειμένων). Αναφέρουμε συνοπτικά, ότι η αναπαράσταση των εικόνων γίνεται με "ιστογράμματα" οπτικών λέξεων, και η σύγκριση εικόνων με σύγκριση ιστογραμμάτων. Οι οπτικές λέξεις είναι παράσταση ομοειδής με τα χαρακτηριστικά. Διαισθητικά, δύο περιοχές ενδιαφέροντος είναι όμοιες, όταν περιγράφουν την ίδια οπτική λέξη.
- Χρήση περιγραφέων για ολόκληρες εικόνες (π.χ. GIST [165]) και σύγκριση αυτών για την εύρεση σχετικών εικόνων. Η ιδέα αυτή εφαρμόζεται και στην δημοφιλή δημοσίευση [58].
- Επειδή κάθε εικόνα μπορεί να ταιριαστεί επιτυχώς (ζευγαρωθεί) με έναν απεριόριστο αριθμό (άλλων) εικόνων, θα μπορούσαμε να ελέγξουμε το μέγεθος του προβλήματος θέτοντας περιορισμό στον μέγιστο αριθμό ζευγαριών που μία εικόνα μπορεί να συμμετέχει. Αν θεωρήσουμε τον γράφο με κόμβους τις εικόνες και ακμές μεταξύ των ζευγαρωμένων κόμβων, θέλουμε ουσιαστικά να θέσουμε όριο στον μέγιστο βαθμό των κόμβων. Ιδέες για την υλοποίηση μπορούν να ληφθούν από την βιβλιογραφία [202, 72].
- Πρόσφατες δημοσιεύσεις σχετικές με την ανάλυση video, προτείνουν γρήγορες μεθόδους ταιριάσματος εικόνων [183]. Προτείνουμε την μελέτη της δυνατότητας αξιοποίησης αυτής της μεθόδου.

- Τέλος, περισσότερες ιδέες μπορούν να αναζητηθούν στην σχετική βιβλιογραφία [91, 58].

Επίσης, θα μπορούσαμε να δοκιμάσουμε διαφορετικό αλγόριθμο tracking χαρακτηριστικών σε πολλές εικόνες. Με την τωρινή προσέγγιση, μια ακολουθία ταιριασμένων χαρακτηριστικών θεωρείται ασυνεπής μόνο όταν περιέχει δύο χαρακτηριστικά από την ίδια εικόνα. Όμως, διαφορετικές έννοιες της συνέπειας μπορούν να χρησιμοποιηθούν. Για παράδειγμα μπορούμε να ελέγξουμε την ιδιότητα

$$(\alpha \otimes \beta) \wedge (\beta \otimes \gamma) \Rightarrow \alpha \otimes \gamma$$

με το σύμβολο \otimes συμβολίζουμε την αντιστοιχισιο-ταίριασμα χαρακτηριστικών και με ελληνικά γράμματα συμβολίζουμε χαρακτηριστικά εικόνων.

Για την βελτίωση των αντιστοιχιών και τον περιορισμό των εσφαλμένα ταιριασμένων χαρακτηριστικών, μπορούμε να χρησιμοποιήσουμε την ιδέα της γεωμετρικής επαλήθευσης των χαρακτηριστικών που ταιριάζαμε. Στην ιδέα αυτή, χρησιμοποιούμε ένα απλοποιητικό γεωμετρικό μοντέλο μεταξύ των εικόνων (π.χ μετασχηματισμός ομοιότητας μεταξύ επιπέδων) το οποίο μπορεί να προσδιοριστεί από μικρό αριθμό αντιστοιχιών (στην περίπτωση του Μετασχηματισμού ομοιότητας μόνο μία αντιστοιχία) και ελέγχουμε αν οι υπόλοιπες αντιστοιχίες μπορούν να εξηγηθούν με το μοντέλο αυτό. Ένα παράδειγμα εφαρμογής του γεωμετρικού αυτού ελέγχου δίνεται στο [208].

Αναφέρουμε τέλος την δυνατότητα επιτάχυνσης της ανίχνευσης των χαρακτηριστικών, με χρήση για παράδειγμα χαρακτηριστικών SURF. Η αντικατάσταση των κυτιοειδών φίλτρων των ανιχνευτών SURF με αυτά που περιγράφονται στο [118] παρουσιάζει επιπλέον ενδιαφέρον.

Εκτίμηση μητρών περιστροφής για ζεύγη εικόνων και υπολογισμός των περιστροφών σε κοινό σύστημα συντεταγμένων

Ξεκινώντας από την εκτίμηση των μητρών περιστροφής για ζεύγη εικόνων, παρατηρούμε ότι ο υπολογισμός του Θεμελιώδους πίνακα F έχει καθοριστική σημασία στην ακρίβεια των αποτελεσμάτων. Για τον υπολογισμό του θεμελιώδους πίνακα λοιπόν σημειώνουμε

- την εξερεύνηση της δυνατότητας υπολογισμού του F-matrix με επίλυση SOCP προβλημάτων, αντίστοιχα με την επίλυση του προβλήματος SfM με γνωστές τις μήτρες περιστροφών. Σημειώνουμε ότι θα πρέπει είτε να προηγηθεί αξιόπιστος προσδιορισμός των inliers είτε να γίνει διαμόρφωση προβλήματος βελτιστοποίησης με slack variables, προκειμένου να χειριστούμε κατάλληλα τους outliers.
- αντί του υπολογισμού των f_1, f_2, R, t από την τελική μήτρα F , μπορούμε να χρησιμοποιήσουμε μια RANSAC τύπου διαδικασία στην οποία θα προσδιορίζουμε τις τιμές των f, R, t σε κάθε επανάληψη. Σημειώνουμε ότι 7 αντιστοιχείς αρκούν για τον υπολογισμό αυτό. Στην συνέχεια οι τελικές τιμές των f_1, f_2 μπορούν να προκύψουν με τον υπολογισμό κάποιου σταθμισμένου (από τους inliers) μέσου και του R από κατάλληλο rotation averaging αλγόριθμο [89]. Υπενθυμίζουμε ότι η εκτίμηση για το t δεν χρησιμοποιείται στα επόμενα στάδια του αλγορίθμου.

Για τον προσδιορισμό των μητρών περιστροφής σε κοινό πλαίσιο συντεταγμένων, αλλά και για τον προσδιορισμό των μητρών σχετικής περιστροφής,

Θεωρούμε σημαντική την εφαρμογή του αλγορίθμου του Weiszfeld, όπως αυτός παρουσιάζεται στην πρόσφατη δημοσίευση [86].

Τέλος, ενδιαφέρον παρουσιάζει η δυνατότητα αντικατάστασης πιθανοτικών αλγορίθμων όπως ο RANSAC και οι παραλλαγές του, με πρόσφατες προσεγγίσεις προσδιορισμού του βέλτιστου συνόλου inliers [19].

Υπολογισμός Δομής (σημείων του χώρου) και κίνησης (θέσεις καμερών)

Structure from Motion

Η πρώτη μας παρατήρηση, αφορά τον προσδιορισμό των σημείων του χώρου (αραιό σύννεφο σημείων). Με το συγκεκριμένο σύστημα MVS (PMVS) που χρησιμοποιούμε, τα σημεία αυτά δεν χρησιμοποιούνται πουθενά. Συνεπώς θα μπορούσαμε να παραλείψουμε τον υπολογισμό τους ολοκληρωτικά, ακολουθώντας κάποια προσέγγιση επηρεασμένη από το [199]. Σημειώνουμε όμως ότι τα σημεία του χώρου είναι απαραίτητα για το clustering των όψεων (π.χ CMVS, βλ. Κεφάλαιο 3), όποτε θα πρέπει να αναζητηθούν διαφορετικές μέθοδοι για τον χωρισμό της εισόδου σε υποσύνολα. Ο χωρισμός αυτός είναι σημαντικός για την επιτάχυνση των υπολογισμών σε μεγάλα σύνολα φωτογραφιών.

Ένα λεπτό σημείο στην διαμόρφωση του προβλήματος βελτιστοποίησης, είναι η κατανόηση της επόμενης διπλής επιλογής

- Επιλογής νόρμας για το σφάλμα στο διάνυσμα των μετρήσεων. Μέτρηση είναι κάθε αντιστοιχία σημείων μεταξύ εικόνων. Οι επιλογές που εξετάσαμε ήταν L_2 νόρμα (κλασσική προσέγγιση) και L_∞ νόρμα.
- Επιλογή νόρμας για το σφάλμα κάθε μέτρησης, δηλαδή για το σφάλμα επαναπροβολής. Για παράδειγμα μπορεί να χρησιμοποιηθεί η L_2 νόρμα (όπως στην μέθοδο που ακολουθήσαμε), η ανά-συντεταγμένη νόρμα-άπειρο $L_{\infty,c} = \max(\delta_x, \delta_y)$ ([44, 169]) κλπ.

Μια διαφωτιστική παρουσίαση του προηγούμενου θέματος δίνεται στο [44]. Πρόσφατη έρευνα [12], ενθαρρύνει την χρήση της περικεκομμένης (truncated) L_2 νόρμας για το σφάλμα επαναπροβολής (σφάλμα κάθε μέτρησης). Προτείνουμε την διερεύνηση της δυνατότητας επαναδιατύπωσης του προβλήματος βελτιστοποίησης, ώστε να χρησιμοποιείται η truncated- L_2 νόρμα των μετρήσεων.

Τέλος, παρατηρήσαμε την δυσκολία που προκαλούν επαναλαμβανόμενες δομές και συμμετρίες. Τα χαρακτηριστικά αυτά είναι συχνά σε αρχιτεκτονικές σκηνές. Πρόσφατη έρευνα για την αντιμετώπιση των παραπάνω φαινομένων [101, 184], μπορεί να αξιοποιηθεί κατάλληλα.

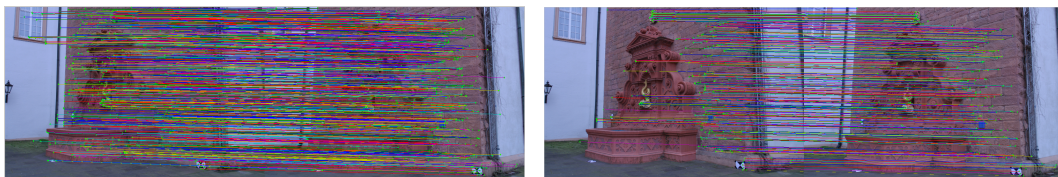
ΒΕΛΤΙΩΣΕΙΣ ΣΤΗΝ ΒΑΣΙΚΗ ΠΡΟΣΕΓΓΙΣΗ: ΓΕΩΜΕΤΡΙΚΗ ΕΠΑΛΗΘΕΥΣΗ ΑΝΤΙΣΤΟΙΧΙΩΝ ΕΙΚΟΝΩΝ & ΒΕΛΤΙΩΣΕΙΣ ΣΤΗΝ ΕΚΤΙΜΗΣΗ ΤΩΝ ΣΧΕΤΙΚΩΝ ΠΕΡΙΣΤΡΟΦΩΝ

7.1 Γεωμετρική επαλήθευση των αρχικών αντι- στοιχιών

7.1.1 Γενικά για την Γεωμετρική Επαλήθευση

Κατά το αρχικό ταίριασμα των διαθέσιμων εικόνων, βασιστήκαμε στην *τοπική εμφάνιση* των εικόνων για να εξάγουμε περιγραφείς (π.χ SIFT) τους οποίους στην συνέχεια ταιριάζαμε για να καταλήξουμε στις αρχικές αντιστοιχίες μεταξύ ζευγών εικόνων.

Εξαιτίας της τοπικής φύσης της προηγούμενης διαδικασίας, είναι σύνηθες να καταλήγουμε σε αντιστοιχίες εικόνων οι οποίες είναι εσφαλμένες και οφείλονται στην ομοιότητα στην τοπική εμφάνιση των περιοχών που αντιστοιχούνται. Για να απορρίψουμε τις αντιστοιχίες αυτές, έχουν προταθεί διάφοροι τρόποι γεωμετρικής επαλήθευσης των αρχικών αντιστοιχιών. Χρησιμοποιούμε στοιχεία από την γεωμετρία του εκάστοτε προβλήματος, προκειμένου να ανιχνεύσουμε λάθη στις αντιστοιχίες της *τοπικής εμφάνισης* των εικόνων. Αναφέρουμε στην συνέχεια διάφορες προσεγγίσεις από την βιβλιο-



Σχήμα 7.1: Αρχικές αντιστοιχίες εικόνων (αριστερά) και αντιστοιχίες μετά την γεωμετρική τους επαλήθευση (δεξιά)

γραφία, κατά κανόνα από την περιοχή της αναγνώρισης αντικειμένων (object recognition). Από την εισαγωγή ακόμα των χαρακτηριστικών SIFT[133], επιχειρήθηκε η βελτίωση του αρχικού ταιριάσματος των εικόνων με χρήση του μετασχηματισμού Hough, χρησιμοποιώντας τον προσανατολισμό των ανιχνευθέντων χαρακτηριστικών.

Μια μεγάλη κατηγορία μεθόδων[235, 175, 173, 38] χρησιμοποιεί έναν μικρής πολυπλοκότητας μετασχηματισμό (π.χ αφινικό, ομοιότητας) προκειμένου να απορρίψει κάποιες αντιστοιχίες πριν εξαχθεί το πλήρες μοντέλο που συνδέει τις εικόνες. Σχετικά με αυτές τις μεθόδους, αναφέρουμε:

- Έχουν το μειονέκτημα ότι συχνά εξαρτώνται από το είδος των χαρακτηριστικών που έχουν εξαχθεί, προκειμένου να επιταχύνουν τους υπολογισμούς.
- Συχνά οδηγούν (ή και ακόμα στοχεύουν) στην απόρριψη μεγάλου αριθμού χαρακτηριστικών
- Όταν δοκιμάσαμε να χρησιμοποιήσουμε μετασχηματισμό ομοιότητας για την γεωμετρική επαλήθευση των αντιστοιχιών, τα αποτελέσματα (σφάλμα στην εκτίμηση για την σχετική περιστροφή μεταξύ των εικόνων), χειρότερουσαν. Το γεγονός αυτό υποδεικνύει ότι οι συγκεκριμένοι μέθοδοι είναι ακατάλληλες για το πρόβλημα της ανακατασκευής.

Οι προηγούμενες μέθοδοι, συχνά συνδυάζονται και με μεθόδους βελτίωσης της συμμεταβλητότητας (covariance) των τοπικών περιγραφών. Σε αυτή την κατεύθυνση, έχουν χρησιμοποιηθεί μετασχηματισμοί των περιοχών που ανιχνεύτηκαν πριν εξάγουμε από αυτές τους περιγραφείς[173], αντιστοιχίες ελλείψεων αντί σημείων[37], χρήση περιγραφών όπως τα LAFs[36] (Local Affine Frames) κ.α.

Τέλος, έχει προταθεί η χρήση της γειτονιάς κάθε χαρακτηριστικού για την επαλήθευση της αντιστοιχίας. Για κάθε ζεύγος χαρακτηριστικών που έχουν ταιριαστεί, εξάγουμε τα γειτονικά τους χαρακτηριστικά σε κάθε εικόνα. Οι αρχικές αντιστοιχίες απορρίπτονται ή επαληθεύονται ανάλογα με το πλήθος των αντιστοιχιών ανάμεσα στις δύο γειτονιές[203] και την σχετική διάταξη των χαρακτηριστικών στις δύο γειτονιές[250]. Για το τελευταίο αυτό σημείο, χρησιμοποιείται το πλήθος των χαρακτηριστικών που βρίσκονται εκτός διάταξης (μέγεθος συγγενές με την edit distance) . Η μέθοδος που προτείνουμε στην επόμενη ενότητα σχετίζεται με την αυτή την τελευταία ιδέα.

7.1.2 Αναγωγή του προβλήματος στο πρόβλημα της μέγιστης κοινής υπακολουθίας-Longest Common Subsequence

Λόγω της ανεπάρκειας αναφορών σχετικών με την γεωμετρική επαλήθευση αντιστοιχιών εικόνων για το πρόβλημα της ανακατασκευής, καθώς και της ανεπάρκειας (ή και ασυμβατότητας) της βασικής μεθόδου επαλήθευσης μέσω απλούστερων μετασχηματισμών, προτείνουμε μια νέα μέθοδο επαλήθευσης των αρχικών αντιστοιχιών.

Για να προσαρμόσουμε την μέθοδο επαλήθευσης στο πρόβλημα της ανακατασκευής αρχιτεκτονικών σκηνών, αναζητήσαμε ιδιότητες που έχουν τέτοιες σκηνές, οι οποίες είναι εμφανείς σε έναν ανθρώπινο παρατηρητή και του επιτρέπουν άμεσα να εντοπίσει εσφαλμένες αντιστοιχίες μεταξύ ενός ζεύγους

εικόνων.

Η θεμελιώδης ιδιότητα την οποία σκοπεύουμε να χρησιμοποιήσουμε, είναι αυτή της διάταξης των απεικονιζόμενων σημείων, σε σχέση με την οριζόντια και την κατακόρυφη διεύθυνση. Για να το διατυπώσουμε απλούστερα, έχουμε

- Αν ένα σημείο απεικονίζεται πιο αριστερά από ένα άλλο στην πρώτη εικόνα, το ίδιο θα πρέπει να ισχύει και στην δεύτερη. Αυτό είναι αποτέλεσμα του γεγονότος ότι δεν εξετάζουμε καθρεπτισμένες εικόνες (*Συνέπεια-x*)
- Αντίστοιχα, αν ένα σημείο απεικονίζεται πιο ψηλά από ένα άλλο στην πρώτη εικόνα, το ίδιο θα πρέπει να ισχύει και στην δεύτερη. Αυτό είναι αποτέλεσμα του γεγονότος ότι οι παρατηρητές-φωτογράφοι, βρίσκονται κατά κανόνα σε ένα κοινό επίπεδο, το επίπεδο του εδάφους, από το οποίο προκύπτει φυσικά η διάταξη των σημείων στον κάθετο άξονα (*Συνέπεια-y*)

Σημειώνουμε ότι τα προηγούμενα δύο σημεία, παρουσιάζονται απλουστευμένα εδώ, ως ο πυρήνας της μεθόδου που θα αναπτυχθεί. Αναλυτικότερος σχολιασμός των δύο αυτών υποθέσεων θα γίνει στην επόμενη Ενότητα.

Ας εξετάσουμε τώρα κάθε μία από τις δύο προηγούμενες ιδιότητες συνέπειας *ανεξάρτητα*, για να δούμε πως μπορούμε να οδηγηθούμε στο πρόβλημα της *Μέγιστης αύξουσας υπακολουθίας* ή στο ισοδύναμο πρόβλημα της *Μέγιστης κοινής υπακολουθίας*.

Έστω, για διευκόλυνση της ανάπτυξης, ότι εξετάζουμε το πρόβλημα της Συνέπειας ως προς την ιδιότητα του δεξιά-αριστερά, δηλαδή το πρόβλημα της Συνέπειας-x. Ένα σύνολο αντιστοιχιών $S = \{(p_1^i, p_2^i)\}$ είναι συνεπές υπό αυτή την έννοια, αν για κάθε σημείο p_1^i της Εικόνας 1 που ανήκει σε αυτό το σύνολο:

1. Κάθε σημείο της Εικόνας 1 που ανήκει στο συνεπές σύνολο αντιστοιχιών και είναι αριστερά από το p_1^i αντιστοιχίζεται στην Εικόνα 2 σε σημεία που είναι αριστερά από το p_2^i . Δηλαδή

$$S = \{(p_1^i, p_2^i)\} : \forall i \forall j p_1^i \leq p_1^j \implies p_2^i \leq p_2^j, (p_1^i, p_2^i), (p_1^j, p_2^j) \in S$$

2. Το αντίστροφο της προηγούμενης σχέσης. Κάθε σημείο της Εικόνας 1 που ανήκει στο συνεπές σύνολο αντιστοιχιών και είναι δεξιά από το p_1^i αντιστοιχίζεται στην Εικόνα 2 σε σημεία που είναι δεξιά από το p_2^i . Δηλαδή

$$S = \{(p_1^i, p_2^i)\} : \forall i \forall j p_1^i \geq p_1^j \implies p_2^i \geq p_2^j, (p_1^i, p_2^i), (p_1^j, p_2^j) \in S$$

Για να λύσουμε αυτό το πρόβλημα, μπορούμε:

1. Να ταξινομήσουμε τις αντιστοιχίες στην Εικόνα 1 σύμφωνα με την διάταξη τους στον άξονα x . Η ταξινόμηση αυτή μπορεί να θεωρηθεί μία αναδιάταξη (permutation) της ακολουθίας των αντιστοιχιών. Εφαρμόζουμε την ίδια αναδιάταξη στην Εικόνα 2 και παίρνουμε μία ακολουθία από τις τετμημένες x_i των αντιστοιχιών. Αναζητούμε την *Μέγιστη Αύξουσα Υπακολουθία* της ακολουθίας αυτής.

2. Δίνουμε σε κάθε αντιστοιχία των δύο εικόνων (p_1^i, p_2^i) μία μοναδική ταμπέλα, για παράδειγμα $1, 2, \dots$. Στην συνέχεια, ταξινομούμε σε κάθε Εικόνα τις αντιστοιχίες σύμφωνα με την διάταξη τους στον άξονα x . Αναζητούμε την *Μέγιστη Κοινή Υπακολουθία* των δύο αυτών "ταξινομημένων" ακολουθιών.

Ο λόγος που επιλέξαμε να διατυπώσουμε το πρόβλημα στο πλαίσιο της αναζήτησης της Μέγιστης Κοινής/Αύξουσας υπακολουθίας είναι γιατί το πρόβλημα αυτό μπορεί να λυθεί αποδοτικά. Συγκεκριμένα, είναι δυνατή η επίλυση του προβλήματος με ασυμπτωτική πολυπλοκότητα $\mathcal{O}(n \log n)$ [8, 96, 59] ή ακόμα και $\mathcal{O}(n \log \log n)$ [236] με χρήση κατάλληλων δομών δεδομένων. Σημειώνουμε βέβαια ότι για την τάξη των χιλιάδων αντιστοιχιών μεταξύ δύο εικόνων, δηλαδή το συνηθισμένο μέγεθος του προβλήματος, η πολυπλοκότητα $\mathcal{O}(n \log n)$ είναι ικανοποιητική.

Για την επίλυση του προβλήματος, χρησιμοποιούμε τον αλγόριθμο *patience sorting*[8]. Λεπτομέρειες μπορούν να αναζητηθούν στην βιβλιογραφία.

Οι δυσκολίες του συνδυασμένου προβλήματος και μία αποδοτική προσεγγιστική μέθοδος

Ο λόγος που εξετάσαμε ανεξάρτητα τα προβλήματα της Συνέπειας- x και της Συνέπειας- y , τα οποία δείξαμε ότι επιδέχονται βέλτιστη λύση όπως αυτή δίνεται από το πρόβλημα της Μέγιστης αύξουσας υπακολουθίας, είναι γιατί το αρχικό-συνδυασμένο πρόβλημα δεν επιδέχεται αποδοτική λύση. Συγκεκριμένα, μπορούμε να δούμε ότι δεν είναι δυνατή η διατύπωση του συνδυασμένου προβλήματος ως πρόβλημα μέγιστης αύξουσας υπακολουθίας γιατί η σχέση "Συνέπεια- x και Συνέπεια- y " δεν είναι μεταβατική. Αυτό μπορεί να δειχθεί εύκολα με ένα αντιπαράδειγμα. Συνεπώς δεν είναι σχέση μερικής ούτε ολικής διάταξης, και οι αλγόριθμοι που χρησιμοποιούμε δεν μπορούν να τροποποιηθούν κατάλληλα για να εφαρμοστούν στο συνδυασμένο πρόβλημα.

Ακόμα, μπορούμε να διατυπώσουμε το συνδυασμένο πρόβλημα ως εξής. Αναζητούμε ακολουθίες S_1, S_2 , τέτοιες ώστε:

- S_1 να είναι κοινή υπακολουθία του προβλήματος (Μέγιστης) Κοινής Υπακολουθίας για την ιδιότητα *Συνέπεια- x* , όπως αυτό διατυπώθηκε στην προηγούμενη Ενότητα
- S_2 να είναι κοινή υπακολουθία του προβλήματος (Μέγιστης) Κοινής Υπακολουθίας για την ιδιότητα *Συνέπεια- y* , όπως αυτό διατυπώθηκε στην προηγούμενη Ενότητα
- Η υπακολουθία S_1 να είναι μία αναδιάταξη της υπακολουθίας S_2 , δηλαδή $S_1 = \text{Permutation}(S_2)$.
- Το μήκος της υπακολουθίας $n = l(S_1) = l(S_2)$ να είναι το μεγαλύτερο δυνατό.

Η προηγούμενη διατύπωση μας διευκολύνει να δούμε ότι δεν υπάρχει κάποια προφανής αποδοτική λύση του συνδυασμένου προβλήματος, στα πλαίσια τουλάχιστον που το εξετάσαμε.

Για να λύσουμε το πρόβλημα *προσεγγιστικά*, προτείνουμε την ακόλουθη διαδικασία

1. Βρίσκουμε το μεγαλύτερο Συνεπές- x σύνολο αντιστοιχιών S_x , με λύση του προβλήματος Μέγιστης Αύξουσας Υπακολουθίας
2. Βρίσκουμε το μεγαλύτερο Συνεπές- y υποσύνολο του S_x , με λύση του προβλήματος Μέγιστης Αύξουσας Υπακολουθίας

Για την λύση αυτή παρατηρούμε:

- Η λύση ικανοποιεί και τις δύο ιδιότητες της Συνέπειας- x και Συνέπειας- y , είναι δηλαδή *λύση του συνδυασμένου προβλήματος*
- Η πληθικότητα n του συνόλου αντιστοιχιών που λαμβάνουμε ως λύση, δεν είναι πάντα η μεγαλύτερη δυνατή. Εύκολα μπορεί να βρεθεί αντιπαράδειγμα για την επαλήθευση της παρατήρησης αυτής. Συνεπώς, δεν παίρνουμε πάντοτε την *βέλτιστη λύση*.

Σχολιάζοντας τις προηγούμενες παρατηρήσεις, αναφέρουμε ότι τα χαρακτηριστικά αυτά της λύσης ταιριάζουν με τις προσδοκίες μας. Συγκεκριμένα, μας *ενδιαφέρει περισσότερο* να μην υπάρχουν outliers στο τελικό σύνολο αντιστοιχιών, πράγμα που *ικανοποιείται* από την λύση αφού ισχύει η ιδιότητα της Συνέπειας- x και της Συνέπειας- y . Η μικρότερη πληθικότητα του συνόλου των αντιστοιχιών, δείχνει ότι απορρίψαμε ενδεχομένως σωστές αντιστοιχίες. Το αποτέλεσμα αυτό είναι μεν αρνητικό, δεν επηρεάζει όμως την ποιότητα της επίλυσης του προβλήματος της ανακατασκευής όσο ή διατήρηση εσφαλμένων αντιστοιχιών (outliers) .

