

# **TEHNICI DE ALINIERE A IMAGINILOR UTILIZÂND ESTIMARE NEPARAMETRICĂ DE DENSITATE DE PROBABILITATE**

Teză destinată obținerii  
titlului științific de doctor inginer  
la

Universitatea "Politehnica" din Timișoara  
în domeniul INGINERIE ELECTRONICĂ ȘI  
TELECOMUNICAȚII  
de către

**Ing. Daniela Fuiorea-Bulucea**

Conducător științific: prof.univ.dr.ing. Corneliu I. Toma  
Referenți științifici: prof.univ.dr.ing. Mircea Petrescu  
prof.univ.dr.ing. Gavril Todorean  
prof.univ.dr.ing. Vasile Gui

Ziua susținerii tezei: 10.07.2009

Seriile Teze de doctorat ale UPT sunt:

- |                        |   |
|------------------------|---|
| 1. Automatică          | 7. Inginerie Electronică și Telecomunicații |
| 2. Chimie              | 8. Inginerie Industrială                    |
| 3. Energetică          | 9. Inginerie Mecanică                       |
| 4. Ingineria Chimică   | 10. Știința Calculatoarelor                 |
| 5. Inginerie Civilă    | 11. Știința și Ingineria Materialelor       |
| 6. Inginerie Electrică |   |

Universitatea „Politehnica” din Timișoara a inițiat seriile de mai sus în scopul diseminării expertizei, cunoștințelor și rezultatelor cercetărilor întreprinse în cadrul școlii doctorale a universității. Seriile conțin, potrivit H.B.Ex.S Nr. 14 / 14.07.2006, tezele de doctorat susținute în universitate începând cu 1 octombrie 2006.

Copyright © Editura Politehnica – Timișoara, 2009

Această publicație este supusă prevederilor legii dreptului de autor. Multiplicarea acestei publicații, în mod integral sau în parte, traducerea, tipărirea, reutilizarea ilustrațiilor, expunerea, radiodifuzarea, reproducerea pe microfilme sau în orice altă formă este permisă numai cu respectarea prevederilor Legii române a dreptului de autor în vigoare și permisiunea pentru utilizare obținută în scris din partea Universității „Politehnica” din Timișoara. Toate încălcările acestor drepturi vor fi penalizate potrivit Legii române a drepturilor de autor.

România, 300159 Timișoara, Bd. Republicii 9,  
tel. 0256 403823, fax. 0256 403221  
e-mail: editura@edipol.upt.ro

## Cuvânt înainte

Teza de doctorat a fost elaborată pe parcursul activității mele în cadrul Departamentului de Comunicații al Universității „Politehnica” din Timișoara. Aceasta constituie rezultatul unui efort susținut de-a lungul unei perioade de patru ani.

Lucrarea este o abordare nouă a procesului de aliniere a imaginilor, utilizând metode neparametrice de densitate de probabilitate și metoda mean shift pentru estimarea parametrilor transformării de similaritate.

Teza își propune să demonstreze în ce măsură metoda de estimare neparametrică este mai avantajoasă în raport cu metodele de eșantionare aleatoare, considerate de referință, în contextul aplicațiilor de rețele de senzori video fără fir. Scopul utilizării acestor metode neparametrice este de a determina câmpul vizual comun dintre senzorii video cât și orientările relative ale acestora. Domeniul de aplicabilitate este unul de mare interes în cadrul conflictelor militare și teroriste de dată recentă.

Cercetările efectuate au condus la realizarea unei aplicații cu operare în timp real care realizează alinierea optimă a unor perechi de imagini care conțin un câmp vizual comun, captat din diferite poziționări și orientări ale senzorului.

În încheiere doresc să aduc mulțumiri deosebite conducătorului de doctorat prof.dr.ing. Corneliu I. Toma, care prin supraveghere constantă și sfaturi neprețuite a contribuit la realizarea prezentei lucrări.

Timișoara, 06 2009

Daniela Fuiorea-Bulucea

Destinatarii dedicației.

Fuiorea-Bulucea, Daniela Narcisa

**Tehnici de aliniere a imaginilor utilizând estimare neparametrică de densitate de probabilitate**

Teze de doctorat ale UPT, Seria 7, Nr. 14, Editura Politehnica, 2009, 150 pagini, 71 figuri, 22 tabele.

ISSN:1842-7014

ISBN:978-973-625-908-1

**Cuvinte cheie:**

alinierea imaginilor, metode robuste, mean shift, eşantionare aleatoare, rețele de senzori video fără fir.

**Rezumat:**

Teza de doctorat este dedicată unei probleme de mare interes în vederea artificială, și anume, alinierea imaginilor. Problema constă în determinarea transformării geometrice care aliniază partea comună a unei perechi de imagini ale aceleiași scene captate cu camere poziționate diferit, pe baza corespondențelor dintre două seturi de puncte, extrase din cele două imagini. Aplicația principală care a inspirat soluțiile adoptate este localizarea senzorilor într-o rețea fără fir.

Perechile de imagini ce intervin într-o aplicație de localizare a senzorilor de imagine conțin frecvent un câmp vizual comun ce ocupă un procentaj redus din imagine și obiectele comune apar la scări diferite. În consecință, punctele caracteristice pe baza cărora se poate face estimarea transformării geometrice sunt puține, iar cele din imaginea aflată la distanță mai mare de cameră sunt afectate de erori de localizare mai mari. Metodele tradiționale bazate pe algoritmul RANSAC sau oricare din variantele sale ameliorate întâmpină dificultăți, pentru că sunt utilizate în condiții în care ipotezele avute în vedere de autorii acestora își pierd valabilitatea. Metodele de aliniere a imaginilor bazate pe estimare neparametrică de densitate de probabilitate, propuse în teză, sunt mai bine adaptate aplicației menționate și oricărei alte aplicații cu condiții similare.

Performanțele estimatorului robust bazat pe algoritmul mean shift au fost evaluate prin experimente, efectuate comparativ cu metoda clasică RANSAC și cu metoda ameliorată, MLESAC. Pe baza rezultatelor experimentale s-a demonstrat eficacitatea metodei mean shift în contextul alinierii imaginilor.



## CUPRINS

CUPRINS.....	5
LISTĂ DE TABELE.....	7
LISTĂ DE FIGURI.....	8
1. MOTIVAȚIE .....	10
1.1. Considerații generale privind alinierea imaginilor.....	10
1.2. Actualitatea temei.....	14
1.3. Structura tezei de doctorat.....	15
1.4. Lucrări publicate de doctorand .....	17
2. TEHNICI DE ALINIERE A IMAGINILOR .....	18
2.1. Considerații generale.....	18
2.2. Extragerea caracteristicilor .....	20
2.3. Corespondența caracteristicilor .....	22
2.4. Alegerea modelului de transformare și estimarea .....	23
parametrilor modelului .....	23
2.5. Transformarea și reeșantionarea imaginii .....	25
2.6. Categoriile de aliniere și problemele care apar .....	25
2.6.1. Aliniere prin tehnici multimodale.....	25
2.6.2. Alinierea modelelor din imagini.....	26
2.6.3. Alinierea imaginilor 3D .....	27
2.6.4. Alinierea temporală.....	28
2.7. Caracteristicile metodelor de aliniere .....	28
2.8. Aplicații .....	30
2.8.1. Problematika realizării alinierii imaginilor .....	30
2.8.2. Localizarea în rețele de senzori fără fir .....	31
2.9. Concluzii .....	34
3. FUNDAMENTELE TEORETICE PRIVIND TEHNICILE DE ALINIERE A IMAGINILOR.....	35
3.1. Transformări geometrice.....	35
3.1.1. Transformări geometrice bidimensionale (2D) .....	35
3.1.2. Transformările geometrice tridimensionale (3D) .....	41
3.2. Tehnici de estimare robustă .....	50
3.3. Estimatorii M .....	52
3.3.1. Legătura cu estimatorii de verosimilitate maximă .....	52
3.3.2. Funcția de influență și funcția de pondere a estimatorilor M .....	54
3.4. Minimizarea mediane pătratelor rezidurilor (Least Median of Squares – LMeS).....	55
3.5. Algoritmul RANSAC .....	56
3.6. Algoritmul MSAC .....	59
3.7. Algoritmul MLESAC.....	59
3.8. Algoritmul NAPSAC.....	62
3.9. Concluzii .....	63
4. METODE PROPUSE DE ALINIERE A IMAGINILOR BAZATE PE ESTIMAREA NEPARAMETRICĂ A DENSITĂȚII DE PROBABILITATE .....	64
4.1. Estimarea densității de probabilitate folosind funcții nucleu .....	65
4.1.1. Algoritmul neparametric mean shift .....	65
4.1.2. Selecția scării algoritmului .....	67

4.2. Metoda robustă propusă de aliniere a imaginilor .....	69
4.3. Metoda de asociere prin compararea punctelor mediane .....	72
4.4. Utilizarea algoritmului de aliniere a imaginilor în rețele de senzori video fără fir .....	74
4.5. Discuție teoretică .....	80
4.6. Concluzii .....	81
5. STUDIU EXPERIMENTAL .....	82
5.1. Detalii de implementare .....	82
5.2. Alinierea 2D a imaginilor utilizând metoda robustă meanshift .....	89
5.3. Utilizarea metodei de aliniere a imaginilor pentru rețele de senzori video fără fir .....	92
5.4. Metoda de asociere prin compararea punctelor mediane .....	95
5.5. Studiul selecției punctelor caracteristice pentru localizarea senzorilor din rețele fără fir .....	103
5.6. Studiu comparativ (I): Mean shift versus RANSAC .....	108
5.7. Studiu comparativ (II): Mean shift versus RANSAC versus MLESAC. ....	113
5.7.1. Experimente cu o cameră fixă .....	113
5.7.2. Experimente cu o cameră mobilă .....	118
5.8. Concluzii .....	128
6. CONCLUZII GENERALE ȘI CONTRIBUȚII .....	130
6.1. Contribuții teoretice .....	131
6.2. Contribuții aplicative .....	132
6.3. Direcții de cercetare. Obiective pentru viitor .....	133
BIBLIOGRAFIE .....	134

## LISTĂ DE TABELE

Tabel 1. Deviația standard a erorilor componentelor vectorului de parametrii .....	92
Tabel 2. Deviația standard a erorilor componentelor vectorului de parametri .....	98
Tabel 3. Deviația standard a erorii vectorului timp .....	99
Tabel 4. Eroarea medie pătratică pentru 25 de puncte caracteristice selectate automat .....	104
Tabel 5. Eroarea medie pătratică pentru 25 de puncte caracteristice selectate manual .....	104
Tabel 6. Timpul de execuție pentru selecția manuală și automată .....	107
Tabel 7. Eroarea medie pătratică pentru RANSAC - T=1 puncte puține .....	109
Tabel 8. Eroarea medie pătratică pentru RANSAC - T=2 puncte puține .....	109
Tabel 9. Eroarea medie pătratică pentru RANSAC - T=3 puncte puține .....	109
Tabel 10. Eroarea medie pătratică pentru Mean Shift puncte puține .....	109
Tabel 11. Eroarea medie pătratică pentru puține și selecție manuală - MEAN SHIFT .....	113
Tabel 12. Eroarea medie pătratică pentru puncte puține și selecție manuală - RANSAC .....	114
Tabel 13. Eroarea medie pătratică pentru puncte puține și selecție manuală - MLESAC .....	114
Tabel 14. Eroarea medie pătratică pentru puncte multe și selecție manuală - MEAN SHIFT .....	115
Tabel 15. Eroarea medie pătratică pentru puncte multe și selecție manuală - RANSAC .....	115
Tabel 16. Eroarea medie pătratică pentru puncte multe și selecție manuală - MLESAC .....	115
Tabel 17. Eroarea medie pătratică pentru puncte puține și selecție automată - MEAN SHIFT .....	116
Tabel 18. Eroarea medie pătratică pentru puncte puține și selecție automată - RANSAC .....	116
Tabel 19. Eroarea medie pătratică pentru puncte puține și selecție automată - MLESAC .....	116
Tabel 20. Eroarea medie pătratică pentru puncte multe și selecție automată - MEAN SHIFT .....	117
Tabel 21. Eroarea medie pătratică pentru puncte multe și selecție automată - RANSAC .....	117
Tabel 22. Eroarea medie pătratică pentru puncte multe și selecție automată - MLESAC .....	118

## LISTĂ DE FIGURI

Fig. 1.1. (a) Imaginea de referință. (b) Imaginea captată. (c) Rezultatul punerii în corespondență a imaginilor. (d) Suprapunerea imaginii transformate în urma corespondenței, peste cea de referință. ....	11
Fig. 1.2. Etapele parcurse în realizarea unei metode de aliniere a imaginilor. Imaginile din coloana 2 sunt supuse procesului de aliniere. Scopul este asocierea acestora cu imaginile din coloana 1. La ieșire se poate observa rezultatul metodei de aliniere.....	13
Fig. 2.1.Parcurgerea metodelor de aliniere a imaginilor.....	20
Fig. 2.2.Spațiul caracteristicilor utilizat în metodele de aliniere a.....	21
Fig. 2.3.Descriptorul local generat de SIFT.....	22
Fig. 2.4.Exemple de diferite funcții de asociere: a) transformarea de similaritate; b) transformarea afină; c) perspectiva de proiecție; d) transformarea elastică.....	24
Fig. 3.1. Translația unui obiect bidimensional. ....	36
Fig. 3.2. Scalarea unui obiect bidimensional. ....	37
Fig. 3.3. Rotația unui obiect față de originea sistemului de coordonate. ....	37
Fig. 3.4. Homografia indusă de un plan.....	39
Fig. 3.5. Transformarea homografică plană. ....	39
Fig. 3.6. Imaginea de intrare a) este procesată în funcție de imaginea transformată b) cu ajutorul transformării homografice, în imaginea de ieșire c). ....	40
Fig. 3.7. Translația unui obiect tridimensional.....	42
Fig. 3.8. Scalarea unui obiect tridimensional. ....	43
Fig. 3.9. Rotația unui obiect tridimensional.....	44
Fig. 3.10. Rotația după axa x. ....	44
Fig. 3.11. Rotația după axa y. ....	45
Fig. 3.12. Rotația după axa z.....	46
Fig. 3.13. Proiecția de perspectivă .....	47
Fig. 3.14. Proiecția de perspectivă. ....	48
Fig. 3.15. Proiecția unui punct din lumea reală în planul imaginii.....	48
Fig. 4.1. Reprezentarea scării kernel pentru $h=0.01$ (lățime de bandă mică)[217].	68
Fig. 4.2. Reprezentarea scării kernel pentru $h=0.2$ (lățime de bandă mare)[217].	68
Fig. 4.3. Vectorii de translație înainte (stânga) și după (dreapta) rotație. ....	70
Fig. 4.4. Reprezentarea punctelor caracteristice și a punctului median. ....	73
Fig. 4.5. Câmpurile vizuale comune ale camerelor.....	74
Fig. 4.6. Transformarea imaginii bazată pe estimarea unghiului de rotație: (a) Imaginea originală din senzorul 1; (b) Imaginea originală din senzorul 2; (c) Imaginea din senzorul 2 după transformare. ....	76
Fig. 4.7. Post-procesare prin asociere Chamfer: (a) Diferența imaginilor înainte de post-procesare; (b) Diferența imaginilor după de post-procesare. ....	78
Fig. 4.8. (a) Imaginea de la Nodul 1; (b) Imaginea de la Nodul 2; .....	79
Fig. 5.1. Primul set de imagini folosit în procesul de aliniere. ....	83
Fig. 5.2. Al doilea set de imagini folosit în procesul de aliniere. ....	86
Fig. 5.3. Captură de ecran cu imaginile care vor fi supuse procesului de aliniere. ..	87
Fig. 5.4. Reprezentarea selecției caracteristicilor.....	87
Fig. 5.5. Apelarea metodelor de estimare robuste. ....	88
Fig. 5.6. Rezultatul procesului de aliniere a imaginilor. ....	89
Fig. 5.7. a) Imaginea de referință; b) Imaginea țintă; c) Imaginea aliniată cu metoda robustă mean shift; d) Imaginea aliniată cu estimatorul least-squares; e) Eroarea de asociere pentru metoda robustă mean shift; f) Eroarea de asociere pentru estimatorul least-squares. ....	90

Fig. 5.8. Erorile de estimare ale metodelor robuste mean shift și least-squares cu un procentaj de erori de 10%.....	91
Fig. 5.9. Erorile de estimare ale metodelor robuste mean shift și least-squares cu un procentaj de erori de 33%.....	91
Fig. 5.10. Eroarea estimării translației pe orizontală.....	93
Fig. 5.11. Eroarea estimării translației pe verticală.....	93
Fig. 5.12. Eroarea estimării factorului unghi.....	94
Fig. 5.13. Eroarea estimării factorului scară.....	94
Fig. 5.14. Eroarea estimării translației pe orizontală.....	96
Fig. 5.15. Eroarea estimării translației pe verticală.....	96
Fig. 5.16. Eroarea estimării factorului unghi.....	97
Fig. 5.17. Eroarea estimării factorului scară.....	97
Fig. 5.18. Eroarea de estimare factorului de timp.....	98
Fig. 5.19. Imaginile utilizate în algoritmul de punere în corespondență.....	100
Fig. 5.20. Eroarea estimării translației pe orizontală.....	100
Fig. 5.21. Eroarea estimării translației pe verticală.....	101
Fig. 5.22. Eroarea estimării factorului unghi.....	101
Fig. 5.23. Eroarea estimării factorului scară.....	102
Fig. 5.24. Numărul de soluții acceptate.....	102
Fig. 5.25. Fotografii din diferite medii:.....	107
Fig. 5.26. (a) Eroarea de translație pe orizontală, (b) Eroarea de translație pe verticală,.....	112
Fig. 5.27. Eroarea de translație pe orizontală pentru puncte puține și selecție manuală.....	119
Fig. 5.28. Eroarea de translație pe verticală pentru puncte puține și selecție manuală.....	119
Fig. 5.29. Eroarea de scară pentru puncte puține și selecție manuală.....	120
Fig. 5.30. Eroarea de unghi pentru puncte puține și selecție manuală.....	120
Fig. 5.31. Eroarea de translație pe orizontală pentru puncte multe și selecție manuală.....	121
Fig. 5.32. Eroarea de translație pe verticală pentru puncte multe și selecție manuală.....	121
Fig. 5.33. Eroarea de scară pentru puncte multe și selecție manuală.....	122
Fig. 5.34. Eroarea de unghi pentru puncte multe și selecție manuală.....	122
Fig. 5.35. Eroarea de translație pe orizontală pentru puncte puține și selecție automată.....	123
Fig. 5.36. Eroarea de translație pe verticală pentru puncte puține și selecție automată.....	123
Fig. 5.37. Eroarea de scară pentru puncte puține și selecție automată.....	124
Fig. 5.38. Eroarea de unghi pentru puncte puține și selecție automată.....	124
Fig. 5.39. Eroarea de translație pe orizontală pentru puncte multe și selecție automată.....	125
Fig. 5.40. Eroarea de translație pe verticală pentru puncte multe și selecție automată.....	126
Fig. 5.41. Eroarea de scară pentru puncte multe și selecție automată.....	126
Fig. 5.42. Eroarea de unghi pentru puncte multe și selecție automată.....	127

# 1. MOTIVAȚIE

Capitolul introductiv al tezei face o prezentare a problematicii realizării alinierii dintre două sau mai multe imagini și plasează în context estimarea neparametrică robustă, problemă asupra căreia se concentrează cea mai mare parte a contribuțiilor din teză. Primul paragraf prezintă considerații generale privind metodele de punere în corespondență a imaginilor. Sunt trecute în revistă și principalele domenii de cercetare în care se utilizează punerea în corespondență a imaginilor. În paragraful al doilea se argumentează actualitatea temei de punere în corespondență a imaginilor. Paragraful final prezintă structura tezei de doctorat.

## 1.1. Considerații generale privind alinierea imaginilor

Oamenii primesc pe cale vizuală cea mai mare parte din informația pe care sistemul lor senzorial o achiziționează. Oamenii preistorici au pictat, mult înainte de a scrie, transmitând în acest mod informații selectate. Probabil că am putea identifica aceste picturi din peșteri ca primele imagini artificiale cu aplicabilitate grafică. Apariția și ulterior dezvoltarea calculatoarelor a oferit posibilitatea ca imaginile artificiale să fie create și memorate sub formă numerică. Noul domeniu de grafică pe calculator este și în prezent foarte cunoscut și utilizat. Informația, sub formă de imagine de sinteză, furnizată de o aplicație de grafică pe calculator este legată de lumea înconjurătoare numai în măsura în care se dorește ca această realitate să fie mai mult sau mai puțin copiată. Decodificarea informației transmise prin aceste imagini este făcută de oamenii care le privesc și care sunt utilizatorii finali ai aplicației. Din acest punct de vedere, imaginile au o rezoluție din ce în ce mai bună pentru a putea reprezenta cât mai bine detaliile. Numărul de culori este de asemenea din ce în ce mai mare pentru ca fidelitatea reprezentărilor să fie susținută și din acest punct de vedere. Se creează în acest mod o realitate virtuală aflată sub controlul total al programatorului care a proiectat-o. Imaginile achiziționate pot fi afișate, ca și imaginile de sinteză, pentru a fi privite și analizate de un observator uman. Dacă se dorește să se determine conținutul imaginii achiziționate fără a face apel la un observator uman atunci apare o problemă nouă, interpretarea sau înțelegerea imaginilor folosind tehnici de inteligență artificială.

În prezent, în domeniul vederii artificiale, una din problemele care ridică semne de întrebare și necesită un interes aparte din partea cercetătorilor, o reprezintă alinierea imaginilor. Această problemă este un subiect foarte important în domeniul prelucrării de imagini și realizează suprapunerea, asocierea a două imagini captate la momente diferite de timp, din diferite unghiuri sau cu diferiți senzori. Pentru a realiza alinierea a două imagini, trebuie găsită o transformare pentru care fiecare punct dintr-o imagine să se potrivească (să aibă corespondent) cu un punct din a doua imagine. Această asociere (în engleză mapping) trebuie să realizeze alinierea optimală a două imagini, optimalitatea depinzând de ce trebuie "potrivit". Procesul de suprapunere a imaginilor se numește *punerea în corespondență a imaginilor sau alinierea imaginilor*.

Prin alinierea a două imagini, se pot determina schimbările scenei și se pot recunoaște obiectele. Un exemplu de aliniere 2D a imaginilor poate fi observat în figura 1.1. Figura 1.1.a prezintă imaginea captată de o cameră video și poartă numele de imagine de referință. Figura 1.1.b prezintă imaginea captată de a doua cameră video și care poartă numele de imagine captată (țintă), cea care urmează a fi transformată. Procesul de aliniere găsește o transformare care să realizeze alinierea imaginilor luate de cei doi senzori. Imaginea care se obține este imaginea transformată, prezentată în figura 1.1.c. Figura 1.1.d arată suprapunerea imaginii transformate peste cea de referință.

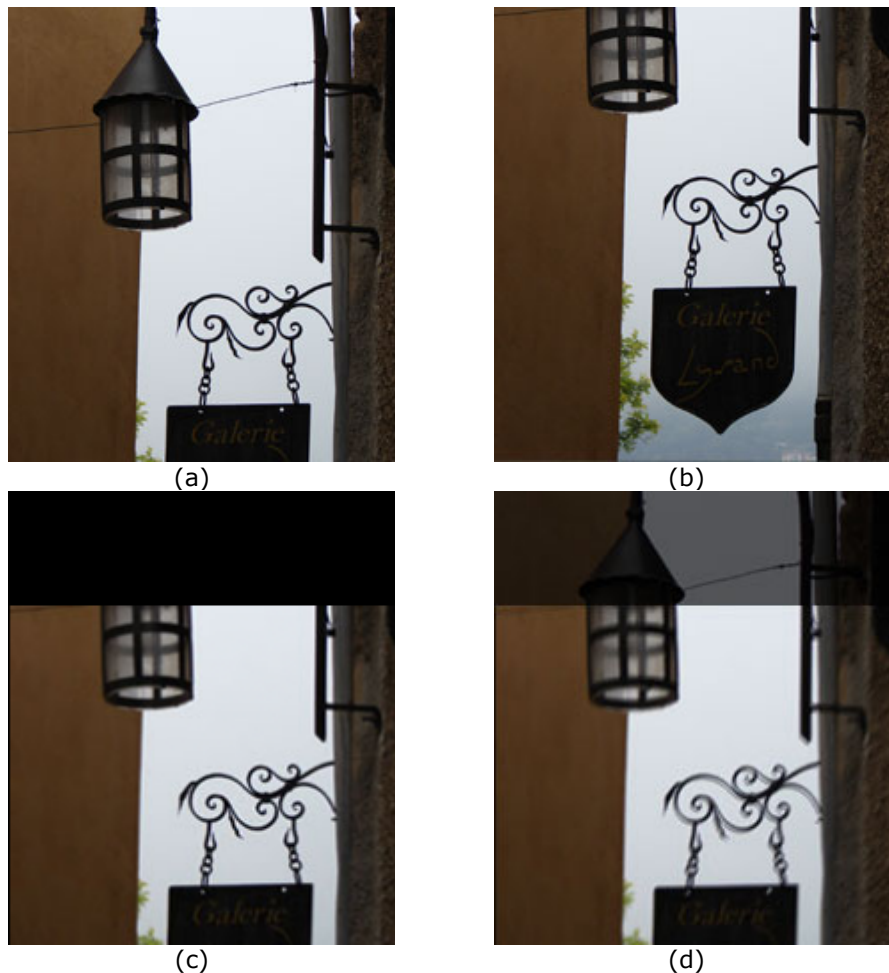


Fig. 1.1. (a) Imaginea de referință. (b) Imaginea captată. (c) Rezultatul punerii în corespondență a imaginilor. (d) Suprapunerea imaginii transformate în urma corespondenței, peste cea de referință.

Datorită diversității de imagini care urmează a fi supuse procesului de aliniere, și datorită diverselor tipuri de transformări, este imposibil de a dezvolta o metodă aplicabilă pentru toate problemele de corespondență. Fiecare metodă ar

trebuie să țină seama nu numai de tipul de transformare geometrică dintre imagini, ci și de deformațiile radiometrice și factorul zgomot.

Metodele de aliniere a imaginilor se realizează în patru etape:

- Extragerea caracteristicilor;
- Stabilirea corespondenței caracteristicilor;
- Estimarea modelului de transformare;
- Reșantionarea imaginii și transformarea ei.

Unul din motivele pentru care alinierea imaginilor reprezintă o problemă de interes este determinarea transformării geometrice optime pe baza căreia se realizează alinierea imaginilor. Întrebarea principală este de unde se poate lua informația necesară găsirii transformatei și cum se poate extrage ea din datele existente. Aceste date poartă numele de *caracteristici*. În realizarea procesului de punere în corespondență se disting două metode: metode bazate pe regiuni și metode bazate pe caracteristici. Caracteristicile trebuie să fie ușor de detectat în ambele imagini, împrăștiate pe toată imaginea și stabile în timp, pe tot parcursul procesului de punere în corespondență. Ele pot fi margini ale obiectelor, linii, intersecții de linii, puncte, regiuni, marker-ere extrinseci etc.

Punerea în corespondență este îngreunată adesea de prezența unor distorsiuni și erori ce afectează transformarea propriu-zisă ce trebuie estimată și inversată. Un factor comun care dă naștere la erori este zgomotul care apare în procesul de achiziție al imaginii și al extragerii caracteristicilor. Prezența zgomotului are ca rezultat o aliniere neexactă a caracteristicilor care rezultă în urma procesului de extragere. Un alt factor care duce la erori este prezența eșantioanelor aberante, ceea ce înseamnă că multe caracteristici pot exista într-un set dintr-o imagine fără corespondent în cel de al doilea set, din cea de a doua imagine. Aceste eșantioane trebuie eliminate în timpul procesului de aliniere.

Un algoritm de punere în corespondență trebuie să fie capabil să rezolve următoarele probleme:

- să determine corespondența dintre două seturi de caracteristici;
- să elimine datele neconforme și să determine o transformare care să alinieze un set de caracteristici peste celălalt;
- să necesite o intervenție umană cât mai puțină și să reducă timpul de calcul.

În realizarea unui algoritm de punere în corespondență trebuie avut în vedere parcurgea următoarelor etape, prezentate în figura 1.2.

- Prima etapă constă în extragerea caracteristicilor principale și speciale (colțuri, margini, contururi, linii de intersecție, etc.). Această detectare se face manual sau, de preferat, automat.

În teza de doctorat se pune accentul pe metodele bazate pe caracteristici, în special acelea care utilizează puncte, și pe extragerea automată a acestora, utilizând algoritmul de detectare automată SIFT.

- Etapa următoare este asocierea acestor caracteristici prin care se stabilește corespondența caracteristicilor pe baza unor măsuri de similaritate.
- Pasul următor constă în estimarea parametrilor modelului de transformare utilizând metode de estimare robuste.
- În final imaginea captată este transformată și reșantionată astfel încât să fie optim aliniată cu imaginea de referință.



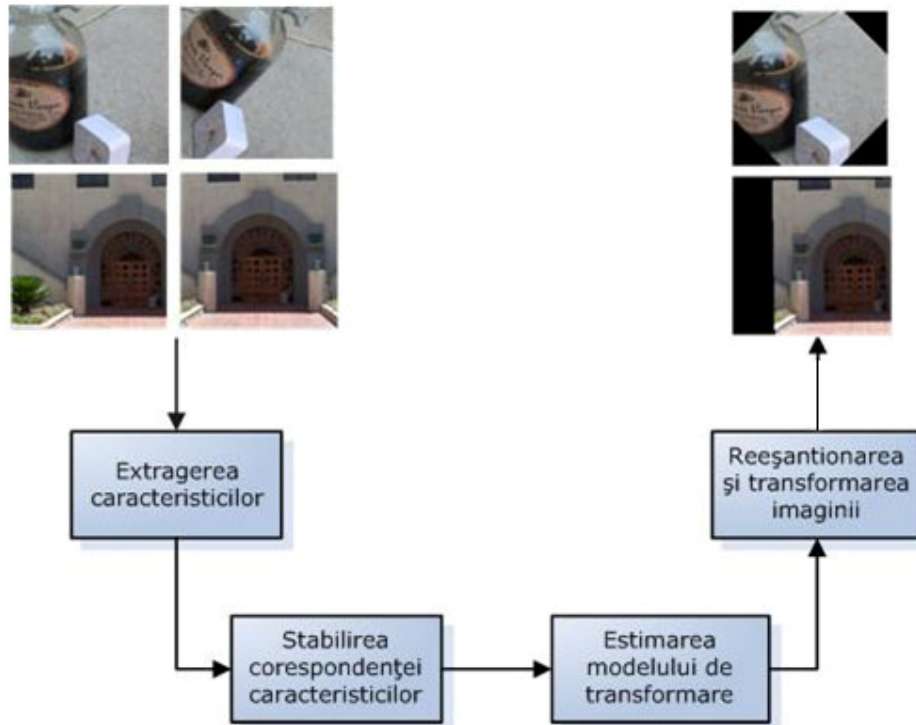


Fig. 1.2. Etapele parcurse în realizarea unei metode de aliniere a imaginilor. Imaginile din coloana 2 sunt supuse procesului de aliniere. Scopul este asocierea acestora cu imaginile din coloana 1. La ieșire se poate observa rezultatul metodei de aliniere.

Majoritatea algoritmilor folosiți în inteligența artificială includ o etapă de estimare robustă, în care parametrii modelului sunt calculați dintr-un set de date care conțin o cantitate considerabilă de date neconforme. Prin termenul de *estimare robustă* se înțeleg acele tehnici care sunt robuste în raport cu prezența erorilor grosiere din date. În acest context, erorile grosiere sunt definite ca și date care nu se potrivesc cu modelul stohastic al estimării parametrilor.

Soluțiile robuste ocupă un loc important în vederea artificială și preocuparea pentru cercetarea lor s-a dezvoltat rapid în ultimul deceniu [1].

**Teza este legată de studiul metodelor de aliniere a imaginilor utilizând metode robuste. Se studiază în special utilitatea tehnicilor neparametrice de estimare a densității de probabilitate folosind funcții nucleu în alinierea imaginilor. Explorarea acestei direcții în alinierea imaginilor a fost de la bun început soluția care a stat la baza temei de doctorat și rămâne actuală și la finalizarea ei, așa cum o atestă numărul de articole publicate.**

Așa cum se cunoaște din literatura de specialitate, o bine cunoscută tehnică, des utilizată în vederea artificială, mai ales pentru alinierea imaginilor, este metoda RANSAC [2]. Această metodă lucrează în spațiul coordonatelor, dar soluția sa nu ia în considerare întotdeauna toată informația disponibilă, de aceea precizia acestei metode nu este maximizată. Adoptând principiul generării de soluții parțiale pe baza unor seturi de date minimale folosit de RANSAC, **teza de doctorat propune o**

**abordare diferită a determinării consensului soluțiilor, bazată pe maximizarea densității de probabilitate și algoritmul mean shift, folosit până acum mai ales în segmentarea imaginilor [3].**

În ciuda numărului mare de studii și cercetări întreprinse pentru a construi algoritmi cât mai eficienți în punerea în corespondență a imaginilor, există totuși obstacole care trebuie depășite și întrebări care așteaptă răspuns.

## 1.2. Actualitatea temei

Metodele de aliniere a imaginilor sunt necesare pentru:

- integrarea informației luată de la diferiți senzori;
- găsirea schimbărilor care au loc în imagini realizate la momente diferite de timp sau sub anumite condiții;
- deducția informației 3D din imagini unde sau camera sau obiectele din scenă au fost mutate;
- pentru recunoașterea formelor obiectului.

În general, metodele de aliniere sunt supuse erorilor care provin fie de la măsurările eronate sau de la zgomotul introdus de senzori. Acest fapt poate conduce la degradarea performanțelor metodelor. O moderare a acestora se poate realiza prin utilizarea de metode robuste. O analiză detaliată a soluțiilor robuste o întâlnim în [1]. Explorarea acestor soluții în estimarea modelului parametrilor și utilizarea lor în stabilirea robustă a punerii în corespondență a imaginilor reprezintă o temă de actualitate așa cum o atestă numărul mare de articole ce continuă să fie publicate în literatură pe acest subiect.

Domeniile de aplicabilitate ale metodelor de punere în corespondență a imaginilor, conform [4], se împart în trei domenii mari de cercetare:

1. Vederea artificială și recunoașterea de modele. Metodele de punere în corespondență se pot utiliza în segmentarea imaginilor, recunoașterea de obiecte, reconstrucția de forme, urmărirea mișcării, asocierea stereo și recunoașterea caracterelor.
2. Analiza imaginilor medicale. Metodele de punere în corespondență se pot utiliza în diagnosticare, cum ar fi detecția tumorii și localizarea bolii, și în cercetare biomedicală.
3. Prelucrarea datelor în teledetecție. Metodele de punere în corespondență se pot utiliza în aplicații militare, civile, geologice, oceanografice, monitorizare urbană, monitorizare a resurselor naturale.

**Teza de doctorat își propune să studieze în ce măsură și în ce tip de aplicații, metodele de estimare neparametrice pot fi avantajoase în raport cu metodele robuste utilizate curent în alinierea imaginilor. În plus, se prezintă un studiu al eficacității celor mai utilizate estimatoare robuste folosite în tehnicile de punere în corespondență a imaginilor. Se compară rezultatele acestor metode cu cele obținute prin metoda de aliniere a imaginilor utilizând tehnica de estimare robustă mean shift și se identifică caracteristicile problemelor de aliniere a imaginilor. Se determină experimental că metodele de aliniere a imaginilor propuse în cadrul tezei de doctorat sunt mai exacte.**

De asemenea, se identifică aplicațiile ce pot beneficia de avantajele noii metode. Între acestea, se numără cea a localizării senzorilor din rețele fără fir, care a inspirat și motivat în bună parte lucrarea prezentă. Este

**adresată și problema selecției asocierilor de puncte valide, pentru care se propune o metodă eficientă, necesitând un efort de calcul redus.**

### **1.3. Structura tezei de doctorat**

În **capitolul 1 „Motivație”** se prezintă o privire de ansamblu a domeniului de aliniere a imaginilor cu toate componentele sale. Sunt introduse pe scurt etapele acestei tehnici și problemele care le ridică fiecare etapă. Tot în cadrul acestui capitol sunt prezentate motivele care au condus la tratarea acestei teme de cercetare ce face obiectivul tezei de doctorat.

În **capitolul 2 „Tehnici de aliniere a imaginilor”** se prezintă stadiul actual al cercetării în problema alinierii imaginilor bazată pe tehnici robuste. Se prezintă o clasificare a metodelor de punere în corespondență existente și etapele care duc la alinierea imaginilor. Se menționează, de asemenea, tipurile de corespondență și problemele pe care acestea le ridică, împreună cu caracteristicile lor. De asemenea este prezentat domeniul de aplicație pentru care se aplică metodele de aliniere a imaginilor propuse în teza de doctorat.

**Capitolul 3 „Fundamentele teoretice privind tehnicile de aliniere a imaginilor”** este o analiză teoretică a metodelor de aliniere și a metodelor de estimare robuste utilizate în vederea artificială. În acest capitol se prezintă:

- Etapele unei metode de aliniere, fiecare etapă fiind explicată pe larg.
- Tipurile de transformări geometrice utilizate în alinierea imaginilor, cu exemplele de rigoare.
- Tehnicile robuste de estimare utilizate la alinierea imaginilor.
- Se realizează o prezentare a metodelor de estimare robuste existente în literatura de specialitate și care sunt utilizate la alinierea imaginilor.

În **capitolul 4 „Metode propuse de aliniere a imaginilor bazate pe estimarea neparametrică a densității de probabilitate”** sunt trecute în revistă metodele propuse de doctorand privind alinierea imaginilor. Din necesitatea unei estimări a parametrilor modelului, în prezența unui număr mare de date neconforme, am elaborat un algoritm de aliniere care utilizează ca metodă robustă de estimare a parametrilor modelului mean shift-ul. Acest capitol descrie:

- Propunerea și implementarea unei metode de aliniere a imaginilor, bazată pe estimare neparametrică de densitate și algoritmul mean shift.
- Propunerea și compararea a două versiuni de estimare a parametrilor transformării de similaritate prin algoritmul mean shift: secvențial și simultan;
- Propunerea unei metode de asociere a punctelor mediane pentru selecția punctelor corespondente și eliminarea corespondențelor incompatibile, folosind similaritatea culorilor și un prag de decizie;
- Utilizarea algoritmului de aliniere a imaginilor bazat pe estimarea neparametrică, în rețele de senzori video fără fir.

În **capitolul 5 „Studiu experimental”** se testează eficacitatea metodelor propuse și robustețea acestora și se prezintă rezultatele obținute. Se studiază posibilitatea utilizării algoritmului propus în cazul localizării senzorilor într-o rețea de senzori video fără fir. Se urmărește:

- eliminarea eficientă a posibilelor necorespondențe, nepotriviri dintre perechile de puncte;
- creșterea acurateței estimării;
- reducerea timpului de calcul;
- estimarea acurateței și eficiența timpului metodei automate în comparație cu metoda manuală;
- testarea algoritmului în localizarea unei rețele de senzori video fără fir.
- studiu comparativ între metoda de aliniere bazată pe estimare non-parametrică de densitate de probabilitate și metodele de referință RANSAC și MLESAC.

Toate aceste experimente au fost realizate cu ajutorul unei aplicații de aliniere a imaginilor, implementată de doctorand în mediul de programare C++.

**Capitolul 6 „Concluzii generale și contribuții”** conține principalele concluzii rezultate din teza de doctorat și contribuțiile originale ale doctorandului, raportate în această teză.

### 1.4. Lucrări publicate de doctorand

Soluțiile și metodele noi dezvoltate în cadrul acestei teze au fost publicate în următoarele articole:

1. D.Fuiorea, V.Gui, D.Pescaru, C.Toma, "Kernel based image registration versus MLESAC: a comparative study ", Proceedings of the 5th International Symposium on Applied Computational Intelligence and Informatics, Timisoara, mai 2009, trimis spre publicare, cotate ISI Proceedings și IEEE Explore.
2. D. Fuiorea, V. Gui, D. Pescaru, P. Paraschiv, I. Codruta, D. Curiac, C. Volosencu, "Video-based Wireless Sensor Networks Localization Technique Based on Image Registration and SIFT Algorithm", World Scientific and Engineering Academy and Society (WSEAS), Vol. 7, Issue 7, 2008, pp. 990-999, ISSN:1109-2750 (SCOPUS).
3. D. Fuiorea, V. Gui, D. Pescaru, P. Paraschiv, I. Codruta, D. Curiac, C. Volosencu, "Sensor node localization using SIFT algorithm", Proceedings of the 9th WSEAS Int. Conf. on AUTOMATION and INFORMATION (ICAI'08), București, Romania, June 24-26, 2008, pp.436-441, ISBN: 978-960-6766-77-0, ISSN: 1790-5117, cotate ISI Proceedings și IEEE Explore.
4. D. Fuiorea, V. Gui, D. Pescaru, C. Toma, "Comparative study on RANSAC and Meanshift algorithm", Buletinul Științific al Universității "Politehnica" Timișoara, Transactions on Electronics and Telecommunications, ISSN 1583-3380, sept. 2008.
5. D.Fuiorea, V.Gui, F.Alexa, T.Corneliu, "A New Point Matching Method for Image Registration", Proceedings of the 6th WSEAS Int. Conf. on COMPUTATIONAL INTELLIGENCE, MAN-MACHINE SYSTEMS and CYBERNETICS (CIMMACS '07), Tenerife, Canary Islands, Spania, dec. 14-16, 2007, pp.135-139, ISSN 1790-5117, cotate ISI Proceedings și IEEE Explore (INSPEC).
6. D. Fuiorea, "A New Point Matching Method for Registration-Based Sensor Localization", Sesiunea de comunicării științifice "Doctor ETc 2007", 20 sept 2007, ISBN 978-973-625-494-9.
7. D. Fuiorea, "A New Point Matching Method for Image Registration Using Pixel Color Information", Buletinul Științific al Universității "Politehnica" Timișoara, pg.29-33, 2007.
8. D. Fuiorea, D. Pescaru, V. Gui, C. Toma, "Using Registration Algorithms for Wireless Sensor Network Node Localization", Proceedings of the 4th International Symposium on Applied Computational Intelligence and Informatics, mai 17-18, 2007, pp. 209-214, ISBN 1-4244-1234-X, cotate ISI Proceedings și IEEE Explore (INSPEC).
9. D. Fuiorea, D. Pescaru, V. Gui, C. Toma, "Feature Based 2D Image Registration Using Mean Shift Parameter Estimation", Scientific Bulletin of the "Politehnica" University, Timișoara, Tom 51(65), Fasc. 2, 2006, pp. 77-80, ISSN 1583-3380, Simpoziomul Internațional de Electronică și Telecomunicații, Universitatea "Politehnica" Timișoara, sept. 21-23, 2006.

Activitatea doctorală de cercetare a fost valorificată și prin câștigarea unui grant CNCIS TD nr. 546/2007, intitulat Tehnici robuste de punere în corespondență a imaginilor. Proiectul a fost câștigat de doctorand în anul 2007 și s-a extins pe perioada unui an.

## 2. TEHNICI DE ALINIERE A IMAGINILOR

Acest capitol este structurat pe opt subcapitole și urmărește prezentarea stadiului actual al metodelor de aliniere a imaginilor. În primul paragraf sunt prezentate și discutate metodele de aliniere a imaginilor folosite în vederea artificială și clasificarea acestora pe baza unor factori care determină tipul corespondenței. În paragrafele 2.2-2.4 se prezintă etapele de parcurgere ale unei metode de aliniere a imaginilor. În paragraful 2.5 se tratează categoriile de corespondență și problemele care apar. Caracteristicile metodelor de aliniere a imaginilor sunt prezentate în paragraful 2.6. În paragraful următor sunt prezentate domeniul de aplicabilitate al metodelor de aliniere a imaginilor. Paragraful final face o trecere în revistă a concluziilor rezultate din studiul prezentat în acest capitol.

### 2.1. Considerații generale

Tehnica de aliniere a imaginilor prezintă un interes deosebit în teledetectie, vederea artificială, analiza imaginilor medicale, supraveghere, reconstrucția scenei, navigarea automată, procesarea imaginilor de la distanță, etc. În general, utilizarea acestei metode poate fi împărțită în patru grupe în funcție de metoda de achiziție existentă [5]. Putem avea *analiza mai multor câmpuri vizuale*, și anume imagini ale aceleiași scene fiind captate din diferite unghiuri. Scopul este de a obține o vedere 2D sau 3D a scenei. Mai putem avea imagini ale aceleiași scene, dar care sunt captate la momente diferite de timp. Scopul acestei *analize multitemporale* este de a evalua schimbările care au loc în scenă și care apar între două imagini achiziționate consecutiv. O altă categorie este *analiza multimodală* care este reprezentată de imagini ale aceleiași scene dar achiziționate de senzori diferiți. Scopul este de a integra informația obținută de la surse diferite pentru a obține o reprezentare a scenei complexă și detaliată. Tot cu ajutorul metodelor de aliniere se poate realiza alinierea dintre imaginile unei scene și imaginile unui model. Modelul poate fi o reprezentare pe calculator a scenei, de exemplu o hartă. Scopul este de a compara imaginile achiziționate cu modelul existent. Printre aplicațiile care folosesc metodele de aliniere a imaginilor amintim de fuziunea imaginilor care provin de la surse multiple, analiza schimbărilor care au loc în imagini [6] și asamblarea de imaginilor pentru vedere panoramică. Alte aplicații curente sunt: teledetectia [7], analiza imaginilor [4], [8], [9], stereo viziunea [10], [11], cartografia automată [12] și analiza mișcării [13], [14]. Recuperarea caracteristicilor 3D ale unei scene [15] în recunoașterea de forme este o altă aplicație relevantă. Roberts [16] a descris localizarea și recunoașterea obiectelor poliedrice aliniind proiecțiile muchiilor modelelor cu muchiile imaginii. Alinierea imaginilor în teledetectie este descrisă în Anuta [17], [18], Barnea și Silverman [19], Leese [20] și Pratt [21].

Utilizarea alinierii imaginilor în vederea artificială a fost pionerată de Mori [22], Levine [23] și Nevatia [24]. Această tehnică a apărut în domeniul imaginilor medicale atunci când au apărut și imaginile digitale multiple de la diferiți senzori, care măsurau anatomia și funcționarea organismului [25], [26], [27]. Studiul acestor metode de aliniere a imaginilor în domeniul imaginilor medicale a fost

prezentat de Hajnal și colaboratorii în [28]. Colecții de lucrări privind acest domeniu și utilizarea metodelor de aliniere au fost editate de Pernus și colaboratorii [29] și Pluim și Fitzpatrick [30]. Altă colecție de lucrări mai întâlnim la Goshtasby și LeMoigne [31], lucrări care cuprind metodologii generale ale corespondenței imaginilor. Subiecte legate de alinierea imaginilor nerigidă se întâlnesc în colecția de lucrări a lui Goshtasby și colaboratorii [32].

Studii și clasificări ale metodelor de aliniere se pot găsi în Gerlot [33], Brown [4], Van den Elsen [34], Mauer [35], Maintz [36], Lester [37], Pluim [38] și Zitova [5].

O clasificare a metodelor de aliniere a imaginilor [39] se poate realiza pe baza unei serii de factori, care determină tipul corespondenței.

1. **Dimensionalitatea.** În domeniul vederii artificiale întâlnim alinierea *bidimensională* folosită pentru imaginile 2D și alinierea *tridimensională* folosită pentru imaginile 3D. Un studiu al acestor metode îl întâlnim Goshtasby [40].
2. **Domeniul transformatei.** Metoda de aliniere mai poate fi *globală* sau *locală*, în funcție de modul de operare a transformării utilizate, pe toată imaginea sau parțial.
3. **Tipul transformatei.** În funcție de tipul transformatei, alinierea poate fi *rigidă* sau *nerigidă*.
4. **Densitatea caracteristicilor.** Transformarea poate fi cu *interpolare* sau *aproximativă*. În primul caz caracteristicile obiectelor dintr-o imagine sunt transformate exact în caracteristicile obiectului din a doua imagine. În al doilea caz apare o eroare de asociere, care afectează caracteristicile obiectului.
5. **Măsurarea calității corespondenței.** Diferite măsuri de similaritate sunt aplicate, folosind fie date, fie caracteristici care derivă din date.
6. **Metode de determinare a parametrilor.** În metodele *directe* parametrii transformării sunt calculați direct din date, iar în metodele de *căutare orientată* sunt găsiți prin cercetarea spațiului de valori posibile în procesul de optimizare.
7. **Obiectele din scenă.** Alinierea poate fi *intrasubiect* dacă mai multe imagini conțin același subiect, sau mai poate fi *intersubiect* dacă subiecții sunt diferiți.
8. **Tipul sursei.** Procesele de aliniere, în funcție de tipul de surse, se împart în două categorii: proces de aliniere *intramodal* și proces de aliniere *intermodal*. Procesul de aliniere *intramodal* se referă la imaginile care provin de la aceeași sursă, iar procesul de aliniere *intermodal* se referă la imaginile care provin din surse diferite. În medicină se utilizează cel mai des alinierea intermodală, și anume alinierea imaginilor de la CT/MRI (Computing Tomography/Magnetic Resonance Imaging) pentru scanarea creierului sau PET/CT (Positron Emission Tomography/Computed Tomography) pentru scanarea întregului corp, scopul final fiind localizarea tumorii, sau corespondența imaginilor CT cu contrast puternic cu imaginile CT cu contrast slab pentru segmentarea unor părți specifice ale anatomiei organismului, etc. Alinierea intramodală se utilizează în special în cazul imaginilor captate la momente diferite cu același echipament.
9. **Tipul informației.** Informațiile care controlează procesul de punere în corespondență pot fi date brute (pixels sau voxels), caracteristici extrase (muchii sau regiuni convexe), markeri (pixeli marcați manual).

10. **Sursa de caracteristici.** Algoritmii de punere în corespondență pot folosi caracteristici *intrinsece* sau *extrinsece*. Cele intrinsece sunt cele prezente în informație, implicit (marginii, forme geometrice ale obiectelor, suprafețe) sau explicit (nivelurile de gri ale imaginii). Cele extrinsece sunt adăugate informației din exterior, cum ar fi marcaje fizice, ex. buline colorate etc.
11. **Seleția caracteristicilor.** Putem întâlni corespondență manuală și corespondență automată [41].
12. **Corespondența caracteristicilor.** Se poate realiza corespondența punctelor caracteristice din imagini sau corespondența regiunilor din imagini. În funcție de acestea, corespondența imaginilor poate fi împărțită în corespondență bazată pe caracteristici și corespondență bazată pe regiuni.
- Datorită diversității imaginilor care urmează a fi aliniate și datorită diferitelor tipuri de distorsiuni, este imposibil de a crea o metodă aplicabilă pentru toate aceste probleme. Fiecare metodă de aliniere trebuie să ia în calcul o serie de factori, ca deformările geometrice ale imaginilor, transformările radiometrice, zgomotul etc. Cu toate acestea, etapele de prelucrare sunt asemănătoare [4], [5]. Majoritatea metodelor de aliniere constau din patru pași prezentați în diagrama de mai jos.

