



**UNIVERSITATEA
TEHNICĂ
DIN CLUJ-NAPOCA**

FACULTATEA DE AUTOMATICĂ ȘI CALCULATOARE
DEPARTAMENTUL DE CALCULATOARE

Ing. Silviu Bota

**TEZĂ DE DOCTORAT
DETECTIA ȘI URMĂRIREA MIȘCĂRII DIN IMAGINI
TRIDIMENSIONALE**

Conducător științific
Prof. Dr. Ing. Sergiu NEDEVSCI

Cuprins

1	Introducere	5
1.1	Motivație	5
1.2	Obiectivele tezei	6
1.3	Structura tezei	7
1.4	Mulțumiri	7
I	Studiu bibliografic	8
2	Studiul fluxului optic și al mișcării camerelor	8
2.1	Mișcarea în imagini	8
2.2	Fundamentele teoretice ale fluxului optic	9
2.3	Lucas–Kanade	9
2.4	Horn–Shunk	9
2.5	Flux optic piramidal	10
2.6	Selecția trăsăturilor	10
2.7	Flux optic 3D	10
2.8	Potrivirea blocurilor	11
2.9	Flux optic bazat pe corelarea fazei	11
2.10	Estimarea mișcării camerelor	11
2.11	Concluzii	11
3	Studiul algoritmilor de urmărire în medii de trafic	12
3.1	Camere montate pe autovehicule	12
3.2	Camere fixe	13
3.3	Filtrul Kalman	13
3.4	Filtrul de particule	13
3.5	Concluzii	14
4	Studiul algoritmilor de recunoaștere a obiectelor din trafic	15
4.1	Necesitatea recunoașterii obiectelor din trafic	15
4.2	Papageorgiou et al.	16
4.3	Gavrila și Grupul Daimler	16
4.4	Universitatea din Parma (VISLAB)	16
4.5	Ghandi et al.	17
4.6	Alte abordări	17
4.7	Concluzii	17

II	Contribuții teoretice și aplicative	19
5	Fluxul optic și estimarea mișcării camerelor	19
5.1	Algoritm în timp real pentru calcularea fluxului optic dens	19
5.2	Segmentarea imaginilor bazată pe mișcare, intensitate și profunzime	20
5.3	Algoritm original de selecție a trăsăturilor pentru fluxul optic	20
5.4	Modelul camerei și al mișcării	21
5.5	Algoritm original pentru estimarea mișcării camerei 2D și detecția unei noi secvențe .	21
5.6	Algoritm original pentru estimarea mișcării camerei 3D calibrate	22
5.7	Algoritm original pentru estimarea mișcării camerelor 3D stereo	22
5.8	Concluzii	23
6	Urmărirea obiectelor în medii de trafic urban	26
6.1	Agregarea datelor zgomotoase din mai multe surse	26
6.2	Arhitectura sistemului de urmărire	27
6.3	Sursele erorilor care trebuie filtrate de sistemul de urmărire	27
6.4	Model original al obiectelor	28
6.4.1	Modelul static al obiectelor	28
6.4.2	Modelul dinamic al obiectelor	30
6.5	Predicția	31
6.6	Algoritmul original de asociere a obiectelor	31
6.7	Filtrarea Kalman	32
6.8	Rezultate experimentale	32
6.9	Concluzii	33
7	Recunoașterea obiectelor din trafic	34
7.1	Introducere	34
7.2	Trăsături originale de mișcare pentru recunoașterea pietonilor	35
7.2.1	Semnătura de mișcare	35
7.2.2	Periodicitatea semnăturii de mișcare	36
7.2.3	Rezultate experimentale	37
7.3	Trăsături simple	37
7.4	Clasificatorul Bayesian	37
7.4.1	Rezultate experimentale	39
7.5	Arhitectura originală de clasificare	39
7.6	Baza de date pentru antrenarea și evaluarea clasificării	39
7.7	Concluzii	42
III	Concluzii	44
8	Concluzii	44
A	Articole publicate	58
A.1	În jurnale ISI cu factor de impact	58
A.2	În volumele unor conferințe ISI Proceedings, IEEE Xplore, SpringerLink	58
A.3	Capitole de carte	60
A.4	În volumele altor conferințe	60

Capitolul 1

Introducere

1.1 Motivație

Mișcarea reprezintă o sursă de informații deosebit de importantă și valoroasă în secvențele de imagini. Putem spune că o secvență de imagini conține mult mai multă informație decât informația care poate să fie extrasă din cadrele sale individuale. Lumea în care trăim este tridimensională. De aceea este extrem de important ca sistemele de viziune artificiale să perceapă lumea în trei dimensiuni. Din păcate, senzorii obișnuiți folosiți în camerele video sunt doar bidimensionali. Totuși, combinând două sau mai multe camere într-o configurație stereoscopică, putem capta parțial anumite aspecte tridimensionale ale lumii. Combinarea mișcării bidimensionale extrasă din secvențele de imagini cu informația tridimensională extrasă folosind stereoviziunea ne permite să obținem o reprezentare în 6 dimensiuni a lumii, captând nu doar aspectele spațiale cât și aspectele temporale ale sale.

O țintă importantă a algoritmilor de viziune computerizată este înțelegerea scenei. Un prim pas pentru înțelegerea scenei este segmentarea acesteia. Segmentarea scenei se poate face relativ ușor atunci când camerele sunt fixe. Totuși, în afară de cazul aplicațiilor de supraveghere video, camerele fixe nu sunt de mare utilitate. De aceea, determinarea mișcării camerei constituie un pas preliminar al multor algoritmi de viziune artificială.

Vom enumera doar două aplicații ale detecției mișcării camerei. Prima aplicație se referă la domeniul extragerii bazate pe context a imaginilor, mai precis la adnotarea automată a filmelor în format digital. Unele din cerințele adnotării automate se referă la detecția mișcărilor specifice ale camerei video, cum ar fi rotațiile în plan orizontal și vertical, zoom-in, zoom-out, precum și detecția începutului unei noi secvențe. Nu este în general necesară o estimare cantitativă a acestor mișcări, ci doar de una calitativă. Deoarece majoritatea filmelor în format digital nu conțin informații 3D, soluționarea estimării acestor mișcări trebuie să se limiteze la o abordare bidimensională.

O altă aplicație a detecției mișcării camerei este în domeniul sistemelor de asistență a conducerii rutiere. Aceste sisteme au nevoie de obicei, ca date de intrare, de informații asupra mișcării autovehiculului propriu. De obicei există posibilitatea obținerii acestor date de la senzorii proprii ai autovehiculului. Totuși, acești senzori nu sunt întotdeauna accesibili, nu sunt montați pe toate autovehiculele, protocoalele de comunicație cu ei diferă de la producător la producător și ei necesită întotdeauna un proces de calibrare, al cărui protocol nu este disponibil. Totodată, în cazul producerii unor derapaje, senzorii de odometrie nu furnizează mișcarea corectă, senzorii de ESP deducând mișcarea din unghiul volanului, accelerația unghiulară și odometrie. Folosirea unei camere mono sau a unei perechi de camere în configurație stereo ca senzori pasivi prezintă avantaje relativ la problemele menționate mai sus. Astfel, camerele își pot găsi aplicații, în afară de asistența conducerii rutiere

propriu zise, și în sistemele de prevenire a blocării roților la frânare (ABS), de control automat al demarajului (ASR), controlul electronic al stabilității (ESP) etc.

Stereoviziunea artificială oferă cantități mari de informații, fapt avantajos pentru extragerea informațiilor semantice din scenă. Din păcate, aceste informații sunt afectate de zgomot, cauzat în special de erori inerente procesului de stereo reconstrucție. Informația temporală poate fi utilizată pentru a minimiza aceste erori. Totuși, este dificilă utilizarea directă a unei reprezentări la nivel de pixel și vectori de mișcare a lumii. O reprezentare bazată pe obiecte ierarhice a scenei este mai utilă din punct de vedere computațional. Obiectele trebuie să își mențină identitatea în timp. Din această cauză este necesar un proces de urmărire a obiectelor. Cunoașterea identității obiectelor din scenă permite integrarea în timp a diverselor măsurători, cu scopul de a obține o descriere de acuratețe mai mare acestora.

Sistemele tradiționale de asistență a conducerii rutiere, cum ar fi tempomatele adaptive sunt limitate la scenarii de autostradă. Scenariile de autostradă sunt relativ simple, singurele obiecte în mișcare fiind celelalte autovehicule, obiectele statice vizibile fiind doar parapetii laterali. Pentru aceste medii, senzorii activi foarte preciși dar care furnizează cantități limitate de informație, cum ar fi radarul, sunt suficienți pentru construirea unor sisteme de asistență a conducerii rutiere. Pe de altă parte, sistemele de asistență a conducerii pentru medii urbane, se confruntă cu problema mult mai dificilă de a percepe și înțelege scenarii mult mai complexe. În scenariile urbane își fac apariția alți participanți la trafic, cum ar fi pietonii, bicicliștii. Totodată apar și alte elemente fixe, cum ar fi mașinile parcate, stâlpi, indicatoare. Prezența intersecțiilor, a drumurilor cu curbura mare, lățimea variabilă a benzilor de circulație face ca abordările bazate pe radar ale sistemelor de asistență a conducerii să nu funcționeze întotdeauna corect. Totuși, aceste sisteme, atât orientate pe confort cât și pe siguranță sunt mult mai utile în medii urbane, unde solicitarea la care este supus șoferul este mult mai mare. Deoarece senzorii bazați pe stereoviziune furnizează o cantitate mare de informație, ei tind să devină o nouă și importantă categorie de senzori folosiți de către sistemele de siguranță și confort. Unele dintre problemele pe care este necesar să le rezolve sistemele de asistare a conducerii rutiere urbane sunt: detecția obstacolelor statice (în scopul evitării acestora) în cadrul funcțiilor de asistență longitudinală, laterală, frânare de urgență sau protecție activă (declanșarea airbag-urilor); detecția și urmărirea obiectelor în mișcare, pentru funcțiile de asistență în traficul de tip oprit – pornit (urmărirea autovehiculului din față), frânare de urgență, limitarea efectelor coliziunilor frontale. O funcție extrem de importantă în limitarea efectelor coliziunii o reprezintă detecția părții autovehiculului corespunzătoare habitaculului, pentru a limita impactul asupra persoanelor aflate în interior în cazul unei coliziuni. Recunoașterea și protecția pietonilor reprezintă o provocare importantă pentru sistemele de asistență a conducerii în mediul urban, deoarece pietonii sunt cei mai vulnerabili participanți la trafic. O alta provocare o reprezintă detecția și recunoașterea semnelor de circulație și a semafoarelor.

Din cele discutate mai sus, reiese faptul că detecția obiectelor, urmărirea și clasificarea lor sunt primii pași care trebuie făcuți în dezvoltarea unui sistem de asistență a conducerii rutiere în medii de trafic urbane.

1.2 Obiectivele tezei

Obiectivele acestei teze sunt detecția mișcării din imagini tridimensionale, urmărirea acestei mișcări și utilizarea ei pentru extragerea de informații despre scenă. Majoritatea algoritmilor actuali nu utilizează informația 3D la capacitate maximă. Chiar și acei algoritmi care utilizează informații 3D le utilizează doar în anumiți pași. Această teză investighează potențialul utilizării complete a

informației de mișcare tridimensionale.

Primul obiectiv al tezei este determinarea mișcării camerei. Determinarea acestei mișcări este utilă în multe domenii, cum ar fi adnotarea automată a secvențelor de imagini sau înlocuirea senzorilor de mișcare tradiționali ai autovehiculelor.

Al doilea obiectiv este dezvoltarea unui algoritm de urmărire a obiectelor prezente în scenarii dificile și complexe de trafic urban. Urmărirea acestor obiecte reprezintă o cerință majoră a sistemelor de asistență avansată a conducerii în medii urbane.

Al treilea obiectiv este reprezentat de folosirea mișcării, precum și a altor trăsături, în recunoașterea pietonilor și a altor clase de obiecte în medii de trafic urban. Recunoașterea obiectelor este la rândul ei o cerință esențială a sistemelor de asistență avansată a conducerii în medii de trafic urban și este doar parțial rezolvată de sistemele actuale.

1.3 Structura tezei

Teza este structurată în trei părți. Prima parte reprezintă un studiu bibliografic al următoarelor trei domenii:

- Extragerea fluxului optic și a mișcării camerei (descrisă în capitolul 2);
- Urmărirea obiectelor în mediul de trafic urban (descrisă în capitolul 3);
- Recunoaștere pietonilor și a altor obiecte din traficul urban (descrisă în capitolul 4);

A doua parte prezintă contribuții originale, teoretice și aplicative în următoarele trei domenii:

- Determinarea mișcării camerei pentru adnotarea automată a filmelor în format digital și calcularea mișcării vehiculului propriu (descrisă în capitolul 5);
- Urmărirea obiectelor din mediul de trafic urban (descrisă în capitolul 6);
- Recunoașterea pietonilor și a altor obiecte din trafic (descrisă în capitolul 7);

A treia parte prezintă concluziile precum și o trecere în revistă a contribuțiilor originale. Este prezentat impactul științific al tezei (reprezentat prin lucrările publicate și citările independente)

1.4 Mulțumiri

În primul rând doresc să îi mulțumesc îndrumătorului științific, domnul prof. dr. ing. Sergiu Nedevschi, pentru îndrumarea și suportul său pe durata realizării acestei teze. Sfaturile și opiniile sale obiective au avut un rol extrem de important pentru îmbunătățirea și extinderea contribuțiilor acestei teze. Doresc totodată să le mulțumesc colegilor mei din colectivul de procesarea imaginilor și recunoașterea formelor pentru ajutorul acordat. Doresc să le mulțumesc și partenerilor de la Volkswagen AG, Dr. Marian Andrzej Obojski, Dr. Thorsten Graf, Dr. Marc-Michael Meinecke, Dr. Thanh Binh To și Matthias König, care au finanțat munca mea și a colegilor prin contracte de cercetare. Nu în ultimul rând, doresc să le mulțumesc soției și părinților mei pentru încrederea, răbdarea și ajutorul acordat.

Partea I

Studiu bibliografic

Capitolul 2

Studiul fluxului optic și al mișcării camerelor

În acest capitol vom prezenta studiul celor mai relevanți algoritmi pentru calcul fluxului optic și pentru determinarea mișcării camerelor. Secțiunea 2.1 oferă două exemple de extragere a unor informații simple direct din mișcare, similare probabil cu procesele care au loc în sistemele vizuale biologice. Secțiunea 2.2 prezintă fundamentele teoretice ale fluxului optic, reprezentate de ecuația de iluminare constantă. Secțiunea 2.3 descrie algoritmi de flux optic de tip Lucas–Kanade care se bazează pe faptul că fluxul optic este constant pe o regiune mică. Secțiunea 2.4 descrie un algoritm alternativ, care se bazează pe netezimea fluxului optic, algoritmul Horn–Shunk. Secțiunea 2.5 descrie algoritmul lui Bouguet care extinde fluxul optic de la subpixel la deplasamente mai mari. Secțiunea 2.6 descrie algoritmi de detecție a trăsăturilor de tip colț folosiți în algoritmi de flux optic bazați pe trăsături. Secțiunea 2.7 descrie o extensie a fluxului optic prezentată de Harville care folosește și informații de adâncime. Secțiunea 2.8 descrie o altă abordare pentru calculul fluxului optic, bazată pe potrivirea blocurilor. Secțiunea 2.9 descrie o altă metodă bazată pe corelația fazelor. Secțiunea 2.10 prezintă studiul literaturii în domeniul determinării mișcării camerelor. În final, secțiunea 2.11 prezintă concluziile acestui studiu.

2.1 Mișcarea în imagini

Dimensiunea temporală a sistemelor de viziune este importantă din două motive: în primul rând, mișcarea aparentă a obiectelor în planul imaginilor aduce informații despre structura 3D a scenei și a mișcării din scenă. În al doilea rând, sistemele vizuale biologice sunt capabile să înțeleagă anumite scene folosind doar mișcarea. Vom prezenta două exemple în acest sens, preluate din [1]).

Primul exemplu constă în secvențe de imagini cu puncte generate aleator, mișcarea constând în deplasarea unei mici regiuni de la un cadru la altul. Dacă imaginile sunt prezentate într-o succesiune rapidă, sistemul vizual uman percepe această mișcare, chiar dacă punctele din imagine nu au niciun fel de altă structură. Exemplul al doilea se referă la calcul timpului de impact. Analiza mișcării permite calculul acestui timp doar din percepția mișcării, fără nicio informație suplimentară despre structura scenei, calibrarea camerei sau amplitudinea mișcării.

Problemele fundamentale ale detecție mișcării sunt:

1. Corespondența: care elemente din cadrul curent corespund căror elemente din cadrul următor al unor secvențe de imagini?
2. Reconstrucția: având un număr de corespondențe și eventual informații de calibrare a camerelor, ce informații putem extrage despre structura 3D a scenei?

Diferența principală dintre determinarea mișcării și stereoviziune este legată de faptul că în determinarea mișcării avem deplasamente relativ mici dar neconstrânse la un spațiu unidimensional al liniilor epipolare.

2.2 Fundamentele teoretice ale fluxului optic

Ecuția fundamentală pentru determinarea fluxului optic este constrângerea de iluminare constantă. Ea limitează modificările de la un cadru la altul la translații, și fără schimbarea iluminării:

$$I(x, y, t) = I(x + v_x(x, y), y + v_y(x, y), t + 1), \quad (2.1)$$

Printr-o aproximare Taylor de ordinul întâi obținem:

$$I(x, y, t + 1) = I'_x(x, y, t)v_x(x, y, t) + I'_y(x, y, t)v_y(x, y, t) + I_t(x, y, t) \quad (2.2)$$

Câmpul vectorial care este soluția acestei ecuații este fluxul optic. Deoarece avem o singură ecuație și două necunoscute, trebuie să impunem anumite constrângeri suplimentare. Constrângerile impuse generează două tipuri principale de algoritmi, de tip Horn–Shunk și Lucas–Kanade, descrise în secțiunile următoare.

2.3 Lucas–Kanade

Metodele de tip Lucas–Kanade generează mai multe ecuații, considerând ca fluxul optic este constant pe o fereastră în jurul punctului considerat, și rezolvă aceste ecuații prin metoda celor mai mici pătrate:

$$\begin{cases} I'_x v_x + I'_y v_y = -I'_t \\ I''_x v_x + I''_y v_y = -I''_t \\ \dots \\ I'^n_x v_x + I'^n_y v_y = -I'^n_t \end{cases} \quad (2.3)$$

Stabilitatea rezolvării acestor ecuații este dată de matricea $A^T A$ (unde A reprezintă matricea sistemului de mai sus. Dacă matricea nu are cele două valori proprii suficient de mari, nu vom putea rezolva sistemul. Condiția aceasta este echivalentă cu existența unor muchii pe două direcții independente în fereastra considerată.

2.4 Horn–Shunk

O altă metodă diferențială de calculare a fluxului optic este metoda Horn–Shunk [2]. Aceasta impune o constrângere de netezime a câmpului vectorial de flux optic:

$$\vec{v} = \arg \min \left\{ \iint_w \|\nabla \vec{v}\| dx dy \right\} \quad (2.4)$$

Împreună cu ecuația de iluminare constantă,

$$\nabla I \cdot \vec{v} = -I'_t \quad (2.5)$$

ele formează o ecuație de minimizare funcțională. Dacă discretizăm aceste ecuații, atunci ele formează un sistem Euler-Lagrange (având o matrice mare, rară). În [3] se prezintă un studiu al convergenței rezolvării acestui sistem, arătând că atât metodele de tip Gauss-Seidel cât și Jacobi sunt convergente și permit rezolvarea acestui sistem.

2.5 Flux optic piramidal

Din cauza aproximării de ordinul întâi metodele diferențiale prezentate mai sus nu sunt aplicabile unor deplasamente mai mari de un pixel. Neputând obține o rată de achiziție suficient de mare, este imposibil ca toate deplasamentele să fie la nivel de subpixel, ceea ce face ca acestea să nu poată fi aplicate direct.

Metoda introdusă de Bouguet [4] încearcă să rezolve aceste limitări, folosind o piramidă de imagini. Mișcările la subpixel sunt determinate pe o imagine mult micșorată, ceea ce asigură deplasamente mici. Apoi, se trece la nivelurile inferioare ale piramidei, eliminând la fiecare pas imaginea estimată anterior. Algoritmul furnizează o acuratețe și o precizie crescută.