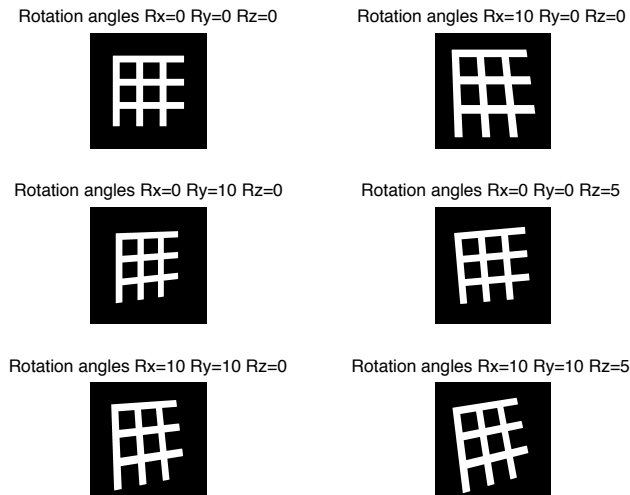
7.1.3 Αποκλίσεις από το αρχικό μοντέλο και βελτιώσεις στην βασική μέθοδο

Τώρα που διατυπώσαμε την μέθοδο γεωμετρικής επαλήθευσης των αντιστοιχιών, μπορούμε να εξετάσουμε αναλυτικότερα τις *παραδοχές* στις οποίες βασίζεται καθώς και τον τρόπο που αυτές *παραβιάζονται*.

Αρχικά, ως αναπτύξουμε την γεωμετρική δομή στην οποία βασίζεται η μέθοδος. Είναι σύνηθες για τις φωτογραφικές απεικονίσεις αρχιτεκτονικών σκηνών, ο άξονας y των συντεταγμένων της κάμερας να ταυτίζεται με το κάθετο στο έδαφος διάνυσμα. Η παραδοχή αυτή, είναι μαθηματικά ισοδύναμη με $R_z = R_x = 0$. Σε πιο ειδικές περιπτώσεις, όπως για παράδειγμα αν φωτογραφίζονται σπίτια σε έναν δρόμο, το ίδιο μπορεί να λεχθεί για την σχέση του άξονα x του συστήματος συντεταγμένων της κάμερας και του διανύσματος που είναι ομόρροπο με την διεύθυνση του δρόμου.

Συχνά όμως, οι προηγούμενες παραδοχές παραβιάζονται μεταξύ διαφορετικών όψεων. Στην Εικόνα 7.2 παρουσιάζονται τα αποτελέσματα τέτοιων παραμορφώσεων. Παρατηρούμε ότι:

- Οι ευθείες που είναι παράλληλες με τον άξονα x (αντίστοιχά y) σε μία Εικόνα, μπορεί να εμφανίζονται με κλίση σε κάποια άλλη. Αυτό είναι αποτέλεσμα τόσο της απουσίας ευθυγράμμισης μεταξύ των συστημάτων αναφοράς των δύο καμερών, όσο και των διαβαθμίσεων στο βάθος της σκηνής.
- Αλλαγές στην σχετική διάταξη των σημείων ανάμεσα σε δύο εικόνες είναι δυνατόν να συμβούν. Συγκεκριμένα, παρατηρούμε ότι είναι εύκολο να αλλάξει η διάταξη ως προς x (αντ. y) δύο σημείων τα οποία είναι κοντά στον άξονα x (αντ. y) και απομακρυσμένα στον άξονα y (αντ. x) . Δηλαδή όταν ισχύει $\Delta x \ll \Delta y$.



Σχήμα 7.2: Παραμορφώσεις παράλληλων ευθειών λόγω περιστροφών του συστήματος συντεταγμένων της κάμερας. Οι Εικόνες έχουν προκύψει από προβολή μιας τρισδιάστατης σκηνής σε σταθερό βάθος $z = z_{const}$. Πρόκειται δηλαδή για εικόνα *επίπεδης σκηνής*. Αντίστοιχες παραμορφώσεις προκύπτουν και από διαφορά στο βάθος των αντικειμένων που απεικονίζονται.

Γενικά, μπορούμε να υποθέσουμε ότι οι κάμερες *δεν θα σχηματίζουν μεγάλες γωνίες* με τους άξονες x, y του φυσικού συστήματος συντεταγμένων, δηλαδή του συστήματος συντεταγμένων του χώρου στο οποίο ο άξονας y έχει την διεύθυνση της βαρύτητας και ο άξονας x είναι παράλληλος με το έδαφος. Αυτό είναι συνέπεια τόσο του περιορισμού στην θέση του φωτογράφου, όσο και σε αισθητικά κριτήρια που κάνουν τις φωτογραφίες μη-ελκυστικές. Όμως, σε περιπτώσεις που η πολυπλοκότητα της σκηνής είναι αυξημένη, οι διάφορα στον προσανατολισμό των κτηρίων και τα διαφορετικά βάθη της σκηνής εντείνουν τα ανεπιθύμητα φαινόμενα.

Τέλος, μπορούμε να *αγνοήσουμε* την περιστροφή ως προς το *επίπεδο της εικόνας* R_z , αφού ξεφεύγει από την συνηθισμένη φωτογραφική πρακτική και μπορεί επίσης να διορθωθεί από κατάλληλα αυτόματα συστήματα[70].

Προσεγγίσεις στην σχέση Συνέπειας και συγκρίσεις αποτελεσμάτων

Όπως αναφέραμε στην προηγούμενη Ενότητα, προβολικά φαινόμενα μπορεί να προκαλέσουν την παραβίαση των ιδιοτήτων Συνέπειας που εισάγαμε νωρίτερα. Ακόμα, παρατηρήσαμε ότι η *αρχική μέθοδος*, *απορρίπτει* ασύμφορα *μεγάλο αριθμό σωστών αντιστοιχιών*. Στην συνέχεια παρουσιάζουμε τους τρόπου μεταβολής της μεθόδου που εξετάσαμε.

Ανοχή ως ποσοστό της Εικόνας

Σε αυτή την προσέγγιση, επιτρέπουμε σημεία εκτός διάταξης (δηλαδή φθίνουσα τμήματα στην αύξουσα υπακολουθία), εφόσον η παραβίαση είναι εντός μιας προκαθορισμένης ανοχής. Ελέγχουμε δηλαδή την σχέση

$$s_i - T \leq s_{i+1}$$

όπου s_i, s_{i+1} διαδοχικά σημεία της υπακολουθίας. Η επιλογή της ανοχής γίνεται ως ποσοστό της μέγιστης απόστασης δύο αρχικών αντιστοιχιών, δηλαδή

$$T = \alpha(y_{max} - y_{min})$$

όταν ελέγχουμε την Συνέπεια-x και αντίστοιχα όταν ελέγχουμε την Συνέπεια-y. Η επιλογή αυτή ως ποσοστό της "δεύτερης" διάστασης είναι σε συμφωνία με τις προηγούμενες παρατηρήσεις μας σχετικά με τα σημεία που μπορεί να βρεθούν εκτός διάταξης. Σημειώνουμε ότι λύνουμε ίδιο αριθμό προβλημάτων μέγιστης αύξουσας υπακολουθίας με την βασική προσέγγιση

Αναδρομική Μέθοδος

Στην αναδρομική μέθοδο, λύνουμε μια ακολουθία προβλημάτων μέγιστης αύξουσας υπακολουθίας, χρησιμοποιώντας πάλι ανοχή ως ποσοστό του μεγέθους της Εικόνας. Αντί όμως να λύσουμε ένα πρόβλημα όπως στην μέθοδο που περιγράψαμε αμέσως προηγούμενα, χωρίζουμε την εικόνα σε μικρότερα τμήματα και επαναλαμβάνουμε. Συγκεκριμένα έχουμε:

1. Επιλύουμε το πρόβλημα μέγιστης αύξουσας υπακολουθίας με ανοχή ως ποσοστό του μεγέθους της εικόνας. Λαμβάνουμε ως λύση συνεπές σύνολο αντιστοιχιών
2. Διαιρούμε την εικόνα σε δύο τμήματα, ώστε κάθε τμήμα να περιέχει ίσο αριθμό αντιστοιχιών. Αν εξετάσουμε την Συνέπεια-x, η διαίρεση γίνεται σε σχέση με τον άξονα y , δηλαδή ως τομή $y = y_{const}$. Αντίστοιχα και για την Συνέπεια-y
3. (Αναδρομή) Επαναλαμβάνουμε την ίδια διαδικασία σε κάθε ένα από τα δύο τμήματα που προκύπτουν. Η διαδικασία τερματίζεται αν το μέγεθος της εικόνας είναι μικρότερο από μια προκαθορισμένη σταθερά (200 pixels)

Η μέθοδος αυτή έχει το πλεονέκτημα ότι επιτρέπει μικρότερες παραβιάσεις στην διάταξη στον x άξονα για σημεία που είναι κοντά στον y άξονα και μεγαλύτερες παραβιάσεις μεταξύ σημείων που είναι απομακρυσμένα στον y άξονα. Η διάκριση αυτή είναι σε συμφωνία με τις προηγούμενες παρατηρήσεις μας.

Η ασυμπτωτική πολυπλοκότητα της μεθόδου δεν αλλάζει, και πρακτικά οι αλλαγές στον χρόνο εκτέλεσης δεν γίνονται αισθητές. Αναλυτικά για τον χρόνο εκτέλεσης της μεθόδου έχουμε:

$$\begin{aligned} \text{Complexity} &= n \log n + \sum_L \sum n_i \log n_i \\ &\leq n \log n + \sum_L \sum n_i \log n, \text{ αφού } n_i \leq n \\ &\leq n \log n + \sum_L n \log n, \text{ αφού } \sum n_i \leq n \\ &= (L + 1)n \log n \end{aligned}$$

όπου η σταθερά L είναι συνάρτηση του αρχικού μεγέθους της εικόνας και του ελάχιστου μεγέθους εικόνας στο οποίο λύνουμε το πρόβλημα.

Υπολογισμός ανοχής από το σφάλμα του ιδανικού μοντέλου

Θα αναλύσουμε την μέθοδο αυτή, εξετάζοντας την Συνέπεια- y . Η αντίστοιχη διαδικασία ακολουθείται και για την σχέση της Συνέπειας- x .

Αρχικά, βρίσκουμε ένα σύνολο αξιόπιστων αντιστοιχιών S_1 , λύνοντας το (ακριβές) πρόβλημα μέγιστης αύξουσας υπακολουθίας. Στην συνέχεια, χρησιμοποιούμε το σύνολο S_1 , ώστε να προσδιορίσουμε κατάλληλη τιμή ανοχής. Λύνουμε ξανά, με είσοδο τις αρχικές αντιστοιχίες των εικόνων, το πρόβλημα μέγιστης υπακολουθίας με την τιμή ανοχής που προσδιορίσαμε.

Για να προσδιορίσουμε την τιμή ανοχής, ακολουθούμε την επόμενη διαδικασία. Στην ιδανική περίπτωση, οι ευθείες $y_{IM1} = c$, της Εικόνας 1, θέλουμε να απεικονίζονται σε ευθείες $y_{IM2} = c'$ στην Εικόνα 2. Αν ισχύει αυτό, τότε δεν θα υπάρχουν αλλαγές στην διάταξη στον άξονα y των έγκυρων αντιστοιχιών στις δύο Εικόνες, συνεπώς η αυστηρή λύση της μέγιστης αύξουσας υπακολουθίας δίνει το σωστό αποτέλεσμα.

Για να ισχύει η προηγούμενη ιδιότητα, θα πρέπει καταρχάς να ισχύει

$$R_y = R_z = 0;$$

δηλαδή $R_2 = R_x$.

Σημειώνουμε ακόμα ότι μπορούμε (και έχουμε) να υποθέσουμε ότι η πρώτη κάμερα βρίσκεται στην αρχή των αξόνων και έχει μήτρα περιστροφής την μοναδιαία. Με αυτό το σύστημα αναφοράς, έχουμε

$$y_{IM1} = const. \iff y/z = const.$$

Αν προβάλλουμε τώρα ένα σημείο της ευθείας της Εικόνας 1 $y_{IM1} = c$ στην Εικόνα 2, μετά από πράξεις και με τον ορισμό κατάλληλων σταθερών έχουμε:

$$y_{IM2} = \frac{By/z + \Gamma + D/z}{B'y/z + \Gamma' + D'/z}$$

Αν τώρα στην προηγούμενη σχέση υποθέσουμε επιπλέον ότι $z = const.$, καταλήγουμε στην σχέση

$$y_{IM2} = \frac{Ey_{IM1} + F}{E'y_{IM1} + F'}$$

Αν έχουμε 4 αντιστοιχίες μεταξύ των δύο εικόνων, μπορούμε να λύσουμε (με svd) το 4×4 ομογενές σύστημα εξισώσεων. Στην συνέχεια, χρησιμοποιώντας το μοντέλο αυτό, και δοσμένη μιας 5ης αντιστοιχίας $y_{IM1} \leftrightarrow y_{IM2}$ μπορούμε να προβλέψουμε την δοσμένου του y_{IM1} το y_{IM2} . Το μέγιστο τέτοιο *σφάλμα εκτίμησης* δίνει μια εκτίμηση για την απόκλιση των εικόνων από τις υποθέσεις μας. Έτσι, χρησιμοποιούμε *αυτή την τιμή ως ανοχή* στην εύρεση της μέγιστης αύξουσας υπακολουθίας.

Συγκρίσεις των μεθόδων

Προκειμένου να συγκρίνουμε τις μεθόδους που παρουσιάσαμε παραπάνω, εντάσσουμε το πρόβλημα της γεωμετρικής επαλήθευσης αντιστοιχιών στο πλαίσιο της ταξινόμησης (classification). Θεωρούμε δηλαδή, ότι οι αρχικές αντιστοιχίες ανήκουν σε δύο κλάσεις, ψευδείς και έγκυρες αντιστοιχίες, και θέλουμε να ποσοτικοποιήσουμε τα αποτελέσματα των διάφορων μεθόδων-ταξινόμητων.

Στην βιβλιογραφία, πιο διαδεδομένα είναι ολοκληρωτικά μέτρα όπως τα AUC,

AP [55]. Τα μέτρα αυτά, επικρίθηκαν πρόσφατα [172] και προϋποθέτουν την ύπαρξη κάποιας μεταβαλλόμενης παραμέτρου που επηρεάζει την ταξινόμηση, η οποία δεν είναι διαθέσιμη για κάθε μέθοδο.

Ακόμα, όπως αναφέραμε και προηγούμενα, θέλουμε να δώσουμε *μεγαλύτερο βάρος στο precision* του ταξινομητή από το recall του ταξινομητή. Δηλαδή, ενδιαφερόμαστε περισσότερο να επιστρέψουμε ένα σύνολο αντιστοιχιών στο οποίο δεν υπάρχουν εσφαλμένες αντιστοιχίες σε σχέση με την επιστροφή όλων των σωστών αντιστοιχιών. Τα ολοκληρωτικά μέτρα δεν δίνουν την δυνατότητα αυτή.

Έτσι, ακολουθώντας και την ανάλυση του [172], επιλέγουμε να συγκρίνουμε τους ταξινομητές με βάση τον συντελεστή ϕ . Ο συντελεστής ϕ ορίζεται ως

$$\phi \equiv \frac{\text{cov}(X, Y)}{\sqrt{\sigma_X \sigma_Y}} \quad (7.1)$$

όπου X, Y δυαδικές τυχαίες μεταβλητές που αντιστοιχούν στην πραγματική κατηγορία (X) και την απόφαση του ταξινομητή (Y). Για τον υπολογισμό του συντελεστή ϕ , μπορούμε να χρησιμοποιήσουμε την ισοδύναμη σχέση

$$\phi = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}$$

Τέλος, αναφέρουμε και μία τελευταία σχέση η οποία βοηθάει στην ερμηνεία του συντελεστή ϕ

$$\phi^2 = \frac{\chi^2}{n}$$

Ο συντελεστής ϕ εκφράζει την ανεξαρτησία των τυχαίων μεταβλητών X , και λαμβάνει τιμή 0 για ανεξάρτητες τυχαίες μεταβλητές. Συνεπώς ένας καλός ταξινομητής δίνει τιμές ϕ κοντά στο 1. Για μια πλήρη ανάλυση του θέματος, παραπέμπουμε τον αναγνώστη στο [42, σ.282].

Επίσης, από την Σχέση 7.1, μπορούμε να αντικαταστήσουμε την ετεροσυσχέτιση με *σταθμισμένη ετεροσυσχέτιση*, προκειμένου να δώσουμε μεγαλύτερο βάρος σε κάποια κατηγορία λάθους ταξινόμησης. Στους Πίνακες 7.1, 7.2, 7.3, 7.4, 7.5, 7.6 παρουσιάζουμε μαζί με τα άλλα μεγέθη και τον *σταθμισμένο συντελεστή* ϕ . Τα βάρη που χρησιμοποιήθηκαν στον υπολογισμό του ήταν 1 για σωστή ταξινόμηση, 2 για ταξινόμηση εσφαλμένης αντιστοιχίας ως έγκυρης και 0.5 για το αντίστροφο είδος λάθους.

Για την αξιολόγηση των μεθόδων, χρησιμοποιήσαμε το σύνολο δεδομένων [216], που έχει χρησιμοποιηθεί και στα Κεφάλαια 5,6 (Εισήχθει στην Ενότητα 5.3.3).

Για την εύρεση των έγκυρων αντιστοιχιών, χρησιμοποιήσαμε τους δοσμένους, ορθούς πίνακες προβολής P και υπολογίσαμε από αυτούς τους Θεμελιώδεις πίνακες F_{ij} για κάθε ζεύγος φωτογραφιών. Οι έγκυρες αντιστοιχίες βρέθηκαν υπολογίζοντας το γεωμετρικό σφάλμα (προσέγγιση Sampson) για κάθε αντιστοιχία από τον γνωστό πίνακα επιτολικής γεωμετρίας F_{ij} .

Επίσης χρησιμοποιήθηκε ένα νέο σύνολο Δεδομένων [244]. Το σύνολο αυτό περιέχει περιέχει αρχιτεκτονικές σκηνές καθώς και σκηνές με αντικείμενα. Τα αποτελέσματα στο σύνολο των δεδομένων και στο υποσύνολο που περιέχει αρχιτεκτονικές σκηνές δίνονται ξεχωριστά. Σημειώνουμε ότι οι αντιστοιχίες σε αυτό το σύνολο δεδομένων έχουν ταξινομηθεί σε εσφαλμένες και ορθές με το χέρι.

Γενικά παρατηρούμε:

- Η "πολυκλιμακωτή" ανάλυση των δεδομένων που επιτυγχάνεται με την *αναδρομική μέθοδο* έχει εμφανή πλεονεκτήματα. Το γεγονός αυτό επιβεβαιώνει την ανάλυση που δώσαμε σχετικά με την αλλαγή της διάταξης σημείων ως προς τον έναν άξονα και την απόσταση των σημείων στον δεύτερο άξονα.
- Η αμέσως επόμενη μέθοδος σε επιδόσεις βασίζεται επίσης σε θεωρητικές επισημάνσεις που κάναμε. Επιβεβαιώνεται έτσι, δια της βελτίωσης των αποτελεσμάτων, η ορθότητα των συλλογισμών μας. Υστερεί όμως σε σχέση με την αναδρομική μέθοδο
- Οι μέθοδοι με σταθερή ανοχή ή και μηδενική ανοχή, σημειώνουν εξαιρετικά χαμηλές τιμές recall, γεγονός που διαπιστώσαμε πως δυσχεραίνει της επιτυχή ανακατασκευή.
- Η *επιτυχία* της μεθόδου εξαρτάται από το *είδος* της σκηνής. Η εξάρτηση αυτή είναι συνέπεια της κατασκευής της μεθόδου. Όπως φαίνεται από τα αποτελέσματα στα δεδομένα [244], η μέθοδος εφαρμόζεται καλύτερα στις *αρχιτεκτονικές σκηνές* από ότι στις σκηνές αντικειμένων. Αυτό οφείλεται τόσο στις διαφορές στην γεωμετρία της σκηνής, όσο και στην μεγαλύτερη ελευθερία στην θέση του *παρατηρητή* στην δεύτερη κατηγορία σκηνών. Για τα δεδομένα [216], η επίδοση του συστήματος (precision) διαφέρει σημαντικά στα διάφορα σύνολα φωτογραφιών. Για σύνολα όπως τα **Fountain-P11,entry-P10,Herz-Jesu-P8** η precision λαμβάνει τιμές > 0.99 . Στα σύνολα αυτά, η γεωμετρία της σκηνής, η οποία αποτελείται από ένα κύριο οικοδόμημα και συνεπώς έχει μοναδική οριζόντια και κάθετη κυρίαρχη διεύθυνση, συμφωνεί πλήρως με τις υποθέσεις μας. Αντίθετα, σε σύνολα όπως το **castle-P30**, στα οποία εικονίζονται και κάθετα μεταξύ τους κτίσματα, η precision του συστήματος πέφτει στο $0.7 - 0.8$.

| Μέθοδος | T=0.02 | T=0.04 | T=0.06 | T=0.08 | T=0.10 | T=0.15 | T=0.20 |
|--------------------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| Βασική | 0.47 ^{0.40} | | | | | | |
| Ανοχή T | 0.46 ^{0.40} | 0.46 ^{0.40} | 0.46 ^{0.40} | 0.47 ^{0.40} | 0.47 ^{0.40} | 0.47 ^{0.40} | 0.48 ^{0.40} |
| Αναδρομική με ανοχή T | 0.77 ^{0.64} | 0.84 ^{0.70} | 0.88 ^{0.73} | 0.88 ^{0.74} | 0.89 ^{0.75} | 0.84 ^{0.71} | 0.80 ^{0.67} |
| Υπολογισμός μέγιστου σφάλματος | 0.81 ^{0.69} | | | | | | |

Πίνακας 7.1: Συντελεστής ϕ για τις μεθόδους γεωμετρικής επαλήθευσης. Οι μέθοδοι παρουσιάζονται στον πίνακα με την σειρά που παρουσιάστηκαν στο κείμενο. Για την Πρώτη και την Τέταρτη μέθοδο, δεν χρησιμοποιείται ανοχή T. Με *μικρούς πλάγιους* χαρακτήρες δίνονται τα αποτελέσματα σε ολόκληρο το σύνολο δεδομένων και με την συνήθη γραμματοσειρά τα αποτελέσματα στο υποσύνολο των αρχιτεκτονικών σκηνών.

| Μέθοδος | T=0.02 | T=0.04 | T=0.06 | T=0.08 | T=0.10 | T=0.15 | T=0.20 |
|--------------------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| Βασική | 0.60 ^{0.54} | | | | | | |
| Ανοχή T | 0.60 ^{0.53} | 0.60 ^{0.53} | 0.60 ^{0.53} | 0.60 ^{0.53} | 0.61 ^{0.53} | 0.61 ^{0.54} | 0.61 ^{0.54} |
| Αναδρομική με ανοχή T | 0.84 ^{0.74} | 0.88 ^{0.77} | 0.88 ^{0.79} | 0.86 ^{0.77} | 0.85 ^{0.77} | 0.78 ^{0.71} | 0.72 ^{0.65} |
| Υπολογισμός μέγιστου σφάλματος | 0.81 ^{0.74} | | | | | | |

Πίνακας 7.2: Σταθμισμένος συντελεστής ϕ για τις μεθόδους γεωμετρικής επαλήθευσης. Οι μέθοδοι παρουσιάζονται στον πίνακα με την σειρά που παρουσιάστηκαν στο κείμενο. Για την Πρώτη και την Τέταρτη μέθοδο, δεν χρησιμοποιείται ανοχή T. Με *μικρούς πλάγιους* χαρακτήρες δίνονται τα αποτελέσματα στο υποσύνολο των αρχιτεκτονικών σκηνών.

| Μέθοδος | T=0.02 | T=0.04 | T=0.06 | T=0.08 | T=0.10 | T=0.15 | T=0.20 |
|--------------------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| Βασική | 0.98 ^{0.97} | | | | | | |
| Ανοχή T | 0.98 ^{0.97} | 0.98 ^{0.97} | 0.98 ^{0.97} | 0.98 ^{0.97} | 0.98 ^{0.97} | 0.98 ^{0.97} | 0.98 ^{0.97} |
| Αναδρομική με ανοχή T | 0.99 ^{0.98} | 0.98 ^{0.97} | 0.97 ^{0.96} | 0.96 ^{0.96} | 0.95 ^{0.95} | 0.91 ^{0.91} | 0.87 ^{0.87} |
| Υπολογισμός μέγιστου σφάλματος | 0.97 ^{0.95} | | | | | | |

Πίνακας 7.3: Precision για τις μεθόδους γεωμετρικής επαλήθευσης. Οι μέθοδοι παρουσιάζονται στον πίνακα με την σειρά που παρουσιάστηκαν στο κείμενο. Για την Πρώτη και την Τέταρτη μέθοδο, δεν χρησιμοποιείται ανοχή T. Με *μικρούς πλάγιους* χαρακτήρες δίνονται τα αποτελέσματα σε ολόκληρο το σύνολο δεδομένων και με την συνήθη γραμματοσειρά τα αποτελέσματα στο υποσύνολο των αρχιτεκτονικών σκηνών

| Μέθοδος | T=0.02 | T=0.04 | T=0.06 | T=0.08 | T=0.10 | T=0.15 | T=0.20 |
|--------------------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| Βασική | 0.43 ^{0.34} | | | | | | |
| Ανοχή T | 0.42 ^{0.34} | 0.42 ^{0.34} | 0.43 ^{0.34} | 0.43 ^{0.34} | 0.43 ^{0.34} | 0.43 ^{0.35} | 0.44 ^{0.35} |
| Αναδρομική με ανοχή T | 0.80 ^{0.64} | 0.89 ^{0.72} | 0.93 ^{0.76} | 0.94 ^{0.78} | 0.96 ^{0.80} | 0.95 ^{0.80} | 0.96 ^{0.81} |
| Υπολογισμός μέγιστου σφάλματος | 0.90 ^{0.74} | | | | | | |

Πίνακας 7.4: Recall για τις μεθόδους γεωμετρικής επαλήθευσης. Οι μέθοδοι παρουσιάζονται στον πίνακα με την σειρά που παρουσιάστηκαν στο κείμενο. Για την Πρώτη και την Τέταρτη μέθοδο, δεν χρησιμοποιείται ανοχή T. Με *μικρούς πλάγιους* χαρακτήρες δίνονται τα αποτελέσματα σε ολόκληρο το σύνολο δεδομένων και με την συνήθη γραμματοσειρά τα αποτελέσματα στο υποσύνολο των αρχιτεκτονικών σκηνών

| Μέθοδος | T=0.02 | T=0.04 | T=0.06 | T=0.08 | T=0.10 | T=0.15 | T=0.20 |
|--------------------------------|--------|--------|--------|--------|--------|--------|--------|
| Βασική | 0.85 | | | | | | |
| Ανοχή T | 0.86 | 0.86 | 0.86 | 0.86 | 0.86 | 0.86 | 0.86 |
| Αναδρομική με ανοχή T | 0.85 | 0.85 | 0.85 | 0.85 | 0.85 | 0.85 | 0.84 |
| Υπολογισμός μέγιστου σφάλματος | 0.85 | | | | | | |

Πίνακας 7.5: Recall για τις μεθόδους γεωμετρικής επαλήθευσης. Οι μέθοδοι παρουσιάζονται στον πίνακα με την σειρά που παρουσιάστηκαν στο κείμενο. Για την Πρώτη και την Τέταρτη μέθοδο, δεν χρησιμοποιείται ανοχή T. Τα Αποτελέσματα αφορούν τα σύνολα δεδομένων που εισήχθησαν στην Ενότητα 5.3.3

| Μέθοδος | T=0.02 | T=0.04 | T=0.06 | T=0.08 | T=0.10 | T=0.15 | T=0.20 |
|--------------------------------|--------|--------|--------|--------|--------|--------|--------|
| Βασική | 0.30 | | | | | | |
| Ανοχή T | 0.30 | 0.30 | 0.30 | 0.30 | 0.30 | 0.31 | 0.30 |
| Αναδρομική με ανοχή T | 0.60 | 0.70 | 0.76 | 0.79 | 0.82 | 0.86 | 0.88 |
| Υπολογισμός μέγιστου σφάλματος | 0.68 | | | | | | |

Πίνακας 7.6: Precision για τις μεθόδους γεωμετρικής επαλήθευσης. Οι μέθοδοι παρουσιάζονται στον πίνακα με την σειρά που παρουσιάστηκαν στο κείμενο. Για την Πρώτη και την Τέταρτη μέθοδο, δεν χρησιμοποιείται ανοχή T. Τα Αποτελέσματα αφορούν τα σύνολα δεδομένων που εισήχθησαν στην Ενότητα 5.3.3

7.2 Βελτιώσεις στις εκτιμήσεις των εξωτερικών και εσωτερικών παραμέτρων των όψεων

7.2.1 Εισαγωγή: κίνητρα και στόχοι

Στα προηγούμενα Κεφάλαια, παρουσιάσαμε εκτενώς μεθόδους από την διβλιογραφία και σχηματίσαμε ένα βασικό σύστημα. Όμως, το σύστημα που αναπτύξαμε, συναντά εντονότερα τους περιορισμούς της *σταθερής επίδοσης* και της *ικανότητας γενίκευσης*. Τα προβλήματα αυτά, γίνονται καλύτερα αντιληπτά με τις επόμενες παρατηρήσεις.