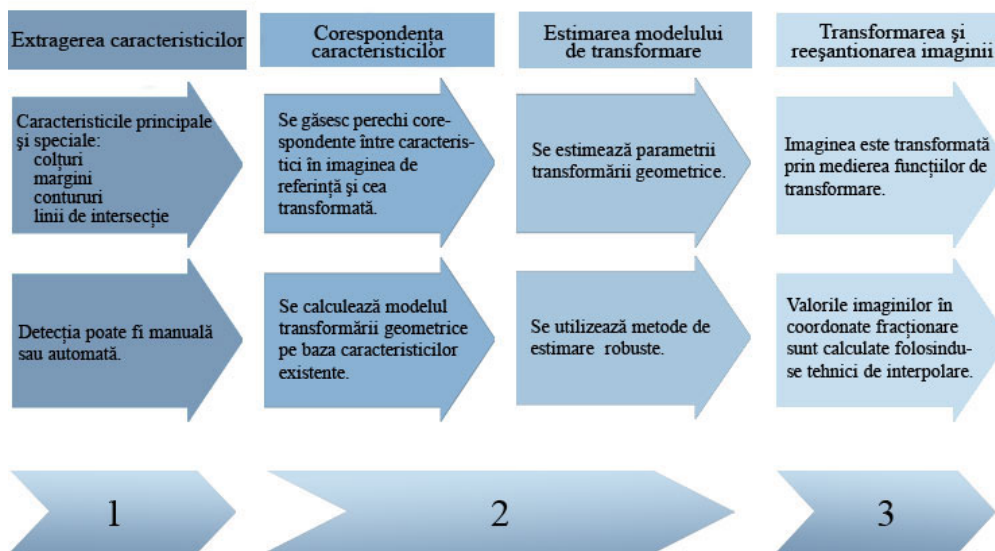


Fig. 2.1. Parcurgerea metodelor de aliniere a imaginilor

## 2.2. Extragerea caracteristicilor

Primul pas în determinarea alinierea a două imagini este de a se decide asupra spațiului caracteristicilor folosit pentru asocierea celor două imagini. Spațiul caracteristicilor poate fi reprezentat de imaginea în sine, muchii, contururi, suprafețe, colțuri, intersecții de linii, caracteristici statistice ca momente invariante sau centroizi, descriptori. Prin alegerea acestor caracteristici, are loc reducerea efectelor produse de zgomotul de senzor, sau a altor distorsiuni, având ca rezultat forțarea corespondenței pentru a optimiza și reduce spațiul de căutare.



O clasificare a spațiului caracteristicilor o întâlnim în [4] și face referire la tipurile de caracteristici utilizate la alinierea imaginilor (figura 2.15). Alegerea spațiului caracteristicilor poate influența procesul de aliniere. Spațiul caracteristicilor prin caracteristicile alese, poate reduce zgomotul sau alte distorsiuni, cum ar fi iluminarea și condițiile atmosferice; poate stabili care proprietăți ale imaginii vor fi potrivite sau poate reduce timpul de calcul prin reducerea căutării spațiului.

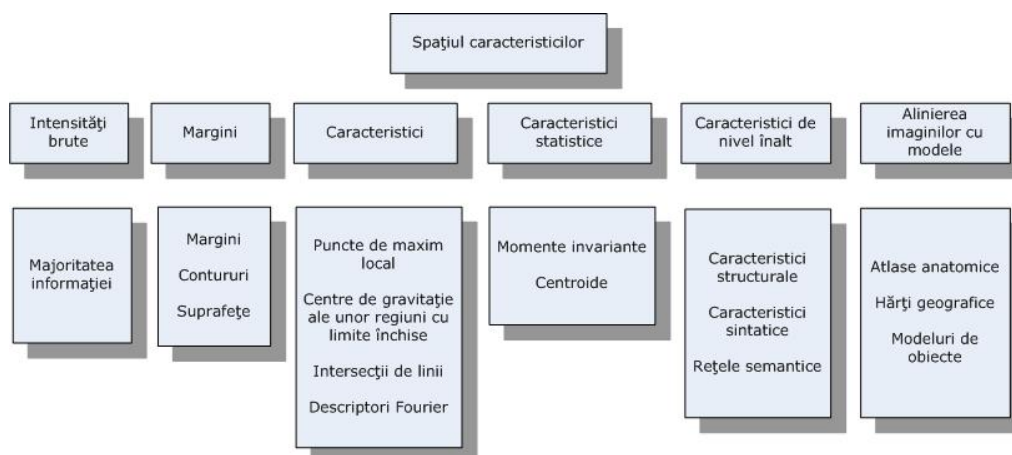


Fig. 2.2. Spațiul caracteristicilor utilizat în metodele de aliniere a imaginilor

Muchiile reprezintă majoritatea structurii intrinsece a unei imagini. Bay [42] extrage muchiile și realizează asocierea segmentelor de linii dintre două imagini. În [43] se propune asocierea curbelor (muchiiilor) prin „super curbe”. Aceste curbe sunt introduse în [44] în procesul de aliniere și se obține o soluție globală prin minimizarea erorii de corespondență pentru toate punctele și curbele. În [45] se propune o metodă de aliniere bazată pe muchii și colțuri. Muchiile sunt detectate și împărțite în segmente, iar colțurile corespondente se determină prin utilizarea informației globale în aflarea parametrilor transformării. Rezultatele experimentale [45] au demonstrat eficacitatea metodei pentru muchiile neideale, algoritmul fiind folositor pentru alinierea multimodală a imaginilor.

Alinierea robustă a imaginilor se poate realiza și cu ajutorul centroizilor [46]. Aceștia sunt considerați [4] a fi mai ușor de calculat, intoleranți la zgomot și pot fi utilizați ca și puncte de control în metodele de corespondență bazate pe puncte. O combinație între punctele caracteristice și trăsăturile caracteristice le întâlnim în [47] unde se propune un algoritm de extragere automată a acestor caracteristici pentru realizarea corespondenței.

SIFT este un algoritm propus de Lowe [48] pentru extragerea și descrierea caracteristicilor locale a imaginilor, care și-a demonstrat utilitatea în numeroase aplicații de punere în corespondență, precum și în recunoașterea formelor. Aceste caracteristici locale sunt invariante la rotația și scalarea imaginilor. Ele sunt, de asemenea, robuste schimbărilor de iluminare, zgomot, schimbări minore ale punctului de observație. În plus, caracteristicile locale date de SIFT sunt distinctive, relativ ușor de extras și permit identificarea corectă a obiectelor cu o probabilitate ridicată. Algoritmul se bazează pe extragerea punctelor de extrem în domeniul spațiu-scară (scale-space) [49], [50]. În prima etapă se realizează convoluția dintre

imagine și filtrele Gaussiene la diferite scări. Apoi calculează diferența dintre imaginile cu rezoluții vecine (DoG), în care se identifică punctele cheie ca maxime/minime locale.

Algoritmul SIFT generează și un vector descriptor al orientării locale estimate la opt scări de analiză. Descriptorul este calculat ca un set de histogramme de orientare (4 x 4) utilizând vecinătăți de pixeli (Fig. 2.8). Se obține un vector de caracteristici SIFT cu  $4 \times 4 \times 8 = 128$  elemente. Acest vector este normalizat pentru a mări invarianța la iluminare.

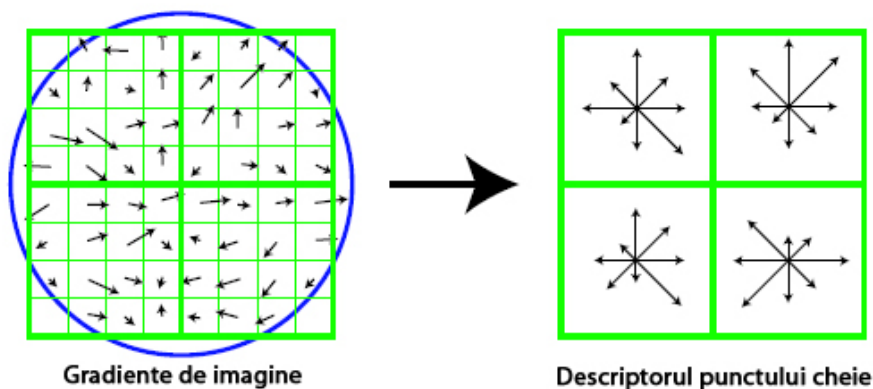


Fig. 2.3.Descriptorul local generat de SIFT

O bună parte din experimentele descrise în teză au la bază puncte caracteristice generate prin algoritmul SIFT.

### 2.3. Corespondența caracteristicilor

Procesul de stabilire a corespondenței caracteristicilor are rolul de a găsi perechi corespondente între caracteristicile din imaginea de referință și caracteristicile din cea transformată, pe baza informației luată de la fiecare vecinătate a punctului caracteristic. Procesul mai poate fi privit ca și un pas de preselecție. Principala responsabilitate este să detecteze nepotrivirile dintre perechile de caracteristici. Corespondența actuală poate să nu fie disponibilă fără estimarea transformării geometrice care aliniază cele două imagini și validează sau invalidează corespondența. Prin urmare, procesul de asociere a caracteristicilor de obicei este combinat cu estimarea parametrilor transformării geometrice.

Procesul este complicat și datorită faptului că există puncte caracteristice care prezente numai într-una din cele două imagini pentru care se urmărește punerea în corespondență. Acele puncte nu au corespondent real și acționează ca date aberante în procesul de estimare a transformării de aliniere a imaginilor. Hsieh [51] a adresat problema pentru caracteristici bazate pe transformarea wavelet. Corespondența dintre caracteristici a fost stabilită prin intercorelație. Un algoritm pentru stabilirea corespondenței punctelor în prezența unei transformări geometrice non-rigide este descris în [52] de Chui. Algoritmul se bazează pe corespondența unu

la unu, dar se poate extinde în cazul potrivii mai multor caracteristici [53]. În [54] sunt propuși doi algoritmi pentru a rezolva problemele modelului de asociere a punctelor folosind căutarea ramificată cu salt. O altă abordare a fost recent propusă în Belongie și colaboratorii [55] și adoptă o altă strategie. Un nou descriptor de formă, numit "shape context", este definit pentru stabilirea corespondenței. Pentru fiecare punct ales, se trasează linii de conexiune cu celelalte puncte. Distribuția lungimii și orientării pentru toate liniile este utilizată pentru stabilirea corespondenței fiecărui punct. Distanța Housdorff a fost utilizată în pentru stabilirea corespondențelor în [56] și [57]. În [58] problema potrivirii punctelor pentru estimarea mișcării obiectelor s-a adresat ca o problemă de clasificare. Fiecare punct din imaginea de antrenament este o clasă. În general, metoda oferă potriviri puține și are o rată a erorilor mai ridicată decât SIFT [59].

Utilizarea metodelor de grupare (în engleză clustering) în problema corespondenței a fost introdusă pentru prima dată de Stockman și colaboratorii [60]. Metodele de acumulare au fost utilizate în [61], [62], [63] și reprezintă o generalizare a transformatei Hough [64], [65]. În [66] Goshtasby și colaboratorii prezintă tehnici de acumulare de puncte care încearcă asocierea punctelor conectate prin linii de segment. Informații mai multe despre metodele de acumulare și aplicațiile acestora se mai găsesc în Dhome [67], Linnainmaa [68], Silberberg [69] și Thompson [70].

## **2.4. Alegerea modelului de transformare și estimarea parametrilor modelului**

După realizarea corespondenței caracteristicilor, se construiește o funcție de transformare care aliniază optimal punctele caracteristice și cele două imagini implicate. Rolul ei este de a transforma imaginea țintă peste imaginea de referință. În realizarea procesului, trebuie ales de tipul funcției de transformare geometrică și metoda de estimare a parametrilor transformării. De exemplu, care se cunoaște că imaginile aparțin unei scene plate, o transformare afină este suficientă pentru a stabili corespondența imaginilor. Dacă nu există informație despre diferențele geometrice dintre imagini, se poate folosi o transformare care se va adapta la diferențele geometrice locale dintre imagini. Tipul funcției de transformare trebuie să corespundă așa presupusei deformări geometrice a imaginii captată, deformare dată de procesul de achiziție al imaginii (Fig. 2.4).

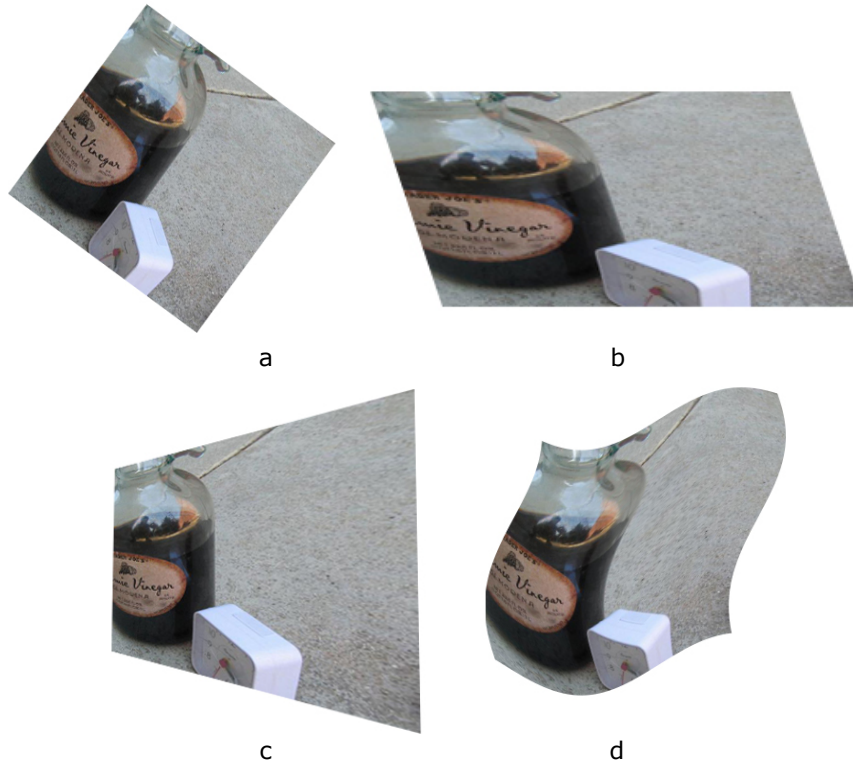


Fig. 2.4. Exemple de diferite funcții de asociere: a) transformarea de similaritate; b) transformarea afină; c) perspectiva de proiecție; d) transformarea elastică.

Frecvent se folosește transformarea de similaritate sau transformarea sistemului de coordonate carteziene, care reprezintă diferența globală de translație, rotație și scalare dintre două imagini. Transformarea este definită de

$$\begin{aligned} X &= S(x \cos \theta + y \sin \theta) + T_x \\ Y &= S(x \sin \theta + y \cos \theta) + T_y \end{aligned} \quad (2.1)$$

unde  $T_x, T_y, S, \theta$  reprezintă parametrii transformării de similaritate, și anume translația pe verticală și orizontală, scară și unghi. Acești patru parametri ai transformării pot fi determinați pe baza combinării dintre două seturi de puncte (4 puncte) din imaginile supuse procesului de punere în corespondență.

Data fiind incertitudinea rezultatelor procesului de asociere a caracteristicilor, existentă în problemele de aliniere a imaginilor abordate în teză, pentru a se găsi parametrii transformării geometrice trebuie să fie folosite metode de estimare robuste [71]. Robuștețea este esențială în corespondență deoarece numărul de corespondențe de puncte false poate să depășească numărul celor corecte. În literatură au fost propuse multe metode robuste de corespondență bazate pe caracteristici [4], [72], [73], [5]. Datorită importanței acestora, în contextul tezei, în capitolul 3 este inclusă o prezentare teoretică a metodelor de estimare robuste.

## 2.5. Transformarea și reeșantionarea imaginii

Transformarea

$$\begin{aligned} X &= f(x, y) \\ Y &= g(x, y) \end{aligned} \quad (2.2)$$

face legătura coordonatelor punctelor  $(x,y)$  din imaginea de referință cu coordonatele  $(X,Y)$  ale punctelor corespondente din imaginea captată. De observat că  $(x,y)$  sunt coordonate reprezentate pe numere întregi, în timp ce coordonatele  $(X,Y)$  sunt numere reale, găsite prin calculul transformării. Din moment ce numai coordonate întregi sunt disponibile în imaginea captată, intensitatea punctului  $(X,Y)$  trebuie să fie estimat din intensitățile unui număr mic de pixeli vecini. Acest lucru îl realizează o metodă de **reeșantionare cu interpolare** [74]. Cea mai simplă metodă este de a atribui valoarea (intensității sau culorii) celui mai apropiat pixel din imaginea originală față de imaginea căutată. Această metodă se numește *metoda celui mai apropiat vecin*. Folosește mai puțin timp de calcul și evită modificarea valorii originale de strălucire. Nu asigură însă întotdeauna o precizie suficientă. Rezultate progresiv mai bune se obțin prin interpolare biliniară, respectiv prin interpolare bicubică [75]. Tehnici de interpolare a imaginilor, prezentate sintetic în Zitova [5] pot fi bazate pe funcții spline pătratice [76], [77], funcții spline pătratice cubice Bsplines [78], funcții B-spline de ordin superior [79], funcții spline cardinale Catmull-Rom [80], [81], gaussiene [82] sau funcții sinus cardinal trunchiate [83].

## 2.6. Categoriile de alinieri și problemele care apar

Punerea în corespondență a punctelor caracteristice, și prin urmare a alinierii imaginilor, este îngreunată adesea de prezența unor distorsiuni ce afectează transformarea propriu-zisă ce trebuie estimată și inversată:

- Zgomotul introdus de senzor;
- Schimbările de perspectivă care au loc din cauza senzorului sau perturbațiile platformei;
- Schimbările obiectului cum ar fi mișcarea, deformarea sau creșterile;
- Schimbările de iluminare și atmosferice;
- Senzori diferiți.

Conform Gottesfeld Brown [4], există patru categorii de corespondență. Cele patru categorii sunt prezentate succint mai jos. Transformările sunt folosite pentru a alinia o pereche de imagini fie local, fie global. O *transformare globală* este dată de către o singură ecuație care transformă întreaga imagine. *Transformarea locală* transformă imaginea diferit, depinzând de locația spațială.

### 2.6.1. Aliniere prin tehnici multimodale

Problema care se prezintă aici este realizarea alinierii imaginilor care provin din aceeași scenă, achiziționate cu diferiți senzori. Această metodă se utilizează în

general pentru integrarea informației cu scopul de a îmbunătăți segmentarea și clasificarea pixelilor.

Un exemplu de aliniere prin tehnici multimodale este integrarea informației structurale care provine de la senzori CT și MRI, sau care provine de la scanerile radio nucleice PET și SPECT, pentru evaluarea funcțiilor metabolice, și se folosește în domeniul analizei imaginilor biomedicale.

În [84] se descrie o nouă metodă de aliniere a imaginilor medicale, care provin de la diferiți senzori, folosind informația furnizată de histogramă și de expresia țesutului din diferite modalități. Simplitatea acestei metode devine potrivită pentru a putea fi implementată hardware. Un alt algoritm de aliniere multimodală este prezentat în [85], și anume algoritmul de corespondență HAMMER care demonstrează acuratețe ridicată în asocierea imaginilor medicale care provin de la senzori diferiți. Din cauza segmentării țesutului, care se dovedește a fi o metodă dificilă, nestabilă sau chiar imposibilă, metoda HAMMER este limitată prin utilizarea în aplicații mai generalizate. Datorită acestui lucru, se utilizează histograme de intensitate pentru a crea un nou tip de vectori atribuite pentru fiecare intensitate a unui punct. Acești vectori sunt invarianți, conțin informația spațială și realizează corespondența punctelor din imagini. Pe baza experimentelor efectuate, metoda propusă s-a dovedit a avea performanțe bune în realizarea alinierii dintre imaginile MR și CT.

În [86] se prezintă un algoritm pentru alinierea automată a imaginilor care prezintă transformări rigide. Acest algoritm utilizează o nouă metrică de similaritate pentru imaginile cu diferențe mari de contrast. Metoda mai folosește o segmentare la scară multiplă și minimizează metrica aleasă prin metoda Powell. În [87] se utilizează informația mutuală (MI) sau entropia relativă ca și criteriu de asociere. Metoda prezentată în lucrare folosește informația mutuală pentru a măsura dependența statistică sau informația de redundanță dintre intensitățile pixelilor din ambele imagini. Această informație mutuală se presupune a fi maximă atunci când imaginile sunt aliniate. Metoda este comparată cu alinierea stereotactică, iar rezultate arată că soluția este obținută automat și fără segmentare în prealabil sau extracție de caracteristici. În [88] se propune o nouă metodă de aliniere automată a imaginilor care utilizează combinația dintre metodele bazate pe caracteristici și cele bazate pe regiuni.

### 2.6.2. Alinierea modelelor din imagini

Alinierea modelelor din imagini ridică problema găsirii forme de referință într-o imagine. Aplicațiile tipice în care se folosește această aliniere sunt aplicații de recunoaștere sau localizare de forme, cum ar fi, localizarea poziției și a orientării pistelor, terminalelor, și a locurilor de parcare în aeroporturi. Alte exemple sunt recunoașterea caracterelor, verificarea semnăturii, corespondența amprentelor [89], [90], [91]. De asemenea, alinierea cu tipar se mai poate folosi și în cazul imaginilor biomedicale. Un exemplu îl prezintă Jain și Chen în [92] în care corespondența unei radiografii dentare cu o colecție dentară dezvăluie poziția și indexul fiecărui dinte din radiografie. Acest lucru ajută la realizarea identificării dinților atunci când se aliniază două radiograme. Pentru reprezentarea colecției dentare, în lucrare se propune un model Markov.

O metodă de aliniere a modelelor din imagini, utilizată larg în localizarea obiectelor 3D pentru manipularea robotizată este algoritmul ICP, propus de către Besl și McKay [93]. Scopul algoritmului este de a găsi o transformare rigidă  $T$  care

aliniază punctele din setul  $S$  cu un model geometric  $M$ . Algoritmul minimizează distanța medie pătratică dintre toate punctele din imaginea captată (scenă), dar și a punctelor cele mai apropiate din imaginea referință (model). Componenta principală a algoritmului calculează corespondențele dintre punctele cele mai apropiate și apoi minimizează eroarea medie pătratică dintre corespondențe. Aceste puncte calculate se utilizează apoi pentru a forma o matrice de rotație și de translație, care se va aplica la rândul ei peste toate punctele din imaginea captată. Acest proces se repetă de un număr de ori, devenind astfel un algoritm iterativ. Diferite versiuni ale algoritmului au fost propuse pentru creșterea performanțelor și a vitezei [94], [95], [96], [97], [98], [99], [100], [101], [102], [103], [104], [105], [106], [107], [108], [109], [110], [111], [112].

### 2.6.3. Alinierea imaginilor 3D

Alinierea imaginilor 3D realizează alinierea imaginilor luate din diferite unghiuri. Se utilizează la reconstrucția formei și a intensității. Metoda se caracterizează prin transformări locale, necesare pentru a corecta distorsiunile de perspectivă. Corespondența de poziție se utilizează în vederea artificială, în special în realizarea urmăririi (tracking) mișcării obiectelor, persoanelor, și stereo, rezolvând asocierea pentru recuperarea formei și a intensității din disparități. Camerele multiple sunt necesare pentru a acoperi o suprafață largă a mediului înconjurător pentru a monitoriza activitatea. Pentru a urmări obiecte sau persoane captate cu camere multiple, este necesar a stabili corespondența dintre acestea.

Conform [113], provocările impuse de o aliniere cât mai precisă, vin de la trei factori: prezența fundalului parazit, ocluziunea tiparului care urmează a fi suprapus și schimbările valorilor caracteristicilor de-a lungul întregii imagini. Pentru a adresa aceste probleme, în [113] se propune o metodă de estimare robustă prevăzută pe reprezentările obiectului modelului și ale imaginii captate, utilizând o estimată de densitate cu nucleu.

Detectarea obiectelor aflate în mișcare, dintr-o secvență video, se realizează de obicei prin alinierea perechilor de imagini și prin crearea imaginii diferență. Obiectele aflate în mișcare se pot detecta din imaginea diferență prin metode de acumulare [114], [115], [116], [117]. În [118] se descrie un sistem în timp real pentru detectarea mișcării obiectului utilizând o tehnică de punere în corespondență ce se bazează pe matricele de distribuție de probabilitate a  $N$  puncte. Detectarea mișcării obiectelor dintr-o secvență video se utilizează în aplicații de supraveghere, misiuni de recuperare, urmărire de rachete. Algoritmii tradiționali sunt creați pentru a urmări un singur obiect din câmpul vizual comun al camerei sau să detecteze multiple obiecte ale aceleiași scene, amândouă cazurile utilizând o singură cameră. În [119] este prezentat un sistem cu două camere folosite pentru a detecta mișcarea obiectului din scenă. Tot în această direcție, în [120] se utilizează pentru urmărirea mișcării obiectelor a unui sistem format din mai multe camere video necalibrate cu câmpuri vizuale care se suprapun. În acest scop, obiectele care se văd în camere se etichetează și se utilizează limite ale FOV (Field Of View) pentru fiecare cameră pentru a diferenția capturile ale aceluiași obiect, dar văzut de camere diferite, creându-se astfel un sistem de detectare a mișcării fără a se utiliza corespondența caracteristicilor. În [121], [122] se descriu sisteme de detectare a mișcării persoanelor utilizând camere multiple necalibrate.

Alinierea imaginilor 3D se mai utilizează în domeniul vederii binoculare având scopul de a găsi perechi de puncte corespondente pe baza cărora se pot



determina disparități. Acestea pot fi convertite la o structură 3D a scenei. Algoritmii de aliniere stereo bazați pe regiuni divizează imagine în subregiuni și aplică măsuri fotometrice de similaritate pentru toate regiunile [23]. Scopul algoritmului este de a returna o hartă a adâncimii, unde pentru fiecare pixel din schemă există o estimată de adâncime. Descrierea algoritmilor bazați pe regiuni utilizați în alinierea stereo se poate găsi în lucrările [23], [22], [123], [124], [125], [126]. În contextul stereo viziunii, metodele bazate pe caracteristici utilizează muchii și intersecții de muchii (colțuri, puncte). Din acest motiv, majoritatea metodelor bazate pe caracteristici implică algoritmi de extragere de colțuri. O cercetare amănunțită a acestor metode o realizează Lane și Thacker în [127], făcând referire la [128], [129], [130], [131], [132]. În comparație cu metodele bazate pe regiuni, metodele bazate pe muchii oferă o informație de adâncime mai precisă și mai stabilă. Acest lucru se datorează faptului că acești algoritmi nu utilizează un model planar bazat pe regiuni. În schimb modelează lumea cu ajutorul muchiilor, ceea ce îl face aproximativ invariant la schimbările de iluminare dintre pixelii aceleiași scene văzuți din unghiuri vizuale diferite.

#### 2.6.4. Alinierea temporală

Conform [4] alinierea temporală este definită ca o aliniere a imaginilor din aceeași scenă filmate la momente de timp diferite sau condiții diferite. Un exemplu de aliniere temporală utilizată în prelucrarea datelor în teledetecție, este monitorizarea naturii, supravegherea centralelor nucleare și monitorizarea dezvoltării orașelor. Lucrarea [133] prezintă un sistem de construire automată a unui model al activității utilizând camere multiple. Conform [134], alinierea temporală reprezintă alinierea a două sau mai multe cadre temporale ale unei secvențe video sau volumul 3D al datelor. În lucrarea citată, se propune utilizarea evenimentelor dinamice. Pentru S. Ravela [135], alinierea temporală se obține vizual prin utilizarea unei perechi de camere, combinând urmărirea caracteristicilor punctelor modelului și calibrarea slabă cameră-obiect. Tehnica de urmărire utilizată în această lucrare este relativ sensibilă la schimbările de iluminare.

O altă aplicabilitate a alinierii temporale se găsește în domeniul analizei imaginilor medicale. Xinyu Guo [136] propune utilizarea acestei corespondențe pentru imaginile de retină. Se precizează că imaginile de retină furnizează o bază pentru a se monitoriza evoluția unor afecțiuni. Alinierea imaginilor de retină este dificilă datorită faptului că intensitatea retinei și a structurii vasculare poate varia în timp. Datorită acestui lucru, fiecare vas este modelat ca și un arbore, purtând numele de Vessel Feature Tree (VFT). Se utilizează o funcție pentru a calcula similaritatea perechilor de vase pe baza VFT-urilor lor.

### 2.7. Caracteristicile metodelor de aliniere

Acuratețea unui algoritm de aliniere depinde în principal de doi factori: selecția spațiului caracteristicilor și de selecția măsurii de similaritate pentru a potrivi caracteristicile imaginii. Metodele de punere în corespondență pot fi definite folosind combinații diferite ale următoarelor componente:



- *Spațiul caracteristicilor* care este definit de informația extrasă din imaginile care vor fi folosite pentru realizarea corespondenței.
- *Spațiul de căutare*, care este o clasă de transformări capabilă să realizeze alinierea imaginilor. Căutarea continuă în conformitate cu strategia de căutare până când este găsită o transformare ce conduce la o soluție ale cărei măsurări similare sunt satisfăcătoare.
- *Strategia de căutare*, care decide cum să se aleagă următoarea transformare optimală pentru spațiul în cauză, în cazul tehnicilor iterative.
- *Măsura de similaritate*, care determină valorile de calitate relative ale fiecărei soluții.

În [4] sunt enumerate câteva exemple de măsuri de similaritate utilizate în alinierea imaginilor, dintre care amintim de cross-corelația, coeficientul de corelație, informația mutuală, diferența sumei pătratelor, corelația de fază. Intercorelația poate fi eficient implementată în domeniul transformatei Fourier [137], dar forma normalizată, preferată pentru aplicațiile de asociere a caracteristicilor, nu are o expresie simplă în domeniul frecvență și este calculată în de regulă domeniul spațial. Pentru procesul de aliniere al imaginilor, în domeniul frecvență, cea mai utilizată măsură de similaritate este corelația de fază, care realizează translația a două imagini [138]. Corelația de fază se bazează pe teorema de deplasare Fourier, care este ușor de implementat și necesită un timp de calcul scăzut datorită eficienței înalte a transformatei Fourier (FFT). Corelația de fază este robustă la interferențe și poate trata translația dintre două imagini numai pe axa  $x$ - $y$ .

Tot în [4] găsim o familie de metode bazate pe corelația valorilor intensităților din cele două imagini, care a fost propusă pentru alinierea imaginilor. Metodele bazate pe suma pătratelor diferențelor (SSD) calculează suma diferențelor absolute pe baza luminozității pixelului dintre imaginea de referință și imaginea captată. Această tehnică este foarte sensibilă la ocluzii. O altă metodă care nu face față la ocluzii este cea a lui Venot și colaboratorii [139] care utilizează suma numărului de schimbări de semn dintre pixelii corespondenți dintr-o pereche de imagini.

Extragerea corespondențelor corecte sau a eșantioanelor conforme modelului (în engleză inlier) din date care conțin date neconforme (în engleză outliers) este o problemă dificilă. Există două tipuri de erori în procedura de corespondență: locația greșită a caracteristicilor, și corespondențele false ale caracteristicilor. Locația greșită înseamnă că punctele caracteristice nu sunt localizate corect și se datorează în principal tehnicii de detecție a punctelor caracteristice. Corespondență falsă înseamnă stabilirea eronată a corespondențelor între punctele extrase. Erorile produse de corespondențe false pot afecta sever precizia și eficacitatea procesului de estimare a parametrilor transformării geometrice. O metodă robustă trebuie să determine corespondențele plauzibile simultan cu estimarea parametrilor modelului de transformare. Robustețea unui estimator este dependentă de alegerea funcției de cost optimizate. O funcție de cost robustă acordă ponderi însemnate eșantioanelor conforme și ponderi nesemnificative sau nule celor aberante. După ce estimatorul clasifică datele în date conforme și neconforme, soluțiile obținute pot fi post-procesate cu o metodă care nu este robustă.

În ultimii douăzeci de ani, s-au aplicat în problemele de vedere artificială o mulțime de tehnici robuste. Există abordări robuste pentru aliniere de suprafețe [140], estimarea mișcării [141], [142], geometria epipolară [143], [144], [145], [146], [147], alinierea imaginilor [148], segmentarea mișcării [149], [150], [151], detectarea de muchii [152], [153], selecția modelului [154], etc.

În domeniul problemelor de aliniere a imaginilor, algoritmul RANSAC [138] a fost utilizat intens și constituie metoda de referință. Algoritmul generează soluții ipotetice pe baza unor seturi minimale de puncte și alege dintre acestea soluția care întrunește cel mai bun consens. De-a lungul timpului s-au propus îmbunătățiri ale metodei RANSAC. Metode mai noi derivate din RANSAC sunt MSAC (M-estimator sample consensus) [155], [144], MLESAC (Maximum Likelihood Estimation Sample Consensus) [144], [156], [157], [158], [159], [160], [161], [162]. Îmbunătățiri ale algoritmului MLESAC le putem găsi în [163], în care Tordoff încearcă să reducă timpul de calcul și să obțină o cale ușoară de accesibilitate pentru modele greu accesibile, și în [164], în care se prezintă o variantă adaptivă la zgomot a lui MLESAC. O altă îmbunătățire a lui RANSAC o putem găsi în [165], și anume NAPSAC (N Adjacent Points Sample Consensus), în care se schimbă strategia de selecție a seturilor de puncte, sub presupunerea că datele conforme vor tinde să fie mai aproape unele de altele decât datele neconforme. De asemenea, metoda robustă mean shift a suferit la rândul său îmbunătățiri de-a lungul timpului. La început metoda a fost adaptată de Cheng [166] pentru analiza imaginilor și recent îmbunătățită [167], [168], [169] și extinsă de Comaniciu, Meer și Ramesh pentru problemele de vedere artificială, și anume segmentare [3], netezire adaptivă [3], urmărire [170].

Deși la început înlocuirea estimării nerobuste a parametrilor cu estimarea robustă a avut succesele ei, s-au evidențiat curând și limitările metodei și eșecuri [171]. În prezent se cunosc motivele acestor eșecuri. Majoritatea estimatorilor robusți sunt incapabili de a lucra cu date atunci când se utilizează mai mult de o structură [172], [173], situație des întâlnită în vederea artificială.

## 2.8. Aplicații

### 2.8.1. Problematika realizării alinierii imaginilor

Metodele de aliniere ale imaginilor propuse în această teză, se pot utiliza în aplicații de **rețele de senzori fără fir**. Acestea se utilizează în aplicații din domenii variate, cum ar fi monitorizarea mediului înconjurător, acțiuni de salvare de urgență, operații militare, monitorizare medicală și sisteme industriale [174]. Multe dintre aceste aplicații depind de capacitatea rețelelor de a-și determina corect locațiile nodurilor. Mai mult decât atât, locația se presupune a fi cunoscută în operații ca rutare sau protocoale de securitate. Soluțiile bazate de GPS nu sunt posibile datorită faptului că sunt scumpe și necesită echipamente hardware de putere mare [175]. Metode de localizare au fost studiate și propuse în [176], [177], [178].

Rețele de senzori video fără fir sunt un caz special al rețelelor de senzori fără fir în care o cantitate mare de date video sunt captate, procesate în timp real și apoi transferate către celelalte noduri din rețea [179]. Una dintre aplicațiile în care sunt folosite, monitorizarea video de supraveghere și controlare a mediului înconjurător acoperă o arie de mare interes datorită multitudinilor de amenințări ca uragane, inundații, cutremure, violarea securității sau atacuri teroriste [180]. Elementul video este o componentă importantă în observarea unor fenomene din lumea fizică. De exemplu, camerele pot fi folosite pentru a supraveghea diferite activități, pentru a evalua eroziunea solului și observa diferite specii de animale. Datorită faptului că într-o rețea de senzori fără fir pot exista în număr foarte mare de camere, imaginile captate de unele perechi de camere pot conține un câmp vizual comun. Imaginile

vor fi captate din diferite poziții și unghiuri. O reconstrucție a scenei ar implica combinarea acestor imagini cu scopul de a oferi o vedere panoramică a mediului înconjurător. De aceea, se utilizează metoda de aliniere a imaginilor. Cu toate acestea, la informația adunată ce conține date despre localizarea nodurilor, se va mai adăuga și direcția camerelor și câmpul vizual comun (Field of View). În [181] s-a propus o metodă de localizare împreună cu estimarea suprapunerii de câmpuri vizuale comune pentru aplicațiile de supraveghere care implică corespondența imaginilor cu scopul de a alinia imaginile simultan dobândite de la diferiți senzori video. Toate imaginile sunt procesate central și informația legată de localizare este distribuită în toată rețeaua.

Administrarea acestor rețele este o provocare destul de mare începând cu faza de amplasare și până la întreaga activitate a rețelei. Scopurile sunt de a îmbunătăți acoperirea, de a crește timpul de viață al rețelei și a obține o sincronizare între senzori. Un număr mare de senzori pot fi distribuiți în masă prin împrăștierea lor din avioane sau rachete teleghidate [174]. În acest caz, amplasarea inițială este dificil de controlat. Cu toate acestea, în ciuda acestor dificultăți, este necesară o amplasare bună. Cercetările recente s-au axat pe metode de îmbunătățire a acestei amplasări inițiale. O metodă posibilă este de a adăuga senzori după ce amplasarea inițială a avut loc. Acest proces poartă numele de reamplasare [175]. O altă metodă este de a utiliza senzori mobili [176]. Această metodă permite o realocare a senzorilor. În timp ce prima metodă implică în plus costuri de reamplasare, marele dezavantaj al celei de a doua metode este complexitatea soluției de mișcare, care reduce drastic aria de aplicare și duce la creșterea prețului inițial.

Multe dintre aplicațiile cu aceste rețele depind de capacitatea rețelei de a determina locația nodurilor. Mai mult decât atât, locația se presupune a fi cunoscută atunci când se realizează operațiile de rutare sau protocoale de securitate. Într-adevăr o cercetare semnificată s-a concentrat pe dezvoltarea unor tehnici de localizare eficiente a nodurilor [177].

Monitorizarea video pentru o infrastructură de supraveghere și control este un caz special al aplicațiilor cu rețele de senzori fără fir [179] în care o cantitate mare de date este adunată și procesată în timp real și apoi transmisă în întreaga rețea. Aceste rețele se numesc rețele de senzori video fără fir (Video Wireless Sensor Networks (VWSN)).

### 2.8.2. Localizarea în rețele de senzori fără fir

Luând în considerare natura unei rețele de senzori fără fir, senzorii sunt amplasați într-o manieră ad-hoc și nu se cunoaște locația în cele mai multe cazuri. Din păcate, multe dintre aplicațiile cu acești senzori se bazează pe corectitudinea locației nodurilor.

Diferite soluții de localizare au fost propuse în literatura de specialitate. Ele pot fi clasificate în soluții de localizare bazate pe timp sau pe puterea semnalului și soluții bazate pe apropierea de punctul de referință [182], [183], [184]. În plus, o următoare problemă ar fi aria de acoperire a suprafeței de supravegheat. Acest fapt depinde de localizare, dar nu în totalitate. Depinde de asemenea de condițiile diferite ale metodei de amplasare [185].

Problema de localizare a nodurilor și procedurile de amplasare sunt mult mai complexe în cazul rețelelor de senzori video fără fir. Pe lângă coordonatele nodurilor, informația referitoare la topologie mai include unghiul dintre camere și câmpul vizual

comun. Într-adevăr, acești parametri care sunt în plus, sunt esențiali pentru majoritatea aplicațiilor de supraveghere video.

### 2.8.2.1. Algoritmi de măsurare a distanței

Acești algoritmi sunt folosiți pentru estimarea poziției relative a nodurilor rețelei. O clasificare a acestor algoritmi conține trei mari categorii: măsurători ale unghiului de sosire, măsurători referitoare la distanță și măsurători referitoare la puterea semnalului recepționat (RSS- Received Signal Strength).

Tehnica de măsurare a unghiului de sosire poate să folosească răspunsul în amplitudine sau în fază a receptorului antenei. Acuratețea acestor tehnici este limitată de directivitatea antenei și de reflexiile pe mai multe căi ale acesteia. O componentă cu mai multe căi poate fi privită ca un semnal care vine dintr-o direcție total diferită și poate să ducă la erori foarte mari în măsurătorile unghiului de sosire. Problemele de cale multiplă în măsurarea unghiului de sosire pot fi exprimate folosind algoritmul de verosimilitate maximă. Au fost dezvoltate mai multe tipuri de astfel de algoritmi. Cele mai cunoscute exemple sunt MUSIC (MULTiple Signal Classification) [186] și C-ESPRIT (Conjugate Estimation of Signal Parameters by Rotational Invariance Techniques) [187].

Măsurătorile de distanță se bazează pe propagarea în timp sau pe puterea semnalului radio. În primul caz o posibilitate este oferită de estimarea distanței dintre senzorii vecini folosind timpul de sosire. Acesta reprezintă timpul de propagare a unui semnal între transmițător și receptor. Drept urmare, necesită ca timpul local de la transmițător și cel de la receptor să fie precis sincronizat. Acest dezavantaj face ca măsurătorile timpului de sosire să constituie o metodă mai puțin agreată [184]. O tehnică îmbunătățită este tehnica dus-întors a propagării timpului. Aceasta măsoară diferența dintre momentul de timp când un senzor transmite un semnal și când răspunsul este recepționat de la senzorul corespondent. Deoarece pentru calcule este folosită aceeași bază de timp se evită problema sincronizării temporale. În orice caz o mare sursă generatoare de erori în măsurătorile propagării temporale dus-întors este întârzierea necesară analizei semnalului de către cel de-al doilea senzor [188].

O altă abordare interesantă în măsurătorile distanței este abordarea *lighthouse* care derivă distanța dintre un receptor și un transmițător optic. Ea presupune măsurarea duratei de timp în care o rază aparținând unei surse de lumină, din transmițător, aflată în rotație este prelucrată de receptor. Punctele slabe ale acestei metode sunt: costul ridicat al senzorului optic și cerința de linie directă de vedere între receptorul optic și transmițător.

Altă categorie în tehnica de măsurare a distanței estimează distanța dintre senzorii vecini folosindu-se de puterea semnalului recepționat. Aceasta reprezintă o tehnică agreată deoarece nu necesită echipamente adiționale. Propagarea semnalului este afectată de reflexii, difracții și împrăștiere [189]. În ceea ce privește localizarea bazată pe puterea semnalului recepționat, aceasta funcționează prin formarea unei hărți a comportamentului puterii semnalului în aria de acoperire. Harta este obținută fie offline prin măsurători apriorice, sau online folosind dispozitive de absorbție amplasate în locații cunoscute [190].

### 2.8.2.2. Localizarea în contextul WSN-video

În [181] se propune o nouă soluție de localizare a nodurilor și de estimare a câmpului vizual comun. Se pornește de la imagini video captate de la diferite noduri și se calculează câmpul vizual comun cu ajutorul unui server central (poate fi un PC sau un notebook). Apoi se calculează parametri de translație, rotație și scalare și se transmite informația obținută în întreaga rețea.

Localizarea nodurilor se bazează pe estimarea acestor parametri dintre fiecare pereche de senzori vecini. Procesul implică o aliniere a imaginilor aplicată tuturor imaginilor din rețea. Pentru realizarea procesului de aliniere pentru o pereche de imagini, un pas important este selecția caracteristicilor. Considerându-se selectate caracteristicile, o transformare este găsită și fiecare punct dintr-o imagine (primul senzor) este potrivit cu un punct din a doua imagine (al doilea senzor). În lucrările [181], [191], selecția caracteristicilor a fost manuală și s-a bazat pe extragerea de colțuri printr-un simplu click al mouse-ului. Pentru aceasta era necesar un observator uman pentru a interacționa cu mouse-ul în alegerea perechilor de corespondente dintre punctele caracteristice, în fiecare imagine.

Acest algoritm propus a fost totuși afectat de mici erori (1 până la 5 pixeli) specifice procedurii de corespondență. Pentru a trata erorile, s-a considerat un pas de post procesare numit *asociere chamfer* [192]. Totuși și această metodă prezintă un dezavantaj și anume timpul mare de calcul. În studiile ce urmează, se va studia problema de reducere a timpului de calcul, creșterii acurateții procesului de punere în corespondență a imaginilor pentru rețelele de senzori fără fir și stabilirea unor corespondențe de precizie ridicate.

În [191] se continuă studiul început în [181] și se studiază posibilitatea utilizării metodelor de aliniere a imaginilor în cazul localizării senzorilor într-o rețea de senzori fără fir. Pe lângă acestea, soluția propusă adaugă estimarea suprapunerii câmpurilor vizuale ale senzorilor la localizarea spațială clasică. Se pune accentul de studiul transformărilor de similaritate simultană și secvențială. Se dorește stabilirea unei transformări de similaritate potrivite pentru realizarea cu succes a procesului de aliniere. Ca urmare, acestea sunt analizate și comparate, urmând ca testele experimentale să fie prezentate în capitolul 5.

Pentru a crește simultan acuratețea estimării și a reduce timpul de calcul, se studiază o soluție robustă bazată pe estimatorul mean shift, care elimină posibilele necorespondențe dintre perechile de puncte [193], [194]. Această metodă este prezentată în subcapitolul 4.3. Metoda prezintă un dezavantaj, și anume utilizarea unui prag în procesul de eliminare al eșantioanelor aberante. Alegerea pragului optim este o problemă studiată ulterior în [195]. Toate acestea duc la studierea necesității unei selecții automate a caracteristicilor în procesul de corespondență [196], [197]. În subcapitolul 2.7 este prezentat principiul algoritmului de selecție automată SIFT.

O situație asemănătoare o întâlnim în [183], în care este prezentată o metodă de localizare a nodurilor împreună cu estimarea câmpului vizual comun pentru aplicațiile de supraveghere care implică o metodă de aliniere a imaginilor obținute de la diferiți senzori video. Acest proces se bazează pe o estimare mean shift [198] pentru o estimare robustă a parametrilor rotație, translație și scalare. Estimarea robustă se bazează pe algoritmul mean shift, o metodă robustă și rapidă de detectare al maximumului local al unei densități multivariată de probabilitate.

## **2.9. Concluzii**

În acest capitol sunt trecute în revistă metode de aliniere a imaginilor prezentate în literatură. Sunt prezentate considerațiile generale privind metodele de aliniere a imaginilor, precum și etapele de parcurgere în realizarea acestor metode.

Este argumentată necesitatea dezvoltării unor tehnici de estimare robuste și sunt exemplificate aplicații ale metodelor de aliniere a imaginilor. Ca domeniu de aplicabilitate al acestor metode, se pune accentul în cadrul tezei de doctorat pe rețelele de senzori video fără fir, utilizate în aplicații din domeniile: monitorizarea mediului înconjurător, acțiuni de salvare de urgență, operații militare, monitorizare medicală și sisteme industriale.

## 3. FUNDAMENTELE TEORETICE PRIVIND TEHNICILE DE ALINIERE A IMAGINILOR

În acest capitol se prezintă fundamentele teoretice ale metodelor de aliniere a imaginilor. Capitolul este structurat pe nouă subcapitole și prezintă aspectele importante dintr-un proces de aliniere a imaginilor. Primul paragraf prezintă tipurile de transformări geometrice utilizate în corespondența imaginilor, cu exemplele reprezentative. În paragrafele următoare se tratează teoretic tehnicile robuste de estimare care se folosesc în vederea artificială: metodele parametrice și metodele de eșantionare aleatoare. Concluziile capitolului sunt sintetizate în paragraful final.

### 3.1. Transformări geometrice

În acest paragraf se introduc succint transformările geometrice 2D și 3D relevante pentru conținutul tezei.

#### 3.1.1. Transformări geometrice bidimensionale (2D)

O transformare geometrică proiectează fiecare pixel, de coordonate  $(x, y)$ , al imaginii de intrare, la o noua poziție,  $(x', y')$  în imaginea de ieșire. Transformarea poate fi descrisă prin ecuațiile:

$$\begin{aligned}x' &= T_x(x, y), \\y' &= T_y(x, y).\end{aligned}\tag{3.1}$$

Transformările  $T_x$ ,  $T_y$  pot fi cunoscute dinainte sau pot fi estimate pe baza imaginii inițiale și a celei transformate.

O transformare  $T$  este lineară dacă pentru fiecare constantă  $C$

$$T(x_1 + x_2) = T(x_1) + T(x_2)\tag{3.2}$$

și

$$cT(x) = T(cx)\tag{3.3}$$

Transformările liniare se împart în transformări afine și de proiecție. Transformările de interes care vor fi prezentate în această lucrare, potrivesc un set de puncte peste un alt set de puncte, și includ: rotație, translație, scalare și distorsiunea de perspectivă (proiecția unui punct).

### 3.1.1.1. Transformările afine

O transformare este afină dacă relația  $T(x) - T(0)$  este lineară. Această transformare este suficientă pentru a potrivi două imagini dintr-o scenă captată din poziții diferite ale camerelor foto. Transformările afine includ operații ca translația, scalarea și rotația. O transformare se numește globală dacă se aplică pe întreaga imagine, și locală dacă porțiunile din imagine au fiecare transformata lor.

**Translația** este transformarea ce deplasează toate punctele unui obiect cu un anumit deplasament față de poziția inițială. Translația poate fi exprimată fie sub forma unui sistem de ecuații, fie (cel mai adesea) sub formă matriceală

$$\begin{cases} x' = x + T_x \\ y' = y + T_y \end{cases} \Leftrightarrow \begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} T_x \\ T_y \end{bmatrix} \quad (3.4)$$

unde  $(x', y')$  sunt coordonatele punctului care se translatează,  $T_x, T_y$  sunt deplasamentele pe abscisă și ordonată (figura 3.1) și  $(x, y)$  sunt coordonatele punctului.

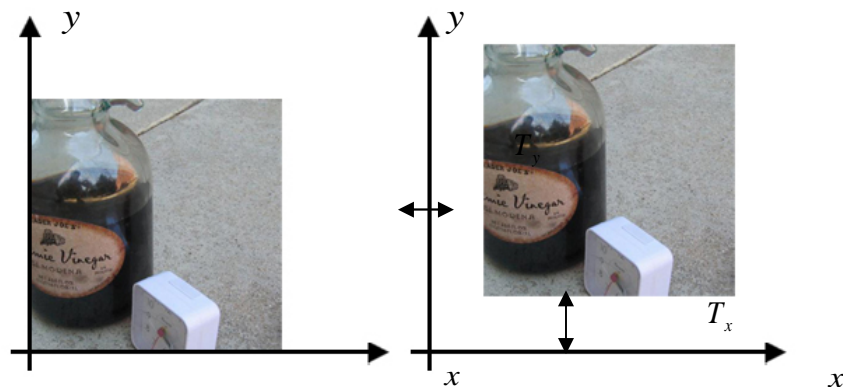


Fig. 3.1. Translația unui obiect bidimensional.

**Scalarea** se realizează modificând coordonatele pixelilor după regula:

$$\begin{cases} x' = S_x x \\ y' = S_y y \end{cases} \Leftrightarrow \begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} S_x & 0 \\ 0 & S_y \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \quad (3.5)$$

unde  $S_x, S_y$  sunt factorii de scalare. Dacă factorii de scalare sunt negativi, atunci rezultatul scalării este **reflexia** față de axa corespunzătoare factorilor de scalare negativi.





Fig. 3.2. Scalarea unui obiect bidimensional.

Factorii de scalare supraunitari aplicați mulțimii de puncte ale unui obiect conduc la supradimensionarea obiectului în fereastra de reprezentare și depărtarea acestuia de originea sistemului de coordonate, în timp ce factorii de scalare subunitari conduc la micșorarea obiectului și apropierea lui de originea sistemului de coordonate.