2.6 Selecția trăsăturilor

O problemă importantă legată de algoritmi de flux optic o reprezintă găsirea acelor puncte în jurul cărora există suficiente informații în imagine pentru a permite calcularea fluxului optic. Aceste puncte sunt colțurile.

O metodă imediată este algoritmul prezentat de Harris [5] abordând problema direct prin prisma valorilor proprii ale matricei de structură a imaginii, și este o dezvoltare a unui algoritm anterior prezentat de Morevac [6] care definește un colț ca fiind un punct de autosimilaritate scăzută. Algoritmul Harris evită calcularea valorilor proprii, aproximând magnitudinea acestora prin calcule mai simple, algoritmul Shi-Tomasi [7] în schimb le calculează direct.

Alte abordări se bazează pe curbura locală [8][9], filtre LoG [10][11] DoG [12] sau DoH [10]. O abordare diferită este filtrul SUSAN [13]. [14] propune algoritmi genetici pentru generarea automată a extractorilor de trăsături.

2.7 Flux optic 3D

O abordare similară ecuației de iluminare constantă este ecuația de adâncime propusă de Harville [15] și aplicabilă imaginilor de profunzime:

$$Z(x, y, t) + V_z(x, y, t) = Z(x + v_x(x, y, t) + v_y \quad (2.6)$$

Ecuația se rezolvă similar cu ecuația de iluminare.

2.8 Potrivirea blocurilor

O metodă care nu este de tip diferențial folosită în calcularea fluxului optic în sisteme de compresie video este metoda bazată pe potrivirea blocurilor. Practic, imaginea se împarte într-un număr de blocuri de dimensiune fixă, care sunt apoi căutate exhaustiv (dar folosind anumite euristici) în cadrul anterior. Aceste metode se pretează la implementări hardware, au un grad mare de paralelism și nu conțin calcule foarte complexe.

2.9 Flux optic bazat pe corelarea fazei

O altă abordare pentru calcul fluxului optic o reprezintă metodele bazate pe corelarea fazei, [16] și [17].

Imaginile sunt transformate în spațiul Fourier, unde mișcările de translație se traduc prin schimbare fazei componentelor de frecvență. Prin folosirea corelației normalizate:

$$R = \frac{G_a G_b^*}{|G_a G_b|} \quad (2.7)$$

urmată de transformarea Fourier inversă, se poate obține translația în domeniul spațial

Deși interesante, metodele bazate pe corelarea fazei sunt costisitoare și sunt mai degrabă aplicabile potrivirii imaginilor.

2.10 Estimarea mișcării camerelor

Există metode de estimare a mișcării camerelor atât pentru camere mono cât și stereo. În [18], autorii prezintă o parametrizare a modelului perspectivă folosind 8 parametri. În [19] autorii propun o estimare cu 6 grade de libertate a mișcării unui robot folosind camere stereo. O abordare similară este prezentată în [20]. O altă abordare stereo este prezentată în [21]. O abordare bazată pe stereoviziune trinoculară este prezentată în [22].

2.11 Concluzii

Studiul acestor algoritmi precum și implementarea unora dintre ei ne-a permis să tragem următoarele concluzii:

- Cei mai utili algoritmi de extragere a colțurilor pentru calcularea fluxului optic sunt algoritmi Harris și Shi–Tomasi;
- Algoritmii de detecție a colțurilor se limitează doar la imagini 2D și nu folosesc informații 3D care pot fi furnizate de stereoviziune;
- Detecția colțurilor poate să consume mai mult timp decât calculul fluxului optic bazat pe trăsături, deoarece detecția colțurilor trebuie aplicată pe întreaga imagine, pe când calculul fluxului optic se aplică doar unui număr relativ mic de colțuri. Din această cauză, pentru a îmbunătăți timpul de procesare al algoritmilor de flux optic, este necesară optimizarea algoritmilor de detecție a colțurilor;

- Cel mai util algoritm existent pentru calculul fluxului optic bazat pe trăsături este algoritmul de tip piramidal Lucas–Kanade prezentat de Bouguet. Această are o precizie și o acuratețe sporită și este capabil să urmărească deplasamente mari pentru colțuri;
- Chiar dacă algoritmi de flux optic bazați pe potrivirea de blocuri sunt ușor de implementat și de optimizat, ei nu oferă suficientă acuratețe spațială.
- Algoritmi de flux optic bazați pe corelare de fază necesită o putere de calcul mare și sunt mai utili pentru potrivirea imaginilor decât pentru calculul fluxului optic.
- Algoritmi existenți pentru determinarea mișcării camerelor folosesc o parametrizare relativ generică, care deși poate să trateze foarte multe tipuri de mișcare este mai puțin utilă în scenarii cu mai puține grade de libertate.

Pe baza concluziilor de mai sus, am descris în capitolul 5 contribuțiile noastre legate de îmbunătățirea algoritmilor existenți precum și algoritmi originali pentru calculul fluxului optic dens, detecția colțurilor și estimarea mișcării camerelor.

Capitolul 3

Studiul algoritmilor de urmărire în medii de trafic

Acest capitol reprezintă un studiu al literaturii existente în domeniul urmăririi obiectelor în medii de trafic precum și a celor doi algoritmi de filtrare larg utilizați, filtrul Kalman și filtrul de particule. În secțiunea 3.1 vom descrie algoritmi de urmărire existenți care folosesc camere montate pe autovehicule pentru urmărire, în secțiunea 3.2 algoritmi care folosesc camere fixe, în secțiunea 3.3 vom prezenta filtrul Kalman iar în secțiunea 3.4 filtrul de particule.

3.1 Camere montate pe autovehicule

Există multe abordări pentru urmărirea obiectelor bazate pe senzori de viziune și eventual pe flux optic. Unele dintre ele utilizează modele pentru obiecte [23] [24], altele reprezintă obiectele mai flexibil, prin seturi de puncte [25]. Unele abordări folosesc filtre de particule [26] deoarece acestea permit descrierea unor densități de probabilitate complexe multi-modale. Modelul static al obiectelor poate fi mai relaxat până la abandonarea totală a conceptului de model [27]. Unele modele pot fi descrise mai degrabă ca hărți de ocupare Bayesiene [28][29][30][31].

Am studiat pe larg abordările propuse de [32], [33], [23], [24], [34],[35], [36], [37], [38], [39]

3.2 Camere fixe

Există mulți algoritmi pentru supravegherea și urmărirea traficului folosind camere fixe. Majoritatea se bazează pe asumția fundalul fix. Am studiat pe larg algoritmi prezentați în [40], [41], [42], [43], [44], [45].

3.3 Filtrul Kalman

În statistică filtrul Kalman [46] reprezintă o metodă matematică prezentată de Rudolf E. Kalman. Scopul ei este utilizarea măsurărilor achiziționate în timp, conținând zgomot pentru a produce valori filtrate mai apropiate de cele reale. Filtrul Kalman produce estimări ale valorilor reale și a incertitudinilor acestora prin agregare. Valorile agregate mai puțin zgomotoase au pondere mai mare. Ecuațiile filtrului Kalman sunt ecuația de tranziție:

$$x_k = F_k x_{k-1} + B_k u_k + w_k \quad (3.1)$$

Ecuația procesului de observație:

$$z_k = H_k x_k + v_k \quad (3.2)$$

Ecuația de predicție:

$$\hat{x}_{k|k-1} = F_k \hat{x}_{k-1|k-1} + B_k u_k \quad (3.3)$$

$$P_{k|k-1} = F_k P_{k-1|k-1} F_k^T + Q_k \quad (3.4)$$

Ecuațiile de filtrare:

$$\tilde{y}_k = z_k - H_k \hat{x}_{k|k-1} \quad (3.5)$$

$$S_k = H_k P_{k|k-1} H_k^T + R_k \quad (3.6)$$

$$K_k = P_{k|k-1} H_k^T S_k^{-1} \quad (3.7)$$

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k \tilde{y}_k \quad (3.8)$$

$$P_{k|k} = (I - K_k H_k) P_{k|k-1} \quad (3.9)$$

3.4 Filtrul de particule

În statistică, filtrele de particule, numite și metode Monte-Carlo secvențiale, reprezintă metode sofisticate bazate pe simulare. Ele sunt folosite pentru a estima modele Bayesiene în care variabilele latente sunt conectate printr-un lanț Markov. Metodele bazate particule presupun că variabilele de stare x_k și observațiile y_k pot fi modelate printr-un proces Markov de ordinul 1. Ele folosesc un set

de particule pentru aproximarea probabilității condiționate $p(x_k|y_0, y_1, \dots, y_k)$. Având P particule probabilitatea se aproximează prin:

$$\int f(x_k|y_0, y_1, \dots, y_k)dx_k \approx \frac{1}{P} \sum_{L=1}^P f(x_k^L) \quad (3.10)$$

3.5 Concluzii

Am studiat literatura existentă, prezentând în acest capitol cele mai relevante contribuții din domeniu precum și cei doi algoritmi de filtrare larg utilizați, filtrul Kalman și filtrul de particule. În secțiunea 3.1 am descris algoritmi de urmărire existenți care folosesc camere montate pe autovehicule pentru urmărire, în secțiunea 3.2 algoritmi care folosesc camere fixe, în secțiunea 3.3 am prezentat filtrul Kalman iar în secțiunea 3.4 filtrul de particule.

Studiul literaturii din domeniu a permis formularea următoarelor concluzii:

- Majoritatea abordărilor existente se concentrează pe urmărirea autovehiculelor individuale sau ușor separabile;
- Algoritmii de urmărire folosiți sunt influențați de algoritmi de detecție utilizați și de modelul obiectelor;
- Majoritatea algoritmilor care folosesc camere fixe se bazează în mare măsură doar pe asumția unui fundal fix, obținându-se rezultate destul de bune, dar făcând acești algoritmi inutili pentru sistemele de asistență a conducerii;
- Cea mai dificilă problemă în urmărirea obiectelor multiple o reprezintă asocierea, nu filtrarea efectivă.
- Filtrele de tip Kalman sunt rapide, precise, și de aceea sunt preferabile din punct de vedere computațional. Totuși ele nu pot modela procese cu zgomot de tip non-Gaussian.
- Filtrele Kalman lucrează recursiv, integrând la fiecare pas informațiile de stare prezise cu cele observate. Ele pot integra și observații multiple, dar, din cauza pierderilor de precizie, aceasta poate duce la acumularea erorilor
- Filtrele de particule sunt mai costisitoare din punct de vedere computațional, dar ele pot modela procese mai complexe. Totuși acestea sunt mai dificil de controlat din cauza comportamentului lor aleator.

Pornind de la aceste concluzii, vom descrie în capitolul 6 un algoritm de urmărire original care rezolvă problema urmăririi obiectelor multiple, ierarhice, în medii de trafic urban.

Capitolul 4

Studiul algoritmilor de recunoaștere a obiectelor din trafic

În acest capitol vom prezenta un studiu al literaturii din domeniul recunoașterii obiectelor din trafic, concentrându-ne în principal pe recunoașterea pietonilor. În secțiunea 4.1 am prezentat necesitatea clasificării obiectelor, în secțiunile 4.2, 4.3, 4.4, 4.6 am prezentat contribuțiile aduse de Papageorgiou, Gavrilă, Broggi și alți autori în acest domeniu.

4.1 Necesitatea recunoașterii obiectelor din trafic

Sistemele tradiționale de asistență a conducerii rutiere, cum ar fi tempomatele adaptive sunt limitate la scenarii de autostradă. Scenariile de autostradă sunt relativ simple, singurele obiecte în mișcare fiind celelalte autovehicule, obiectele statice vizibile fiind doar parapetii laterali. Pentru aceste medii, senzorii activi foarte preciși dar care furnizează cantități limitate de informație, cum ar fi radarul, sunt suficienți pentru construirea unor sisteme de asistență a conducerii rutiere. Pe de altă parte, sistemele de asistență a conducerii pentru medii urbane, se confruntă cu problema mult mai dificilă de a percepe și înțelege scenarii mult mai complexe. În scenariile urbane își fac apariția alți participanți la trafic, cum ar fi pietonii, bicicliștii. Totodată apar și alte elemente fixe, cum ar fi mașinile parcate, stâlpi, indicatoare. Prezența intersecțiilor, a drumurilor cu curbă mare, lățimea variabilă a benzilor de circulație face ca abordările bazate pe radar ale sistemelor de asistență a conducerii să nu funcționeze întotdeauna corect. Totuși, aceste sisteme, atât orientate pe confort cât și pe siguranță sunt mult mai utile în medii urbane, unde solicitarea la care este supus șoferul este mult mai mare. Deoarece senzorii bazați pe stereoviziune furnizează o cantitate mare de informație, ei tind să devină o nouă și importantă categorie de senzori folosiți de către sistemele de siguranță și confort de asistență a conducătorului auto. Unele dintre problemele pe care este necesar să le rezolve sistemele de asistare a conducerii rutiere urbane sunt: detecția obstacolelor statice (în scopul evitării acestora) în cadrul funcțiilor de asistență frontală, laterală, frânare de urgență sau protecție activă (declanșarea airbag-urilor); detecția și urmărirea obiectelor în mișcare, pentru funcțiile de asistență în traficul de tip oprit – pornit (urmărirea autovehiculului din față), frânare de urgență, limitarea efectelor coliziunilor frontale. O funcție extrem de importantă în limitarea efectelor coliziunii o reprezintă detecția părții autovehiculului corespunzătoare habitaculului, pentru a limita impactul asupra persoanelor aflate în interior în cazul unei coliziuni. Recunoașterea și protecția pietonilor reprezintă o provocare importantă pentru sistemele de asistență a conducerii în mediul urban, deoarece pietonii sunt cei mai vulnerabili participanți la trafic. O alta provocare o reprezintă detecția și recunoașterea semnelor de circulație și a semafoarelor.

4.2 Papageorgiou et al.

Una dintre primele abordări în domeniul recunoașterii pietonilor reprezintă un sistem de recunoaștere a pietonilor bazat pe wavelet-uri Haar, prezentat în [47], [48] și [49]. Dintr-un total de 1326 wavelet-uri posibile autorii selectează un număr de 29 relevante. Un algoritm SVM este folosit pentru clasificare. În [50] abordările anterioare sunt extinse pe mai multe canale de imagine.

4.3 Gavrilă și Grupul Daimler

Într-o serie de articole începând cu [51], Gavrilă și grupul de la Daimler prezintă un sistem de recunoaștere a pietonilor bazat pe potrivirea conturilor unor modele de pietoni folosind distanța Chamfer. Contribuția principală a autorului o reprezintă folosirea unui arbore de modele, reducând astfel timpul de calcul de la $O(N)$ la $O(\log N)$. În afară de distanța Chamfer, autorii propun folosirea RBF [52], trăsături Haar în cascadă [53], HPN [54]. Am studiat diverse îmbunătățiri ulterioare ale algoritmilor, prezentate în [55], [56], [57],[58], [59], [60]. Un sistem monocular bazat pe flux optic pentru detecția pietonilor este prezentat în [61]. Alte îmbunătățiri sunt prezentate în [62]. În [63] și [64] sunt prezentate metode pentru recunoașterea posturii umane. Un studiu al trăsăturilor pentru recunoașterea monoculară a pietonilor este prezentat în [65]. Un sistem folosind trei camere este prezentat în [66]. Investigații legate de folosirea informațiilor de profunzime sunt prezentate în [67] și [68].

În [69] sunt prezentate metode pentru recunoașterea pietonilor și a posturii lor dintr-un singur cadru. În [70] sunt prezentate metode de recunoaștere a pietonilor vizibili doar parțial. În [71] autorii investighează o mixtură de experți pentru recunoașterea pietonilor. În [72] se prezintă metode pentru recunoașterea posturii 3D a pietonilor folosind modele prezentate în [63]. În [73] se prezintă un sistem complet de recunoaștere și urmărire a pietonilor bazat pe stereoviziune densă. În [74] este prezentat un sistem de frânare și de evitare a pietonilor. În [75] este prezentat un set de teste pentru detecția pietonilor.

4.4 Universitatea din Parma (VISLAB)

În [76] autorii propun un sistem de detecție a pietonilor bazat pe muchii verticale și simetrie. Sistemul este rafinat în [77].

În [78] și [79] se prezintă folosirea imaginilor din spectrul infraroșu îndepărtat pentru detecția pietonilor. În [80] se prezintă o metodă de preprocesare stereo pentru recunoașterea pietonilor. În [81] și [82] se extind abordările anterioare prin folosirea rezoluției multiple. Un algoritm simplu de urmărire este prezentat în [83]. În [84] este prezentat un sistem de evaluare a algoritmilor de recunoaștere a pietonilor. În [85] și [86] este prezentată o abordare stereo în infraroșu. În [87] se discută generarea de modele pentru pietoni. În [88] autorii folosesc trăsături Haar în spectrul infraroșu. În [89] este prezentată o abordare cu 4 camere, reprezentând două sisteme stereo, unul bazat pe intensitate celălalt în spectrul infraroșu. În [90] se combină câteva abordări anterioare.

În [91] autorii investighează spectrul infraroșu apropiat. Alte abordări privind spectrul infraroșu îndepărtat sunt prezentate în [92], [93],[94]. În [95] sunt investigate contururi active. În [96] este prezentată o nouă abordare care detectează pietoni doar în zonele periculoase (descrise în [97]).

4.5 Ghandi et al.

În [98] autorii propun un sistem stereo omnidirecțional, care oferă vederi asupra întregului spațiu din jurul vehiculului. În [99] este prezentat un studiu al domeniului recunoașterii pietonilor, extins apoi în [100]. În [101] este prezentată o metodă de determinare a orientării pietonilor folosind HOG și SVM.

4.6 Alte abordări

Un sistem stereo este prezentat în [102]. În [103] se prezintă un sistem bazat pe detecția părților pietonilor. În [104], [105] și [106] se prezintă un sistem monocular în infraroșu îndepărtat. În [107] se prezintă un sistem pentru intersecții folosind fundalul fix. În [108] se prezintă un sistem stereo în infraroșu. În [109] este prezentat un sistem de urmărire a pietonilor cu camere fixe. În [110] se prezintă o metodă de detecție a pietonilor cu camere fixe pentru scenarii aglomerate.

4.7 Concluzii

În acest capitol am prezentat un studiu al literaturii în care am prezentat tehnicile actuale de clasificare a obiectelor din trafic, concentrându-ne în principal pe recunoașterea pietonilor. În secțiunea 4.1 am prezentat necesitatea clasificării obiectelor, în secțiunile 4.2, 4.3, 4.4, 4.6 am prezentat contribuțiile aduse de Papageorgiou, Gavrilă, Broggi și alți autori în acest domeniu. Studiul lucrărilor publicate în acest domeniu ne-a permis formularea următoarelor concluzii:

- Există multiple metode de detecție a obiectelor și de formulare a ipotezelor pentru pietoni. Unele se bazează pe informații 2D (de exemplu pe densitatea de muchii verticale), altele se bazează pe imagini în domeniul infraroșu îndepărtat, altele pe informații date de stereoviziune.
- Majoritatea metodelor bazate pe stereoviziune se folosesc doar de imagini de disparitate, folosind sau metode de segmentare sau metoda de v-disparitate.
- Abordările bazate pe v-disparitate sunt mai rapide, dar nu sunt precise pentru medii complexe.
- Dezavantajul sistemelor de recunoaștere bazate doar pe informații 2D este faptul că nu au informații despre scara obiectelor, trebuind să caute modele în întreaga imagine la diverse scale.
- Informația 3D generată de sistemele stereo generează adâncimi pentru anumite puncte din imagine. Există câteva metode de clasificare bazate exclusiv pe adâncime [111], [112]. Erorile și lipsa reconstrucției stereo fac dificilă folosirea directă a informației de adâncime.
- O trăsătură importantă pentru recunoașterea pietonilor în mișcare este modelul lor de mers. Există mai multe lucrări (de exemplu [113], [103], [114]) care folosesc trăsături de mișcare pentru recunoașterea pietonilor. A abordare tipică constă în determinarea faptului că mișcarea este periodică. Din păcate diferențierea mișcării obiectelor față de mișcarea fundalului este dificilă din cauza complexității mediului de trafic urban.
- Informațiile 3D nu sunt folosite nici pentru mișcare în algoritmi descriși în literatură.

- Există puține articole legate de recunoașterea altor obiecte în afară de pietoni.
- Antrenarea și testarea algoritmilor necesită seturi de date extinse. Majoritatea datelor disponibile sunt doar imagini individuale, reprezentând exemple pozitive și negative. Din această cauză ele nu pot fi folosite pentru sisteme de clasificare bazate pe informații 3D și de mișcare.

În urma acestor concluzii, vom prezenta în capitolul 7 contribuțiile noastre originale în domeniul recunoașterii obiectelor în medii de trafic urban.