Αναφέραμε διάφορες *μεθόδους* υπολογισμού του *δεμελιώδους Πίνακα F* από αντιστοιχίες σημείων. Όμως, δεν καταφέραμε να λάβουμε σταθερά, αξιόπιστα και ακριβή αποτελέσματα χρησιμοποιώντας την *ίδια* μέθοδο σε *όλα* τα σύνολα φωτογραφιών που χρησιμοποιήσαμε. Αντίστοιχα και για τις παραμέτρους των μεθόδων αυτών.

Η εκτίμηση μας για τις *εσωτερικές παραμέτρους*, δηλαδή το εστιακό μήκος f , προέρχονται από έναν μόνο πίνακα F (για κάθε ζεύγος καμερών). Η επιλογή αυτή, *αφαιρεί* από το πρόβλημα το στοιχείο του υπερρορισμένου συστήματος. Δηλαδή, το πλήθος των σημείων χρησιμοποιείται για την εκτίμηση της ενδιάμεση αναπαράστασης του Θεμελιώδους πίνακα και όχι για την εκτίμηση των τελικών ποσοτήτων, δηλαδή των μητρών εσωτερικής και εξωτερικής βαθμονόμησης. *Όμως*, θα μπορούσαμε να εκτιμήσουμε και τις τελικές ποσότητες από ένα ελάχιστο δείγμα (7ή8 σημείων), χωρίς σημαντικό υπολογιστικό φόρτο. Συνεπώς, πιστεύουμε ότι η υπολογιστική και μαθηματική ευκολία στον προσδιορισμό των τελικών ποσοτήτων οφείλει να αξιοποιηθεί.

Τέλος, για την καταχώρηση όλων των όψεων σε κοινό σύστημα αναφοράς (*rotation registration*), εφαρμόζουμε διαδικασία RANSAC με δειγματοληψία σε συνδεδετικά δέντρα στον γράφο των σχετικών περιστροφών. Η ιδέα αυτή, δεν *κλιμακώνει* καλά με την αύξηση των εικόνων εισόδου, λόγω της εκθετικής αύξησης του αριθμού των υποψηφίων συνδεδετικών δέντρων. Επίσης, εξαιτίας της *τυχαιότητας* της μεθόδου, τα αποτελέσματα δεν είναι αξιόπιστα και σταθερά. Πιστεύουμε ότι θα ήταν προτιμότερο να αντικαταστήσουμε την μέθοδο με άλλη, ντετερμινιστική και περισσότερο αξιόπιστη.

Η κατεύθυνση που θα ακολουθήσουμε στηρίζεται πρωταρχικά στους ακόλουθους δύο άξονες:

- Αξιοποίηση της δυνατότητας εκτίμησης των εξωτερικών και εσωτερικών παραμέτρων για ένα ζεύγος καμερών από *ελάχιστα δείγματα* 7 – 8 σημείων, με χαμηλής πολυπλοκότητας μεθόδους. Προτείνουμε την συλλογή *πολλών εκτιμήσεων* και την ακόλουθη *στάθμιση* (με κατάλληλες μεθόδους) των εκτιμήσεων για την παραγωγή της τελικής λύσης
- Αξιοποίηση της πλήρους μορφής των αρχικών εκτιμήσεων με την χρήση *γεωμετρικών επιχειρημάτων* για να καθοδηγήσουμε την εξαγωγή των τελικών λύσεων σε πιο ακριβείς τιμές

7.2.2 Σκιαγραφώντας την Γενική Μέθοδο

Στην ενότητα αυτή, θα περιγράψουμε με μεγαλύτερη σαφήνεια την γενική μέθοδο που εφαρμόσαμε. Συγκεκριμένα, αλλάξαμε τα επόμενα τμήματα της διαδικασίας ανακατασκευής:

- *Εκτίμηση σχετικών περιστροφών*: Χρησιμοποιούμε διαδικασία RANSAC. Σε κάθε επανάληψη k του αλγορίθμου, από την εκτίμηση για τον θεμελιώδη πίνακα F_{ij}^k , λαμβάνουμε χρησιμοποιώντας την μέθοδο που αναπτύχθηκε στο Κεφάλαιο 5 και εκτιμήσεις f_i^k, f_j^k, R_{ij}^k . Η τελική εκτίμηση των παραμέτρων αυτών λαμβάνεται χρησιμοποιώντας τον αλγόριθμο του Weiszfeld (Ενότητα 7.2.3) για την στάθμιση των εκτιμήσεων R_{ij}^k . Σταθμίζουμε επίσης και τις βαθμωτές ποσότητες f_i^k, f_j^k χρησιμοποιώντας απλές τεχνικές (π.χ. μέσο, διάμεσο, ευριστικές κλπ). Υπενθυμίζουμε ότι οι εκτιμήσεις για την μεταφορά t_{ij} δεν είναι σημαντική για την μέθοδο του Κεφαλαίου 6. Η ιδιότητα αυτή είναι καθοριστικής σημασίας για την δυνατότητα εφαρμογής των τροποποιήσεων μας
- *Καταχώρηση περιστροφών σε κοινό σύστημα αναφοράς*: Αντί να εφαρμόσουμε RANSAC, χρησιμοποιούμε πάλι μορφή του αλγορίθμου Weiszfeld, αυτή την φορά για το πρόβλημα της στάθμισης πολλών περιστροφών (multiple rotation averaging)

7.2.3 Μια σύντομη περιγραφή της στάθμισης περιστροφών με τον αλγόριθμο του Weiszfeld

Σε αυτή την Ενότητα περιγράφουμε σύντομα την μέθοδο στάθμισης περιστροφών, εκτιμήσεων για την ίδια περιστροφή και πολλαπλών περιστροφών, που εφαρμόσαμε. Αναλυτική περιγραφή μπορεί να αναζητηθεί στα [89, 86]. Ο αλγόριθμος του Weiszfeld είναι ένας επαναληπτικός αλγόριθμος για την εύρεση του L_1 -μέσου ενός συνόλου σημείων στον χώρο \mathcal{R}^n . Για ένα σύνολο σημείων \mathbf{x}_i , ο L_1 μέσος τους είναι το σημείο \mathbf{y} που ελαχιστοποιεί την ποσότητα:

$$\sum_{i=1}^n \|\mathbf{y} - \mathbf{x}_i\|$$

Στον χώρο των περιστροφών, μπορούν να οριστούν διάφορα μέτρα απόστασης. Γενικά, τα μέτρα αυτά συνδέονται μεταξύ του με μη-γραμμικές (τριγωνομετρικές) συναρτήσεις. Αναφέρουμε δύο από αυτά, καθώς θα τα χρησιμοποιήσουμε αργότερα:

$$d_{\text{geometric}}(R, S) = \text{γωνία της περιστροφής } RS^{-1}$$

$$d_{\text{chordal}}(R, S) = \|R - S\|_F = 2\sqrt{2} \sin(\theta/2)$$

Ο αλγόριθμος του Weiszfeld είναι ένας επαναληπτικός αλγόριθμος τύπου Gradient-descent, αφού δεδομένης μιας εκτίμησης η επόμενη προκύπτει με βήμα στην διεύθυνση της κλίσης. Σημειώνουμε ότι η σύγκλιση του αλγορίθμου στο σημείο ελαχίστου είναι εγγυημένη. Επίσης, τονίζουμε ότι αυτό ισχύει για την περίπτωση της απλής στάθμισης περιστροφών, δεν επεκτείνεται όμως στο πρόβλημα της στάθμισης πολλών περιστροφών (καταχώρηση περιστροφών σε κοινό σύστημα).

Συγκεκριμένα, όταν εφαρμοστεί στο πρόβλημα της στάθμισης εκτιμήσεων R_i για μία περιστροφή, δίνει την λύση R_y που ελαχιστοποιεί την ποσότητα:

$$\sum_{i=1}^n d_{\text{geometric}}(R_i, R_y)$$

Η εφαρμογή αυτού του αλγορίθμου, γίνεται για να σταθμίσουμε τις διάφορες εκτιμήσεις που έχουμε μέσω RANSAC για τον σχετικό προσανατολισμό δυο

όψεων, R_{ij} . Λαμβάνουμε έτσι την τελική εκτίμηση για τον σχετικό προσανατολισμό των όψεων R_{ij}^* . Αντί να τρέξουμε τον αλγόριθμο έως της σύγκλιση, εκτελούμε μικρό σταθερό αριθμό επαναλήψεων (10 – 20).

Για το πρόβλημα της καταχώρησης των σχετικών περιστροφών σε κοινό σύστημα αναφοράς, εφαρμόζουμε τον ίδιο αλγόριθμο, ως εξής

1. Δειγματοληπτούμε ένα συνδεδετικό δέντρο στον γράφο των σχετικών περιστροφών. Με εφαρμογή της σχέσης $R_j = R_{ij}R_i$, λαμβάνουμε αρχικές εκτιμήσεις για τους προσανατολισμούς των όψεων R_i
2. Για κάθε κόμβο του γράφου R_i , χρησιμοποιούμε τις διαθέσιμες εκτιμήσεις R_{ij} για να λάβουμε (μη-συμβατές) εκτιμήσεις για την προσανατολισμό της όψης i , $R_i^k, k = 1, 2, \dots$. Σταθμίζουμε τις εκτιμήσεις αυτές με *μία* εκτέλεση του αλγορίθμου Weiszfeld.
3. Το προηγούμενο βήμα, επαναλαμβάνεται για έναν σταθερό αριθμό επαναλήψεων (10 – 20)

Για το Βήμα 1, την αρχικοποίηση δηλαδή του αλγορίθμου, θα αναφέρουμε περισσότερα σε επόμενες Ενότητες

7.2.4 Πρώτη υλοποίηση

Στην αρχική ενσωμάτωση των προηγούμενων αλλαγών στον υπολογισμό των σχέσεων μεταξύ ζευγών φωτογραφιών, κάναμε τις επόμενες επιλογές:

1. Για κάθε ζεύγος (ταιριασμένων) φωτογραφιών, εκτελέσαμε ένα σταθερό αριθμό επαναλήψεων RANSAC. Για να διαλέξουμε ανάμεσα στις δύο λύσεις για την σχετική θέση και προσανατολισμό των καμερών που επιστρέφει η μέθοδος του Κεφαλαίου 5, απαιτούμε τα σημεία του δείγματος RANSAC να είναι *όλα μπροστά από τις κάμερες*.
2. Από κάθε ζεύγος φωτογραφιών ij λαμβάνουμε εκτιμήσεις για τις παραμέτρους f_i, f_j , τις οποίες συμβολίζουμε f_i^{ij}, f_j^{ij} . Για την λήψη των f_i^{ij}, f_j^{ij} , συγκεντρώνουμε τις n εκτιμήσεις για το εστιακό μήκος από τις n επαναλήψεις του RANSAC στο ζεύγος ij και λαμβάνουμε κάποιο στατιστικό μέτρο των εκτιμήσεων αυτών όπως ο μέσος ή η διάμεσος.
3. Αντίστοιχα, ο σχετικός προσανατολισμός R_{ij} προκύπτει με εφαρμογή του αλγορίθμου του Weiszfeld στις εκτιμήσεις για τον προσανατολισμό από τις επαναλήψεις του RANSAC

Στα προηγούμενα, αναφέρουμε για το σημείο 1 ότι παρατηρήθηκε εύκολος διαχωρισμός των δύο λύσεων. Συγκεκριμένα, στις περισσότερες περιπτώσεις, τα σημεία του δείγματος είναι *όλα* μπροστά από την μία λύση και *κανένα* μπροστά από την έτερη λύση. Επίσης, συχνή είναι η περίπτωση που *όλα* τα σημεία είναι πίσω και από τις δύο κάμερες. Γενικά, όταν δεν υπάρχει λύση για την οποία όλα τα σημεία του δείγματος είναι μπροστά από τις κάμερες, απορρίπτουμε και τις δύο λύσεις. Η απόφαση αυτή, *μειώνει τον αριθμό των διαδέσιμων λύσεων σε σχέση με τον αριθμό των επαναλήψεων RANSAC*.

Συνοπτικά, αναφέρουμε ότι η αρχική αυτή προσέγγιση αποτυγχάνει να προσδιορίσει με ακρίβεια τις τιμές των εσωτερικών και εξωτερικών παραμέτρων. Από την πρώτη αυτή υλοποίηση, καταλήξαμε στα επόμενα συμπεράσματα, σχετικά με την δυνατότητα εφαρμογής της μεθόδου:

- Προκειμένου να βελτιώσουμε την επίδοση του συστήματος, εφαρμόσαμε διάφορες μεθόδους μη-γραμμικής βελτιστοποίησης των εκτιμήσεων των μητρών F που λαμβάνουμε από ελάχιστα δείγματα, πριν την εξαγωγή των παραμέτρων f_i, f_j, R_{ij} . Οι μέθοδοι μη-γραμμικής βελτιστοποίησης που εφαρμόσαμε προέρχονται από το Κεφάλαιο 5. Αντίθετα από τις προσδοκίες μας, παρατηρήθηκε ότι τα αποτελέσματα *χειροτερεύουν* και τα σημεία του δείγματος συχνά μετακινούνται πίσω από τις κάμερες μετά την μη-γραμμική βελτιστοποίηση. Θα πρέπει επομένως να εγκαταλείψουμε τέτοιες κατευθύνσεις
- Σε συνέχεια του προηγούμενου σημείου, προτείνουμε τον *γεωμετρικό έλεγχο* των μερικών λύσεων. Συγκεκριμένα, για κάθε λύση f_i, f_j, R_{ij} που προκύπτει από ένα δείγμα του RANSAC, υπολογίζουμε αρχικά ένα σύνολο inliers από τον θεμελιώδη πίνακα F . Στην συνέχεια, απαιτούμε συγκεκριμένο ποσοστό $T_{support}$ των inliers να *βρίσκεται μπροστά από την κάμερα*, προκειμένου να αποδεχθούμε την λύση
- Για τον υπολογισμό των παραμέτρων f_i, f_j , θα πρέπει πάλι να εκμεταλλευθούμε κατάλληλα την *δομή του προβλήματος*. Η προσέγγιση που εφαρμόσαμε θα αναλυθεί σε επόμενη Ενότητα

7.2.5 Βελτιώσεις στην πρώτη υλοποίηση

Εκτίμηση του Εστιακού Μήκους

Παρατηρήσαμε στην αρχική υλοποίηση, ότι η κατανομή των εκτιμήσεων f_i^{ij} δεν είναι αρκετά *συγκεντρωμένη* γύρω από μία τιμή, ώστε να μπορούμε να επιλέξουμε από αυτήν μία τιμή του εστιακού μήκους, για παράδειγμα την μέση τιμή ή την διάμεσο. Αυτό μας επιβάλλει να κρατήσουμε όλες τις τιμές $\{f_i^{ij}\}$ και να αναζητήσουμε κατάλληλη μέθοδο για την στάθμιση όλων των εκτιμήσεων για το εστιακό μήκος f_i , ανεξάρτητα από το ζεύγος από το οποίο προέρχονται.

Έτσι, εισάγουμε την επόμενη διαδικασία. Αρχικά ορίζουμε ένα νέο μέτρο, το οποίο ονομάζουμε *άθροισμα εμπιστοσύνης*. Για να υπολογίσουμε το άθροισμα εμπιστοσύνης, ακολουθούμε την εξής διαδικασία, την οποία εφαρμόζουμε σε κάθε εικόνα.

1. Συγκεντρώνουμε όλες τις εκτιμήσεις f_i^n , ανεξάρτητα από το ζεύγος εικόνων ij από το οποίο προέκυψαν, για την τιμή του εστιακού μήκους f_i
2. Για κάθε εκτίμηση f_i^n , αθροίζουμε τις εκτιμήσεις f_i^k που απέχουν μέχρι 10% από την f_i^n . Το άθροισμα αυτό είναι το *άθροισμα εμπιστοσύνης* cc_i^n της εκτίμησης f_i^n
3. Αφού ολοκληρωθεί η προηγούμενη διαδικασία, κανονικοποιούμε τις τιμές cc_i^n με την μέγιστη τιμή $\max_n cc_i^n$
4. Η εκτίμηση μας για την τιμή του f_i , λαμβάνεται από την μέση τιμή των f_i^n που έχουν άθροισμα εμπιστοσύνης μεγαλύτερη από ένα ποσοστό (95%) του μέγιστου $\max_n cc_i^n$

Ο υπολογισμός του αθροίσματος εμπιστοσύνης για όλες τις εκτιμήσεις f_i^n μπορεί να γίνει σε *γραμμικό χρόνο*, αν έχουμε ταξινομήσει τις εκτιμήσεις

αυτές.

Η προηγούμενη προσέγγιση, βασίζεται στην ακόλουθη προσδοκία. Παρότι οι κατανομές των f από κάθε ζευγάρι δεν είναι αποκαλυπτικές για την τιμή της παραμέτρου, περιμένουμε οι λάθος εκτιμήσεις να διαφέρουν ανάμεσα σε κάθε ζεύγος ενώ ταυτόχρονα κάθε ζεύγος να παρέχει έναν αριθμό σωστών εκτιμήσεων. Έτσι, όταν εξετάζουμε μαζί όλες τις εκτιμήσεις του f , περιμένουμε να παρατηρήσουμε διασπορά των λάθος εκτιμήσεων και συγκέντρωση σχετικά περισσότερων εκτιμήσεων κοντά στην σωστή τιμή της παραμέτρου. Παρόμοιες τεχνικές είναι ήδη αρκετά διαδεδομένες στην βιβλιογραφία, ως για παράδειγμα ο μετασχηματισμός Hough.

Αν και τα αποτελέσματα βελτιώνονται χρησιμοποιώντας το άθροισμα εμπιστοσύνης, ακόμα δεν είναι ικανοποιητικά. Για αυτόν τον λόγο χρησιμοποιούμε την επόμενη τροποποίηση, την οποία ονομάζουμε *συζευγμένο άθροισμα εμπιστοσύνης*. Για να υπολογίσουμε το συζευγμένο άθροισμα εμπιστοσύνης για μία εικόνα i , έχουμε

1. Υποθέτουμε ότι τα απλά αθροίσματα εμπιστοσύνης cc_i^n έχουν υπολογιστεί για όλες τις εικόνες
2. Έστω ότι η εικόνα i έχει ταιριάζει με άλλες m εικόνες. Για να υπολογίσουμε το *συζευγμένο άθροισμα εμπιστοσύνης* Jcc_i^n , κάνουμε τα εξής για κάθε άλλη από τις ταιριασμένες εικόνες $k = 1 \dots m$
 - Από τις εκτιμήσεις για τα εστιακά μήκη που προέρχονται από το ζεύγος ik , f_i^{ik} , επιλέγουμε όσες βρίσκονται αρκετά κοντά (σε απόσταση 10%) από το f_i^n . Ας συμβολίσουμε αυτές τις τιμές $f_{il}^{ik}, l = 1 \dots L$.
 - Κάθε εκτίμηση για το f_i στο ζεύγος ik συνοδεύεται από μια εκτίμηση για το f_k . Παίρνουμε έτσι τις L εκτιμήσεις f_{kl}^{ik} . Βρίσκουμε για τις τιμές αυτές το άθροισμα εμπιστοσύνης από την κατανομή του f_k , και λαμβάνουμε την μέση τιμή αυτών των αθροισμάτων εμπιστοσύνης
3. Το συζευγμένο άθροισμα εμπιστοσύνης Jcc_i^n υπολογίζεται από το άθροισμα των m μέσων τιμών που περιγράψαμε προηγουμένως.
4. Η τελική εκτίμηση του f_i προκύπτει όπως και προηγουμένως, χρησιμοποιώντας όμως τα συζευγμένα αθροίσματα εμπιστοσύνης αντί για τα απλά αθροίσματα εμπιστοσύνης

Ας σχολιάσουμε σύντομα την τελευταία αυτή μέθοδο των συζευγμένων αθροισμάτων. Αν μία εκτίμηση για το f_i δεν είναι κοντά στην πραγματική του τιμή, υποθέσαμε ότι δεν θα υπάρχει μεγάλη συγκέντρωση εκτιμήσεων κοντά σε αυτή την τιμή από το σύνολο των φωτογραφιών, και έτσι προτείνουμε την εισαγωγή του αθροίσματος εμπιστοσύνης. Στην συνέχεια, επεκτείνουμε την συλλογιστική αυτή, με χρήση του γεγονότος ότι κάθε εκτίμηση του εστιακού μήκους f_i έχει γίνει ταυτόχρονα με την εκτίμηση ενός άλλου εστιακού μήκους f_m , αφού χρησιμοποιούμε σχέσεις ζευγών εικόνων. Αν η εκτίμηση για το f_i είναι λάθος, περιμένουμε και η εκτίμηση για το f_m να είναι λάθος. Την ιδιότητα αυτή εκμεταλλευόμαστε για να εισάγουμε το συζευγμένο άθροισμα εμπιστοσύνης.

Στον Πίνακα 7.7, παρουσιάζουμε κάποια αποτελέσματα, ενδεικτικά για την

κακή ποιότητα των αρχικών αποτελεσμάτων και για την επίδραση της εισαγωγής του απλού και συζευγμένου αθροίσματος εμπιστοσύνης. Για τον ίδιο σκοπό, παραθέτουμε και τα διαγράμματα 7.3,7.4,7.5.

| Μέθοδος | Διάμεσος | Άθροισμα Εμπιστοσύνης | Συζευγμένο Άθροισμα Εμπιστοσύνης |
|------------------------|----------|--------------------------|--|
| Μέσο σφάλμα Δf | 0.28 | 0.17 | 0.07 |

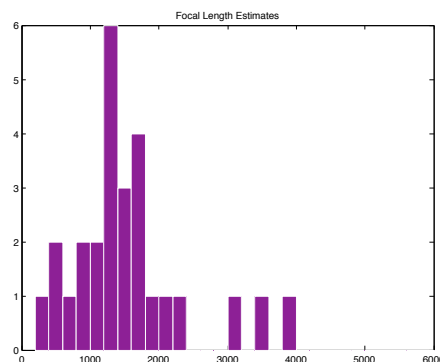
Πίνακας 7.7: Αποτελέσματα στάθμισης εστιακού μήκους στο σύνολο castle-P30. Έγιναν 300 Επαναλήψεις του αλγορίθμου RANSAC, χρησιμοποιήθηκε ο αλγόριθμος 8 σημείων για την εύρεση του θεμελιώδους πίνακα και θέσαμε $T_{support} = 0.8$

Υπολογισμός Θεμελιώδους Πίνακα

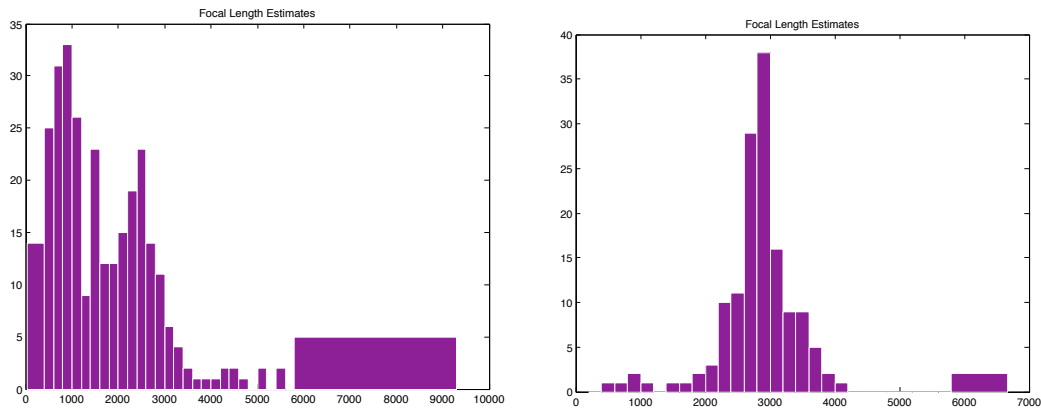
Στην ενότητα αυτή, συγκρίνουμε τα αποτελέσματα στην εκτίμηση του εστιακού μήκους, για διαφορετικές επιλογές των επόμενων παραμέτρων που επηρεάζουν τον υπολογισμό του θεμελιώδους πίνακα F

- Τον αλγόριθμο υπολογισμού του Θεμελιώδους Πίνακα F . Συγκρίνουμε τους αλγορίθμους 7 και 8 σημείων (Εικόνα 7.6)
- Τον μέγιστο αριθμό επαναλήψεων RANSAC. Εναλλακτικά, η διαδικασία σταματάει όταν επιστραφούν περισσότερες λύσεις από έναν προκαθορισμένο αριθμό (Εικόνα 7.6)
- Την τιμή του κατωφλιού $T_{support}$ (Εικόνα 7.7)

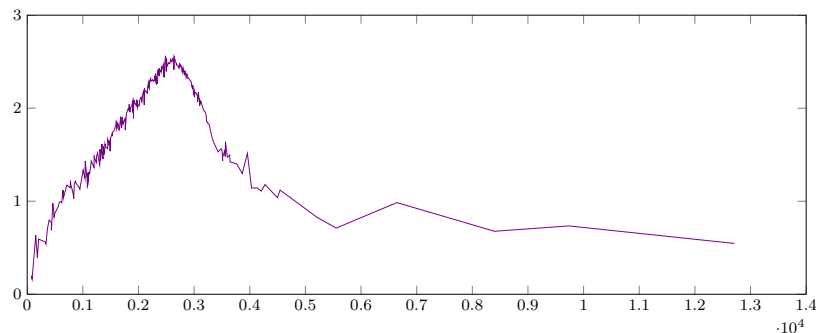
Σχολιάζουμε ότι ο αλγόριθμος των 8 σημείων δίνει ελαφρώς καλύτερα αποτελέσματα σε σχέση με τον αλγόριθμο των 7 σημείων, κυρίως ως προς το εστιακό μήκος, ενώ συνήθως οδηγεί στην ανεύρεση λιγότερων σχέσεων μεταξύ ζευγών φωτογραφιών. Στις επόμενες ενότητες χρησιμοποιούμε τον αλγόριθμο των 8 σημείων.