Următoarea transformare afină este **rotația** de unghi  $\theta$  a unui obiect față de origine conduce la păstrarea distanței obiectului față de origine și modificarea coordonatelor în sensul pozitiv (trigonometric) pentru  $\theta > 0$  sau în sensul orar, negativ, pentru  $\theta < 0$ .

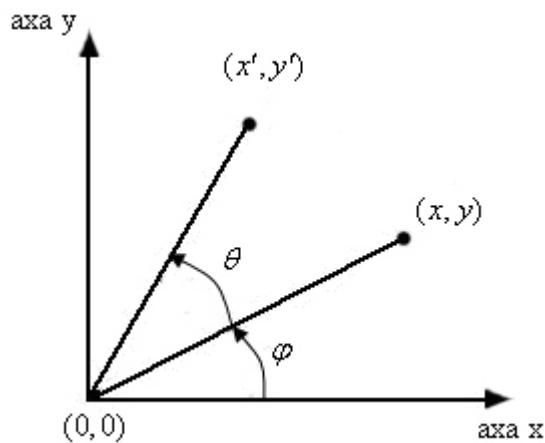


Fig. 3.3. Rotația unui obiect față de originea sistemului de coordonate.

Pentru exprimarea rotației este preferabilă reprezentarea în sistemul de coordonate polare (figura 3.3). Rotația este astfel caracterizată de sistemul:

$$R_{\theta} : \begin{cases} x' = x \cos \theta - y \sin \theta \\ y' = x \sin \theta + y \cos \theta \end{cases} \Leftrightarrow \begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \quad (3.6)$$

### 3.1.1.2. Principiul proiecției de perspectivă în 2D

Proiecția de perspectivă mai poartă numele de **homografie** și este modelată ca și o deformare 2D folosind coordonate omogene. Imaginile  $\mathbf{X}$  și  $\mathbf{x}'$  în două câmpuri vizuale a unui punct  $X$  care se afla pe un plan (figura 3.4), satisfac relația  $\mathbf{x}' = H\mathbf{x}$ , în coordonate omogene. Matricea de proiecție 3x3  $H$  poartă numele de homografie. Homografia poate fi recuperată dintr-un număr suficient de puncte și corespondențe de linii [199]. Matricea de proiecție poate fi descrisă sub forma:

$$H = \begin{bmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & 1 \end{bmatrix} \quad (3.7)$$

Iar relația  $\mathbf{x}' = H\mathbf{x}$  se scrie sub forma:

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (3.8)$$

Cele 8 necunoscute ale matricei  $H$  se pot afla cu setul de ecuații:

$$\begin{cases} x' = \frac{h_1x + h_2y + h_3}{h_7x + h_8y + 1} \\ y' = \frac{h_4x + h_5y + h_6}{h_7x + h_8y + 1} \end{cases} \quad (3.9)$$

unde  $x', y'$  reprezintă coordonatele punctului din imaginea inițială;

$x, y$  reprezintă coordonatele punctului din imaginea transformată;

$h_1 \dots h_8$  reprezintă parametrii matricei  $H$ .

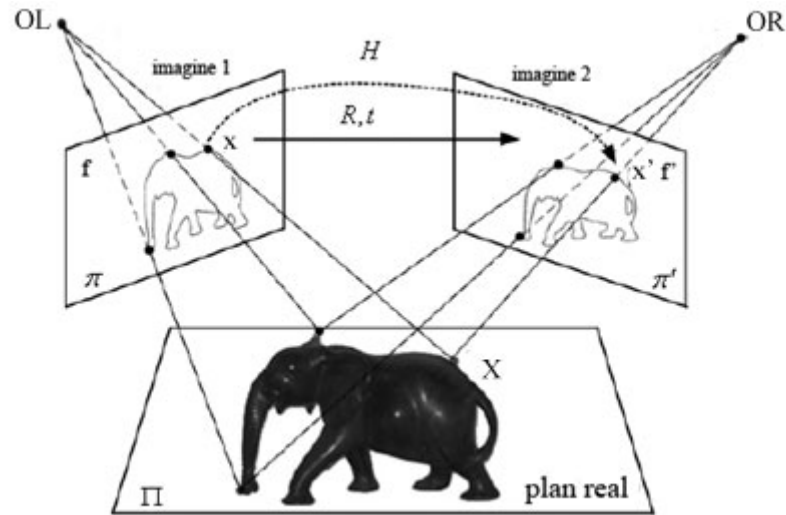


Fig. 3.4. Homografia indusă de un plan.

Raza care intersectează prima imagine de plan în punctul  $x$  întâlnește planul real  $\Pi$  în punctul  $X$ . Punctul  $X$  este proiectat într-un punct  $x'$  din a doua imagine. Asocierea din punctul  $x$  în  $x'$  este homografia, descrisă de  $H$ , indusă de planul  $\Pi$ .  $R$ , respectiv  $t$ , reprezintă rotația și translația relativă dintre camere.

Operația fundamentală în inteligența artificială implică capturarea de imagini ale obiectelor 3D și proiecția lor într-o perspectivă 2D. Presupunem existența a patru puncte  $(x_1, y_1), (x_2, y_2), (x_3, y_3)$  și  $(x_4, y_4)$  ale unui obiect rectangular. Se utilizează metoda de homografie plană pentru a proiecta aceste puncte într-un dreptunghi având colțurile  $(X_1, Y_1), (X_2, Y_2), (X_3, Y_3)$  și  $(X_4, Y_4)$  ( figura 3.5).

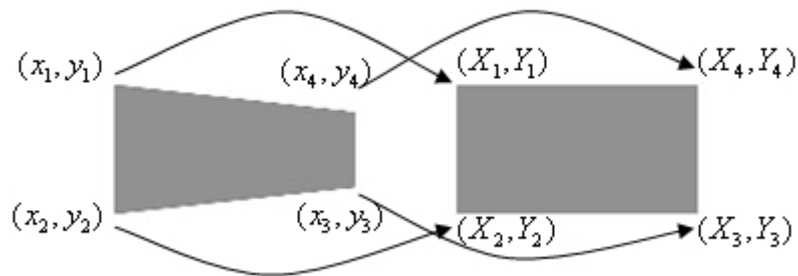


Fig. 3.5. Transformarea homografică plană.

Rezultatul transformării homografice plană poate fi observată în figura de mai jos, fiind aplicată pe o pereche de imagini 2D. Se poate observa corecția distorsiunii de perspectivă în imaginea de ieșire (c) și alinierea acesteia cu imaginea de referință (a).

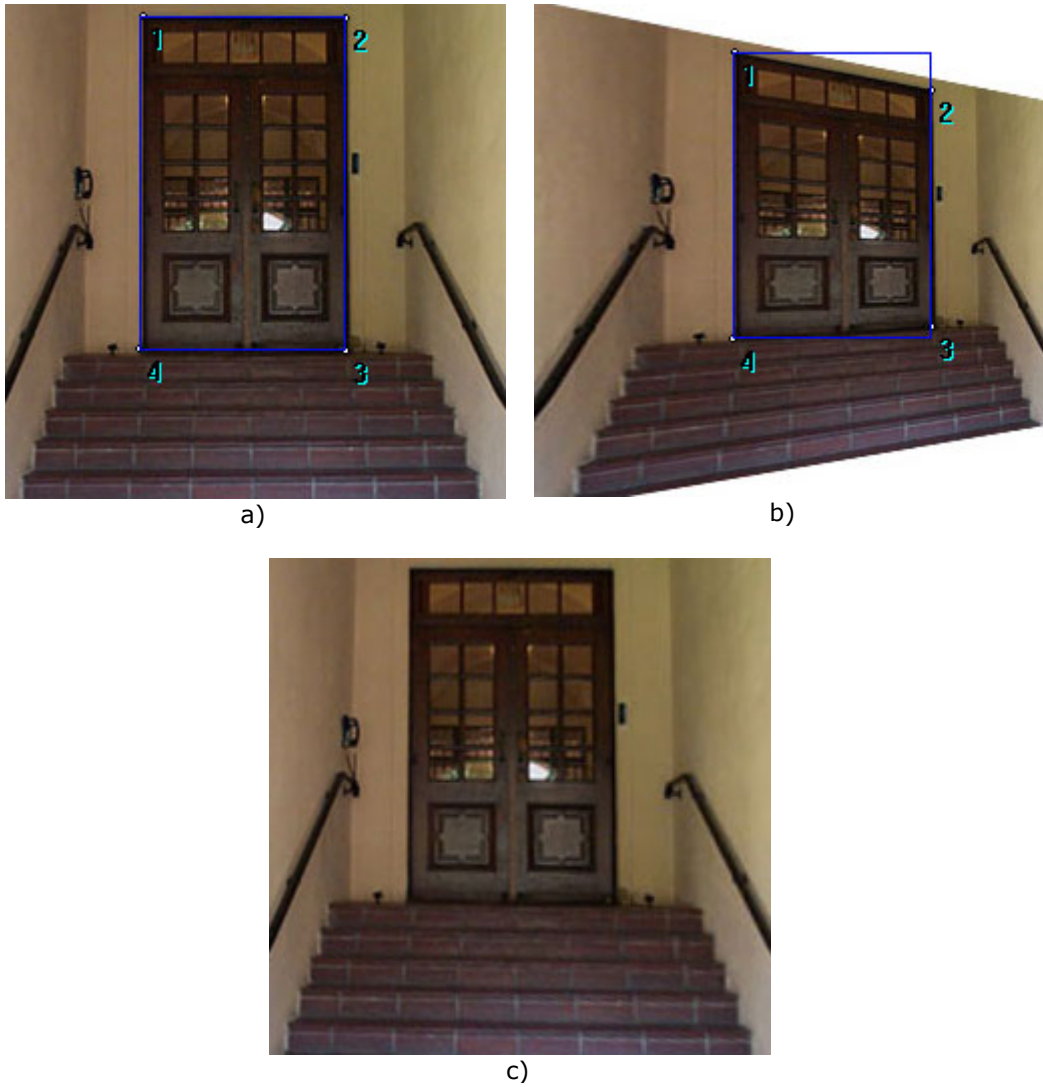


Fig. 3.6. Imaginea de intrare a) este procesată în funcție de imaginea transformată b) cu ajutorul transformării homografice, în imaginea de ieșire c).

Această transformare este dată de soluția următoarei ecuații:

$$A * T = B \quad (3.10)$$

unde A este

$$\begin{pmatrix} -x_1 & -y_1 & -1 & 0 & 0 & 0 & X_1x_1 & X_1y_1 \\ 0 & 0 & 0 & -x_i & -y_i & -1 & Y_1x_1 & Y_1y_1 \\ -x_2 & -y_2 & -1 & 0 & 0 & 0 & X_2x_2 & X_2y_2 \\ 0 & 0 & 0 & -x_i & -y_i & -1 & Y_2x_2 & Y_2y_2 \\ -x_3 & -y_3 & -1 & 0 & 0 & 0 & X_3x_3 & X_3y_3 \\ 0 & 0 & 0 & -x_i & -y_i & -1 & Y_3x_3 & Y_3y_3 \\ -x_4 & -y_4 & -1 & 0 & 0 & 0 & X_4x_4 & X_4y_4 \\ 0 & 0 & 0 & -x_4 & -y_4 & -1 & Y_4x_4 & Y_4y_4 \end{pmatrix} \quad (3.11)$$

și vectorii

$$T = (h_1, h_2, h_3, h_4, h_5, h_6, h_7, h_8)^T$$

$$B = (X_1, Y_1, X_2, Y_2, X_3, Y_3, X_4, Y_4)^T$$

Vectorul  $T$  reprezintă cele 8 necunoscute ale matricii homografice  $3 \times 3$  care cuprinde rotația, scalarea, translația, distorsionarea, întinderea și distorsiunea de perspectivă.

### 3.1.2. Transformările geometrice tridimensionale (3D)

Pentru cazul reprezentărilor tridimensionale (pe scurt 3D), matricea de reprezentare va avea dimensiunea  $4 \times 4$  și se vor utiliza coordonate omogene. Prin transformarea de omogenizare a unui punct  $P(x, y, z)$ , definit în coordonate 3D, se asociază un punct  $(x, y, z, w)$  în coordonate omogene. Reprezentarea standard a unui punct în coordonate omogene  $(x, y, z, w)$  este, pentru toate punctele cu  $w \neq 0$ :  $(x/w, z/w, y/w, 1)$ .

Vom trece în continuare în revistă și vom caracteriza principalele transformări definite anterior pentru spațiul bidimensional.

### 3.1.2.1. Transformările afine

**Translația** tridimensională este caracterizată în spațiul tridimensional de matricea:

$$T(T_x, T_y, T_z) = \begin{bmatrix} 1 & 0 & 0 & T_x \\ 0 & 1 & 0 & T_y \\ 0 & 0 & 1 & T_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (3.12)$$

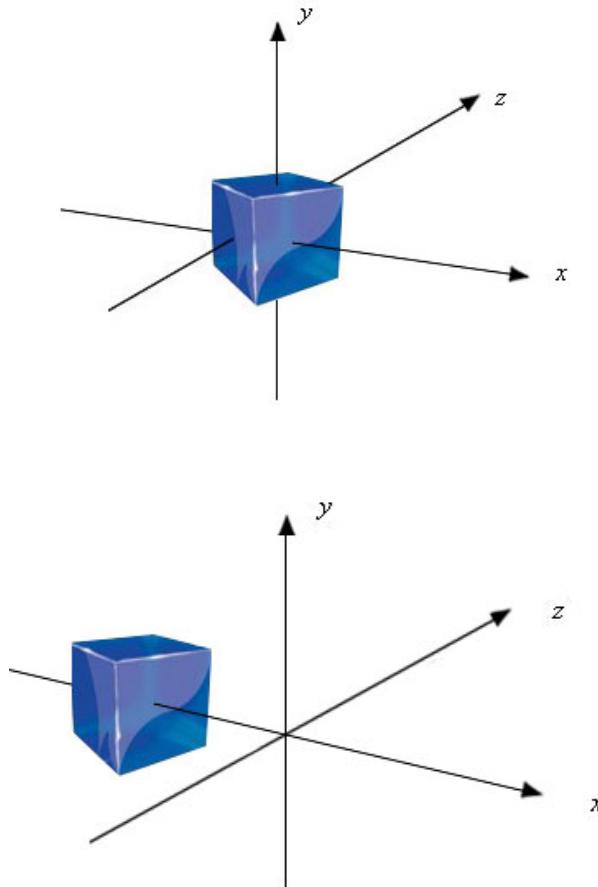


Fig. 3.7. Translația unui obiect tridimensional.

**Scalarea** tridimensională este caracterizată în spațiul tridimensional de matricea:

$$S(S_x, S_y, S_z) = \begin{bmatrix} S_x & 0 & 0 & 0 \\ 0 & S_y & 0 & 0 \\ 0 & 0 & S_z & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (3.13)$$

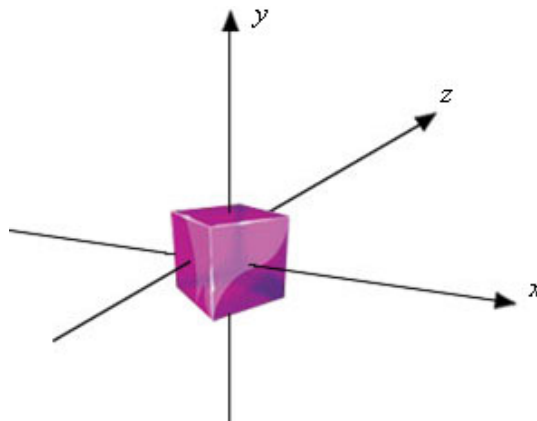
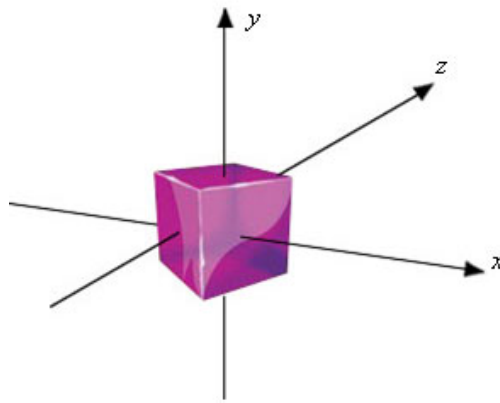


Fig. 3.8. Scalarea unui obiect tridimensional.

**Rotația** în spațiul tridimensional se descompune în maximum trei rotații după fiecare axa a sistemului 3D. Convenția de semn (pozitiv) este pentru axa  $Z$  de la axa  $X$  către axa  $Y$ , pentru axa  $Y$  de la axa  $Z$  către axa  $X$ , pentru axa  $X$  de la axa  $Y$  către axa  $Z$ .

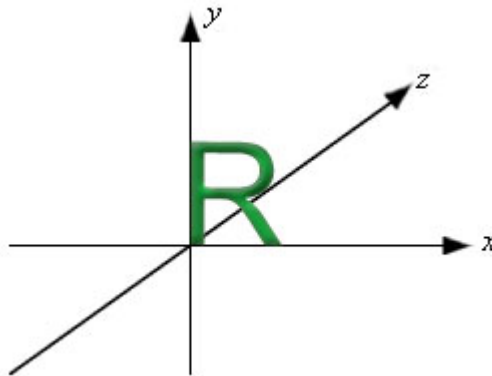


Fig. 3.9. Rotația unui obiect tridimensional.

Rotația după axa  $x$  este reprezentată de ecuațiile:

$$\begin{aligned} x' &= x \\ y' &= y \cos(\theta) - z \sin(\theta) \\ z' &= y \sin(\theta) + z \cos(\theta) \end{aligned} \quad (3.14)$$

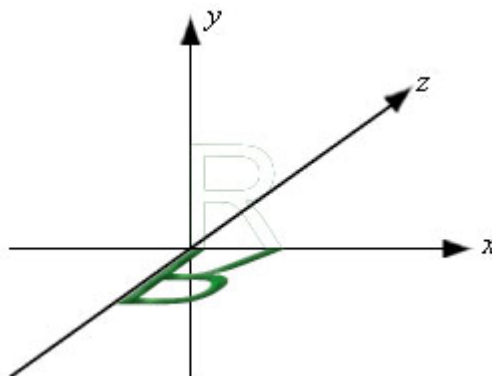


Fig. 3.10. Rotația după axa  $x$ .



$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta & 0 \\ 0 & \sin \theta & \cos \theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (3.15)$$

unde  $R_x = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta & 0 \\ 0 & \sin \theta & \cos \theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$  reprezintă matricea de rotație după  $x$ .

Rotația după axa  $y$  este reprezentată de ecuațiile:

$$\begin{aligned} x' &= x \cos(\theta) + z \sin(\theta) \\ y' &= y \\ z' &= z \cos(\theta) - x \sin(\theta) \end{aligned} \quad (3.16)$$

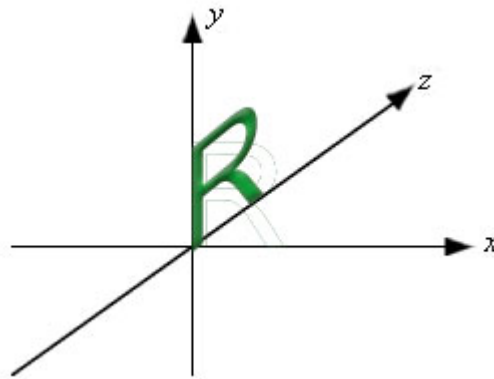


Fig. 3.11. Rotația după axa  $y$ .

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} \cos \theta & 0 & -\sin \theta & 0 \\ 0 & 1 & 0 & 0 \\ \sin \theta & 0 & \cos \theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (3.17)$$

unde  $R_y = \begin{bmatrix} \cos \theta & 0 & -\sin \theta & 0 \\ 0 & 1 & 0 & 0 \\ \sin \theta & 0 & \cos \theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$  reprezintă matricea de rotație după  $y$ .

Rotația după axa z este reprezentată de ecuațiile:

$$\begin{aligned}x' &= x \cos(\theta) - y \sin(\theta) \\y' &= x \sin(\theta) + y \cos(\theta) \\z' &= z\end{aligned}\quad (3.18)$$

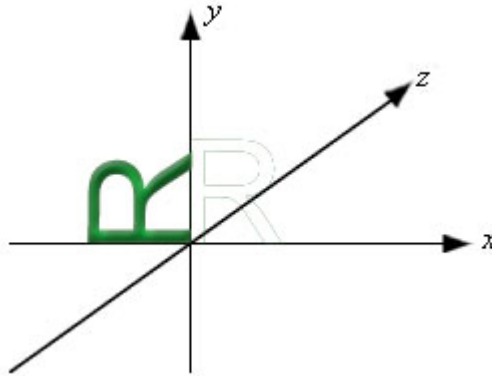


Fig. 3.12. Rotația după axa z.

$$(3.19) \quad \begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta & 0 & 0 \\ \sin \theta & \cos \theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

unde  $R_z = \begin{bmatrix} \cos \theta & -\sin \theta & 0 & 0 \\ \sin \theta & \cos \theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$  reprezintă matricea de rotație după z.

### 3.1.2.2. Proiecția unui obiect tridimensional

Razele de proiecție (numite *proiectori*) trec printr-un punct dat al spațiului (numit *centru de proiecție*) și prin fiecare punct al obiectului, intersectând planul de proiecție pe care se realizează proiecția.

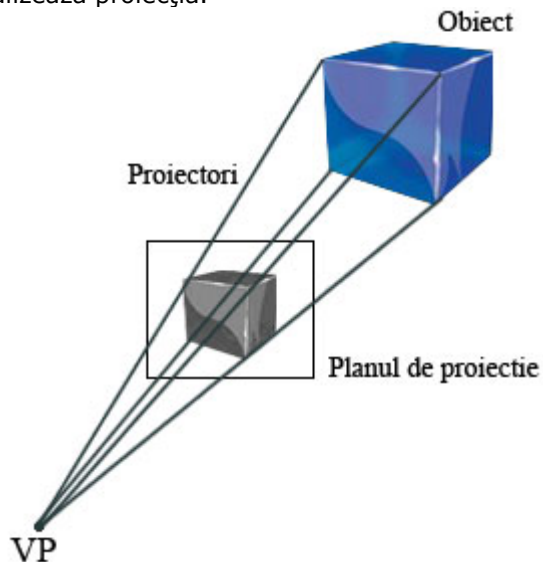


Fig. 3.13. Proiecția de perspectivă

Proiecțiile plane pot fi *paralele* sau *perspective*. Acestea diferă prin distanța dintre centrul de proiecție și planul de proiecție. Dacă distanța este finită se obține **proiecția perspectivă** (Figura 3.13); dacă distanța este infinită se obține **proiecția paralelă** (Figura 3.14). Proiecția perspectivă este caracterizată de poziția centrului de proiecție și de planul de proiecție. Proiecția paralelă este caracterizată de direcția de proiecție și de planul de proiecție.

În cazul proiecției de perspectivă, dacă liniile sunt paralele între ele și neparalele cu planul de proiecție converg spre un punct de fugă sau limită (în limba engleză *vanishing point*) (Figura 3.13). Dacă liniile sunt paralele cu axele principale, punctul de fugă se numește *punct de fugă axial*. Într-o asemenea proiecție pot exista cel mult trei puncte de fugă.

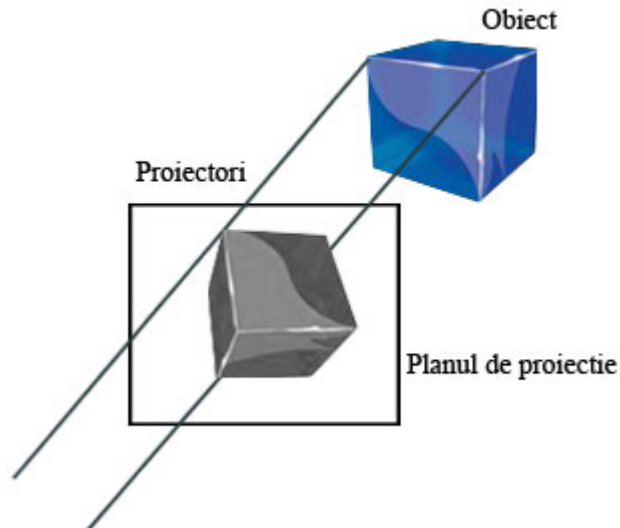


Fig. 3.14. Proiecția de perspectivă.

Proiecția de perspectivă calculează coordonatele 2D ale unui obiect în planul imaginii din coordonatele 3D ale acestuia din lumea reală. Un punct din lumea reală este proiectat printr-un punct în planul imaginii  $Z = 0$ . În figura 3.15 avem  $(x_r, y_r, z_r)$  - coordonatele punctului curent din spațiu și  $(x_i, y_i)$  - coordonatele proiecției lui în planul imaginii, iar  $f$  este distanța focală a sistemului de lentile.

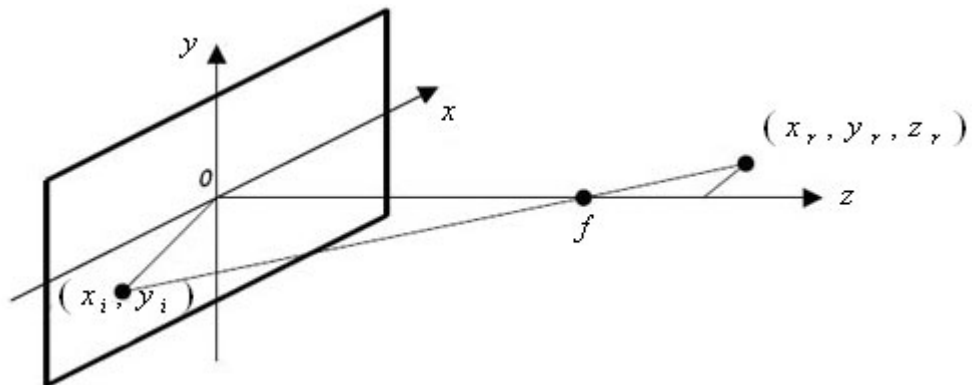


Fig. 3.15. Proiecția unui punct din lumea reală în planul imaginii.

Distorsiunea de perspectivă este dată de matricea:

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -\frac{1}{f} & 1 \end{bmatrix} \quad (3.20)$$

Coordonatele omogene ale proiecției punctului  $\mathbf{p} = [wx_r, wy_r, wz_r, w]^T$  pe planul de imagine sunt [200]:

$$c_h = P\mathbf{p} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -1/f & 1 \end{bmatrix} \begin{bmatrix} wx_r \\ wy_r \\ wz_r \\ w \end{bmatrix} = \begin{bmatrix} wx_r \\ wy_r \\ wz_r \\ w - (wz_r / f) \end{bmatrix} \quad (3.21)$$

Coordonatele carteziene ale proiecției sunt:

$$c = \begin{bmatrix} x_i \\ y_i \\ z_i \end{bmatrix} = \begin{bmatrix} \frac{fx_r}{f - z_r} \\ \frac{fy_r}{f - z_r} \\ \frac{fz_r}{f - z_r} \end{bmatrix} \quad (3.22)$$

Din asemănarea triunghiurilor formate în figura 3.15 se deduc ecuațiile:

$$x_i = \frac{-fx_r}{z_r - f}, \quad (3.23)$$

$$y_i = \frac{-fy_r}{z_r - f} \quad (3.24)$$

Ecuțiile de mai sus permit calculul coordonatelor punctului din planul de imagine în funcție de coordonatele punctului din lumea reală care l-a generat.

### 3.2. Tehnici de estimare robustă

Estimarea robustă este o problemă cheie în multe domenii din inteligența artificială, în special vederea artificială. Prin termenul de estimare robustă se înțeleg acele tehnici de estimare care sunt robuste în prezența datelor aberante, date care sunt definite ca niște observații care nu se potrivesc cu modelul stohastic al estimării parametrilor. Ele pot apărea din cauza zgomotului sau din cauza erorilor rezultate din urma măsurărilor și interpretărilor greșite ale datelor. Tehnicile robuste sunt fie adoptate din statistică sau create special pentru domeniul de vedere artificială. Aceste tehnici se împart în două mari direcții: metode parametrice, care depind de forma funcției probabilitate de densitate (pdf) a datelor, și metodele neparametrice, al căror scop este de a estima funcția de densitate de probabilitate (pdf) prin presupuneri minime despre structura funcției.

Tehnicile robuste se pot caracteriza printr-o serie de măsuri de robustețe. Una din ele este *punctul de ruptură* (breakdown point) [201], reprezentând procentajul maxim de date eronate arbitrar de mult, fără ca soluția să fie deplasată arbitrar de mult. De exemplu, metoda celor mai mici pătrate (LS) [201] are punctul de ruptură la 0 din cauză că o singură dată neconformă poate compromite complet estimarea, indiferent de numărul de date conforme. Pe de alta parte, LmeS [202] are punctul de ruptură la 0.5 din cauza că jumătate din date pot fi aberante fără ca aceasta să compromită eșantionarea. Cu toate acestea, tehnicile de estimare robustă necesită un punct de ruptură având valoarea mai mare decât 50%.

O altă măsură de robustețe este *funcția de influență* [146], [203], care reprezintă schimbarea unei estimate, schimbare cauzată de inserarea unei funcții de distanță dintre un eșantion aberant și estimată (se înlocuiește valoarea  $i$  din distribuție cu o valoare arbitrară, urmărindu-se ieșirea estimatorului). De exemplu, funcția de influență a estimatorului celor mai mici pătrate este proporțională cu distanța dintre punct și estimată. Pentru a obține robustețea, funcția de influență ar trebui să tindă la 0 pe măsură ce distanța se mărește.

În [204] estimarea robustă este descrisă ca și o *extensie a estimării parametrilor* pentru a compensa sensibilitatea erorii grosiere a pătratelor reziduurilor (least squares). Se presupune că în seturile de date există o amestecare de două sau mai multe distribuții. Una din distribuții poate proveni dintr-o mulțime de date  $M$ , iar cealaltă poate proveni din afara mulțimii  $M$ . Se presupune că datele au o dimensiune  $d$  și că distribuția de probabilitate ca datele să conțină un amestec de  $k$  distribuții se poate exprima în felul următor: fie  $p_i$  probabilitatea de distribuție  $i$  și  $\mu_i$  coeficientul de amestecare.

$$p = \mu_1 p_1 + \mu_2 p_2 + \dots + \mu_k p_k \quad (3.25)$$

$$p = \sum_{i=1}^k \mu_i p_i \quad (3.26)$$

Din cauză că probabilitatea trebuie să fie normalizată

$$\sum_{i=1}^k \mu_i = 1 \quad (3.27)$$

Se mai definește un indicator de variabila  $\gamma_i$

$$\begin{cases} \gamma_i = 1 \text{ dacă data } i \text{ din } x \text{ este generată de } p_1 \\ \gamma_i = 0 \text{ dacă data } i \text{ din } x \text{ este generată de } p_2 \end{cases}$$

(3.28)

Probabilitatea de selecție a datelor generate de  $p_1$  este

$$\underline{\gamma} = 1/n \sum \gamma_i \quad (3.29)$$

Estimarea robustă mai poate fi definită ca și o *problema de separare a distribuțiilor* [204]. Există două distribuții în estimarea robustă, și anume distribuția plauzibilă ( $p_i$ ) și o distribuție necunoscută, aberantă ( $p_o$ ). Folosind ecuația (3.25) putem deduce că:

$$p(\alpha) = \mu_i p_i(\alpha) + \mu_o p_o(\alpha)$$

(3.30)

Această metodă de clasificare conduce la două tipuri de erori.

- Tipul I - eroarea apare atunci când un eșantion aberant este considerat plauzibil.
- Tipul II - eroarea apare când un eșantion plauzibil este considerat aberant.

În aplicațiile de aliniere a imaginilor, de cele mai multe ori sunt presupuse corespondenței punctele caracteristice reprezentate prin colțuri, și datorită acestui fapt, metodele utilizate a stabili corespondența, cum ar fi cross-corelația, sunt nesigure, iar seturile de corespondențe care apar pot fi aleatoare [205]. Aceste date incorecte formează datele neconforme care trebuie filtrate cu ajutorul metodelor robuste. Majoritatea algoritmilor din inteligența artificială includ o etapă de estimare robustă, în care parametrii modelului sunt estimați dintr-un set de date conforme, ale căror structură se apropie de structura impusă de date, care conține o proporție considerabilă de date neconforme.

Valorile neconforme sunt observații depărtate de structura impusă de partea cea mai mare a datelor. În prezența valorilor neconforme, metodele statistice clasice pot deveni complet ineficiente. De exemplu, o singură valoare neconformă poate face ca media de selecție să dea rezultate complet eronate. Dacă oricare din valorile de selecție tinde către infinit, atunci media de selecție tinde către infinit. Spre deosebire de media de selecție, mediana de selecție este mai puțin afectată de prezența unei valori neconforme. Spunem că mediana este rezistentă sau robustă la prezența valorilor neconforme, în timp ce media de selecție nu are această proprietate.

În literatura de specialitate se întâlnesc două metode importante de estimare de parametri. Ele sunt estimarea celor mai mici pătrate (LSE) și estimarea prin verosimilitatea maximă (MLE).

De exemplu, metoda minimizării mediane pătratelor rezidurilor LMeS (Least Median of Squares) [202], [206] are la bază principiul celor mai mici pătrate. Un alt estimator de parametri înrudit cu LMeS, dar cu robustețe îmbunătățită în prezența erorilor grosiere prezente în date, sunt estimatorii M [146] care au la bază principiul

verosimilității maxime. Ambele metode au fost importate din statistică. LMeS face față unui număr mai mare de erori, până la 50% dar prezintă eficiență statistică scăzută. Estimatorii M prezintă eficiență statistică ridicată dar tolerează un număr mai mic de erori.

Un alt estimator des utilizat în estimarea de parametri este estimatorul RANSAC [2], dezvoltat pentru punerea în corespondență a seturilor de puncte 2D și 3D. Algoritmul are la bază principiul verosimilității maxime și este un estimator de eșantionare aleatoare. El este capabil de a face față unui număr mare de erori grosiere, și se poate utiliza și în probleme de estimare a modelului parametrilor, cum ar fi alinierea caracteristicilor, estimarea geometriei epipolare sau detectarea primitivelor geometrice. Principiul acestei metode a dus la dezvoltarea unei serii de algoritmi robusți care au la bază eșantionarea aleatoare, fiind utilizați să minimizeze diferite criterii: MSAC [155], [144], MLESAC [144], NAPSAC [165], MINPRAN [207].

### 3.3. Estimatorii M

#### 3.3.1. Legătura cu estimatorii de verosimilitate maximă

Estimarea de verosimilitate maximă determină parametrii care maximizează probabilitatea datelor. Din [208] putem deduce că MLE prezintă multe proprietăți optime în estimare: suficiență (informație completă despre parametrii estimatorului MLE); consistență (valoarea parametrului care generează datele recuperate asimptotic) și invarianța la parametrizare (aceeași soluție MLE obținută independent). Aceste lucruri nu pot fi găsite la LSE. Drept urmare, majoritatea statisticienilor nu privesc LSE ca o metodă generală pentru estimarea parametrilor, ci mai degrabă ca o abordare inițială folosită împreună cu modele de regresie liniară.

Considerăm  $n$  valori,  $x_1, x_2, \dots, x_n$ , generate aleator de o distribuție cu densitatea de probabilitate  $p(x | \theta)$ , unde

$$p(x | \theta) \geq 0 \quad (3.31)$$

ce depinde de un parametru  $\theta$ . Din acest set de  $n$  valori trebuie să se găsească valoarea optimă a lui  $\theta$ . Se poate defini funcția de verosimilitate:

$$L(\theta) = \prod_{i=1}^n p(x_i | \theta) \quad (3.32)$$

reprezentând probabilitatea de a se observa setul de  $n$  date, pentru valoarea specificată a parametrului  $\theta$ . Din motive teoretice este mai ușor de lucrat cu forma logaritmică a funcției  $L(\theta)$  sau cu forma negativă a logaritmului:

$$F(\theta) = -\ln L(\theta) = -\sum_{i=1}^n \ln p(x_i | \theta) \quad (3.33)$$



Metoda de maximă verosimilitate estimează  $\theta$  prin găsirea valorii lui  $\hat{\theta}$  care maximizează  $L(\theta)$  sau minimizează forma negativă a lui  $L(\theta)$ . Estimatorul de maximă verosimilitate (MLE) al lui  $\theta$  este reprezentat de ecuația:

$$\hat{\theta} = \underset{\theta}{\operatorname{argmin}} F(\theta) \quad (3.34)$$

Prin urmare, legătura estimatorilor M cu estimatorii de maximă verosimilitate se poate face prin intermediul unei funcții de pierdere  $\rho(\cdot)$ , denumită și normă a erorii. Estimatorii M pot fi definiți ca și o generalizare a estimatorilor de maximă verosimilitate și sunt soluția ecuației:

$$\min \left( \sum_{i=1}^n \rho(r_i) \right) \quad (3.35)$$

unde  $r_i$  reprezintă rezidualul experimental  $i$ , adică diferența dintre observația  $i$  și valoarea dată de un model. Metoda standard a celor mai mici pătrate (LS), minimizează suma pătratelor reziduurilor

$$\sum_i r_i^2 \quad (3.36)$$

care devine instabilă atunci când datele conțin date neconforme. Estimatorii M încearcă să reducă efectul reziduurilor generate de date neconforme prin înlocuirea pătratelor ecuației (3.36) cu ecuația (3.35). Soluția estimatorului M este definită de o ecuație de forma:

$$\hat{\theta} = \underset{\theta}{\operatorname{argmin}} \sum_{i=1}^n \rho(r_i | \theta). \quad (3.37)$$

Pentru corespondența cu estimatorii de maximă verosimilitate, funcțiile de pierdere pot fi interpretate ca logaritmi ai funcțiilor densitate de probabilitate a erorii:

$$\rho(r) = -\ln p(r) \quad (3.38)$$

Semnul minus din ecuația de mai sus are ca efect transformarea problemei de maxim rezolvate de estimatorii de maximă verosimilitate în probleme de minim rezolvate de estimatorii M.

### 3.3.2. Funcția de influență și funcția de pondere a estimatorilor M

Ecuția  $\min\left(\sum_{i=1}^n \rho(r_i)\right)$  poate fi rezolvată sub forma estimării iterative ponderate a celor mai mici pătrate (iterated reweighted least mean squares, IRLS). Fie  $p = [p_1, \dots, p_m]^T$  vectorul parametrilor modelului care urmează a fi estimat. Produsul unui estimator M care utilizează funcția  $\rho(r_i)$  este vectorul  $p$ , soluție a următoarelor  $m$  ecuații:

$$\sum_i \psi(r_i) \frac{\partial r_i}{\partial p_j} = 0, \quad j = 1, \dots, m \quad (3.39)$$

unde derivata  $\psi(x) = \frac{d\rho(x)}{dx}$  se numește funcția de influență. Dacă se definește o funcție de pondere

$$w(r) = \frac{\psi(r)}{r}, \quad (3.40)$$

atunci ecuația (3.39) devine

$$\sum_i w(r_i) r_i \frac{\partial r_i}{\partial p_j} = 0, \quad j = 1, \dots, m. \quad (3.41)$$

Funcția de influență  $\psi(r)$  este o măsură a influenței pe care o observație o are asupra parametrilor estimați. Dacă  $r_i = C(x_i - v_j)$  atunci din ecuația (3.41) obținem:

$$v_j = \frac{\sum_i w(r_i) x_i}{\sum_i w(r_i)}. \quad (3.42)$$

Factorul  $v_j$  devine media ponderată a lui  $x_i$  și poate fi rezolvată iterativ.

Un estimator este **robust** dacă influența unei singure observații este insuficientă pentru a produce o deplasare semnificativă în estimății. Un estimator M robust trebuie să satisfacă un număr de condiții:

- Să aibă o funcție de influență mărginită;
- Să asigure unicitatea soluției. Funcția obiectiv prin minimizarea căreia se obține vectorul de parametri  $p$  trebuie să fie unimodală (cu un singur extrem) ceea ce matematic corespunde unei convexități a funcției  $\rho$  în variabila  $p$ ;

- Ori de câte ori derivata a doua  $\frac{\partial^2 \rho(\cdot)}{\partial p^2}$  este singulară, funcția obiectiv trebuie să aibă un gradient nenul,  $\frac{\partial \rho(\cdot)}{\partial p} \neq 0$ , pentru a evita căutarea prin întreg spațiul parametrilor.

### 3.4. Minimizarea medianeii pătratelor rezidurilor (Least Median of Squares – LMeS)

LMeS este un estimator robust dezvoltat de către Rousseeuw [202] pentru probleme de regresie. Estimatorul estimează parametrii unui model prin rezolvarea problemei de minimizare neliniară.

Modelul clasic liniar este dat de  $y_i = x_{i1}\theta_1 + \dots + x_{ip}\theta_p + e_i (i = 1, \dots, n)$ , unde eroarea  $e_i$  are o distribuție cu media 0 și deviația standard  $\sigma$ . Scopul regresiei este de a estima  $\theta = (\theta_1, \dots, \theta_p)$  din datele  $(x_{i1}, \dots, x_{ip}, y_i)$ . Principiul acestei metode pleacă de la metoda celor mai mici pătrate LS (least squares):

$$\min_{\hat{\theta}} \sum_{i=1}^n r_i^2, \quad (3.43)$$

unde reziduurile  $r_i$  sunt egale cu  $y_i - x_{i1}\hat{\theta}_1 - \dots - x_{ip}\hat{\theta}_p$ . În ciuda avantajului simplității matematice care caracterizează metoda celor mai mici pătrate, acest estimator poate conduce la lipsă de robustețe atunci când setul de date este contaminat cu date neconforme. De exemplu, o problemă generală în statistică este caracterizarea unui set de puncte  $P$  de către o linie dreaptă. Metoda prezentată mai sus, cea a celor mai mici pătrate, minimizează suma pătratelor distanțelor a punctelor de pe linie. Dacă un singur eșantion este deplasat la infinit, eroarea de estimare devine, la rândul ei, infinită.

Plecând de la aceste considerente, așa cum Rousseeuw a sugerat în cartea sa [202]: „Cei mai mulți oameni au încercat să facă acest estimator unul robust prin înlocuirea pătratului cu altceva... Dar de ce să nu înlocuim suma cu o mediană, care este foarte robustă?”. În consecință el definește funcția de pierderi prin ecuația:

$$C(\hat{\theta}) = \min \text{med } r_i^2 \quad (3.44)$$

Metoda LS se bazează pe calculul mediilor, iar mediile sunt sensibile la extreme. Prin înlocuirea acestei medii cu o mediană, se face ca estimatorul LMeS să fie unul mai robust.

Estimatorul LMeS minimizează mediana pătratelor rezidurilor, dovedindu-se a fi foarte robust în probleme de aliniere a imaginilor atunci când există potriviri false sau date neconforme datorită localizării proaste. Robustețea acestui estimator se datorează utilizării medianeii. Acest estimator are punctul de ruptură de 0.5. Teoretic, 0.5 este limita maximă pentru orice estimator robust. Ea poate fi, totuși depășită dacă eșantioanele neconforme nu conspiră. Asemenea situații nu sunt excepționale, ci mai degrabă curente în problemele de aliniere a imaginilor.

Problema estimării minimului mediane pătratelor rezidurilor nu poate fi redusă la o problemă a criteriului celor mai mici pătrate ponderat, pentru că funcția de pierdere nu este derivabilă. Este aproape imposibil a scrie o formulă exactă pentru estimatorul LMeS. Aceasta trebuie rezolvată prin căutarea spațiului de posibile estimate generate din date. Din moment ce spațiul este prea mare, numai un subset de date ales aleatoriu poate fi analizat.

Un dezavantaj al acestui estimator este durata timpului de calcul. Din [71] reiese ca LmeS nu a prezentat rezultate bune pentru aplicațiile complexe din vederea artificială, unde adesea procentul de contaminare cu date neconforme depășește 50%.

### 3.5. Algoritmul RANSAC

Una dintre cele mai puternice metode de estimare robustă este RANSAC [2]. RANSAC este acronimul de la "RANdom SAMple Consensus". Este o metodă iterativă de estimare a parametrilor unui model matematic dintr-un set de date care conține date neconforme. El a fost dezvoltat pentru corespondența seturilor de date 3D cu 2D, dar și utilizat pentru sarcini ca estimarea geometriei epipolare [144], selecția modelului de mișcare [209], estimarea matricei fundamentale [210], estimarea tensorului trifocal [211], estimarea poziției camerei [212], estimarea de structuri prin mișcare [212]. Alte aplicații ale metodei RANSAC se pot găsi în [199]. De la introducerea sa de acum două decade de către Fischler și Bolles [2], și stabilirea sa mai târziu în literatura statistică de către Rousseeuw și Leroy [201], paradigma de consens cu eșantionare aleatoare (random sampling consensus) a fost folosită în multe aplicații de inteligență artificială. Algoritmul RANSAC funcționează în doi pași, fiind o asociere între o metodă de generare de soluții-ipoteză, și un algoritm de verificare.

Metodele de eșantionare aleatoare generează soluțiile-ipoteză din seturi minimale de date, selectate astfel încât fiecare punct să aibă aceeași verosimilitate de selecție. Ideea de bază este de a se alege  $L$  seturi minimale de puncte corespondente, cu presupunerea că cel puțin unul din seturi va conține exclusiv puncte corespondente. Un set minimal conține numărul minim de date ce permite calculul unei soluții, respectiv generarea unei ipoteze pentru parametrii modelului,  $\theta$ . Rațiunea principală pentru care se preferă un set minimal, în locul unui extins, care ar conduce la un sistem de ecuații supradeterminat, nu constă în evitarea complicațiilor de calcul aferente unei asemenea opțiuni ci în minimizarea șanselor ca setul să conțină date neconforme.

Etapa a doua a algoritmului, cea de verificare, folosește o funcție de cost  $C(\hat{\theta})$  pentru a determina soluția care asigură cea mai bună asociere cu modelul. Funcția cost se evaluează pentru toate soluțiile generate și constă în numărarea perechilor de puncte al căror vecin cel mai apropiat în imaginea corespondentă după aplicarea transformării estimate se află la o distanță mai mică decât un anumit prag. Ipoteza care primește cel mai mare număr de puncte suport, respectiv cel mai mic număr de puncte ce nu verifică modelul furnizează estimata algoritmului. De remarcat faptul că identificarea unui singur set de date necontaminat este suficient pentru generarea unei soluții corecte.

Fie  $\theta$  valoarea adevărată a estimatei  $\hat{\theta}$ . Pașii algoritmului sunt prezentați după cum urmează:

- pentru  $i = 0, n$ 
  - se alege un set minimal de date pe baza căruia se realizează estimarea unui vector al parametrilor modelului  $\hat{\theta}_i$ .
  - se calculează valoarea funcției de cost  $C(\hat{\theta}_i)$ , pentru a se determina numărul de puncte care nu sunt în consens cu soluția  $\hat{\theta}_i$ .
- se selectează modelul,  $\hat{\theta}^*$ , pentru care costul  $C(\hat{\theta}_i)$  este minim,  $i = 1 \dots n$
- se folosesc date conforme (inliers) pentru a genera estimata finală  $\hat{\theta}$ .

Se presupune că fiecare coordonată a spațiului parametrilor, poate lua un număr finit de valori,  $q$ , datorită unei cuantizări care are loc, și că probabilitatea de selecție a unei soluții corecte este  $\frac{1}{q}$ . Într-un spațiu al parametrilor de dimensiune  $n$ , probabilitatea de selecție a unei soluții corecte este:

$$p = \left(\frac{1}{q}\right)^n \quad (3.45)$$

Prin urmare, eficiența acestei metode este scăde exponențial cu dimensiunea vectorului parametrilor. În concluzie, selecția aleatoare este ineficientă pentru aplicații cu spații largi de căutare și dimensiune mare, așa cum se regăsesc în estimarea robustă.

Dacă punctele caracteristice generate în imagini nu ar fi perturbate de zgomot sau erori de măsurare, atunci un singur set minimal bine ales ar determina corect parametrii  $\theta$  ai modelului. Cu toate acestea, din cauza că punctele sunt perturbate de zgomot sau de erori de măsurare, când numărul de seturi minimale utilizat pentru a se determina parametrii este mic, discrepanța dintre  $\hat{\theta}$  și  $\theta$  poate fi foarte mare.

Găsirea unui model exact, implică utilizarea unui prag  $T$  care determină corect numărul de puncte caracteristice conforme unei soluții. Eroarea calculată pentru fiecare punct din setul de puncte disponibil este distanța la cel mai apropiat punct din imaginea de referință, după aplicarea transformării geometrice de aliniere estimate. Eroarea se va compara cu acest prag pentru a se clasifica punctul ca fiind conform sau aberant. Procesul se va repeta de  $L$  ori. În final se va alege setul de date care minimizează funcția de cost reprezentată de:

$$C = \sum_i \rho(e_i^2) \quad (3.46)$$

unde  $\rho()$  este definită prin ecuația:

$$\rho(e^2) = \begin{cases} 0 & e^2 < T^2 \\ \text{constant} & e^2 > T^2. \end{cases} \quad (3.47)$$

Cu alte cuvinte, datele conforme nu înregistrează nici o penalizare, pe când datele neconforme înregistrează o penalizare constantă.

Torr, în [144], pune calculează numărul iterații este nevoie pentru ca algoritmul să aibă o soluție cu semnificație statistică. Ideal ar fi să se aleagă toate subseturile posibile din datele existente, dar din punctul de vedere al calcului, este adesea prohibitiv. Fischler și Bolles [2] și Rousseeuw și Leroy [201] au propus câteva moduri de calcul pentru numărul de seturi minime necesare,  $L$ .

Fie  $P$  probabilitatea de a găsi un set de necontaminat de  $S$  date, din  $L$  încercări, având la dispoziție  $N$  date, între care  $M \leq N$  aparțin modelului. Ne interesează să determinăm numărul de seturi de date de care avem nevoie pentru a genera cu probabilitatea dorită,  $P$ , soluția corectă. Procentul de date contaminate este definit prin ecuația:

$$\varepsilon = \frac{N - M}{N} \quad (3.48)$$

Probabilitatea de a avea un set de  $S$  date integral corecte este

$$P_c = (1 - \varepsilon)^S \quad (3.49)$$

în timp ce probabilitatea ca setul să nu fie integral corect este  $1 - P_c$ . Probabilitatea ca niciunul din cele  $L$  seturi de date selectate să nu fie integral conforme modelului este :

$$P_e = [1 - (1 - \varepsilon)^S]^L \quad (3.50)$$

respectiv probabilitatea de a avea cel puțin un set de date necontaminat este

$$p = 1 - p_e \quad (3.51)$$

Impunând probabilitatea  $P$  de a avea cel puțin un set de date necontaminat, respectiv o soluție corectă, deducem numărul minim de seturi de date pe care trebuie să-l folosim:

$$L = \frac{\log(1 - P)}{\log[1 - (1 - \varepsilon)^S]} \quad (3.52)$$

Valorile numărului  $L$  de seturi de eșantioane necesare crește rapid cu dimensiunea setului minimal și cu procentul de date neconforme modelului. constată creșterea exponențială a numărului de seturi de date necesare cu dimensiunea spațiului parametrilor și a setului de date minimal. Similar este și efectul creșterii procentului de contaminare.

Viteza algoritmului RANSAC depinde de doi factori. În primul rând, de numărul de eșantioane aleatoare care trebuie alese pentru a garanta o anumită certitudine în optimalitatea soluției. În al doilea rând, timpul petrecut evaluând calitatea fiecăruia dintre parametrii modelului este proporțional cu dimensiunea setului de date.