Partea II

Contribuții teoretice și aplicative

Capitolul 5

Fluxul optic și estimarea mișcării camerelor

În acest capitol vom descrie contribuțiile noastre originale legate de îmbunătățirea algoritmilor existenți precum și algoritmi originali pentru calcul fluxului optic dens, detecția colțurilor și estimarea mișcării camerelor. În secțiunea 5.1 vom prezenta un algoritm pentru calculul fluxului optic dens. Originalitatea acestui algoritm constă în optimizări care fac algoritmul mai ușor de adaptat la o implementare de tip SIMD. Ca o aplicație a acestui algoritm vom prezenta în secțiunea 5.2 un algoritm original de segmentare a imaginilor bazat pe mișcare, informații 3D și intensitatea imaginii. În secțiunea 5.3 vom prezenta un algoritm original de detecție a colțurilor care este mai rapid decât algoritmii de tip Harris [5] sau Shi–Tomasi [7]. Pentru a obține îmbunătățirea timpului de execuție, am folosit muchiile Canny deja extrase, informații 3D și limitarea detecției colțurilor la părți relevante din imagine. În funcție de acuratețea dorită, algoritmul este de 3–8 ori mai rapid decât Shi–Tomasi, și este la fel de bun pentru extragerea colțurilor pentru fluxul optic. În secțiunea 5.4 vom descrie modelul camerei folosit în algoritmii de detecție a mișcării camerei. În secțiunea 5.5 vom descrie o parametrizare originală în coordonate polare a camerei, care împreună cu un algoritm de tip RANSAC este utilizată la detecția mișcării camerei mono necalibrate. În secțiunea 5.6, profitând de informația de calibrare a camerei și vom prezenta un algoritm original care este capabil să măsoare mișcarea de rotație a camerei și, cu anumite informații în plus și mișcarea de translație. Algoritmul este util pentru sistemele de asistență a conducerii autovehiculelor. În secțiunea 5.7 vom descrie un algoritm original pentru estimarea mișcării camerelor stereo, originalitatea acestui algoritm constând în alegerea punctelor pentru calcularea mișcării camerelor, și determinarea separată a translației printr-o metodă stabilă. În secțiunea 5.8 prezentăm concluziile acestui capitol.

5.1 Algoritm în timp real pentru calcularea fluxului optic dens

În această secțiune vom prezenta un algoritm optimizat pentru calcularea fluxului optic dens în timp real. Optimizările se bazează pe eliminarea redundanțelor în rezolvarea ecuațiilor prin metoda celor mai mici pătrate pentru fluxul optic dens de tip Lucas-Kanade.

Optimizările propuse conduc la o îmbunătățire a vitezei de execuție cu un factor de 5.

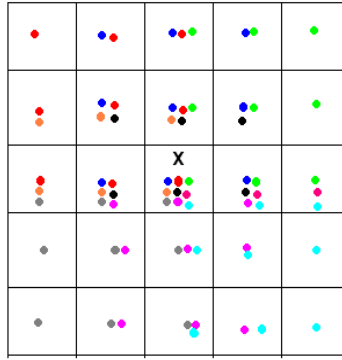


Figura 5.1: Fiecare poziție a ferestrei de 3x3 pixeli pe care se calculează fluxul optic calculează două ecuații pentru 9 puncte. Un calcul prin forță brută are un factor de redundanță de 9

5.2 Segmentarea imaginilor bazată pe mișcare, intensitate și profunzime

În această secțiune prezentăm un algoritm de segmentare a imaginilor bazat pe creșterea regiunilor și o măsură de similaritate compusă din intensitate, profunzime și mișcare 3D:

$$[C(A, B) = \alpha|Z_0^A - Z_0^B| + \beta|I_0^A - I_0^B| + \gamma|V_x^A - V_x^B| + \delta|V_y^A - V_y^B| + \varepsilon|V_z^A - V_z^B|] \quad (5.1)$$



Figura 5.2: Rezultatele algoritmului de segmentare

5.3 Algoritm original de selecție a trăsăturilor pentru fluxul optic

Algoritmul propus are următorii pași:

1. Pentru fiecare punct reconstruit:
2. (Opțional) doar pentru puncte 3D dintr-o anumită clasă;
3. (Opțional) respinge punctele care se află lângă o muchie a imaginii de profunzime;

4. Pe o fereastră de n pe n în jurul fiecărui punct se calculează numărul minim dintre muchiile verticale și orizontale și minimul dintre muchiile orientate pe cele două diagonale. Se calculează apoi maximul acestor muchii.
5. Se face o supresie a non-maximelor pe imaginea anterioară.
6. (Opțional) se creează o histogramă a imaginii anterioare, din care se deduce un prag în funcție de numărul de colțuri dorit.
7. Se aplică pragul, rămânând doar punctele de colț în imagine.

Am testat mai multe versiuni ale acestui algoritm. Rezultatele sunt comparabile calitativ cu algoritmul Shi–Tomas timpul de execuție fiind de 3–8 ori mai rapid în funcție de acuratețea dorită și de numărul de puncte selectat.

5.4 Modelul camerei și al mișcării

Modelul perspectivă este definit prin următoarele ecuații:

$$x = \frac{f_x X}{Z} + c_x \quad (5.2)$$

$$y = \frac{f_y Y}{Z} + c_y \quad (5.3)$$

5.5 Algoritm original pentru estimarea mișcării camerei 2D și detecția unei noi secvențe

În această secțiune vom descrie un algoritm care este capabil să extragă cele mai uzuale mișcări ale camerei din filme în format digital, adică mișcările de panning vertical și orizontal, zoom-in, zoom-out precum și începutul unei noi secvențe. Intrarea algoritmului este constituită dintr-un număr mare vectori 2D de mișcare. Nu cunoaștem parametrii camerei. Pentru a limita numărul de grade de libertate al camerei, vom folosi o parametrizare în coordonate polare a camerei.

$$X = R \cos \beta \sin \alpha; Y = R \sin \beta; Z = R \cos \beta \cos \alpha \quad (5.4)$$

$$x = \frac{f R \cos \beta \sin \alpha}{R \cos \beta \cos \alpha} = f \tan \alpha \quad (5.5)$$

$$y = \frac{f R \sin \beta}{R \cos \beta \cos \alpha} = f \frac{\tan \beta}{\cos \alpha} \quad (5.6)$$

$$\alpha = \arctan \frac{x}{f}; \beta = \arctan \frac{y \cos \alpha}{f} \quad (5.7)$$

Deplasamentele din fluxul optic sunt:

$$x_1 = x_0 + \Delta x; y_1 = y_0 + \Delta y \quad (5.8)$$

$$x_0 + \Delta x = (f_0 + \Delta f) \tan (\alpha_0 + \Delta \alpha) \quad (5.9)$$

$$y_0 + \Delta y = (f_0 + \Delta f) \frac{\tan(\beta_0 + \Delta\beta)}{\cos(\alpha_0 + \Delta\alpha)} \quad (5.10)$$

Printr-o aproximare Taylor obținem:

$$\frac{f_0}{\cos^2 \alpha_0} \Delta\alpha + \frac{x_0}{f_0} \Delta f = \Delta x \quad (5.11)$$

$$\frac{x_0 y_0}{f_0} \Delta\alpha + \frac{f_0}{\cos^2 \beta_0 \cos \alpha_0} \Delta\beta + \frac{y_0}{f_0} \Delta f = \Delta y \quad (5.12)$$

$\Delta\alpha$, $\Delta\beta$ și Δf reprezintă valorile pe care trebuie să le determinăm. Pentru a elimina mișcările cauzate de alte cauze decât mișcarea camerei folosim un algoritm de tip RANSAC [115]. Am testat sistemul pe circa 3 ore, obținând o acuratețe de 91% pentru detectarea unei noi secvențe și 88% pentru detecția calitativă a mișcării.

5.6 Algoritm original pentru estimarea mișcării camerei 3D calibrate

Am extins algoritmul prezentat în secțiunea anterioară pentru a profita de informațiile de calibrare ale camerei. Acest algoritm este util pentru sistemele de viziune pentru asistența conducerii auto, putând detecta mișcarea vehiculului propriu. Folosind informația de calibrare putem extrage mișcările de rotație a camerei în mod exact. Având anumite informații suplimentare, putem calcula și mișcarea de translație a vehiculului propriu.

Pentru calibrarea camerei am folosit algoritmi descriși în [116] și [117]. Știind c_y , f_x și f_y putem determina $\Delta\alpha$ și $\Delta\beta$ exact. Cunoscând distanța dintre două puncte din scenă, putem determina mișcarea de translație cu exactitate:

$$x'_0 - x''_0 = \frac{f_x(X' - X'')}{Z_0}; \quad x'_1 - x''_1 = \frac{f_x(X' - X'')}{Z_1} \quad (5.13)$$

$$y'_0 - y''_0 = \frac{f_y(Y' - Y'')}{Z_0}; \quad y'_1 - y''_1 = \frac{f_y(Y' - Y'')}{Z_1} \quad (5.14)$$

Mișcarea de translație este dată de $\Delta Z = Z_0 - Z_1$.

Algoritmul funcționează la 100ms per cadru pe o configurație 2.6 GHz Intel Core 2 Duo.

5.7 Algoritm original pentru estimarea mișcării camerelor 3D stereo

Algoritmul descris în această secțiune permite determinarea mișcării camerelor stereo. El este de asemenea util pentru sistemele de asistență a conducerii, de această dată însă folosind camere stereo. Pentru a filtra erorile care apar din cauza fluxului optic și a stereoviziunii, avem optat să determinăm mișcarea camerei în doi pași. Primul pas este determinarea tranzației, printr-o metodă stabilă (nu prin metoda celor mai mici pătrate):

$$\Delta Z = \frac{1}{N} \sum_{i=1}^N Z_1^i - Z_0^i \quad (5.15)$$

Apoi determinăm mișcarea folosind:

$$X_1^i = r_1 X_0^i; Y_1^i = r_2 Y_0^i; i = \overline{1, N} \quad (5.16)$$

unde r_1 și r_2 sunt primele două linii ale matricei de rotație. Normalizăm aceste linii la lungime 1 și calculăm $r_3 = r_1 \times r_2$. Obținem apoi unghiurile instantanee de rotație ca fiind:

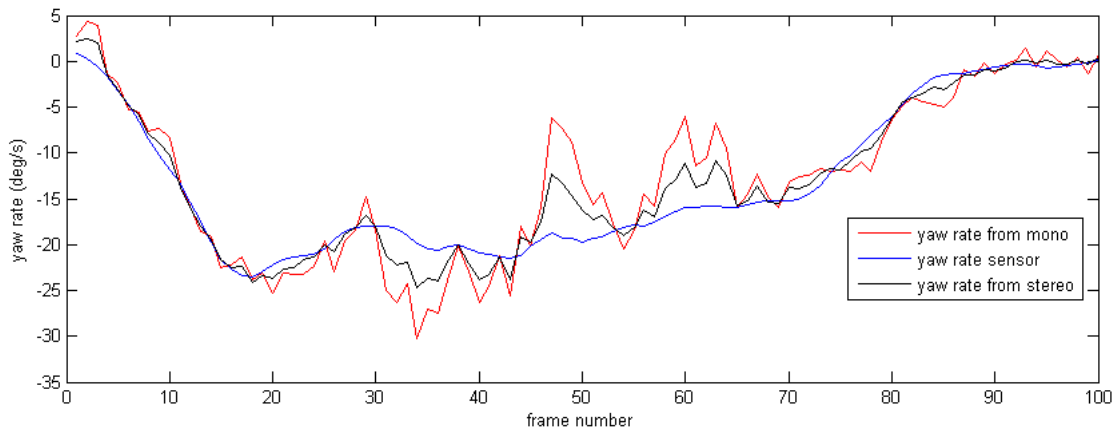
$$\begin{aligned} v_{yaw} &= P_{xOz} R(0 \ 0 \ 1)^T \\ \Delta_{yaw} &= \arccos \frac{v_{yaw}(0 \ 0 \ 1)}{\|v_{yaw}\|} \end{aligned} \quad (5.17)$$

$$\begin{aligned} v_{pitch} &= P_{yOz} R(0 \ 0 \ 1) \\ \Delta_{pitch} &= \arccos \frac{v_{pitch}(0 \ 0 \ 1)}{\|v_{pitch}\|} \end{aligned} \quad (5.18)$$

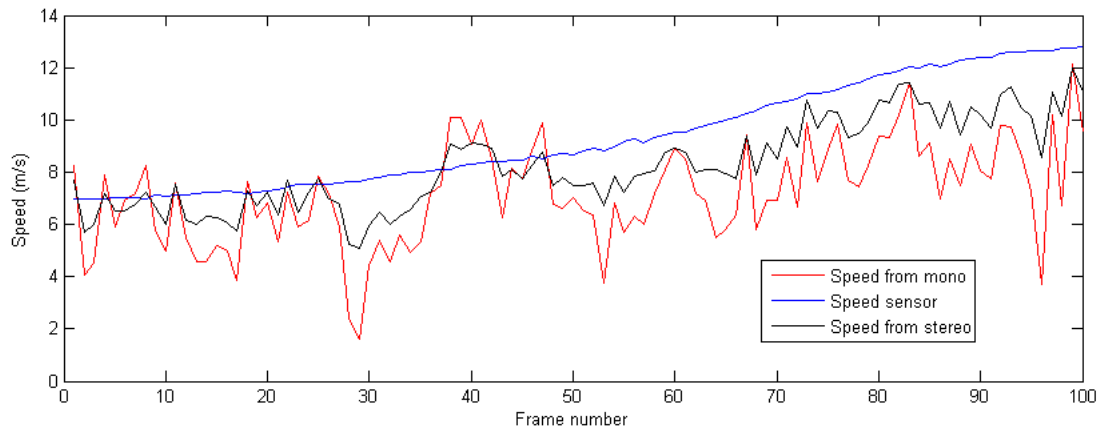
$$\begin{aligned} v_{roll} &= P_{xOy} R(1 \ 0 \ 0) \\ \Delta_{roll} &= \arccos \frac{v_{roll}(1 \ 0 \ 0)}{\|v_{roll}\|} \end{aligned} \quad (5.19)$$

5.8 Concluzii

În acest capitol am descris contribuțiile noastre originale legate de îmbunătățirea algoritmilor existenți precum și algoritmi originali pentru calcul fluxului optic dens, detecția colțurilor și estimarea mișcării camerelor. În secțiunea 5.1 am prezentat un algoritm pentru calculul fluxului optic dens. Originalitatea acestui algoritm constă în optimizări care fac algoritmul mai ușor de adaptat la o implementare de tip SIMD. Ca o aplicație a acestui algoritm am prezentat în secțiunea 5.2 un algoritm original de segmentare a imaginilor bazat pe mișcare, informații 3D și intensitatea imaginii. În secțiunea 5.3 am prezentat un algoritm original de detecție a colțurilor care este mai rapid decât algoritmi de tip Harris [5] sau Shi–Tomasi [7]. Pentru a obține îmbunătățirea timpului de execuție, am folosit muchiile Canny deja extrase, informații 3D și limitarea detecției colțurilor la părți relevante din imagine. În funcție de acuratețea dorită, algoritmul este de 3–8 ori mai rapid decât Shi–Tomasi, și este la fel de bun pentru extragerea colțurilor pentru fluxul optic. În secțiunea 5.4 am descris modelul camerei folosit în algoritmi de detecție a mișcării camerei. În secțiunea 5.5 am descris o parametrizare originală în coordonate polare a camerei, care împreună cu un algoritm de tip RANSAC s-au utilizat la detecția mișcării camerei mono necalibrate. Am testat sistemul nostru folosind circa 3 ore de film în format digital. Acuratețea de detecție a începutului unei noi secvențe a fost de 91%, iar acuratețea estimării calitative a mișcării camerelor a fost de 88%. În secțiunea 5.6 am dezvoltat algoritmul anterior, profitând de informația de calibrare a camerei și am dezvoltat un algoritm original care este capabil să măsoare mișcarea de rotație a camerei și, cu anumite informații în plus și mișcarea de translație. Algoritmul este util pentru sistemele de asistență a conducerii autovehiculelor. Aspectul original al acestui algoritm constă în folosirea coordonatelor polare și în eliminarea mișcării de translație considerând că ea ar fi un efect de zoom. În secțiunea 5.7 am descris un algoritm original pentru estimarea mișcării camerelor stereo care folosește detectorul de colțuri propriu, descris în secțiunea 5.3. Originalitatea acestui algoritm constă în alegerea punctelor pentru calcularea mișcării



(a) Yaw rate

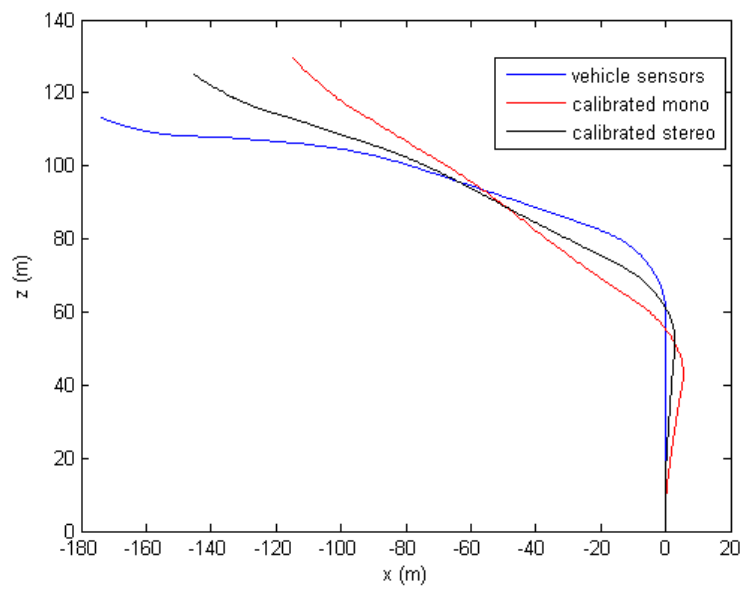


(b) Speed

Figura 5.3: Comparație a vitezei și vitezei de rotație obținute folosind algoritmul de detecție a mișcării camerei mono, stereo și cea determinată din senzorii vehiculului propriu



(a) Fotografii din avion



(b) Traiectorii

Figura 5.4: Comparație în calculul traiectoriei obținute folosind algoritmul de detecție a mișcării camerei mono, stereo și cea determinată din senzorii vehiculului propriu

camerelor, și determinarea separată a translație printr-o metodă stabilă. Rezultatele prezentate în acest capitol au fost publicate într-un articol [118].

Capitolul 6

Urmărirea obiectelor în medii de trafic urban

În acest capitol vom descrie un algoritm de urmărire original care rezolvă problema urmăririi obiectelor multiple, ierarhice, în medii de trafic urban. În secțiunea 6.1 vom prezenta fundamentele matematice ale agregării de date zgomotoase din mai multe surse, în secțiunea 6.2 vom prezenta arhitectura sistemului de stereoviziune pentru detecția și urmărirea obiectelor din trafic, în secțiunea 6.3 vom identifica sursele erorilor care trebuie filtrate de algoritmul de urmărire, și, bazat pe ele vom introduce în secțiunea 6.4 un model original, avansat, probabilistic static și dinamic pentru reprezentarea obiectelor. În secțiunea 6.5 vom descrie faza de predicție a algoritmului de urmărire. În secțiunea 6.6 vom prezenta un algoritm original de asociere a obiectelor, pentru obiecte reprezentate ierarhic. În secțiunea 6.7 vom prezenta arhitectura filtrelor Kalman utilizate. În secțiunea 6.8 am prezenta rezultatele experimentale obținute, iar în secțiune 6.9 vom prezenta concluziile acestui capitol

6.1 Agregarea datelor zgomotoase din mai multe surse

Să considerăm problema agregării datelor zgomotoase din mai multe surse. Având mai multe observații Z_i având un zgomot Gaussian de medie zero descris de covarianțele Σ_i cu $i = 1, \dots, n$. Dorim să calculăm o medie ponderată optimă a acestor observații, cu factorii de mediere K_i :

$$\sum_{i=1}^n K_i = I \quad (6.1)$$

Obținem următoarea expresie pentru coeficienții K_i :

$$K_i = \left(\sum_{j=1}^n \Sigma_j^{-1} \right)^{-1} \Sigma_i^{-1} \quad (6.2)$$

Matricea de covarianță a zgomotului estimării:

$$\Sigma = \left(\sum_{i=1}^n \Sigma_i^{-1} \right)^{-1} \quad (6.3)$$

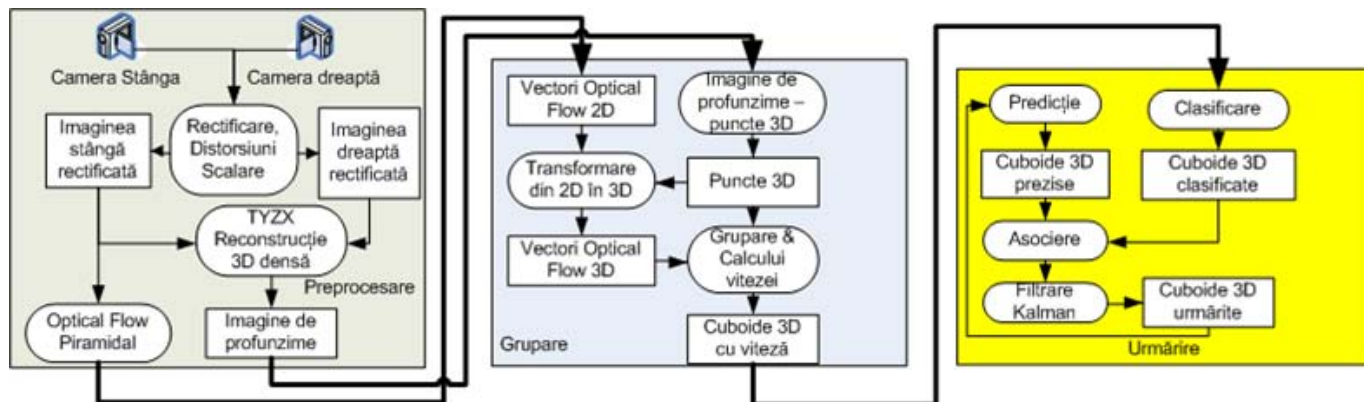


Figura 6.1: Arhitectura sistemului, constând din modulele de preprocesare, grupare și urmărire

6.2 Arhitectura sistemului de urmărire

Arhitectura sistemului este prezentată în figura 6.1.