Σχήμα 7.3: Κατανομή εκτιμήσεων f από ένα ζεύγος καμερών. Η σωστή τιμή του εστιακού μήκους είναι 2760. Το ζεύγος προέρχεται από το σύνολο castle-P30. Παρατηρούμε την δυσκολία ορθής εκτίμησης του f από τα δεδομένα ενός ζεύγους



Σχήμα 7.4: Κατανομή εκτιμήσεων f από όλα τα ταιριασμένα ζεύγη καμερών. Τα δεδομένα προέρχονται από το σύνολο **castle-P30**. Παρατηρούμε ότι για κάποιες κάμερες η τιμή του εστιακού μήκους μπορεί να εκτιμηθεί εύκολα, ενώ για άλλες είναι δύσκολη η εκτίμηση από την κατανομή των f_i^n (βλ. κείμενο)



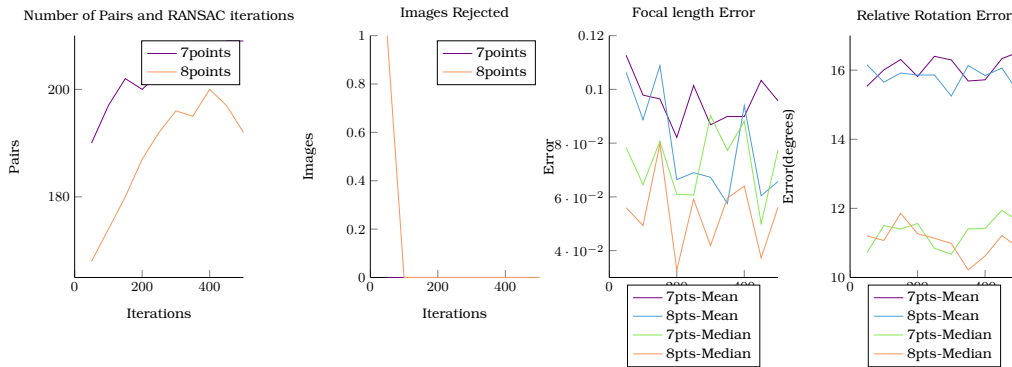
Σχήμα 7.5: Κατανομή συζευγμένου αθροίσματος εμπιστοσύνης Jcc_i^n . Παρατηρούμε ότι υπάρχει έντονη κορυφή της κατανομής κοντά στην πραγματική τιμή του εστιακού μήκους (2760). Τα δεδομένα προέρχονται από το σύνολο δεδομένων **castle-P30**.

Επίσης, όσον αφορά το κατώφλι $T_{support}$ παρατηρούμε ότι αύξηση του βελτιώνει τις εκτιμήσεις προσανατολισμού και εστιακού μήκους, μειώνει όμως τον αριθμό των σχέσεων μεταξύ ζευγών φωτογραφιών. Για να σταθμίσουμε τις δύο αυτές παραμέτρους, στην πλειονότητα των επόμενων πειραμάτων χρησιμοποιούμε $T_{support} = 0.75$

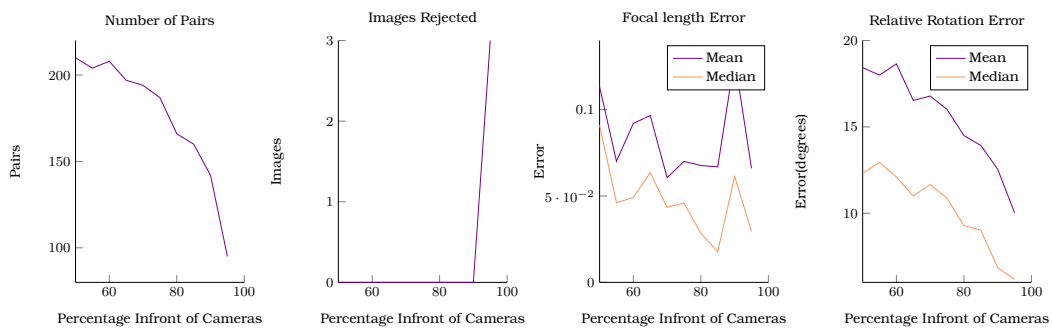
Τέλος, η αύξηση του αριθμού των επαναλήψεων του RANSAC δεν επιδρά με ιδιαίτερα έντονο τρόπο στα αποτελέσματα, πέραν του αριθμού των σχέσεων μεταξύ ζευγών φωτογραφιών. Για το σφάλμα στο εστιακό μήκος, η μετριοπαθής αύξηση των επαναλήψεων δείχνει να βελτιώνει κάπως τα αποτελέσματα, ενώ τέλος το σφάλμα στις εκτιμήσεις του σχετικού προσανατολισμού των εικόνων επηρεάζεται λιγότερο.

Συγκρίσεις με προσέγγιση Κεφαλαίου 6

Στην Ενότητα αυτή, θα συγκρίνουμε την νέα μέθοδο με αυτή του Κεφαλαίου 6. Για τις συγκρίσεις, χρησιμοποιούμε το σύνολο **castle-P30**. Επιλέξαμε το συγκεκριμένο σύνολο, γιατί σε αυτό είχαμε τα χειρότερα αποτελέ-



Σχήμα 7.6: Διάφορα μέτρα επίδοσης για διαφορετικό αριθμό επαναλήψεων RANSAC. Συγκρίνουμε ακόμα τους αλγορίθμους 7 και 8 σημείων για την εύρεση του θεμελιώδους πίνακα. Τα αποτελέσματα αφορούν το σύνολο **castle-P30** και $T_{support} = 0.8$



Σχήμα 7.7: Διάφορα μέτρα επίδοσης για διαφορετικό ποσοστό σημείων μπροστά από τις κάμερες, $T_{support}$. Χρησιμοποιούμε τον αλγόριθμο 8 σημείων για την εύρεση του θεμελιώδους πίνακα. Τα αποτελέσματα αφορούν το σύνολο **castle-P30** και 400 επαναλήψεις RANSAC

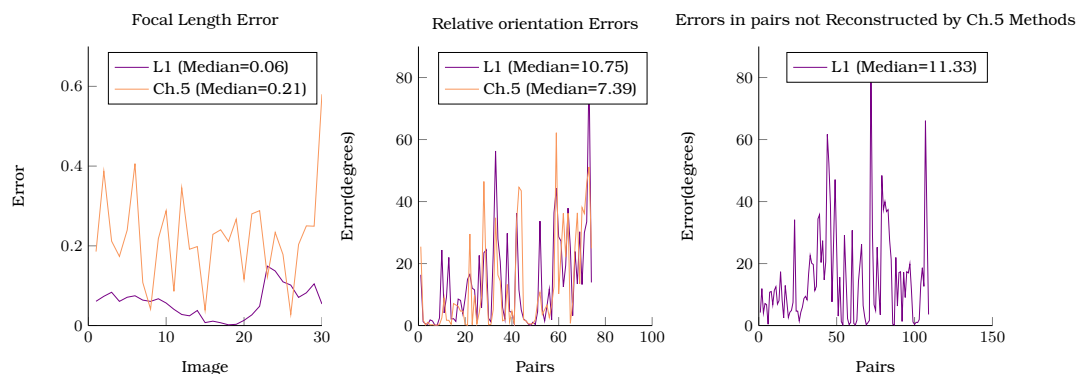
σματα ανάμεσα στα σύνολα δεδομένων [216]. Στα επόμενα, χρησιμοποιούμε για το σύνολο την μέθοδο που περιγράφεται στον Πίνακα 6.2.

Στην εικόνα 7.8, παρουσιάζουμε μια αρχική σύγκριση που αφορά την εκτίμηση του *εστιακού μήκους* και την εκτίμηση του *σχετικού προσανατολισμού* για ζεύγη εικόνων. Σημειώνουμε ότι έχουμε χρησιμοποιήσει μια βασική υλοποίηση του συστήματος, χωρίς τις βελτιώσεις που θα αναφέρουμε αργότερα. Παρατηρούμε ότι

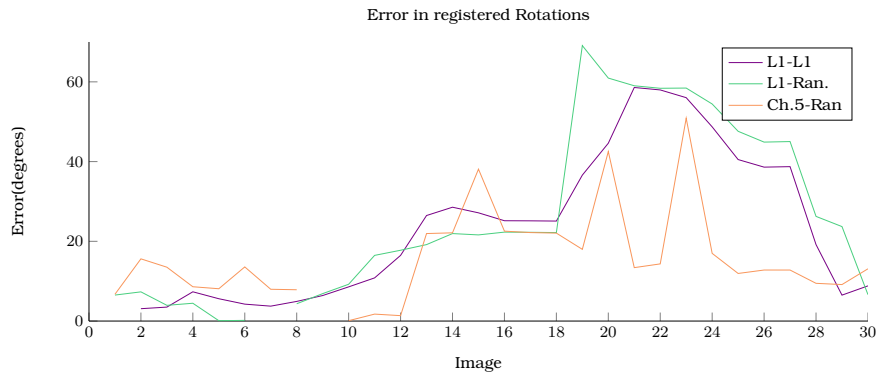
- Όσον αφορά την εκτίμηση του εστιακού μήκους, τα αποτελέσματα είναι ξεκάθαρα βελτιωμένα
- Αντίστοιχα, μεγάλη βελτίωση παρουσιάζει ο αριθμός των σχέσεων μεταξύ φωτογραφιών που ανιχνεύουμε. Η βελτίωση αυτή μπορεί να επηρεάσει τα επόμενα στάδια της ανακατασκευής και να οδηγήσει σε μεγαλύτερη πληρότητα στις ανακατασκευασμένες επιφάνειες
- Ως προς την εκτίμηση του σχετικού προσανατολισμού ζευγών καμερών, στα ζεύγη που ανιχνεύονται και από τις δύο μεθόδους, παρατηρούμε παρόμοια επίδοση στα ζεύγη που έχουν μικρό σφάλμα (και επομένως προσδιορίζονται με ικανοποιητική ακρίβεια), και μεγαλύτερο σφάλμα στα ζεύγη στα οποία η μέθοδος του Κεφαλαίου 6 δίνει εκτιμήσεις με μεγάλο σφάλμα

Στην Εικόνα 7.9, βλέπουμε τα αποτελέσματα της *καταχώρησης* των περιστροφών σε κοινό πλαίσιο αναφοράς. Οι λεπτομέρειες των Μεθόδων δίνονται στην λεζάντα της εικόνας. Παρατηρούμε ότι

- Η εφαρμογή του αλγορίθμου Weiszfeld βελτιώνει τα αποτελέσματα, αν και δεν επικρατεί τελείως στην μέθοδο RANSAC για κάθε κάμερα (Σύγκριση **L1-L1** με **L1-Ran**. Αυτό οφείλεται μερικώς στο μικρό μέγεθος του συνόλου δεδομένων. Αξίζει όμως να παρατηρήσουμε, την αύξηση του σφάλματος μακριά από την κάμερα που έχει τεθεί στην αρχή των αξόνων (κενό στις γραφικές παραστάσεις). Η παρατήρηση αυτή θα οδηγήσει στην βελτίωση της αρχικής υλοποίησης για τον αλγόριθμο του Weiszfeld, όπως θα δούμε επόμενα.
- Επίσης, ή μέθοδος των προηγούμενων Κεφαλαίων δίνει ακόμα καλύτερα αποτελέσματα από τις νέες μεθόδους.



Σχήμα 7.8: Σύγκριση της μεθόδου που εισάγουμε στο παρόν Κεφάλαιο (L1 στα διαγράμματα) και της μεθόδου που αναπτύχθηκε στα προηγούμενα Κεφάλαια 5,6 (Ch.5 στα διαγράμματα)



Σχήμα 7.9: Σύγκριση επίδοσης διαφορετικών μεθόδων για την καταχώρηση των περιστροφών σε κοινό πλαίσιο αναφοράς. Στο διάγραμμα, συμβολίζουμε με **L1-L1** την μέθοδο που εισάγουμε στο παρόν Κεφάλαιο, με **L1-Ran** την μέθοδο που εφαρμόζει τις τεχνικές του παρόντος Κεφαλαίου για τις σχέσεις μεταξύ ζευγών εικόνων και την εκτίμηση των εστιακών μηκών και στην συνέχεια καταχωρεί τις περιστροφές με τυχαία δειγματοληψία σε συνδεδετικά δέντρα (Κεφάλαιο 6). Τέλος με **Ch.5-Ran** συμβολίζουμε την μέθοδο που αναπτύχθηκε στα Κεφάλαια 5,6.

Τα αποτελέσματα αυτά, θα πρέπει να ερμηνευτούν προσεκτικά, ώστε να μην οδηγηθούμε σε εσφαλμένα συμπεράσματα. Στην Εικόνα 7.9, παρουσιάσαμε το σφάλμα *ως προς τις γνωστές, πραγματικές θέσεις* των καμερών. Όμως, όπως οφείλουμε να τονίσουμε, το σφάλμα αυτό δεν μπορεί να υπολογιστεί από τις ίδιες τις μεθόδους υπολογισμού των προσανατολισμών για τις κάμερες. *Αντίθετα*, το κριτήριο το οποίο βελτιστοποιούν οι δύο μέθοδοι είναι το σφάλμα στις σχέσεις $R_i = R_{ij}R_j$. Συγκεκριμένα:

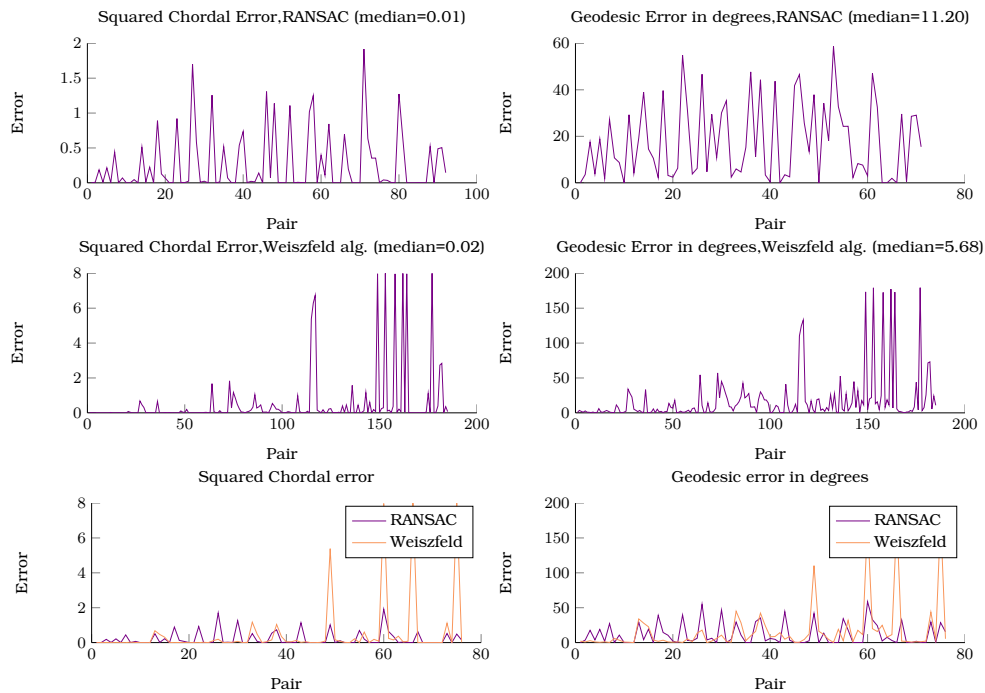
- Η μέθοδος RANSAC χρησιμοποιεί την προηγούμενη σχέση για να κατατάξει τις σχετικές περιστροφές R_{ij} σε outliers και inliers και προσπαθεί να μεγιστοποιήσει το πλήθος των inliers
- Ο Αλγόριθμος Weiszfeld βρίσκει την περιστροφή διάμεσο χρησιμοποιώντας την προηγούμενη σχέση, ελαχιστοποιεί δηλαδή το L_1 σφάλμα ως προς την σχέση αυτή

Η Εικόνα 7.10, συγκρίνει την επίδοση της μεθόδου του Κεφαλαίου 6 με την μέθοδο που εισάγουμε στο παρόν Κεφάλαιο. Από τα διαγράμματα αυτά, είναι σαφής η *υπεροχή της νέας μεθόδου σε σχέση με το ταίριασμα στις διαδέσιμες εκτιμήσεις για τις σχετικές περιστροφές*.

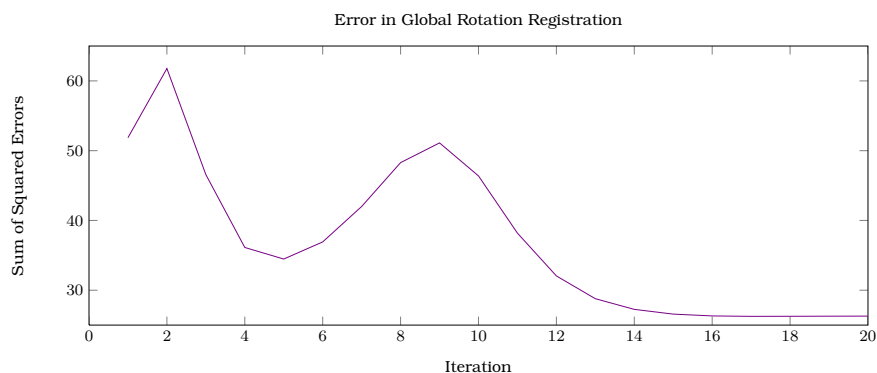
Επίσης, στην Εικόνα 7.11, εξετάζουμε την *σύγκλιση* της μεθόδου Weiszfeld για την καταχώρηση των περιστροφών. Είναι φανερό, ότι 20 επαναλήψεις αρκούν για την σύγκλιση της μεθόδου, και ότι το *συνολικό σφάλμα έχει μειωθεί σημαντικά στο τέλος της εκτέλεσης του αλγορίθμου σε σχέση με την τυχαία αρχικοποίηση του*.

Με την εισαγωγή των προηγούμενων στοιχείων, μπορούμε πλέον να απαντήσουμε στην ερώτηση: *Πως μπορεί να βελτιωθεί η επίδοση της νέας μεθόδου στην καταχώρηση περιστροφών*; Η απάντησή μας είναι ότι θα πρέπει να βελτιώσουμε τις διαθέσιμες σχετικές περιστροφές. Αυτό μπορεί να γίνει είτε με *βελτίωση των εκτιμήσεων των σχετικών περιστροφών* είτε με *απόρριψη των σχετικών περιστροφών που παρουσιάζουν μεγάλο σφάλμα*.

Για την επίτευξη των προηγούμενων στόχων δοκιμάσαμε:



Σχήμα 7.10: Σφάλμα στις σχέσεις $R_i = R_{ij}R_j$. Συγκρίνουμε την καταχώρηση με RANSAC στα συνδεδετικά δέντρα και με εφαρμογή του αλγορίθμου Weiszfeld.



Σχήμα 7.11: Εξέλιξη του σφάλματος στην σχέση $R_i = R_{ij}R_j$ σε σχέση με τις επαναλήψεις του αλγορίθμου Weiszfeld

- Εφαρμογή Bundle Adjustment στις εκτιμήσεις για τις σχετικές περιστροφές R_{ij} . Δοκιμάσαμε να βελτιστοποιήσουμε κατά την εκτέλεση Bundle Adjustment και μήτρες εσωτερικών παραμέτρων 2, 4, 5 βαθμών ελευθερίας ή να κρατήσουμε τις εσωτερικές παραμέτρους σταθερές. Λάβαμε τα καλύτερα αποτελέσματα όταν δώσαμε 2 βαθμούς ελευθερίας στην μήτρα. Η μέθοδος αρχικοποιείται με τις εκτιμήσεις για τα R_{ij} που προκύπτουν με χρήση του αλγορίθμου Weiszfeld (L1-BA)
- Από το σύνολο των εκτιμήσεων R_{ij}^n για την σχέση R_{ij} , απόρριψη όσων (ενδεχομένως και όλων) με κριτήρια που σχετίζονται με το σφάλμα των R_{ij}^n, R_{ij} , το σφάλμα στο εστιακό μήκος και τον αριθμό των αντιστοιχιών στο ζεύγος φωτογραφιών. Στην συνέχεια, εφαρμόζουμε ξανά τον αλγόριθμο του Weiszfeld στα δείγματα R_{ij}^n που απομένουν (L1-Reaveraged)
- Απόρριψη των εκτιμήσεων R_{ij} . Χρήση των εκτιμήσεων για τα εστιακά μήκη f_i και για τις έγκυρες αντιστοιχίες (inliers) κάθε ζεύγους. Εύρεση των R_{ij} με εφαρμογή RANSAC και του καθιερωμένου αλγορίθμου 5-σημείων (calibrated relative rotation problem). Βελτίωση του αποτελέσματος με λίγες επαναλήψεις (20) Bundle Adjustment

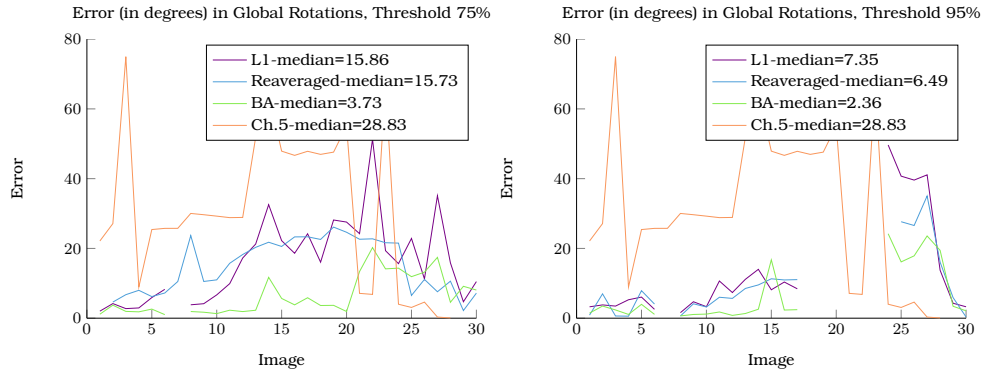
Τα αποτελέσματα των προηγούμενων προσπαθειών φαίνονται στην Εικόνα 7.12. Παρατηρούμε ότι η εφαρμογή Bundle Adjustment βελτιώνει σημαντικά τα αποτελέσματα, ενώ δεν εισάγει επιπλέον παραμέτρους στο πρόβλημα. Έτσι, στα επόμενα πειράματα εφαρμόζουμε Bundle Adjustment στις εκτιμήσεις για τις σχέσεις R_{ij} πριν προχωρήσουμε στην καταχώρηση των περιστροφών. Τέλος, συγκρίσεις των τελικών αποτελεσμάτων μετά την εκτέλεση όλων των σταδίων του SfM, μπορούν να γίνουν με τα στοιχεία που δίνονται στον Πίνακα 7.8. Αντίστοιχα, τα Αποτελέσματα της στάθμισης των σχετικών περιστροφών R_{ij} , δίνονται στον Πίνακα 7.9.

| | L1 | L1-Reav. | L1-BA | L1 | L1-Reav. | L1-BA | Ch.5 |
|---------------------|------|----------|-------|------|----------|-------|-------|
| Μέσο ΔR | 3.05 | 1.06 | 1.66 | 7.57 | 3.12 | 2.35 | 68.14 |
| Διάμεσος ΔR | 0.82 | 0.79 | 0.82 | 0.78 | 0.71 | 0.78 | 55.87 |
| Μέσο Δt | 4.27 | 2.52 | 3.01 | 9.11 | 5.83 | 5.39 | 40.10 |
| Διάμεσος Δt | 2.58 | 2.46 | 2.75 | 3.54 | 2.52 | 2.45 | 33.92 |
| Εικόνες | 27 | 23 | 29 | 24 | 21 | 20 | 24 |

Πίνακας 7.8: Τελικά αποτελέσματα SfM στο σύνολο **castle-P30**, για διαφορετικές μεθόδους εκτίμησης σχετικού προσανατολισμού R_{ij} και καταχώρησης περιστροφών R_i . Στις στήλες 1 – 3 δίνουμε τα αποτελέσματα για $T_{support} = 0.75$ και στις στήλες 4 – 6 για $T_{support} = 0.95$

Σύγκριση με τον αλγόριθμο των 5 σημείων

Στην Ενότητα αυτή συγκρίνουμε δύο μεθόδους για την αρχικοποίηση του αλγορίθμου Bundle Adjustment για την βελτίωση των εκτιμήσεων των σχετικών περιστροφών R_{ij} :



Σχήμα 7.12: Σφάλμα στις εκτιμήσεις των περιστροφών R_i , για διάφορες τροποποιήσεις της μεθόδου εκτίμησης σχετικών περιστροφών R_{ij} . Οι τιμές του $T_{support}$ δίνονται στους τίτλους των διαγραμμάτων. Η μέθοδος *Ch.5*, αν και ανεξάρτητη του $T_{support}$, έχει σχεδιαστεί και στα δύο διαγράμματα για διευκόλυνση της σύγκρισης. Τονίζουμε ότι αν και οι εκτιμήσεις είναι πιο ακριβείς για αυστηρές τιμές $T_{support}$, αυξάνεται παράλληλα και ο αριθμός των καμερών που απορρίπτονται (κενά στις γραφικές παραστάσεις)

| | L1 | L1-Reav. | L1-BA | L1 | L1-Reav. | L1-BA | Ch.5 |
|---------------------|-------|----------|-------|-------|----------|-------|-------|
| Μέσο ΔR | 15.09 | 6.92 | 12.19 | 10.82 | 5.39 | 9.11 | 19.47 |
| Διάμεσος ΔR | 9.63 | 4.36 | 3.98 | 5.97 | 2.50 | 4.34 | 8.62 |
| Ζεύγη | 186 | 54 | 186 | 186 | 54 | 186 | 82 |

Πίνακας 7.9: Αποτελέσματα προσδιορισμού σχετικού προσανατολισμού R_{ij} στο σύνολο *castle-P30*, για διαφορετικές μεθόδους εκτιμήσεις σχετικού προσανατολισμού R_{ij} . Στις στήλες 1–3 δίνουμε τα αποτελέσματα για $T_{support} = 0.75$ και στις στήλες 4 – 6 για $T_{support} = 0.95$

1. Αρχικοποίηση με τον αλγόριθμο του Weiszfeld
2. Αρχικοποίηση με τον αλγόριθμο των 5 σημείων. Για την εφαρμογή του αλγορίθμου, χρησιμοποιούμε τα f_i που επιστρέφει η μέθοδος που αναλύσαμε νωρίτερα στο Κεφάλαιο

Η μέθοδος των 5 σημείων είναι η καθιερωμένη στην βιβλιογραφία μέθοδος για την αρχική εκτίμηση των R_{ij} , θεωρώντας τα f_i γνωστά.