În [213] se descriu metode de oprire a algoritmului RANSAC. Dacă procentul de contaminare este necunoscut, numărul de experimente necesar,  $L$ , se poate recalcula după fiecare experiment, estimând procentul de contaminare prin calculul

raportului între numărul de date conforme corespunzătoare celui mai bun model găsit până la momentul respectiv și numărul total de eşantioane de date.

Metode combinate care utilizează eşantionarea aleatoare și estimatorii M se pot întâlni în [214] și [173].

### 3.6. Algoritmul MSAC

MSAC (estimatorul M cu consens) este o îmbunătățire a algoritmului RANSAC, propusă de către Torr și Zisserman [155], [144]. Diferența dintre RANSAC și MSAC este că RANSAC potrivește parametrii modelului cu scopul de a maximiza numărul de date conforme, iar MSAC potrivește parametrii modelului cu scopul de a maximiza verosimilitatea datelor conforme, care furnizează parametrii modelul dorit.

Metoda utilizează un estimator M redescendent (în engleză redescending) [215], la care funcția de influență este diferită de zero numai pe un suport finit, simetric în raport cu originea, asemenea variantei de bază. La RANSAC datele conforme au o penalizare egală cu 0, în timp ce datele neconforme au o penalizare constantă. La MSAC se utilizează o funcție de cost unde datele conforme au o penalizare diferită de zero, cu o valoare aleasă în funcție de cât de bine satisfac modelul parametrilor. Datele neconforme au o penalizare constantă, la fel ca la RANSAC:

$$C = \sum_i \rho_2(r_i^2) \quad (3.53)$$

unde

$$\rho_2(r_i^2) = \begin{cases} r^2 & r^2 < T^2 \\ T^2 & r^2 \geq T^2 \end{cases} \quad (3.54)$$

Utilizarea metodei MSAC în locul lui RANSAC nu necesită un calcul în plus.

### 3.7. Algoritmul MLESAC

MLESAC (Maximum Likelihood Estimation Sample Consensus) este o metodă de estimare a unei soluții de verosimilitate maximă prin eşantionarea aleatoare de date. Algoritmul este o adaptare a lui RANSAC, și a fost dezvoltat de către Torr și Zisserman [144] pentru estimarea matricilor fundamentale și a transformărilor de perspectivă din corespondența caracteristicilor. MLESAC urmărește ideea de bază, aceea a eşantionării aleatoare în care seturi minimale de seturi de corespondențe sunt utilizate pentru a dezvolta o ipoteză, iar corespondențele rămase se utilizează pentru a evalua calitatea fiecărei ipoteze. MLESAC evaluează verosimilitatea ipotezei prin reprezentarea erorii de distribuție ca un model de mixtură între o distribuție normală și una uniformă, cu un parametru de amestec necunoscut. Datele conforme se consideră că sunt afectate de erori distribuite normal, în timp ce erorile produse de datele neconforme sunt modelate printr-o distribuție uniformă.

Pentru a înțelege mai bine sensul relațiilor care au loc în MLESAC, Torr și Zisserman [144] efectuează câteva notații. Fie  $\mathbf{X}$  un punct dintr-o scenă 2D, unde  $\mathbf{x}$  face parte din prima imagine și  $\mathbf{x}'$  în a doua imagine, fiind vectori omogeni,

având coordonatele  $\mathbf{x} = (x, y, 1)^T$ . Corespondența  $\mathbf{x} \leftrightarrow \mathbf{x}'$  va fi notată  $\mathbf{x}_1 \leftrightarrow \mathbf{x}_2$ . Simbolul  $\underline{x}$  indică valoarea corectă a lui  $x$ , unde  $x = \underline{x} + \Delta_x$  reprezintă valoarea coruptă de zgomot.

Pornind de la corespondența  $\mathbf{x}^{1,2}$ , scopul este de a se găsi estimata de maximă verosimilitate,  $\hat{\mathbf{x}}^{1,2}$  a poziției corecte a lui  $\underline{\mathbf{x}}^{1,2}$ , astfel încât  $\hat{\mathbf{x}}^{1,2}$  satisface și minimizează relația:

$$\mathbf{e}_i^2 = \sum_{j=1,2} (\underline{x}_i^j - x_i^j)^2 + (\underline{y}_i^j + y_i^j)^2 \quad (3.55)$$

Metoda implică un amestec  $\gamma$  de date conforme perturbate gaussian (cu o deviație standard  $\sigma$ ) și date neconforme distribuite uniform (într-o fereastră de dimensiune  $v$ ).

$$\begin{aligned} p_{in}(\mathbf{e}) &= \gamma \left( \frac{1}{\sqrt{2\pi}\sigma} \right)^n \exp\left(-\frac{\mathbf{e}^2}{2\sigma^2}\right) \\ p_{out}(\mathbf{e}) &= (1-\gamma) \frac{1}{v} \end{aligned} \quad (3.56)$$

unde  $\sigma$  este deviația standard a distribuției Gaussiene,  $n$  este dimensiunea datelor măsurate,  $v$  este dimensiunea ferestrei de căutare. Informația legată de raportul între numărul eșantioanelor aberante și al celor plauzibile nu se cunoaște și urmează a fi estimat pe baza datelor experimentale, dar se consideră cunoscută dispersia erorii eșantioanelor conforme,  $\sigma$ . Parametrul  $\gamma$ , se determină din condiția de a se maximiza verosimilitatea datelor  $d$  care constă dintr-o mixtură între  $p_{in}(\mathbf{e})$  și  $p_{out}(\mathbf{e})$  după cum urmează:

$$p(d | \mathbf{M}) = \gamma \left( \frac{1}{\sqrt{2\pi}\sigma} \right)^n \exp\left(-\frac{\mathbf{e}^2}{2\sigma^2}\right) + (1-\gamma) \frac{1}{v} \quad (3.57)$$

unde  $e$  este eroarea distanțelor euclidiene dintre puncte, iar  $\mathbf{M}$  reprezintă estimata de maximă verosimilitate. Verosimilitatea pentru toate datele  $D$  se obține din

$$p(D | \mathbf{M}) = \prod_{i=1}^k p(d_i | \mathbf{M}) \quad (3.58)$$



În acest fel, funcția negativă logaritmică de verosimilitate se poate scrie

$$\begin{aligned}
 FNLV(\mathbf{M}) &= -\log p(D | \mathbf{M}) \\
 &= -\sum_{i=1}^k \log p(d_i | \mathbf{M}) \\
 &= -\sum_i \log \left( \gamma \left( \frac{1}{\sqrt{2\pi}\sigma} \right)^n \exp \left( -e_i / (2\sigma^2) + (1-\gamma) \frac{1}{v} \right) \right).
 \end{aligned} \tag{3.59}$$

Cel mai probabil model este acela care minimizează verosimilitatea negativă logaritmică. MLESAC a fost prezentat în [144] ca fiind superior lui RANSAC în diferite aplicații de vedere artificială.

Pentru estimarea parametrului de mixtură,  $\gamma$ , și al intervalului distribuției, presupuse uniforme a eșantioanelor neconforme, MLESAC, folosește un algoritm EM [216]. Algoritmul este inițializat cu o valoare de 0.5 pentru  $\gamma$  și constă din doi pași.

În pasul E, se estimează probabilitățile  $p_i, p_0$  conform ecuațiilor:

$$p_i = \gamma \left( \frac{1}{\sqrt{2\pi}\sigma} \right)^n \exp \left( - \left( \sum_{j=1,2} (x_i^j - x_i^j)^2 + (y_i^j + y_i^j)^2 \right) / (2\sigma^2) \right) \tag{3.60}$$

și

$$p_0 = (1-\gamma) \frac{1}{v} \tag{3.61}$$

În pasul M, se recalculează parametrul de amestec din ecuațiile

$$z_i = \frac{p_i}{p_i + p_0} \tag{3.62}$$

și

$$\gamma = \frac{1}{n} \sum_i z_i \tag{3.63}$$

De menționat semnificația parametrului de amestec:  $\gamma$  este probabilitatea ca eșantioanele să fie plauzibile. Parametrul  $\sigma$  este deviația standard a distribuției Gaussiene pentru fiecare coordonată și nu este reactualizat la fiecare iterație, contrar uzanțelor de aplicare a algoritmului EM. Parametrul  $v$  al distribuției uniforme este inițializat prin diametrul ferestrei de eșantioane. Pe parcursul iterațiilor, acest parametru este re-estimat pe baza erorii de asociere maxime generate de soluția curentă.

### 3.8. Algoritmul NAPSAC

S-a demonstrat că în cazul aplicațiilor cu zgomot ridicat (peste 40%), RANSAC este superior metodei de minimizare a medianei rezidurilor Least Median of Squares [209]. Această problemă s-a dovedit a fi adevărată pentru cazurile în care 50% din date sunt date neconforme, nivel la care LMedS eșuează (mediana devine un element ineficient de estimare a datelor conforme).

Plecând de la presupunerea că distribuția de date neconforme este necunoscută, algoritmul RANSAC va furniza cele mai precise estimate, deși va suferi probleme din cauza dimensiunii mari a spațiului și a zgomotului. Myatt [165] realizează un studiu în care implică algoritmul RANSAC și arată că acesta devine nepotrivit pentru aplicațiile care lucrează în spații de dimensiuni mari și zgomot ridicat, iar în încercarea de estimare a parametrilor modelului în aceste condiții, probabilitatea ca RANSAC să reușească, fără a necesita un timp destul de mare de procesare, devine zero.

Pentru a contracara eșecul metodei RANSAC și a altor algoritmi asociați, în spații de dimensiuni ridicate, datorită selecției aleatoare de date obiective, se poate îmbunătăți un astfel de algoritm prin luarea în considerare a distribuțiilor datelor în spațiul multidimensional. În acest scop, Myatt a dezvoltat un nou algoritm [165] care se poate utiliza în locul procesului de eșantionare uniformă în orice algoritm de estimare. NAPSAC (N Adjacent Points Sample Consensus) propune următorul algoritm de eșantionare aleatoare eficientă, care exploatează ideea că eșantioanele conforme sunt grupate spațial în imagine:

1. Se selectează aleator din toate punctele un punct inițial  $x_0$ .
2. Se găsește setul de puncte  $S_{x_0}$  care se află într-o zonă cu raza  $r$ , centrată în  $x_0$ .
3. Dacă numărul de puncte din  $S_{x_0}$  este mai mic decât dimensiunea setului minimal, atunci algoritmul se întrerupe.
4. Se selectează uniform puncte din  $S_{x_0}$  pana când setul minimal a fost selectat, inclusiv pe  $x_0$ .

Dacă punctul inițial  $x_0$  se încadrează într-un model, atunci restul de puncte eșantionate adiacente, teoretic, vor avea o probabilitate mai mare de a fi date plauzibile. În caz contrar, dacă un punct aberant este selectat, atunci datele adiacente au șanse mici de a fi date plauzibile. Dacă nu sunt suficiente puncte pentru a estima modelul, atunci eșantionul este abandonat. Acest algoritm s-a intitulat NAPSAC de la N Adjacent Points Sample Consensus.

Pe baza experimentelor [165] care au avut loc între NAPSAC și RANSAC, pentru o problemă de stabilire a corespondenței între imagini, s-a demonstrat că probabilitatea de selecție a unui set de date conforme, în spațiul de dimensiuni mari și zgomot ridicat, este crescută. NAPSAC este mult mai eficient decât RANSAC în spații de dimensiune mare O alternativă a metodei NAPSAC a fost propusă deja de Tordoff în [142].

### 3.9. Concluzii

În multe din aplicațiile de procesare de imagini este necesar să alinieze imagini multiple ale aceleași scene achiziționate de senzori diferiți, sau imagini captate de același senzor dar la momente diferite de timp. Imaginile care urmează a fi aliniate, pot conține distorsiuni. Cunoașterea acestor distorsiuni trebuie aplicată în dezvoltarea sau selectarea unei metode pentru o anumită aplicație. Distorsiunile reprezintă sursa unei alinieri false care are loc. Se pot utiliza pentru a decide care transformare să fie folosită la alinierea cu succes a imaginilor.

În paragraful 3.1 sunt prezentate pe larg transformările geometrice utilizate pentru alinierea imaginilor. Aceste tehnici de modelare matematică sunt utilizate pentru a corecta erorile geometrice, precum rotația, translația și scalarea imaginii de intrare ( imaginea țintă) în raport cu imaginea de referință, astfel încât imaginile să fie utilizate în diverse aplicații: detectarea schimbărilor, fuziunea imaginilor etc. În metodele convenționale, aceste distorsiuni sunt corectate de către punctele de control care se utilizează în stabilirea modelului matematic.

Metodele de aliniere a imaginilor au aplicații în diverse domenii, iar în funcție de domeniu de aplicabilitate, realizarea procesului de aliniere a imaginilor implică cunoașterea următoarelor informații: poziția și identitatea unui obiect, poziția camerei, geometria suprafețelor obiectelor, transformările necesare construirii unei imagini de ansamblu (mozaic) dintr-o serie de imagini ale aceleași scene, care se suprapun. Fiecare proces, utilizat pentru a extrage aceste informații, folosește estimarea parametrilor pentru a descrie transformările afine și de proiecție care au loc între imagini.

Metodele de estimare din aplicațiile de vedere artificială trebuie să estimeze robust modelul parametrilor în prezența erorilor de măsurare (date neconforme) și a zgomotului din date. Prezența datelor neconforme, atunci când metoda de corespondență nu este robustă, poate conduce la o acuratețe de nivel scăzut a metodei de aliniere a imaginilor pentru a gamă largă de aplicații.

În acest scop, în acest capitol sunt descrise pe larg metodele de estimare robuste a modelului parametrilor. Dintre cele mai clasice amintim pe RANSAC, MSAC, MLESAC, NAPSAC, estimatorii M și Lmes, toate acestea fiind utilizate cu succes în diferite aplicații.



## 4.1. Estimarea densității de probabilitate folosind funcții nucleu

### 4.1.1. Algoritmul neparametric mean shift

Mean shift este un algoritm eficient pentru estimarea locației modurilor densității de probabilitate. Este în esență un algoritm de tip gradient ascendent și a fost dezvoltat de către Fukunaga și Hostetler în urmă cu 30 de ani [217] și recent utilizat în probleme de vedere artificială de către Comaniciu și Meer [167], [3].

Mean shift nu prezintă un prag și este o metodă neparametrică din cauza estimatei densității și a gradientului densității folosind funcții nucleu. Funcțiile nucleu pot diferi și acest fapt duce la versiuni diferite de meanshift. De exemplu, poate exista mean shift gaussian (GMS) unde estimata densității kernel este un amestec gaussian.

Estimarea densității folosind funcții nucleu (cunoscută ca și tehnica cu fereastră Parzen) este o metodă populară de estimare a densității de probabilitate. Se dă un număr  $n$  de date  $\mathbf{x}_i, i = 1, \dots, n$  în spațiul  $d$  - dimensional, extrase dintr-o distribuție cu o funcție de densitate de probabilitate  $p(\mathbf{x})$ . O estimată a acestei densități  $\mathbf{x}$  poate fi scrisă sub forma:

$$\hat{p}_{\mathbf{H}}(\mathbf{x}) = \frac{1}{n} \sum_{i=1}^n K_{\mathbf{H}}(\mathbf{x} - \mathbf{x}_i) \quad (4.1)$$

unde

$$K_{\mathbf{H}}(\mathbf{x}) = |\mathbf{H}|^{-1/2} K_{\mathbf{H}}(\mathbf{H}^{-1/2} \mathbf{x}) \quad (4.2)$$

$K$  este funcția nucleu care depinde de matricea simetrică pozitivă  $\mathbf{H}$  cu dimensiunea  $d \times d$ , numită matricea lățime de bandă. Matricea  $\mathbf{H}$  poate fi aleasă ca fiind diagonală  $H = \text{diag}[h_1^2, \dots, h_d^2]$  sau proporțională cu matricea de identitate  $\mathbf{H} = h^2 \mathbf{I}$ , presupunând aceeași scara  $h$  pentru toate dimensiunile.

Profilul unei funcții nucleu simetrice radiale este definită ca:

$$K^R(\mathbf{x}) = c_{k,d} k(\|\mathbf{x}\|^2) \quad (4.3)$$

cu  $c_{k,d}$  o constantă normalizată strict pozitivă și  $k(x)$  ca fiind profilul funcției nucleu, numai pentru  $x \geq 0$ . Folosind un singur parametru de lățime de bandă,  $h > 0$ , estimata densității se poate pune în forma:

$$\hat{f}(x) = \frac{1}{nh^d} \sum_{i=1}^n K\left(\frac{\mathbf{x} - \mathbf{x}^i}{h}\right) \quad (4.4)$$

Calitatea estimatorului densității folosind funcții nucleu, este măsurată de către eroarea medie pătratică dintre densitatea actuală și estimata. Ecuația (4.4) mai poate fi scrisă:

$$\hat{f}(\mathbf{x}) = \frac{c_{k,d}}{nh^d} \sum_{i=1}^n k\left(\left\|\frac{\mathbf{x} - \mathbf{x}^i}{h}\right\|^2\right) \quad (4.5)$$

Primul pas în analiza spațiului caracteristicilor cu densitatea  $f(\mathbf{x})$  este de a găsi maximele acestei densități. Ele se găsesc printre zerourile gradientului  $\nabla f(\mathbf{x}) = 0$ , iar procedura de mean shift este un mod elegant de a localiza aceste zero-uri fără a estima densitatea.

Exploatând liniaritatea lui (4.5) putem obține un estimator al densității de gradient.

$$\hat{\nabla} f_{h,k}(\mathbf{x}) \equiv \nabla \hat{f}_{h,k}(\mathbf{x}) = \frac{2c_{k,d}}{nh^{d+2}} \sum_{i=1}^n (\mathbf{x} - \mathbf{x}^i) k'\left(\left\|\frac{\mathbf{x} - \mathbf{x}^i}{h}\right\|^2\right) \quad (4.6)$$

Se definește funcția nucleu  $G$ :

$$G(\mathbf{x}) = c_{g,d} \times g(\|\mathbf{x}\|^2) \quad (4.7)$$

unde

$$g(x) = -k'(x) = -\frac{dk(x)}{dx} \quad (4.8)$$

Nucleul  $K(x)$  a fost numit „umbra” lui  $G(x)$  în [166]. Astfel, nucleul Epanechnikov este „umbra” nucleului fereastră uniform. Estimata densității a fost definită până acum cu nucleul  $K$ , dar mean shift-ul se va executa cu nucleul  $G$ .

Înlocuind (4.8) în (4.6), se obține

$$\begin{aligned} \hat{\nabla} f_{h,k}(\mathbf{x}) &= \frac{2c_{k,d}}{nh^{d+2}} \sum_{i=1}^n (\mathbf{x}^i - \mathbf{x}) g\left(\left\|\frac{\mathbf{x} - \mathbf{x}^i}{h}\right\|^2\right) \\ &= \frac{2c_{k,d}}{nh^{d+2}} \left[ \sum_{i=1}^n g\left(\left\|\frac{\mathbf{x} - \mathbf{x}^i}{h}\right\|^2\right) \right] \left[ \frac{\sum_{i=1}^n \mathbf{x}^i g\left(\left\|\frac{\mathbf{x} - \mathbf{x}^i}{h}\right\|^2\right)}{\sum_{i=1}^n g\left(\left\|\frac{\mathbf{x} - \mathbf{x}^i}{h}\right\|^2\right)} - \mathbf{x} \right] \end{aligned} \quad (4.9)$$

Din (4.5) se poate observa că primul termen este proporțional cu estimata densității în  $\mathbf{x}$  calculată cu funcția nucleu  $G$ :

$$\hat{f}_{h,G}(\mathbf{x}) = \frac{c_{g,d}}{nh^d} \sum_{i=1}^n g\left(\left\|\frac{\mathbf{x} - \mathbf{x}^i}{h}\right\|^2\right) \quad (4.10)$$

Termenul a doua din ecuația (4.9) reprezintă vectorul mean shift cu funcția nucleu  $G$ :

$$\mathbf{m}(\mathbf{x}) \equiv \frac{\sum_{i=1}^n \mathbf{x}_i g\left(\left\|\frac{\mathbf{x} - \mathbf{x}_i}{h}\right\|^2\right)}{\sum_{i=1}^n g\left(\left\|\frac{\mathbf{x} - \mathbf{x}_i}{h}\right\|^2\right)} - \mathbf{x} \quad (4.11)$$

și definește diferența dintre media ponderată folosind funcția nucleu  $G$ , și  $\mathbf{x}$ , centru ferestrei kernel. Vectorul  $\mathbf{m}(\mathbf{x})$  mai poate fi scris sub forma

$$\mathbf{m}(\mathbf{x}) = \frac{1}{2} h^2 c \frac{\hat{\nabla} f_{h,K}(\mathbf{x})}{\hat{f}_{h,G}(\mathbf{x})} \quad (4.12)$$

Expresia (4.12) indică faptul că la locația  $\mathbf{x}$ , vectorul mean shift estimat cu nucleul  $G$  este proporțional cu estimata gradientului densității obținută cu funcția nucleu  $K$  și invers proporțională cu densitatea la locația  $\mathbf{x}$ , obținută cu funcția nucleu  $G$ . Vectorul  $\mathbf{m}(\mathbf{x})$  indică direcția maximului din densitate. Din moment ce este aliniat cu estimata de gradient local, poate defini o cale spre un punct staționar al densității estimate. Procedura de meanshift constă din pașii:

- calculul vectorului de mean shift
- translația centrului ferestrei cu acest vector,
- procesul se repeta până când procedura converge spre un punct unde estimata are gradientul 0 [3].

#### 4.1.2. Selecția scării algoritmului

Scara (parametrul lățime de bandă  $h$ ) este un parametru important pentru performanța algoritmului mean shift. Dacă  $h$  tinde la 0, atunci numărul de maxime crește până la numărul datelor existente. Dacă  $h$  este prea mic se generează o estimată zgomotoasă, așa cum se ilustrează în figura 4.1, pentru un set de date experimentale ce au fost generate de două distribuții normale, cu mediile 0 și deviația standard 1. Dacă  $h$  tinde la infinit, numărul de maxime scade la 1, generând o estimare cu o formă unimodală (figura 4.2).

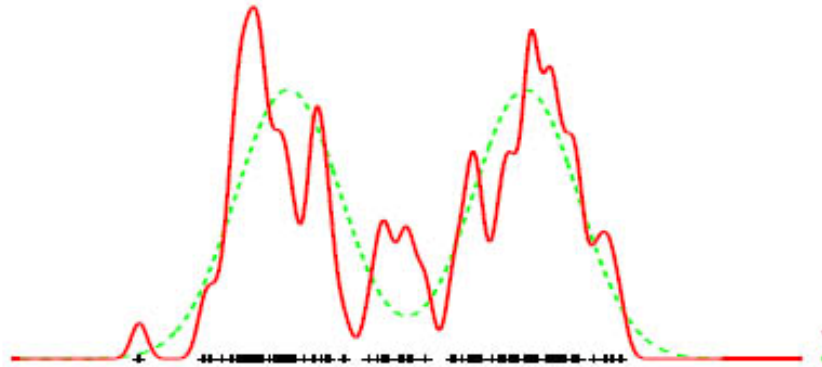


Fig. 4.1. Representarea scării kernel pentru  $h=0.01$  (lățime de bandă mică)[218].

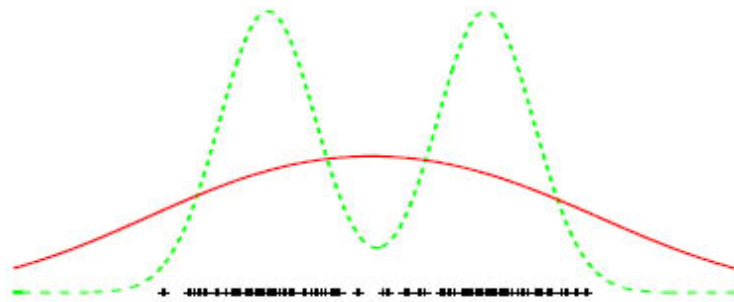


Fig. 4.2. Representarea scării kernel pentru  $h=0.2$  (lățime de bandă mare)[218].

În mod evident, este necesară dezvoltarea unor metode care să permită alegerea parametrului de scară pe baza datelor experimentale disponibile. O lucrare utilă în care sunt discutate numeroase soluții este [219]. Aici găsim diferite metode de selecție a scării  $h$ . Amintind de metodele plug-in, metodele de tipul "quick and dirty": *metoda empirică și principiul netezirii maxime*; metoda empirică și metoda netezirii maxime.

Metodele de selecție a scării orientate pe aplicație, caută un parametru de scară care asigură o estimată constantă pe o paletă cât mai largă de scări (figura 4.3). Stabilitatea unei estimate pentru scări diferite depinde de numărul de partiții generate în setul de date. Parametrul de scară se alege la mijlocul intervalului scărilor corespunzătoare unei partiții constante. În [3], o lucrare recentă a lui Comaniciu, se găsește această abordare.



## 4.2. Metoda robustă propusă de aliniere a imaginilor

În lucrarea [220] am propus o **metodă de aliniere a imaginilor robustă utilizând estimatorul neparametric cu deplasare la medie, mean shift**. Metoda folosește puncte caracteristice, iar extragerea lor s-a realizat, în lucrarea citată, interactiv în ambele imagini. Stabilirea corespondenței acestora se realizează cu ajutorul unei transformări geometrice 2D de similaritate care implică o rotație, translație și scalare și ai cărei parametri sunt supuși unui proces de estimare pentru a se determina estimata de maximă verosimilitate. Modelul de corespondență are la bază sistemul de ecuații:

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} s & 1 \\ 1 & s \end{bmatrix} \begin{bmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix} \quad (4.13)$$

Se asociază coordonatele vechi  $(x, y)$  cu cele noi  $(u, v)$ . Cei patru parametri ai transformării pot fi determinați din corespondența celor 2 seturi de puncte. Numărul corespondențelor posibile dintre perechile de puncte formate pentru a estima parametrii transformării,  $\varphi$ ,  $t_x$  și  $t_y$  este mult mai mare. Majoritatea dintre acestea pot fi eronate. Ca și în cazul estimatorului RANSAC, pentru fiecare subset minimal de puncte, o soluție candidat este generată.

Prin notarea vectorului de parametri,

$$\mathbf{p} = \begin{bmatrix} s \\ \varphi \\ t_x \\ t_y \end{bmatrix} \quad (4.14)$$

problema estimării transformării geometrice poate fi formulată ca o problemă de minimizare a erorii de asociere a datelor existente,

$$\mathbf{p} = \underset{\mathbf{p}}{\operatorname{argmin}} \sum_i \rho(r_i) \quad (4.15)$$

unde reziduurile  $r_i$  reprezintă estimări ale erorii de aliniere dintre perechile de puncte corespondente după realizarea corespondenței. Prin

$$\rho(r) = r^2 \quad (4.16)$$

se obține o aliniere least-squares (metoda celor mai mici pătrate LS). Această problemă a fost intens studiată în problema de stabilire a corespondenței punctelor. Se poate urmări, de exemplu, studiul din [14] cu îmbunătățiri din [221]. Din cauza ridicării la pătrat din ecuația (4.16), alinierea prin metoda celor mai mici pătrate este foarte sensibilă în prezența erorilor, datele incorecte putând devia soluția cu mult de la valoarea corectă. Problema poate fi rezolvată prin utilizarea unei alte funcții, în ecuația (4.15) cu scopul de a reduce influența eșantioanelor aberante din datele existente. Estimatorii M [172] sunt cele mai evidente exemple din această categorie.

În această lucrare, se utilizează o abordare de estimare a densității, bazată pe estimatorul mean shift. Abordarea este strâns legată de estimatorii M, după cum reiese din [3]. Definind

$$\rho(r) = -\log p(r) \quad (4.17)$$

și considerând reziduurile  $r_i$  independent și identic distribuite, se obține un estimator de maximă verosimilitate. În absența unor modele realiste și general valabile pentru distribuția reziduurilor  $r_i$ , este firesc să se opteze pentru o modelare neparametrică a densității. Soluția de verosimilitate maximă se poate obține eficient folosind algoritmul mean shift.

Pornind de la orice eșantion, un maxim al funcției de densitate de probabilitate poate fi obținut prin iterarea ecuației

$$\mathbf{y}_{j+1} = \frac{\sum_{i=1}^n \mathbf{x}_i g\left(\left\|\frac{\mathbf{y}_j - \mathbf{x}_i}{h}\right\|^2\right)}{\sum_{i=1}^n g\left(\left\|\frac{\mathbf{y}_j - \mathbf{x}_i}{h}\right\|^2\right)}, \quad j = 1, 2, \dots \quad (4.18)$$

În estimarea parametrilor transformării geometrice pentru alinierea imaginilor,  $\mathbf{y}_j$  este inițializat de vectorul parametrilor generat de un set de date minimal oarecare, iar  $\mathbf{x}_i$  reprezintă ceilalți vectori parametru generați. Parametrii pot fi estimați simultan, folosind forma vectorială a ecuației (4.18) sau secvențial. Pentru estimarea secvențială, se pornește de la observația că unghiurile dintre segmentele de linie nu sunt influențate de către translație sau rescalare. În consecință, parametrul de rotație poate fi estimat cu ajutorul unghiurilor înainte de estimarea parametrilor de translație sau rescalare. Estimarea parametrului de rescalare poate de asemenea să fie făcută înainte de estimarea translației sau rotației, bazată pe distanțele între puncte. Pe de altă parte, rotația sau rescalarea afectează puternic parametrii de translație, așa cum este ilustrat în exemplu din figura 4.3.

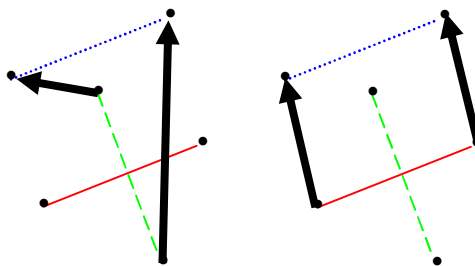


Fig. 4.3. Vectorii de translație înainte (stânga) și după (dreapta) rotație.

Din exemplu de mai sus, este clar că dacă se dorește un vector de translație robust trebuie să se realizeze *după* ce rotația și rescalarea au fost estimate și compensate. Se utilizează  $\{V_i\}$ ,  $i = 1, 2, \dots, N$  un set de puncte din imaginea de

referință și  $\{Q_i\}, i = 1, 2, \dots, N$  punctele corespondente din imaginea captată. Pentru estimarea unghiului de rotație, se alege un set minimal, adică perechile de puncte  $(Q_i, Q_j)$  și  $(V_i, V_j)$ , având vectorii de corespondență  $\mathbf{q}_{ij}$  and  $\mathbf{v}_{ij}$ . Unghiul dintre liniile  $(Q_i, Q_j)$  și  $(V_i, V_j)$  este dat de ecuația:

$$\cos(\varphi) = \frac{\mathbf{q}_{ij}^T \mathbf{v}_{ij}}{\|\mathbf{q}_{ij}\| \|\mathbf{v}_{ij}\|} \quad (4.19)$$

Se notează  $\{\varphi_i\}, i = 1, 2, \dots, M$ , unde  $M$  este setul de unghiuri obținute de către perechile de puncte. Fie  $MS(\varphi_i)$  locația maximumului de densitate găsit de algoritmul mean shift inițializat cu unghiul  $\varphi_i$ . Fie  $\hat{p}(MS(\varphi_i))$  densitatea în punctul de convergență, proporțională cu numitorul expresiei (4.18). Unghiul de rotație estimat este

$$\hat{\varphi} = \underset{MS(\varphi_i)}{\operatorname{argmax}} \{\hat{p}(MS(\varphi_i))\} \quad (4.20)$$

Într-o manieră similară, factorul de scalare estimat poate fi obținut din setul de perechi de puncte folosind ecuația:

$$s = \|\mathbf{v}_{ij}\| / \|\mathbf{q}_{ij}\| \quad (4.21)$$

și folosind algoritmul de mean shift în aceeași manieră ca și estimarea unghiului.

După efectuarea transformării geometrice pentru a se compensa unghiul de rotație și scalare, estimarea vectorului robust de translație este realizată de către corespondența punctelor. Fiind dată o pereche de puncte cu vectorii de poziție  $\mathbf{v}_i$  și  $\mathbf{q}_i$ , se formează parametrii de translație

$$t_x = v_{xi} - q_{xi} \quad (4.22)$$

$$t_y = v_{yi} - q_{yi} \quad (4.23)$$

apoi se va trece la estimarea parametrilor de translație folosind mean shift la fel cum s-a folosit pentru estimarea unghiului și scalării.

Există două abordări de determinare a parametrilor transformării geometrice. Prima se referă la obținerea de densități ridicate prin reducerea dimensiunii spațiului de căutare. Se propune o soluție bazată pe spațiul de căutare 1D, se pot utiliza relațiile (4.19), (4.21), (4.22), (4.23). Pe baza acestor relații, se realizează estimarea secvențială a parametrilor transformării geometrice. O altă soluție propusă este aceea care utilizează spațiul de căutare 4D, se poate utiliza relația (4.13), pentru estimarea simultană a parametrilor transformării geometrice.

Avantajul estimării separate îl reprezintă lucrul cu densități mari, prin reducerea dimensiunii spațiului de căutare. Totuși, experimentele cu imaginile 3D pe care le-am efectuat în cadrul tezei, descrise în Capitolul 5, au condus la concluzia că

este mai avantajoasă estimarea simultană a celor patru parametri ai transformării de similaritate.

Cu scopul de a se compara cele două transformări de similaritate în vederea performanței acestora pe un set de imagini 2D, și de a se evalua câștigul în precizia obținută din estimarea abordării simultane, în capitolul 5 se vor prezenta și discuta rezultatele experimentale comparative ale celor două transformări de similaritate.

### **4.3. Metoda de asociere prin compararea punctelor mediane**

Utilizând transformarea de similaritate simultană s-a trecut la adoptarea unei noi idei, care își propune creșterea acurateții estimării prin eliminarea eșantioanelor aberante dar și scăderea timpului de calcul. În [193], [197] s-a propus **o nouă metodă de corespondență care realizează asocierea punctelor caracteristice utilizând informația de culoare a pixelilor de pe medianele segmentelor ce unesc perechi de puncte caracteristice detectate.**

Metodele tradiționale se bazează pe măsurarea unei informații extrasă dintr-o vecinătate de puncte caracteristice și comparării ei, cu scopul de a elimina potrivirile eronate. Cea mai simplă informație poate fi intensitatea pixelului sau culoarea. Cu toate acestea, fiindcă punctele caracteristice tipice, cum ar fi colțurile, sunt localizate în regiuni cu schimbări rapide, erori mici de poziționare pot rezulta în variații mari în informația vecină. În plus, diferența de scară face ca problema să fie și mai severă. Ideea de bază a abordării noastre este de a compara informația extrasă din pixelii localizați la jumătatea distanței dintre perechile de puncte. De obicei, aceste puncte sunt localizate în zone omogene, în consecință sunt mai puțin afectate de pozițiile exacte ale punctelor caracteristice. Mediana unui segment de linie este invariantă la translație, rotație și rescalare și teoretic informația de culoare a punctelor mediane este de asemenea invariantă la transformările de scară, translație și rotație.

Ideea dezvoltată este de a compara informația imaginii extrasă din pixeli aflați la mijlocul distanței dintre perechile de puncte (figura 4.9). Cea mai simplă metodă de similaritate se referă numai la culorile perechilor pixelilor situați în aceeași poziție în ambele imagini. Informația de culoare este foarte ușor de extras, ținând costul de procesare la un nivel minim. De fapt, din moment ce punctele mediane aparțin zonelor omogene, această informație caracterizează complet vecinătatea. Pentru a studia această abordare, o comparăm în capitolul 5, conținând partea experimentală a tezei, cu abordările tradiționale, bazate pe măsurări în punctele caracteristice.

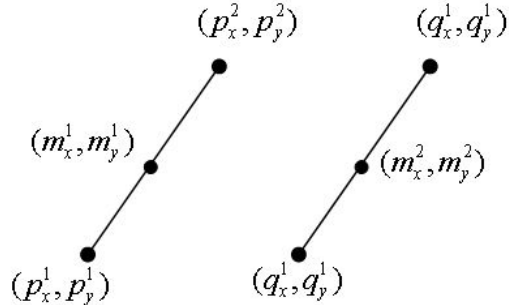


Fig. 4.4. Reprezentarea punctelor caracteristice și a punctului median.

Metoda propusă se numește **metoda de asociere a punctelor mediane** și calculează punctul median  $m_i^j, i = 1, 2, j = 1, 2$  dintre două perechi de puncte  $p_i$  și  $q_i$ . Această metodă utilizează o normă  $D(m_1, m_2)$  care ulterior va fi comparată cu un prag  $T$ . Pentru a calcula norma este nevoie de a găsi informația de culoare a acestor puncte:

$$D(\mathbf{m}_1, \mathbf{m}_2) = \|\mathbf{c}(\mathbf{m}_1) - \mathbf{c}(\mathbf{m}_2)\|^2 \quad (4.24)$$

unde  $\mathbf{c}(\mathbf{m}_i)$  sunt vectorii de culoare ai punctelor  $\mathbf{m}_i$ . O asociere este considerată validă dacă

$$D(\mathbf{m}_1, \mathbf{m}_2) < T \quad (4.25)$$

unde  $T$  reprezintă un prag. Dacă norma diferențelor culorilor este mai mare decât un prag atunci perechea considerată necorespondentă este eliminată.

Dintre metodele tradiționale cu care va fi comparată metoda mai sus menționată, face parte metoda ce utilizează media culorilor pixelului median. Un pixel are opt vecini. Pentru fiecare pixel, vectorii de culoare ai vecinilor săi sunt însumați și divizați cu nouă. La urmă, norma acestor diferențe între informația de culoare este calculată cu ecuația (4.24).

A doua metodă tradițională utilizează culorile punctelor  $p_i = (p_x^i, p_y^i)$  și  $q_i = (q_x^i, q_y^i)$ . Acum vom avea două ecuații ale normei:

$$D(\mathbf{p}_1, \mathbf{q}_1) = \|\mathbf{c}(\mathbf{p}_1) - \mathbf{c}(\mathbf{q}_1)\|^2 \quad (4.26)$$

$$D(\mathbf{p}_2, \mathbf{q}_2) = \|\mathbf{c}(\mathbf{p}_2) - \mathbf{c}(\mathbf{q}_2)\|^2 \quad (4.27)$$

unde  $\mathbf{c}(\mathbf{p}_i)$ ,  $\mathbf{c}(\mathbf{q}_i)$ , sunt vectorii de culoare pentru fiecare punct  $p_i = (p_x^i, p_y^i)$  și  $q_i = (q_x^i, q_y^i)$ . Pentru ultimele două metode, procesul de comparare a normei cu un prag este identic cu cel din prima metodă, numai rezultatele sunt diferite.

Metodele de asociere a punctelor caracteristice folosesc toate combinațiile între perechile de puncte pentru a genera soluții și folosesc metoda robustă mean shift pentru a găsi cea mai bună estimată dintre soluțiile parțiale. *Găsirea unui prag optim* care să elimine posibilele nepotriviri dintre perechile de puncte, cu scopul de a crește acuratețea estimării și de a reduce timpul de procesare, constituie studiul publicat în articolul [195]. Rezultatele experimentale sunt prezentate în capitolul 5.

#### 4.4. Utilizarea algoritmului de aliniere a imaginilor în rețele de senzori video fără fir

Tehnica de aliniere prezentată mai sus poate fi utilizată în *localizarea nodurilor dintr-o rețea de senzori fără fir*. Studiul în acest domeniu este descris pe larg în [181] și [191]. Prima lucrare prezintă o metodă eficientă de localizare împreună cu estimarea câmpului vizual comun pentru aplicațiile de supraveghere care folosesc corespondența imaginilor cu scopul de a alinia imaginile capturate simultan de la diferiți senzori video.

Senzorii unei rețele fără fir fac adesea parte dintr-o infrastructură în care sunt amplasate într-o manieră ad-hoc, drept urmare nu se cunoaște locația lor. Problema estimării coordonatelor spațiale ale unei rețele poartă numele de localizare.

În literatura de specialitate există 2 metode propuse pentru localizarea nodurilor: metodă de localizare bazată pe un punct de referință și metode de localizare bazate pe timp sau forța semnalului. În contextul unei rețele cu senzori video, problema localizării este complexă. Pe lângă coordonatele nodurilor, informația privind topologia rețelei implică și unghiurile dintre camere dar și câmpul vizual comun al camerelor (figura 4.5).

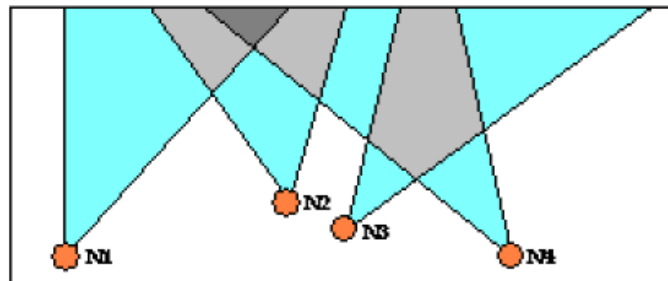


Fig. 4.5. Câmpurile vizuale comune ale camerelor.

Se poate afirma că suprapunerea câmpurilor vizuale conține suficientă informație de localizare pentru majoritatea aplicațiilor de supraveghere.

Algoritmul de aliniere prezentat mai sus este folosit pentru localizarea senzorilor dintr-o rețea fără fir dar și la determinarea câmpului vizual comun al

camerelor. Se pornește de la observația că unghiul dintre două linii de segment nu este schimbat de translație și rotație. De aceea, parametrul de rotație  $\varphi$  poate fi estimat și anterior parametrilor de translație și scalare. Parametrii de scalare pot fi

de asemenea estimați anterior parametrilor de translație și rotație. Componentele vectorului de translație trebuie să fie calculate după calcularea parametrilor de scalare și rotație. Ca și estimatorul RANSAC pentru a estima parametrii este necesar de un set minimal de date de intrare.

După aplicarea transformatei inverse geometrice pentru a compensa scara și unghiul de rotație, estimarea robustă a vectorului de translație este efectuată prin corespondența punctelor. Fiecare pereche de puncte generează un set de parametrii de translație (translația pe orizontală și translația pe verticală). După un număr de  $n$  combinații, vor rezulta doi vectori, fiecare cu  $n$  parametrii de translație. Algoritmul robust mean shift este utilizat pentru a găsi cea mai bună estimată pentru vectorul de translație pe  $x$  și pentru vectorul de translație pe  $y$ .



(a)





(b)



(c)

Fig. 4.6. Transformarea imaginii bazată pe estimarea unghiului de rotație: (a) Imaginea originală din senzorul 1; (b) Imaginea originală din senzorul 2; (c) Imaginea din senzorul 2 după transformare.



Algoritmul propus este totuși afectat de erori (între 1 și 5 pixeli) specifice procedurii de selecție a punctelor. Pentru a trata aceste erori, este necesar de o etapă de post-procesare. În [181] se propune asocierea chamfer. Aceasta este o tehnică de asociere a muchiilor, fiind propusă de Barrow și colaboratorii [222] și revizuită ulterior de Borgefors [192]. Metoda implică calculul unei hărți de distanțe începând cu colțul din dreapta sus. Tiparul care se caută este apoi mutat peste harta de distanțe. Sub acțiunea gravității, tiparul se va muta până găsește altitudinea cea mai joasă posibilă. Dacă altitudinea este 0 sau apropiată de 0, se consideră că s-a găsit un tipar optim. Tiparul se găsește acolo unde corelația atinge un minim absolut. Această metodă se bazează pe o piramidă de rezoluție, asocierea făcându-se nu numai în imaginea originală dar și într-o serie de imagini care reprezintă scena la cea mai mică rezoluție. Minimele locale care sunt prea mici se resping. Pozițiile care rămân sunt luate în considerare la următorul nivel al piramidei și procedura se repetă până când un minim local este găsit la cel mai mare nivel de rezoluție al piramidei. Cu toate acestea, datorită corecției mici – între 1 și 5 pixeli – pentru cazul nostru, nu se justifică utilizarea potrivirii ierarhice.



(a)



(b)

Fig. 4.7. Post-procesare prin asociere Chamfer: (a) Diferența imaginilor înainte de post-procesare; (b) Diferența imaginilor după de post-procesare.

În [191] se continuă studiul posibilității utilizării metodelor de aliniere a imaginilor în cazul localizării senzorilor într-o rețea de senzori fără fir. Se pune accentul pe estimarea simultană (rezolvarea sistemului de ecuații 4.13) și secvențială (ecuațiile 4.19, 4.21, 4.22, 4.23), scopul fiind de a evalua câștigul în precizia obținută din estimarea abordării simultane.

Pe baza imaginilor captate de senzorii video și pe baza punctelor caracteristice alese în fiecare pereche de imagini, metoda robustă shift este folosită pentru a găsi cea mai bună estimată dintre parametrii de translație, rotație și scalare. O etapă finală constă din utilizarea acestor estimate pentru a alinia optim imaginile (figura 4.8).



(a)



(b)



(c)

Fig. 4.8. (a) Imaginea de la Nodul 1; (b) Imaginea de la Nodul 2;  
(c) Imaginea de la Nodul 2 după realizarea procesului de corespondență.

#### 4.5. Discuție teoretică

Obiectivul principal al cercetărilor prezentate în acest capitol a fost acela de a studia utilitatea tehnicilor neparametrice de estimare a densității de probabilitate folosind funcții nucleu în alinierea imaginilor. Cu toate că se cunoaște caracterul lor robust și legătura cu estimatorii  $M$ , acești estimatori nu au mai fost utilizați în aplicații de aliniere a imaginilor, probabil datorită popularității algoritmilor din familia RANSAC, care au fost inventați chiar cu scopul de a fi utilizați în acest tip de probleme.

Problema centrală în estimarea robustă constă în rezolvarea dihotomiei datelor conforme modelului/neconforme modelului estimat. În acest scop, algoritmul RANSAC în forma lui de bază, utilizează un prag  $T$ . Alegerea optimală a pragului ar trebui să beneficieze de o metodă de evaluare a distribuției reziduurilor de estimare pentru aplicația în cauză. Recunoscând existența problemei, o asemenea metodă este propusă în varianta ameliorată, MLESAC. Denumirea metodei evidențiază faptul că în clasificarea datelor conforme/ neconforme, se apelează la o tehnică de maximă verosimilitate. Limitarea estimatorului MLESAC, evidențiată de studiile ulterioare, constă în ipoteza restrictivă că datele conforme au o distribuție normală, în timp ce datele neconforme au o distribuție uniformă. Deși validitatea modelului de mixtură propus pentru aplicațiile descrise de autori este dincolo de orice îndoială, generalizarea pentru toate aplicațiile este dubioasă și conduce uneori la rezultate modeste. Modelele de mixtură sunt o clasă extrem de bogată și fac parte din categoria celor ce sunt estimate cu ajutorul metodelor semi-parametrice. Probleme majore ale acestei abordări sunt stabilirea corectă a numărului de componente și a modelelor individuale ale componentelor.

În contrast cu metodele semi-parametrice, metodele de estimare neparametrice acceptă faptul că, în unele aplicații, distribuția datelor este prea complicată pentru a putea fi modelată cu precizia necesară prin metode parametrice sau semi-parametrice. În asemenea situații, estimarea neparametrică devine o alternativă demnă de interes. Teza prezentă și-a propus să studieze în ce măsură și în ce tip de aplicații, metodele de estimare neparametrice pot fi avantajoase în raport cu metodele robuste utilizate curent în alinierea imaginilor. O asemenea aplicație, identificată și implementată în cadrul tezei este alinierea imaginilor în rețele de senzori video fără fir, în vederea determinării câmpului de vedere comun și a orientării relative a senzorilor, o problemă ce suscită un interes acut, în contextul conflictelor militare și a amenințărilor teroriste de dată recentă.

Metodele din familia RANSAC, indiferent de variantă, *selectează* una din soluțiile ipotetice generate pe baza seturilor minimale de date. În stabilirea soluției câștigătoare, variantele MSAC și MLESAC utilizează criterii de performanță nuanțate, dar în final soluția este tot una dintre soluțiile generate. Când numărul de puncte caracteristice detectate, disponibile pentru estimarea transformării, este redus, când câmpul vizual comun al celor două imagini implicate în estimare este redus, sau când punctele sunt localizate cu erori mari, aceste metode sunt în dificultate evidentă, așa cum o confirmă și experimentele descrise în Capitolul 5. Toți acești factori sunt prezenți în problema alinierii imaginilor captate de rețelele de senzori fără fir, lansate în teren mai mult sau mai puțin la întâmplare. În particular, distanțele diferite față de obiectele ce generează puncte de interes conduc la reprezentări de scări diferite și în consecință la erori de localizare mai mari.

Avantajele potențiale ale estimatorilor neparametrici constau în flexibilitate în raport cu distribuțiile reale ale datelor, absența unor praguri de decizie critice și

dificil de selectat automat și sensibilitatea relativ redusă la forma funcției nucleu. Selecția parametrului de scară al acestei funcții rămâne un aspect important în stabilirea performanțelor, deși mai puțin critic decât pragul utilizat de estimatorul RANSAC. Prin medierea ponderată pe care o operează algoritmul mean shift pentru estimarea locației modurilor densității de probabilitate (model continuu), soluțiile obținute prin metoda propusă în cazul aplicațiilor cu puncte caracteristice puține și erori mari sunt clar mai bune decât cele obținute prin metodele cunoscute, așa cum se demonstrează prin experimentele extensive descrise în Capitolul 5.

Spre deosebire de algoritmi din familia RANSAC, care evaluează soluțiile pe baza unor criterii ce folosesc distanțe măsurate în spațiul coordonatelor, metoda propusă operează în spațiul parametrilor geometrici ai transformării, în care estimează locația de densitate maximă. În general, acest spațiu este de dimensiune mai mare decât spațiul coordonatelor, având un potențial mai mare de separare a datelor conforme de cele neconforme. Un dezavantaj potențial al metodei de estimare a parametrilor transformării geometrice folosind algoritmul mean shift constă în necesitatea de a genera un număr mai mare de soluții corecte pentru a se obține rezultate performante. În experimentele noastre, acest dezavantaj nu s-a manifestat altfel decât prin calcule mai intensive.