6.3 Sursele erorilor care trebuie filtrate de sistemul de urmărire

Algoritmul de grupare generează o serie de erori tipice. Primul tip de eroare este generat de erorile în reconstrucție stereo. Modelăm acest tip de eroare prin asignarea unei matrice de covarianță a colțurilor proporțională cu dimensiunea celulelor spațiului în care se face gruparea punctelor în obiecte. Celelalte tipuri de erori sunt cauzate de divizarea eronată a obiectelor în obiecte mai mici, sau unirea accidentală a obiectelor în obiecte mai mari.

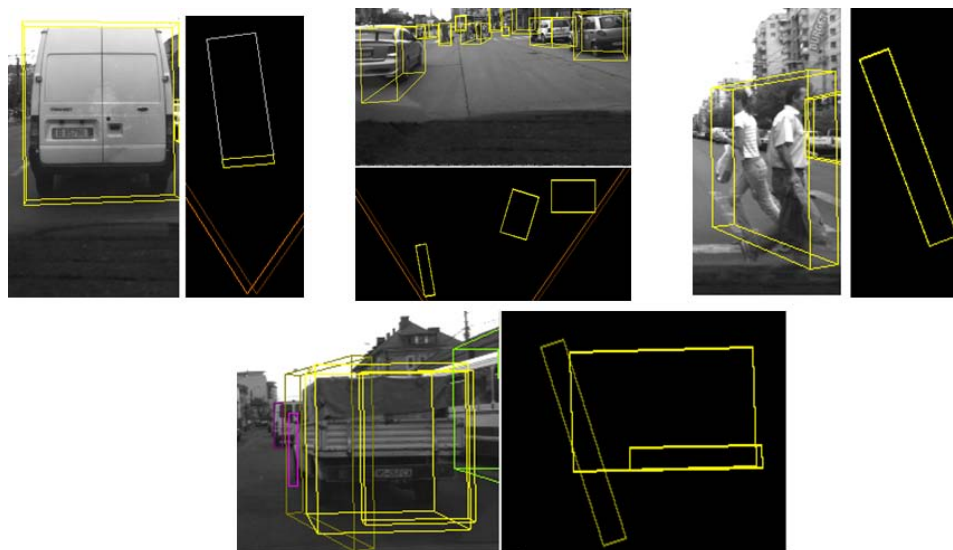


Figura 6.2: Erori de grupare, sus: spatele obiectului este invizibil, orientare greșită, unire greșită a obiectelor, divizarea eronată a obiectelor

6.4 Model original al obiectelor

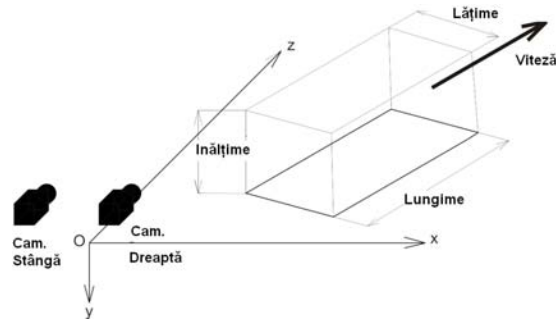


Figura 6.3: Sistemul de coordonate și modelul obiectelor (vedere de perspectivă)

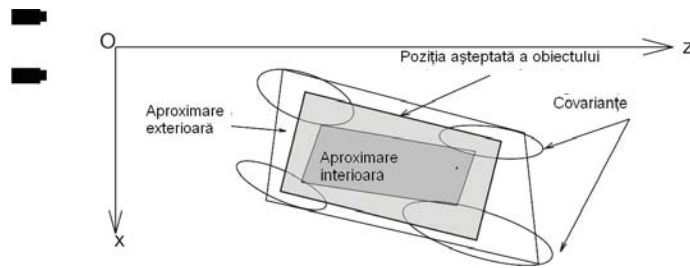


Figura 6.4: Sistemul de coordonate și modelul obiectelor, vedere de sus

6.4.1 Modelul static al obiectelor

Partea statică a modelului obiectelor este descrisă prin următoarele valori: (vezi Fig. 6.4)

- Patru coordonate 2D descriind poziția, în planul xOz a celor patru colțuri ale bazei fiecărui cuboid.
- Patru matrice de covarianță de dimensiune 2 descriind incertitudinile fiecărui colț al bazei.
- Două coordonate y pentru fețele superioare și inferioare ale obiectelor.
- Două varianțe reprezentând incertitudinile legate de coordonatele y .
- Un set de obiecte fiu precum și un obiect părinte.

Covarianțele pozițiilor colțurilor au proprietatea că erorile lor sunt mai mari de-a lungul razelor optice, din cauza modului în care se face reconstrucția stereo (vezi Fig. 6.7, covarianțele fiind reprezentate prin elipse). O problemă suplimentară apare la obiectele care sunt apropiate de marginea câmpului vizual. Deoarece algoritmul de detecție a obiectelor nu poate vedea dincolo de câmpul vizual, obiectele care trec de marginea acestuia pot fi tăiate. Pentru a trata acest caz, creștem covarianțele colțurilor care sunt apropiate de marginea câmpului vizual (vezi fig. 6.8), modelând astfel această situație.

Obiectele nu sunt doar o mulțime de colțuri. Ele trebuie tratate ca niște obiecte solide în spațiu, mai ales în faza de asociere. Am ales să modelăm aspectul de solid ca o funcție de distribuție de

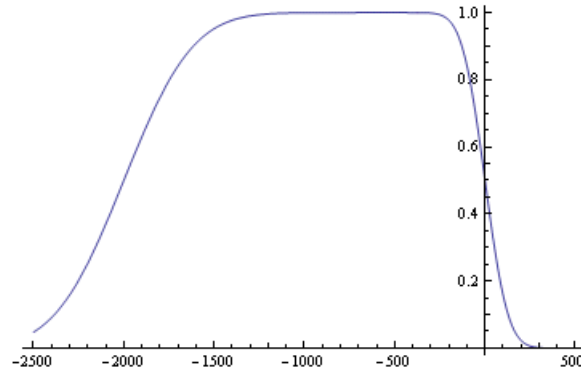


Figura 6.5: Densitatea probabilistică a obiectului pe axa Oy

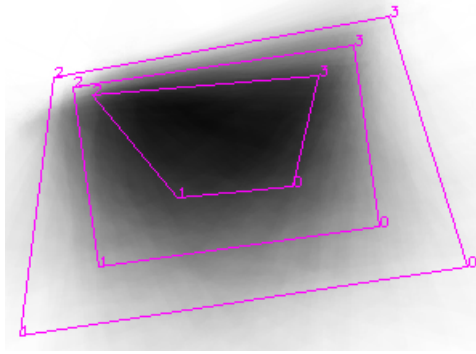


Figura 6.6: Densitatea probabilistică a obiectului în planul xOz

probabilitate cumulată. Punctele aflate mai aproape de centrul cuboidului au o probabilitate mai mare de a aparține obiectului decât cele de la margine. Punctele aflate în afara cuboidului au și ele la rândul lor o mică probabilitate de a aparține de fapt obiectului.

Pe direcția y ne confruntăm cu problema de a modela un interval probabilistic, având ca și capete două variabile aleatoare distribuite normal. Suntem practic interesați de probabilitatea:

$$P(y_1 < y < y_2) = 0.25 * (1 + \operatorname{erf}((y_1 - y)/\sqrt{2}\sigma_1)(\operatorname{erf}(y_2 - y)/(\sqrt{2}\sigma_2))). \quad (6.4)$$

În fig. 6.5 este ilustrată densitatea pe direcția y a unui cuboid având $y_1 = -2000$, $\sigma_1 = 300$, $y_2 = 0$, $\sigma_2 = 150$ (mm). Dacă $y_2 - y_1 \gg \sigma_1 + \sigma_2$ atunci lungimea medie a intervalului probabilistic $E(y_2 - y_1, y_1 < y_2) \approx E(y_2) - E(y_1)$ se apropie de lungimea intervalului neprobabilistic.

În planul xOz situația este mai complicată. Am dori să reprezentăm probabilitatea ca un punct 2D dat să se afle în interiorul unui poligon convex, ale cărui vârfuri sunt puncte 2D aleatoare distribuite normal în jurul unei medii și având o matrice de covarianță cunoscută. Scrierea unei ecuații pentru acest caz este extrem de dificilă. O simulare a acestei funcții de probabilitate este arătată în fig. 6.6. Dreptunghiul din mijloc reprezintă perimetrul obiectului, patruleterele din interior și exterior sunt situate la distanța de o deviație standard (pe direcția centru de masă — colț) de fiecare vârf. Din nou, dacă covarianțele colțurilor nu sunt prea mari, aria medie a poligonului probabilistic se apropie de aria sa non-probabilistică.

Deoarece folosirea directă a densităților de probabilitate ar fi costisitoare din punct de vedere computațional, am ales în primă fază, să aproximăm aceste densități de probabilitate printr-o funcție de tip treaptă, folosind trei cuboide în loc de unul (vezi fig. 6.4). Cele 3 cuboide sunt generate prin

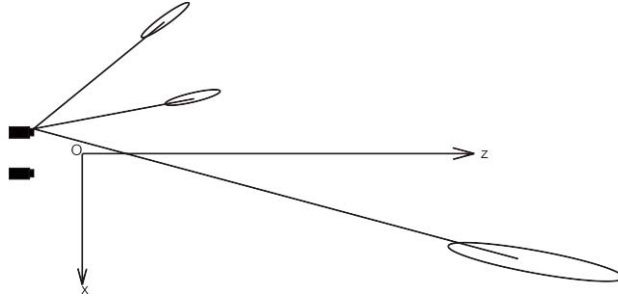


Figura 6.7: Covarianțele pozițiilor în funcție de razele optice

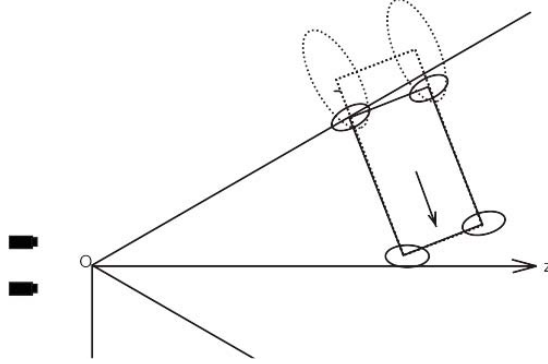


Figura 6.8: Covarianțele colțurilor apropiate de marginea câmpului vizual

mutarea colțurilor pe direcția centru de masă – colț. Amplitudinea acestei translații este dată de deviațiile standard pe această direcție, conform matricelor de covarianță din colțuri, și înmulțite cu următorii factori: -1.69 , -0.79 și 0.79 , conform ecuației:

$$f d^T \Sigma d \quad (6.5)$$

unde Σ reprezintă matricea de covarianță, d direcția și f factorul de mai sus.

Ponderile acestor cuboide sunt 0.206 , 0.4 , 0.206 . Aceste valori au fost alese pentru a obține o aproximare optimă a distribuției de probabilitate cumulative Gaussiene cu o funcție treaptă.

Ca o alternativă mai flexibilă și mai precisă, am propus un al doilea tip de modelare al densității, folosind metode Monte-Carlo. Se generează patrulater aleatoare în planul xOz și intervale aleatoare pe axa Oy , volumul cuboidului fiind reprezentat de media acestora.

6.4.2 Modelul dinamic al obiectelor

Partea dinamică a modelului obiectelor este dată de presupunerea că obiectele se mișcă rectiliniu și uniform. Evident, această presupunere este doar o simplificare, deoarece atât vehiculele cât și pietonii își pot schimba cu ușurință direcția. Pentru a modela aceste erori, introducem o componentă de zgomot la predicția acestora.

Un obiect are asociat deci un vector de viteză în planul xOz și matricea de covarianță de rang 2 atașată. Practic, matricea de covarianță a vitezei este calculată ca o medie armonică armonică a matricelor de covarianță asociate vectorilor de flux optic tridimensionali din a căror medie rezultă calculul vectorului de viteză. Utilizarea acestei medii rezultă din considerații legate de asocierea datelor zgomotoase (similar ecuației Riccati a filtrului Kalman standard).

6.5 Predicția

Pasul de predicție aplică modelul dinamic al vehiculului și al obiectelor din cadrul anterior pentru a prezice starea acestora în cadrul curent. Din cauză că modelul dinamic al obiectelor nu este decât o aproximare, se adaugă un anumit zgomot de model. Acest zgomot depinde de clasa obiectelor. Pentru vehicule modelul zgomotului este o matrice de covarianță alungită pe direcția mișcării, deoarece vehiculele tind să se miște pe traiectorii rectilinii. Matricea de covarianță pentru pietoni și pentru celelalte obiecte este omogenă deoarece aceste clase de obiecte nu au de obicei constrângeri de mișcare.

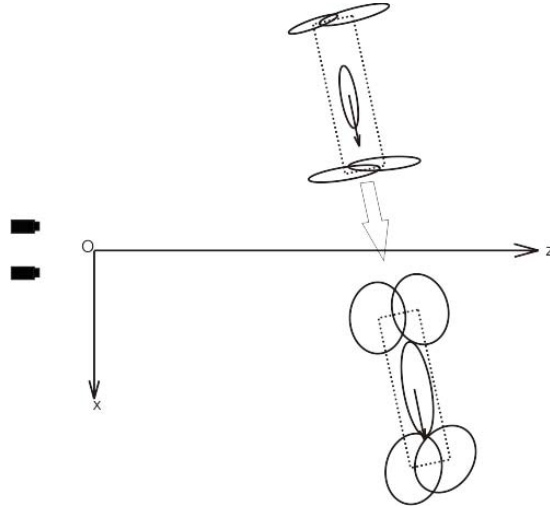


Figura 6.9: Efectele propagării asupra covarianțelor

6.6 Algoritmul original de asociere a obiectelor

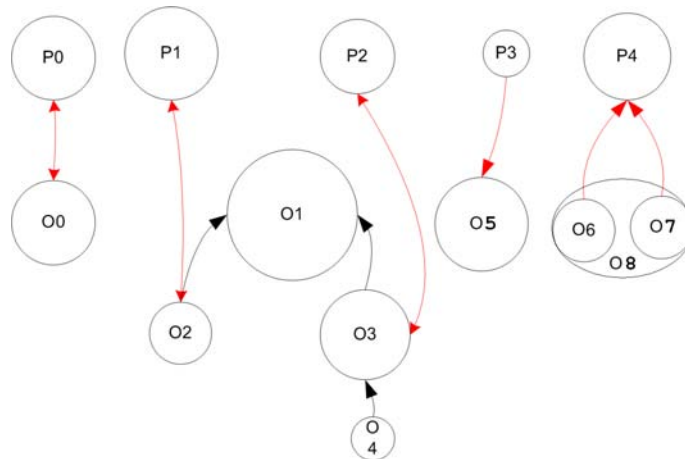


Figura 6.10: Două păduri de arbori, reprezentând obiectele prezise P și obiectele observate O. Săgețile negre indică structura ierarhică, săgețile roșii indică asocierile

Algoritmul de asociere constă din următorii pași:

1. În primul pas se consideră doar rădăcinile arborilor.

2. Se aplică algoritmul lui Kuhn pentru a găsi obiectele asociate unu la unu.
3. Pentru obiectele rămase se determină dacă obiectul prezis este un fiu al obiectului detectat sau viceversa.
4. Se creează obiecte mari conținând copii obiectelor detectați în pasul anterior. Se încearcă potrivirea unu la unu a acestor obiecte noi cu obiectele corespunzătoare.
5. Se aplică recursiv algoritmul pe obiectele copil.

6.7 Filtrarea Kalman

Filtrarea Kalman folosită a fost una standard. Am folosit 2 filtre Kalman pentru coordonatele y ale feței superioare și inferioare a cuboidelor, 4 filtre Kalman pentru pozițiile celor 4 colțuri și un filtru Kalman pentru viteză. Originalitatea filtrelor folosite constă în integrarea a 3 surse de date pentru viteză, viteza calculată din fluxul optic, viteza prezisă și viteza calculată din asocierie.

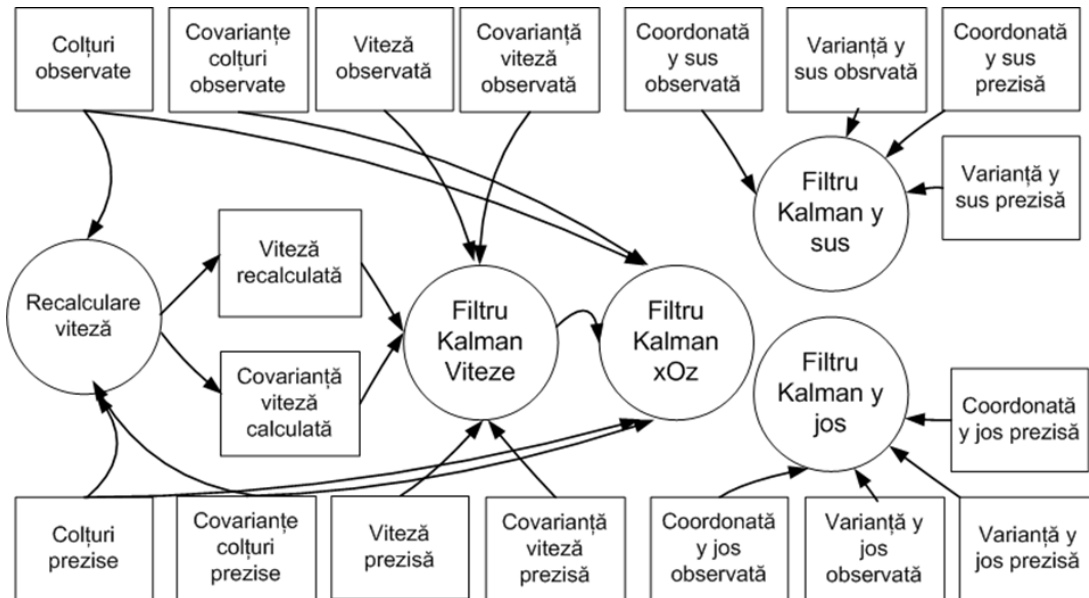


Figura 6.11: Filtrarea Kalman

6.8 Rezultate experimentale

Am testat sistemul nostru folosind atât scenarii reale de trafic urban cât și scenarii controlate. Pentru scenariile controlate am obținut următoarele rezultate:

Am testat sistemul nostru și pe un număr mare de scenarii urbane, complexe. Un procent 71.3% din obiecte au fost urmărite. Durata medie de urmărire a fost de 5.1 secunde. Numărul mediu de obiecte dintr-un cadru a fost de 13.5. Timpul de rulare pe cadru a fost de 1.9 milisecunde pe configurație Intel Core 2 la 2.66GHz. Algoritmul a fost implementat în C++.