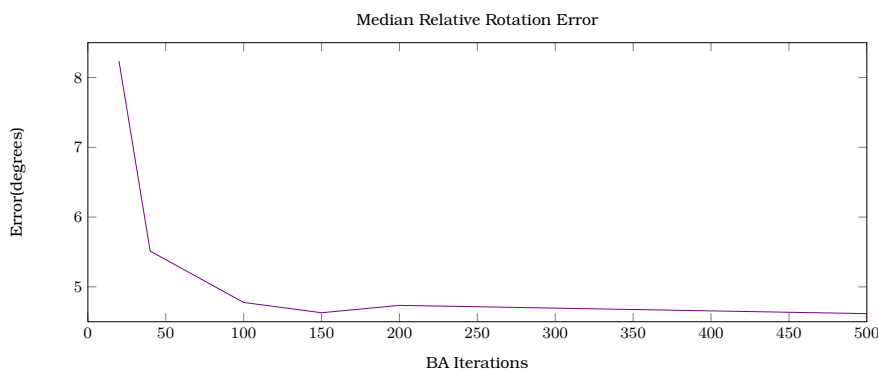
Στον Πίνακα 7.10 δίνονται τα αποτελέσματα των μεθόδων με εφαρμογή 20 επαναλήψεων Β.Α. για την βελτίωση των R_{ij} . Παρατηρούμε ότι οι δύο μέθοδοι έχουν παρόμοια επίδοση. Το αποτέλεσμα αυτό είναι σημαντικό, καθώς δείχνει ότι η μέθοδος που προτείνουμε είναι εφάμιλλη με την καθιερωμένη στην βιβλιογραφία. Συνεπώς, για την βελτίωση των εκτιμήσεων των σχετικών περιστροφών, θα πρέπει να στραφούμε σε άλλες παραμέτρους του προβλήματος, όπως οι αρχικές αντιστοιχίες και οι εκτιμήσεις για τα f_i .

| | Μέσος | Διάμεσος | $\Delta R < 10^\circ$ | $\Delta R < 5^\circ$ |
|------------------|-------|----------|-----------------------|----------------------|
| Weiszfeld | 11.40 | 3.49 | 123 | 106 |
| 5point | 9.99 | 3.77 | 128 | 104 |

Πίνακας 7.10: Αποτελέσματα εύρεσης σχετικού προσανατολισμού R_{ij} στο σύνολο **castle-P30**. Η αρχικοποίηση γίνεται με τον αλγόριθμο Weiszfeld ή τον αλγόριθμο των 5 σημείων. Ακολουθούν 20 επαναλήψεις Weiszfeld, με σταθερές μήτρες K , όπως αυτές υπολογίζονται με τις μεθόδους που περιγράψαμε στο Κεφάλαιο

Αριθμός επαναλήψεων Bundle Adjustment

Οι επιδόσεις του συστήματος για διαφορετικές τιμές επαναλήψεων Bundle Adjustment φαίνονται στην Εικόνα 7.13. Παρατηρούμε ότι γενικά, δεν απαιτείται μεγάλος αριθμός επαναλήψεων της μεθόδου. Αναφέρουμε ότι εφαρμόσαμε Bundle Adjustment, αφήνοντας το εστιακό μήκος να μεταβληθεί, όχι όμως και τις υπόλοιπες εσωτερικές παραμέτρους των καμερών.



Σχήμα 7.13: Διάμεσος του σφάλματος στην εκτίμηση των σχετικών περιστροφών R_{ij} , για διάφορους αριθμούς επαναλήψεων ΒΑ. Τα αποτελέσματα αφορούν το σύνολο **castle-P30**. Κατά την μη-γραμμική βελτιστοποίηση, αφήσαμε το εστιακό μήκος ελεύθερο να μεταβληθεί.

Αρχικοποίηση αλγορίθμου Weiszfeld και σύγκλιση επαναληπτικής μεθόδου

Στην βιβλιογραφία [89, 86], το θέμα της αρχικοποίησης του αλγορίθμου Weiszfeld για το πρόβλημα της στάθμισης πολλών περιστροφών (καταχώρηση περιστροφών), δεν αναπτύσσεται ικανοποιητικά. Στην συγκεκριμένη Ενότητα, συγκρίνουμε διαφορετικές μεθόδους επιλογής ενός αρχικού συνδετικού δέντρου για τον γράφο των σχετικών περιστροφών. Επίσης, συγκρίνουμε και με την μέθοδο τυχαίας δειγματοληψίας συνδετικών δένδρων του Κεφαλαίου 6.

Για τα σύνολα δεδομένων [216], εφαρμόσαμε τις επόμενες μεθόδους στάθμισης περιστροφών

- Δειγματοληψία ενός τυχαίου συνδετικού δένδρου. Ως ρίζα διαλέγουμε τον κόμβο με τους περισσότερους γείτονες. Η πιθανότητα επιλογής της ακμής e_{ij} είναι ανάλογη του αριθμού των αντιστοιχιών για το ζεύγος καμερών ij . Η κάμερα στη ρίζα του συνδετικού δένδρου, η οποία τέθηκε αρχικά στην αρχή των αξόνων, αφήνεται ελεύθερη να μεταβληθεί στις επόμενες επαναλήψεις του αλγορίθμου Weiszfeld (*L1-1*)
- Δειγματοληψία ενός τυχαίου συνδετικού δένδρου. Ως ρίζα διαλέγουμε τον κόμβο με τους περισσότερους γείτονες. Η πιθανότητα επιλογής της ακμής e_{ij} είναι ανάλογη του αριθμού των αντιστοιχιών για το ζεύγος καμερών ij . Η κάμερα στη ρίζα του συνδετικού δένδρου, η οποία τέθηκε αρχικά στην αρχή των αξόνων, δεν αφήνεται να μεταβληθεί (*L1-2*)
- Τυχαία δειγματοληψία συνδετικών δένδρων, όπως στο Κεφάλαιο 6. Τονίζουμε ότι δεν εφαρμόζουμε στάθμιση περιστροφών στην περίπτωση αυτή (*RANSAC*)
- Τυχαία δειγματοληψία συνδετικών δένδρων, όπως στο Κεφάλαιο 6. Σε κάθε επανάληψη της διαδικασίας *RANSAC*, αν το δείγμα είναι το καλύτερο που έχει επιλεγεί ως εκείνη την στιγμή, εφαρμόζουμε στις περιστροφές στάθμιση με τον αλγόριθμο του Weiszfeld. Η μέθοδος αυτή, είναι παρόμοια στην έμπνευση της με την γενική μέθοδο *LO-RANSAC* [38] (*LO-R*)
- Επιλογή κοντούτερου συνδετικού δένδρου (*shortest spanning tree*). Στην συνέχεια εφαρμόζεται στάθμιση των περιστροφών με τον αλγόριθμο του Weiszfeld (*Short*)

Ας σχολιάσουμε σύντομα την τελευταία μέθοδο *Short*. Κατά την αρχική εκτίμηση των περιστροφών R_i , περιμένουμε το σφάλμα να αυξάνεται καθώς απομακρυνόμαστε από την ρίζα του συνδετικού δένδρου, λόγω της επαναληπτικής εφαρμογής της σχέσης $R_i = R_{ij}R_j$. Για αυτό, προσπαθούμε να βρούμε το συνδετικό δένδρο, στο οποίο κάθε κόμβος βρίσκεται στην ελάχιστη απόσταση από την ρίζα, δηλαδή το *shortest spanning tree*. Σύντομα, περιγράφουμε τον αλγόριθμο που εφαρμόσαμε

1. Εύρεση των συντομότερων μονοπατιών στον γράφο των σχετικών περιστροφών (*all pairs shortest paths*), με χρήση του αλγορίθμου των *Floyd-Warshall*
2. Για ρίζα του συνδετικού δένδρου, επιλέγουμε κόμβο ώστε η μέγιστη συντομότερη απόσταση να ελαχιστοποιείται (το μακρύτερο συντομότερο

μονοπάτι που ξεκινάει από τον κόμβο n_i είναι το ελάχιστο δυνατό). Το συνδεδειγμένο δένδρο προκύπτει με *BFS* διάσχιση του γράφου ξεκινώντας από τον κόμβο ρίζα. Οι ισοπαλίες επιλύονται αυθαίρετα, τόσο στην διάσχιση *BFS* όσο και στην επιλογή ρίζας

Για το βήμα 2, δοκιμάσαμε και διαφορετικά κριτήρια για την επίλυση των περιπτώσεων ισοπαλίας. Για παράδειγμα, δοκιμάσαμε να επιλέξουμε το συνδεδειγμένο δένδρο με την μεγαλύτερη (με βάση το πλήθος των αντιστοιχιών) ελάχιστη ακμή ή με το μεγαλύτερο άθροισμα ακμών. Διαπιστώσαμε όμως ότι τα κριτήρια αυτά δεν επηρεάζουν σημαντικά τα αποτελέσματα.

Για να ελέγξουμε την ακρίβεια αλλά και την αξιοπιστία των μεθόδων, εφαρμόσαμε 100 επαναλήψεις της καταχώρησης περιστροφών για κάθε μέθοδο (και σε κάθε σύνολο). Ο μέσος όρος και η τυπική απόκλιση των αποτελεσμάτων δίνονται στους Πίνακες 7.11, 7.12. Παρατηρούμε ότι είναι σημαντικό να μην κρατήσουμε την ρίζα του δένδρου στην αρχή των αξόνων κατά την επαναληπτική στάθμιση των περιστροφών. Επίσης, η στάθμιση περιστροφών δίνει μικρότερο μέσο σφάλμα και με μικρότερη τυπική απόκλιση σε σχέση με την εφαρμογή *RANSAC*. Στην συνέχεια του Κεφαλαίου χρησιμοποιούμε την μέθοδο *L1-1*. Τέλος, στα πειράματά μας χρησιμοποιήσαμε 20 επαναλήψεις του

| Σύνολο | L1-1 | L1-2 | RANSAC | LO-R | Short |
|---------------|-------|-------|--------|-------|-------|
| castle-P30 | 15.05 | 15.71 | 17.41 | 15.58 | 17.27 |
| castle-P19 | 22.13 | 22.61 | 30.94 | 23.61 | 26.60 |
| entry-P10 | 6.64 | 11.16 | 29.16 | 7.26 | 6.63 |
| Herz-Jesu-P8 | 2.22 | 3.99 | 3.18 | 2.12 | 1.70 |
| Herz-Jesu-P25 | 0.66 | 0.90 | 1.69 | 0.63 | 0.69 |
| Fountain-P11 | 0.49 | 0.56 | 0.80 | 0.53 | 0.48 |

Πίνακας 7.11: Σύγκριση μεθόδων αρχικοποίησης αλγορίθμου Weiszfeld για το πρόβλημα της καταχώρησης περιστροφών. Σε κάθε σύνολο δεδομένων, έγιναν 100 επαναλήψεις κάθε μεθόδου, για τις μεθόδους με πιθανοτικά χαρακτηριστικά. Στον Πίνακα δίνουμε την μέση τιμή του σφάλματος ΔR σε σχέση με τις ορθές μήτρες R_i . Για τον αλγόριθμο του Weiszfeld, εκτελέσαμε 20 επαναλήψεις. Οι λεπτομέρειες των μεθόδων δίνονται στο κείμενο

αλγορίθμου.

Τελευταία σχόλια

Προκειμένου να βελτιώσουμε περαιτέρω τα αποτελέσματα, προσπαθήσαμε να συσχετίσουμε διάφορα διαθέσιμα μεγέθη με τα άγνωστα $d(R_{averaged}, R_{GT})$, Δf . Στόχος μας είναι να εισάγουμε κριτήρια για την απόρριψη σχέσεων μεταξύ ζευγών φωτογραφιών. Παρατηρήσαμε ότι:

- μεγέθη όπως το σφάλμα στο οποίο συγκλίνει ο αλγόριθμος του Weiszfeld, ο αριθμός των λύσεων που βρέθηκαν από τις επαναλήψεις *RANSAC* δεν συνδέονται ικανοποιητικά με το σφάλμα $d(R_{averaged}, R_{GT})$
- ο αριθμός των αντιστοιχιών μεταξύ δύο εικόνων συσχετίζεται καλύτερα με την ακρίβεια στην εκτίμηση του εστιακού μήκους. Θα μπορούσαμε

| Σύνολο | L1-1 | L1-2 | RANSAC | LO-R |
|---------------|------|------|--------|------|
| castle-P30 | 3.01 | 3.32 | 8.65 | 7.61 |
| castle-P19 | 3.02 | 1.92 | 10.36 | 2.48 |
| entry-P10 | 0.01 | 2.19 | 12.33 | 0.70 |
| Herz-Jesu-P8 | 0.16 | 0.83 | 0.75 | 0.44 |
| Herz-Jesu-P25 | 0.20 | 0.36 | 0.67 | 0.11 |
| Fountain-P11 | 0.01 | 0.10 | 0.19 | 0.04 |

Πίνακας 7.12: Σύγκριση μεθόδων αρχικοποίησης αλγορίθμου Weiszfeld για το πρόβλημα της καταχώρησης περιστροφών. Στον πίνακα δίνουμε την τυπική απόκλιση του σφάλματος ΔR . Σε κάθε σύνολο δεδομένων, έγιναν 100 επαναλήψεις κάθε μεθόδου. Στον πίνακα δίνουμε την *τυπική απόκλιση* του σφάλματος ΔR . Για τον αλγόριθμο του Weiszfeld, εκτελέσαμε 20 επαναλήψεις. Οι λεπτομέρειες των μεθόδων δίνονται στο κείμενο

να εκμεταλλευτούμε την ιδιότητα αυτή σχηματίζοντας ένα κατάλληλα σταθμισμένο συζευγμένο άθροισμα εμπιστοσύνης

Παρόμοια, αν εφαρμόσουμε Bundle Adjustment για την βελτίωση των εκτιμήσεων των σχετικών περιστροφών R_{ij} , μπορούμε να απορρίψουμε κάποιες σχέσεις ζευγών από την μεταβολή στην τιμή του εστιακού μήκους πριν και μετά την εφαρμογή B.A. Για παράδειγμα, στο σύνολο **castle-P30**, αν απορρίψουμε όλα τα ζεύγη για τα οποία $\Delta(f_{est}, f_{BA}) > 0.6$, λιγότερο από το 5% των ζευγών που απορρίψαμε έχουν σφάλμα $\Delta R < 10^\circ$.

7.3 Αποτελέσματα

7.3.1 Αξιολόγηση με δείκτες σφάλματος

Στην Ενότητα αυτή, παρουσιάζουμε τα Αποτελέσματα στα διάφορα σύνολα του [216].

| Σύνολο | Δt | ΔR | Δf | Εικόνες |
|---------------|---------------------------|---------------------------|--------------------------|---------|
| castle-P30 | 3.1046 ^{13.2945} | 1.0670 ^{22.9021} | 0.0389 ^{0.0504} | 29 |
| castle-P19 | 7.3504 ^{18.9373} | 4.1718 ^{26.5797} | 0.0586 ^{0.0908} | 15 |
| entry-P10 | 4.6299 ^{5.7162} | 4.6702 ^{3.6610} | 0.2118 ^{0.2062} | 10 |
| Herz-Jesu-P8 | 1.0027 ^{0.9962} | 0.6872 ^{0.5674} | 0.0266 ^{0.0255} | 8 |
| Herz-Jesu-P25 | 0.4158 ^{0.6664} | 0.3118 ^{0.3457} | 0.0049 ^{0.0097} | 25 |
| Fountain-P11 | 0.4406 ^{0.6758} | 0.4150 ^{0.3923} | 0.0095 ^{0.0113} | 11 |

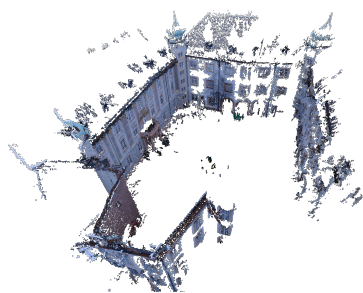
Πίνακας 7.13: Σφάλμα στην εκτίμηση εστιακού μήκους, θέσης και προσανατολισμού, στην τελική ανακατασκευή. Με την συνήθη γραμματοσειρά, δίνεται η διάμεσος των σχετικών μεγεθών, ενώ με *μικρούς πλάγιους* γραμματοσειρά η μέση τιμή. Ο συνδυασμός των δύο αυτών μεγεθών δίνει εικόνα τόσο για το σφάλμα στο σύνολο όσο και για την ύπαρξη καμερών με μεγαλύτερο σφάλμα. Τα ΔR , Δt δίνονται σε μοίρες



Σχήμα 7.14: 2Δ φωτογραφίες που χρησιμοποιήθηκαν σαν είσοδο στο σύστημα για ανακατασκευή Fountain-P11

7.3.2 Οπτική Αξιολόγηση

Πυκνό σύννεφο σημείων MvS



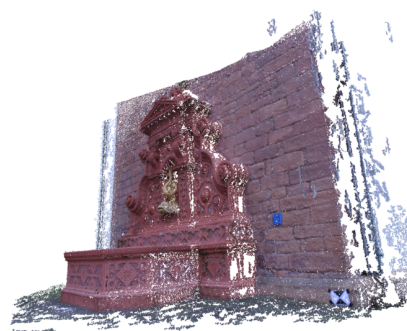
(α) castle-P30



(β) castle-P19



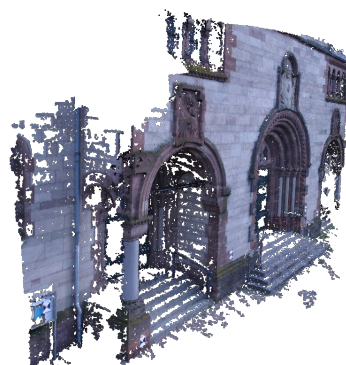
(γ) Fountain-P11



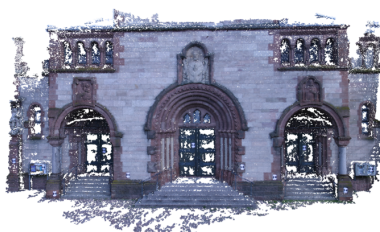
(δ) Fountain-P11



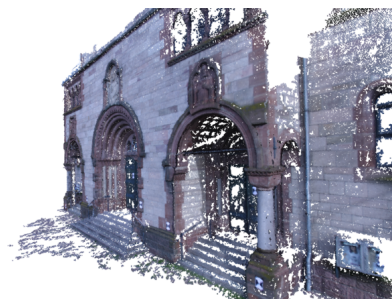
(ε) Herz-Jesu-P8



(στ) Herz-Jesu-P8



(ζ) Herz-Jesu-P25



(η) Herz-Jesu-P25

Σχήμα 7.15: Αποτελέσματα Mvs-Πυκνό σύννεφο σημείων

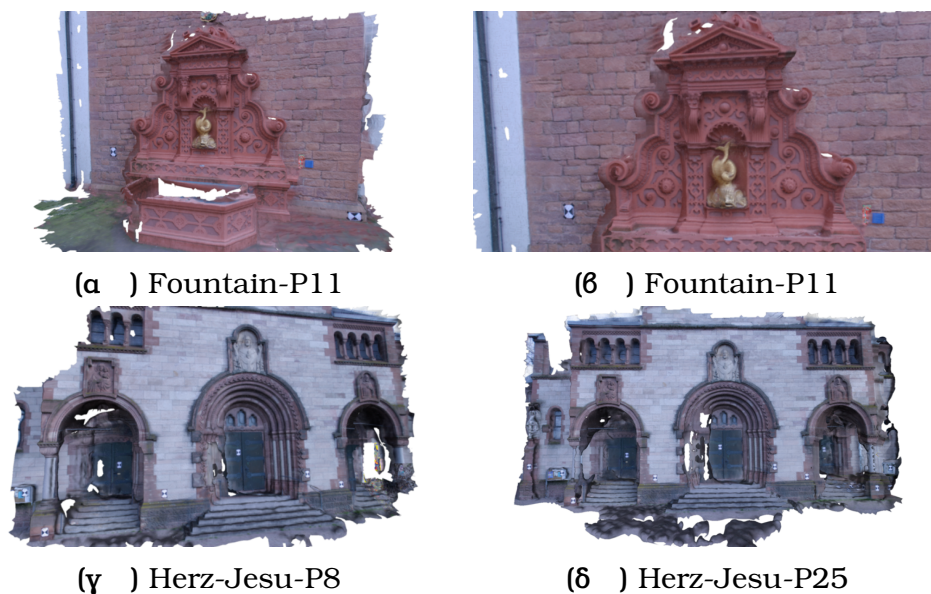


Σχήμα 7.16: 2Δ φωτογραφίες που χρησιμοποιήθηκαν σαν είσοδο στο σύστημα για ανακατασκευή στην πρώτη σειρά της Εικόνας 7.17



Σχήμα 7.17: Αποτελέσματα Mvs-Πυκνό σύννεφο σημείων

Ανακατασκευασμένη επιφάνεια



Σχήμα 7.18: Αποτελέσματα ανακατασκευής επιφάνειας

7.4 Επίλογος

Το πρόβλημα του Θεμελιώδους Πίνακα στο πλαίσιο της L_∞ νόρμας

Στα προηγούμενα Κεφάλαια, παρουσιάσαμε την επίλυση προβλημάτων που προέρχονται από την περιοχή της γεωμετρίας πολλών όψεων με τεχνικές βελτιστοποίησης, χρησιμοποιώντας την L_∞ νόρμα.

Είναι φυσικό, να αναρωτηθούμε αν θα μπορούσαμε να λύσουμε το πρόβλημα του προσδιορισμού του θεμελιώδους πίνακα F , με αντίστοιχες τεχνικές, καθώς και τα οφέλη που θα αποκομίζαμε από αυτό. Για να εξετάσουμε την δυνατότητα μιας τέτοιας προσέγγισης, θα πρέπει να μελετήσουμε την μορφή του σφάλματος *ανά σημείο δεδομένων*, στην προκειμένη περίπτωση ανά αντιστοιχία $(\mathbf{x}_{im1}^i, \mathbf{x}_{im2}^i)$. Στα επόμενα, αναφερόμαστε στο σφάλμα αυτό ως *υπόλοιπο ανά σημείο* R_p^i .

Προκειμένου να είναι υπολογιστικά εύκολη η επίλυση του προβλήματος βελτιστοποίησης που θα θέσουμε, αλλά και κυρίως, να έχει την επιθυμητή ιδιότητα του μοναδικού (ολικού) ελαχίστου, θα πρέπει οι ανισότητες

$$R_p^i < \sigma$$

να είναι περιορισμοί κώνου (Second Order Cone Constraints). Αυτό εξασφαλίζει την εφαρμοσιμότητα υπολογιστικά αποδοτικών αλγορίθμων επίλυσης καθώς και την ικανοποίηση ιδιοτήτων γενικευμένης κυρτότητας για το πρόβλημα.

Στην περίπτωση του Θεμελιώδους πίνακα F , περιπτώσεις υπολοίπων με γε-

ωμετρικό νόημα που κατά κανόνα χρησιμοποιούνται [134] είναι:

$$R_p^i = \frac{|x_{im2}^i{}^T F x_{im1}^i|}{\sqrt{(F x_{im1}^i)_1^2 + (F x_{im1}^i)_2^2}} + \frac{|x_{im2}^i{}^T F x_{im1}^i|}{\sqrt{(F^T x_{im2}^i)_1^2 + (F^T x_{im2}^i)_2^2}}$$

$$R_p^i = \frac{|x_{im2}^i{}^T F x_{im1}^i|}{\sqrt{(F x_{im1}^i)_1^2 + (F x_{im1}^i)_2^2 + (F^T x_{im2}^i)_1^2 + (F^T x_{im2}^i)_2^2}}$$

Μπορεί να δειχθεί, ότι χρησιμοποιώντας τα προηγούμενα υπόλοιπα ανά σημείο παίρνουμε περιορισμούς της μορφής

$$R_p^i < \sigma \iff \|A_i \mathbf{x}\| > \mathbf{y}_1^T \mathbf{x}$$

Οι προηγούμενες ανισότητες δεν δίνουν υποχρεωτικά κυρτό σύνολο λύσεων για το άγνωστο διάνυσμα \mathbf{x} . Συγκεκριμένα, δίνουν το συμπλήρωμα ενός κυρτού συνόλου, δηλαδή ενός κώνου δεύτερης τάξης. Θυμίζουμε ότι συμπληρώματα κυρτών συνόλων είναι κυρτά μόνο στην ειδική περίπτωση των halfspaces. Συνεπώς, το πρόβλημα δεν εντάσσεται ομαλά στο πλαίσιο της L_∞ νόρμας. Συνοπτικά, αυτό οφείλεται στην μη-εφαρμογή της ιδιότητας κυρτής τομής κυρτών συνόλων. Έτσι, εξηγείται και το σχετικό κενό στην βιβλιογραφία.

Κώδικας και σύνολα δεδομένων

| Πακέτο Λογισμικού | Αναφορές | Σχόλια |
|--|----------|---|
| Vincent's SFM Toolbox for Matlab | [182] | Λύση του προβλήματος αραιού Bundle Adjustment, με δυνατότητα εύκολης ρύθμισης των αγνώστων και των παραμέτρων της βελτιστοποίησης |
| AdelaideRMF: Robust Model Fitting Data Set | [244] | Σύνολο δεδομένων που περιέχει ταιριασμένα ζεύγη φωτογραφιών. Οι αντιστοιχίες έχουν κατηγοριοποιηθεί σε ορθές και εσφαλμένες |
| MatlabBGL | [76] | Βιβλιοθήκη αλγορίθμων γράφων |

Πίνακας 7.14: Λογισμικό που χρησιμοποιήσαμε στην διεξαγωγή των πειραμάτων

8.1 Συμβολή της διπλωματικής εργασίας

Στην Διπλωματική εργασία μελετήθηκε το πρόβλημα της τρισδιάστατης ανακατασκευής από σύνολο φωτογραφιών, μη δομημένο και χωρίς επιπλέον πληροφορίες σχετικά με την λήψη των φωτογραφιών και το είδος των φωτογραφικών μηχανών. Συνοπτικά, η συμβολή της Διπλωματικής εργασίας δίνεται στα επόμενα σημεία

- Περιγραφή και οργάνωση του προβλήματος της ανακατασκευής, από την εύρεση χαρακτηριστικών στις εικόνες και το ταίριασμα τους ως την παραγωγή της τελικής τρισδιάστατης επιφάνειας. Η βιβλιογραφική παρουσίαση της προηγούμενης έρευνας στους τομείς αυτούς είναι ενδελεχής, ενώ δίνονται και σχετικές αναφορές για διαθέσιμες υλοποιήσεις αλγορίθμων σχετικών με τα στάδια αυτά
- Γεωμετρική επαλήθευση των αρχικών αντιστοιχιών μεταξύ ζευγών εικόνων. Το πρόβλημα μετασχηματίζεται σε αυτό της εύρεσης της μέγιστης κοινής/αύξουσας υπακολουθίας και λύνεται αποδοτικά. Διερευνώνται και τροποποιήσεις στην βασική μέθοδο, με σκοπό την μείωση των αντιστοιχιών που απορρίπτονται (εσφαλμένα) στο στάδιο της γεωμετρικής επαλήθευσης.
- Εισαγωγή μεθόδου για την βαθμονόμηση και την μετρική ανακατασκευή ζεύγους καμερών με άγνωστα και διαφορετικά εστιακά μήκη. Η μέθοδος θεμελιώνεται θεωρητικά και διερευνάται τόσο θεωρητικά όσο και πειραματικά.
- Εισαγωγή μεθόδου για την βελτίωση των εκτιμήσεων των εστιακών μηκών των καμερών. Η μέθοδος βασίζεται στην κατάλληλη στάθμιση των πολυάριθμων εκτιμήσεων που λαμβάνουμε για κάθε εστιακό μήκος, χρησιμοποιώντας γεωμετρικά επιχειρήματα και την δομή του ολικού προβλήματος ανακατασκευής.
- Εφαρμογή του αλγορίθμου Weiszfeld για την L_1 στάθμιση των περιστροφών, τόσο για τις εκτιμήσεις των σχετικών περιστροφών όσο και για την καταχώρηση των περιστροφών (rotation registration). Εξάλειψη της εξάρτησης από υπολογισμό θεμελιώδους Πίνακα. Μελέτη αρχικοποίησης στάθμισης περιστροφών και άλλων παραμέτρων.