## 4.6. Concluzii

În acest capitol **am elaborat o metodă de aliniere a imaginilor, bazată pe algoritmul mean shift. Metoda calculează o soluție de maximă verosimilitate, prin maximizarea densității de probabilitate în spațiul parametrilor, spre deosebire de metodele de referință, ce folosesc optimizări în spațiul coordonatelor. Algoritmul mean shift folosit în localizarea maximelor densității de probabilitate a vectorului parametrilor este fundamentat teoretic pe estimări neparametrică, folosită în premieră în alinierea imaginilor.**

În teză **am propus și comparat două versiuni de estimare a parametrilor transformării de similaritate prin algoritmul mean shift: secvențial și simultan. Am stabilit și justificat teoretic ordinea de estimare cea mai avantajoasă la estimatorul secvențial.**

**Am propus o nouă metodă pentru selecția punctelor caracteristice și eliminarea corespondențelor incompatibile prin metoda de asociere a punctelor mediane, folosind similaritatea culorilor și un prag de decizie. Am analizat teoretic avantajele potențiale ale noii abordări, comparativ cu soluțiile de referință și identifică aplicațiile ce pot beneficia de aceste avantaje.**

**Am definit ca direcție de aplicabilitate pentru metoda originală de aliniere a imaginilor utilizând estimatorul robust bazat pe estimarea neparametrică de densitate de probabilitate, localizarea senzorilor în rețelele de senzori video fără fir. Procesul de aliniere a imaginilor captate de la întreaga rețea de senzori, este utilizat atât pentru determinarea localizării acestora, cât și pentru determinare câmpului vizual comun dintre perechile de camere.**

**Teza identifică aspectele teoretice pentru care algoritmi din familia RANSAC/MLESAC sunt mai puțin performanți în problemele de aliniere a imaginilor cu puncte caracteristice puține și erori de localizare mari.**

## 5. STUDIU EXPERIMENTAL

În acest capitol se prezintă rezultatele experimentelor efectuate pentru evaluarea performanțelor metodelor de aliniere a imaginilor propuse în capitolul precedent. S-au studiat performanțele metodei de estimare robustă a parametrilor transformării bazate pe algoritmul mean shift în diferite situații. Pentru evaluarea obiectivă a performanțelor metodei robuste de estimare propuse s-au testat pe aceleași seturi de date și metodele de referință RANSAC și MLESAC.

În paragraful 5.1 se prezintă detaliile de implementare a algoritmului de aliniere a imaginilor. În paragraful 5.2 se urmărește realizarea alinierii imaginilor utilizând algoritmul robust mean shift pentru estimarea robustă a parametrilor transformării geometrice care aliează optim imaginile. Utilizarea estimatorului robust bazat pe estimarea densității de probabilitate și algoritmul mean shift pentru problemele de localizare într-o rețea de senzori video fără fir este prezentată în paragraful 5.3. În paragraful 5.4 se testează obținerea unui timp de calcul minim și a unei acurateți de estimare ridicată, urmată de îmbunătățirea detecției caracteristicilor în paragraful 5.5. Paragrafele 5.6-5.7 conțin un studiu comparativ între metoda propusă mean shift și metodele RANSAC și MLESAC. Concluziile finale sunt prezentate în capitolul 5.8.

### 5.1. Detalii de implementare

Pentru a testa performanțele metodelor robuste prezentate în capitolul 4, se folosesc perechi de imagini conținând un câmp vizual obținut din diferite poziționări și orientări ale aparatului de fotografiat.

Cu privire la câmpul vizual captat, în general, distingem trei situații familiare. Prima este reprezentată de scenele din natură, de exemplu, păduri, deșert; a doua se referă la peisajele citadine, precum piețe, intersecții; cea de a treia situație este reprezentată de spațiile închise: camere, interioare. Datorită acestui fapt experimentele au fost realizate folosind imagini ce cuprind toate cele trei ipostaze. Drept urmare, imaginile au fost împărțite în trei categorii: natură, interior și peisaj urban.

Imaginile achiziționate prezintă două tipuri de deformări geometrice. Primul set (figura 5.1) este format din 12 imagini achiziționate cu aparatul de fotografiat reprezentând imaginile de referință. Imaginile țintă sunt reprezentate de imagini generate artificial pe baza imaginilor de referință, folosind transformări geometrice cu parametrii (de rotație, translație și scalare) cunoscuți.



5.1. Detalii de implementare 83

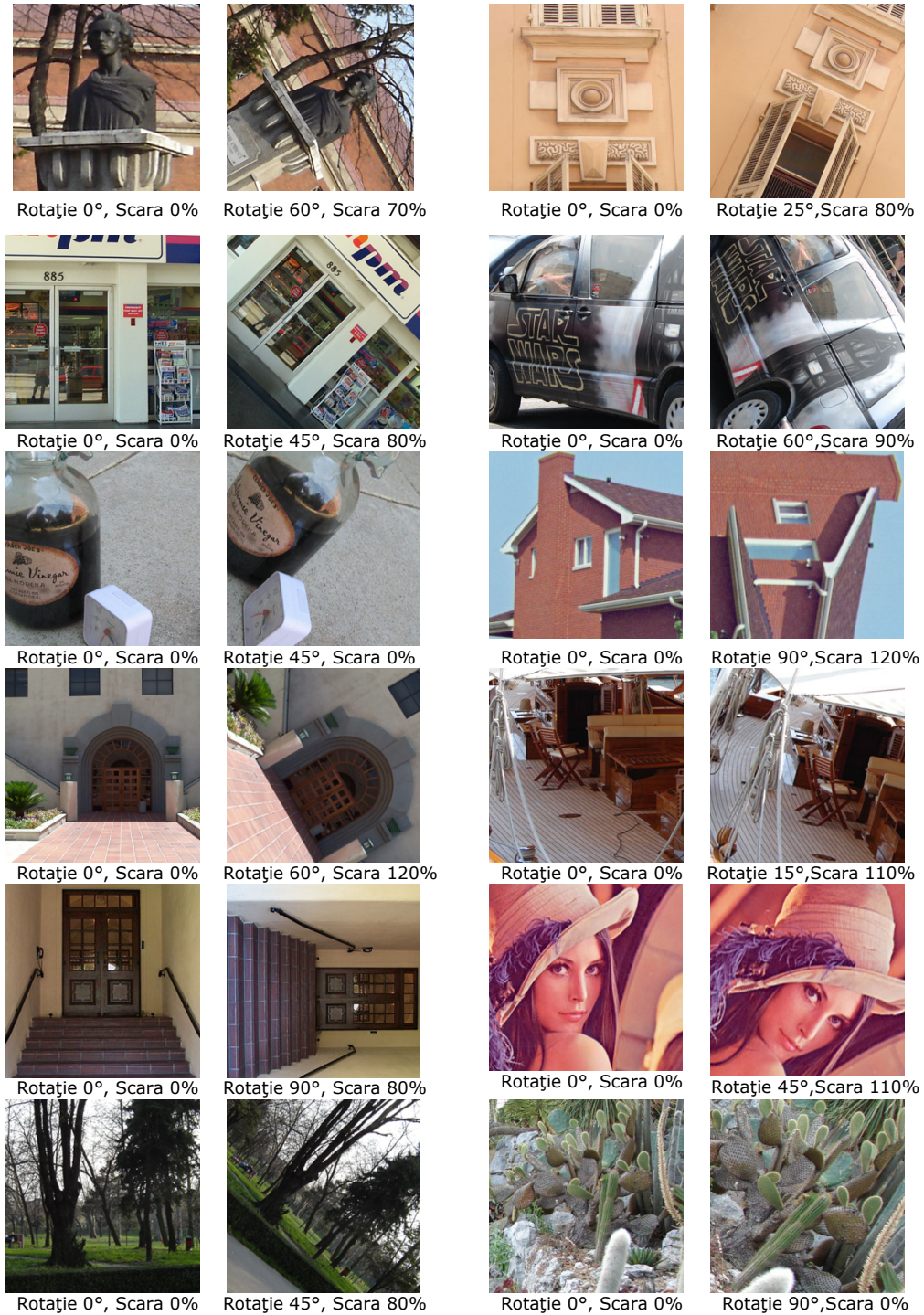


Fig. 5.1. Primul set de imagini folosit în procesul de aliniere.

Al doilea set de imagini realizat cu același aparat de fotografiat, prezintă transformări geometrice introduse prin captarea din poziții diferite a scenelor 3D. Datorită cantității numeroase de imagini care au fost folosite pentru realizarea experimentelor, în figura 5.2 sunt prezentate doar câteva imagini din fiecare categorie.

Pentru fiecare imagine, selecția caracteristicilor a fost făcută atât automat cât și manual, pentru un număr diferit de puncte caracteristice. Aceste puncte sunt formate dintr-un număr necunoscut de puncte false (puncte ce nu au corespondent) și puncte corespondente.

Fiecare metodă a fost testată pe cele două seturi de imagini. Primul set conține 12 imagini, setul doi conține 6 imagini din fiecare categorie (6 imagini  $\times$  3 categorii). În total există un număr de 30 de imagini pe care se realizează experimentele.





5.1. Detalii de implementare 85





Fig. 5.2. Al doilea set de imagini folosit în procesul de aliniere.

Pentru a se putea testa metodele robuste discutate în capitolele anterioare, s-a implementat o aplicație de punere în corespondență a imaginilor în limbajul de programare C++. Aplicația conține implementate cele trei metode robuste, mean shift, RANSAC și MLESAC, precum și cei patru pași de parcurgere ai metodei de punere în corespondență a imaginilor.

Pentru realizarea procesului de aliniere al imaginilor, se încarcă imaginea de referință și imaginea țintă, precum arată și imaginea 5.3.

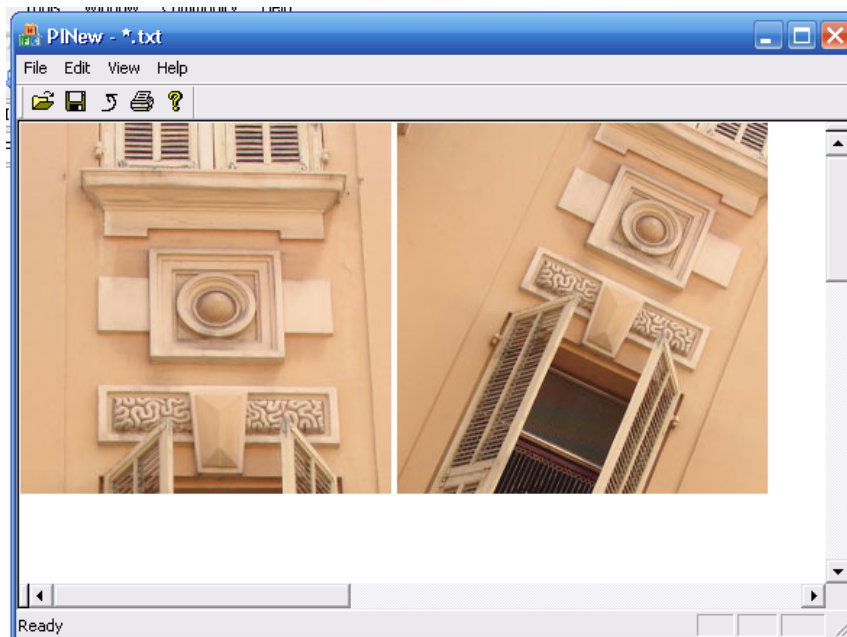


Fig. 5.3. Captură de ecran cu imaginile care vor fi supuse procesului de aliniere.

Selecția caracteristicilor se face în paralel. Poate fi fie manual, punctele caracteristice fiind puse cu mouse-ul de către utilizator, fie automat utilizându-se un program de selecție automată a caracteristicilor (figura 5.4).

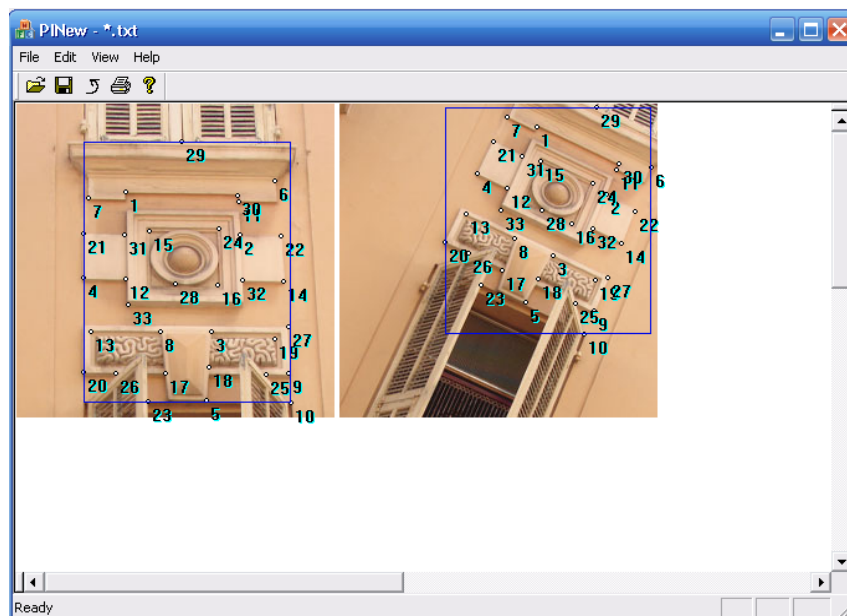


Fig. 5.4. Reprezentarea selecției caracteristicilor.



Pentru stabilirea corespondenței punctelor caracteristice, se pornește de la alegerea de seturi minimale de date, pentru fiecare set calculându-se parametrii transformării de similaritate (translație, rotație și scalare), aceștia fiind memorați într-un tablou unidimensional.

Estimarea robustă a parametrilor (figura 5.5) se realizează diferit în funcție de metoda robustă de estimare utilizată.

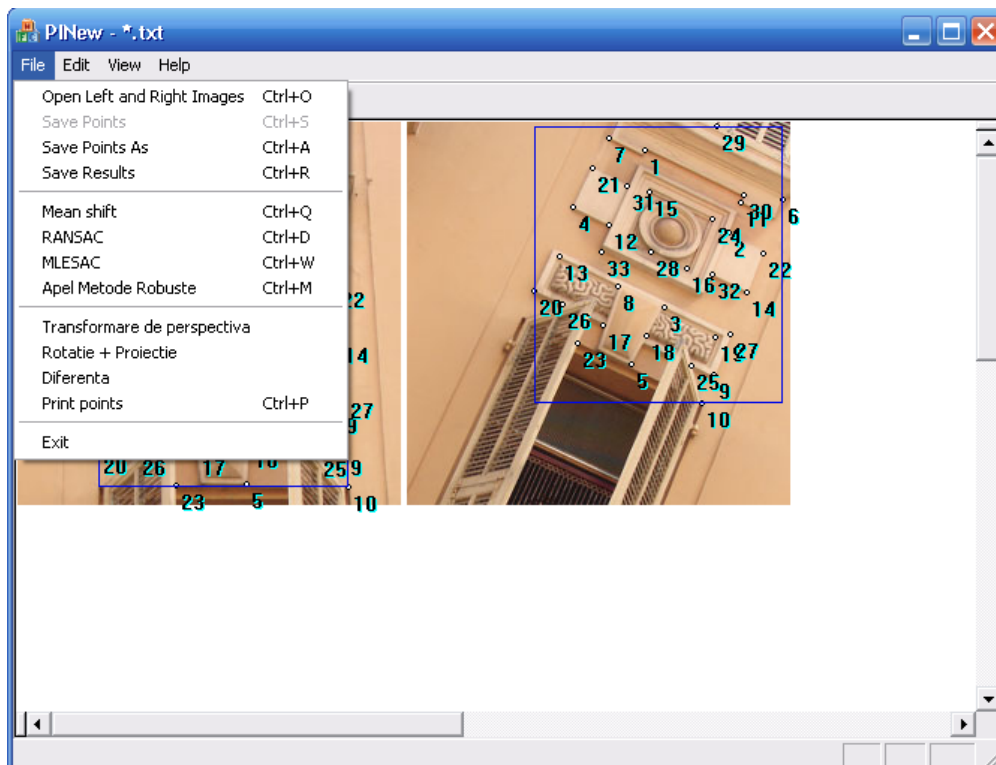


Fig. 5.5. Apelarea metodelor de estimare robuste.

Cu parametrii astfel estimați, se va aplica transformarea de similaritate pe cea de a doua imagine (dreapta). După cum se poate observa în figura 5.6, aplicația găsește partea comună a celor două imagini, o rotește și o translatează, astfel încât să se alinieze cu prima imagine (stânga).

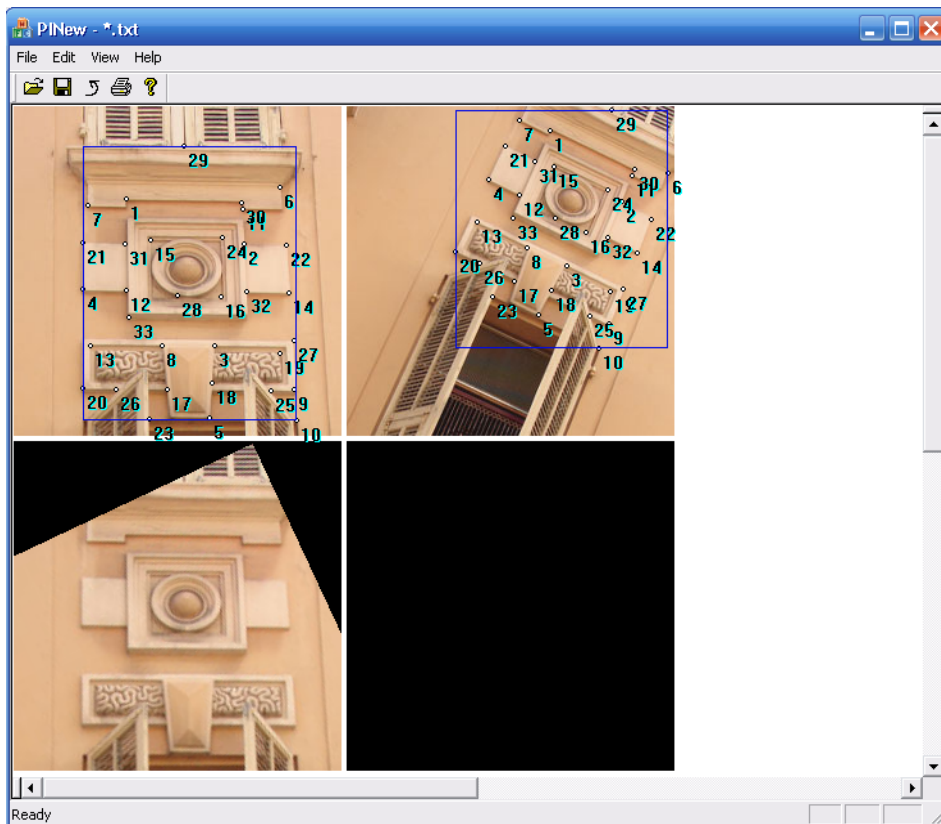


Fig. 5.6. Rezultatul procesului de aliniere a imaginilor.

## 5.2. Alinierea 2D a imaginilor utilizând metoda robustă meanshift

În acest paragraf prezentăm rezultatele experimentale din [220] care compară metoda propusă cu metoda least-squares. Metoda de aliniere utilizată folosește puncte caracteristice, iar extragerea lor s-a realizat interactiv în ambele imagini. Asocierea acestora se realizează cu ajutorul transformării de similaritate care implică o scalare, rotație și translație și ai cărei vectori cu parametri sunt supuși unui proces de estimare a densității, pentru a determina estimata de maximă verosimilitate.

Cu scopul de a se obține o evaluare calitativă a utilizării metodei robuste mean shift în procesul de aliniere, s-au generat imagini artificiale având parametrii transformării geometrice cunoscute. Punctele caracteristice au fost selectate manual în ambele imagini. Pentru referință, un estimator least squares a fost utilizat cu scopul de a se valida performanțele abordării propuse pentru erori mici, unde estimatorul least squares este cel mai bun. Rezultatele, în cazul transformării de similaritate, sunt formate dintr-o translație și o rotație cu  $45^\circ$  după cum reiese din figura 5.7. Imaginea originală se poate vedea în figura 5.7a, în timp ce imaginea transformată se poate vedea în figura 5.7b. În imaginea 5.7c sunt prezentate

rezultatele metodei robuste de aliniere propusă în aceasta lucrare pentru o pereche de imagini din figurile 5.7a și 5.7b. Aceleași rezultate pentru corespondența least-squares sunt ilustrate în figura 5.7d. În figura 5.7e și 5.7f sunt prezentate erorile de asociere pentru metoda de aliniere și pentru cea de least-squares.

Erorile au fost evaluate doar în zona minimă dreptunghiulară care cuprinde punctele caracteristice folosite pentru corespondență, erorile fiind prezentate în contrast negativ pentru o mai bună vizibilitate. O examinare mai atentă a erorilor din imagini din figura 2e și 2f dezvăluie prezența unor erori mari pentru estimatorul de least-squares comparat cu estimatorul robust mean shift.

Cu scopul de a se obține o evaluare bună a performanțelor tehnicii propuse de punere în corespondență a imaginilor, în a doua serie de experimente, am generat seturi 1D de date cu un procent controlat de erori. Atât datele neconforme cât și datele conforme au fost generate ca secvențe uniform distribuite. Datele conforme au fost generate cu o deviație standard de 10, în timp ce datele neconforme au fost generate cu o deviație standard 1.

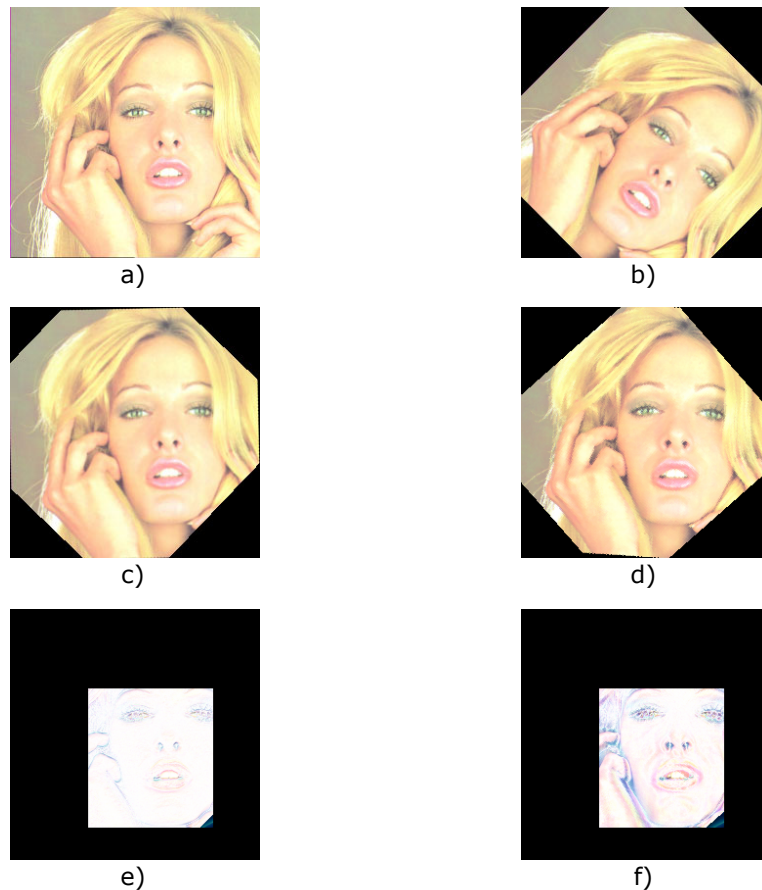


Fig. 5.7. a) Imaginea de referință; b) Imaginea țintă; c) Imaginea aliniată cu metoda robustă mean shift; d) Imaginea aliniată cu estimatorul least-squares; e) Eroarea de asociere pentru metoda robustă mean shift; f) Eroarea de asociere pentru estimatorul least-squares.

Rezultate comparative pentru parametri estimați cu mean shift și least-squares sunt prezentate în figurile 5.8 și 5.9, pentru 10% și respectiv 33% procentaj de erori. Estimatorul mean shift a fost implementat cu ajutorul nucleului Epanechnikov și cu un parametru de scară  $h = 2$ . Așa cum era de așteptat teoretic, erorile metodei robuste mean shift sunt neschimbate și mici, în timp ce erorile estimatorului least-squares sunt în creștere.

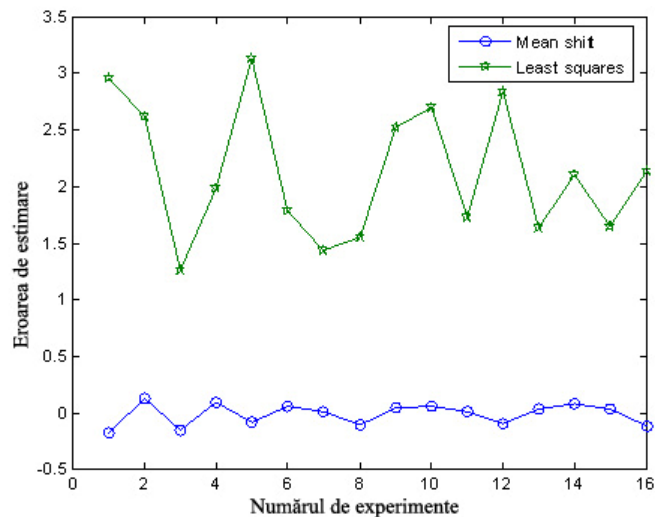


Fig. 5.8. Erorile de estimare ale metodelor robuste mean shift și least-squares cu un procentaj de erori de 10%.

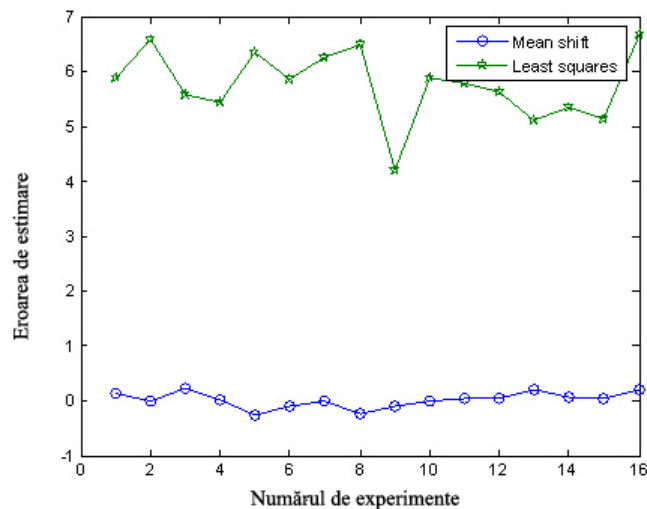


Fig. 5.9. Erorile de estimare ale metodelor robuste mean shift și least-squares cu un procentaj de erori de 33%.

Metoda de aliniere propusă în această lucrare, bazată pe 2D mean shift s-a dovedit a fi solidă și eficientă din punct de vedere al calculului.

### 5.3. Utilizarea metodei de aliniere a imaginilor pentru rețele de senzori video fără fir

Tehnica de aliniere poate fi utilizată în *localizarea nodurilor dintr-o rețea de senzori fără fir*. Studiul în acest domeniu este descris în articolele [181] și [191], ultimul fiind prezentat la conferința internațională Symposium on Applied Computational Intelligence and Informatics 2007, care este cotateă ISI Proceedings și IEEE Explore.

Prima lucrare prezintă o metodă eficientă de localizare împreună cu estimarea câmpului vizual comun pentru aplicațiile de supraveghere care folosesc metoda de aliniere cu scopul de a alinia imaginile capturate simultan de la diferiți senzori video.

Având în vedere că senzorii unei rețele fără fir fac adesea parte dintr-o infrastructură în care sunt repartizate într-o manieră ad-hoc, nu se cunoaște locația lor. Problema estimării coordonatelor spațiale ale unei rețele poartă numele de localizare. În literatura de specialitate există două metode propuse pentru localizarea nodurilor: metode de localizare bazate pe un punct de referință și metode de localizare bazate pe timp sau forța semnalului. În contextul unei rețele de senzori video, problema localizării este complexă. Pe lângă coordonatele nodurilor, informația privind topologia rețelei implică și unghiurile dintre camere dar și câmpul vizual comun al camerelor (figura 4.5). Se poate afirma că suprapunerea câmpurilor vizuale conține suficientă informație de localizare pentru majoritatea aplicațiilor de supraveghere.

Algoritmul de aliniere prezentat mai sus este folosit pentru localizarea senzorilor dintr-o rețea fără fir dar și la determinarea câmpului vizual comun al camerelor. Pe baza imaginilor captate de senzorii video și pe baza punctelor caracteristice alese în fiecare pereche de imagini, un estimator neparametric al densității de probabilitate și metoda robustă mean shift, sunt folosite pentru a găsi cea mai bună estimată pentru parametrii de translație, rotație și scalare. O etapă finală constă din utilizarea acestei estimate pentru a calcula câmpul vizual comun al camerelor (figura 4.5).

În [191] s-au studiat două abordări cu scopul de a genera soluții parțiale pentru estimarea parametrilor transformării de similaritate. Evaluarea performanțelor este interpretată pe baza celor două metode, estimată din 9 experimente. Tabelul 1 prezintă deviația erorii componentelor vectorului de parametri.

Tabel 1. Deviația standard a erorilor componentelor vectorului de parametri

Metoda	Translația pe X (s.d)	Translația pe Y (s.d)	Unghi (s.d)	Scară (s.d)
Secvențială	0.5957	0.4797	0.0022	0.0050
Simultană	0.9978	0.8869	0.0026	0.0084

Rezultatele grafice ale testelor pentru toate componentele vectorilor sunt prezentate în figurile 5.10-5.13. Estimarea simultană a avut performanțe bune în teste, cu excepția estimatei unghiului, dar amândouă metode de estimare au avut rezultate cu precizii ridicate.



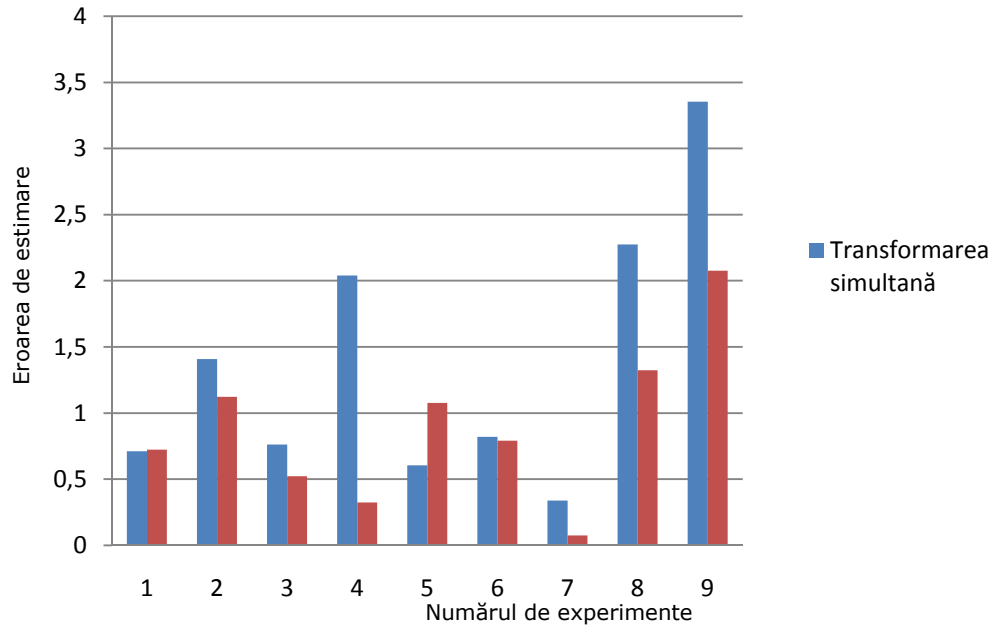


Fig. 5.10. Eroarea estimării translației pe orizontală.

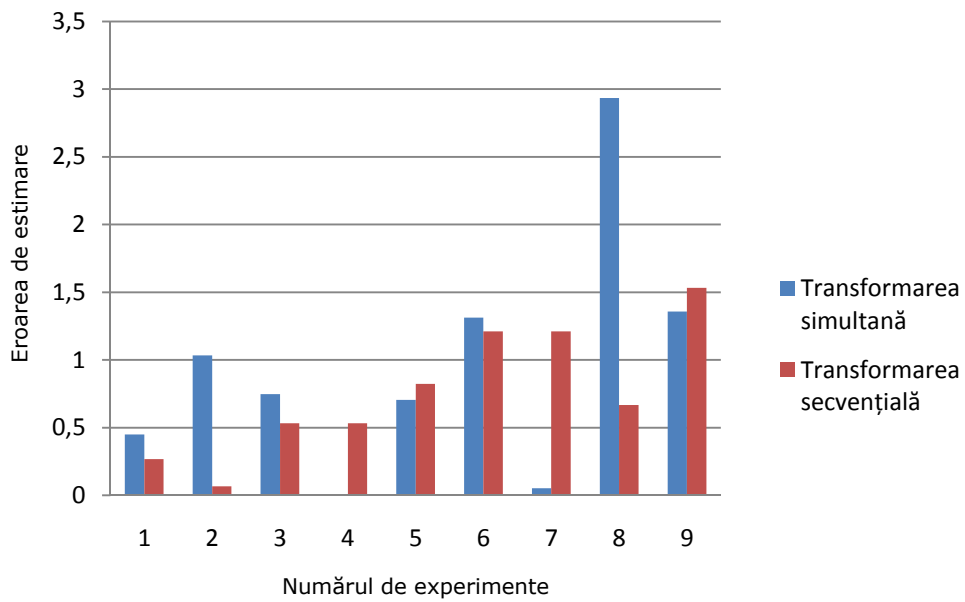


Fig. 5.11. Eroarea estimării translației pe verticală.

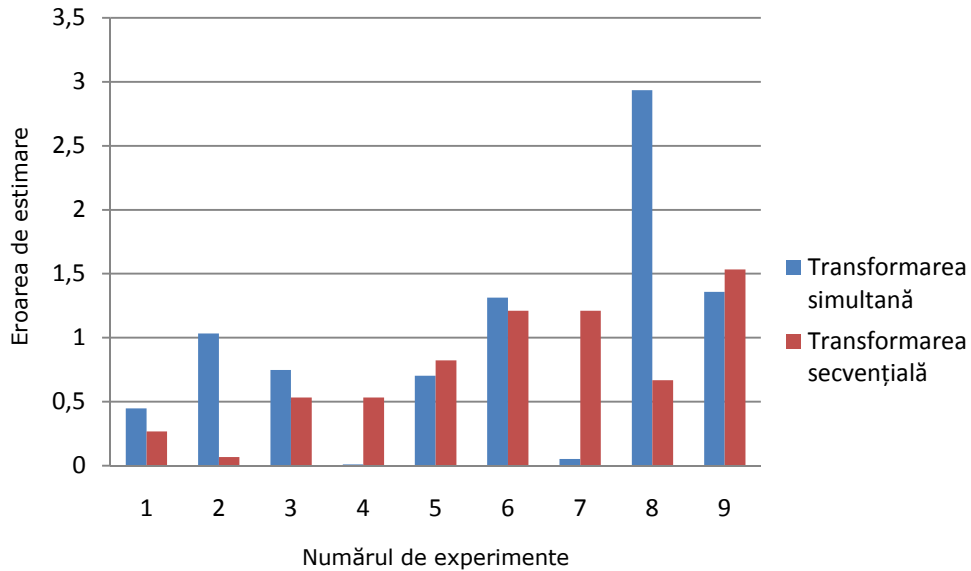


Fig. 5.12. Eroarea estimării factorului unghi.

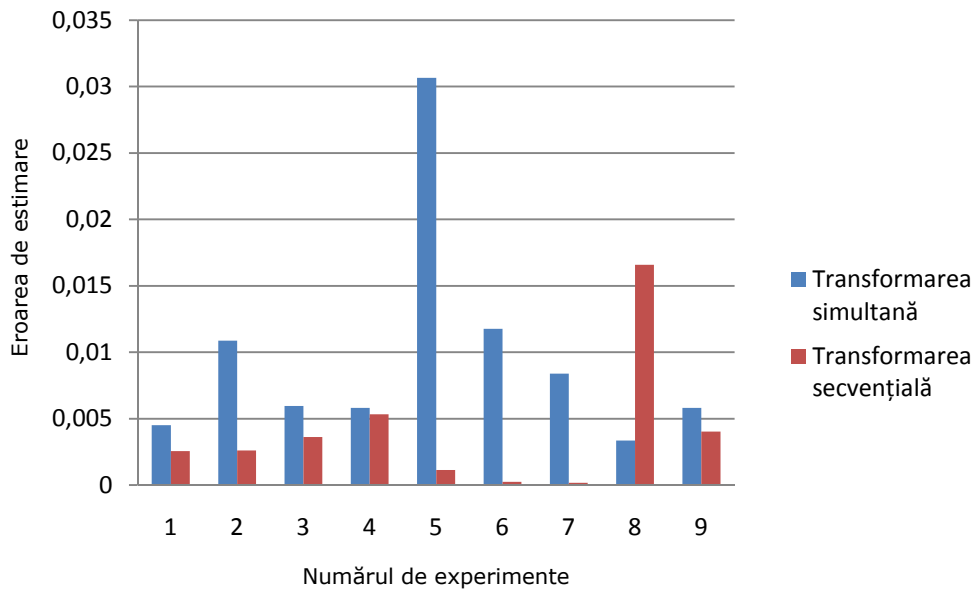


Fig. 5.13. Eroarea estimării factorului scară.

Pe baza experimentelor (figurile 5.10-5.13) s-a constatat că prin metoda simultană a rezultat o eroare de estimare mică și rezultate bune în prezența erorilor determinate exclusiv de posibilele puncte false existente.

## 5.4. Metoda de asociere prin compararea punctelor mediane

În acest paragraf prezentăm rezultatele experimentale din [193], [194] care compară metoda de asociere prin compararea punctelor mediane (MACPM) cu metodele tradiționale

bazate pe măsurarea informației extrase dintr-o vecinătate de puncte caracteristice și a comparării ei, cu scopul de a se elimina asocierile eronate.

Ideea dezvoltată în acest articol este de a compara informația imaginii extrasă din pixeli aflați la mijlocul distanței dintre perechile de puncte (figura 4.4). De obicei, aceste puncte sunt localizate în zone omogene fiind mai puțin afectate de pozițiile exacte ale punctelor caracteristice. Mediana unui segment de linie este invariantă la translație, rotație și rescalare. Teoretic și informația de culoare a punctelor mediane este invariantă la transformările menționate mai sus. Pentru a se studia utilitatea acestei metode, se va compara cu abordările tradiționale, bazate pe măsurări în punctele caracteristice. Acestea sunt metoda de asociere a vecinilor punctelor mediane (MAVPM), metoda de asociere a informației de culoare a punctelor caracteristice (MAICPC) și metoda tradițională de asociere a punctelor caracteristice (MTAPC).

Cea mai simplă metodă de similaritate se referă numai la culorile pixelilor situați în aceeași poziții în ambele imagini. Pentru a calcula norma este nevoie de a se găsi informația de culoare a acestor puncte (relațiile 4.24-4.25). Dacă norma diferențelor culorilor este mai mare decât un prag, perechea considerată corespondentă este eliminată. Rezultatele metodelor menționate mai sus sunt estimate din nouă experimente. Pentru fiecare experiment s-a utilizat un număr diferit de puncte caracteristice. O valoare constantă a pragului,  $T = 10$  pixeli, a fost folosită în toate experimentele, pentru toate metodele în afară de metoda de asociere a informației de culoare a punctelor caracteristice, în care s-au utilizat două norme și ca urmare, două praguri (relațiile 4.26-4.27).

În figurile 5.14-5.18 este reprezentată eroarea de estimare pentru toate soluțiile vectorilor de parametrii.

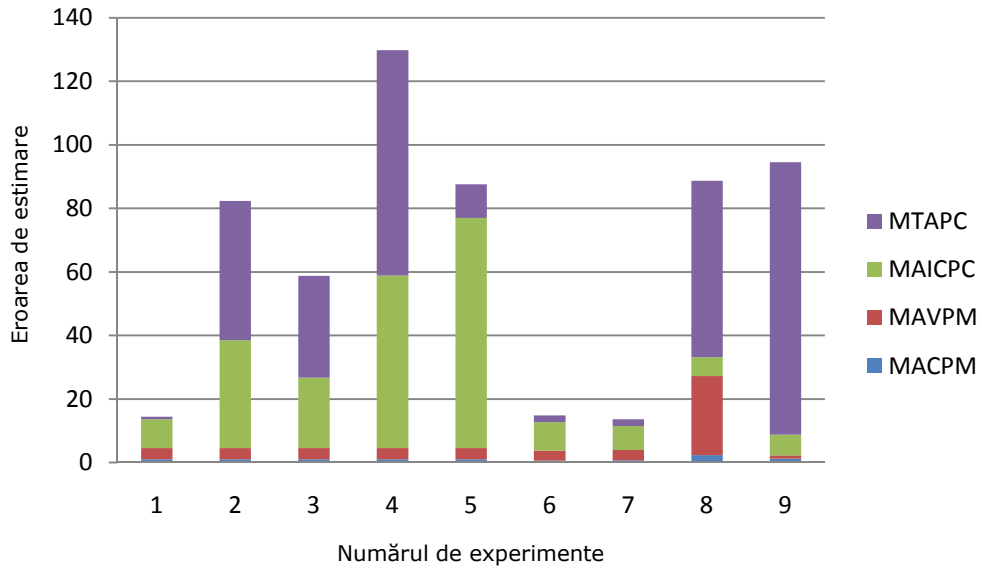


Fig. 5.14. Eroarea estimării translației pe orizontală.

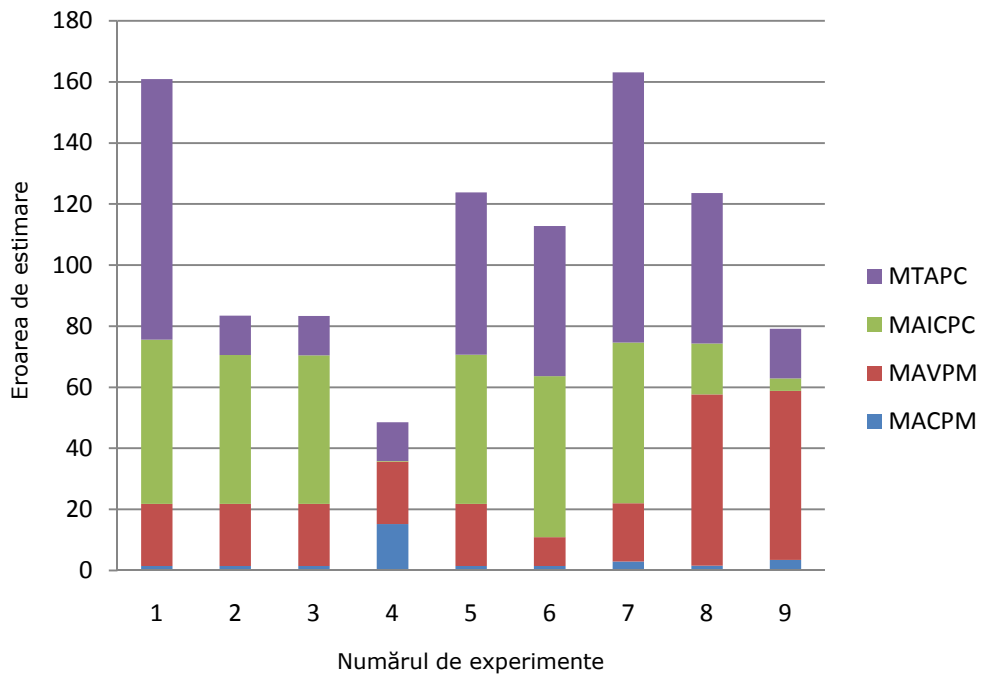


Fig. 5.15. Eroarea estimării translației pe verticală.

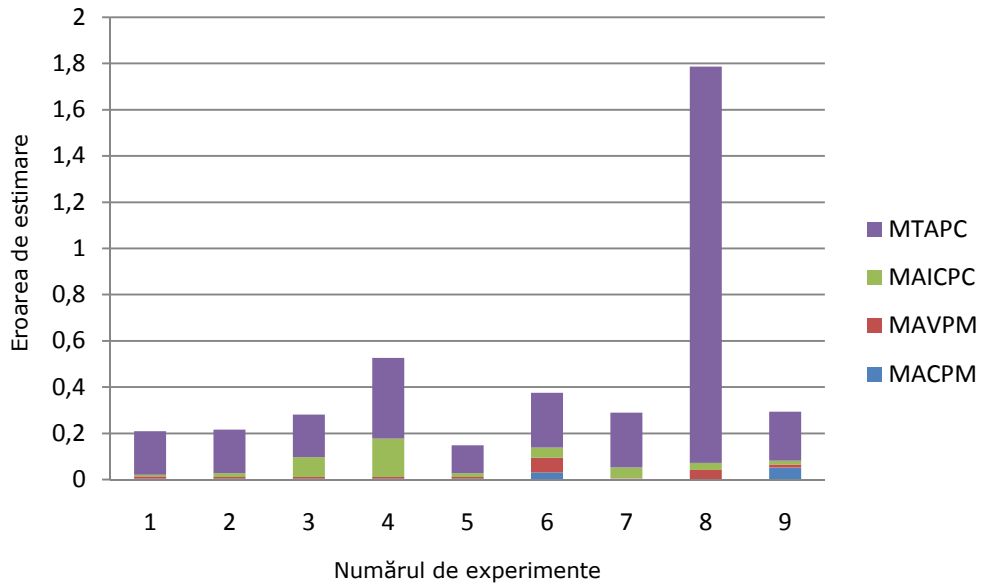


Fig. 5.16. Eroarea estimării factorului unghi.

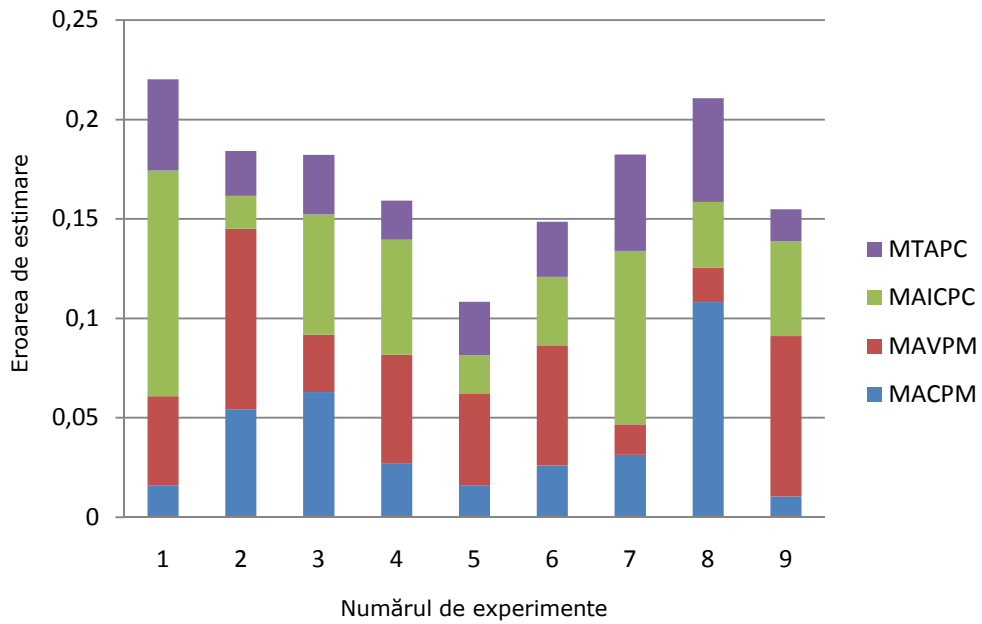


Fig. 5.17. Eroarea estimării factorului scară.

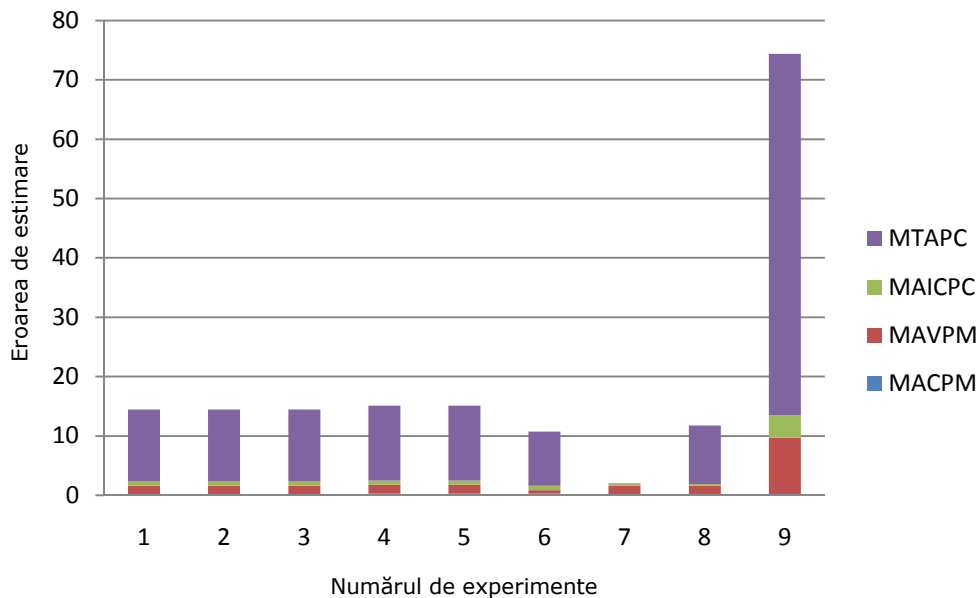


Fig. 5.18. Eroarea de estimare factorului de timp.

Deviațiile standard ale soluțiilor propuse sunt prezentate în tabelul 2.

Tabel 2. Deviația standard a erorilor componentelor vectorului de parametri

Metode	Translația pe X (s.d.)	Translația pe Y (s.d.)	Unghi (s.d.)	Scară (s.d.)
MACPM	0.5027	4.5060	0.0170	0.0313
MAVPM	7.3043	16.7424	0.0206	0.0262
MAICPC	24.2860	22.4363	0.0505	0.0319
MTAPC	32.1852	30.5816	1.1590	0.0133

Prima metodă prezintă rezultate bune în toate testele având cea mai ridicată acuratețe. În special cazul unghiului, translației orizontale și a estimatei timpului. Cel mai neperformant caz pentru metoda propusă este estimata translației pe verticală, unde valoarea erorii este aproximativ 5 pixeli. Cu toate acestea, valoare obținută este acceptabilă. După cum se poate observa din tabelul 4 și din grafice, metoda de asociere a punctelor mediane are avantaj atât față de metoda de asociere a vecinilor punctelor mediane cât și față de celelalte metode.

Tabel 3. Deviația standard a erorii vectorului timp

Metode	Timp T
MACPM	0.0833
MAVPM	2.7565
MAICPC	1.0730
MTAPC	16.2964

Media timpului din experimente este dată în tabelul 3. Se poate observa că metoda propusă prezintă un timp de calcul scăzut și acuratețe de estimare ridicată. Performanțele depind de pragul  $T$ , o problemă studiată în [195] și prezentată la conferința internațională WSEAS, Computational Intelligence, Man-Machine Systems and Cybernetics (CIMMACS '07), cotate ISI Proceedings și IEEE Explore. Acest prag este esențial în eliminarea posibilităților nepotriviri dintre perechile de puncte, cu scopul de a crește acuratețea estimării și de a reduce timpul de procesare.

Pentru a se găsi pragul optim, s-au efectuat 11 experimente, pe un grup de imagini care conțin caracteristici diferite. Fiecare experiment a fost realizat pentru 5 imagini (figura 5.19). Pentru prag s-au folosit valorile constante - 2, 5, 10, 15, 20, 25, 35, 50, 70, 90 și 100 pixeli.



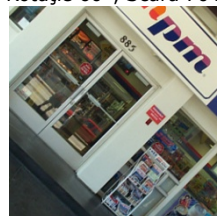
Rotație 0°



Rotație 60°, Scară 70%



Rotație 0°



Rotație 45°, Scară 80%



Rotație 0°



Rotație 45°, Scară 0%



Fig. 5.19. Imaginile utilizate în algoritmul de punere în corespondență.

Rezultatele grafice ale testelor sunt prezentate în figurile 5.20-5.24.