Unghi de apropiere	Viteză	Eroare medie pătratică
45	40kmh	4.2kmh
45	50kmh	4.5kmh
45	60kmh	8.2kmh
90	40kmh	3.1kmh
90	50kmh	4.2kmh
90	60kmh	6.5kmh

Tabela 6.1: Viteza obținută din algoritmul de urmărire

6.9 Concluzii

În acest capitol am descris un algoritm de urmărire original care rezolvă problema urmăririi obiectelor multiple, ierarhice, în medii de trafic urban. Originalitatea algoritmului dezvoltat constă în următoarele aspecte:

- Viteza obiectelor a fost calculată folosind fluxul optic. O combinație de filtre mediane și de medie a fost folosită pentru agregarea vectorilor de flux optic, eliminând erorile majore.
- Obiectele au fost reprezentate printr-un model probabilistic original avansat. Modelul probabilistic ia în considerare caracteristicile algoritmilor de detecție a obiectelor bazați pe stereoviziune.
- Modelul probabilistic al obiectelor conduce la o densitate de probabilitate a obiectului în spațiul 3D. Am aproximat această densitate prin două metode, prima folosind un set de cuboide imbricate, iar a doua printr-o metodă de tip Monte-Carlo, folosind un set de poligoane generate aleator;
- Faza de asociere este implementată folosind un algoritm original, care este capabil să trateze obiecte reprezentate ierarhic, precum și cazul obiectelor divizate sau unite în mod eronat;
- Partea de filtrare Kalman efectivă agregă 3 estimări pentru viteze, viteza calculată din flux optic, viteza prezisă precum și o viteză calculată implicit din faza de asociere a obiectelor.
- S-au făcut numeroase optimizări pentru a face ca algoritmul să funcționeze în timp real.

În secțiunea 6.1 am prezentat fundamentele matematice ale agregării de date zgomotoase din mai multe surse, în secțiunea 6.2 am prezentat arhitectura sistemului de stereoviziune pentru detecția și urmărirea obiectelor din trafic, în secțiunea 6.3 am identificat sursele erorilor care trebuie filtrate de algoritmul de urmărire, și, bazat pe ele am introdus în secțiunea 6.4 un model original, avansat, probabilistic static și dinamic pentru reprezentarea obiectelor. În secțiunea 6.5 am descris faza de predicție a algoritmului de urmărire. În secțiunea 6.6 am prezentat un algoritm original de asociere a obiectelor, pentru obiecte reprezentate ierarhic. În secțiunea 6.7 am prezentat arhitectura filtrelor Kalman utilizate. În secțiunea 6.8 am prezentat rezultatele experimentale obținute, care arată că algoritmul de urmărire este capabil să urmărească majoritatea obiectelor din scenă, pentru un timp suficient de lung, cu un timp de inițializare de un singur cadru, îmbunătățind astfel performanța

detecriei obiectelor. Algoritmul are un timp de rulare mediu de doar 1.9 ms pe cadru. Rezultatele din capitolul 6 au fost publicate în două articole,[119] și [120].

Capitolul 7

Recunoașterea obiectelor din trafic

În acest capitol vom prezenta contribuțiile noastre originale în domeniul recunoașterii obiectelor în medii de trafic urban. În secțiunea 7.2 vom descrie două trăsături originale pentru recunoașterea pietonilor, bazate pe mișcare: semnătura de mișcare și periodicitatea mișcării. Aceste trăsături se bazează pe analiza câmpului de mișcare 3D și pe faptul că pietonii au un model de mișcare articulat. Pentru îmbunătățirea rezultatelor și pentru recunoașterea pietonilor staționari vom introduce în secțiunea 7.3 un număr de trăsături simple și ușor de calculat aplicabile tuturor obiectelor. Pentru agregarea acestor trăsături multiple, vom prezenta în secțiunea 7.4 un prim algoritm, bazat pe un clasificator Bayesian naiv. Acest clasificator este util atunci când datele de antrenare sunt puține și când trăsăturile sunt independente. Pentru folosirea mai multor trăsături, mai complexe și al unui număr mai mare de date vom prezenta arhitectura originală a unui clasificator mai puternic, în secțiunea 7.5. Acest clasificator poate trata mai multe clase de obiecte, mai multe trăsături și poate folosi mai mulți algoritmi de clasificare. Pentru a avea suficiente date pentru antrenarea și testarea clasificatorilor, vom prezenta, în secțiunea 7.6, arhitectura originală a unei baze de date relaționale. Aceasta conține informații spațiale pentru localizarea în spațiu a obiectelor, trăsături și clase etichetate manual.

7.1 Introducere

Unele dintre problemele pe care este necesar să le rezolve sistemele de asistare a conducerii rutiere urbane sunt: detecția obstacolelor statice (în scopul evitării acestora) în cadrul funcțiilor de asistență longitudinală, laterală, frânare de urgență sau protecție activă (declanșarea airbag-urilor); detecția și urmărirea obiectelor în mișcare, pentru funcțiile de asistență în traficul de tip oprit – pornit (urmărirea autovehiculului din față), frânare de urgență, limitarea efectelor coliziunilor frontale. O funcție extrem de importantă în limitarea efectelor coliziunii o reprezintă detecția părții autovehiculului corespunzătoare habitaculului, pentru a limita impactul asupra persoanelor aflate în interior în cazul unei coliziuni. Recunoașterea și protecția pietonilor reprezintă o provocare importantă pentru sistemele de asistență a conducerii în mediul urban, deoarece pietonii sunt cei mai vulnerabili participanți la trafic. O alta provocare o reprezintă detecția și recunoașterea semnelor de circulație și a semafoarelor.

Din cele discutate mai sus, reiese faptul că recunoașterea obiectelor din trafic este esențială în dezvoltarea sistemelor de asistență a conducerii rutiere în medii de trafic urbane.

7.2 Trăsături originale de mișcare pentru recunoașterea pietonilor

Pietonii sunt obiecte articulate. Ei nu se mișcă rigid și, din această cauză, câmpul vectorilor de mișcare asociați lor nu este unul omogen. Majoritatea celorlalte obiecte din trafic au o mișcare de tip obiect rigid, și, din acest motiv, au un câmp al vectorilor de mișcare omogen. În această secțiune vom defini două trăsături originale care captează aceste diferențe.

Primii pași efectuați sunt eliminarea mișcării globale a obiectelor, eliminarea punctelor de fundal, extragerea colțurilor, și calcularea fluxului optic 2D urmat de transformarea acestuia în 3D (vezi fig. reffigoflow).

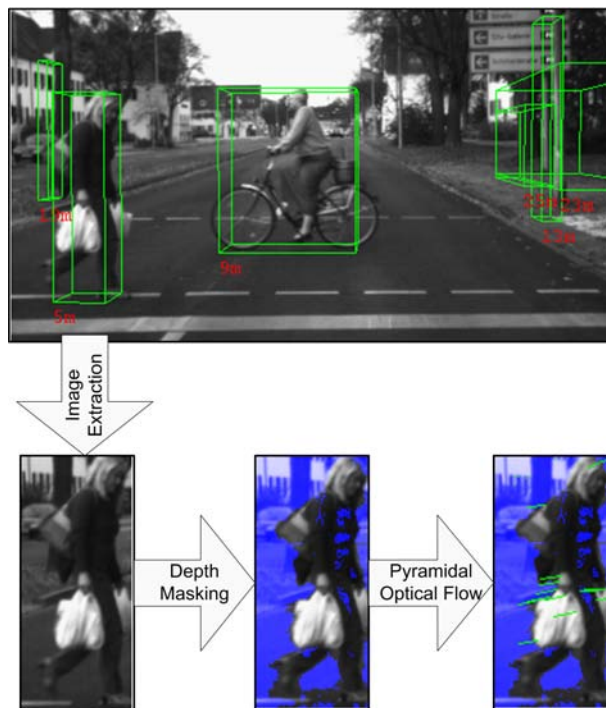


Figura 7.1: Extragerea imaginilor obiectelor, aplicarea măștii de profunzime, calcul fluxului optic. Pixelii mascați (de fundal) sunt desenați cu albastru, vectorii de mișcare sunt desenați cu verde.

7.2.1 Semnătura de mișcare

Pentru a captura variabilitatea câmpului vectorilor de mișcare tridimensională asociat pietonilor calculăm matricea de covarianță a acestuia:

$$C = \frac{1}{n} \sum_{i=1}^n (V_i^w - \mu_V)(V_i^w - \mu_V)^T \quad (7.1)$$

Semnătura de mișcare este constituită de proiecția în plan orizontal a componentei principale a acestei matrice, $\lambda_x y$. Rezultatele experimentale au arătat că aceasta reprezintă o trăsătură utilă pentru a diferenția pietonii în mișcare de celelalte obiecte din trafic. Din cauză că există erori în calculul

fluxului optic, este necesară o filtrare în timp a semnăturii de mișcare. Am realizat această filtrare printr-o metodă simplă:

$$\lambda_{xz}^t = (\alpha - 1)\lambda_{xz} + \alpha\lambda_{xz}^{t-1}. \quad (7.2)$$

. O valoare $\alpha = 0.5$ a dat rezultatele cele mai bune.

7.2.2 Periodicitatea semnăturii de mișcare

O limitare a semnăturii de mișcare este constituită de zgomotele asociate vectorilor de mișcare ale obiectelor rigide (non-pietoni). Aceste zgomote pot genera o semnătură de mișcare eronată de mare. Un altă limitare apare în situația opusă, atunci când, în anumit moment al ciclului de mișcare al pietonilor mișcarea care apare are o amplitudine mică. Figura 7.2 prezintă un grafic în timp al semnăturii de mișcare tipice a pietonilor și a altor obiecte din trafic. Se poate observa că semnătura de mișcare a pietonilor este mare, dar cu un anumit grad de variabilitate periodică, iar semnătura de mișcare a celorlalte obiecte este de obicei mică, cu anumit zgomot de tip impuls.

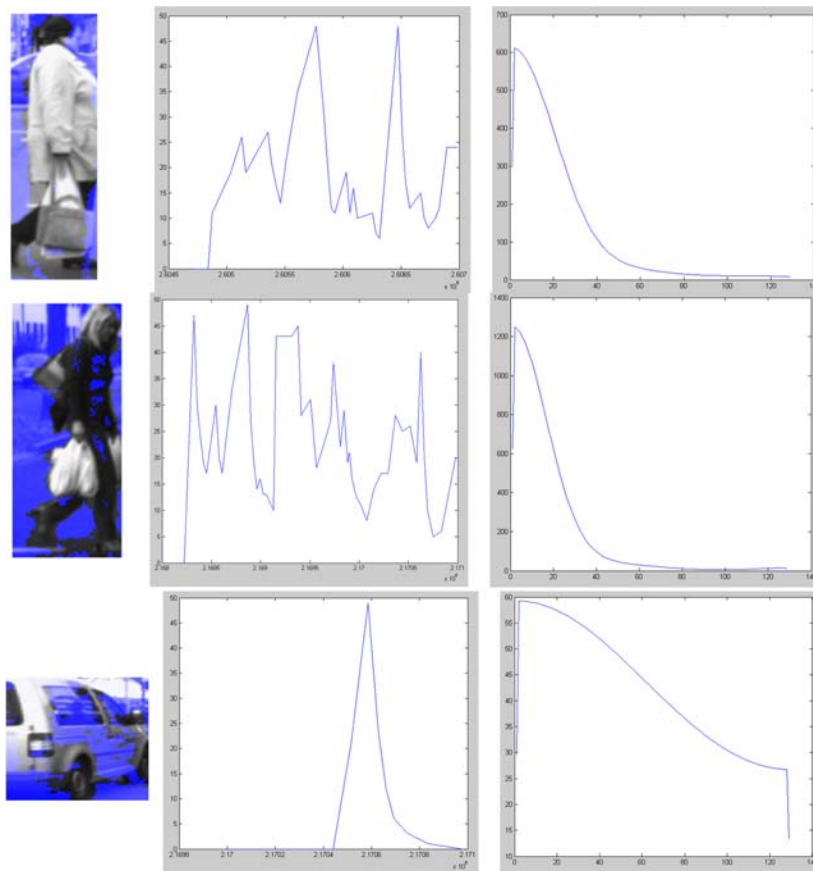


Figura 7.2: Stânga: obiecte, centru: variația semnăturii de mișcare în timp, dreapta: spectrul de frecvențe al semnăturii de mișcare

Pornind de la aceste considerații am introdus a doua trăsătură originală pentru recunoașterea pietonilor în mișcare, periodicitatea semnăturii de mișcare. Pentru a distinge între mișcări periodice și aperiodice determinăm frecvența de tăiere a spectrului semnăturii de mișcare. Dacă frecvența de tăiere este mare, înseamnă că spectrul reprezintă un zgomot alb (întins pe toate frecvențele). Dacă frecvența de tăiere este mică, înseamnă că spectrul reprezintă doar câteva frecvențe bine definite, reprezentând mișcările diferitelor părți ale corpului pietonului în timpul mersului.

Tabela 7.1: Rezultate experimentale ale trăsăturilor bazate pe mișcare

Categorie	Număr	Procent
Obiecte	600	100%
Pietoni	243	40.5%
Alte obiecte	357	59.5%
Pietoni falși	49	8.1%
Pietoni nedetecțați	39	6.5%
Detecție corectă	512	85.3%

7.2.3 Rezultate experimentale

Am evaluat performanța celor două trăsături bazate pe mișcare pe un set de 600 de obiecte de test. Tabelul 7.1 conține rezultatele obținute:

Trăsăturile bazate pe mișcare sunt calculate în timp real (peste 25fps pe o arhitectură Intel Core 2 Duo 2.6GHz). Semnătura de mișcare este calculată începând cu primul cadru în care obiectul este urmărit. Periodicitatea semnăturii de mișcare necesită mai multe cadre în care obiectul este urmărit pentru a putea fi calculată.

7.3 Trăsături simple

Pentru a îmbunătăți rata de recunoaștere a pietonilor obținută folosind doar trăsăturile de mișcare și pentru a trata cazul pietonilor staționari am introdus câteva trăsături simple, ușor de calculat. Trăsăturile se calculează folosind caracteristicile cuboidului asociat obiectului considerat și viteza acestuia. Aceste trăsături sunt: înălțimea obiectului, raza bazei, viteza longitudinală și viteza laterală (orientarea vitezelor este față de vehiculul propriu). Figura 7.3 arată aceste trăsături și logaritmul verosimilităților lor pentru pietoni. Calculul acestor trăsături simple nu afectează timpul de execuție al algoritmului.

7.4 Clasificatorul Bayesian

Având mai multe trăsături, devine necesară introducerea unui clasificator care să le agregheze.

Având clasa C_1 și trăsătura F_1 , avem:

$$P(C_1|F_1) = \frac{P(C_1)P(F_1|C_1)}{P(F_1)} \quad (7.3)$$

În această formulă $P(C_1)$ reprezintă probabilitatea a priori a clasei C_1 , $P(F_1|C_1)$ este o probabilitate care se determină prin antrenare, iar $P(F_1)$ este probabilitatea apariției trăsăturii. Verosimilitatea este definită ca:

$$L(C_1|F_1) = \frac{P(C_1|F_1)}{P(\neg C_1|F_1)} \quad (7.4)$$

Dacă este mai mare ca 1, atunci obiectul este clasificat ca aparținând clasei C_1 . Combinând formula lui Bayes cu logaritmul verosimilității, obținem:

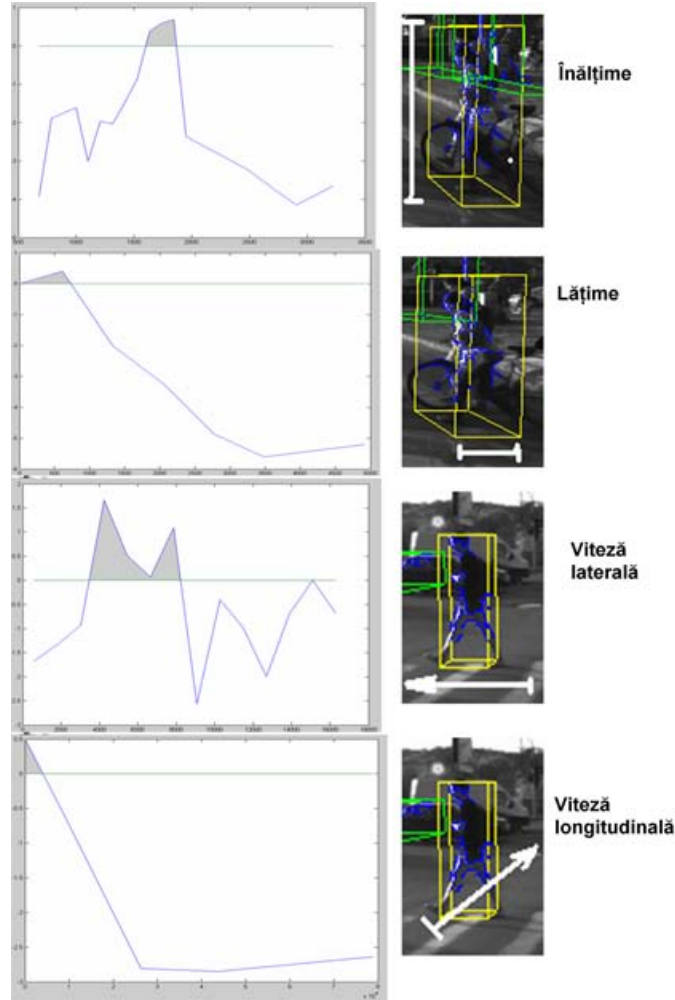


Figura 7.3: Trăsături simple. Pe axa Oy este reprezentat logaritmul verosimilității $\left(\ln \frac{P_{\text{pedon}}}{P_{\text{non-pedon}}}\right)$ pedonilor condiționat de valoarea trăsăturii pe axa Ox (mm pentru dimensiuni, mm/s pentru viteze).

$$\ln(L(C_1|F_1)) = \ln(P(C_1)P(F_1|C_1)) - \ln(P(C_1)P(F_1|C_1)) \quad (7.5)$$

Pentru mai multe trăsături:

$$P(C_1|F_1, F_2, \dots, F_n) = \frac{P(C_1)P(F_1, F_2, \dots, F_n|C_1)}{P(F_1, F_2, \dots, F_n)} \quad (7.6)$$

Făcând presupunerea că trăsăturile sunt independente:

$$P(C_1|F_1, \dots, F_n) = \frac{P(C_1)P(F_1|C_1) \dots P(F_n|C_1)}{P(F_1) \dots P(F_n)} \quad (7.7)$$

și deci:

$$L(C_1|F_1, \dots, F_n) = \frac{P(C_1)P(F_1|C_1) \dots P(F_n|C_1)}{P(\neg C_1)P(F_1|\neg C_1) \dots P(F_n|\neg C_1)} \quad (7.8)$$

Trecând la logaritmi:

$$\ln(L(C_1|F_1, \dots, F_n)) = A + B_1 + \dots + B_n \quad (7.9)$$

unde:

$$A = \ln \frac{P(C_1)}{P(\neg C_1)} \quad (7.10)$$

și:

$$B_i = \ln \frac{P(F_i|C_i)}{P(F_i|\neg C_i)} \quad (7.11)$$

Pentru a antrena clasificatorul Bayesian am implementat un modul de etichetare manuală, folosind și informațiile de identitate a obiectelor obținute din algoritmul de urmărire prezentat în capitolul anterior. Trăsăturile sunt apoi extrase și stocate într-un fișier text. Un script parsează fișierul text, generând clasificatorul ca și cod C.

7.4.1 Rezultate experimentale

Pentru a testa performanțele clasificatorilor am folosit o secvență cu 2600 de cadre etichetate manual.