- Διατύπωση του προβλήματος Structure From Motion στο πλαίσιο της L_∞ νόρμας. Χρήση σχετικής γεωμετρίας (περιστροφών) που υπολογίστηκαν με μεθόδους δύο όψεων (προηγούμενη συνεισφορά) για την αρχικοποίηση του προβλήματος Structure From Motion.
- Ανάλυση διαφόρων παραμέτρων που εμφανίζονται στα επιμέρους προβλήματα που επιλύθηκαν. Σύγκριση διαφορετικών αλγορίθμων για την επίλυση υπό-προβλημάτων.

8.2 Μελλοντικές κατευθύνσεις

Το πρόβλημα της ανακατασκευής, είναι ένα αρκετά διευρυμένο πρόβλημα και για αυτό θα μπορούσε να επεκταθεί σε πολλές κατευθύνσεις. Προσωπική μας πεποίθηση είναι ότι οι μέθοδοι που είναι διαθέσιμες σήμερα, έχουν προχωρήσει αρκετά για να δίνουν ικανοποιητικά αποτελέσματα στο πρόβλημα της απλής ανακατασκευής. Συνεπώς, αισθανόμαστε ότι η μελλοντική έρευνα θα ήταν πιο γόνιμη αν αξιοποιούσε τις διαθέσιμες αυτές μεθόδους και προσανατολιζόταν στο να συνδυάσει το πρόβλημα της ανακατασκευής με άλλα προβλήματα, ή να επεκτείνει το πεδίο εφαρμογής των σημερινών μεθόδων. Συγκεκριμένα προτείνουμε

- Επιτάχυνση των μεθόδων ανακατασκευής. Σε αυτή την κατεύθυνση, θα ήταν ενδιαφέρουσα η προσπάθεια προσαρμογής του μοντέλου bag of visual words, ιδιαίτερα επιτυχημένου στον τομέα της αναγνώρισης αντικειμένων, στο πρόβλημα της ανακατασκευής. Άλλες δυνατότητες, συμπεριλαμβάνουν την χρήση περιγραφών εικόνων για τον περιορισμό του συνόλου των ζευγών φωτογραφιών που θα ταιριαστούν με μεγαλύτερη ακρίβεια και την χρήση εξελιγμένων αλγορίθμων για την περαιτέρω επιτάχυνση της ανίχνευσης και του ταιριάσματος χαρακτηριστικών
- Επέκταση του πεδίου εφαρμογής του προβλήματος της ανακατασκευής σε σκηνές με πολλαπλές και διαφόρων ειδών οπτικές συμμετρίες και επαναλαμβανόμενες δομές ή πολλαπλές εμφανίσεις όμοιων δομών. Πρόσφατη έρευνα έχει κινηθεί στην κατεύθυνση αυτή
- Semantic Structure from Motion. Ταυτόχρονη εξαγωγή πληροφορίας σχετικής με τα αντικείμενα ή τα κτήρια που ανακατασκευάζονται και σύμμειξη με το πρόβλημα της αναγνώρισης αντικειμένων.
- Guided Sampling. Καθοδήγηση της τυχαίας δειγματοληψίας αντιστοιχιών εικόνων, σχετικών περιστροφών κ.λ.π, ώστε να επιλέγονται ελάχιστα δείγματα που οδηγούν σε περισσότερο πιθανές και ορθότερες λύσεις με ταυτόχρονη την πιθανή απόρριψη εσφαλμένων αντιστοιχιών και μερικών ανακατασκευών.

Παραρτήματα

I

ΠΡΟΣΘΗΚΗ ΘΟΡΥΒΟΥ ΣΕ ΕΙΚΟΝΕΣ ΓΙΑ ΠΕΙΡΑΜΑΤΑ ΜΕ ΣΥΝΘΕΤΙΚΑ ΔΕΔΟΜΕΝΑ

Για την πραγματοποίηση πειραμάτων με συνθετικά δεδομένα και την προσομοίωση των συνθηκών που θα συναντηθούν σε περίπτωση εφαρμογής των υπό εξέταση μεθόδων σε πραγματικές φωτογραφίες, απαιτείται η προσθήκη θορύβου στις εικόνες.

Καταρχάς, σημειώνουμε ότι ο θόρυβος πρέπει να προστεθεί στις μετρήσεις που αφορούν τις εικόνες, δηλαδή στις θέσεις των σημείων της εικόνας. Η προσθήκη θορύβου σε άλλες οντότητες, όπως για παράδειγμα σημεία της σκηνής, πίνακες προβολής, θεμελιώδης πίνακας κ.α., δεν ανταποκρίνεται στην διαδικασία σχηματισμού των εικόνων. Για αυτό και θα πρέπει να αποφεύγεται.

Προκειμένου οι μέθοδοι να αξιολογούνται σε ίδια βάση, είναι επίσης θεμιτό να γίνει περιγραφή της ποσότητας του προστιθέμενου θορύβου, με όρους που δεν σχετίζονται με τις διαστάσεις των συνθετικών εικόνων ή άλλες σχετικές ποσότητες.

Στην βιβλιογραφία, μια συνήθης πρακτική είναι η προσθήκη στις θέσεις των σημείων της εικόνας, Γκαουσιανού θορύβου με τυπική απόκλιση που εκφράζεται, ως ποσοστό, συναρτήσει κάποιας διάστασης της εικόνας. Θεωρούμε όμως ότι η πρακτική αυτή δεν είναι αντικειμενική, αφού οι εικόνες δεν έχουν τετράγωνο σχήμα και επομένως η σχέση του μήκους και του πλάτους της εικόνας μπορούν να επηρεάσουν τις ποσοστιαίες αυτές περιγραφές. Συνεπώς, θεωρούμε αναξιόπιστη μια τέτοια ποσοστιαία έκφραση.

Μια ορθότερη προσέγγιση, προτείνει την περιγραφή του επιπέδου του θορύβου με τον λόγο των επόμενων δύο ποσοτήτων

- της διασποράς του προστιθέμενου (Γκαουσιανού) θορύβου
- της μέσης τιμής του μέτρου των σημείων (αντιστοιχιών) των εικόνων

Με μαθηματικό συμβολισμό, υπολογίζουμε το επίπεδο του θορύβου από τον λόγο

$$\frac{JT\sigma^2}{\|\mathbf{p}_{1:T}\|_F}$$

Η περιγραφή αυτή έχει το πλεονέκτημα να υπολογίζεται από το μέτρο των σημείων της εικόνας, το οποίο επηρεάζεται εξίσου από το πλάτος και το ύψος

της εικόνας. Ενσωματώνουμε επομένως στην περιγραφή καλύτερα τις συνολικές διαστάσεις της εικόνας.

II

ΥΠΟΛΟΓΙΣΜΟΣ ΣΦΑΛΜΑΤΟΣ ΣΤΗΝ ΘΕΣΗ ΚΑΙ ΤΟΝ ΠΡΟΣΑΝΑΤΟΛΙΣΜΟ ΚΑΜΕΡΑΣ

Έστω κάμερα και περιγραφή αυτής με την μήτρα προβολής

$$K [R \ t] \quad (B \ .1)$$

Θεωρούμε ότι ο πίνακας R είναι ομαλός (proper) πίνακας στροφής και ότι ο πίνακας εσωτερικών παραμέτρων K είναι διαγώνιος πίνακας, με θετικά στοιχεία στην διαγώνιο του.

Έστω ότι θέλουμε τώρα να υπολογίσουμε το σφάλμα της περιγραφής $B \ .1$ σε σχέση με την "ορθή" περιγραφή

$$K_{GT} [R_{GT} \ t_{GT}]$$

Για το σφάλμα στροφής έχουμε:

$$\begin{aligned} R_{er} &\equiv R R_{GT}^T \\ 1 + 2 \cos \theta_{er} &= \text{trace } R_{er} \\ \Rightarrow \cos \theta_{er} &= \frac{\text{trace } R_{er} - 1}{2} \end{aligned}$$

Αντίστοιχα, και με την υπόθεση μοναδιαίων διανυσμάτων t, t_{GT} ώστε να απλοποιήσουμε τον συμβολισμό μας, το σφάλμα στην διεύθυνση μετατόπισης δίνεται:

$$\cos \phi_{er} = t \cdot t_{GT}^T$$

Όμως, γνωρίζουμε ότι μπορούμε να πολλαπλασιάσουμε την αναπαράσταση $K [R \ t]$ κατά μία σταθερά, και να καταλήξουμε σε μια ισοδύναμη παράσταση. Έχουμε λοιπόν

- Οι παραστάσεις $K [R \ t]$, $-K [R \ t]$, $-IK [R \ t]$ είναι ισοδύναμες.
- Ο πολλαπλασιασμός διαγώνιων τετραγωνικών μητρών είναι αντιμεταθετικός $IK = KI$.
- Συνεπώς οι περιγραφές $K [R \ t]$, $K [-R \ -t]$ είναι ισοδύναμες.

Εφόσον η μήτρα R είναι ομαλή μήτρα στροφής, η μήτρα $-R$ θα είναι μη ομαλή (improper) περιστροφή, δηλαδή έχει διακρίνουσα -1 . Προκύπτει επίσης άμεσα, από ιδιότητες οριζουσών, ότι και το γινόμενο

$$R' R_{GT}^T = -R R_{GT}^T$$

θα είναι μη-ομαλή μήτρα περιστροφής. Αντίστοιχα τώρα με τους υπολογισμούς που προηγήθηκαν, έχουμε για το σφάλμα στον προσανατολισμό και την θέση της κάμερας:

$$\begin{aligned} \cos \theta' &= \frac{\text{trace } R' R_{GT}^T + 1}{2} \\ &= \frac{-\text{trace } R_{er} + 1}{2} \\ &= -\cos \theta_{er} \end{aligned} \quad (\text{B } .2)$$

και αντίστοιχα

$$\begin{aligned} \cos \phi' &= \mathbf{t}' \cdot \mathbf{t}_{GT}^T \\ &= -\cos \phi_{er} \end{aligned} \quad (\text{B } .3)$$

Για τον υπολογισμό της γωνίας στροφής χρησιμοποιήσαμε τους επόμενους τύπους

$$\begin{aligned} \text{trace } R &= 1 + 2 \cos \theta_r, \text{ για } R \text{ ομαλή μήτρα στροφής} \\ \text{trace } R &= -1 + 2 \cos \theta_r, \text{ για } R \text{ μη-ομαλή μήτρα στροφής} \end{aligned}$$

Από τις σχέσεις **B .2B .3** έχουμε ότι, ανάλογα με την αναπαράσταση της ευρεθείσας μήτρας προβολής P και της θεωρούμενης ορθής μήτρας προβολής P_{GT} , το σφάλμα στον προσανατολισμό θ και στην θέση ϕ θα δοθούν είτε ως οξείες γωνίες, είτε ως οι παραπληρωματικές τους. Το γεγονός αυτό πρέπει να ληφθεί υπόψιν κατά τον υπολογισμό του σφάλματος, προκειμένου να γίνει σωστά ο υπολογισμός του.

ΚΑΤΑΛΟΓΟΣ ΣΧΗΜΑΤΩΝ

| | |
|--|----|
| 1.1. Το μοντέλο της προοπτικής προβολής | 20 |
| 1.2. Χρησιμοποιώντας τα όμοια τρίγωνα, υπολογίζουμε τις συντεταγμένες της προβολής στο επίπεδο της εικόνας | 20 |
| 1.3. Η διαδικασία σχηματισμού εικόνας χωρισμένη σε βήματα. Αναπαραγωγή απο [234] | 26 |
| 2.1. Συγκεντρωτική μορφή ενός συστήματος Structure from Motion | 38 |
| 2.2. Χωρισμός προβλήματος SfM σε υποχάρτες. Ο διαχωρισμός φαίνεται με τα διαφορετικά χρώματα. Από [160] | 40 |
| 2.3. Παράδειγμα γράφου για πρόβλημα SfM. Οι συμπαγείς γραμμές ενώνουν σημεία με κάμερες που βλέπουν αυτά τα σημεία. Οι διακεκομμένες γραμμές ενώνουν κάμερες που βλέπουν ένα κοινό σημείο στον χώρο. | 43 |
| 2.4. Κατανομή χρόνου υπολογισμού μεθόδου Bundle Adjustment | 48 |
| 2.5. Διάφορες προσεγγίσεις για την περάτωση μιας επανάληψης του Bundle Adjustment. | 50 |
| 3.1. Αβεβαιότητα στην θέση του τρισδιάστατου σημείου, σε σχέση με την γωνία τριγωνισμού | 57 |
| 3.2. Αποτελέσματα ανακατασκευής [78], σε σχέση με το σύστημα επιλογής όψεων και το κριτήριο βελτιστοποίησης. | 62 |
| 3.3. Εικόνες εισόδου και αποτελέσματα ανακατασκευής [78] | 63 |
| 3.4. Σύγκριση ανακατασκευής και ground truth μοντέλου. Από [78] | 63 |
| 3.5. Σύνολο δεδομένων [69]. Αντικείμενα. | 71 |
| 3.6. Αποτελέσματα [69]. Αντικείμενα. | 72 |
| 3.7. Εικόνες συνόλου δεδομένων [69]. Εξωτερική σκηνή. | 72 |
| 3.8. Αποτελέσματα [69]. Εξωτερική σκηνή. | 72 |
| 3.9. Διάγραμμα ροής του αλγορίθμου Clustering. Αναπαραγωγή από [65] | 77 |
| 3.10 Αποτελέσματα clustering όψεων CMVS | 79 |
| 3.11 Συνεισφορά χρήσης επικαλυπτόμενων clusters στην μέθοδο CMVS. | 79 |
| 3.12 Ανακατασκευασμένο μοντέλο με laser scanner. Από [216]. | 82 |
| 3.13 Αριθμός σωστά ανακατασκευασμένων σημείων ανά μέθοδο, συναρτήσει του κατωφλιού σφάλματος | 83 |
| 3.14 Σωρευτικό σφάλμα για διάφορες μεθόδους MVS και σύνολα δεδομένων. | 84 |

| | |
|---|-----|
| 3.15.Τριγωνισμός Delauny ενός συνόλου σημείων στο επίπεδο. Οι περιγεγραμμένοι κύκλοι των τριγώνων που σχηματίζουν τα σημεία έχουν σχεδιαστεί. | 85 |
| 3.16.Μη ομοιόμορφη δειγματοληψία μιας καμπύλης (Αριστερά) και Delauny τριγωνισμός τους (Δεξιά). Γειτονικά σημεία σε όλες τις διευθύνσεις ανιχνεύονται. | 85 |
| 4.1. Εντοπισμός ακροτάτων για τον ανιχνευτή DoG. | 90 |
| 4.2. Επίδοση ανιχνευτών SIFT, σε σχέση με την συχνότητα δειγματοληψίας στον χώρο κλίμακας. | 91 |
| 4.3. Χωρική άθροιση κλίσεων και εξαγωγή περιγραφέα για τα χαρακτηριστικά SIFT. | 93 |
| 4.4. Δομή log-polar grid με διακριτές τις περιοχές συμψηφισμού των τιμών των κλίσεων για την δημιουργία ιστογραμμάτων. . . | 98 |
| 4.5. Χαρακτηριστικά SURF. Προσέγγιση δεύτερης τάξης παραγωγών Γκαουσιανών από κυτιοειδή φίλτρα. | 101 |
| 4.6. Στάδια αλγορίθμου κατασκευής περιγραφέα. | 105 |
| 4.7. Παρουσίαση μεθόδων χωρικού συμψηφισμού χαρακτηριστικών για τον υπολογισμό του περιγραφέα. | 107 |
| 4.8. Δομή περιγραφέα Daisy. | 110 |
| 4.9. Σύγκριση Γεωμετρικής και ομοιόμορφης (Γκαουσιανής) εξομάλυνσης. | 113 |
| 5.1. Σχηματική παρουσίαση Θεωρημάτων 1,2. | 131 |
| 5.2. Περιστροφή για ευθυγράμμιση διανυσμάτων. Σχέση διάταξης διανυσμάτων και γωνίας πίνακα στροφής. | 143 |
| 5.3. Παρουσίαση κατασκευής για την Απόδειξη του Θεωρήματος 14. | 147 |
| 5.4. Σχηματική παρουσίαση όρων που σχηματίζουν τον όρο A της Εξίσωσης 5.59. | 148 |
| 5.5. Ισοσταθμικές της $ \theta_2 - \theta_1 $ για διάφορες τιμές της γωνίας ψ και ενδεικτικό εύρος γωνιών $\theta_1, 2\phi$ | 150 |
| 5.6. Εξέλιξη του σφάλματος στις εκτιμήσεις των εστιακών μηκών, της περιστροφής και της μετατόπισης για διαφορετικά επίπεδα Θορύβου. Ο Θόρυβος περιγράφεται με την τυπική του απόκλιση ως ποσοστό των συντεταγμένων των σημείων της εικόνας. | 158 |
| 5.7. Εξέλιξη της διασποράς του σφάλματος στις εκτιμήσεις των εστιακών μηκών και της περιστροφής για διαφορετικά επίπεδα Θορύβου. Ο Θόρυβος περιγράφεται με την τυπική του απόκλιση ως ποσοστό των συντεταγμένων των σημείων της εικόνας. | 158 |
| 5.8. Εξέλιξη του ποσοστού επιτυχίας της μεθόδου, για διαφορετικά επίπεδα θορύβου. Ο Θόρυβος περιγράφεται με την τυπική του απόκλιση ως ποσοστό των συντεταγμένων των σημείων της εικόνας. | 159 |
| 5.9. Συσσωρευτικά ιστογράμματα για την κατανομή του σφάλματος. Το επίπεδο του Θορύβου είναι 1%. | 159 |
| 5.10.Εξέλιξη του σφάλματος στις εκτιμήσεις των εστιακών μηκών, της περιστροφής και της μετατόπισης για διαφορετικά επίπεδα Θορύβου. Ο Θόρυβος περιγράφεται με την τυπική του απόκλιση ως ποσοστό των συντεταγμένων των σημείων της εικόνας. | 161 |

| | |
|--|-----|
| 5.11 Εξέλιξη της διαφοράς του σφάλματος στο εστιακό μήκος Δf_1 , για διαφορετικά επίπεδα Θόρυβου. Ο Θόρυβος περιγράφεται με την τυπική του απόκλιση ως ποσοστό των συντεταγμένων των σημείων της εικόνας. | 161 |
| 5.12 Σύγκριση των αποτελεσμάτων της βασικής μεθόδου, και της μεθόδου με τις πολλαπλές εκτιμήσεις του f_1 . Ο αριθμός των καμέρων δίνεται στο υπόμνημα κάθε γραφήματος. Ο Θόρυβος περιγράφεται με την τυπική του απόκλιση ως ποσοστό των συντεταγμένων των σημείων της εικόνας. | 162 |
| 5.13 Σφάλμα και ποσοστό επιτυχίας για σχεδόν εκφυλισμένες διατάξεις. | 167 |
| 5.14 Σφάλμα σε πειράματα με πραγματικά δεδομένα. | 172 |
| 5.15 Πειράματα σε πραγματικά δεδομένα. Απόδοση διαφορετικών ανιχνευτών, για διάφορες τιμές του T_{ratio} | 173 |
| 5.16 Πειράματα σε πραγματικά δεδομένα. Μέσο σφάλμα Δf_1 , ρυθμός αποτυχίας και διάμεσος του σφάλματος, για μεταβαλλόμενα T_u, T_l | 175 |
| 6.1. Οργάνωση του προβλήματος σε στάδια. Σχετικές τεχνικές και αλγόριθμοι. | 181 |
| 6.2. Τα στάδια της μεθόδου που ακολουθήσαμε, από την καταχώριση των περιστροφών σε κοινό πλαίσιο έως την εκτίμηση της τελικής επιφάνειας | 185 |
| 6.3. Εικόνες συνόλου δεδομένων | 190 |
| 6.4. Άραιο σύννεφο σημείων-έξοδος SfM. Τα κόκκινα σημεία αντιστοιχούν στις θέσεις των καμερών | 190 |
| 6.5. Πυκνό σύννεφο σημείων, έξοδος MnS | 191 |
| 6.6. Τελική ανακατασκευή της επιφάνειας, όπως προέκυψε από τον αλγόριθμο PSR | 191 |
| 6.7. Εικόνες συνόλου δεδομένων | 192 |
| 6.8. Άραιο σύννεφο σημείων-έξοδος SfM. Τα κόκκινα σημεία αντιστοιχούν στις θέσεις των καμερών | 192 |
| 6.9. Πυκνό σύννεφο σημείων, έξοδος MnS | 193 |
| 6.10. Τελική ανακατασκευή της επιφάνειας, όπως προέκυψε από τον αλγόριθμο PSR | 193 |
| 6.11. Εικόνες συνόλου δεδομένων | 194 |
| 6.12. Άραιο σύννεφο σημείων-έξοδος SfM. Τα κόκκινα σημεία αντιστοιχούν στις θέσεις των καμερών | 194 |
| 6.13. Πυκνό σύννεφο σημείων, έξοδος MnS | 195 |
| 6.14. Τελική ανακατασκευή της επιφάνειας, όπως προέκυψε από τον αλγόριθμο PSR | 195 |
| 6.15. Εικόνες συνόλου δεδομένων | 196 |
| 6.16. Άραιο σύννεφο σημείων-έξοδος SfM. Τα κόκκινα σημεία αντιστοιχούν στις θέσεις των καμερών | 196 |
| 6.17. Πυκνό σύννεφο σημείων, έξοδος MnS | 197 |
| 6.18. Τελική ανακατασκευή της επιφάνειας, όπως προέκυψε από τον αλγόριθμο PSR | 197 |
| 6.19. Εικόνες συνόλου δεδομένων | 198 |
| 6.20. Άραιο σύννεφο σημείων-έξοδος SfM. Τα κόκκινα σημεία αντιστοιχούν στις θέσεις των καμερών | 198 |
| 6.21. Πυκνό σύννεφο σημείων, έξοδος MnS | 198 |

| | |
|---|-----|
| 6.22.Τελική ανακατασκευή της επιφάνειας, όπως προέκυψε από τον αλγόριθμο PSR | 199 |
| 6.23.Εικόνες συνόλου δεδομένων | 200 |
| 6.24.Άραιο σύννεφο σημείων-έξοδος SfM. Τα κόκκινα σημεία αντιστοιχούν στις θέσεις των καμερών | 200 |
| 6.25.Πυκνό σύννεφο σημείων, έξοδος MvS | 201 |
| 6.26.Τελική ανακατασκευή της επιφάνειας, όπως προέκυψε από τον αλγόριθμο PSR | 201 |
| 6.27.Αποτελέσματα ανακατασκευής (ως πύκνο σύννεφο σημείων και ως textured επιφάνεια) για τις Καρυάτιδες. | 202 |
| 6.28.Αποτελέσματα ανακατασκευής (ως πύκνο σύννεφο σημείων και ως textured επιφάνεια) για το Ερέχθειο. | 203 |
| 6.29.Αποτελέσματα ανακατασκευής (ως πύκνο σύννεφο σημείων και ως textured επιφάνεια) για τον Παρθενώνα. | 204 |
| 6.30.Μερική ανακατασκευή μνημείου του Λυσικράτη | 205 |
| 6.31.Μερική ανακατασκευή μνημείου Φιλοπάππου | 206 |
| 7.1. Αντιστοιχίες Εικόνων | 210 |
| 7.2. Επίδραση περιστροφών συστήματος συντεταγμένων κάμερας στην απεικόνιση παράλληλων ευθειών | 215 |
| 7.3. Κατανομή εκτιμήσεων f από ένα ζεύγος καμερών | 227 |
| 7.4. Κατανομή εκτιμήσεων f από όλα τα ταιριασμένα ζεύγη καμερών | 228 |
| 7.5. Κατανομή συζευγμένου αθροίσματος εμπιστοσύνης | 228 |
| 7.6. Σφάλμα μεταβάλλοντας τον αριθμό των επαναλήψεων RANSAC. Αλγόριθμοι 7 και 8 σημείων. | 229 |
| 7.7. Σφάλμα μεταβάλλοντας το ποσοστό των σημείων μπροστά από τις κάμερες. | 229 |
| 7.8. Σύγκριση επίδοσης στην εκτίμηση του εστιακού μήκους και του σχετικού προσανατολισμού | 230 |
| 7.9. Σύγκριση επίδοσης στην καταχώρηση των περιστροφών σε κοινό πλαίσιο αναφοράς | 231 |
| 7.10.Υπόλοιπο στις σχέσεις συμβατότητας καταχωρημένων περιστροφών και σχετικών περιστροφών | 232 |
| 7.11.Σύγκλιση αλγορίθμου Weiszfeld | 232 |
| 7.12.Σύγκριση μεθόδων εκτίμησης σχετικών περιστροφών R_{ij} με κριτήριο το σφάλμα στις τελικές εκτιμήσεις R_i | 234 |
| 7.13.Επαναλήψεις Bundle Adjustment | 235 |
| 7.14.2Δ Εικόνες Εισόδου | 239 |
| 7.15.Αποτελέσματα MvS | 240 |
| 7.16.2Δ Εικόνες Εισόδου | 241 |
| 7.17.Αποτελέσματα MvS | 241 |
| 7.18.Αποτελέσματα ανακατασκευής επιφάνειας | 242 |