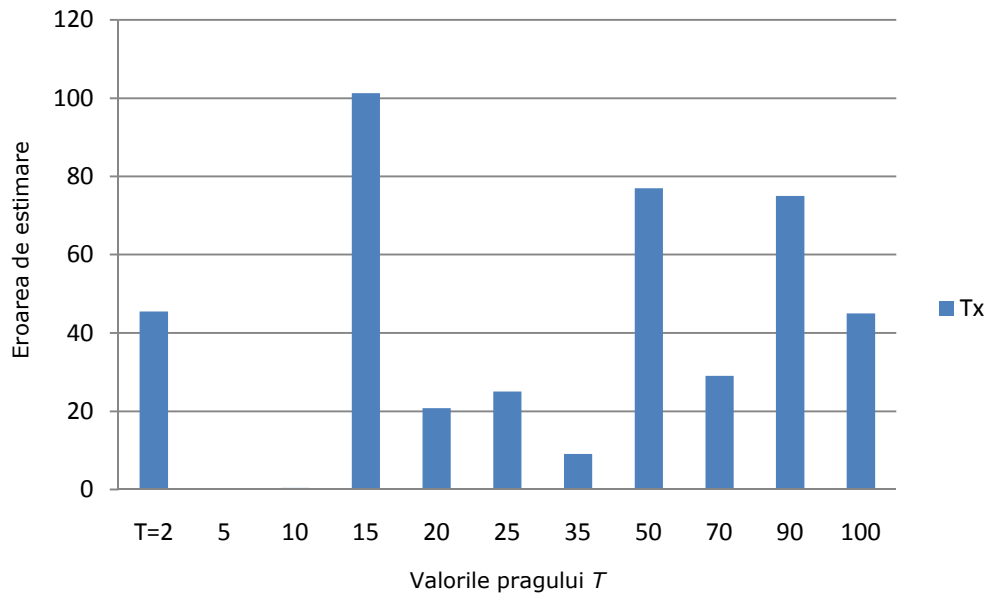


Fig. 5.20. Eroarea estimării translației pe orizontală.



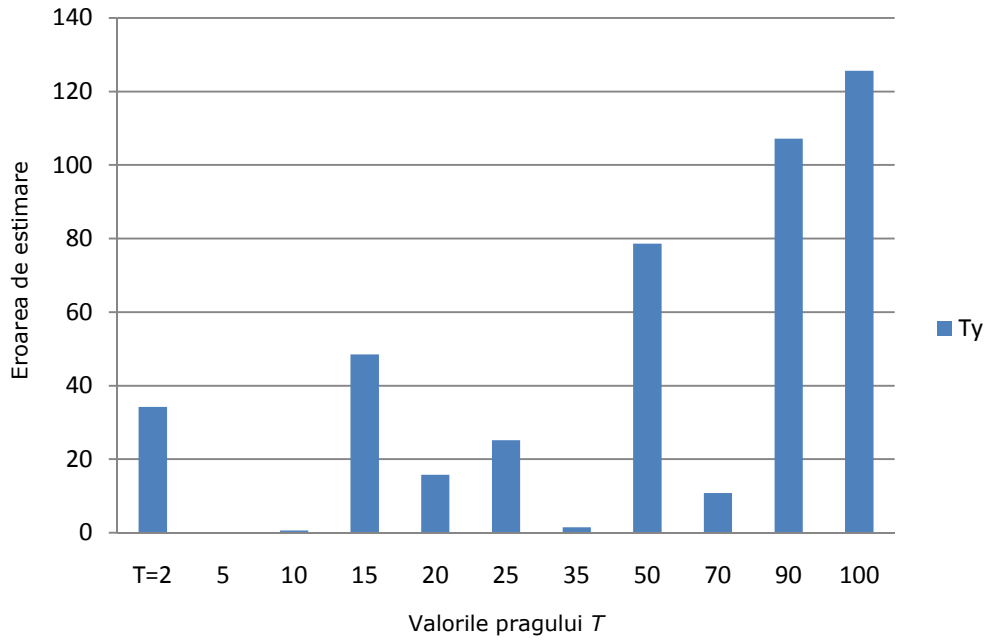


Fig. 5.21. Eroarea estimării translației pe verticală.

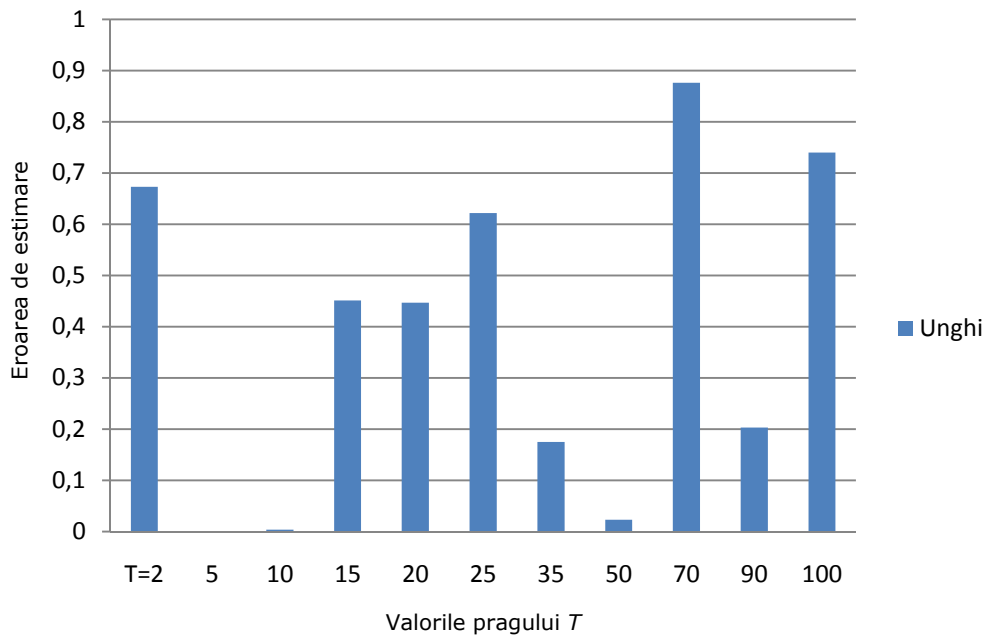


Fig. 5.22. Eroarea estimării factorului unghi.

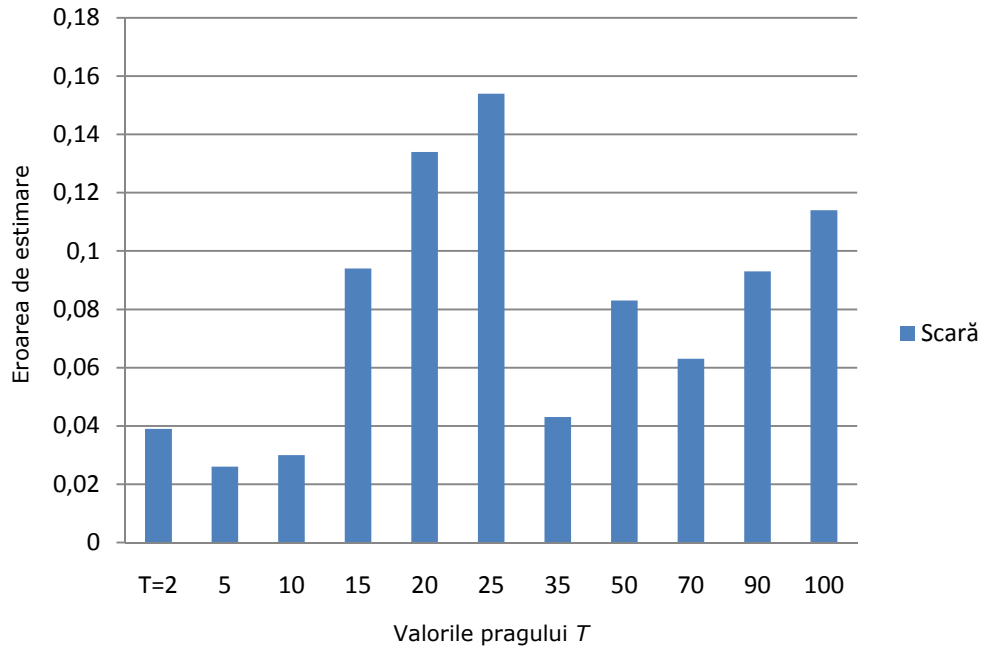


Fig. 5.23. Eroarea estimării factorului scară.

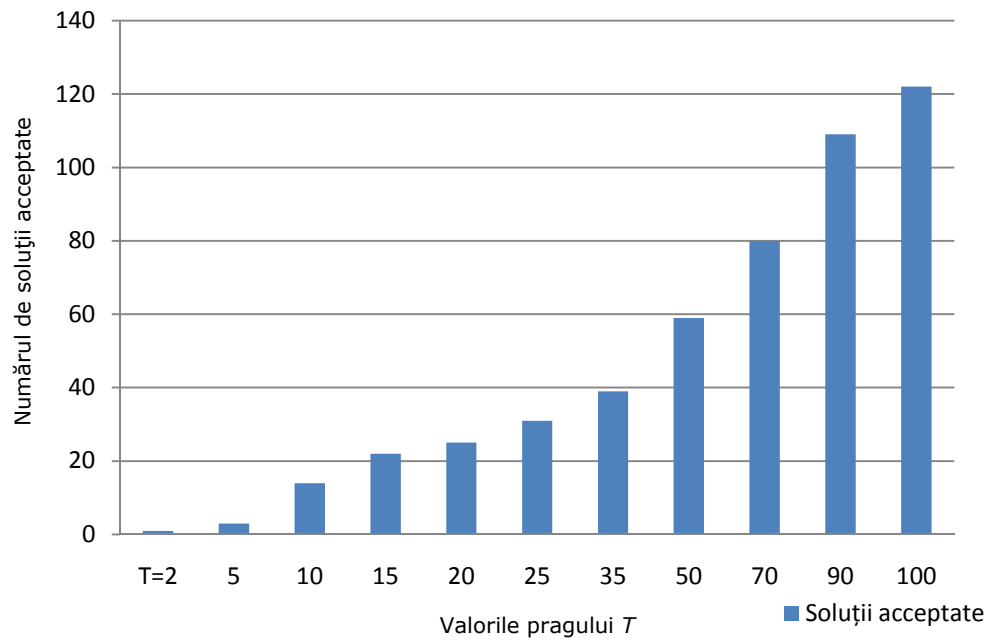


Fig. 5.24. Numărul de soluții acceptate.

Din analiza rezultatelor, se constată că pragul optim are valori reduse, între 2 și 3. Pragul optim găsit în experimente este acela care permite trecerea unui număr de soluții apropiat de numărul real de corespondențe. Prin găsirea unui prag optim în asocierea caracteristicilor, numărul de puncte potrivite devine mic, de asemenea și numărul de puncte eronate și timpul de calcul devine mic. Pragul scăzut găsit în experimente sugerează faptul că o rată de rejecție mare este mai bună decât una mai mică. Algoritmul poate fi modificat atunci când numărul de puncte disponibile este mic. În acest tip de imagini, pragul poate fi ridicat până când este găsit un număr suficient de puncte corespondente.

### 5.5. Studiul selecției punctelor caracteristice pentru localizarea senzorilor din rețele fără fir

Pornind de la ideea de a **îmbunătății extragerea caracteristicilor** pentru a **crește acuratețea procesului de corespondență**, se studiază în cele două metode de selecție a caracteristicilor. Acestea pot fi *manual* sau *automat*. Putem alege corespondențele între caracteristici, *manual*, prin selectarea cu mouse-ul a perechilor de puncte corespondente, sau *automat* folosind un algoritm pentru extragerea caracteristicilor.

Localizarea senzorului este cerută de cele mai multe aplicații WSN (Wireless Sensor Networks – rețele de senzor fără fir). Considerând o aplicație de supraveghere video, localizarea include nu numai coordonate spațiale, ci și direcțiile camerelor și estimarea parțială a câmpului vizual. În [196], [197] se prezintă o tehnică inedită pentru problema localizării în rețele video de senzori fără fir (WSN) folosind punerea în corespondență a imaginilor. Această tehnică implică algoritmul SIFT pentru selectarea automată a caracteristicilor. Rezultatele experimentale arată acuratețea estimării și eficiența temporală în comparație cu soluțiile manuale. Studiul cuprins în [196] a fost prezentat la conferința internațională WSEAS, AUTOMATION and INFORMATION (ICAI'08), cotate ISI Proceedings și IEEE Explore.

Pentru a se testa performanțele celor două metode de selecție a punctelor se folosesc perechi de imagini conținând un câmp vizual obținut din diferite poziționări și orientări ale aparatului de fotografiat. Cu privire la câmpul vizual captat, în general, distingem trei situații familiare. Prima este reprezentată de către scenele din natură, de exemplu, păduri, deșert; a doua se referă la peisajele citadine, precum piețe, intersecții; cea de a treia situație este reprezentată de spațiile închise: camere, interioare. Datorită acestui fapt, experimentele au fost realizate folosind imagini ce cuprind toate cele trei ipostaze. Drept urmare, imaginile au fost împărțite în trei categorii: natură, interior și peisaj urban (figura 5.25).

Pentru fiecare imagine selecția a fost făcută automat și manual pentru un număr diferit de puncte de caracteristici. Fiecare metodă este testată pe 15 imagini (5 imagini  $\times$  3 categorii). De exemplu, pentru cazul în care alegem 25 de perechi de puncte corespondente, selecția automată a punctelor corespondente în perechi de imagini aparținând celor trei categorii de medii este prezentată în figura 5.25. Eroarea medie pătratică de estimare pentru cei patru parametri ai transformării geometrice este reprezentată în tabelul 6. Valorile obținute sunt pentru un număr de 25 de puncte caracteristice ale imaginilor din cele trei categorii (figura 5.25).

Tabel 4. Eroarea medie pătratică pentru 25 de puncte caracteristice selectate automat

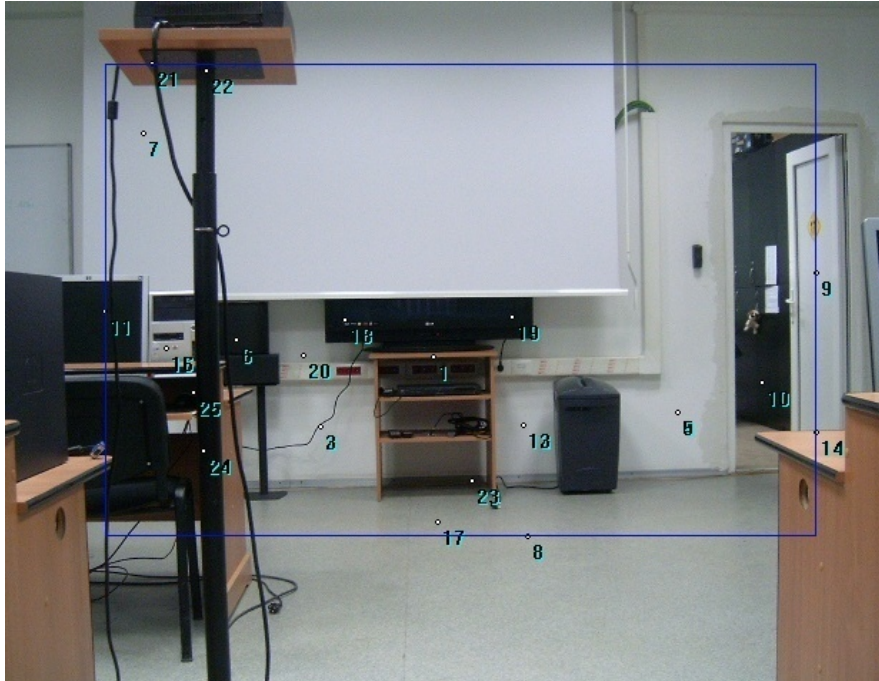
<b>Automat</b>	Tx	Ty	Scară	Unghi
<b>Spațiu închis</b>	9.2210	7.3854	0.0491	0.0247
<b>Natura</b>	0.8135	1.6600	0.0119	0.0259
<b>Peisaj urban</b>	2.5516	2.2762	0.0147	0.0378

Tabel 5. Eroarea medie pătratică pentru 25 de puncte caracteristice selectate manual

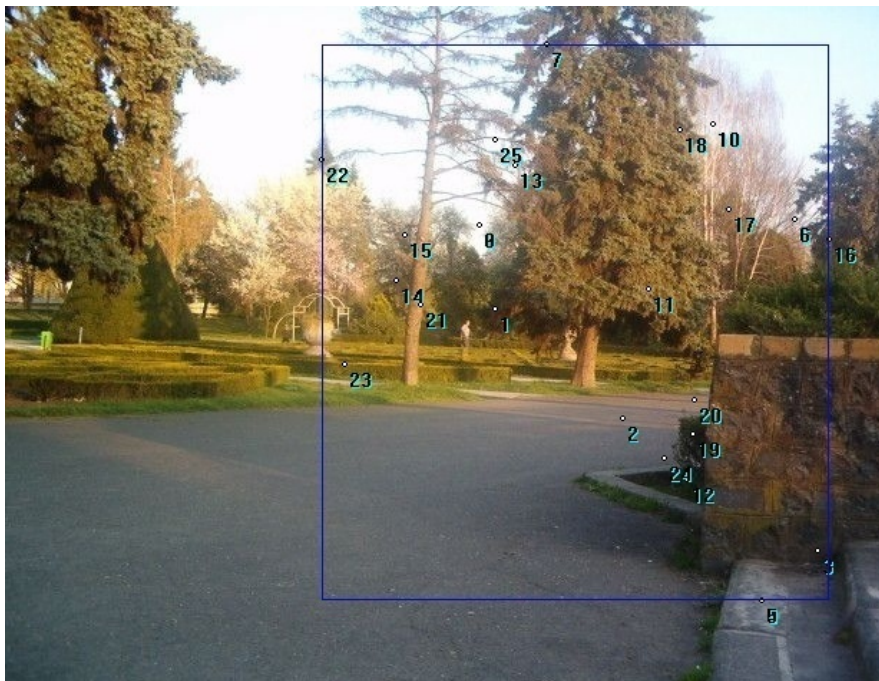
<b>Manual</b>	Tx	Ty	Scară	Unghi
<b>Spațiu închis</b>	13.2963	5.6287	0.0266	0.0233
<b>Natura</b>	5.0616	4.5964	0.0140	0.0333
<b>Peisaj urban</b>	2.8484	2.3618	0.0145	0.0396



a)

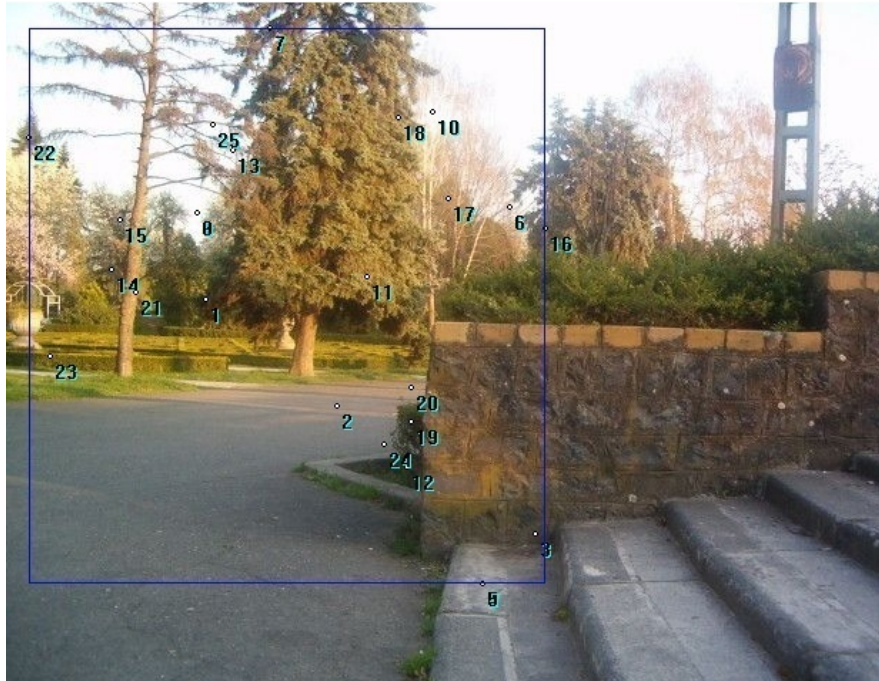


b)

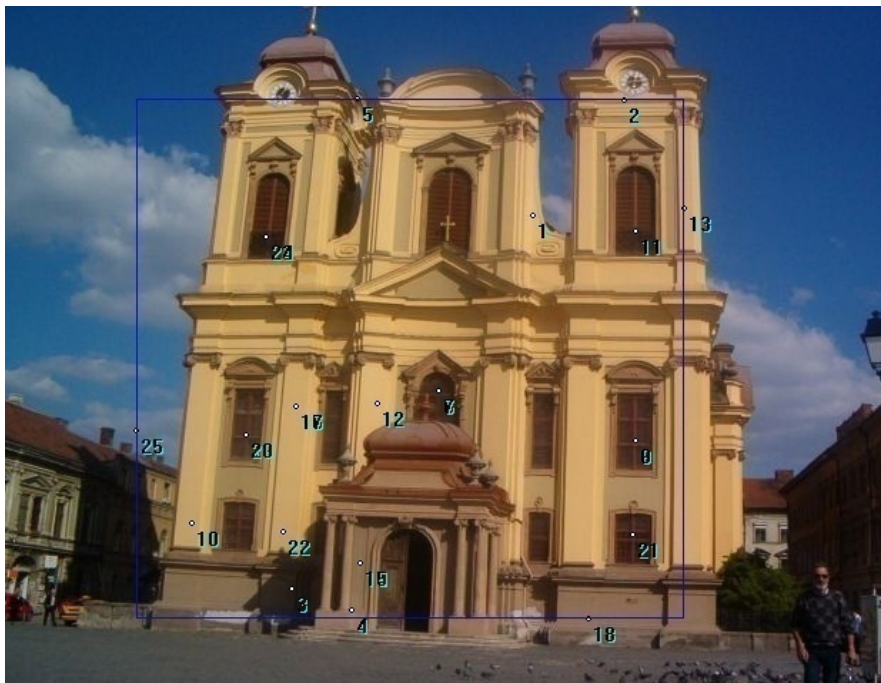


c)

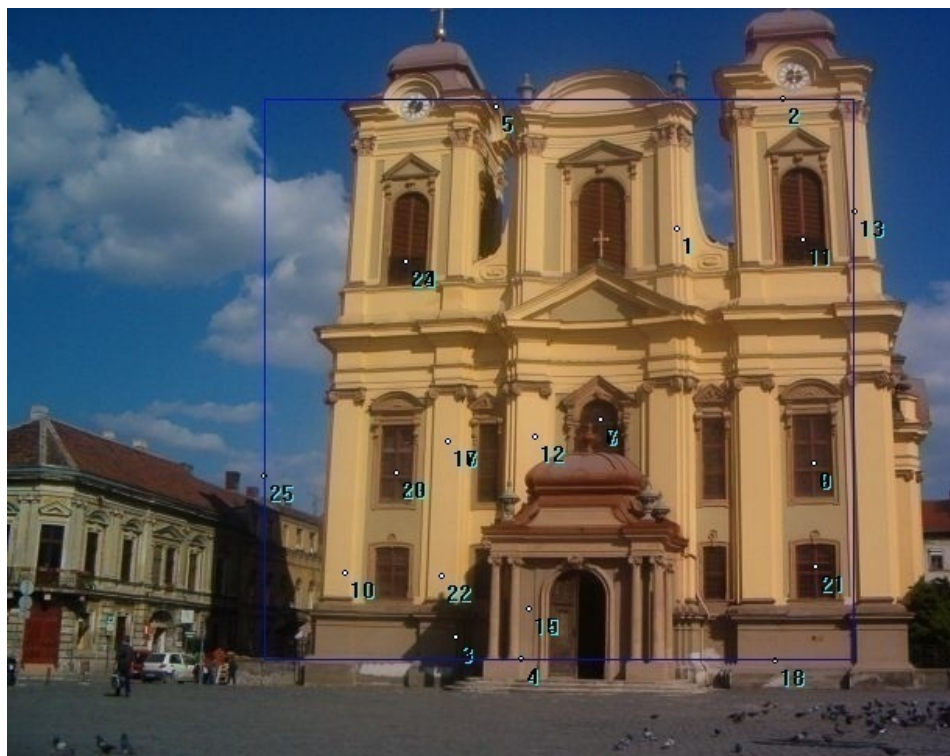




d)



e)



f)

Fig. 5.25. Fotografii din diferite medii:  
 (a), (b)- Spațiu închis; (c), (d)- Natură; (e), (f)-Urban.

Tabelul 6 prezintă măsurările de timp pentru fiecare categorie în cazul cu 25 de puncte și în cazul cu 15 puncte.

Tabel 6. Timpul de execuție pentru selecția manuală și automată

Medii	Manual		Automat	
	25 puncte	15 puncte	25 puncte	15 puncte
<b>Natură</b>	4'10"	3'5"	15"	14"
<b>Spațiu închis</b>	3'10"	1'30"	12"	11"
<b>Peisaj urban</b>	2'10"	1'30"	9"	8"

Într-o rețea de senzori fără fir, numărul de imagini obținut este mare. Bazându-ne pe măsurările din tabelul 8, putem estima timpul necesar unei aplicații reale. Presupunem că avem 500 de perechi de imagini pentru care considerăm 25 de puncte corespondente, și selecția este manuală. În acest caz, procesul de punere

în corespondență va folosi aproximativ  $4'10''$  pentru fiecare imagine în parte. Pentru un număr de 500 de imagini, rezultă un timp total de 35 ore ( $4'10'' * 500$ ). Dacă selecția de caracteristici este automată timpul estimat pentru procesul de punere în corespondență este de  $15'' * 500$  perechi de imagini, care înseamnă aproximativ 2 ore. Chiar dacă considerăm cel mai bun caz, adică peisajul urban, pentru selecția manuală localizarea în timp este inacceptabilă (18 ore), în timp ce selecția automată durează doar o oră.

Rezultatele experimentale arată acuratețea estimării și eficiența temporală a selecției automate în comparație cu selecția manuală. Rezultatele testelor demonstrează avantajele acestei tehnici în termeni ai timpului de execuție, pierderile de precizie fiind acceptabile.

## 5.6. Studiu comparativ (I): Mean shift versus RANSAC

Obiectivul acestui studiu este de a pune în evidență care dintre cei doi algoritmi de estimare robustă prezintă rezultate mai bune atunci când este vorba de punerea în corespondență a imaginilor.

Studiul comparativ a fost publicat în lucrarea [223], cotate ISI Proceedings și IEEE Explore. S-au utilizat imagini ce conțin un câmp vizual comun obținut din diferite poziții și orientări ale aparatului de fotografiat (figura 5.21).

Pragul  $T$  utilizat în algoritmul RANSAC este un parametru nespecificat. Pentru a testa performanța algoritmului, trebuie să fie determinat un prag optim. Nivelul acestui prag influențează eliminarea corespondențelor considerate false. Dacă pragul este prea mare, atunci prea multe date neconforme pot fi considerate ca fiind plauzibile. Aceste rezultate conduc la creșterea timpului de procesare și a procentajului de date neconforme. Rezultatul afectează procesul de punere în corespondență. În caz contrar, dacă pragul este prea mic, se ajunge la lipsa unui număr suficient de corespondențe. Prin urmare, alegerea unui prag optim reprezintă o problemă care trebuie studiată.

Pentru a găsi pragul optim, s-au efectuat 8 experimente. Fiecare experiment s-a realizat pentru câte 6 imagini pentru diferite valori constante ale pragului  $T=1,2,3$  pixeli. S-au utilizat combinații dintre diferite puncte corespondente și puncte false.

Numărul total de puncte din fiecare experiment este de 6, 7, 10, 15, 21, 25 și 35. Algoritmul mean shift este testat pentru același număr de puncte ca și algoritmul RANSAC. Amândouă metodele lucrează cu informație necunoscută despre numărul de puncte corespunde și false.

Experimentele sunt organizate în două grupuri. Prima parte a experimentelor a fost realizată cu un set de puncte puține 6,7 și 10. Eroarea medie pătratică (în engleză mean square error) a vectorului parametrilor transformării de similaritate pentru aceste experimente, este prezentată în tabelele. Seturile de puncte între 15 și 35 s-au utilizat în partea a doua a experimentelor. Datorită numărului ridicat de puncte, rezultatele acestor teste, pentru toți vectorii cu parametri, sunt prezentate sub formă grafică în imaginea.



Tabel 7. Eroarea medie pătratică pentru RANSAC - T=1 puncte puține

Numărul de puncte	RANSAC			
	Tx	Ty	Scară	Unghi
3 corecte – 3 false	5.1239	41.1440	0.1302	0.0041
3 corecte – 4 false	24.3461	132.908	0.3041	1.4116
4 corecte – 6 false	157.041	78.6692	0.2384	1.4861
6 corecte – 4 false	41.1761	162.312	0.1831	1.4781

Tabel 8. Eroarea medie pătratică pentru RANSAC - T=2 puncte puține

Numărul de puncte	RANSAC			
	Tx	Ty	Scară	Unghi
3 corecte – 3 false	36.9889	133.087	0.4155	1.3192
3 corecte – 4 false	94.3188	202.099	0.3319	2.0442
4 corecte – 6 false	94.0327	80.8905	0.1846	0.9050
6 corecte – 4 false	1.1350	1.5727	0.0071	0.0089

Tabel 9. Eroarea medie pătratică pentru RANSAC - T=3 puncte puține

Numărul de puncte	RANSAC			
	Tx	Ty	Scară	Unghi
3 corecte – 3 false	49.7444	173.672	0.4912	1.5493
3 corecte – 4 false	98.5106	202.133	0.4232	2.0506
4 corecte – 6 false	32.5009	74.142	0.0580	0.6647
6 corecte – 4 false	1.3544	1.7269	0.0070	0.0072

Tabel 10. Eroarea medie pătratică pentru Mean Shift puncte puține

Numărul de puncte	MEAN SHIFT			
	Tx	Ty	Scară	Unghi
3 corecte – 3 false	0.9358	1.2581	0.0310	0.0197
3 corecte – 4 false	1.1969	1.3238	0.0273	0.0109
4 corecte – 6 false	8.1557	2.0958	0.0506	0.0411
6 corecte – 4 false	1.4601	1.4834	0.0158	0.0108

Cu privire la algoritmul RANSAC, rezultatele din tabelele 7-10 prezintă eroarea medie pătratică pentru parametrii transformării de similaritate. Erorile de estimare a factorilor de translație de orizontală și verticală sunt diferite de la un prag la altul. Valorile sunt foarte ridicate pentru oricare valoare a pragului (1,2 sau 3) pornind de la 1.13 până la 202. Cele mai mici valori ale erorii medii pătratice le avem pentru  $T=2$  și  $T=3$ , când numărul de puncte este 10 (numărul de puncte eronate este mai mic decât numărul de puncte corecte). Valorile obținute sunt 1.3544 pentru  $T=3$  și 1.5727 pentru  $T=2$ . La  $T=1$ , se prezintă o creștere substanțială a valorii erorii medii pătratice pentru cele 10 puncte, și anume 41.1761.

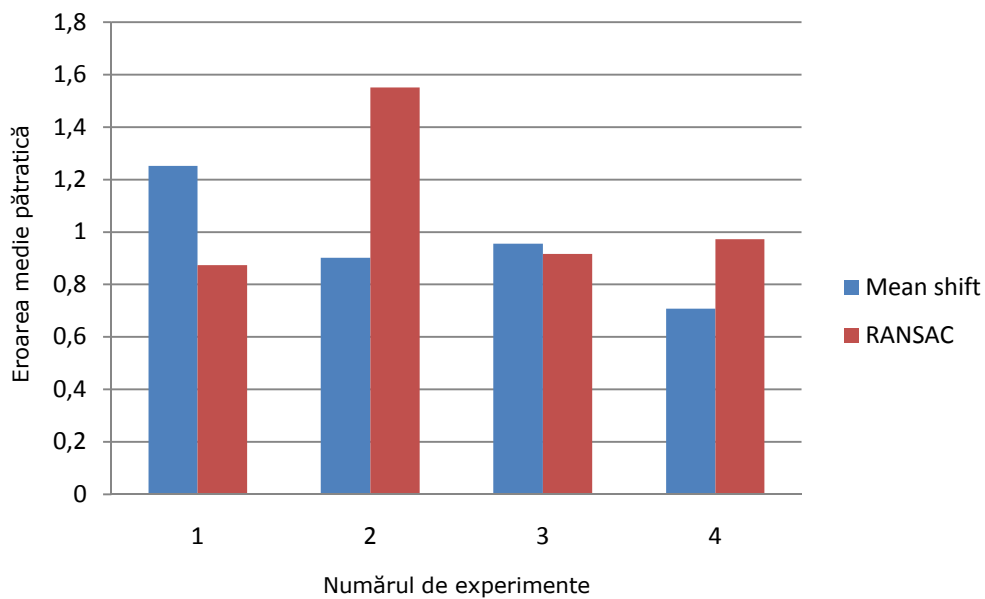
Eroarea medie pătratică pentru scalare prezintă valori apropiate de 0 pentru  $T=2$  și  $T=3$ , când numărul de puncte este 10 (numărul de puncte eronate este mai mic decât numărul de puncte corecte). Valorile obținute sunt 0.0072 pentru  $T=3$  și

0.0071 pentru  $T=2$ . În cazul lui  $T=1$ , valoarea erorii medii pătratice pentru 10 puncte crește, ea fiind de 0.1831.

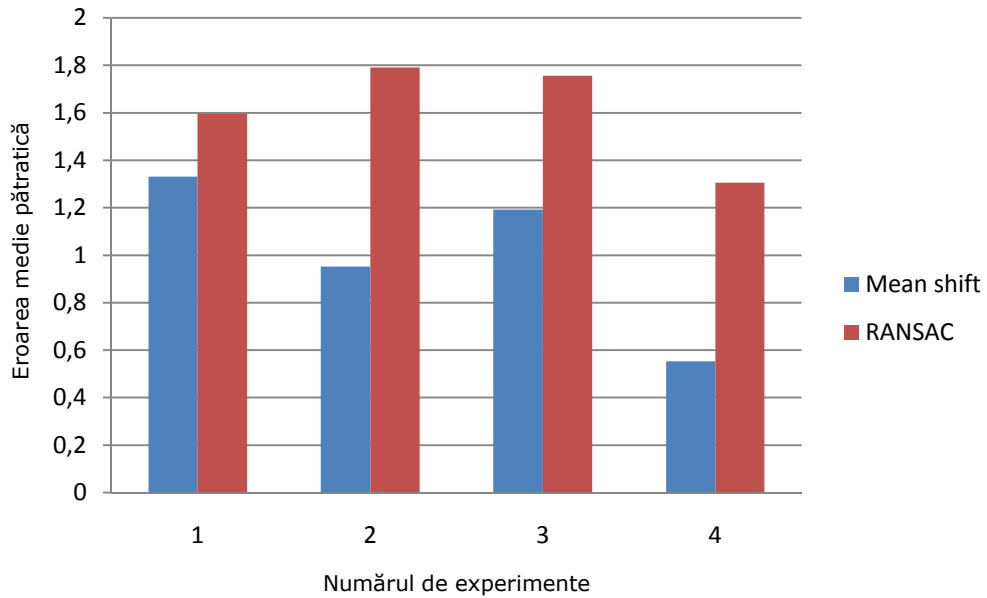
Eroarea medie pătratică pentru unghi prezintă valori apropiate de 0 pentru  $T=2$  și  $T=3$ , atunci când numărul de puncte este 10 (4 corecte, 6 greșite și 6 corecte, 4 greșite). La  $T=1$  aceste valori cresc. Din aceste tabele se poate observa că erorile pentru  $T=2$  sunt cele mai bune, având cele mai mici valori.

În tabelul 10, se prezintă eroarea medie pătratică pentru toți parametrii transformării de similaritate în cazul algoritmului mean shift. Aceste erori sunt mici în raport cu erorile prezente la RANSAC, pentru toate valorile pragului.

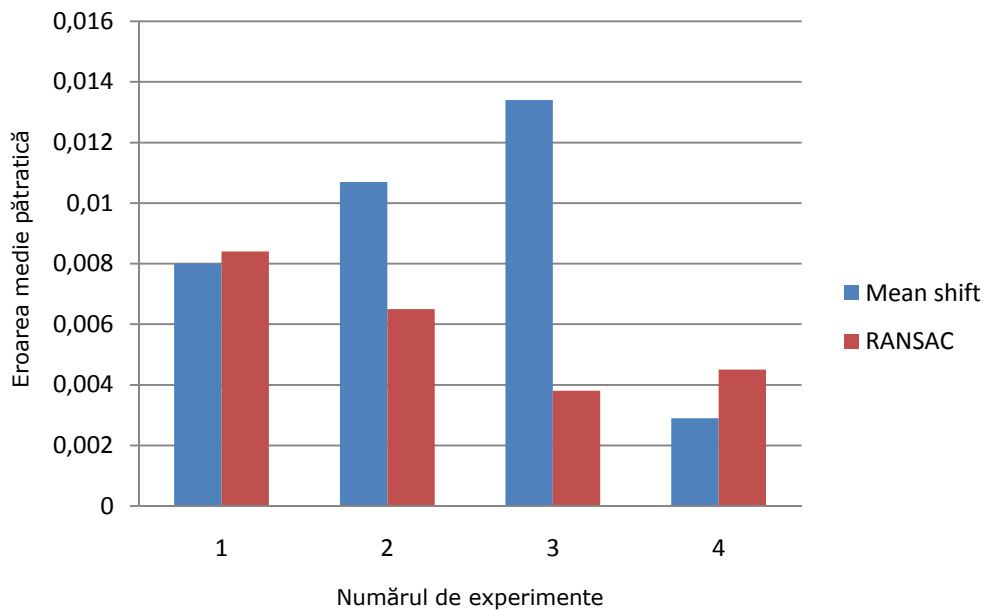
În figura 5.26 sunt reprezentate grafic erorile de estimare a parametrilor pentru metoda mean shift și RANSAC. În cazul ultimului algoritm, pe baza rezultatelor obținute în tabele 7-9, se alege pragul  $T=2$ . Numărul de puncte utilizat în acest set de experimente este între 15 și 35. Algoritmul de mean shift are performanțe mai bune pentru eroarea de estimare a translației pe verticală (a). Erorile prezente fiind mai mici decât cele prezente la RANSAC. Pentru erorile de estimare ale unghiului, scării și translației pe orizontală (b,c,d), amândouă metodele prezintă erori similare.



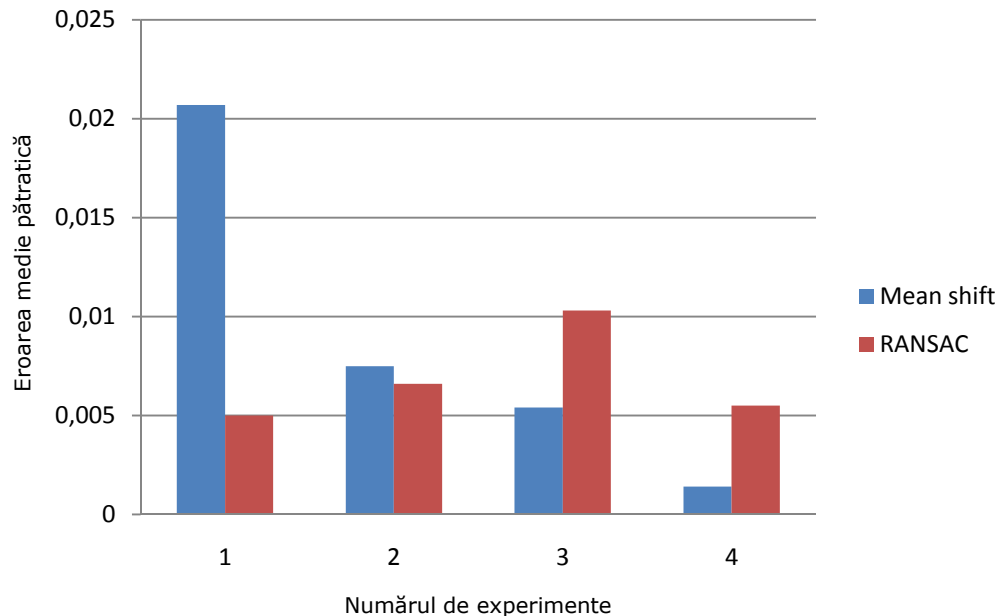
(a)



(b)



(c)



(d)

Fig. 5.26. (a) Eroarea de translație pe orizontală, (b) Eroarea de translație pe verticală, (c) Eroarea de scară, (d) Eroarea de unghi pentru puncte setul doi de puncte (între 15 și 35)

În concluzie, metoda neparametrică a fost testată pentru început pe un set de imagini având parametrii transformării generați artificial de către un program de grafică. Metoda a fost testată și îmbunătățită, fiind comparată cu metoda robustă clasică RANSAC. Testele au arătat că pentru un număr mic de puncte, algoritmul mean shift prezintă mult mai bune rezultate decât algoritmul RANSAC. Atunci când numărul de puncte crește, cele două metode robuste prezintă rezultate similare. Aceste teste au fost extinse pe un număr de imagini mult mai mare, testându-se în paralel metoda mean shift, RANSAC și MLESAC, studiu prezentat în paragraful următor.

## 5.7. Studiu comparativ (II): Mean shift versus RANSAC versus MLESAC

Acest studiu, publicat în lucrarea [224] cotate ISI Proceedings și IEEE Explore, și-a propus analizarea modului în care cele trei metode robuste estimează parametrii transformării de similaritate pentru punerea în corespondență a imaginilor pentru o rețea de senzori video fără fir.

Imaginile pe care se realizează experimentele se regăsesc în figurile 5.1-5.2 și provin din trei medii: natură, urban și interior. Acestea se împart în două grupuri. Primul grup (figura 5.1) este format din 12 imagini captate cu aparatul de fotografiat dintr-o singură poziție și orientare a aparatului. Imaginile țintă au fost transformate artificial având parametrii transformării de similaritate cunoscuți. Al doilea grup (figura 5.2) conține imagini ale căror transformări au fost generate cu aparatul de fotografiat cu diferite orientări și poziționări ale aparatului.

Pentru fiecare imagine selecția punctelor caracteristice a fost făcută automat și manual pentru un număr diferit de puncte de caracteristici. Fiecare metodă este testată pe cele două grupuri de imagini. Numărul total de puncte din fiecare experiment este de 6, 7, 10, 15, 21, 25 și 35. Algoritmul mean shift este testat pentru același număr de puncte ca și algoritmul RANSAC și MLESAC. Toate cele trei metode lucrează cu informație necunoscută despre numărul de puncte corespondente și necorespondente (false).

Experimentele sunt împărțite în două seturi principale: manual și automat. Pentru fiecare dintre aceste experimente, pentru grupul 1 de imagini sunt generate tabele (13-24) care conțin eroarea medie pătratică a vectorului parametrilor transformării de similaritate. Experimentele se realizează pentru două seturi de puncte, primul set de puncte este cel de puncte puține 6,7 și 10, iar cel de-al doilea set este setul de puncte multe, între 15 și 35. Pentru grupul 2 de imagini, datorită numărului ridicat de puncte, rezultatele acestor teste, pentru toți vectorii cu parametri, sunt prezentate sub formă grafică în imaginile 5.29-5.44.

### 5.7.1. Experimente cu o cameră fixă

#### 5.7.1.1. Experimente cu selecție manuală și puncte puține

Pentru început, în tabelele 11-13 se prezintă eroarea medie pătratică pentru parametrii transformării de similaritate atunci când numărul punctelor caracteristice este mic. Selecția punctelor caracteristice s-a realizat manual.

Tabel 11. Eroarea medie pătratică pentru puține și selecție manuală - MEAN SHIFT

Numărul de puncte	MEAN SHIFT			
	Tx	Ty	S	A
3 corecte – 3 false	2.9462	16.0074	0.0631	0.0433
3 corecte – 4 false	2.9817	15.5827	0.0638	0.0872
3 corecte – 6 false	6.0977	4.3274	0.0312	0.0396
4 corecte – 6 false	2.9904	5.4968	0.0335	0.0303
4 corecte – 8 false	4.4074	2.3294	0.0389	0.0304
6 corecte – 4 false	1.2754	2.2099	0.0236	0.0227

Tabel 12. Eroarea medie pătratică pentru puncte puține și selecție manuală – RANSAC

Numărul de puncte	RANSAC			
	T <sub>x</sub>	T <sub>y</sub>	S	A
3 corecte – 3 false	278.6025	9.4674	282.9184	1.1130
3 corecte – 4 false	161.7350	27.0964	152.6533	0.6152
3 corecte – 6 false	133.4776	27.8586	170.5217	0.6182
4 corecte – 6 false	169.8881	24.1892	170.6640	0.4932
4 corecte – 8 false	162.3438	34.6518	186.3342	0.4666
6 corecte – 4 false	201.4516	28.1351	168.7397	0.4644

Tabel 13. Eroarea medie pătratică pentru puncte puține și selecție manuală – MLESAC

Numărul de puncte	MLESAC			
	T <sub>x</sub>	T <sub>y</sub>	S	A
3 corecte – 3 false	50.2161	133.9510	0.3965	1.1466
3 corecte – 4 false	90.6904	163.6598	0.3878	1.5401
3 corecte – 6 false	73.6061	136.2036	0.2999	1.3260
4 corecte – 6 false	31.7354	31.6531	0.1148	0.4898
4 corecte – 8 false	40.7955	118.3194	0.2367	1.2264
6 corecte – 4 false	1.1647	1.9910	0.0149	0.0078

Se poate observa din tabele de mai sus, că pentru un număr mic de puncte (între 6 și 10) mean shift înregistrează cele mai mici erori medii pătratice, apropiate de 0 pentru translația pe orizontală, verticală, unghi și scară. Mai mult decât atât, se obțin erori medii pătratice mici atunci când *numărul de puncte corespondente este mai mic decât numărul de puncte necorespondente*.

Mean shift păstrează valoarea erorii medii pătratice, pentru translația pe orizontală și verticală, sub valoarea 16. Pe de altă parte erorile medii pătratice furnizate de RANSAC și MLESAC depășesc cu mult această valoare. Cele mai mari valori se găsesc la RANSAC, cuprinse între 1.1647 și 278.6025.

Eroarea medie pătratică pentru scară are cele mai mari valori pentru RANSAC, între 152.6533 și 282.9184 și cele mai mici pentru mean shift, între 0.0236 și 0.0631. MLESAC prezintă erori medii pătratice pentru estimarea scării mai bune decât RANSAC, dar mai slabe decât mean shift.

În cazul erorii medii pătratice pentru unghi, cele mai mari valori se înregistrează la MLESAC, între 0.0078 și 1.5401, și cele mai mici la mean shift, între 0.0227 și 0.0872.

Se poate concluziona că atunci când *numărul de puncte corespondente este mai mare decât numărul de puncte necorespondente* (6 corecte/4 false), erorile medii pătratice ai celor patru parametri prezintă valori mici, similare pentru mean shift și MLESAC.

### 5.7.1.2. Experimente cu selecție manuală și puncte multe

În tabelele 14-16 sunt prezentate erorile medii pătratice pentru parametrii transformării de similaritate, în cazul selecției manuale de caracteristici. Numărul de puncte caracteristice ales este mai mare decât cel prezentat în paragraful anterior, cuprins între 15 și 33 de puncte.

Tabel 14. Eroarea medie pătratică pentru puncte multe și selecție manuală - MEAN SHIFT

Numărul de puncte	MEAN SHIFT			
	Tx	Ty	S	A
8 corecte – 7 false	1.2644	1.6844	0.0175	0.0186
6 corecte – 12 false	3.9003	3.4114	0.0159	0.0238
14 corecte – 6 false	3.2048	1.3995	0.0150	0.0101
10 corecte – 5 false	1.0901	1.3700	0.0189	0.0256
13 corecte – 15 false	0.7994	1.2420	0.0169	0.0220
33 corecte	0.7325	1.5976	0.0116	0.0023

Tabel 15. Eroarea medie pătratică pentru puncte multe și selecție manuală – RANSAC

Numărul de puncte	RANSAC			
	Tx	Ty	S	A
8 corecte – 7 false	192.0633	24.3172	155.2438	0.4829
6 corecte – 12 false	134.0414	44.6136	149.2207	0.4935
14 corecte – 6 false	159.4740	19.9669	117.7932	1.0261
10 corecte – 5 false	97.9928	21.7177	142.0029	0.7819
13 corecte – 15 false	105.1682	20.0483	142.1891	0.7655
33 corecte	175.3608	32.8549	143.0885	0.4701

Tabel 16. Eroarea medie pătratică pentru puncte multe și selecție manuală – MLESAC

Numărul de puncte	MLESAC			
	Tx	Ty	S	A
8 corecte – 7 false	1.4338	1.8982	0.0157	0.0078
6 corecte – 12 false	52.2271	72.5579	0.3988	0.0046
14 corecte – 6 false	47.9509	46.3017	0.3426	0.5106
10 corecte – 5 false	84.9578	118.6735	0.5726	0.5106
13 corecte – 15 false	75.7682	104.6528	0.4972	1.4797
33 corecte	43.7680	67.5444	0.4053	0.5841

Din analiza tabelelor 14-16 rezultă că atunci când numărul de puncte caracteristice crește (între 15 și 33), și numărul de puncte corespondente este mai mic decât numărul de puncte necorespondente, pentru translația pe orizontală și pe verticală, eroarea medie pătratică are cele mai mici valori pentru algoritmul mean shift. La o diferență destul de mare (50 pixeli) este urmat de algoritmul MLESAC și RANSAC. Eroarea medie pătratică pentru scară și unghi prezintă erori similare pentru MLESAC și mean shift.

Din tabelul 14 rezultă că atunci când numărul de puncte corespondente este mai mare decât numărul de puncte necorespondente, eroarea medie pătratică

pentru translația pe orizontală și verticală are cele mai mici valori pentru mean shift, între 3.2048 și 0.7325. Cele mai mari valori le înregistrează RANSAC, 192.0633 și 19.9669. Eroarea medie pătratică pentru scară înregistrează cele mai mari valori pentru RANSAC, între 117.7932 și 155.2438. Mean shift și MLESAC înregistrează valori ale erorii medii pătratice similare. Totuși, cele mai mici valori se găsesc la mean shift.

Eroarea medie pătratică pentru unghi prezintă valori apropiate de 0 în toate cele trei metode robuste, dar cele mai mici valori se găsesc la mean shift, fiind urmat de MLESAC și RANSAC.

### 5.7.1.3. Experimente cu selecție automată și puncte puține

În tabelele 17-19 se prezintă eroarea medie pătratică pentru parametrii transformării de similaritate atunci când numărul de puncte caracteristice este mic. Extragerea caracteristicilor s-a realizat automat cu un program de selecție automată SIFT. Numărul de puncte detectate este cuprins între 5 și 12.