7.5 Arhitectura originală de clasificare

Pentru a putea trata cazul unor trăsături multiple, neindependente, mai multe clase de obiecte și mai mulți algoritmi de clasificare, am propus o arhitectură originală de clasificare:

7.6 Baza de date pentru antrenarea și evaluarea clasificării

Am constatat că pentru antrenarea și evaluarea clasificării sunt necesare volume mari de date. Totodată, detectorii de obiecte, extractorii de trăsături și din acest motiv am propus arhitectura originală a unei baze de date relaționale care să conțină:

- Secvențele de imagini etichetate manual;
- Clasele obiectelor etichetate;

Tabela 7.2: Rezultate experimentale

Scenariu		Rezultate		
Scenariu de complexitate scăzută 300 cadre		Nr. obiecte	720	100%
		Pietoni	610	85.7%
		Alte obiecte	110	15.3%
		Pietoni falși	0	0%
		Pietoni nedetecțați	0	0%
		Detecție incorectă	0	0%
		Detecție corectă	720	100%
Scenariu de complexitate scăzută 800 cadre		Total obiecte	1003	100%
		Pietoni	721	71.9%
		Alte obiecte	282	28.1%
		Pietoni falși	23	2.3%
		Pietoni nedetecțați	20	2%
		Detecție incorectă	43	4.3%
		Detecție corectă	960	95.7%
Scenariu de complexitate crescută 1500 cadre		Total obiecte	5923	100%
		Pietoni	1231	21.8%
		Alte obiecte	3692	79.2%
		Pietoni falși	411	6.9%
		Pietoni nedetecțați	392	6.6%
		Detecție incorectă	803	13.6%
		Detecție corectă	5120	86.4%

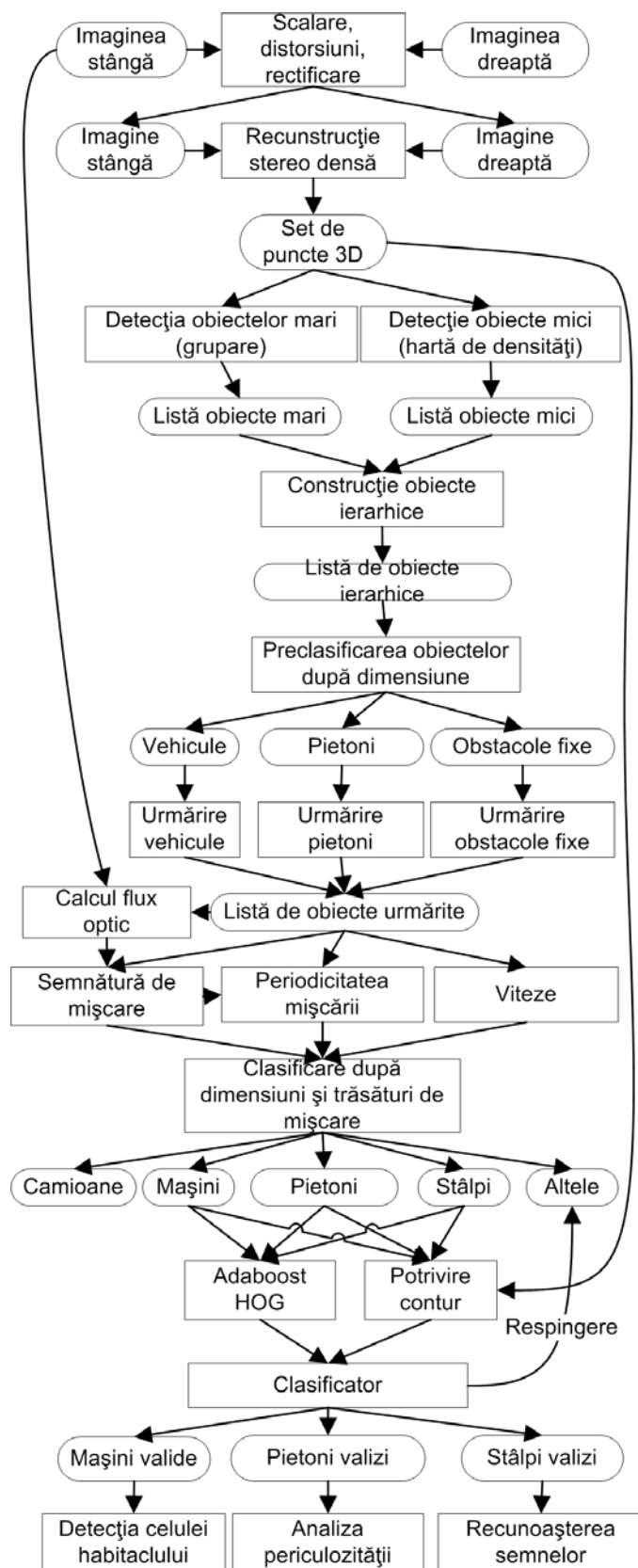


Figura 7.4: Arhitectura clasificatorului

- Trăsăturile extrase pentru fiecare obiect;
- Localizarea în spațiu a obiectelor etichetate;
- Contururile obiectelor.

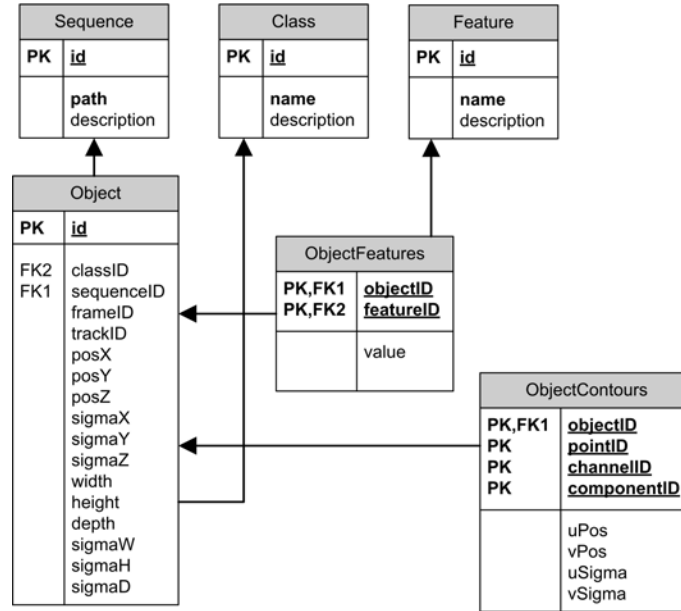


Figura 7.5: Schema bazei de date pentru antrenarea și evaluarea clasificării

Am reușit să etichetăm manual 107719. Pe aceste instanțe am obținut următoarele rezultate pentru clasificare:

Tabela 7.3: Rezultatele clasificării

Instanțe clasificate corect	94955	88.15%
Instanțe clasificate incorect	12764	11.85%

7.7 Concluzii

În acest capitol am prezentat contribuțiile noastre originale în domeniul recunoașterii obiectelor în medii de trafic urban. În secțiunea 7.2 am descris două trăsături originale pentru recunoașterea

Tabela 7.4: Rezultatele clasificării

Clasă	Rata TP	Rată FP	Precizie	Aria ROC
Camioane	0.894	0.063	0.879	0.977
Mașini	0.877	0.071	0.856	0.968
Pietoni	0.909	0.048	0.869	0.974
Stâlpi	0.814	0.01	0.844	0.978
Altele	0.706	0.096	0.797	0.895

pietonilor, bazate pe mișcare: semnătura de mișcare și periodicitatea mișcării. Aceste trăsături se bazează pe analiza câmpului de mișcare 3D și pe faptul că pietonii au un model de mișcare articulat. Rezultatele prezentate în această secțiune au fost publicate în [121]. Pentru îmbunătățirea rezultatelor și pentru recunoașterea pietonilor staționari am introdus în secțiunea 7.3 un număr de trăsături simple și ușor de calculat aplicabile tuturor obiectelor. Pentru agregarea acestor trăsături multiple, am prezentat în secțiunea 7.4 o primă încercare, bazată pe un clasificator Bayesian naiv. Acest clasificator este util atunci când datele de antrenare sunt puține și când trăsăturile sunt independente. Rezultatele obținute au fost prezentate într-un număr de articole [122][118][123][124] și [125]. Pentru folosirea mai multor trăsături, mai complexe și al unui număr mai mare de date am trecut la un clasificator mai puternic, a cărui arhitectură originală este prezentată în secțiunea 7.5. Acest clasificator poate trata mai multe clase de obiecte, mai multe trăsături și poate folosi mai mulți algoritmi de clasificare. Rezultatele obținute au fost publicate în [126]. Pentru a avea suficiente date pentru antrenarea și testarea clasificatorilor, am introdus arhitectura originală a unei baze de date relaționale, prezentate în secțiunea 7.6. Aceasta conține informații spațiale pentru localizarea în spațiu a obiectelor, trăsături și clase etichetate manual. Baza de date permite evoluția detectorilor de obiecte, a algoritmilor de extragere de trăsături și a algoritmilor de clasificare, fără a fi necesară reetichetarea manuală a obiectelor (care este foarte dificilă). Etichetarea manuală a unui număr mare de obiecte a permis îmbunătățirea rezultatelor de clasificare.

Partea III

Concluzii

Capitolul 8

Concluzii

Obiectivul principal al acestei teze a fost detecția mișcării din imagini tridimensionale, urmărirea acestei mișcări și folosirea ei pentru a extrage informații utile. Majoritatea algoritmilor existenți care utilizează mișcarea nu folosesc în mod consistent informații 3D. Chiar și acei algoritmi care folosesc informații 3D, le folosesc doar pentru anumiți pași. Această teză a investigat întreg potențialul folosirii informațiilor 3D în detecția și în urmărirea mișcării.

Primul obiectiv al acestei teze a fost constituit de determinarea mișcării camerelor. Determinarea mișcării camerelor este utilă în multe domenii, cum ar fi adnotarea automată a secvențelor video sau înlocuirea senzorilor inerțiali și de odometrie ai vehiculelor cu senzori bazați pe viziune artificială.

Primul pas în atingerea acestui obiectiv a fost studiul bibliografic (prezentat în capitolul 2) al celor mai relevante tehnici pentru calcularea fluxului optic, (atât 2D cât și 3D), ai algoritmilor de găsierea a colțurilor pentru calcularea fluxului optic precum și a algoritmilor de determinare a mișcării camerelor. În secțiunea 2.1 am prezentat două exemple simple de extragere de informații utile din mișcare, în secțiunea 2.2 am prezentat fundamentele teoretice ale fluxului optic, în secțiunea 2.3 am descris algoritmul Lucas–Kanade, în secțiunea 2.4 am descris algoritmul Horn–Shunk, în secțiunea 2.5 am descris un algoritm pentru calcul fluxului optic prin metode piramidale descris de Bouguet, în secțiunea 2.6 am descris algoritmi utilizați pentru detecția colțurilor, în secțiunea 2.7 am descris o extensie în 3D a ecuației de intensitate luminoasă constantă, în secțiunea 2.8 am descris algoritmul de flux optic bazat pe potrivirea blocurilor, în secțiunea 2.9 am descris algoritmi de flux optic bazați pe corelarea fazei, iar în secțiunea 2.10 am prezentat câteva abordări existente pentru determinarea mișcării camerelor.

Studiul acestor algoritmi precum și implementarea unora dintre ei ne-a permis să tragem următoarele concluzii:

- Cei mai utili algoritmi de extragere a colțurilor pentru calcularea fluxului optic sunt algoritmi Harris și Shi–Tomasi;
- Algoritmi de detecție a colțurilor se limitează doar la imagini 2D și nu folosesc informații 3D care pot fi furnizate de stereoviziune;

- Detectia colturilor pate sa consume mai mult timp decat calculul fluxului optic bazat pe trasaturi, deoarece detectia colturilor trebuie aplicata pe intreaga imagine, pe cand calculul fluxului optic se aplica doar unui numar relativ mic de colturi. Din aceasta cauza, pentru a imbunatati timpul de procesare al algoritmilor de flux optic, este necesara optimizarea algoritmilor de detectie a colturilor;
- Cel mai util algoritm existent pentru calculul fluxului optic bazat pe trasaturi este algoritmul de tip piramidal Lucas-Kanade prezentat de Bouguet. Acesta are o precizie si o acuratețe sporita si este capabil sa urmareasca deplasamente mari pentru colturi;
- Chiar daca algoritmii de flux optic bazati pe potrivirea de blocuri sunt usor de implementat si de optimizat, ei nu ofera suficienta acuratețe spațială.
- Algoritmii de flux optic bazati pe corelare de faza necesita o putere de calcul mare si sunt mai utili pentru potrivirea imaginilor decat pentru calculul fluxului optic.
- Algoritmii existenti pentru determinarea miscării camerelor folosesc o parametrizare relativ generica, care desi poate sa trateze foarte multe tipuri de miscare este mai puțin utilă în scenarii cu mai puține grade de libertate.

Pe baza concluziilor de mai sus, am descris în capitolul 5 contribuțiile noastre legate de îmbunătățirea algoritmilor existenți precum și algoritmi originali pentru calculul fluxului optic dens, detectia colturilor și estimarea mișcării camerelor. În secțiunea 5.1 am prezentat un algoritm pentru calculul fluxului optic dens. Originalitatea acestui algoritm constă în optimizări care fac algoritmul mai ușor de adaptat la o implementare de tip SIMD. Ca o aplicație a acestui algoritm am prezentat în secțiunea 5.2 un algoritm original de segmentare a imaginilor bazat pe mișcare, informații 3D și intensitatea imaginii. În secțiunea 5.3 am prezentat un algoritm original de detectie a colturilor care este mai rapid decât algoritmii de tip Harris [5] sau Shi-Tomasi [7]. Pentru a obține îmbunătățirea timpului de execuție, am folosit muchiile Canny deja extrase, informații 3D și limitarea detectiei colturilor la părți relevante din imagine. În funcție de acuratețea dorită, algoritmul este de 3–8 ori mai rapid decât Shi-Tomasi, și este la fel de bun pentru extragerea colturilor pentru fluxul optic. În secțiunea 5.4 am descris modelul camerei folosit în algoritmii de detectie a mișcării camerei. În secțiunea 5.5 am descris o parametrizare originală în coordonate polare a camerei, care împreună cu un algoritm de tip RANSAC s-au utilizat la detectia mișcării camerei mono necalibrate. Am testat sistemul nostru folosind circa 3 ore de film în format digital. Acuratețea de detectie a începutului unei noi secvențe a fost de 91%, iar acuratețea estimării calitative a mișcării camerelor a fost de 88%. În secțiunea 5.6 am dezvoltat algoritmul anterior, profitând de informația de calibrare a camerei și am dezvoltat un algoritm original care este capabil să măsoare mișcarea de rotație a camerei și, cu anumite informații în plus și mișcarea de translație. Algoritmul este util pentru sistemele de asistență a conducerii autovehiculelor. Aspectul original al acestui algoritm constă în folosirea coordonatelor polare și în eliminarea mișcării de translație considerând că ea ar fi un efect de zoom. În secțiunea 5.7 am descris un algoritm original pentru estimarea mișcării camerelor stereo care folosește detectorul de colturi propriu, descris în secțiunea 5.3. Originalitatea acestui algoritm constă în alegerea punctelor pentru calcularea mișcării camerelor, și determinarea separată a translației printr-o metodă stabilă. Rezultatele prezentate în acest capitol au fost publicate într-un articol [118].

Al doilea obiectiv al tezei a fost dezvoltarea unui algoritm pentru urmărirea obiectelor în medii complexe și dificile de trafic urban, urmărirea reprezentând o cerință majoră a sistemelor de asistență a conducerii.

În primul rând am studiat literatura existentă, prezentând în capitolul 3 cele mai relevante contribuții din domeniu precum și cei doi algoritmi de filtrare larg utilizați, filtrul Kalman și filtrul de particule. În secțiunea 3.1 am descris algoritmi de urmărire existenți care folosesc camere montate pe autovehicule pentru urmărire, în secțiunea 3.2 algoritmi care folosesc camere fixe, în secțiunea 3.3 am prezentat filtrul Kalman iar în secțiunea 3.4 filtrul de particule.

Studiul literaturii din domeniu a permis formularea următoarelor concluzii:

- Majoritatea abordărilor existente se concentrează pe urmărirea autovehiculelor individuale sau ușor separabile;
- Algoritmi de urmărire folosiți sunt influențați de algoritmi de detecție utilizați și de modelul obiectelor;
- Majoritatea algoritmilor care folosesc camere fixe se bazează în mare măsură doar pe asumția unui fundal fix, obținând rezultate destul de bune, dar făcând acești algoritmi inutili pentru sistemele de asistență a conducerii;
- Cea mai dificilă problemă în urmărirea obiectelor multiple o reprezintă asocierea, nu filtrarea efectivă.
- Filtrele de tip Kalman sunt rapide, precise, și de aceea sunt preferabile din punct de vedere computațional. Totuși ele nu pot modela procese cu zgomot de tip non-Gaussian.
- Filtrele Kalman lucrează recursiv, integrând la fiecare pas informațiile de stare prezise cu cele observate. Ele pot integra și observații multiple, dar, din cauza pierderilor de precizie, aceasta poate duce la acumularea erorilor
- Filtrele de particule sunt mai costisitoare din punct de vedere computațional, dar ele pot capta procese mai complexe. Totuși acestea sunt mai dificil de controlat din cauza comportamentului lor aleator.

Pornind de la aceste concluzii, am descris în capitolul 6 un algoritm de urmărire original care rezolvă problema urmăririi obiectelor multiple, ierarhice, în medii de trafic urban. Originalitatea algoritmului dezvoltat constă în următoarele aspecte:

- Viteza obiectelor a fost calculată folosind fluxul optic. O combinație de filtre mediane și de medie a fost folosită pentru agregarea vectorilor de flux optic, eliminând erorile majore.
- Obiectele au fost reprezentate printr-un model probabilistic original avansat. Modelul probabilistic ia în considerare caracteristicile algoritmilor de detecție a obiectelor bazați pe stereoviziune.
- Modelul probabilistic al obiectelor conduce la o densitate de probabilitate a obiectului în spațiul 3D. Am aproximat această densitate prin două metode, prima folosind un set de cuboide imbricate, iar a doua printr-o metodă de tip Monte-Carlo, folosind un set de poligoane generate aleator;
- Faza de asociere este implementată folosind un algoritm original, care este capabil să trateze obiecte reprezentate ierarhic, precum și cazul obiectelor divizate sau unite în mod eronat;
- Partea de filtrare Kalman efectivă agregă 3 estimări pentru viteze, viteza calculată din flux optic, viteza prezisă precum și o viteză calculată implicit din faza de asociere a obiectelor.

- S-au făcut numeroase optimizări pentru a face ca algoritmul să funcționeze în timp real.

În secțiunea 6.1 am prezentat fundamentele matematice ale agregării de date zgomotoase din mai multe surse, în secțiunea 6.2 am prezentat arhitectura sistemului de stereoviziune pentru detecția și urmărirea obiectelor din trafic, în secțiunea 6.3 am identificat sursele erorilor care trebuie filtrate de algoritmul de urmărire, și, bazat pe ele am introdus în secțiunea 6.4 un model original, avansat, probabilistic static și dinamic pentru reprezentarea obiectelor. În secțiunea 6.5 am descris faza de predicție a algoritmului de urmărire. În secțiunea 6.6 am prezentat un algoritm original de asociere a obiectelor, pentru obiecte reprezentate ierarhic. În secțiunea 6.7 am prezentat arhitectura filtrelor Kalman utilizate. În secțiunea 6.8 am prezentat rezultatele experimentale obținute, care arată că algoritmul de urmărire este capabil să urmărească majoritatea obiectelor din scenă, pentru un timp suficient de lung, cu un timp de inițializare de un singur cadru, îmbunătățind astfel performanța detecției obiectelor. Algoritmul are un timp de rulare mediu de doar 1.9 ms pe cadru. Rezultatele din capitolul 6 au fost publicate în două articole, [119] și [120].

Al treilea obiectiv al tezei a constat în utilizarea mișcării, precum și a altor trăsături pentru recunoașterea pietonilor și a altor clase de obiecte din mediul de trafic urban. Aceasta reprezintă o altă cerință majoră a sistemelor avansate de asistență a conducerii în medii de trafic urban. Am început printr-un studiu al literaturii din domeniu prezentat în capitolul 4 în care am prezentat tehnicile actuale de clasificare a obiectelor din trafic, concentrându-ne în principal pe recunoașterea pietonilor. În secțiunea 4.1 am prezentat necesitatea clasificării obiectelor, în secțiunile 4.2, 4.3, 4.4, 4.6 am prezentat contribuțiile aduse de Papageorgiou, Gavrila, Broggi și alți autori în acest domeniu. Studiul lucrărilor publicate în acest domeniu ne-a permis formularea următoarelor concluzii:

- Există multiple metode de detecție a obiectelor și de formulare a ipotezelor pentru pietoni. Unele se bazează pe informații 2D (de exemplu pe densitatea de muchii verticale), altele metode se bazează pe imagini în domeniul infraroșu îndepărtat, altele pe informații date de stereoviziune.
- Majoritatea metodelor bazate pe stereoviziune se folosesc doar de imagini de disparitate, folosind sau metode de segmentare sau metoda de v-disparitate.
- Abordările bazate pe v-disparitate sunt mai rapide, dar nu sunt precise pentru medii complexe.
- Dezavantajul sistemelor de recunoaștere bazate doar pe informații 2D este faptul că nu au informații despre scara obiectelor, trebuind să caute modele în întreaga imagine la diverse scale.
- Informația 3D generată de sistemele stereo generează adâncimi pentru anumite puncte din imagine. Există câteva metode de clasificare bazate exclusiv pe adâncime [111], [112]. Erorile și lipsa reconstrucției stereo fac dificilă folosirea directă a informației de adâncime.
- O trăsătură importantă pentru recunoașterea pietonilor în mișcare este modelul lor de mers. Există mai multe lucrări (de exemplu [113], [103], [114]) care folosesc trăsături de mișcare pentru recunoașterea pietonilor. A abordare tipică constă în determinarea faptului că mișcarea este periodică. Din păcate diferențierea mișcării obiectelor față de mișcarea fundalului este dificilă din cauza complexității mediului de trafic urban.
- Informațiile 3D nu sunt folosite nici pentru mișcare în algoritmii descriși în literatură.
- Există puține articole legate de recunoașterea altor obiecte în afară de pietoni.