ΚΑΤΑΛΟΓΟΣ ΠΙΝΑΚΩΝ

| | | |
|------|--|-----|
| 1. | Βασική Βιβλιογραφία πρώτου μέρους | 13 |
| 2. | Βασικές συνεισφορές Κεφαλαίων δεύτερου μέρους | 13 |
| 1.1. | Τρισδιάστατοι Μετασχηματισμοί | 18 |
| 2.1. | Απαιτούμενος αριθμός επαναλήψεων RANSAC για 0.01 πιθανότητα αποτυχίας. Από [212]. | 37 |
| 2.2. | Αριθμητικές μέθοδοι επίλυσης προβλήματος Bundle Adjustment | 48 |
| 2.3. | Επιλογή κατάλληλων αριθμητικών μεθόδων για επίλυση Bundle Adjustment | 48 |
| 3.1. | Παράμετροι PMVS και τιμές τους. Οι τιμές είναι από [69]. | 71 |
| 3.2. | Χρόνος εκτέλεσης σε λεπτά επιμέρους σταδίων PMVS. | 72 |
| 3.3. | Παράμετροι CMVS και τιμές τους. Οι τιμές είναι από [65]. | 80 |
| 3.4. | Επιδόσεις συστήματος CMVS | 80 |
| 3.5. | Σύγκριση μεθόδων MVS. | 83 |
| 4.1. | Χρόνοι υπολογισμού για τα χαρακτηριστικά SURF και SIFT. | 103 |
| 4.2. | Παράμετροι προς μάθηση για την διαμόρφωση περιγραφέα. | 108 |
| 4.3. | Παράμετροι περιγραφέα Daisy. Από [225]. | 109 |
| 4.4. | Πλήθος πράξεων (ανά κατηγορία) για τον υπολογισμό των περιγραφέων DAISY και SIFT . Από [225]. | 112 |
| 4.5. | Χρόνοι (σε second) για τον υπολογισμό των περιγραφέων DAISY και SIFT. Από [225]. | 112 |
| 5.1. | Αποτελέσματα για επίπεδο Θορύβου 1%. | 157 |
| 5.2. | Αποτελέσματα σύγκρισης της προτεινομένης μεθόδου και της μεθόδου που βασίζεται στις Εξισώσεις Kruppa. | 163 |
| 5.3. | Περιγραφή συνόλων δεδομένων | 166 |
| 5.4. | Παράμετροι προβλήματος | 168 |
| 5.5. | Πειράματα σε πραγματικά δεδομένα. Σφάλμα στην εκτίμηση του εστιακού μήκους f_1 | 169 |
| 5.6. | Πειράματα σε πραγματικά δεδομένα. Σφάλμα στην εκτίμηση της διεύθυνσης μετατόπισης και της περιστροφής. | 170 |
| 5.7. | Αποτελέσματα για εκτίμηση των μητρών F στο σύνολο δεδομένων Fountain-P11. Χρησιμοποιήσαμε κατώφλι $T_{inliers} = 30$ και $T_{ratio} = 0.4$ | 174 |
| 5.8. | Σφάλμα στην εκτίμηση του εστιακού μήκους Δf . Δίνουμε επίσης και το πλήθος των φωτογραφιών στις οποίες αποτυγχάνει κάθε μέθοδος | 176 |

| | |
|---|-----|
| 5.9. Σφάλμα στην εκτίμηση της περιστροφής ΔR . Οι τιμές είναι σε μοίρες. Η παράμετρος T_f για την διαλογή των αξιόπιστων μητρώων F τέθηκε στην τιμή 0.1. | 176 |
| 5.10 Σφάλμα στην εκτίμηση της διεύθυνσης μετατόπισης Δt . Οι τιμές είναι σε μοίρες. Η παράμετρος T_f για την διαλογή των αξιόπιστων μητρώων F τέθηκε στην τιμή 0.1. | 177 |
| 5.11 Λογισμικό που χρησιμοποιήσαμε στην διεξαγωγή των πειραμάτων | 178 |
| 6.1. Σφάλμα (σε μοίρες) στις εκτιμήσεις της θέσης και του προσανατολισμού των καμερών. | 189 |
| 6.2. Σφάλμα (σε μοίρες) στις εκτιμήσεις της θέσης και του προσανατολισμού των καμερών και σφάλμα Δf | 189 |
| 7.1. Συντελεστής ϕ στο πρόβλημα της γεωμετρικής επαλήθευσης αρχικών αντιστοιχιών | 219 |
| 7.2. Σταθμισμένος συντελεστής ϕ στο πρόβλημα της γεωμετρικής επαλήθευσης αρχικών αντιστοιχιών | 220 |
| 7.3. Precision στο πρόβλημα της γεωμετρικής επαλήθευσης αρχικών αντιστοιχιών | 220 |
| 7.4. Recall στο πρόβλημα της γεωμετρικής επαλήθευσης αρχικών αντιστοιχιών | 220 |
| 7.5. Recall στο πρόβλημα της γεωμετρικής επαλήθευσης αρχικών αντιστοιχιών | 221 |
| 7.6. Precision στο πρόβλημα της γεωμετρικής επαλήθευσης αρχικών αντιστοιχιών | 221 |
| 7.7. Σύγκριση μεθόδων στάθμισης εστιακού μήκους | 227 |
| 7.8. Σύγκριση μεθόδων εκτίμησης σχετικών περιστροφών και καταχώρησης περιστροφών | 233 |
| 7.9. Σύγκριση μεθόδων εκτίμησης σχετικών περιστροφών και καταχώρησης περιστροφών | 234 |
| 7.10 Σύγκριση αλγορίθμου Weiszfeld και αλγορίθμου 5 σημείων | 235 |
| 7.11 Σύγκριση μεθόδων αρχικοποίησης αλγορίθμου Weiszfeld | 237 |
| 7.12 Σύγκριση μεθόδων αρχικοποίησης αλγορίθμου Weiszfeld. Τυπική απόκλιση. | 238 |
| 7.13 Σφάλμα στην εκτίμηση εστιακού μήκους, θέσης και προσανατολισμού, στην τελική ανακατασκευή | 239 |
| 7.14 Λογισμικό | 243 |

ΒΙΒΛΙΟΓΡΑΦΙΑ

- [1] A comparison and evaluation of multi-view stereo reconstruction algorithms. <http://vision.middlebury.edu/mview/>. [Online; accessed 11-March-2013]. 71, 81, 86
- [2] Gallery(pmvs-version 2). <http://www.di.ens.fr/pmvs/gallery.html>. [Online; accessed 11-March-2013]. 71
- [3] Multi-view stereo reconstruction algorithms evaluation on outdoor datasets. <http://cvlabwww.epfl.ch/~strecha/multiview/denseMVS.html>. [Online; accessed 11-March-2013]. 71, 81, 86
- [4] Sameer Agarwal, Noah Snavely, Steven Seitz, and Richard Szeliski. Bundle adjustment in the large. *Computer Vision--ECCV 2010*, pages 29--42, 2010. 39, 47, 48
- [5] Sameer Agarwal, Noah Snavely, and Steven M Seitz. Fast algorithms for 1 problems in multiview geometry. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pages 1--8. IEEE, 2008. 181
- [6] Sameer Agarwal, Noah Snavely, Ian Simon, Steven M Seitz, and Richard Szeliski. Building rome in a day. In *Computer Vision, 2009 IEEE 12th International Conference on*, pages 72--79. IEEE, 2009. 73
- [7] Pablo F Alcantarilla, Luis M Bergasa, and Andrew J Davison. Gauge-surf descriptors. *Image and Vision Computing*, 2012. 113
- [8] David Aldous and Persi Diaconis. Longest increasing subsequences: from patience sorting to the baik-deift-johansson theorem. *Bulletin of the American Mathematical Society*, 36(4):413--432, 1999. 213
- [9] Matthew Antone and Seth Teller. Scalable extrinsic calibration of omni-directional image networks. *International Journal of Computer Vision*, 49:143--174, 2002. 37
- [10] Xavier Armangué and Joaquim Salvi. F-matrix toolbox. <http://eia.udg.es/~qsalvi/recerca.html>. [Online; accessed 08-July-2013]. 178
- [11] Xavier Armangué and Joaquim Salvi. Overall view regarding fundamental matrix estimation. *Image and vision computing*, 21(2):205--220, 2003. 171, 178

- [12] Erik Ask, Olof Enqvist, and Fredrik Kahl. Optimal geometric fitting under the truncated l2-norm. **209**
- [13] Various Authors. Affine covariant features. <http://www.robots.ox.ac.uk/~vgg/research/affine/index.html>. [Online; accessed 28-March-2013]. **97, 103, 115**
- [14] PE Bahsavias. *Multi-Photo Geometrically Constraint Matching*. PhD thesis, PH D Dissertation. ETH Zur ich, 1991.[2j Norihiko Nakano, et al. Digital Still Camera System for Megapixel CCD. *IEEE Transaction on Consumer Electronics*, 1998, 44 (3): 581~586, 1991. **58**
- [15] Christian Bailer, Manuel Finckh, and Hendrik PA Lensch. Scale robust multi view stereo. In *Computer Vision-ECCV 2012*, pages 398--411. Springer, 2012. **84**
- [16] Arindam Banerjee, Chase Krumpelman, Joydeep Ghosh, Sugato Basu, and Raymond J Mooney. Model-based overlapping clustering. In *Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining*, pages 532--537. ACM, 2005. **73**
- [17] Adam Baumberg. Reliable feature matching across widely separated views. In *Computer Vision and Pattern Recognition, 2000. Proceedings. IEEE Conference on*, volume 1, pages 774--781. IEEE, 2000. **95**
- [18] Herbert Bay, Tinne Tuytelaars, and Luc Van Gool. Surf: Speeded up robust features. In *Computer Vision-ECCV 2006*, pages 404--417. Springer, 2006. **13, 87, 99, 101, 103**
- [19] Jean-Charles Bazin, Yongduek Seo, and Marc Pollefeys. Globally optimal consensus set maximization through rotation search. In *Computer Vision-ACCV 2012*, pages 539--551. Springer, 2013. **209**
- [20] Serge Belongie, Jitendra Malik, and Jan Puzicha. Shape matching and object recognition using shape contexts. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 24(4):509--522, 2002. **97**
- [21] Alexander C Berg and Jitendra Malik. Geometric blur for template matching. In *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, volume 1, pages 1--607. IEEE, 2001. **109, 112, 113**
- [22] Rahul Bhotika, David Fleet, and Kiriakos Kutulakos. A probabilistic theory of occupancy and emptiness. *Computer Vision-ECCV 2002*, pages 112--130, 2002. **52, 55**
- [23] Stanley Bileschi and Lior Wolf. Image representations beyond histograms of gradients: The role of gestalt descriptors. In *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*, pages 1--8. IEEE, 2007. **114**

- [24] Ake Bjorck. *Numerical methods for least squares problems*. Number 51. Society for Industrial Mathematics, 1996. 51
- [25] Thomas Bonfort and Peter Sturm. Voxel carving for specular surfaces. In *Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on*, pages 591--596. IEEE, 2003. 52, 53, 54
- [26] Andrian Broadhurst, Tom W Drummond, and Roberto Cipolla. A probabilistic framework for space carving. In *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on*, volume 1, pages 388--393. IEEE, 2001. 52
- [27] GN Newsam DQ Huynh MJ Brooks and HP Pan. Recovering unknown focal lengths in self-calibration: An essentially linear algorithm and degenerate configurations. In *Proc. ISPRS-Congress*, volume 31, pages 575--580. Citeseer, 1996. 162, 164
- [28] Michael J Brooks, Lourdes de Agapito, DQ Huynh, and Luis Baumela. Direct methods for self-calibration of a moving stereo head. In *Computer Vision—ECCV'96*, pages 413--426. Springer, 1996. 164
- [29] D Brown. A solution to the general problem of multiple station analytical stereo-triangulation, rca data reduction. Technical report, Tech. Rept, 43, 1958. 44
- [30] Matthew Brown and David G Lowe. Unsupervised 3d object recognition and reconstruction in unordered datasets. In *3-D Digital Imaging and Modeling, 2005. 3DIM 2005. Fifth International Conference on*, pages 56--63. IEEE, 2005. 58
- [31] Gustavo Carneiro and Allan Jepson. Phase-based local features. *Computer Vision—ECCV 2002*, pages 282--296, 2002. 89
- [32] Manmohan Chandraker, Sameer Agarwal, Fredrik Kahl, David Nistér, and David Kriegman. Autocalibration via rank-constrained estimation of the absolute quadric. In *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*, pages 1--8. IEEE, 2007. 155, 156
- [33] Manmohan Chandraker, D Kriegman, S Belongie, et al. Globally optimal affine and metric upgrades in stratified autocalibration. In *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on*, pages 1--8. IEEE, 2007. 156
- [34] Yanqing Chen, Timothy A Davis, William W Hager, and Sivasankaran Rajamanickam. Algorithm 887: Cholmod, supernodal sparse cholesky factorization and update/downdate. *ACM Transactions on Mathematical Software (TOMS)*, 35(3):22, 2008. 39, 45
- [35] Ondrej Chum and Jiri Matas. Matching with prosac-progressive sample consensus. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 1, pages 220--226. IEEE, 2005. 37

- [36] Ondrej Chum and Jiri Matas. Geometric hashing with local affine frames. In *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, volume 1, pages 879--884. IEEE, 2006. 211
- [37] Ondrej Chum and Jiri Matas. Homography estimation from correspondences of local elliptical features. In *Pattern Recognition (ICPR), 2012 21st International Conference on*, pages 3236--3239. IEEE, 2012. 211
- [38] Ondrej Chum, Jiri Matas, and Stepan Obdrzalek. Enhancing ransac by generalized model optimization. In *Proc. of the ACCV*, volume 2, pages 812--817, 2004. 211, 236
- [39] Various Contributors. dlib c++ library. <http://dlib.net/>. [Online; accessed 29-March-2013]. 114
- [40] Various Contributors. Respatializing photographs — a field tool. <http://openendedgroup.com/field/ReconstructionDistribution>. [Online; accessed 29-March-2013]. 86
- [41] Various Contributors. Surf: speeded up robust features. <http://www.vision.ee.ethz.ch/~surf/index.html>. [Online; accessed 29-March-2013]. 114
- [42] Harald Cramér. *Mathematical methods of statistics*, volume 9. Princeton university press, 1999. 218
- [43] Li Zhang Noah Snavely Brian Curless and Steven M Seitz. Spacetime faces: High resolution capture for modeling and animation. 57
- [44] Arnak Dalalyan and Renaud Keriven. l_1 -penalized robust estimation for a class of inverse problems arising in multiview geometry. In *Advances in Neural Information Processing Systems*, pages 441--449, 2009. 180, 184, 209
- [45] Jeremy S De Bonet and Paul Viola. Poxels: Probabilistic voxelized volume reconstruction. In *Proceedings of International Conference on Computer Vision (ICCV)*, pages 418--425, 1999. 52
- [46] JE Dennis. Nonlinear least-squares, state of the art in numerical analysis, 269-312, 1977. 51
- [47] M-A Drouin, Martin Trudeau, and Sebastien Roy. Geo-consistency for wide multi-camera stereo. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 1, pages 351--358. IEEE, 2005. 53, 54
- [48] Ye Duan, Liu Yang, Hong Qin, and Dimitris Samaras. Shape reconstruction from 3d and 2d data using pde-based deformable surfaces. *Computer Vision-ECCV 2004*, pages 238--251, 2004. 52
- [49] Peter Eisert, Eckehard Steinbach, and Bernd Girod. Multi-hypothesis, volumetric reconstruction of 3-d objects from multiple calibrated camera views. In *Acoustics, Speech, and Signal Processing, 1999. Proceedings., 1999 IEEE International Conference on*, volume 6, pages 3509--3512. IEEE, 1999. 52, 55

- [50] Ludwig Elsner. A note on optimal block-scaling of matrices. *Numerische Mathematik*, 44(1):127--128, 1984. 47
- [51] Chris Engels, Henrik Stewénus, and David Nistér. Bundle adjustment rules. *Photogrammetric Computer Vision*, 2, 2006. 13, 27, 41, 42, 46, 48
- [52] Chris Evans. Open surf. <http://www.chrisevansdev.com/computer-vision-opensurf.html>. [Online; accessed 29-March-2013]. 114
- [53] Olivier Faugeras, Renaud Keriven, et al. Variational principles, surface evolution, pde's, level set methods and the stereo problem. 1996. 52, 53, 54
- [54] Olivier D. Faugeras, E Le Bras-Mehlman, and Jean-Daniel Boissonnat. Representing stereo data with the delaunay triangulation. *Artificial Intelligence*, 44(1):41--87, 1990. 55
- [55] Tom Fawcett. An introduction to roc analysis. *Pattern recognition letters*, 27(8):861--874, 2006. 218
- [56] Vittorio Ferrari, Tinne Tuytelaars, and Luc Van Gool. Simultaneous object recognition and segmentation by image exploration. *Computer Vision-ECCV 2004*, pages 40--54, 2004. 64
- [57] Martin A Fischler and Robert C Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381--395, 1981. 13, 27, 36
- [58] Jan-Michael Frahm, Pierre Fite-Georgel, David Gallup, Tim Johnson, Rahul Raguram, Changchang Wu, Yi-Hung Jen, Enrique Dunn, Brian Clipp, Svetlana Lazebnik, et al. Building rome on a cloudless day. *Computer Vision-ECCV 2010*, pages 368--381, 2010. 73, 207, 208
- [59] Michael L Fredman. On computing the length of longest increasing subsequences. *Discrete Mathematics*, 11(1):29--35, 1975. 213
- [60] William T. Freeman and Edward H. Adelson. The design and use of steerable filters. *IEEE Transactions on Pattern analysis and machine intelligence*, 13(9):891--906, 1991. 95, 96, 105
- [61] Thomas Fromherz and Martin Bichsel. Shape from multiple cues: Integrating local brightness information. In *Proceedings of the Fourth International Conference for Young Computer Scientists, ICYCS*, volume 95, pages 855--862. Citeseer, 1995. 52, 54, 55
- [62] Pascal Fua and Yvan G Leclerc. Object-centered surface reconstruction: Combining multi-image stereo and shading. *International Journal of Computer Vision*, 16(1):35--56, 1995. 53
- [63] Yasatuka Furukawa. Patch-based bundle adjustment software (pba). <http://homes.cs.washington.edu/~furukawa/research/pba/index.html>. [Online; accessed 22-February-2013]. 86

- [64] Yasutaka Furukawa. Clustering views for multi-view stereo (cmvs) software. <http://www.di.ens.fr/cmvs/>. [Online; accessed 11-March-2013]. 13, 52, 73, 86
- [65] Yasutaka Furukawa, Brian Curless, Steven M Seitz, and Richard Szeliski. Towards internet-scale multi-view stereo. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pages 1434--1441. IEEE, 2010. 73, 74, 77, 78, 79, 80, 251, 255
- [66] Yasutaka Furukawa and Jean Ponce. Patch-based multi-view stereo software (pmvs - version 2). <http://www.di.ens.fr/pmvs/>. [Online; accessed 8-March-2013]. 13, 52, 64, 86, 186
- [67] Yasutaka Furukawa and Jean Ponce. Carved visual hulls for image-based modeling. *Computer Vision--ECCV 2006*, pages 564--577, 2006. 70
- [68] Yasutaka Furukawa and Jean Ponce. Accurate camera calibration from multi-view stereo and bundle adjustment. *International Journal of Computer Vision*, 84(3):257--268, 2009. 85
- [69] Yasutaka Furukawa and Jean Ponce. Accurate, dense, and robust multiview stereopsis. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 32(8):1362--1376, 2010. 56, 57, 64, 65, 66, 67, 70, 71, 72, 73, 83, 186, 251, 255
- [70] Andrew C Gallagher. Using vanishing points to correct camera rotation in images. In *Computer and Robot Vision, 2005. Proceedings. The 2nd Canadian Conference on*, pages 460--467. IEEE, 2005. 215
- [71] David Gallup, J-M Frahm, Philippos Mordohai, Qingxiong Yang, and Marc Pollefeys. Real-time plane-sweeping stereo with multiple sweeping directions. In *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*, pages 1--8. IEEE, 2007. 73
- [72] Rahul Garg, Hao Du, Steven M Seitz, and Noah Snavely. The dimensionality of scene appearance. In *Computer Vision, 2009 IEEE 12th International Conference on*, pages 1917--1924. IEEE, 2009. 207
- [73] Pau Gargallo and Peter Sturm. Bayesian 3d modeling from images using multiple depth maps. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 2, pages 885--891. IEEE, 2005. 53, 54, 55
- [74] Riccardo Gherardi and Andrea Fusiello. Practical autocalibration. In *Computer Vision--ECCV 2010*, pages 790--801. Springer, 2010. 155, 156
- [75] Frédéric Cazals—Joachim Giesen. Delaunay triangulation based surface reconstruction: Ideas and algorithms. 2004. 85
- [76] David Gleich. Matlabagl. a matlab graph library. *Institute for Computational and Mathematical Engineering, Stanford University*. Available: http://www.stanford.edu/~dgleich/programs/matlab_bgl, 2008. 243

- [77] Michael Goesele, Brian Curless, and Steven M Seitz. Multi-view stereo revisited. In *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, volume 2, pages 2402--2409. IEEE, 2006. 55, 57
- [78] Michael Goesele, Noah Snavely, Brian Curless, Hugues Hoppe, and Steven M Seitz. Multi-view stereo for community photo collections. In *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on*, pages 1--8. IEEE, 2007. 13, 52, 56, 57, 62, 63, 73, 79, 85, 251
- [79] Venu Madhav Govindu. Combining two-view constraints for motion estimation. In *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, volume 2, pages II--218. IEEE, 2001. 39, 183
- [80] Venu Madhav Govindu. Lie-algebraic averaging for globally consistent motion estimation. In *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, volume 1, pages I--684. IEEE, 2004. 183
- [81] Andreas Griewank, David Juedes, and Jean Utke. Algorithm 755: Adol-c: a package for the automatic differentiation of algorithms written in c/c++. *ACM Transactions on Mathematical Software (TOMS)*, 22(2):131--167, 1996. 51
- [82] A Gruen and Emmanuel P Baltsavias. Geometrically constrained multiphoto matching. *Photogrammetric Engineering and Remote Sensing*, 54(5):633--641, 1988. 58, 61
- [83] Martin Habbecke and Leif Kobbelt. Iterative multi-view plane fitting. In *Proc. Vision, Modeling, and Visualization Conference*, pages 73--80, 2006. 56, 64
- [84] Martin Habbecke and Leif Kobbelt. A surface-growing approach to multi-view stereo reconstruction. In *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*, pages 1--8. IEEE, 2007. 57
- [85] Chris Harris and Mike Stephens. A combined corner and edge detector. In *Alvey vision conference*, volume 15, page 50. Manchester, UK, 1988. 64, 92, 96
- [86] Richard Hartley, Khurram Aftab, and Jochen Trunpf. L1 rotation averaging using the weiszfeld algorithm. In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pages 3041--3048. IEEE, 2011. 209, 223, 236
- [87] Richard Hartley and Fredrik Kahl. Optimal algorithms in multiview geometry. *Computer Vision--ACCV 2007*, pages 13--34, 2007. 39, 155, 180
- [88] Richard Hartley and Frederik Schaffalitzky. l_∞ minimization in geometric reconstruction problems. In *Computer Vision and*

- Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, volume 1, pages 1--504. IEEE, 2004. 184
- [89] Richard Hartley, Jochen Trunpf, Yuchao Dai, and Hongdong Li. Rotation averaging. *International Journal of Computer Vision*, pages 1--39, 2013. 183, 208, 223, 236
- [90] Richard Hartley and Andrew Zisserman. *Multiple view geometry in computer vision*, volume 2. Cambridge Univ Press, 2000. 13, 15, 57, 117, 132, 150, 170, 171, 178
- [91] Michal Havlena, Akihiko Torii, and Tomáš Pajdla. Efficient structure from motion by graph optimization. In *Computer Vision--ECCV 2010*, pages 100--113. Springer Berlin Heidelberg, 2010. 208
- [92] Carlos Hernández Esteban and Francis Schmitt. Silhouette and stereo fusion for 3d object modeling. *Computer Vision and Image Understanding*, 96(3):367--392, 2004. 53, 54, 55, 57, 73
- [93] KL Hiebert. An evaluation of mathematical software that solves nonlinear least squares problems. *ACM Transactions on Mathematical Software (TOMS)*, 7(1):1--16, 1981. 51
- [94] Alexander Hornung and Leif Kobbelt. Robust and efficient photo-consistency estimation for volumetric 3d reconstruction. *Computer Vision--ECCV 2006*, pages 179--190, 2006. 57
- [95] Xiaoyan Hu and Philippos Mordohai. Least commitment, viewpoint-based, multi-view stereo. In *3D Imaging, Modeling, Processing, Visualization and Transmission (3DIMPVT), 2012 Second International Conference on*, pages 531--538. IEEE, 2012. 84
- [96] James W Hunt and Thomas G Szymanski. A fast algorithm for computing longest common subsequences. *Communications of the ACM*, 20(5):350--353, 1977. 213
- [97] J Isidro and Stan Sclaroff. Stochastic refinement of the visual hull to satisfy photometric and consistency constraints. In *Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on*, pages 1335--1342. IEEE, 2003. 53, 54
- [98] Michal Jancosek and Tomas Pajdla. Cmp-mvs software. <http://ptak.felk.cvut.cz/sfmservice/websfm.pl?menu=cmpmvs>. [Online; accessed 22-February-2013]. 86
- [99] Michal Jancosek, Alexander Shekhovtsov, and Tomas Pajdla. Scalable multi-view stereo. In *Computer Vision Workshops (ICCV Workshops), 2009 IEEE 12th International Conference on*, pages 1526--1533. IEEE, 2009. 73
- [100] Yekeun Jeong, David Nister, Drew Steedly, Richard Szeliski, and In-So Kweon. Pushing the envelope of modern methods for bundle adjustment. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pages 1474--1481. IEEE, 2010. 39

- [101] Nianjuan Jiang, Ping Tan, and Loong-Fah Cheong. Seeing double without confusion: Structure-from-motion in highly ambiguous scenes. In *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, pages 1458--1465, 2012. 209
- [102] Hailin Jin, Stefano Soatto, and Anthony J Yezzi. Multi-view stereo beyond lambert. In *Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on*, volume 1, pages I--171. IEEE, 2003. 52, 53, 57
- [103] Hailin Jin, Stefano Soatto, and Anthony J Yezzi. Multi-view stereo reconstruction of dense shape and complex appearance. *International Journal of Computer Vision*, 63(3):175--189, 2005. 52
- [104] T Kadir, M Brady, and A Zisserman. An affine invariant method for selecting salient regions in images. In *Proceedings of the European Conference on Computer Vision*, pages 345--457, 2004. 96
- [105] Fredrik Kahl. Critical motions and ambiguous euclidean reconstructions in auto-calibration. In *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*, volume 1, pages 469--475. IEEE, 1999. 162, 163
- [106] Fredrik Kahl and Richard Hartley. Multiple-view geometry under the $\{L_{\infty}\}$ -norm. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 30(9):1603--1617, 2008. 180, 184
- [107] George Kamberov, Gerda Kamberova, O Chum, Š Obdržálek, D Martinec, J Kostkova, T Pajdla, J Matas, and R Šára. 3d geometry from uncalibrated images. *Advances in Visual Computing*, pages 802--813, 2006. 57, 58
- [108] Sing Bing Kang, Richard Szeliski, and Jinxiang Chai. Handling occlusions in dense multi-view stereo. In *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, volume 1, pages I--103. IEEE, 2001. 53, 54
- [109] Michael Kazhdan. Screened poisson surface reconstruction (version 5). <http://www.cs.jhu.edu/~misha/Code/PoissonRecon/Version5/>. [Online; accessed 06-September-2013]. 186
- [110] Michael Kazhdan, Matthew Bolitho, and Hugues Hoppe. Poisson surface reconstruction. In *Proceedings of the fourth Eurographics symposium on Geometry processing*, 2006. 62, 70, 186
- [111] Yan Ke and Rahul Sukthankar. Pca-sift: A more distinctive representation for local image descriptors. In *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, volume 2, pages II--506. IEEE, 2004. 96, 98
- [112] Junhwan Kim, Vladimir Kolmogorov, and Ramin Zabih. Visual correspondence using energy minimization and mutual information. In *Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on*, pages 1033--1040. IEEE, 2003. 57

- [113] Jan J Koenderink and Andrea J van Doorn. Representation of local geometry in the visual system. *Biological cybernetics*, 55(6):367--375, 1987. 95, 96
- [114] Iasonas Kokkinos, Michael Bronstein, Alan Yuille, et al. Dense scale invariant descriptors for images and surfaces. 2012. 113
- [115] Iasonas Kokkinos and Alan Yuille. Scale invariance without scale selection. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pages 1--8. IEEE, 2008. 113
- [116] Vladimir Kolmogorov and Ramin Zabih. Multi-camera scene reconstruction via graph cuts. *Computer Vision—ECCV 2002*, pages 8--40, 2002. 53, 55
- [117] Peter Kovesi. Matlab and octave functions for computer vision and image processing. <http://www.csse.uwa.edu.au/~pk/Research/MatlabFns/index.html>. 178
- [118] Peter Kovesi. Fast almost-gaussian filtering. In *Digital Image Computing: Techniques and Applications (DICTA), 2010 International Conference on*, pages 121--125. IEEE, 2010. 208
- [119] Zuzana Kukelova, Martin Bujnak, and Tomas Pajdla. Polynomial eigenvalue solutions to the 5-pt and 6-pt relative pose problems. In *BMVC*, pages 1--10, 2008. 155, 156
- [120] Kiriakos Kutulakos. Approximate n-view stereo. *Computer Vision—ECCV 2000*, pages 67--83, 2000. 52, 53, 54, 55
- [121] Kiriakos N Kutulakos and Steven M Seitz. A theory of shape by space carving. *International Journal of Computer Vision*, 38(3):199--218, 2000. 52, 53, 54, 55
- [122] Aldo Laurentini. The visual hull concept for silhouette-based image understanding. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 16(2):150--162, 1994. 54, 55, 70
- [123] Svetlana Lazebnik, Cordelia Schmid, and Jean Ponce. A sparse texture representation using affine-invariant regions. In *Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on*, volume 2, pages II--319. IEEE, 2003. 95, 96, 106
- [124] Maxime Lhuillier and Long Quan. A quasi-dense approach to surface reconstruction from uncalibrated images. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 27(3):418--433, 2005. 56, 64
- [125] Tony Lindeberg. *Discrete Scale-space Theory and Scale-space Primal Sketch*. PhD thesis, Royal Institute of Technology, 1991. 100
- [126] Tony Lindeberg. Feature detection with automatic scale selection. *International journal of computer vision*, 30(2):79--116, 1998. 95