Tabel 17. Eroarea medie pătratică pentru puncte puține și selecție automată - MEAN SHIFT

Numărul de puncte	MEAN SHIFT			
	Tx	Ty	S	A
5	2.6603	1.4789	0.0110	0.0099
6	1.5015	1.4709	0.0117	0.0090
7	1.4173	5.4001	0.0107	0.0082
9	1.6482	1.8621	0.0092	0.0064
10	1.3656	1.8991	0.0088	0.0050
12	1.4106	1.4378	0.0089	0.0047

Tabel 18. Eroarea medie pătratică pentru puncte puține și selecție automată - RANSAC

Numărul de puncte	RANSAC			
	Tx	Ty	S	A
5	194.4472	30.2436	155.0535	0.4642
6	203.9307	33.4054	162.8310	0.4578
7	201.4773	35.3275	163.7917	0.4392
9	182.8679	26.8289	156.9107	0.4543
10	172.9629	27.7970	153.9221	0.4952
12	165.7938	29.0890	148.5937	0.4940

Tabel 19. Eroarea medie pătratică pentru puncte puține și selecție automată - MLESAC

Numărul de puncte	MLESAC			
	Tx	Ty	S	A
5	3.8269	134.8428	0.0175	0.0106
6	4.0772	134.3383	0.0205	0.0099
7	4.1320	134.6278	0.0215	0.0104
9	1.0754	136.9639	0.0118	0.0049
10	1.1025	136.8424	0.0103	0.0047
12	1.2635	137.0024	0.0099	0.0047



Din tabele 17-19 rezultă că pentru selecția automată de puncte caracteristice, cele mai mici valori ale erorii medii pătratice pentru translația pe orizontală și verticală se înregistrează pentru metoda robustă mean shift. Cele mai mari valori le prezintă metoda robustă RANSAC, fiind cuprinse între 26.8289 și 203.9307, fiind urmată de MLESAC cu valori între 137.0024 și 4.1320.

Pentru eroarea medie pătratică la estimarea scării, cele mai mari erori le prezintă metoda RANSAC. Metoda mean shift și MLSEAC prezintă erori similare de valori mici, apropiate de 0.

Eroarea medie pătratică pentru unghi prezintă valori similare în cazul celor trei metode robuste, apropiate de 0. Totuși cele mai mici valori se înregistrează pentru mean shift, urmată de MLESAC și RANSAC.

#### 5.7.1.4. Experimente cu selecție automată și puncte multe

În tabelele 20-22 se prezintă eroarea medie pătratică pentru parametrii transformării de similaritate atunci când numărul de puncte caracteristice este mai mare. Punctele sunt detectate automat și numărul lor este cuprins între 15 și 35 de puncte.

Tabel 20. Eroarea medie pătratică pentru puncte multe și selecție automată - MEAN SHIFT

Numărul de puncte	MEAN SHIFT			
	Tx	Ty	S	A
15	1.1468	1.2837	0.0080	0.0032
18	1.0211	1.3006	0.0094	0.0026
20	1.0840	0.9090	0.0092	0.0030
25	1.0463	0.9217	0.0097	0.0025
28	0.8356	0.9462	0.0095	0.0024
35	0.7805	0.9381	0.0099	0.0018

Tabel 21. Eroarea medie pătratică pentru puncte multe și selecție automată - RANSAC

Numărul de puncte	RANSAC			
	Tx	Ty	S	A
15	165.2306	30.3556	149.6154	0.4716
18	167.3553	28.7279	152.3852	0.4715
20	164.9997	29.4404	154.7114	0.4653
25	158.9121	26.4449	148.1608	0.5028
28	169.2773	30.6602	154.3732	0.4781
35	162.0931	34.2302	140.5795	0.6208

Tabel 22. Eroarea medie pătratică pentru puncte multe și selecție automată – MLESAC

Numărul de puncte	MLESAC			
	Tx	Ty	S	A
15	0.8402	137.1830	0.0082	0.0045
18	1.1971	136.7422	0.0072	0.0058
20	0.5767	137.1140	0.0092	0.0045
25	1.0216	137.2934	0.0068	0.0057
28	48.2297	141.5019	0.3317	0.0453
35	59.7164	164.0765	0.4976	1.6037

Se observă din tabelul 20 că în cazul selecției automate de caracteristici, atunci când numărul punctelor este mare, valorile pentru eroarea medie pătratică a celor patru parametrii sunt mici pentru metoda robustă mean shift. Eroarea medie pătratică pentru translația pe orizontală și pe verticală prezintă cele mai mari erori pentru RANSAC, valori între 26.4449 și 169.2773. Tot la capitolul valori mari este urmat de MLESAC cu valori între 0.8402 și 164.0765.

Eroarea medie pătratică pentru scară prezintă cele mai mari valori pentru RANSAC. Erorile medii pătratice înregistrate de mean shift și MLESAC sunt apropiate de 0. Dar cele mai mici se înregistrează la mean shift.

Valorile cele mai mici ale erorii medii pătratice pentru unghi sunt prezente în cazul metodei mean shift, urmate de MLESAC și RANSAC. RANSAC prezintă cele mai mari erori medii pătratice pentru translația pe orizontală, verticală și scară.

## 5.7.2. Experimente cu o cameră mobilă

### 5.7.2.1. Experimente cu selecție manuală și puncte puține

În figurile 5.27-5.30 sunt reprezentate erorile de estimare ale celor 4 parametrii ai transformării de similaritate. Analizând aceste figuri se poate observa că în cazul unui număr de puncte corespondente care este mai mic sau egal decât numărul de puncte necorespondente, valorile erorii medii pătratice pentru translația pe orizontală și verticală sunt foarte mari pentru toate cele trei metode robuste.

Eroarea medie pătratică pentru translația pe orizontală și pe verticală are cele mai reduse valori pentru algoritmul mean shift, cu valori între 3.7188 și 72.7405. Cea mai mare valoare pentru eroarea medie pătratică este pentru translație în cazul algoritmului RANSAC, unde valorile se încadrează între 84.9249 și 323.0432. Pentru algoritmul MLESAC, valorile sunt între 4.0907 și 168.7461.

La estimarea factorului de scară, estimatorul mean shift prezintă valori apropiate de 0 pentru eroarea medie pătratică, între 0.0113 și 0.0288. Pentru metoda MLESAC, valorile erorii medii pătratice sunt între 0.0189 și 0.3189. Prin urmare, cele mai mici valori pentru eroarea medie pătratică le întâlnim la algoritmul mean shift. RANSAC prezintă cele mai mari erori pentru estimarea scării. Eroarea medie pătratică pentru estimarea unghiului prezintă valori mici (apropiate de 0) în toate cazurile. Chiar și așa, cele mai bune rezultate le întâlnim la metoda mean shift, chiar și atunci când numărul de puncte corespondente este depășit de numărul de puncte necorespondente.

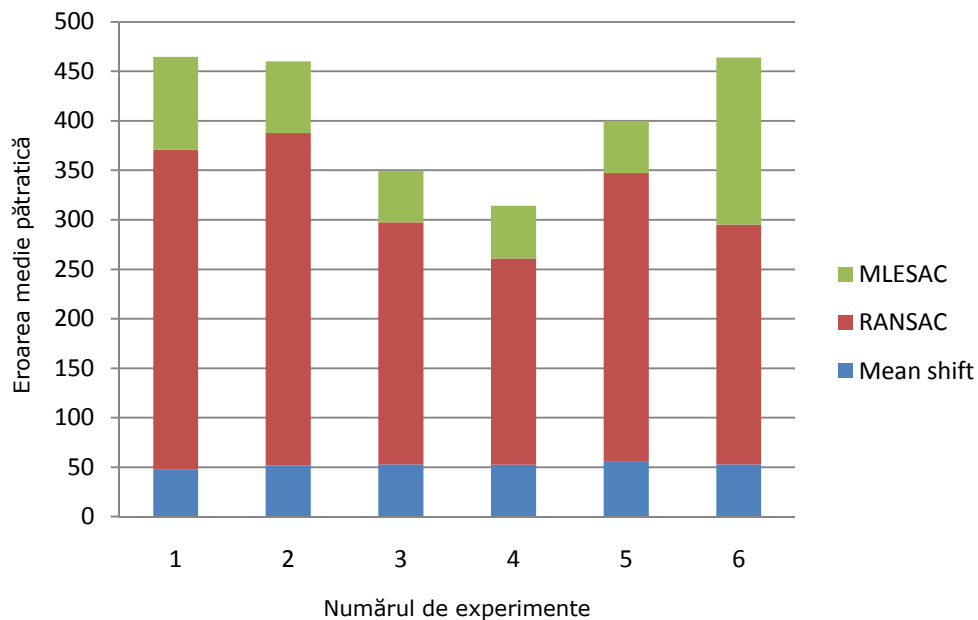


Fig. 5.27. Eroarea de translație pe orizontală pentru puncte puține și selecție manuală

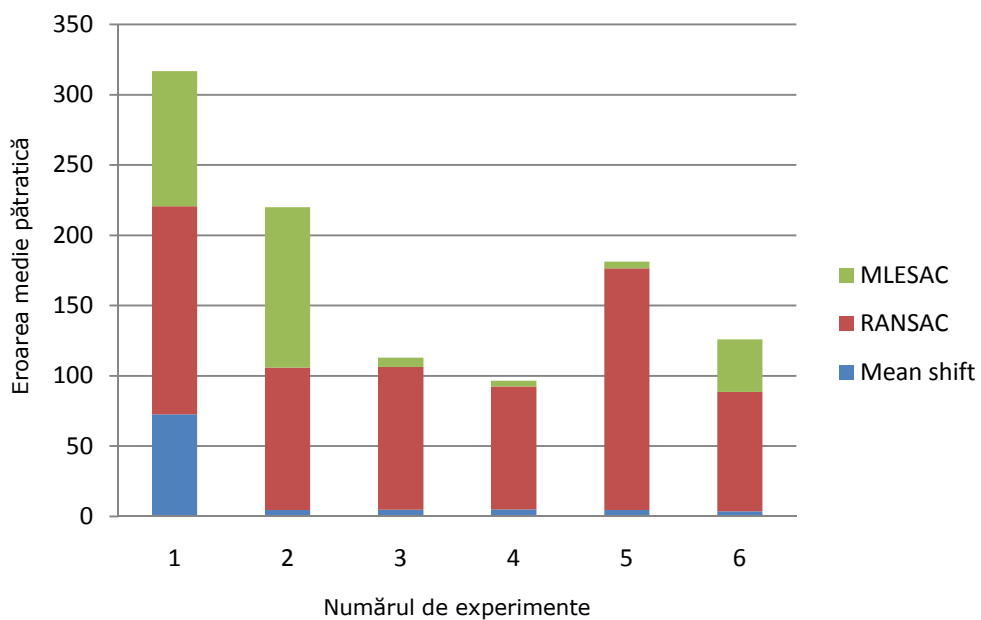


Fig. 5.28. Eroarea de translație pe verticală pentru puncte puține și selecție manuală.

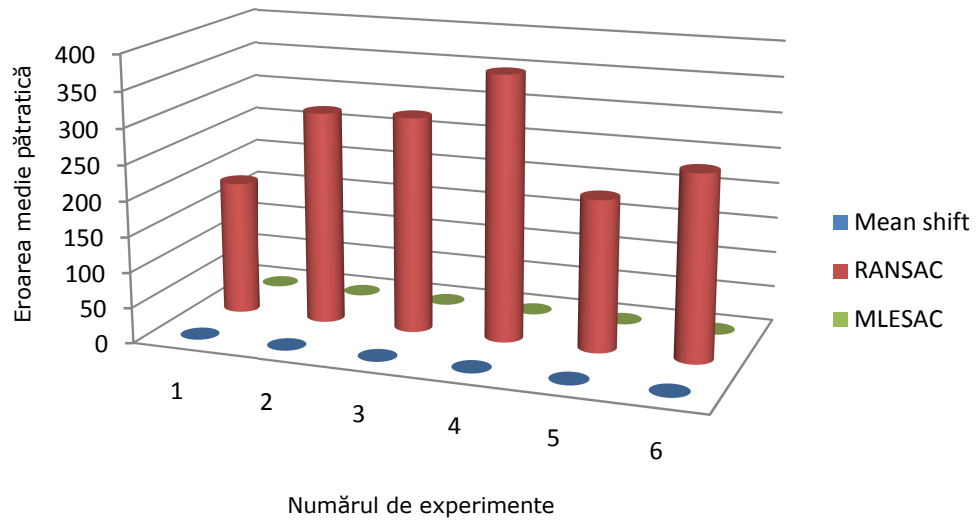


Fig. 5.29. Eroarea de scară pentru puncte puține și selecție manuală

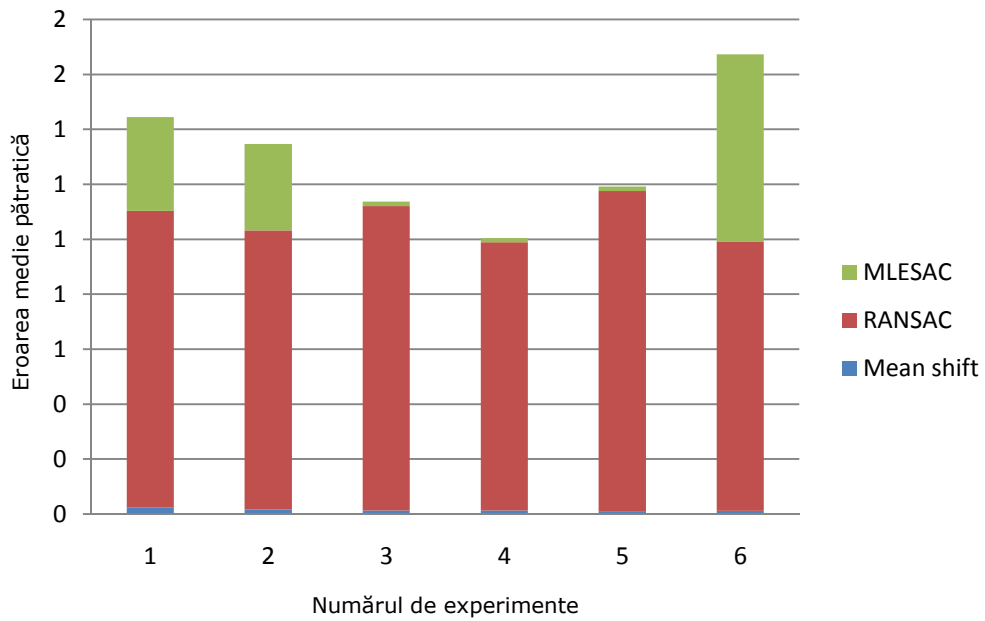


Fig. 5.30. Eroarea de unghi pentru puncte puține și selecție manuală

### 5.7.2.2. Experimente cu selecție manuală și puncte multe

În graficele 5.31-5.34 sunt prezentate erorile medii pătratice pentru toți vectorii de parametri. Numărul de puncte utilizat este între 15 și 33. Rezultatele experimentelor sunt similare cu cele de la selecția manuală și puncte puține pentru imaginile realizate cu o cameră mobilă. Algoritmul mean shift prezintă cele mai bune rezultate ale erorilor medii pătratice pentru estimarea parametrilor transformării de similaritate, decât MLESAC și RANSAC.

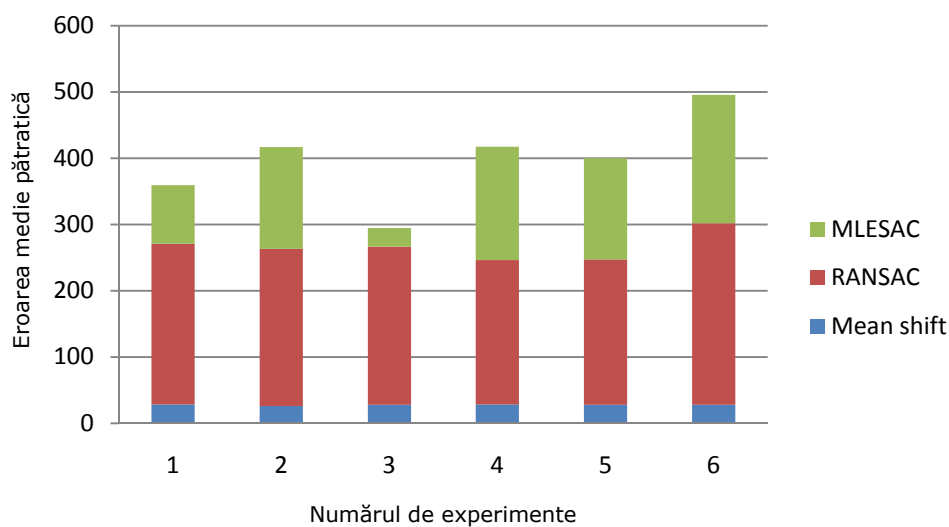


Fig. 5.31. Eroarea de translație pe orizontală pentru puncte multe și selecție manuală

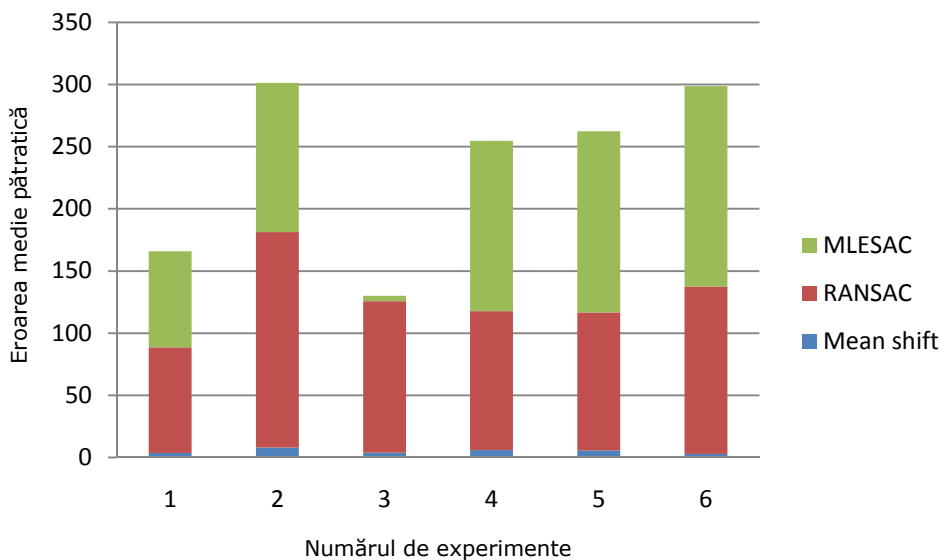


Fig. 5.32. Eroarea de translație pe verticală pentru puncte multe și selecție manuală.

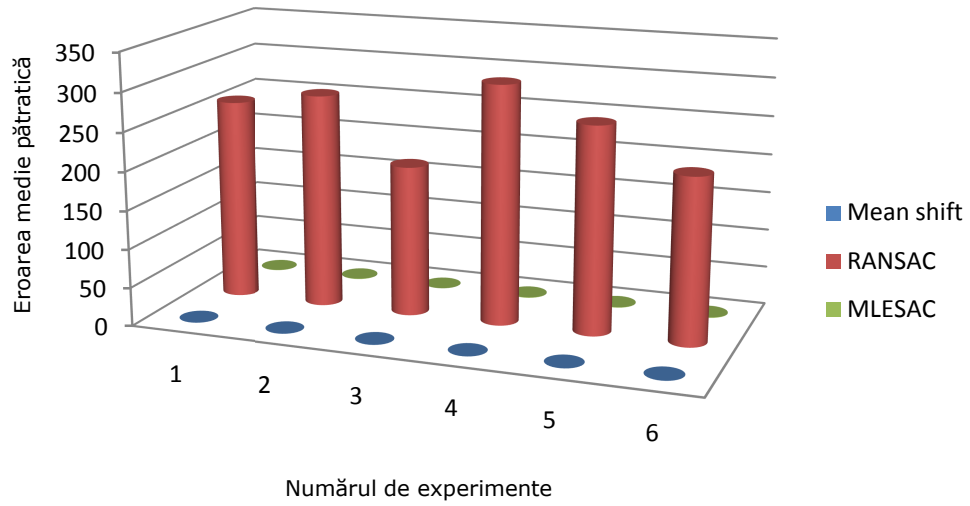


Fig. 5.33. Eroarea de scară pentru puncte multe și selecție manuală

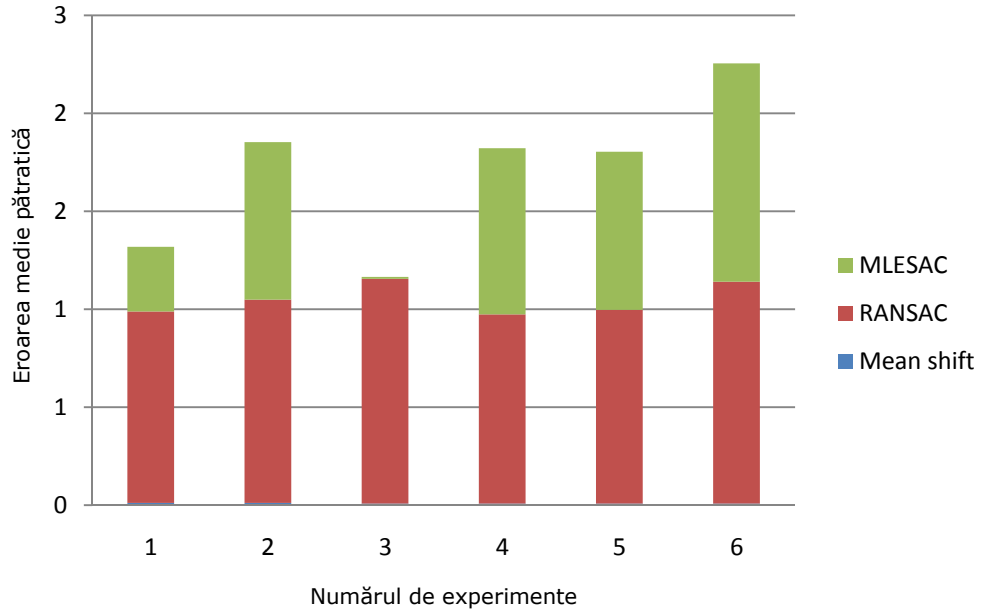


Fig. 5.34. Eroarea de unghi pentru puncte multe și selecție manuală.

### 5.7.2.3. Experimente cu selecție automată și puncte puține

Selecția punctelor caracteristice s-a realizat cu un program de selecție automată, SIFT, numărul punctelor generate având valori între 5 și 35. Primul set de experimente se realizează cu un număr mic de puncte, cuprins între 5 și 12.

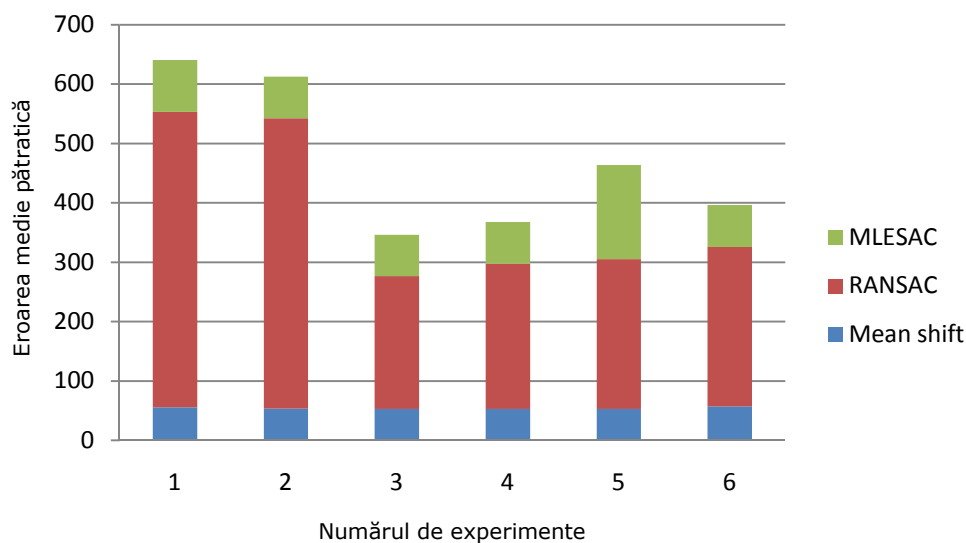


Fig. 5.35. Eroarea de translație pe orizontală pentru punte puține și selecție automată.

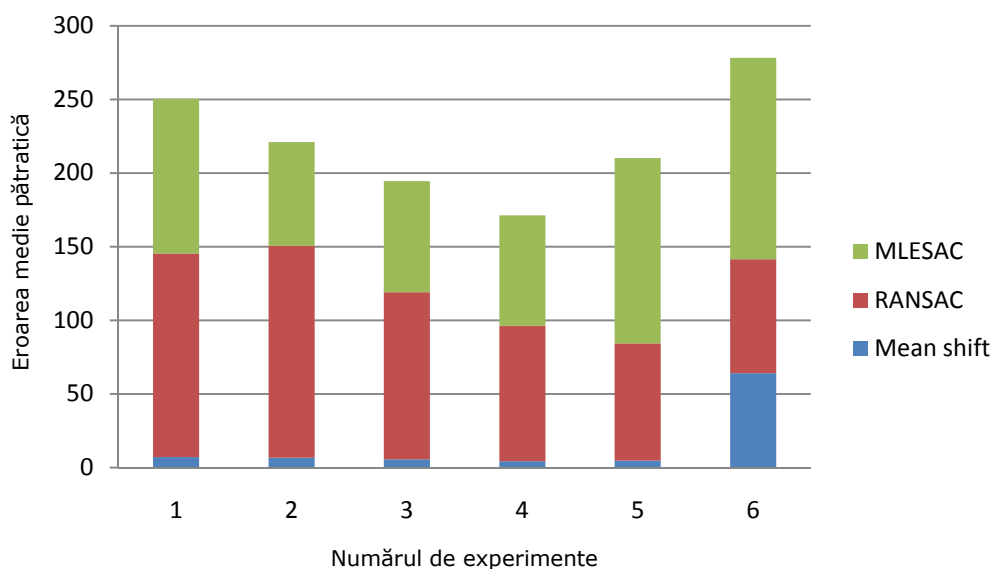


Fig. 5.36. Eroarea de translație pe verticală pentru punte puține și selecție automată.

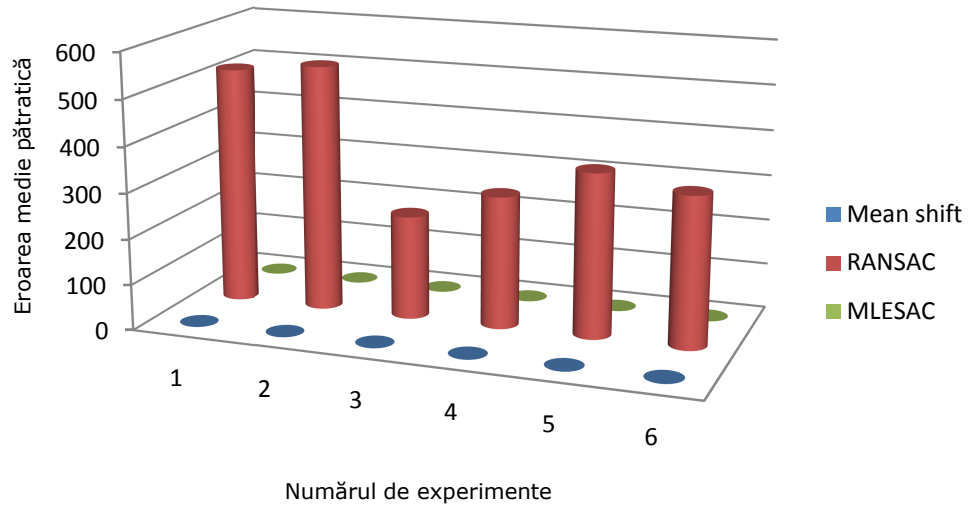


Fig. 5.37. Eroarea de scară pentru punte puține și selecție automată

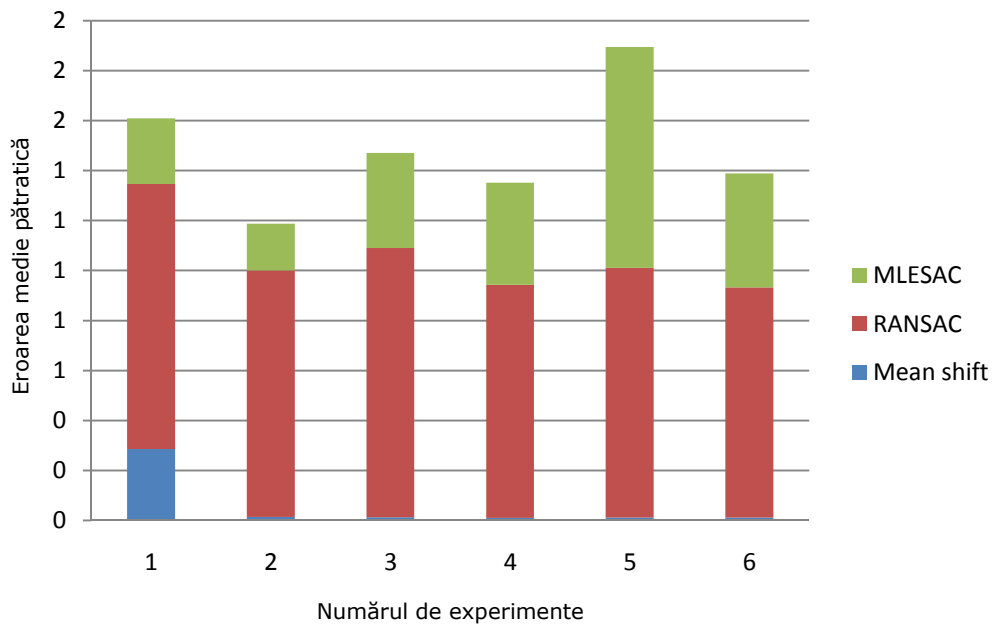


Fig. 5.38. Eroarea de unghi pentru punte puține și selecție automată.



Din figurile 5.35-5.38 se poate observa că eroarea de translație pe orizontală prezintă cele mai mici valori pentru mean shift, și cele mai mari pentru RANSAC. Eroarea de translație pe verticală prezintă erori similare pentru RANSAC și MLESAC, dar mult mai mari decât cele prezentate de mean shift. Eroarea de scară are valori mici pentru mean shift și erori mari pentru RANSAC, iar eroarea de unghi are valori mici pentru toate cele trei metode. Mean shift furnizează cele mai mici valori pentru eroarea de unghi.

#### 5.7.2.4. Experimente cu selecție automată și puncte multe

În graficele 5.39-5.42 sunt prezentate erorile medii pătratice pentru toți vectorii de parametrii. Numărul de puncte utilizat este între 15 și 35. Rezultatele experimentelor sunt similare cu cele de la selecția automată și puncte puține pentru imaginile realizate cu o cameră mobilă. Algoritmul mean shift prezintă cele mai bune rezultate ale erorilor medii pătratice pentru estimarea parametrilor transformării de similaritate, decât MLESAC și RANSAC.

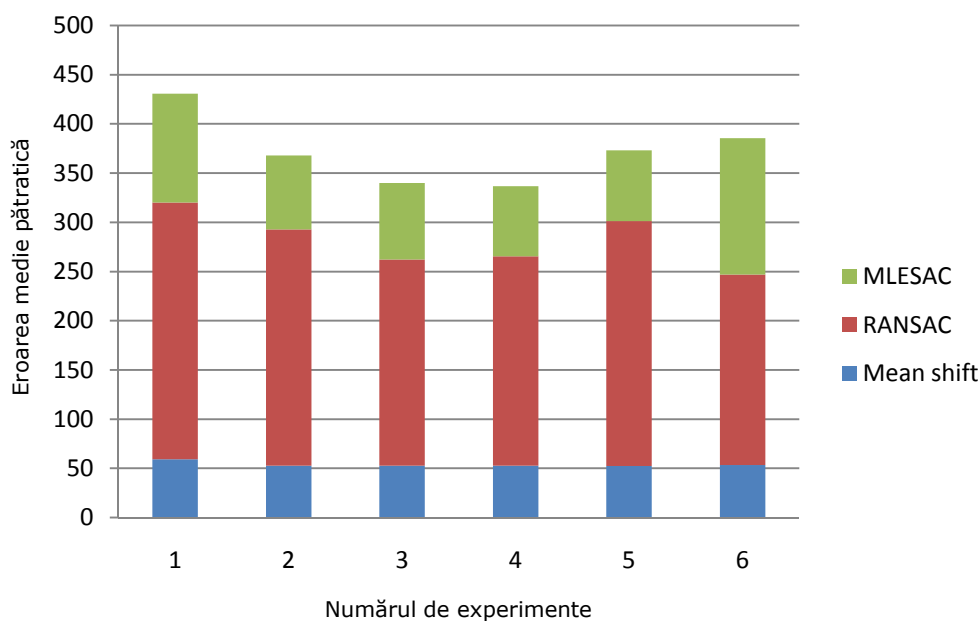


Fig. 5.39. Eroarea de translație pe orizontală pentru puncte multe și selecție automată.

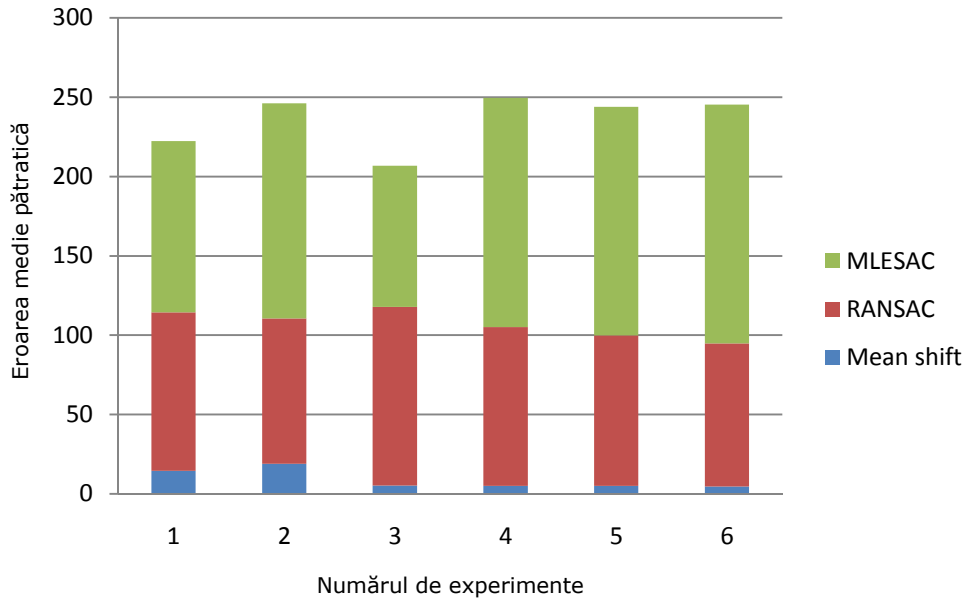


Fig. 5.40. Eroarea de translație pe verticală pentru puncte multe și selecție automată

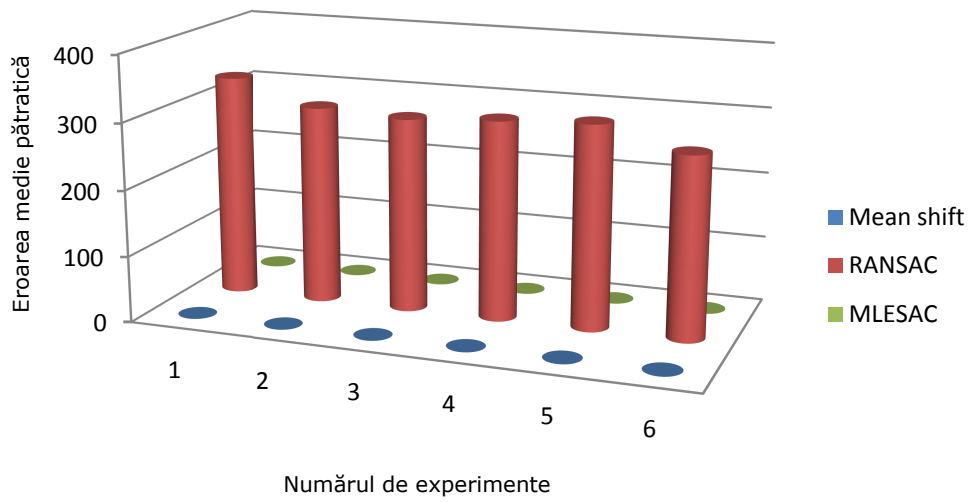


Fig. 5.41. Eroarea de scară pentru puncte multe și selecție automată

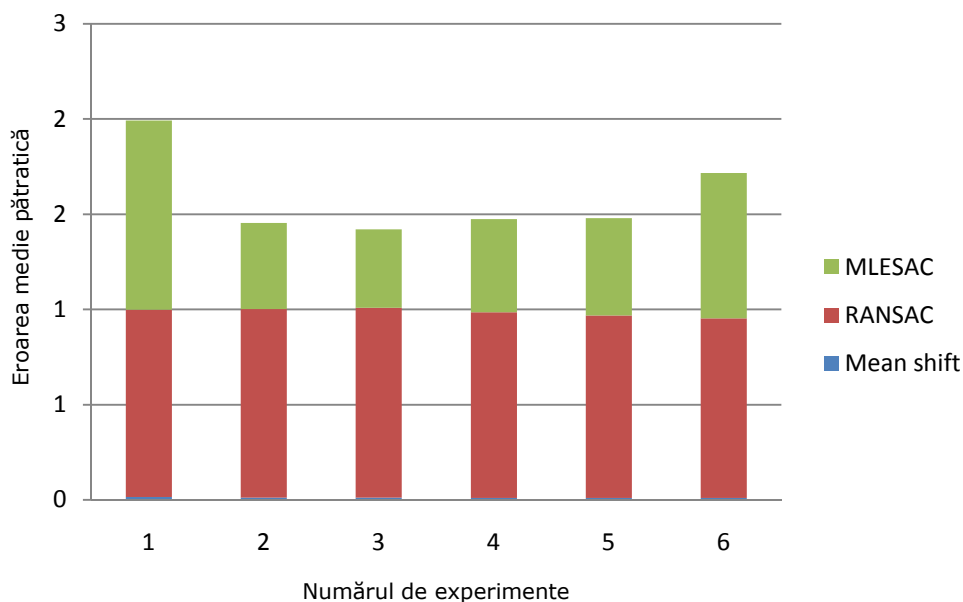


Fig. 5.42. Eroarea de unghi pentru puncte multe și selecție automată.

În concluzie, s-au efectuat diferite teste care presupun selecție manuală și automată de caracteristici. Numărul de caracteristici pe care se efectuează testele sunt diferite, fiind formate dintr-o combinație de date conforme și date neconforme. Imaginile utilizate la punerea în corespondență a acestora aparțin a două grupuri de imagini, prezentate în paragrafele anterioare.

Pentru ca algoritmul mean shift să realizeze alinierea robustă a două imagini necesită un număr minim de puncte corespondente. Dacă pe lângă numărul minim de puncte corespondente există și un număr, fie mare, fie mic, de puncte false, algoritmul mean shift nu întâmpină probleme în majoritatea cazurilor de a furniza un set de parametri apropiați de valorile reale. Un caz contrar îl reprezintă metodele de eșantionare aleatoare care în prezența unui minim de puncte corespondente însoțite de puncte false, furnizează în majoritatea cazurilor erori destul de mari. De asemenea, spre deosebire de mean shift, aceste metode oferă erori mari și pentru cazurile în numărul de date conforme este depășit de numărul de date neconforme.

Procesul de estimare robust realizat cu metodele de eșantionare aleatoare poate fi serios influențat de erorile cu care sunt amplasate punctele false. Chiar dacă numărul de puncte de corespondență este suficient de mare pentru a se putea determina un set de parametri corecți, metodele nu pot face față erorilor foarte mari ale punctelor false chiar dacă ele sunt puține.

În cazul selecției automate de puncte caracteristice, dacă acestea se găsesc într-un număr mare, atunci există posibilitatea, ca datorită scalării, ca unele puncte caracteristice să nu fie corespondente. Chiar dacă erorile de corespondență sunt mici, numărul mare de puncte influențează rezultatul metodelor de eșantionare aleatoare care poate fi deviat de la rezultatul real. Selecția automată poate prezenta erori de amplasare a punctelor necorespondente destul de mari atunci când se

lucrează cu imaginile captate de o cameră mobilă. Se poate evita acest lucru prin selecția manuală a unor puncte false cu erori de corespondență mici.

Trebuie menționat că datorită distorsiunilor introduse de camera foto, și anume rotație, scalare, translație, distorsiuni de perspectivă, alegerea unor puncte corespondente este un proces care presupune experiență din partea utilizatorului. Alegerea punctelor corespondente pentru grupul 1 nu este asemănătoare cu alegerea punctelor caracteristice la imaginile din grupul 2, datorită distorsiunilor de perspectivă introduse de camera foto. Datorită acestor distorsiuni, unele obiecte care apar într-o imagine, pot apărea diferite în cealaltă imagine, ca urmare, ca procesul de aliniere să fie reușit, aceste puncte trebuie alese pentru obiectele care rămân neschimbate în imagini. Acest procedeu poate influența și algoritmul mean shift.

## 5.8. Concluzii

Am demonstrat experimental că metoda de aliniere 2D a imaginilor propusă, care implică algoritmul neparametric mean shift, se poate folosi cu succes la alinierea robustă a două imagini. Mai mult decât atât, metoda propusă se poate folosi cu succes la estimarea câmpului vizual comun dintre senzorii video ai unei rețele de senzori video fără fir. În urma procesului de aliniere se extrag câmpurile vizuale comune dintre imaginile achiziționate de senzori, care se pot utiliza pentru urmărirea țintei în aplicațiile de supraveghere.

Am studiat două abordări cu scopul de a genera soluții parțiale pentru estimarea parametrilor transformării de similaritate. Printr-o analiză mai atentă, căutarea într-un spațiu 1D permite lucrul cu densități mari. Densitatea fiind invers proporțională cu volumul, va rezulta o dimensiune a spațiului mai mică. **Am constatat experimental că metoda de similaritate cea mai potrivită pentru un spațiu larg, în realizarea potrivirii caracteristicilor, este transformarea de similaritate simultană.** Astfel, se optează pe viitor lucrul cu un spațiu 4D, estimarea parametrilor realizându-se cu ajutorul unei transformări de similaritate simultane estimată pentru două perechi de puncte.

Am optimizat selecția caracteristicilor în cadrul metodei de aliniere a imaginilor prin propunerea unei noi metode. Aceasta realizează asocierea punctelor caracteristice, utilizând informația de culoare a pixelilor de pe medianele segmentelor care unesc perechi de puncte caracteristice detectate. În urma experimentelor, metoda de asociere a punctelor mediane pe baza culorii pixelilor s-a dovedit a fi mai puțin sensibilă la zgomot și oferă rezultate mai bune decât metodele tradiționale bazate pe măsurări făcute în jurul vecinătăților punctelor caracteristice. În plus metoda de asociere a punctelor mediane prezintă un timp de calcul scăzut și este văzută ca un test adițional care completează metodele de aliniere a imaginilor tradiționale. Cu toate acestea, prezintă și un dezavantaj, și anume utilizarea unui prag în procesul de eliminare al datelor aberante. **Astfel, am studiat experimental influența pragului de similaritate folosit în validarea corespondențelor și se stabilesc valori optime pentru un set de imagini. Se evaluează câștigul în viteza de calcul obținut prin aplicarea metodei de asociere a punctelor mediane.** Pragul optim găsit în urma experimentelor este acela care permite trecerea unui număr aproximativ de soluții apropiat de numărul real de corespondențe. Un prag cu o valoare mică sugerează că un număr mai mare de eşantioane eliminate este indicat decât un prag care permite eliminarea unui număr mic de eşantioane. Astfel, pragul se poate crește până când se obține un

număr rezonabil de puncte corespondente. Analizele teoretice și testele practice demonstrează că metodele de aliniere bazate pe extragerea caracteristicilor și metoda propusă pentru asocierea punctelor mediane, reprezintă o soluție fezabilă pentru rezolvarea problemei localizării. Mai mult, metoda de aliniere bazată pe estimarea neparametrică de densitate și algoritmul mean shift, demonstrează o acuratețe înaltă și un nivel de complexitate rezonabil.

În continuare, am **implementat un algoritm automat de extragere a caracteristicilor SIFT, care este folosit la localizarea senzorilor dintr-o rețea de senzori video fără fir. În aplicația de localizare a senzorilor dintr-o rețea fără fir, s-au făcut experimente pentru diferite medii: natură, clădiri, spații închise. S-au testat performanțele metodei automate SIFT și a celei manuale. Acestea s-au aplicat pe seturi de imagini reale, fiind supuse procesului de corespondență. În urma experimentelor, am comparat estimarea acurateței și eficiența timpului metodei automate în comparație cu metoda manuală.** Rezultatele experimentale arată acuratețea estimării și eficiența temporală a selecției automate în comparație cu selecția manuală. Rezultatele testelor demonstrează avantajele acestei tehnici în termeni ai timpului de execuție, pierderile de precizie fiind acceptabile.

Scopul acestei teze este de a demonstra că metoda neparametrică mean shift poate înlocui cu succes metodele de eșantionare aleatoare în estimarea parametrilor de translație, rotație și scalare pentru punerea în corespondență a imaginilor. Pentru a observa performanța algoritmului mean shift în metodele de punere în corespondență a imaginilor, se va compara cu metodele robuste de eșantionare aleatoare, RANSAC și MLESAC.

**Pe baza concluziilor rezultate în urma experimentelor, rezultă că pentru un număr mic de puncte și un număr mare de date neconforme, algoritmul mean shift obține cele mai mici erori pentru toți parametrii estimați decât metodele de eșantionare aleatoare RANSAC și MLESAC. Acest tip de situație intervine de obicei în rețelele de senzori video fără fir, în probleme de punere în corespondență a imaginilor, atunci când câmpul vizual comun dintre imagini este scăzut și punctele caracteristice sunt puține. Metoda mean shift s-a dovedit a fi stabilă în prezența unui procentaj variabil de mare de date neconforme și un număr mic de puncte caracteristice. Ca o concluzie, se poate spune că metoda robustă neparametrică mean shift poate înlocui cu succes metodele de eșantionare aleatoare RANSAC și MLESAC în cadrul procesului de aliniere a imaginilor, în special când numărul de puncte caracteristice este mic și conține date neconforme.**

Pentru testarea metodelor de estimare propuse în alinierea optimă a imaginilor, **am implementat o aplicație cu operare în timp real în VC++ destinată alinierii imaginilor având în vedere următoarele aspecte:**

- **selecția caracteristicilor manuală sau automată;**
- **asocierea caracteristicilor cu ajutorul unei transformări de similaritate simultane;**
- **estimarea parametrilor transformării cu ajutorul estimatorilor robuști. Drep urmare se implementează și se testează metodele robuste parametrice (RANSAC și MLESAC) și non-parametrice (mean shift).**
- **transformarea și reeșantionarea imaginii;**
- **implementarea algoritmului de asociere a punctelor mediane;**

**Softul a fost verificat pe aplicația de localizare a senzorilor dintr-o rețea de senzori video fără fir.**

## 6. CONCLUZII GENERALE ȘI CONTRIBUȚII

Teza de doctorat este dedicată unei probleme de mare interes în vederea artificială, și anume, alinierea imaginilor. Problema constă în determinarea transformării geometrice care aliniază partea comună a unei perechi de imagini ale aceleiași scene captate cu camere poziționate diferit, pe baza corespondențelor dintre două seturi de puncte, extrase din cele două imagini. Aplicația principală care a inspirat soluțiile adoptate este localizarea senzorilor într-o rețea fără fir. Multe dintre aplicațiile în care se utilizează aceste rețele depind de capacitatea lor în a-și determina locația nodurilor. Soluțiile bazate pe GPS nu sunt posibile datorită faptului că sunt scumpe și necesită echipamente hardware de putere mare. Un alt domeniu de aplicabilitate, vizat ca direcție viitoare, este cel al imaginilor biomedicale.

Perechile de imagini ce intervin într-o aplicație de localizare a senzorilor de imagine conțin frecvent un câmp vizual comun ce ocupă un procentaj redus din imagine și obiectele comune apar la scări diferite. În consecință, punctele caracteristice pe baza cărora se poate face estimarea transformării geometrice sunt puține, iar cele din imaginea aflată la distanță mai mare de cameră sunt afectate de erori de localizare mai mari. Metodele tradiționale bazate pe algoritmul RANSAC sau oricare din variantele sale ameliorate întâmpină dificultăți, pentru că sunt utilizate în condiții în care ipotezele avute în vedere de autorii acestora își pierd valabilitatea. Metodele de aliniere a imaginilor bazate pe estimare neparametrică de densitate de probabilitate, propuse în teză, sunt mai bine adaptate aplicației menționate și oricărei alte aplicații cu condiții similare, cum ar fi alinierea imaginilor stomatologice în vederea calculului procentului de remineralizare la tratamente cu laseri.

Spre deosebire de tehnicile de aliniere folosite anterior, optimizarea soluției prin metodele propuse în teză de doctorand, are loc în spațiul parametrilor, nu în cel al coordonatelor imaginii. Mai mult, având la bază o metodă neparametrică, performanțele metodelor nu depind de ipoteze legate de legile de distribuții ale erorilor eșantioanelor conforme modelului sau al celor neconforme modelului, spre deosebire de situația estimatorului MLESAC. Validitatea modelelor la estimarea parametrică devine dificil de verificat când eșantioanele de date sunt extrem de sărace. Metodele propuse beneficiază de eficiența algoritmului mean shift pentru determinarea maximelor estimatei densității de probabilitate, fără estimarea densității aferente.

Performanțele estimatorului robust bazat pe algoritmul mean shift au fost evaluate prin experimente, efectuate comparativ cu metoda clasică RANSAC și cu metoda ameliorată, MLESAC. Pe baza rezultatelor experimentale s-a demonstrat eficacitatea metodei mean shift în contextul alinierii imaginilor.

## 6.1. Contribuții teoretice

1. Principala contribuție a tezei este elaborarea și implementarea unei metode de aliniere a imaginilor, bazate pe algoritmul mean shift publicat de autor în [220], [225], [226], [227]. Contribuțiile inovatoare ale metodei se referă la următoarele aspecte:
  - a. Metoda calculează o soluție de maximă verosimilitate, prin maximizarea densității de probabilitate în spațiul parametrilor, spre deosebire de metodele de referință, ce folosesc optimizări în spațiul coordonatelor.
  - b. Algoritmul mean shift folosit în localizarea maximelor densității de probabilitate a vectorului parametrilor este fundamentat teoretic pe estimatie neparametrică, folosită în premieră în alinierea imaginilor.
  - c. Teza analizează teoretic avantajele potențiale ale noii abordări, comparativ cu soluțiile de referință și identifică aplicațiile ce pot beneficia de aceste avantaje.
  - d. Teza identifică aspectele teoretice pentru care algoritmi din familia RANSAC/MLESAC sunt mai puțin performanți în problemele de aliniere a imaginilor cu puncte caracteristice puține și erori de localizare mari [223].
  
2. Am realizat un studiu comparativ între metoda de estimare robustă neparametrică mean shift de punere în corespondență a imaginilor, utilizată în teză, și metodele clasice de eșantionare aleatoare RANSAC și MLESAC. Acest studiu comparativ realizat și publicat în [223], [224], a stabilit performanțele și acuratețea estimatorului mean shift în cazul imaginilor 2D care au un grad diferit de distorsiune și un factor de zgomot mic. Experimentele realizate confirmă avantajele metodei bazate pe estimarea densității de probabilitate folosind funcții nucleu.
  - a. Am demonstrat că metoda robustă neparametrică mean shift poate înlocui cu succes metodele de eșantionare aleatoare RANSAC și MLESAC în cadrul procesului de aliniere a imaginilor.
  - b. Testele au arătat că pentru un număr mic de puncte și un număr mare de date neconforme, algoritmul mean shift obține cele mai mici erori pentru toți parametrii estimați decât metodele de eșantionare aleatoare RANSAC și MLESAC. Acest tip de situație intervine de obicei în rețelele de senzori video fără fir, în probleme de punere în corespondență a imaginilor, atunci când câmpul vizual comun dintre imagini este scăzut și punctele caracteristice sunt puține.
  - c. Am constatat experimental stabilitatea rezultatelor estimatorului mean shift în prezența unui procentaj variabil mare de date neconforme și un număr mic de puncte caracteristice.