- Antrenarea și testarea algoritmilor necesită seturi de date extinse. Majoritatea datelor disponibile sunt doar imagini individuale, reprezentând exemple pozitive și negative. Din această cauză ele nu pot fi folosite pentru sisteme de clasificare bazate pe informații 3D și de mișcare.

În urma acestor concluzii, am prezentat în capitolul 7 contribuțiile noastre originale în domeniul recunoașterii obiectelor în medii de trafic urban. În secțiunea 7.2 am descris două trăsături originale pentru recunoașterea pietonilor, bazate pe mișcare: semnătura de mișcare și periodicitatea mișcării. Aceste trăsături se bazează pe analiza câmpului de mișcare 3D și pe faptul că pietonii au un model de mișcare articulat. Rezultatele prezentate în această secțiune au fost publicate în [121]. Pentru îmbunătățirea rezultatelor și pentru recunoașterea pietonilor staționari am introdus în secțiunea 7.3 un număr de trăsături simple și ușor de calculat aplicabile tuturor obiectelor. Pentru agregarea acestor trăsături multiple, am prezentat în secțiunea 7.4 o primă încercare, bazată pe un clasificator Bayesian naiv. Acest clasificator este util atunci când datele de antrenare sunt puține și când trăsăturile sunt independente. Rezultatele obținute au fost prezentate într-un număr de articole [122][118][123][124] și [125]. Pentru folosirea mai multor trăsături, mai complexe și al unui număr mai mare de date am trecut la un clasificator mai puternic, a cărui arhitectură originală este prezentată în secțiunea 7.5. Acest clasificator poate trata mai multe clase de obiecte, mai multe trăsături și poate folosi mai mulți algoritmi de clasificare. Rezultatele obținute au fost publicate în [126]. Pentru a avea suficiente date pentru antrenarea și testarea clasificatorilor, am introdus arhitectura originală a unei baze de date relaționale, prezentate în secțiunea 7.6. Aceasta conține informații spațiale pentru localizarea în spațiu a obiectelor, trăsături și clase etichetate manual. Baza de date permite evoluția detectorilor de obiecte, a algoritmilor de extragere de trăsături și a algoritmilor de clasificare, fără a fi necesară reetichetarea manuală a obiectelor (care este foarte dificilă). Etichetarea manuală a unui număr mare de obiecte a permis îmbunătățirea rezultatelor de clasificare.

Obiectivele principale ale acestei teze au fost îndeplinite. Impactul științific se regăsește într-un număr mare de articole publicate, în jurnale ISI cu factor de impact, volumele unor conferințe indexate ISI, BDI, IEEEExplore și într-un număr de citări independente. Acestea vor fi prezentate în anexele următoare.

Bibliografie

- [1] E. Trucco and A. Verri, *Introductory Techniques for 3-D Computer Vision*. Upper Saddle River, NJ, USA: Prentice Hall PTR, 1998.
- [2] B. K. Horn and B. G. Schunck, “Determining optical flow,” *Artificial Intelligence*, vol. 17, pp. 185–203, 1981.
- [3] A. Mitiche and A.-R. Mansouri, “On convergence of the horn and schunck optical-flow estimation method,” *IEEE Transactions on Image Processing*, vol. 13, no. 6, June 2004. [Online]. Available: http://externe.emt.inrs.ca/users/mitiche/IP/_Horn-and-Shunck.pdf
- [4] J.-Y. Bouguet. (2000) Pyramidal implementation of the Lucas Kanade feature tracker. [Online]. Available: http://mrl.nyu.edu/~bregler/classes/vision/_spring06/bouget00.pdf

- [5] C. Harris and M. Stephens, "A combined corner and edge detector," in *Proceedings of the 4th Alvey Vision Conference*, 1988, pp. 147–151.
- [6] H. Moravec, "Obstacle avoidance and navigation in the real world by a seeing robot rover, tech report cmu-ri-tr-3," Carnegie-Mellon University, Robotics Institute, Tech. Rep., 1980.
- [7] J. Shi and C. Tomasi, "Good features to track," in *IEEE Conference on Computer Vision and Pattern Recognition*, 1994, pp. 593–600.
- [8] L. Kitchen and A. Rosenfeld, "Gray-level corner detection," *Pattern Recognition Letters*, 1982.
- [9] J. J. Koenderink and W. Richards, "Two-dimensional curvature operators," *Journal of the Optical Society of America: Series A* 5, no. 7, pp. 1136–1141, 1988.
- [10] T. Lindeberg, "Feature detection with automatic scale selection," *International Journal of Computer Vision*, vol. 2, no. 30, pp. 77–116, 1998.
- [11] K. Mikolajczyk, K. Schmid, and C. Schmid, "Scale and affine invariant interest point detectors," *International Journal of Computer Vision*, vol. 1, no. 60, pp. 63–86, 2004.
- [12] D. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 2, no. 60, 2004.
- [13] S. M. Smith and J. M. Brady, "Susan - a new approach to low level image processing," *International Journal of Computer Vision*, vol. 1, no. 23, pp. 45–78, 1997.
- [14] L. Trujillo and G. Olague, "Automated design of image operators that detect interest points," *Evolutionary Computation*, vol. 4, no. 16, pp. 483–507, 2008.
- [15] M. Harville, A. Rahimi, T. Darrell, G. Gordon, and J. Woodfill, "3d pose tracking with linear depth and brightness constraints," in *Proceedings of the IEEE International Conference on Computer Vision*, 1999. [Online]. Available: <http://www.tyxx.com/pubs/ICCV99.pdf>
- [16] E. de Castro and C. Morandi, "Registration of translated and rotated images using finite fourier transforms," *IEEE Transactions on pattern analysis and machine intelligence*, September, 1987.
- [17] B. S. Reddy and B. N. Chatterji, "An fft-based technique for translation, rotation, and scale-invariant image registration," *IEEE Transactions on Image Processing*, no. 8, pp. 1266–1271, 1996.
- [18] L.-Y. Duan, J. Wang, Y.-T. Zheng, C. Xu, Q. Tian, J. S. Jin, and H. Lu, "Shot-level camera motion estimation based on a parametric model," in *TRECVID*, 2005.
- [19] A. Mallet, S. Lacroix, and L. Gallo, "Position estimation in outdoor environments using pixel tracking and stereovision," in *Proceedings of IEEE International Conference of Robotics and Automation (ICRA)*, vol. 4, 2000, pp. 3519–3524.
- [20] H. Hirschmuller, P. R. Innocent, and J. M. Garibaldi, "Fast, unconstrained camera motion estimation from stereo without tracking and robust statistics," in *Proceedings of the 7th International Control, Automation, Robotics and Vision Conference (ICARCV)*, vol. 2, December 2002, pp. 1099–1104.

- [21] N. Molton and M. Brady, “Practical structure and motion from stereo when motion is unconstrained,” *International Journal of Computer Vision*, vol. 39, pp. 5–23, August 2000.
- [22] P. Saeedi, P. Lawrence, and D. Lowe, “3d motion tracking of a mobile robot in a natural environment,” in *Proceedings of the IEEE International Conference on Robotics and Automation*, vol. 2, April 2000, pp. 1682–1687.
- [23] R. Danescu, S. Nedevschi, M. M. Meinecke, and T. Graf, “Stereovision based vehicle tracking in urban traffic environments,” in *Proceedings of the IEEE Intelligent Transportation Systems Conference (ITSC 2007), Seattle USA*, vol. 2, September 2007, pp. 400–404.
- [24] B. Barrois, S. Hristova, C. Woehler, F. Kummert, and C. Hermes, “3d pose estimation of vehicles using a stereo camera,” in *IEEE Intelligent Vehicles Symposium*, 2009.
- [25] A. Barth and U. Franke, “Tracking oncoming and turning vehicles at intersections,” in *Intelligent Transportation Systems, IEEE Conference on*, Madeira Island, Portugal, 2010, pp. 861–868.
- [26] M. Isard and A. Blake, “Condensation – conditional density propagation for visual tracking,” *International Journal of Computer Vision*, vol. 29, no. 1, pp. 5–28, 1998.
- [27] R. Danescu, F. Oniga, and S. Nedevschi, “Particle grid tracking system for stereovision based environment perception,” in *Proc. of the IEEE Intelligent Vehicles Symposium (IEEE-IV 2010), San Diego, USA*, June 2010, pp. 987–992.
- [28] C. Coue, C. Pradalier, C. Laugier, T. Fraichard, and P. Bessiere, “Bayesian occupancy filtering for multitarget tracking: An automotive application,” *The International Journal of Robotics Research*, vol. 25, no. 1, 2006.
- [29] C. Chen, C. Tay, K. Mekhnacha, and C. Laugier, “Dynamic environment modeling with grid-map: a multiple-object tracking application,” in *Proceedings of International Conference on Automation, Robotics and Computer Vision (ICARCV)*, December 2006, pp. 1–6.
- [30] H. Badino, U. Franke, and R. Mester, “Free space computation using stochastic occupancy grids and dynamic programming,” in *Workshop on Dynamical Vision, ICCV*, 2007.
- [31] T. Gindele, S. Brechtel, J. Schroeder, and R. Dillmann, “Bayesian occupancy grid filter for dynamic environments using prior map knowledge,” in *Proceedings of IEEE Intelligent Vehicles Symposium*, July 2009, pp. 669–676.
- [32] T. Dang, C. Hoffmann, and C. Stiller, “Fusing optical flow and stereo disparity for object tracking,” in *The 5th IEEE International Conference on Intelligent Transportation Systems*, 2002, pp. 112–117.
- [33] H. B. Uwe Franke, Clemens Rabe and S. Gehrig, “6d-vision: Fusion of stereo and motion for robust environment perception,” in *Pattern recognition – Lecture Notes in Computer Science*, vol. 3663, 2005, pp. 216–223.
- [34] A. Barth and U. Franke, “Estimating the driving state of oncoming vehicles from a moving platform using stereo vision,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 10, pp. 560–571, 2009.

- [35] —, “Where will the oncoming vehicle be the next second?” in *Proceedings of the IEEE Intelligent Vehicles Symposium*, 2008, pp. 1068–1073.
- [36] A. Barth, D. Pfeiffer, and U. Franke, “Vehicle tracking at urban intersections using dense stereo,” in *3rd Workshop on Behaviour Monitoring and Interpretation, BMI*, Ghent, Belgium, 11 2009, pp. 47–58.
- [37] A. Fossati, P. Schönmann, and P. Fua, “Real-time vehicle tracking for driving assistance,” *Machine Vision and Applications*, vol. 22, no. 2, pp. 439–448, 2010.
- [38] M. Manz, T. Luettel, F. von Hundelshausen, and H.-J. Wuensche, “Monocular model-based 3d vehicle tracking for autonomous vehicles in unstructured environment,” in *IEEE International Conference on Robotics and Automation (ICRA)*, 2010, pp. 2465–2471.
- [39] S. Sivaraman and M. M. Trivedi, “A general active-learning framework for on-road vehicle recognition and tracking,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 11, pp. 267–276, 2010.
- [40] Z. Kim and J. Malik, “Fast vehicle detection with probabilistic feature grouping and its application to vehicle tracking,” in *Proceedings Ninth IEEE International Conference on Computer Vision*, 2003.
- [41] D. Koller, J. Weber, and J. Malik, “Robust multiple car tracking with occlusion reasoning,” in *Computer Vision—ECCV*, 1994, pp. 189–196.
- [42] D. Roller, K. Daniilidis, and H. H. Nagel, “Model-based object tracking in monocular image sequences of road traffic scenes,” *International Journal of Computer Vision*, vol. 10, pp. 257–281, 1993.
- [43] J. Lou, T. Tan, W. Hu, H. Yang, and S. J. Maybank, “3-d model-based vehicle tracking,” *IEEE Transactions on Image Processing*, pp. 1561–1569, 2005.
- [44] N. K. Kanhere and S. T. Birchfield, “Real-time incremental segmentation and tracking of vehicles at low camera angles using stable features,” *IEEE Transactions on Intelligent Transportation Systems*, pp. 148–160, 2008.
- [45] Z. Zhang, K. Huang, T. Tan, and Y. Wang, “3d model based vehicle tracking using gradient based fitness evaluation under particle filter framework,” in *Proceedings of the 20th International Conference on Pattern Recognition*, 2010.
- [46] R. E. Kalman, “A new approach to linear filtering and prediction problems,” *Journal of Basic Engineering*, vol. 82, no. Series D, pp. 35–45, 1960.
- [47] C. Papageorgiou, T. Evgeniou, and T. Poggio, “A trainable pedestrian detection system,” in *Proceedings of Intelligent Vehicles*, 1998, pp. 241–246.
- [48] M. Oren, C. Papageorgiou, P. Shina, E. Osuna, and T. Poggio, “Pedestrian detection using wavelet templates,” in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1997, pp. 193–199.

- [49] C. Papageorgiou and T. Poggio, "A pattern classification approach to dynamical object detection," in *Proceedings of the 7th IEEE International Conference on Computer Vision*, 1999, pp. 1223–1228.
- [50] F. Miau, C. Papageorgiou, and L. Itti, "Neuromorphic algorithms for computer vision and attention," in *Proceedings of the 46th Annual International Symposium on Optical Science and Technology (SPIE)*, vol. 4479, 2001, pp. 12–23.
- [51] D. M. Gavrila, "Pedestrian detection from a moving vehicle," in *Proceedings of the European Conference on Computer Vision, Dublin*, 2000, pp. 37–49.
- [52] F. Girosi, T. Poggio, and B. Caprile, "Extensions of a theory of networks for approximation and learning: . . ." in *Advances in Neural information processings systems 3*. Morgan Kaufmann Publishers, 1990, pp. 597–603.
- [53] P. Viola and M. Jones, "Robust real-time object detection," in *International Journal of Computer Vision*, 2001.
- [54] M. Oberlaender, "Hyperpermutation networks - a discrete approach to machined perception," in *Third Weightless Neuronal Networks Workshop, University of York*, 1999.
- [55] S. Hezel, A. Kugel, R. Manner, and D. M. Gavrila, "Fpga-based template matching using distance transforms," in *Proceedings of the 10th Annual IEEE Symposium on Field-Programmable Custom Computing Machines*, 2002, pp. 89–97.
- [56] D. M. Gavrila, "A bayesian exemplar-based approach to hierarchical shape matching," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 8, August 2007.
- [57] D. M. Gavrila and J. Giebel, "Virtual sample generation for template-based shape matching," *Computer Vision and Pattern Recognition, IEEE Computer Society Conference on*, vol. 1, p. 676, 2001.
- [58] M. Mählich, M. Oberländer, O. Löhlein, D. Gavrila, and W. Ritter, "A multiple detector approach to low-resolution fir pedestrian recognition," in *Proceedings of IEEE Intelligent Vehicles Symposium 2005*, 2005, pp. 23–28.
- [59] S. Munder and D. M. Gavrila, "An experimental study on pedestrian classification," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 11, pp. 1863–1868, 2006.
- [60] D. M. Gavrila and S. Munder, "Multi-cue pedestrian detection and tracking from a moving vehicle," *International Journal of Computer Vision, Springer Verlag*, vol. 73, no. 1, pp. 41–59, June 2007.
- [61] M. Enzweiler, P. Kanter, and D. M. Gavrila, "Monocular pedestrian recognition using motion parallax," in *Proceedings of the IEEE Intelligent Vehicles Symposium*, 2008, pp. 792–797.
- [62] S. Munder, C. Schnörr, and D. M. Gavrila, "Pedestrian detection and tracking using a mixture of view-based shape-texture models," *IEEE Transactions on Intelligent Transportation Systems*, vol. 9, no. 2, pp. 333–343, 2008.

- [63] M. Hofmann and D. M. Gavrilu, “Single-frame 3d human pose recovery from multiple views,” in *Proceedings of the Deutsche Arbeitsgemeinschaft für Mustererkennung (DAGM) Jena*, 2009.
- [64] —, “Multi-view 3d human pose estimation in complex environment,” *International Journal of Computer Vision*, 2011.
- [65] M. Enzweiler and D. M. Gavrilu, “Monocular pedestrian detection: Survey and experiments,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 12, pp. 2179–2195, 2009.
- [66] M. Liem and D. M. Gavrilu, “Multi-person tracking with overlapping cameras in complex, dynamic environments,” in *Proceedings of the British Machine Vision Conference (BMVC), London*, 2009.
- [67] C. G. Keller, D. F. Llorca, and D. M. Gavrilu, “Dense stereo-based roi generation for pedestrian detection,” in *Proceedings of the Deutsche Arbeitsgemeinschaft für Mustererkennung (DAGM), Jena*, 2009.
- [68] M. Rohrbach, M. Enzweiler, and D. M. Gavrilu, “High-level fusion of depth and intensity for pedestrian classification,” in *Proceedings of the Deutsche Arbeitsgemeinschaft für Mustererkennung (DAGM), Jena*, 2009.
- [69] M. Enzweiler and D. M. Gavrilu, “Integrated pedestrian classification and orientation estimation,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, San Francisco (USA)*, 2010.
- [70] M. Enzweiler, A. Eigenstetter, B. Schiele, and D. M. Gavrilu, “Multi-cue pedestrian classification with partial occlusion handling,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, San Francisco (USA)*, 2010.
- [71] M. Enzweiler and D. M. Gavrilu, “A multi-level mixture-of-experts framework for pedestrian classification,” in *IEEE Transactions on Image Processing*, 2011.
- [72] M. Hofmann and D. M. Gavrilu, “3d human shape model adaptation by automatic frame selection and batch-mode optimization,” *Computer Vision Image Understanding*, 2011.
- [73] C. Keller, M. Enzweiler, C. Schnör, M. Rohrbach, D. F. Llorca, and D. M. Gavrilu, “The benefits of dense stereo for pedestrian detection,” *IEEE Transactions on Intelligent Transportation Systems*, 2011.
- [74] C. Keller, T. Dang, A. Joos, C. Rabe, H. Fritz, and D. M. Gavrilu, “Active pedestrian safety by automatic braking and evasive steering,” *IEEE Transactions on Intelligent Transportation Systems*, 2011.
- [75] C. Keller, M. Enzweiler, and D. M. Gavrilu, “A new benchmark for stereo-based pedestrian detection,” in *Proceedings of the IEEE Intelligent Vehicles Symposium, Baden-Baden*, 2011.
- [76] A. Broggi, M. Bertozzi, A. Fascioli, and M. Sechi, “Shape-based pedestrian detection,” in *Proceedings of the IEEE Intelligent Vehicles Symposium Dearbon (MI), USA*, October 2000.