- [127] Tony Lindeberg and Lars Bretzner. Real-time scale selection in hybrid multi-scale representations. In *Scale Space Methods in Computer Vision*, pages 148--163. Springer, 2003. 101
- [128] Tony Lindeberg and Jonas Gårding. Shape-adapted smoothing in estimation of 3-d shape cues from affine deformations of local 2-d brightness structure. *Image and vision computing*, 15(6):415--434, 1997. 96
- [129] Manolis Lourakis. Sparse non-linear least squares optimization for geometric vision. *Computer Vision--ECCV 2010*, pages 43--56, 2010. 39
- [130] Manolis I. A. Lourakis and Antonis A. Argyros. Sba: a software package for generic sparse bundle adjustment. *ACM Transactions on Mathematical Software*, pages 1--30, 2009. 13, 27, 39, 45, 51
- [131] M.I. A. Lourakis and A.A. Argyros. Sparse bundle adjustment. <http://www.ics.forth.gr/~lourakis/sba/index.html>. [Online; accessed 22-February-2013]. 51
- [132] David Lowe. Demo software: Sift keypoint detector. <http://www.cs.ubc.ca/~lowe/keypoints/>. [Online; accessed 29-March-2013]. 114
- [133] David G Lowe. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2):91--110, 2004. 13, 73, 87, 88, 90, 91, 93, 96, 99, 103, 104, 108, 111, 166, 167, 168, 211
- [134] Quan-Tuan Luong and Olivier D Faugeras. The fundamental matrix: Theory, algorithms, and stability analysis. *International Journal of Computer Vision*, 17(1):43--75, 1996. 170, 243
- [135] Yi Ma, René Vidal, Jana Košecká, and Shankar Sastry. Kruppa equation revisited: Its renormalization and degeneracy. In *Computer Vision--ECCV 2000*, pages 561--577. Springer, 2000. 162, 165
- [136] Kaj Madsen, Hans Bruun, and Ole Tingleff. Methods for non-linear least squares problems. 1999. 51
- [137] Jan Mandel. On block diagonal and schur complement preconditioning. *Numerische Mathematik*, 58(1):79--93, 1990. 47
- [138] Anastasios Manassis, Adrian Hilton, Phil Palmer, Phil McLauchlan, and Xinquan Shen. Reconstruction of scene models from sparse 3d structure. In *Computer Vision and Pattern Recognition, 2000. Proceedings. IEEE Conference on*, volume 2, pages 666--671. IEEE, 2000. 55
- [139] Gian Luca Mariottini and Domenico Prattichizzo. Egt for multiple view geometry and visual servoing: robotics vision with pinhole and panoramic cameras. *Robotics & Automation Magazine, IEEE*, 12(4):26--39, 2005. 178
- [140] David Marr and Ellen Hildreth. Theory of edge detection. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, 207(1167):187--217, 1980. 64

- [141] Daniel Martinec and Tomáš Pajdla. Robust rotation and translation estimation in multiview reconstruction. In *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*, pages 1-8. IEEE, 2007. 39, 181, 182
- [142] Jiri Matas, Ondrej Chum, Martin Urban, and Tomáš Pajdla. Robust wide baseline stereo from maximally stable extremal regions. In *British machine vision conference*, volume 1, pages 384--393, 2002. 96
- [143] Jiri Matas, Ondrej Chum, Martin Urban, and Tomáš Pajdla. Robust wide-baseline stereo from maximally stable extremal regions. *Image and Vision Computing*, 22(10):761--767, 2004. 89
- [144] Nikos Melanitis-Paraskevas. Youtube channel. <http://www.youtube.com/channel/UCp5O-gzDEOZ77QpQl6BTQUg>. [Online; accessed 06-September-2013]. 190, 202
- [145] Krystian Mikolajczyk and Cordelia Schmid. Indexing based on scale invariant interest points. In *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on*, volume 1, pages 525--531. IEEE, 2001. 96, 100
- [146] Krystian Mikolajczyk and Cordelia Schmid. An affine invariant interest point detector. In *Computer Vision—ECCV 2002*, pages 128-142. Springer, 2002. 96, 98
- [147] Krystian Mikolajczyk and Cordelia Schmid. Scale & affine invariant interest point detectors. *International journal of computer vision*, 60(1):63--86, 2004. 96
- [148] Krystian Mikolajczyk and Cordelia Schmid. A performance evaluation of local descriptors. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 27(10):1615--1630, 2005. 87, 88, 96, 97, 108
- [149] Krystian Mikolajczyk, Tinne Tuytelaars, Cordelia Schmid, Andrew Zisserman, Jiri Matas, Frederik Schaffalitzky, Timor Kadir, and L Van Gool. A comparison of affine region detectors. *International journal of computer vision*, 65(1):43--72, 2005. 13, 87, 95
- [150] Krystian Mikolajczyk, Andrew Zisserman, Cordelia Schmid, et al. Shape recognition with edge-based features. In *British Machine Vision Conference (BMVC'03)*, volume 2, pages 779--788, 2003. 89
- [151] Hossein Mobahi, C Lawrence Zitnick, and Yi Ma. Optimisation by gaussian smoothing software. <http://perception.csl.illinois.edu/smoothing/>. [Online; accessed 29-March-2013]. 115
- [152] Hossein Mobahi, C Lawrence Zitnick, and Yi Ma. Seeing through the blur. In *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, pages 1736--1743. IEEE, 2012. 113, 115
- [153] Theo Moons, Maarten Vergauwen, and Luc Van Gool. 3d reconstruction from multiple images. 2008. 57

- [154] Jorge More. The levenberg-marquardt algorithm: implementation and theory. *Numerical analysis*, pages 105--116, 1978. 51
- [155] Daniel D Morris and Takeo Kanade. Image-consistent surface triangulation. In *Computer Vision and Pattern Recognition, 2000. Proceedings. IEEE Conference on*, volume 1, pages 332--338. IEEE, 2000. 55
- [156] Marius Muja and David Lowe. Fast library for approximate nearest neighbors. <http://www.cs.ubc.ca/~mariusm/index.php/FLANN/FLANN>. [Online; accessed 08-July-2013]. 178
- [157] Marius Muja and David Lowe. *FLANN - Fast Library for Approximate Nearest Neighbors*, January 2005. 178
- [158] PJ Narayanan, Peter W Rander, and Takeo Kanade. Constructing virtual worlds using dense stereo. In *Computer Vision, 1998. Sixth International Conference on*, pages 3--10. IEEE, 1998. 54, 57
- [159] Will Naylor and Bill Chapman. Library with efficient numerical algorithms. <http://www.willnaylor.com/wnlib.html>. [Online; accessed 10-March-2013]. 66
- [160] Kai Ni, D. Steedly, and F. Dellaert. Out-of-core bundle adjustment for large-scale 3d reconstruction. In *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on*, pages 1 --8, oct. 2007. 39, 40, 251
- [161] David Nistér. An efficient solution to the five-point relative pose problem. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 26(6):756--770, 2004. 46
- [162] David Nistér. Preemptive ransac for live structure and motion estimation. *Machine Vision and Applications*, 16(5):321--329, 2005. 37
- [163] Jorge Nocedal and Stephen J Wright. *Numerical optimization*. Springer verlag, 1999. 47, 51
- [164] John Oliensis and Richard Hartley. Iterative extensions of the sturm/triggs algorithm: convergence and nonconvergence. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 29(12):2217--2233, 2007. 29
- [165] Aude Oliva and Antonio Torralba. Modeling the shape of the scene: A holistic representation of the spatial envelope. *International journal of computer vision*, 42(3):145--175, 2001. 207
- [166] Enqvist Olof, Kahl Fredrik, and Ramalingam Srikumar. Non-sequential structure from motion. 2011. 181
- [167] Carl Olsson. Stable structure from motion matlab code. http://www.maths.lth.se/matematiklth/personal/calle/sys_paper/sys_paper.html. [Online; accessed 05-September-2013]. 182

- [168] Carl Olsson and Olof Enqvist. Stable structure from motion for unordered image collections. In *Image Analysis*, pages 524--535. Springer, 2011. 180, 181, 182, 183, 185, 187
- [169] Carl Olsson, Anders Eriksson, and Richard Hartley. Outlier removal using duality. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pages 1450--1457. IEEE, 2010. 181, 209
- [170] Carl Olsson and Fredrik Kahl. Generalized convexity in multiple view geometry. *Journal of Mathematical Imaging and Vision*, 38(1):35--51, 2010. 180, 184
- [171] GP Otto and TKW Chau. 'region-growing' algorithm for matching of terrain images. *Image and vision computing*, 7(2):83--94, 1989. 58, 64
- [172] Charles Parker. An analysis of performance measures for binary classifiers. In *Data Mining (ICDM), 2011 IEEE 11th International Conference on*, pages 517--526. IEEE, 2011. 218
- [173] Michal Perd'och, Ondrej Chum, and Jiri Matas. Efficient representation of local geometry for large scale object retrieval. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, pages 9--16. IEEE, 2009. 211
- [174] Ákos Pernek and Levente Hajder. Automatic focal length estimation as an eigenvalue problem. *Pattern Recognition Letters*, 2013. 154
- [175] James Philbin, Ondrej Chum, Michael Isard, Josef Sivic, and Andrew Zisserman. Object retrieval with large vocabularies and fast spatial matching. In *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*, pages 1--8. IEEE, 2007. 211
- [176] Marc Pollefeys, Reinhard Koch, and Luc Van Gool. Self-calibration and metric reconstruction inspite of varying and unknown intrinsic camera parameters. *International Journal of Computer Vision*, 32(1):7--25, 1999. 117
- [177] Marc Pollefeys, David Nistér, J-M Frahm, Amir Akbarzadeh, Phillipos Mordohai, Brian Clipp, Christopher Engels, David Gallup, S-J Kim, Paul Merrell, et al. Detailed real-time urban 3d reconstruction from video. *International Journal of Computer Vision*, 78(2):143--167, 2008. 73
- [178] J-P Pons, Renaud Keriven, O Faugeras, and Gerardo Hermosillo. Variational stereovision and 3d scene flow estimation with statistical similarity measures. In *Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on*, pages 597--602. IEEE, 2003. 52
- [179] J-P Pons, Renaud Keriven, and Olivier Faugeras. Modelling dynamic scenes by registering multi-view image sequences. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 2, pages 822--827. IEEE, 2005. 52, 53

- [180] Jean-Philippe Pons, Renaud Keriven, and Olivier Faugeras. Multi-view stereo reconstruction and scene flow estimation with a global image-based matching score. *International Journal of Computer Vision*, 72(2):179--193, 2007. 73
- [181] Domenico Prattichizzo and Gian Luca Mariottini. Epipolar geometry toolbox. <http://egt.dii.unisi.it/>. [Online; accessed 08-July-2013]. 178
- [182] Vincent Rabaud. Vincent's Structure from Motion Toolbox. <http://vision.ucsd.edu/~vrabaud/toolbox/>. 243
- [183] Jérôme Revaud, Matthijs Douze, Schmid Cordelia, Hervé Jégou, et al. Event retrieval in large video collections with circulant temporal encoding. In *CVPR-International Conference on Computer Vision and Pattern Recognition*, 2013. 207
- [184] R. Roberts, S.N. Sinha, R. Szeliski, and D. Steedly. Structure from motion for scenes with large duplicate structures. In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pages 3137--3144, 2011. 209
- [185] Alyn P Rockwood and Jim Winget. Three-dimensional object reconstruction from two-dimensional images. *Computer-Aided Design*, 29(4):279--285, 1997. 53
- [186] Carsten Rother and Stefan Carlsson. Linear multi view reconstruction and camera recovery using a reference plane. *International Journal of Computer Vision*, 49(2):117--141, 2002. 30, 37
- [187] Peter J Rousseeuw. Least median of squares regression. *Journal of the American statistical association*, 79(388):871--880, 1984. 168
- [188] Sébastien Roy and Ingemar J Cox. A maximum-flow formulation of the n-camera stereo correspondence problem. In *Computer Vision, 1998. Sixth International Conference on*, pages 492--499. IEEE, 1998. 52, 54
- [189] Szymon Rusinkiewicz and Marc Levoy. Qsplat: A multiresolution point rendering system for large meshes. In *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, pages 343--352. ACM Press/Addison-Wesley Publishing Co., 2000. 78
- [190] Yousef Saad and Y Saad. *Iterative methods for sparse linear systems*, volume 620. PWS publishing company Boston, 1996. 47
- [191] Hideo Saito and Takeo Kanade. Shape reconstruction in projective grid space from large number of images. In *Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on*, volume 2. IEEE, 1999. 55
- [192] Frederik Schaffalitzky and Andrew Zisserman. Multi-view matching for unordered image sets, or "how do i organize my holiday snaps?". In *Computer Vision—ECCV 2002*, pages 414--431. Springer, 2002. 95, 96

- [193] Steven M Seitz, Brian Curless, James Diebel, Daniel Scharstein, and Richard Szeliski. A comparison and evaluation of multi-view stereo reconstruction algorithms. In *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, volume 1, pages 519--528. IEEE, 2006. 13, 52, 70, 81, 86
- [194] Steven M Seitz and Charles R Dyer. Photorealistic scene reconstruction by voxel coloring. *International Journal of Computer Vision*, 35(2):151--173, 1999. 52, 53, 54
- [195] Yongduek Seo, Anders Heyden, and Roberto Cipolla. A linear iterative method for auto-calibration using the dac equation. In *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, volume 1, pages I--880. IEEE, 2001. 117
- [196] Fumin Shen, Chunhua Shen, Rhys Hill, Anton van den Hengel, and Zhenmin Tang. Fast approximate l_{∞} minimization: Speeding up robust regression. *arXiv preprint arXiv:1304.1250*, 2013. 181
- [197] Shuhan Shen. Accurate multiple view 3d reconstruction using patch based stereo for large-scale scenes. 2013. 64, 82, 83
- [198] Jianbo Shi and Jitendra Malik. Normalized cuts and image segmentation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 22(8):888--905, 2000. 75
- [199] Kristy Sim and Richard Hartley. Recovering camera motion using l_{∞} minimization. In *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, volume 1, pages 1230--1237. IEEE, 2006. 184, 209
- [200] Kristy Sim and Richard Hartley. Removing outliers using the l_{∞} norm. In *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, volume 1, pages 485--494. IEEE, 2006. 181
- [201] Sudipta N Sinha and Marc Pollefeys. Multi-view reconstruction using photo-consistency and exact silhouette constraints: A maximum-flow formulation. In *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on*, volume 1, pages 349--356. IEEE, 2005. 54, 55
- [202] Sudipta N Sinha, Drew Steedly, and Richard Szeliski. A multi-stage linear approach to structure from motion. In *ECCV 2010 Workshop on Reconstruction and Modeling of Large-Scale 3D Virtual Environments*, volume 3002, pages 3003--3005, 2010. 39, 181, 182, 207
- [203] Josef Sivic and Andrew Zisserman. Video google: A text retrieval approach to object matching in videos. In *Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on*, pages 1470--1477. IEEE, 2003. 211

- [204] Gregory Slabaugh, B Culbertson, T Malzbender, and Ron Schafer. Improved voxel coloring via volumetric optimization. *Center for Signal and Image Processing, Georgia Institute of Technology, Tech. Rep*, 2000. 55
- [205] Gregory G Slabaugh, W Bruce Culbertson, Thomas Malzbender, Mark R Stevens, and Ronald W Schafer. Methods for volumetric reconstruction of visual scenes. *International Journal of Computer Vision*, 57(3):179--199, 2004. 52, 53, 54, 55
- [206] Keith N Snavely. *Scene reconstruction and visualization from internet photo collections*. ProQuest, 2008. 168
- [207] Noah Snavely. Bundler: Structure from motion (sfm) for unordered image collection. <http://phototour.cs.washington.edu/bundler/>. [Online; accessed 22-February-2013]. 51, 58, 85, 180, 182
- [208] Noah Snavely. Bundler: Structure from motion (sfm) for unordered image collection. <http://www.di.ens.fr/willow/events/cvml2013/materials/practicals/instance-level-recognition.html>. [Online; accessed 06-September-2013]. 208
- [209] Noah Snavely. Scene reconstruction and visualization from internet photo collections: A survey. *IPSJ Transactions on Computer Vision and Applications*, 3(0):44--66, 2011. 27
- [210] Noah Snavely, Steven M Seitz, and Richard Szeliski. Photo tourism: exploring photo collections in 3d. In *ACM Transactions on Graphics (TOG)*, volume 25, pages 835--846. ACM, 2006. 39, 49, 51, 58, 167, 168, 180
- [211] Stefano Soatto, Anthony J Yezzi, and Hailin Jin. Tales of shape and radiance in multiview stereo. In *Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on*, pages 974--981. IEEE, 2003. 52, 53
- [212] Charles V Stewart. Robust parameter estimation in computer vision. *Siam Review*, 41(3):513--537, 1999. 37, 255
- [213] Henrik Stewenius, Christopher Engels, and David Nistér. Recent developments on direct relative orientation. *ISPRS Journal of Photogrammetry and Remote Sensing*, 60(4):284--294, 2006. 183, 185
- [214] Henrik Stewenius, David Nistér, Fredrik Kahl, and Frederik Schaffalitzky. A minimal solution for relative pose with unknown focal length. *Image and Vision Computing*, 26(7):871--877, 2008. 182
- [215] Christoph Strecha, Rik Fransens, and Luc Van Gool. Combined depth and outlier estimation in multi-view stereo. In *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, volume 2, pages 2394--2401. IEEE, 2006. 55

- [216] Christoph Strecha, Wolfgang Von Hansen, Luc Van Gool, Pascal Fua, and Ulrich Thoennessen. On benchmarking camera calibration and multi-view stereo for high resolution imagery. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pages 1--8. IEEE, 2008. 70, 81, 82, 83, 86, 166, 218, 219, 230, 236, 238, 251
- [217] Peter Sturm. Critical motion sequences for the self-calibration of cameras and stereo systems with variable focal length. In *The 10th British Machine Vision Conference (BMVC'99)*, pages 63--72, 1999. 162, 163, 164
- [218] Peter Sturm. A case against kruppa's equations for camera self-calibration. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 22(10):1199--1204, 2000. 162, 164, 165
- [219] Richard Szeliski. A multi-view approach to motion and stereo. In *Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on.*, volume 1. IEEE, 1999. 53, 55
- [220] Richard Szeliski. Prediction error as a quality metric for motion and stereo. In *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*, volume 2, pages 781--788. IEEE, 1999. 53
- [221] Richard Szeliski and Polina Golland. Stereo matching with transparency and matting. In *Computer Vision, 1998. Sixth International Conference on*, pages 517--524. IEEE, 1998. 52
- [222] Camillo J Taylor. Surface reconstruction from feature based stereo. In *Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on*, pages 184--190. IEEE, 2003. 55
- [223] Sergios Theodoridis and Konstantinos Koutroumbas. *Pattern Recognition, Third Edition*. Academic Press, Inc., Orlando, FL, USA, 2006. 95, 100, 106, 107
- [224] David Tingdahl and Luc Van Gool. Arc 3d webservice. <http://homes.esat.kuleuven.be/~visit3d/webservice/v3/links.php>. [Online; accessed 29-March-2013]. 86
- [225] Engin Tola, Vincent Lepetit, and Pascal Fua. Daisy: An efficient dense descriptor applied to wide-baseline stereo. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 32(5):815--830, 2010. 13, 82, 87, 108, 109, 110, 111, 112, 255
- [226] Engin Tola, Christoph Strecha, and Pascal Fua. Efficient large scale multi-view stereo for ultra high resolution image sets. <http://www.engintola.com/research/emvs/>. [Online; accessed 22-February-2013]. 82, 114, 178
- [227] Engin Tola, Christoph Strecha, and Pascal Fua. Efficient large-scale multi-view stereo for ultra high-resolution image sets. *Machine Vision and Applications*, pages 1--18, 2011. 82, 83, 172

- [228] Philip H. S. Torr. Bayesian model estimation and selection for epipolar geometry and generic manifold fitting. *International Journal of Computer Vision*, 50(1):35--61, 2002. 174
- [229] Philip HS Torr and David W Murray. The development and comparison of robust methods for estimating the fundamental matrix. *International journal of computer vision*, 24(3):271--300, 1997. 168, 171
- [230] Lloyd N Trefethen and David Bau III. *Numerical linear algebra*. Number 50. Society for Industrial Mathematics, 1997. 44
- [231] Adrien Treuille, Aaron Hertzmann, and Steven Seitz. Example-based stereo with general brdfs. *Computer Vision-ECCV 2004*, pages 457--469, 2004. 52, 53, 54
- [232] Bill Triggs. Bill triggs' software. <http://ljk.imag.fr/membres/Bill.Triggs/src/>. [Online; accessed 08-July-2013]. 178
- [233] Bill Triggs, Philip McLauchlan, Richard Hartley, and Andrew Fitzgibbon. Bundle adjustment—a modern synthesis. *Vision algorithms: theory and practice*, pages 153--177, 2000. 39
- [234] Roger Tsai. A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf tv cameras and lenses. *Robotics and Automation, IEEE Journal of*, 3(4):323--344, 1987. 24, 26, 251
- [235] Panu Turcot and David G Lowe. Better matching with fewer features: The selection of useful features in large database recognition problems. In *Computer Vision Workshops (ICCV Workshops), 2009 IEEE 12th International Conference on*, pages 2109--2116. IEEE, 2009. 211
- [236] Peter van Emde Boas. Preserving order in a forest in less than logarithmic time. In *FOCS*, pages 75--84, 1975. 213
- [237] Luc Van Gool, Theo Moons, and Dorin Ungureanu. Affine/photometric invariants for planar intensity patterns. In *Computer Vision—ECCV'96*, pages 642--651. Springer, 1996. 95, 96
- [238] Manik Varma and Debajyoti Ray. Learning the discriminative power-invariance trade-off. In *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on*, pages 1--8. IEEE, 2007. 114
- [239] A. Vedaldi and B. Fulkerson. VLFeat: An open and portable library of computer vision algorithms. <http://www.vlfeat.org/>, 2008. 114, 178
- [240] George Vogiatzis, Philip HS Torr, and Roberto Cipolla. Multi-view stereo via volumetric graph-cuts. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 2, pages 391--398. IEEE, 2005. 52, 54, 55, 70, 73

- [241] H-H Vu, Patrick Labatut, J-P Pons, and Renaud Keriven. High accuracy and visibility-consistent dense multiview stereo. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 34(5):889-901, 2012. **83, 84**
- [242] Simon Winder, Gang Hua, and Matthew Brown. Picking the best daisy. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, pages 178--185. IEEE, 2009. **114**
- [243] Simon AJ Winder and Matthew Brown. Learning local image descriptors. In *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*, pages 1--8. IEEE, 2007. **13, 87, 103, 105, 106, 107, 207**
- [244] H. S. Wong, T.-J. Chin, J. Yu, and D. Suter. Dynamic and hierarchical multi-structure geometric model fitting. In *International Conference on Computer Vision (ICCV)*, 2011. **218, 219, 243**
- [245] SJ Wright and John Norman Holt. An inexact levenberg-marquardt method for large sparse nonlinear least squares. *The Journal of the Australian Mathematical Society. Series B. Applied Mathematics*, 26(04):387--403, 1985. **47**
- [246] Changchang Wu. Siftgpu: A gpu implementation of scale invariant feature transform (sift). <http://www.cs.unc.edu/~ccwu/siftgpu/>. [Online; accessed 29-March-2013]. **114**
- [247] Changchang Wu. Visualsfm: A visual structure from motion system. <http://homes.cs.washington.edu/~ccwu/vsfm/>. [Online; accessed 22-February-2013]. **51, 58, 86**
- [248] Changchang Wu. SiftGPU: A GPU implementation of scale invariant feature transform (SIFT). <http://cs.unc.edu/~ccwu/siftgpu>, 2007. **88**
- [249] Changchang Wu, Sameer Agarwal, Brian Curless, and Steven M Seitz. Multicore bundle adjustment. In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pages 3057--3064. IEEE, 2011. **13, 27, 47, 51, 73, 180**
- [250] Zhong Wu, Qifa Ke, Michael Isard, and Jian Sun. Bundling features for large scale partial-duplicate web image search. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, pages 25--32. IEEE, 2009. **211**
- [251] Ruigang Yang, Marc Pollefeys, and Greg Welch. Dealing with textureless regions and specular highlights-a progressive space carving scheme using a novel photo-consistency measure. In *Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on*, pages 576--584. IEEE, 2003. **52, 53, 55**
- [252] Lifan Yao, Hao Feng, Yiqun Zhu, Zhiguo Jiang, Danpei Zhao, and Wenquan Feng. An architecture of optimised sift feature detection for an fpga implementation of an image matcher. In *Field-Programmable Technology, 2009. FPT 2009. International Conference on*, pages 30--37, 2009. **88**

- [253] Tianli Yu, Ning Xu, and Narendra Ahuja. Shape and view independent reflectance map from multiple views. *International journal of computer vision*, 73(2):123--138, 2007. 53
- [254] Ze-Huan Yuan, Lu Tong, Hao-Yi Zhou, Chen Bin, and Jia-Ning Li. Incremental 3d reconstruction using bayesian learning. *Advanced Research in Applied Artificial Intelligence*, pages 754--763, 2012. 84
- [255] Ramin Zabih and John Woodfill. Non-parametric local transforms for computing visual correspondence. In *Computer Vision--ECCV'94*, pages 151--158. Springer, 1994. 95
- [256] Christopher Zach and Marc Pollefeys. Practical methods for convex multi-view reconstruction. In *Computer Vision--ECCV 2010*, pages 354--367. Springer, 2010. 180, 184
- [257] Gang Zeng, Sylvain Paris, Long , and François Sillion. Progressive surface reconstruction from images using a local prior. In *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on*, volume 2, pages 1230--1237. IEEE, 2005. 54, 55
- [258] Li Zhang and Steven M Seitz. Image-based multiresolution shape recovery by surface deformation. In *Photonics West 2001-Electronic Imaging*, pages 51--61. International Society for Optics and Photonics, 2000. 53, 54
- [259] Bin Fan Zhenhua Wang and Fuchao Wu. Local intensity order pattern for feature description. In *IEEE International Conference on Computer Vision (ICCV)*, pages 603--610, Nov. 2011. 207
- [260] Andrew Zisserman. Matlab functions for multiple view geometry. <http://www.robots.ox.ac.uk/~vgg/hzbook/code/>. 178
- [261] C Lawrence Zitnick, Sing Bing Kang, Matthew Uyttendaele, Simon Winder, and Richard Szeliski. High-quality video view interpolation using a layered representation. In *ACM Transactions on Graphics (TOG)*, volume 23, pages 600--608. ACM, 2004. 53, 55