3. Teza propune selecția punctelor corespondente și eliminarea corespondențelor incompatibile prin metoda de asociere a punctelor mediane, folosind similaritatea culorilor și un prag de decizie. Metoda propusă a fost publicată în [193], [194], [195].
  - a. În cadrul metodei propuse, am studiat experimental influența pragului de similaritate folosit în validarea corespondențelor și am stabilit valori optime pentru un set de imagini.
  - b. Am evaluat câștigul în viteza de calcul obținut prin aplicarea metodei de asociere a punctelor mediane.
  
4. Teza propune și compară în [191] două versiuni de estimare a parametrilor transformării de similaritate prin algoritmul mean shift: secvențial și simultan.
  - a. Am stabilit și justificat teoretic ordinea de estimare cea mai avantajoasă la estimatorul secvențial.
  - b. Am constatat experimental că metoda de calcul cea mai potrivită pentru un spațiu larg, în realizarea corespondenței caracteristicilor, este transformarea de similaritate secvențială.

## 6.2. Contribuții aplicative

1. Am implementat în metoda de aliniere a imaginilor algoritmul automat de extragere a caracteristicilor SIFT, care poate fi folosit la localizarea senzorilor dintr-o rețea de senzori video fără fir, utilizând corespondența imaginilor. Studiul a fost realizat și publicat în [196], [197].
  - a) În aplicația de localizare a senzorilor dintr-o rețea fără fir, am efectuat experimente pentru diferite medii: natură, clădiri, spații închise.
  - b) Am testat performanțele metodei automate SIFT și a celei manuale. Acestea s-au aplicat pe seturi de imagini reale, fiind supuse procesului de corespondență.
  - c) În urma experimentelor, am comparat estimarea acurateții și eficiența timpului metodei automate în comparație cu metoda manuală.
  
2. Am definit ca direcție de aplicabilitate pentru metoda originală de aliniere a imaginilor, utilizând estimatorul robust bazat pe estimarea neparametrică de densitate de probabilitate, localizarea senzorilor în rețelele de senzori video fără fir [191], [193], [194], [196], [197], [181]. Procesul de aliniere a imaginilor captate de la întreaga rețea de senzori, este utilizat atât pentru determinarea localizării acestora, cât și pentru determinarea câmpului vizual comun dintre perechile de camere.
  
3. Am implementat o aplicație cu operare în timp real în VC++ destinată alinierii imaginilor având în vedere următoarele aspecte:
  - a) selecția caracteristicilor manuală sau automată;
  - b) asocierea caracteristicilor cu ajutorul unei transformări de similaritate simultane;



- c) estimarea parametrilor transformării cu ajutorul estimatorilor robuști. Drept urmare se implementează și se testează metodele robuste de eșantionare aleatoare (RANSAC și MLESAC) și neparametrice (mean shift).
  - d) transformarea și reeșantionarea imaginii;
  - e) implementarea algoritmului de asociere a punctelor mediane;
- Softul a fost verificat pe aplicația de localizare a senzorilor dintr-o rețea de senzori video fără fir.

### 6.3. Direcții de cercetare. Obiective pentru viitor

Metodele de aliniere originale propuse în această teză se pot utiliza în aplicații ca rețele de senzori fără fir. La scară largă, aceste rețele sunt compuse din sute sau mii de noduri senzori autonome. Aceștia operează în absența unui infrastructuri stabilite de dinainte, se configurează singure, de cost scăzut și pot lucra în medii ostile. Senzorii pot interactiona cu mediul inconjurator prin monitorizare și măsurare a intensității luminii, căldurii, poziția, mișcarea și prezența chimică. Informația colectată de la senzori este apoi transmisă către celelate noduri ale rețelei. Alinierea imaginilor intervine în cazul acestor rețele pentru localizarea lor, pe baza informației colectată de la aceștia, și pentru a estima suprapunerea câmpului vizual comun.

Ca viitoare direcție de cercetare se urmărește combinarea metodelor bazate pe puncte cu cele pe regiuni folosind metode de segmentare care generează regiuni bune cu aplicare în domeniul medical. Metodele de aliniere a imaginilor fac posibilă suprapunerea caracteristicilor luate de la un studiu cu cele care aparțin altui studiu. De exemplu, structurile scheletice și ariile de contrast văzute în imaginile CT (computed tomography) se pot suprapune peste imaginile MR (magnetic resonance) care prezintă anatomia țesutului moale. Acest proces se poate aplica pe multiple seturi de date obținute din aceeași senzori dar la momente de timp diferite. Imaginile pot fi supuse procesului de aliniere, cu scopul unei comparații cantitative care va crește precizia monitorizării tratamentului.

## BIBLIOGRAFIE

- [1] P. Meer, "Robust techniques for computer vision," in *Emerging Topics in Computer Vision*: G. Medioni and S. B. Kang (Eds.), Prentice Hall, 2004, pp. 107-190.
- [2] M. A. Fischler and R. C. Bolles, "Random consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, pp. 381-395, 1981.
- [3] D. Comaniciu and P. Meer, "Mean Shift: A robust approach toward feature space analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, pp. 603-619, 2002.
- [4] L. G. Brown, "A Survey of Image Registration Techniques," *ACM Computing Surveys*, vol. 24, pp. 325-376, 1992.
- [5] B. Zitova and J. Flusser, "Image registration methods: a survey," *Department of Image Processing, Institute of Information theory and Automation*, 2003.
- [6] X. Dai and S. Khorram, "The effects of image misregistration on the accuracy of remotely sensed change detection," *IEEE Trans. Geosci. Remote Sensing*, vol. 36, pp. 1566-1577, Sept. 1998.
- [7] L. M. G. Fonseca and B. S. Manjunath, "Registration techniques for multisensor remotely sensed imagery," *Photogramm. Eng. Remote Sensing*, vol. 562, pp. 1049-1056, Sept. 1996.
- [8] A. D. Ventura, A. Rampini, and R. Schettini, "Image registration by recognition of corresponding structures," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 28, pp. 305 - 314, 1990.
- [9] Q. Zheng and R. Chellappa, "A computational vision approach to image registration," in *Computer Vision and Applications, Proceedings, 11th IAPR International Conference on Pattern Recognition*. vol. 1, 1992, pp. 193-197.
- [10] S. T. Barnard and M. A. Fischler, "Computational Stereo," *ACM Computer Surveys* 14, vol. 4, pp. 553-572, 1982.
- [11] H. J. Lee and W. L. Lei, "Region matching and depth finding for 3D objects in stereo aerial photographs," *Pattern recognition*, vol. 23, pp. 81-94, 1990.
- [12] Y. C. Hsieh, D. M. McKeown, and F. P. Perlant, "Performance evaluation of scene registration and stereo matching for cartographic feature extraction," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, pp. 214-237, 1992.
- [13] Q. Zheng and R. Chellapa, "Automatic feature point extraction and tracking in image sequences for arbitrary camera motion," *International Journal of Computer Vision*, vol. 15, pp. 31 - 76 1995.
- [14] K. S. Arun, T. S. Huang, and S. D. Blostein, "Least-squares fitting of two 3-D point sets," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 9, pp. 698 - 700, 1987.
- [15] J. J. Rodriguez and J. K. Aggarwal, " Matching aerial images to 3-D terrain maps," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, pp. 1138-1149, 1990.
- [16] L. G. Roberts, "Machine Perception of 3-D Solids," in *Ph.D. Thesis, MIT*, 1963.
- [17] P. E. Anuta, "Registration of multispectral video imagery," *Society Photo-Optical Instrum. Eng. J.*, vol. 7, pp. 168-175, 1969.

- [18] P. E. Anuta, "Spatial registration of multispectral and multitemporal digital imagery using fast Fourier transform techniques," *IEEE Trans. Geoscience Electronics*, vol. 8, pp. 353-368, 1970.
- [19] D. Barnea and H. Silverman, "A class of algorithms of fast digital image registration," *IEEE Trans. Comput.*, vol. 21 (2), pp. 179-186, 1972.
- [20] J. A. Leese, G. S. Novak, and B. B. Clark, "An automatic technique for obtaining cloud motion from geosynchronous satellite data using cross correlation," *A. Applied Meteorology*, vol. 10, pp. 110-132, 1971.
- [21] W. K. Pratt, "Correlation techniques for image registration," *IEEE Trans. Aerospace and Electronic Systems*, vol. 10, pp. 353-358, 1974.
- [22] K. Mori, M. Kidode, and H. Asada, "An Iterative Prediction and Correction Method for Automatic Stereocomparison," *Computer Graphics and Image Processing*, vol. 2, pp. 393-401, 1973.
- [23] M. D. Levine, D. A. O'Handley, and G. M. Yagi, "Computer Determination of Depth Maps," *Computer Graphic and Images Processing*, vol. 2, pp. 131-150, 1973.
- [24] R. Nevatia, "Depth measurement by motion stereo," *Computer Graphics and Image Processing*, vol. 15, pp. 203-214, 1976.
- [25] D. C. Barber, "Automatic alignment of radionuclide images," *Physics in Medicine and Biology*, vol. 27, pp. 387-396, 1982.
- [26] M. Singh, W. Frei, T. Shibata, G. C. Huth, and N. E. Telfer, "A Digital Technique for Accurate Change Detection in Nuclear Medical Images – with application to Myocardial Perfusion Studies using Thallium-201," *IEEE Trans. On Nuclear Science NS-26*, vol. 1, Feb. 1979.
- [27] A. Venot and V. Leclerc, "Automated correlation of patient motion and gray values prior to subtraction in digitized angiography," *IEEE Trans. Medical Imaging*, vol. 8, pp. 179-186, 1984.
- [28] J. V. Hajnal, D. L. G. Hill, and D. J. Hawkes, "Medical Image Registration," *CRS Press*, 2001.
- [29] F. Pernus, S. H. Siegfried, and M. Vierveger, " Biomedical Image Registration," *a special issue of Image and Vision Computing*, January 2001.
- [30] J. P. W. Pluim and J. M. Fitzpatrick, "Image Registration," *a special issue of IEEE Trans. Medical Imaging*, Nov. 2003.
- [31] A. Goshtasby and J. LeMoigne, "Image Registration," *a special issue of Pattern Recognition*, Jan. 1999.
- [32] A. Goshtasby, L. Staib, C. Studholme, and D. Terzopoulos, "Non-Rigid Image Registration," *a special issue of Computer Vision and ImageUnderstanding*, Feb. 2003.
- [33] P. Gerlot and Y. Bizais, "Image registration: A review and a strategy for medical application," in *Information Processing in Medical Imaging*, New York, 1988, pp. 81-89.
- [34] P. A. V. d. Elsen, E.-J. D. Pol, and M. A. Viergever, "Medical image matching: A review with classification," *IEEE Eng. Med. Biol.*, vol. 12, pp. 26-39, 1993.
- [35] C. R. Maurer and J. M. Fitzpatrick, "A review of medical image registration," *Interactive Image Guided Neurosurgery*, pp. 17-44, 1996.
- [36] J. B. A. Maintz and M. A. Viergever, "A survey of medical image registration," *Medical Image Analysis (in press)*, vol. 2, pp. 1-36, 1988.
- [37] H. Lester and S. R. Arridge, "A survey of hierarchical non-linear medical image registration," *Pattern Recognition*, vol. 32, pp. 129-149, 1999.

- [38] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever, "Mutual-information-based image registration of medical images: A survey," *IEEE Trans. Medical Imaging*, vol. 22, pp. 986-1004 2003.
- [39] L. Chmielewski and D. Kozinska, "Image Registration," in *Computer Pattern Recognition Systems KOSYR Poland*, May 26-29, 2001.
- [40] A. Goshtasby and A. Ardesb, "2-D and 3-D Image Registration: For Medical, Remote Sensing, and Industrial Applications," N. Y. Wiley, Ed., 2005.
- [41] C. V. Rao, K. M. M. Rao, A. S. Manjunath, and R. V. N. Srinivas, "OPTIMIZATION OF AUTOMATIC IMAGE REGISTRATION ALGORITHMS AND CHARACTERIZATION."
- [42] H. Bay, V. Ferrari, and L. V. Gool, "Wide-baseline stereo matching with line segments," in *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, 2005, pp. 329-336.
- [43] M. Xia and B. Liu, "Image registration by "super-curve"," *IEEE Transactions on Image Processing*, vol. 13, pp. 720-732, May 2004.
- [44] A. Bartoli, E. v. Tunzelmann, and A. Zisserman, "Augmenting images of non-rigid scenes using point and curve correspondences," in *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'04)*, 2004, pp. 699-706.
- [45] Y. Li, R. L. Stevenson, and J. Gai, "Corner-Guided Image Registration by using Edges," *IEEE International Conference on Image Processing (ICIP)* vol. 5, pp. 361-364, 2007.
- [46] S. Kruger and A. Calway, "Image Registration Using Multiresolution Frequency Domain Correlation," in *Proceedings British Machine Vision Conference*, 1998, pp. 316-325.
- [47] J. Yang, R. S. Blum, J. P. Williams, Y. Sun, and C. Xu, "Non-rigid Image Registration Using Geometric Features and Local Salient Region Features," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. 825- 832, June 2006
- [48] D. G. Lowe, "Object recognition from local scale-invariant features," in *Proc. 7th Int'l Conf. Computer Vision*, Kerkyra, Greece, 1999, pp. 1150-1157.
- [49] J. J. Koenderink, "The structure of images," *Biological Cybernetics*, vol. 50, pp. 363-396, 1984.
- [50] T. Lindeberg, "Scale-space theory: A basic tool for analysing structures at different scales," *Journal of Applied Statistics*, vol. 21, pp. 224-270, 1994.
- [51] J.-W. Hsieh, H.-Y. M. Liao, K.-C. Fan, M.-T. Ko, and Y.-P. Hung, "Image Registration Using a New Edge-Based Approach," *COMPUTER VISION AND IMAGE UNDERSTANDING*, vol. 67, pp. 112-130, August 1997.
- [52] H. Chui and A. Rangarajan, "A new point matching algorithm for non-rigid Registration," *Elsevier Science*, October 2002.
- [53] H. Chui, L. Win, J. Duncan, R. Schultz, and A. Rangarajan, "A unified non-rigid feature registration method for brain mapping," *Medical Image Analysis (in press)*, 2002.
- [54] D. M. Mount, N. S. Netanyahu, and J. L. Moigne, "Efficient Algorithms for Robust Feature Matching," in *Proceedings of the CESDIS Image Registration Workshop NASA Goddard Space Flight (GSFC)*, 1997, pp. 247-256.
- [55] S. Belongie, J. Malik, and J. Puzicha, "Shape matching and object recognition using shape contexts," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24 pp. 509-522, 2002.

- [56] D. P. Huttenlocher, G. A. Klanderman, and W. J. Rucklidge, "Comparing images using the Hausdorff distance," *IEEE Trans. Pattern Intelligence and Machine Intelligence*, vol. 9, pp. 850-863, 1993.
- [57] J. W. Rucklidge, "Efficiently locating objects using the Hausdorff distance," *Int'l J. Computer Vision*, vol. 24, pp. 251-270, 1997.
- [58] V. Lepetit, J. Pilet, and P. Fua, "Point Matching as a Classification Problem for Fast and Robust Object Pose Estimation," *Conference on Computer Vision and Pattern Recognition*, 2004.
- [59] E. Delponte, F. Isgrò, F. Odone, and A. Verri, "SVD-matching using SIFT features," in *Proceedings of the of the International Conference on Vision, Video and Graphics*, Edinburgh, UK, 2005, pp. 125-132.
- [60] G. Stockman, S. Kopstein, and S. Bennett, "Matching images to models for registration and object detection via clustering," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 4, pp. 229-241, 1982.
- [61] T. J. Fang, Z. H. Huang, L. N. Kanal, B. D. Lavine, G. Stockman, and F. L. Xiong, "Three-dimensional object recognition using a transformation clustering technique," *6th Int'l Conf. Pattern Recognition*, pp. 678-681, 1982.
- [62] G. Stockman, "Object recognition and localization via pose clustering," *Computer Vision, Graphics, and Image Processing*, vol. 40, pp. 361-387, 1987.
- [63] G. Stockman and J. C. Esteva, "Use of geometrical constraints and clustering to determine 3-D object pose," *7th International Conference Pattern Recognition*, vol. 2, pp. 742-744, 1984.
- [64] D. H. Ballard, "Generalizing the Hough transform to detect arbitrary shapes," *Pattern Recognition*, vol. 13, pp. 111-122, 1981.
- [65] W. E. L. Grimson and D. P. Huttenlocher, "On the sensitivity of the Hough transform for object recognition," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 12, pp. 255-276, 1990.
- [66] A. Goshtasby and G. C. Stockman, "A region-based approach to digital image registration with subpixel accuracy," in *IEEE Transactions Geosci. Remote Sensing*, vol. 24 (3), 1986, pp. 390-399.
- [67] M. Dhome and T. Kasvand, "Polyhedra recognition by hypothesis accumulation," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 9, pp. 429-438, 1987.
- [68] S. Linnainmaa, D. Harwood, and L. S. Davis, "Pose determination of a three-dimensional object using triangle pairs," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 10, pp. 634-647, 1988.
- [69] T. M. Silberberg, D. A. Harwood, and L. S. Davis, "Object recognition using oriented model points," *Computer Vision, Graphics, and Image Processing*, vol. 35, pp. 47-71, 1986.
- [70] D. W. Thompson and J. L. Mundy, "Three-dimensional model matching from an unconstrained viewpoint," *Proc. IEEE Conf. Robotics and Automation*, pp. 208-220, 1987.
- [71] C. V. Stewart, "Robust parameter estimation in computer vision," *SIAM Reviews*, pp. 513-537, September 1999.
- [72] X. Dai and S. Khorram, "A feature-based image registration algorithm using improved chain-code representation combined with invariant moments," *IEEE Trans. Geosci. Remote Sens.*, vol. 37 (5), pp. 2351-2362, 1999.

- [73] B. K. Ghaffary and A. A. Sawechuk, "A survey of new techniques for image registration and mapping," in *Proceedings SPIE: Applications of Digital Image Processing 432*, 1983, pp. 222-239.
- [74] A. Goshtasby, "Image Resampling," *Wright State University and Image Fusion Systems Reseach*, June 2004.
- [75] J. Sachs, "Image Resampling," 2001.
- [76] N. A. Dodgson, "Quadratic interpolation for image resampling," *IEEE Transactions on Image Processing*, vol. 6, pp. 1322-1326, 1997.
- [77] K. Toraichi, S. Yang, and R. Mori, "Two-dimensional spline interpolation for image reconstruction," *Pattern Recognition Letters*, vol. 21, pp. 275-284, 1988.
- [78] H. S. Hou and H. C. Andrews, "Cubic splines for image interpolation and digital filtering," *IEEE Transactions on Acoustic, Speech and Signal Processing*, vol. 26, pp. 508-517, 1978.
- [79] T. M. Lehmann, C. Go"nner, and K. Spitzer, "Addendum: B-spline interpolation in medical image processing," *IEEE Transaction on Medical Imaging*, vol. 20 pp. 660-665, 2001.
- [80] R. G. Keys, "Cubic convolution interpolation for digital image processing," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 29 pp. 1153-1160, 1981.
- [81] P. Thevenaz, U. E. Ruttimann, and M. Unser, "A pyramidal approach to subpixel registration based on intensity," *IEEE Transactions on Image Processing*, vol. 7, pp. 27-41, 1998.
- [82] C. R. Appledorn, "A new approach to the interpolation of sampled data," *IEEE Transactions on Medical Imaging*, vol. 15, pp. 369-376, 1996.
- [83] P. Thevenaz, T. Blu, and M. Unser, "Image interpolation and resampling," in *Handbook of Medical Image Processing*. vol. Academic Press, New York, 2003.
- [84] J. Meyer, "Histogram Transformation for Inter-Modality Image Registration," in *IEEE 7th International Symposium on BioInformatics and BioEngineering (IEEE BIBE 2007)*, Oct. 2007, pp. 1118-1123.
- [85] D. Shen, "Image Registration by Hierarchical Matching of Local Spatial Intensity Histograms," in *7th International Conference Medical Image Computing and Computer-Assisted Intervention - MICCAI'04*, Saint-Malo, France, Sept. 2004, pp. 582-590.
- [86] Z. P. Liang, H. Pan, R. L. Magin, N. Ahuja, and T. S. Huang, "Automated registration of multimodality images by maximization of a region similarity measure," in *IEEE International Conference on Image Processing Santa Barbara, California*, Oct. 1997, pp. 272-275.
- [87] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, "Multimodality image registration by maximization of mutual information," *IEEE transactions on Medical Imaging*, pp. 187-198, 1997.
- [88] G. Hong and Y. Zhang, "An Image Registration Technique of High Resolution Remote Sensing Image in Hilly Area," in *IEEE/ISPRS 3rd International Symposium Remote Sensing and Data Fusion Over Urban Areas (URBAN 2005)* Tempe, Arizona, USA, March 2005.
- [89] Y. Chen and A. K. Jain, "DOTS AND INCIPIENTS: Extended Features For Partial Fingerprint Matching," in *Proceedings of Biometric Symposium, Biometric Consortium Conference*, Baltimore, 2007.



- [90] C. Hong, J. Tian, and X. Yang, "Fingerprint matching with registration pattern inspection " in *4th International Conference on Audio-and Video-Based Biometric Person Authentication (AVBPA)*, Guilford, UK, June 2003.
- [91] J. Tian, Y. He, X. Yang, L. Li, and X. Chen, "Improving Fingerprint Recognition Performance Based on Feature Fusion and Adaptive Registration Pattern," in *5th Proceedings Chinese Conference on Biometric Recognition, SINOBIOOMETRICS 2004*, Guangzhou, China, 2004, pp. 57-66.
- [92] A. Jain and H. Chen, "Registration of dental atals to radiographs for human indetification," in *Proceedings of SPIE Conference on Biometric Technilogy for Human Identification II*, Orlando, Florida, 2005, pp. 292-298.
- [93] P. Besl and N. McKay, "A Method for Registration of 3-D Shapes," *IEEE Trans. Pattern Analysis and Machine Intelligence* 14, pp. 239-256, 1992.
- [94] J. Pita, "An Implementation of the Iterative Closest Point Algorithm," 2007.
- [95] T. Jost and H. Hugli, "A multi-resolution ICP with heuristic closest point search for fast and robust 3-D registration of range images," in *Proc. 4th International Conference 3-D Digital Imaging and Modeling*, Banff, Alberta, Canada, Oct. 2003, pp. 427-433.
- [96] H. Schonfeld, G. Hausler, and S. Karbacher, "Reverse engineering using optical 3-D sensors, 3-D Image Capture and Applications," *Procedeedings SPIE*, vol. 3313, 1998.
- [97] N. Gelfand, L. Ikemoto, S. Rusinkiewicz, and M. Levoy, "Geometrically stable sampling for the ICP algorithm," in *Proc. 4th Int'l Conf. 3-D Digital Imaging and Modeling*, Banff, Alberta, Canada, Oct. 2003, pp. 260-267.
- [98] L. Silva, O. R. P. Bellon, and K. L. Boyer, "Enhanced, robust genetic algorithms formultiviewrange image registration," in *Proc. 4th Int'lConf. 3-D Digital Imaging and Modeling*, Banff, Alberta, Canada, Oct. 2003, pp. 268-275.
- [99] D. Tubic, P. Hebert, and D. Laurendeau, "A volumetric approach for interactive 3-D modeling," in *Proc. 1st Int'l Sym. 3-D Data Processing Visualization and Transmission*, Padova, Italy, June 2002, pp. 150-158
- [100] S. El-Hakim, J. A. Beraldin, and J. F. Lapointe, "Toward automatic modeling of monuments and towers," in *Proc. 1st Int'l Sym. 3-D Data Processing Visualization and Transmission*, Padova, Italy, June 2002, pp. 526-531.
- [101] J. V. Wyngaerd and L. V. Gool, "Combining texture and shape for automatic crude patch registration," in *Proc. 4th Int'lConf. 3-D Digital Imaging and Modeling*, Banff, Alberta, Canada, Oct. 2003, pp. 179-186.
- [102] P. W. Smith and M. D. Elstrom, "Automatic feature correspondence for scene reconstruction from multiple views," in *Proc. 2nd Int'l Conf. 3-D Digital Imaging and Modeling*, Ottawa, Canada, Oct. 1999, pp. 463-472.
- [103] L. Ikemoto, N. Gelfand, and M. Levoy, "A hierarchical method for aligning warped meshes," in *Proc. 4th Int'l Conf. 3-D Digital Imaging and Modeling*, Banff, Alberta, Canada, Oct. 2003, pp. 434-441.
- [104] K. Pulli, "Multiview registration for large data sets," in *Proc. 2nd Int'l Conf. 3-D Digital Imaging and Modeling*, Ottawa, Canada, Oct. 1999, pp. 160-168.
- [105] Z. Y. Zhang, "Iterative point matching for registration of free-form curves and surfaces," *International Journal of Computer Vision*, vol. 13, pp. 119-152, October 1994.
- [106] T. Masuda and N. Yokoya, "A robust method for registration and segmentation of multiple range images," *Computer Vision and Image Understanding*, vol. 61, pp. 295-307, May 1995.

- [107] Y. Chen and G. G. Medioni, "Object modeling by registration of multiple range images," *Image and Vision Computing*, vol. 10, pp. 145-155, 1992.
- [108] C. Dorai, J. Weng, and A. K. Jain, "Optimal registration of object views using range data," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, pp. 1131-1138, October 1997.
- [109] R. Bergevin, M. Soucy, H. Gagnon, and D. Laurendeau, "Towards a general multi-view registration technique," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 18, pp. 540-547, May 1996.
- [110] D. W. Eggert, A. W. Fitzgibbon, and R. B. Fisher, "Simultaneous registration of multiple range views for use in reverse engineering of cad models," *Computer Vision and Image Understanding*, vol. 69, pp. 253-272, March 1998.
- [111] A. J. Stoddart and A. Hilton, "Registration of multiple point sets. In International Conference on Pattern Recognition," 1996.
- [112] K. E. Boehnke, "Hierarchical Object Localization for Robotic Bin Picking," 2008.
- [113] M. Singh, H. Arora, and N. Ahuja, "Robust Registration and Tracking Using Kernel Density Correlation," in *Proc. Conf. Computer Vision and Pattern Recognition Workshop (CVPRW '04)*, 2004, p. 174.
- [114] P. Burt, J. Bergen, R. Hingorani, R. Kolczynski, W. Lee, A. Leung, J. Lubin, and H. Shvaytser, "Object tracking with a moving camera," in *IEEE Workshop on Visual Motion*, 1989, pp. 2-12.
- [115] M. Irani, B. Rousso, and S. Peleg, "Detecting and tracking multiple moving objects using temporal integration," in *Second European Conference on Computer Vision Santa Margherita, Italy*, 1992, pp. 282-287.
- [116] P. Nordlund and T. Uhlin, "Closing the loop: Detection and pursuit of a moving object by a moving observer.," *Image and Vision Computing*, vol. 14, pp. 265-275, 1996.
- [117] Y. Rosenberg and M. Werman, "Representing local motion as a probability distribution matrix and object tracking," in *DARPA Image Understanding Workshop New Orleans, Louisiana*, 1997, pp. 153-158.
- [118] Y. Rosenberg and M. Werman, "Real-Time Object Tracking from a Moving Video Camera: A Software Approach on a PC," in *IEEE Workshop on Applications of Computer Vision Princeton*, 1998, pp. 238-239.
- [119] H. Eltoukhy and K. Salama, "Multiple Camera Tracking," Stanford University 2002.
- [120] S. Khan and M. Shah, "Consistent labeling of tracked objects in multiple cameras with overlapping fields of view," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, pp. 1355 - 1360, Oct. 2003.
- [121] S. Khan, O. Javed, Z. Rasheed, and M. Shah, "Human Tracking in Multiple Cameras," in *Proceedings Eighth IEEE International Conference on Computer Vision (ICCV)*, 2001, pp. 331-336.
- [122] S. Khan, O. Javed, and M. Shah, "Tracking in Uncalibrated Cameras with Overlapping Field of View," in *Proc. IEEE Int'l Workshop Performance Evaluation of Tracking and Surveillance*, 2001, pp. 84-91.
- [123] M. J. Hannah, "SRI's Baseline Stereo System," in *Proc. of DARPA Image Understanding Workshop*, 1985, pp. 149-155.
- [124] M. J. Hannah, "A System for Digital Stereo Image Matching," *Photogrammetric Engineering and Remote Sensing*, pp. 1765-1770, 1989.
- [125] I. R. L. France, "A Parallel Stereo Algorithm that Produces Dense Depth Maps and Preserves Image Features," 1991.



- [126] G. P. Otto and T. K. W. Chau, "A Region Growing Algorithm for Matching of Terrain Images," *Proc. 4th Alvey Vision Conference*, pp. 123-128, 1988.
- [127] R.A.Lane and N.A.Thacker, "Tutorial: Overview of Stereo Matching Research," Imaging Science and Biomedical Engineering Division, Medical School, University of Manchester 1998.
- [128] S. T. Barnard and W. B. Thompson, "Disparity Analysis of Images," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*. vol. 4, 1980, pp. 333-340.
- [129] S. B. Pollard, "Identifying Correspondences in Binocular Stereo," in *AIVRU: University of Sheffield*, 1985.
- [130] Y. Ohta and T. Kanade, "Stereo by Intra- and Inter-Scanline Search Using Dynamic Programming," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*. vol. 2, 1985, pp. 139-154.
- [131] H. H. Baker, "Depth from Edge and Intensity Based Stereo," Stanford AI Laboratory, Stanford University 1982.
- [132] P. F. McLauchlan, "Recovery of Textured Surfaces using Stereo Vision," in *AIVRU: University of Sheffield*, 1990.
- [133] G. Stein, R. Romano, and L. Lee, "Monitoring Activities from Multiple Video Streams: Establishing a Common Coordinate Frame," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, pp. 758-767, Aug 2000.
- [134] M. Singh, A. Basu, and M. Mandal, "Event Dynamics Based Temporal Registration," *IEEE Transactions on Multimedia*, vol. 9, pp. 1004-1015, Aug. 2007.
- [135] S. S. Ravela, T. J. Schnackertz, J. R. A. Grupen, R. S. Weiss, E. M. Riseman, and A. R. Hanson, "Temporal registration for assembly," in *Proceedings Workshop on Vision for Robots*, Pittsburgh, August 6, 1995, pp. 64-69.
- [136] X. Guo, W. Hsu, M. L. Lee, and T. Y. Wong, "A Tree Matching Approach for the Temporal Registration of Retinal Images," in *18th IEEE International Conference on Tools with Artificial Intelligence (ICTAI'06)*, 2006, pp. 632-642.
- [137] M. D. Pritt, "Automated subpixel image registration of remotely sensed imagery," *IBM Journal of Research and Development*, vol. 38, pp. 157-166, 1994.
- [138] A.Averbuch and Y. Keller, "FFT Based Image Registration," in *IEEE International Conference on Acoustics, Speech and Signal Processing*. vol. 4 Orlando, USA, May 2002, pp. 3608-3601.
- [139] A. Venot, J. F. Lebruchec, and J. C. Roucayrol, "A new class of similarity measure for robust image registration," in *Proceedings of Computer Vision, Graphics, and Image Processing*, Nov: 1984, pp. 176-184.
- [140] J. V. Miller and C. V. Stewart, "MUSE: Robust surface fitting using unbiased scale estimates," in *IEEE Conference on Computer Vision and Pattern Recognition* San Francisco, CA, 1996 pp. 300-306.
- [141] E.-Y. Kang, I. Cohen, and G. Medioni, "Robust affine motion estimation in joint image space using tensor voting," in *16th Intern. Conf. on Computer Vision and Pattern Recog.* vol. 4 Quebec City, Canada, Aug. 2002, pp. 256-259.
- [142] B. Tordoff and D. Murray, "Guided sampling and consensus for motion estimation," in *7th European Conference on Computer Vision*. vol. 1 Copenhagen, Denmark, May 2002, pp. 82-96.

- [143] P. Pritchett and A. Zisserman, "Wide baseline stereo matching," in *6th International Conference on Computer Vision* Bombay, India, January 1998, pp. 754-760.
- [144] P. H. S. Torr and A. Zisserman, "MLESAC: A new robust estimator with application to estimating image geometry," *Computer Vision and Image Understanding*, vol. 78, pp. 138-156, 2000.
- [145] R. M. Haralick and H. Joo, "2D-3D pose estimation," in *Proceedings of the 9th International Conference on Pattern Recognition*, Rome, Italy, Nov. 1988, pp. 385-391.
- [146] F. R. Hampel, P. J. Rousseeuw, E. Ronchetti, and W. A. Stahel, "Robust Statistics: The Approach Based on Influence Functions," N. Y. John Wiley, Ed., 1986.
- [147] G. H. Golub and C. Reinsch, "Singular value decomposition and least squares solutions," *Number. Math.*, vol. 14, pp. 403-420, 1970.
- [148] M. Irani and P. Anandan, "Robust multi-sensor image alignment," in *4th International Conference on Computer Vision*. vol. 959-966 Bombay, India, January 1998.
- [149] S. Ayer, P. Schroeter, and J. Bigum, "Segmentation of moving objects by robust motion parameter estimation over multiple frames," in *Third European Conference on Computer Vision*. vol. 2 Stockholm, Sweden, May 1994, pp. 316-327.
- [150] D. Comaniciu and P. Meer, "Robust Analysis of Image Spaces: Color Image Segmentation," in *Proc. IEEE Conf on Comp. Vis. And Pattern Recognition*, Puerto-Rico, 1997, pp. 750-755.
- [151] B. Georgescu, I. Shimshoni, and P. Meer, "Mean shift based clustering in high dimensions: A texture classification example," in *Proc. 9th Intl. Conf. on Computer Vision*, Nice, France.
- [152] P. Meer, S. Wang, and H. Wechsler, "Edge detection by associative mapping," *Pattern Recognition*, vol. 22, pp. 491-503, 1989.
- [153] P. Meer and B. Georgescu, "Edge detection with embedded confidence," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 23, pp. 1351-1365, 2001.
- [154] R. M. Haralick and L. G. Shapiro, "Computer and Robot Vision," *Addison-Wesley*, 1992.
- [155] P. H. S. Torr and A. Zisserman, "Robust computation and parametrization of multiple view relations," *ICCV 1998*, pp. 727-732, 1998.
- [156] P. Beardsley, P. H. S. Torr, and A. Zisserman, "3D model acquisition from extended image sequences," in *Proc. 4th European Conference on Computer Vision*, B. B. a. C. R., Ed. Cambridge: Springer-Verlag, 1996, pp. 683-695.
- [157] R. I. Hartley, "Euclidean reconstruction from uncalibrated views," in *Proc. 2nd European US Workshop on Invariance*, Azores, 1993, pp. 187-202.
- [158] C. Tomasi and T. Kanade, "Shape and motion from image streams under orthography: A factorisation approach," *International Journal of Computer Vision*, vol. 9, pp. 137-154, 1992.
- [159] Z. Zhang and O. Faugeras, "3D Dynamic Scene Analysis," *SpringerVerlag*, 1992.
- [160] P. H. S. Torr and D. W. Murray, "Stochastic motion clustering," in *Proc. 3rd European Conference on Computer Vision and Pattern Recognition*, Stockholm, 1994, pp. 328-338.
- [161] P. H. S. Torr, A. Zisserman, and D. W. Murray, "Motion clustering using the trilinear constraint over three views," in *Workshop on Geometrical Modelling*

- and Invariants for COMPUTER VISION AND IMAGE UNDERSTANDING*, R. M. a. C.Wu, Ed.: Springer-Verlag, 1995, pp. 118-125.
- [162] P. H. S. Torr, "An assessment of information criteria for motion model selection," *CVPR97*, pp. 47-53, 1997.
  - [163] B. J. Tordoff and D. W. Murray, "Guided-MLESAC: Faster Image Transform Estimation by Using Matching Priors," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, pp. 1523-1535, Oct. 2005.
  - [164] A. Konushin, V. Gaganov, and V. Veznevets, "AMLESAC: A New Maximum Likelihood Robust Estimator," in *Graphicon*, Novosibirsk Akademgorodok, Russia, 2005.
  - [165] D. R. Myatt, P. H. S. Torr, S. J. Nasuto, J. M. Bishop, and R. Craddock, "NAPSAC: High Noise, High Dimensional Robust Estimation - it's in the Bag," *Proc. British Machine Vision Conf.*, pp. 458-467, 2002.
  - [166] Y. Cheng, "Mean shift, Mode Seeking, and Clustering," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 17, pp. 790-799, Aug. 1995.
  - [167] D. Comaniciu and P. Meer, "Mean shift analysis and applications," in *7th Proceedings IEEE International Conference on Computer Vision (ICCV'99)*, Kerkyra, Greece, 1999.
  - [168] D. Comaniciu, V. Ramesh, and P. Meer, "The Variable Bandwidth Mean Shift and Data-Driven Scale Selection," in *IEEE Int. Conf. Computer Vision (ICCV'01)*, vol. 1 Vancouver, Canada, 2001, pp. 438-445.
  - [169] D. Comaniciu, "An Algorithm for Data-Driven Bandwidth Selection," *IEEE Trans. Pattern Analysis Machine Intell.*, vol. 25, pp. 281-288, 2003
  - [170] D. Comaniciu, V. Ramesh, and P. Meer, "Kernel-based object tracking," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, pp. 564 - 577, 2003.
  - [171] D. Mintz, P. Meer, and A. Rosenfeld, "Consensus by decomposition: A paradigm for fast high breakdown point robust estimation," in *Proceedings 1991 DARPA Image Understanding Workshop*, La Jolla, CA, 1992, pp. 345-362.
  - [172] H. Chen, P. Meer, and D. E. Tyler, "Robust regression for data with multiple structures," in *IEEE Conference on Computer Vision and Pattern Recognition*, vol. 1 Kauai, HI, 2001 pp. 1069-1075.
  - [173] C. V. Stewart, "Bias in robust estimation caused by discontinuities and multiple structures," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 19, pp. 818-833, 1997.
  - [174] I. Akyildiz, S. Weilian, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *IEEE Communications Magazine*, vol. 40/8, pp. 102-114, 2002.
  - [175] A. H. Sayed, A. Tarighat, and N. Khajehnouri, "Network-based wireless location: challenges faced in developing techniques for accurate wireless location information," *IEEE Signal Processing Magazine*, vol. 22, pp. 24-40, 2005.
  - [176] C. Buschmann, H. Hellbruck, S. Fischer, A. Kroller, and S. P. Fekete, "Radio Propagation-Aware Distance Estimation Based on Neighborhood Comparison," in *Proc. of the 4th European Conference Wireless Sensor Networks*, Delft, Netherlands, 2007, pp. 325-340.
  - [177] G. Mao, B. Fidan, and B. Anderson, "Wireless sensor network localization techniques," *ACM Computer Networks Journal*, vol. 51, pp. 2529-2553, July 2007.

- [178] D. Pescaru, V. Gui, and C. Toma, "Analysis of Post-Deployment Sensing Coverage for Video Wireless Sensor Networks," in *Proceedings of the 6th RoEduNet International Conference*, Craiova, Romania, 2007, pp. 109-112.
- [179] Y. Weng and A. Daboli, "Smart Sensor Architecture Customized for Image Processing Applications," in *Proc. 10th IEEE Real-Time and Embedded Technology and Applications*, Toronto, 2004.
- [180] H. M. Chen, S. Lee, R. M. Rao, M. A. Slaman, and P. K. Varshney, "Imaging for Concealed Weapon Detection," *IEEE Signal Processing Magazine*, vol. 22, pp. 52-61, 2005.
- [181] D. Pescaru, **D. Fuiorea**, V. Gui, C. I. Toma, G. M. Muntean, and A. Daboli, "Image-based Node Localization Algorithm for Wireless Video Sensor Networks," in *The Sixth Information Technology and Telecommunication Conference, IT&T'06, "Digital Convergence in a Knowledge Society"* Carlow, Ireland, Oct. 2006, pp. 139-147.
- [182] A. Savvides, C. C. Han, and M. B. Srivastava, "Dynamic Fine-Grained Localization in Ad-Hoc Networks of Sensors," in *Proceedings of the 7th International Conference on Mobile Computing and Networking (MobiCom)*, Rome, Italy,, 2001.
- [183] T. He, C. Huang, B. Blum, J. A. Stankovic, and T. Abdelzaher, "Range-free localization schemes for large scale sensor networks," in *Proceedings of the 9th International Conference on Mobile Computing and Networking (MobiCom)*, San Diego, CA, 2003.
- [184] E. Elnahrawy, X. Li, and R. P. Martin, "The Limits of Localization Using Signal Strength: A Comparative Study," in *Proceedings of The First IEEE International Conference on Sensor and Ad hoc Communications and Networks (SECON 2004)*, Santa Clara, CA, USA, 2004.
- [185] C.-F. Huang and Y.-C. Tseng, "The Coverage Problem in a Wireless Sensor Network," in *Proceedings of the 2nd ACM International Conference on Wireless Sensor Networks and Applications, WSNA '03*, ACM Press, 2003, pp. 115-121.
- [186] M. Kaveh and A. Bassias, "Threshold extension based on a new paradigm for MUSIC-type estimation," *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing*, vol. 5, pp. 2535-2538, 1990.
- [187] N. Tayem and H. M. Kwon, "Conjugate ESPRIT (C-SPRIT)," *IEEE Transactions on Antennas and Propagation*, vol. 52, pp. 2618-2624, 2004.
- [188] D. McCrady, L. Doyle, H. Forstrom, T. Dempsey, and M. Martorana, "Mobile ranging using low-accuracy clocks," *IEEE Transactions on Microwave Theory and Techniques*, vol. 48, pp. 951-958, 2000.
- [189] K. Romer, "The lighthouse location system for smart dust," *Proceedings of ACM Conference on Mobile Systems, Applications, and Services*, pp. 15-30, 2003.
- [190] P. Krishnan, A. Krishnakumar, W.-H. Ju, C. Mallows, and S. Gamt, "A system for LEASE: location estimation assisted by stationary emitters for indoor RF wireless networks," *Proceedings of IEEE INFOCOM'04*, vol. 2, pp. 1001-1011, 2004.
- [191] **D. Fuiorea**, D. Pescaru, V. Gui, and C. I. Toma, "Using Registration Algorithms for Wireless Sensor Network Node Localization," in *Proceedings of the 4th International Symposium on Applied Computational Intelligence and Informatics*, Mai 2007, pp. 209-214.

- [192] G. Borgefors, "Hierarchical Chamfer Matching: A parametric edge matching algorithm," *IEEE Transaction on Transactions on Pattern Analysis and Machine Intelligence*, vol. 10, pp. 849-865, 1988.
- [193] **D. Fuiorea**, "A New Point Matching Method for Registration-Based Sensor Localization," in *Sesiunea de comunicări științifice "Doctor ETc 2007"* Timișoara, Sept. 2007, pp. 29-33.
- [194] **D. Fuiorea**, "A New Point Matching Method for Image Registration Using Pixel Color Information," in *Buletinul Științific al Universității "Politehnica" Timișoara, International Symposium on Electronics and Telecommunications* vol. 52(66) Timișoara, Sept. 2007, pp. 77-80.
- [195] **D. Fuiorea**, V. Gui, F. Alexa, and C. Toma, "A New Point Matching Method for Image Registration," in *Proceedings of the 6th WSEAS Int. Conf. on COMPUTATIONAL INTELLIGENCE, MAN-MACHINE SYSTEMS and CYBERNETICS (CIMMACS '07)*, Tenerife, Canary Islands, Spain, Dec. 2007, pp. 135-139.
- [196] **D. Fuiorea**, V. Gui, D. Pescaru, P. Paraschiv, I. Codruta, C. D, and C. Volosencu, "Sensor node localization using SIFT algorithm," in *Proceedings of the 9th WSEAS Int. Conf. on AUTOMATION and INFORMATION (ICAI'08)*, București, Romania, Iunie 2008, pp. 436-441.
- [197] **D. Fuiorea**, V. Gui, D. Pescaru, P. Paraschiv, I. Codruta, C. D, and C. Volosencu, "Video-based Wireless Sensor Networks Localization Technique Based on Image Registration and SIFT Algorithm," *WSEAS TRANSACTIONS on COMPUTERS*, vol. 7, pp. 990-999, Iulie 2008.
- [198] D. Pescaru, B. Ciubotar, D. Chiciudean, and A. Doboli, "Experimenting Motion Detection Algorithms for Sensor Network Video Surveillance Applications," *Scientific Bulletin of "Politehnica" University of Timisoara, Romania, Transactions on Automatic Control and Computer Science*, vol. 50(64), pp. 39-44, 2005.
- [199] R. Hartley and A. Zisserman, "Multiple View Geometry in Computer Vision," *Cambridge University Press*, 2000.
- [200] V. Gui, D. Lacrămă, and D. Pescaru, "Prelucrarea Imaginilor," T. Politehnica, Ed., 1999.
- [201] P. J. Rousseeuw and A. M. Leroy, "Robust Regression and Outlier Detection," N. Y. Wiley, Ed., 1987.
- [202] P. J. Rousseeuw, "Least median of squares regression," *J. American Statistical Assoc.*, vol. 79, pp. 871-880, 1984.
- [203] P. J. Huber, "Robust Statistics," N. John Wiley & Sons, Ed., 1981.
- [204] D. R. Myatt, "Robust estimation in high noise and highly dimensional data sets with applications to machine vision," Department of Cybernetics Oct 2002.
- [205] P. Hall, T. C. Hui, and J. S. Marron, "Improved variable window kernel estimates of probability densities," *Annals of Statistics*, vol. 23, pp. 1-10, 1995.
- [206] P. Meer, D. Mintz, D. Kim, and A. Rosenfeld, "Robust regression methods in computer vision: A review," *International Journal of Computer Vision*, vol. 6, pp. 59-70, 1991.
- [207] C. V. Stewart, "MINPRAN: A new robust operator for computer vision," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17(10), pp. 925-938, 1995.
- [208] I. Myung, "Tutorial on maximum likelihood estimation," *Journal of Mathematical Psychology*, vol. 47, pp. 90-100, 2003.

- [209] P. H. S. Torr and D. W. Murray, "Outlier detection and motion segmentation," in *Proc. Sensor Fusion VI*, Boston, 1993, pp. 432-443.
- [210] P. H. S. Torr and D. W. Murray, "The development and comparison of robust methods for estimating the fundamental matrix," *Int. Journal of Computer Vision*, vol. 24, pp. 271-300, Sept. 1997.
- [211] P. H. S. Torr and A. Zisserman, "Robust parameterization and computation of the trifocal tensor," *Image and Vision Computing*, vol. 15, pp. 591-605, August 1997.
- [212] D. Nister, "Preemptive RANSAC for live structure and motion estimation," in *Proc. 9th Intl. Conf. on Computer Vision*, Nice, France, 2003, pp. 199-206.
- [213] A. Konouchine, V. Gaganov, and V. Vezhnevets, "AMLESAC: A New Maximum Likelihood Robust Estimator," *Proc. Graphicon'05*, pp. 93-100, 2005.
- [214] P. H. S. Torr, P. A. Beardsley, and D. W. Murray, "Robust vision," *Proc. 5th British Machine Vision Conference*, pp. 145-155, 1994.
- [215] P. J. Huber, "Projection Pursuit," *Annals of Statistics*, vol. 13, pp. 433-475, 1985.
- [216] S. Borman, "The expectation and maximization algorithm, A short tutorial," 2005.
- [217] K. Fukunaga and L. D. Hostetler, "The estimation of the gradient of a density function, with applications in pattern recognition," *IEEE Trans. Information Theory*, vol. 21, pp. 32-40, 1975.
- [218] C. V. Raykar and R. Duraiswami, "Fast optimal bandwidth selection for kernel density estimation," in *Proceedings of the sixth SIAM International Conference on Data Mining*, Bethesda, 2006, pp. 524-528.
- [219] B. A. Turlach, "Bandwidth selection in kernel density estimation: A review," C.O.R.E and Institut de Statistique, Universite Catholique de Louvain, Louvain-la-Neuve 1993.
- [220] **D. Fuiorea**, "Feature Based 2D Image Registration Using Mean Shift Parameter Estimation " *Buletinul Științific al Universității "Politehnica" Timișoara*, vol. 51(65), pp. 77-80, Sept. 2006.
- [221] S. Umeyama, "Least-squares estimation of transformation parameters between two point patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, pp. 376-380, 1991.
- [222] H. G. Barrow, J. M. Tanenbaum, R. C. Balles, and H. C. Wolf, "Parametric correspondence and chamfer matching: Two new techniques for image matching," in *Proceedings Fifth Internat. Joint Conf. on Artificial Intelligence*, 1977, pp. 659-663.
- [223] **D. Fuiorea**, V. Gui, D. Pescaru, and C. Toma, "Comparative study on RANSAC and Mean shift algorithm," in *Simpoziomul Internațional de Electronică și Telecomunicații, Ediția 8*. vol. 53(67) Timisoara: Scientific Bulletin of the "Politehnica" University of Timisoara, Sept. 2008, pp. 80-85.
- [224] **D. Fuiorea**, V. Gui, D. Pescaru, and C. Toma, "Kernel based image registration versus MLESAC: a comparative study " in *Proceedings of the 5th International Symposium on Applied Computational Intelligence and Informatics* Timisoara, 2009.
- [225] **D. Fuiorea**, "Tehnici de punere in corespondenta a imaginilor (I)," in *Referatul de doctorat nr.1* Universitatea "Politehnica" Timișoara, Martie 2006.



- [226] **D. Fuiorea**, "Tehnici robuste de punere in corespondenta a imaginilor (II)," in *Referatul de doctorat nr. 2* Universitatea "Politehnica" Timișoara, Iulie 2006.
- [227] **D. Fuiorea**, "Folosirea algoritmilor de corespondență pentru localizarea nodurilor dintr-o rețea de senzori fără fir," in *Referatul de doctorat nr. 3* Universitatea "Politehnica" Timișoara, Noiembrie 2007.
- [228] V. Gui, F. Alexa, C. Căleanu, and D. Fuiorea, "Motion Segmentation and Analysis in Video Sequences," *WSEAS Transaction on Circuits and Systems*, vol. 6, pp. 142-148, Ianuarie 2007.
- [229] V. Gui, F. Alexa, C. Căleanu, and D. Fuiorea, "Robust 2D Moving Object Segmentation and Tracking in Video Sequences," in *Proceedings of the 5th WSEAS Int. Conf. on System Science and Simulation in Engineering*, Tenerife, Canary Islands, Spania, 2006, pp. 232-236.
- [230] V. Gui, D. Fuiorea, C. I. Toma, and F. Alexa, "Video Object Segmentation and Tracking using Kernel Density Estimation," in *Proceedings of the 12th NETTIES International Conference, NETTIES 2006*, Timisoara, Romania, 2006, pp. 237- 240.
- [231] D. Fuiorea, "Grant PN II, CNCSIS - TD nr. 546/2007, Tehnici robuste de punere în corespondență a imaginilor," 2007-2008.