- [77] M. Bertozzi, A. Broggi, R. Chapuis, F. Chausse, A. Fascioli, and A. Tibaldi, "Shape-based pedestrian detection and localization," in *Proceedings of the IEEE Intelligent Transportation Systems*, vol. 1, 2003, pp. 328–333.
- [78] M. Bertozzi, A. Broggi, P. Grisleri, T. Graf, and M. M. Meinecke, "Pedestrian detection in infrared images," in *IEEE Proceedings of the Intelligent Vehicles Symposium*, 2003, pp. 662–667.
- [79] M. Bertozzi, A. Broggi, M. Carletti, A. Fascioli, T. Graf, P. Grisleri, and M. Meinecke, "Ir pedestrian detection for advanced driver assistance systems," in *Pattern Recognition – Lecture Notes in Computer Science*, vol. 2781, 2003, pp. 582–590.
- [80] A. Broggi, A. Fascioli, I. Fedriga, A. Tibaldi, and M. D. Rose, "Stereo-based preprocessing for human shape localization in unstructured environments," in *IEEE Proceedings of the Intelligent Vehicles Symposium*, 2003, pp. 410–415.
- [81] M. Bertozzi, A. Broggi, A. Fascioli, T. Graf, and M. M. Meinecke, "Pedestrian detection for driver assistance using multiresolution infrared vision," *IEEE Transactions on Vehicular Technology*, vol. 53, pp. 1666–1678, 2004.
- [82] A. Broggi, A. Fascioli, M. Carletti, T. Graf, and M. M. Meinecke, "A multi-resolution approach for infrared vision-based pedestrian detection," in *IEEE Intelligent Vehicles Symposium*, June 2004, pp. 7–12.
- [83] M. Bertozzi, A. Broggi, A. Fascioli, A. Tibaldi, R. Chapuis, and F. Chausse, "Pedestrian localization and tracking system with kalman filtering," in *IEEE Intelligent Vehicles Symposium*, June 2004, pp. 584–589.
- [84] M. Bertozzi, A. Broggi, A. Fascioli, A. Tibaldi, and M. D. Rose, "A tool for vision based pedestrian detection performance evaluation," in *IEEE Intelligent Vehicles Symposium*, June 2004, pp. 784–789.
- [85] M. Bertozzi, E. Binelli, A. Broggi, and M. D. Rose, "Stereo vision-based approaches for pedestrian detection," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR) - Workshops*, June 2005, p. 16.
- [86] M. Bertozzi, A. Broggi, A. Lasagni, and M. D. Rose, "Infrared stereo vision-based pedestrian detection," in *Proceedings of the IEEE Intelligent Vehicles Symposium*, June 2005, pp. 24–29.
- [87] A. Broggi, A. Fascioli, P. Grileri, T. Graf, and M. M. Meinecke, "Model-based validation approaches and matching techniques for automotive vision based pedestrian detection," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR) - Workshops*, June 2005.
- [88] F. Suard, A. Rakotomamonjy, A. Bensrhair, and A. Broggi, "Pedestrian detection using infrared images and histograms of oriented gradients," in *Proceedings of the IEEE Intelligent Vehicles Symposium, Tokyo*, September 2006, pp. 206–212.
- [89] M. Bertozzi, A. Broggi, M. Felisa, G. Vezzoni, and M. D. Rose, "Pedestrian detection using infrared images and histograms of oriented gradients," in *Proceedings of the IEEE Intelligent Vehicles Symposium, Tokyo*, September 2006, pp. 231–236.

- [90] M. Bertozzi, A. Broggi, M. D. Rose, M. Felisa, A. Rakotomamonjy, and F. Suard, "A pedestrian detector using histograms of oriented gradients and a support vector machine classifier," in *Proceedings of the IEEE Intelligent Transportation Systems Conference (ITSC)*, 2007, pp. 143–148.
- [91] A. Broggi, R. I. Fedriga, A. Tagliati, T. Graf, and M. M. Meinecke, "Pedestrian detection on a moving vehicle: an investigation about near infra-red images," in *Proceedings of the IEEE Intelligent Vehicles Symposium, Tokyo*, September 2006, pp. 431–436.
- [92] M. Bertozzi, A. Broggi, C. Caraffi, M. D. Rose, M. Felisa, and G. Vezzoni, "Pedestrian detection by means of far-infrared stereo vision," *Computer Vision and Image Understanding*, vol. 106, no. 2-3, pp. 194 – 204, 2007, special issue on Advances in Vision Algorithms and Systems beyond the Visible Spectrum. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S1077314206001858>
- [93] M. Bertozzi, A. Broggi, S. Ghidoni, and M. M. Meinecke, "A night vision module for the detection of distant pedestrians," in *Proceedings of the IEEE Intelligent Vehicles Symposium*, 2007, pp. 25–30.
- [94] M. Bertozzi, A. Broggi, M. D. Rose, M. Felisa, A. Rakotomamonjy, and F. Suard, "A symmetry-based validator and refinement system for pedestrian detection in far infrared images," in *Proceedings of the IEEE Intelligent Transportation Systems Conference (ITSC)*, 2007, pp. 155–160.
- [95] S. G. Massimo Bertozzi, Alberto Broggi and M. D. Rose, "Pedestrian shape extraction by means of active contours," *Springer Tracts in Advanced Robotics*, vol. 42, 2008.
- [96] A. Broggi, P. Cerri, S. Ghidoni, P. Grisleri, and H. G. Jung, "A new approach to urban pedestrian detection for automatic braking," *IEEE Transactions on Intelligent Transportation Systems*, vol. 10, pp. 594–605, 2009.
- [97] —, "Localization and analysis of critical areas in urban scenarios," in *Proceedings of the IEEE Intelligent Vehicles Symposium, Eindhoven, The Netherlands*, June 2008, pp. 1074–1079.
- [98] T. Ghandi and M. M. Trivedi, "Vehicle mounted wide fov stereo for traffic and pedestrian detection," in *Proceedings of the IEEE International Conference on Image Processing (ICIP)*, 2005.
- [99] —, "Pedestrian collision avoidance systems: a survey of computer vision based recent studies," in *Proceedings of the IEEE Intelligent Transportation Systems Conference (ITSC) (ICIP)*, 2006, pp. 976–981.
- [100] —, "Pedestrian protection systems: Issues, survey, and challenges," *IEEE Transactions on Intelligent Transportation Systems*, vol. 8, pp. 413–430, 2009.
- [101] —, "Image based estimation of pedestrian orientation for improving path prediction," in *Proceedings of the IEEE Intelligent Vehicles Symposium*, 2008, pp. 506–511.
- [102] L. Zhao and C. Thorpe, "Stereo and neural network-based pedestrian detection," *IEEE Transactions on Intelligent Transportation Systems*, vol. 1, no. 3, pp. 148–154, September 2000.

- [103] A. Shashua, Y. Gdalyahu, and G. Hayun, "Pedestrian detection for driving assistance systems: Single-frame classification and system level performance," in *IEEE Intelligent Vehicle Symposium*, June 2004, pp. 1–6.
- [104] H. Nanda and L. Davis, "Probabilistic template based pedestrian detection in infrared videos," in *Proceedings of the IEEE Intelligent Vehicle Symposium*, vol. 1, 2002, pp. 15–20.
- [105] Y. Fang, K. Yamada, Y. Ninomiya, B. K. P. Horn, and I. Masaki, "A shape-independent method for pedestrian detection with far-infrared images," *IEEE Transactions on Vehicular Technologies*, pp. 1679–1697, 2004.
- [106] F. Xu, X. Liu, and K. Fujimura, "Pedestrian detection and tracking with night vision," *IEEE Transactions on Intelligent Transportation Systems*, vol. 6, pp. 63–71, 2005.
- [107] C.-J. Paia, H.-R. Tyanb, Y.-M. Liangc, H.-Y. M. Liaoc, and S.-W. Chen, "Pedestrian detection and tracking at crossroads," *Pattern Recognition*, vol. 37, pp. 1025–1034, 2004.
- [108] X. Liu and K. Fujimura, "Pedestrian detection using stereo night vision," *IEEE Transactions on Vehicular Technologies*, pp. 1657–1665, 2004.
- [109] O. Sidla, Y. Lypetsky, N. Brande, and S. Seer, "Pedestrian detection and tracking for counting applications in crowded situations," in *Proceedings of the IEEE International Conference on Video and Signal Based Surveillance (AVSS '06)*, 2006, p. 70.
- [110] B. Leibe, E. Seemann, and B. Schiele, "Pedestrian detection in crowded scenes," *Computer Vision and Pattern Recognition, IEEE Computer Society Conference on*, vol. 1, pp. 878–885, 2005.
- [111] D. Huber, A. Kapuria, R. Donamukkala, and M. Herbert, "Parts-based 3d object classification," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2004.
- [112] R. Osada, T. Funkhouser, B. Chazelle, and D. Dobkin, "Matching 3d models with shape distributions," in *Shape Modeling International, Genova, Italy*, May 2001.
- [113] L. Havasi, Z. Székely, and T. Szirényi, "Pedestrian detection using derived third-order symmetry of legs," in *Proceedings of the International Conference on Computer Vision and Graphics*, 2004.
- [114] P. Viola, M. J. Jones, and D. Snow, "Detecting pedestrians using patterns of motion and appearance," in *Proceedings of the Ninth IEEE International Conference on Computer Vision*, vol. 1, 2003. [Online]. Available: <http://www.merl.com/papers/docs/TR2003-90.pdf>
- [115] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Comm. of the ACM* 24, pp. 381–395, June 1981.
- [116] J.-Y. Bouguet. Camera calibration toolbox for matlab. [Online]. Available: www.vision.caltech.edu/bouguetj
- [117] T. Marita, F. Oniga, S. Nedevschi, T. Graf, and R. Schmidt, "Camera calibration method for far range stereovision sensors used in vehicles," in *Proceedings of IEEE Intelligent Vehicles Symposium, (IV2006), Tokyo, Japan*, June 2006, pp. 356–363.

- [118] S. Bota and S. Nedevschi, “Multi-feature walking pedestrians detection for driving assistance systems,” *IET Intelligent Transportation Systems Journal*, vol. 2, pp. 92–104, June 2008.
- [119] —, “Tracking multiple objects in urban traffic environments using dense stereo and optical flow,” in *Proceedings of the 14th International IEEE Conference On Intelligent Transportation Systems (ITSC)*, Washington DC, USA, 2011.
- [120] —, “Vision based obstacle tracking in urban traffic environment,” in *Proceedings of the 7th IEEE Conference On Intelligent Computer Communication and Processing (ICCP)*, Cluj-Napoca, Romania, 2011.
- [121] —, “Pedestrians detection using motion field and dense stereo,” in *Proceedings of IEEE 2nd International Conference on Intelligent Computer Communication and Processing (ICCP)*, 1–2 September, Cluj-Napoca, Romania, 2006, pp. 45–51.
- [122] —, “Multi-feature walking pedestrian detection using dense stereo and motion,” in *Proceedings of the 4th International Workshop on Intelligent Transportation (WIT2007)*, 20–21 March, Hamburg, Germany, 2007, pp. 113–118.
- [123] —, “Multi-feature walking pedestrians detection for driving assistance systems,” in *Proceedings of the 6th European Congress and Exhibition on Intelligent Transport Systems and Services*, 18–20 June, Aalborg, Denmark, 2007.
- [124] S. Nedevschi, C. Tomiuc, and S. Bota, “Stereo based pedestrian detection for collision avoidance applications,” in *Proceedings of the Workshop on Planning, perception and navigation of intelligent for vehicles, IEEE International Conference on Robotics and Automation (ICRA2007)* 10–14 April, Roma, Italy, 2007.
- [125] S. Nedevschi, S. Bota, and C. Tomiuc, “Stereo-based pedestrian detection for collision-avoidance applications,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 10, no. 3, September 2009.
- [126] S. Bota and S. Nedevschi, “A framework for object detection, tracking and classification in urban traffic scenarios using stereovision,” in *Proceedings of the 5th Conference on International Conference on Intelligent Computer Communication and Processing (ICCP)*, Cluj-Napoca, August 2009, pp. 153–158.

Anexa A

Articole publicate

A.1 În jurnale ISI cu factor de impact

1. S. Nedevschi, **S. Bota**, C. Tomiuc – Stereo-Based Pedestrian Detection for Collision-Avoidance Applications, IEEE Transactions on Intelligent Transportation Systems, September 2009, Volume 10, Number 3, pp. 380–391 ISSN: 1524-9050 (factor de impact 2.234)
2. **S. Bota**, Sergiu Nedevschi – Multi-Feature Walking Pedestrians Detection for Driving Assistance Systems, IET Intelligent Transportation Systems Journal, June 2008, Volume 2, Issue 2, pp. 92–104, ISSN: 1751-956X (factor de impact 0.473)

A.2 În volumele unor conferințe ISI Proceedings, IEEE Xplore, SpringerLink

3. **S. Bota**, S. Nedevschi – Tracking Multiple Objects in Urban Traffic Environments Using Dense Stereo and Optical Flow – Proceedings of the IEEE 14th International Conference On Intelligent Transportation Systems – Washington DC, USA (ITSC 2011)
4. **S. Bota**, S. Nedevschi – Vision Based Obstacle Tracking in Urban Traffic Environment – Proceedings of the 2011 IEEE 7th International Conference On Intelligent Computer Communication and Processing – Cluj-Napoca, pp. 321–238,
5. R. Brehar, C. Fortuna, **S. Bota**, D. Mladenic and S. Nedevschi – Spatio-temporal Reasoning for Traffic Scene Understanding – Proceedings of the 2011 IEEE 7th International Conference On Intelligent Computer Communication and Processing – Cluj-Napoca, pp. 377–384, ISBN: 978-14577-1478-8
6. S. Nedevschi, T. Marita, R. Danescu, F. Oniga, **S. Bota**, I. Haller, C. Pantilie, M. Drulea, C. Golban – On-board 6D Visual Sensor for Intersection Driving Assistance – 14th International Forum on Advanced Microsystems for Automotive Applications – Advanced Microsystems for Automotive Applications 2010, Springer, pp. 253–264, ISBN: 978-3-642-12647-5
7. C. Pantilie, **S. Bota**, I. Haller, S. Nedevschi, Real-time obstacle detection using dense stereo vision and dense optical flow – Proceedings of the 2010 IEEE 6th International Conference on

Intelligent Computer Communication and Processing, Volume 1, pp. 191–196, ISBN: 978-1-4244-8228-3

8. **S. Bota**, S. Nedevschi – A framework for Object Detection, Tracking and Classification in Urban Traffic Scenarios Using Stereovision – Proceedings of the 2009 IEEE 5th Conference on Intelligent Computer Communication and Processing, 27–29 August 2009, Cluj-Napoca, Romania, pp. 153–158 ISBN: 978-1-4244-5007-7
9. I. Giosan, S. Nedevschi, **S. Bota** – Real Time Stereo Vision Based Pedestrian Detection Using Full Body Contours - Proceedings of the 2009 IEEE 5th Conference on Intelligent Computer Communication and Processing, 27–29 August 2009, Cluj-Napoca, Romania, pp. 79–86, ISBN: 978-1-4244-5007-7
10. S. Nedevschi, T. Marita, R. Danescu, F. Oniga, **S. Bota** - Onboard Stereo Sensor for Intersection Driving Assistance. Architecture and Specification – Proceedings of the 2009 IEEE 5th Conference on Intelligent Computer Communication and Processing, 27–29 August 2009, Cluj-Napoca, Romania, pp. 409–416, ISBN: 978-1-4244-5007-7
11. S. Nedevschi, R. Danescu, T. Marita, F. Oniga, C. Pocol, **S. Bota**, M.-M. Meinecke, M. A. Obojski – Stereovision-Based Sensor for Intersection Assistance, Advanced Microsystems for Automotive Applications 2009 – Smart Systems for Safety, Sustainability and Comfort, pp.129–164, ISBN: 978-3-642-00744-6
12. **S. Bota**, S. Nedevschi – Camera Motion Detection Using Monocular and Stereovision, Proceedings of the 2008 IEEE 4th IEEE Conference on Intelligent Computer Communication and Processing, pp. 275–278, Cluj Napoca, Romania, ISBN: 978-1-4244-2673-7
13. **S. Bota**, S. Nedevschi – Multi-Feature Walking Pedestrians Detection for Driving Assistance Systems, Proceedings of the 6th European Congress and Exhibition on Intelligent Transport Systems and Services, 18–20 June, Aalborg, Denmark, 2007
14. S. Nedevschi, C. Tomiuc, **S. Bota** – Stereo Based Pedestrian Detection for Collision Avoidance Applications, Proceedings of the Workshop on Planning, perception and navigation of intelligent for vehicles, IEEE International Conference on Robotics and Automation (ICRA2007) 10–14 April 2007, Roma, Italy ISBN: 1-4244-0601-3
15. **S. Bota**, S. Nedevschi – GLSCENEINT: A Synthetic image generation tool for testing computer vision systems, Proceedings of the 2006 IEEE 2nd International Conference on Intelligent Computer Communication and Processing, pp. 39–44, 1–2 September, Cluj-Napoca, Romania, 2006, ISBN: 973-662-233-5
16. **S. Bota**, S. Nedevschi – Pedestrians Detection Using Motion Field and Dense Stereo, Proceedings of the 2006 IEEE 2nd International Conference on Intelligent Computer Communication and Processing, pp 45–51, 1–2 September, Cluj-Napoca, Romania, 2006, ISBN: 973-662-233-5
17. **S. Bota**, Sergiu Nedevschi – Real-time trinocular vision for driving assistance systems, Proceedings of the 2006 IEEE 2nd International Conference on Intelligent Computer Communication and Processing, 1–2 September, Cluj-Napoca, Romania, 2006, pp. 69–76, ISBN: 973-662-233-9

18. S. Nedevschi, T. Marita, R. Danescu, F. Oniga, C. Pocol, S. Sobol, C. Tomiuc, C. Vancea, **S. Bota**, Stereovision Sensor for Driving Assistance, Proceedings of the IEEE 2nd International Conference on Intelligent Computer Communication and Processing, 1–2 September, Cluj-Napoca, Romania, 2006, pp. 105–112, ISBN: 973-662-233-9
19. S. Nedevschi, **S. Bota**, T. Marita, F. Oniga, C. Pocol - -Real-Time 3D Environment Reconstruction Using High Precision Trinocular Stereovision 2006 IEEE-TTTC International Conference on Automation, Quality and Testing, Robotics AQTR 2006 (THETA 15), pp. 333–338, May 25–28 2006 Cluj-Napoca, Romania

A.3 Capitoile de carte

20. S. Nedevschi, R. Danescu, T. Marita, F. Oniga, C. Pocol, **S. Bota** and C. Vancea – A Sensor for Urban Driving Assistance Systems Based – Chapter 14 of Stereo Vision book, published by I-Tech Education and Publishing KG, Vienna, Austria, pp. 235–258, ISBN: 978-953-7619-22-0

A.4 În volumele altor conferințe

21. **S. Bota**, S. Nedevschi – Multi-Feature Walking Pedestrian Detection Using Dense Stereo and Motion, Proceedings of the 4th International Workshop on Intelligent Transportation(WIT2007), pp 113–118, 20–21 March, Hamburg, Germany, 2007

Anexa B

Citări independente

1. S. Nedevschi, **S. Bota**, C. Tomiuc – Stereo-Based Pedestrian Detection for Collision-Avoidance Applications, IEEE Transactions on Intelligent Transportation Systems, September 2009, Volume 10, Number 3, ISSN: 1524-9050

Citat de:

- (a) C. G. Keller, T. Dang, H. Fritz, A. Joos, C. Rabe, D. M. Gavrila – Active Pedestrian Safety by Automatic Braking and Evasive Steering, IEEE Transactions on Intelligent Transportation Systems, 30 June 2011, Volume PP Issue: 99 1–13, ISSN: 1524-9050
 - (b) H. Menga, K. Appiaha, S. Yuea, A. Huntera, M. Hobdenb, N. Priestleyb, P. Hobdenb and C. Pettitb – A modified model for the Lobula Giant Movement Detector and its FPGA implementation – Computer Vision and Image Understanding, Volume 114, Issue 11, November 2010, Pages 1238-1247
 - (c) M. Enzweiler, D. Gavrila, D. M. Enzweiler – A Multi-Level Mixture-of-Experts Framework for Pedestrian Classification – IEEE Transactions on Image Processing, 11 April 2011, Volume PP Issue: 99 ISSN: 1057-7149
 - (d) C. G. Keller, M. Enzweiler, M. Rohrbach, D. F. Llorca, C. Schnörr, D. M. Gavrila – The Benefits of Dense Stereo for Pedestrian Detection – IEEE Transactions on Intelligent Transportation Systems, 12 May 2011, Volume: PP Issue:99 1–11 ISSN: 1524-9050
 - (e) N. Wu, F. Chu, S. Mammar, M. Zhou – Petri Net Modeling of the Cooperation Behavior of a Driver and a Copilot in an Advanced Driving Assistance System – IEEE Transactions on Intelligent Transportation Systems, 19 May 2011, Volume: PP Issue:99 1–13 ISSN: 1524-9050
2. **S. Bota**, Sergiu Nedevschi – Multi-Feature Walking Pedestrians Detection for Driving Assistance Systems, IET Intelligent Transportation Systems Journal, June 2008, Volume 2, Issue 2, pp. 92–104, ISSN: 1751-956X

Citat de:

- (a) A. Pérez Grassia, V. Frolovb, F. P. Leónb – Information fusion to detect and classify pedestrians using invariant features – Information Fusion, Volume 12, Issue 4, October 2011, 284-292 Special Issue on Information Fusion for Cognitive Automobiles