

**Universitatea Tehnica din Cluj-Napoca,
Facultatea de Automatica si Calculatoare
Catedra de Calculatoare**

**PERCEȚIA MULTI-SCALARĂ ȘI MULTI-MODALĂ A MEDIILOR 3D
DINAMICE PRIN FUZIUNEA DATELOR DE STEREOVIZINE
DENSĂ, FLUX OPTIC DENS ȘI ODOMETRIE VIZUALĂ
(MULTI-SCALE MULTI-MODAL PERCEPTION OF DYNAMIC 3D
ENVIRONMENTS BASED ON THE FUSION OF DENSE STEREO,
DENSE OPTICAL FLOW AND VISUAL ODOMETRY
INFORMATION)**

**Cod proiect: PN-II-ID-PCE-2011-3-1086
Contract 344 din 01/01/2012**

Sinteza lucrării,
an 1 faza unica

Director proiect:

Prof. dr. ing. Sergiu Nedevschi

Colectiv:

Sef lucr. dr. ing. Tiberiu Marita,

Sef. lucr. dr. ing. Radu Danescu,

Asist. dr. ing. Florin Oniga

Doctorand ing. Andrei Vatavu

Doctorand ing. Marius Drulea

Doctorand ing. Ion Giosan

Doctorand ing. Mihai Negru

Doctorand ing. Raluca Brehar

Doctorand ing. Voichita Popescu

Decembrie, 2012

Cluj-Napoca

Raport științific

privind implementarea proiectului în perioada ianuarie – decembrie 2012

Titlul proiectului: Percepția multi-scalară și multi-modală a mediilor 3D dinamice prin fuziunea datelor de stereoviziune densă, flux optic dens și odometrie vizuală.

Obiectivul etapei 2012:

Studiul comparativ, teoretic și experimental al implementării unor soluții de înalta acuratețe și timp real pentru stereoviziune, flux optic, mișcare proprie, fuziune temporală, percepție și urmărirea mediului; dezvoltarea fundamentelor matematice

SO1.1 Stereoviziune densă de înalta acuratețe în timp real

A1.1.1 Studiu comparativ, teoretic și experimental, al stării actuale în domeniu

Stereo reconstrucția este o arie intensă de cercetare. Multe soluții au fost propuse și discutate în două studii asupra domeniului, unul furnizat de către Scharstein și Szeliski [1] și altul de către Brown *et al* [2]. În [1] autorii furnizează de asemenea o taxonomie și o evaluare a algoritmilor de stereoviziune și stabilesc pașii efectuați de majoritatea algoritmilor: calcularea costului de potrivire, agregarea costurilor, calcularea disparităților/optimizarea și rafinarea disparităților.

Performanțele în timp real sunt de dorit pentru majoritatea aplicațiilor. Pentru aplicațiile ADAS execuția în timp real este obligatorie. Au fost propuși mulți algoritmi de stereoviziune care îndeplinesc această constrângere. Una din aceste soluții este sistemul DeepSea [3] din Tyzx. Ei folosesc o corelare locală combinată cu transformata Census pentru îmbunătățirea acurateței și pentru o procesare rapidă. Este una din cele mai rapide metode disponibile, fiind capabilă să calculeze hărți de disparitate dense la viteza de 200fps pentru imagini de 512x480 la disparitate maximă de 52 pixeli.

Gehrig *et al* prezintă o implementare în timp real al algoritmului SGM utilizând o placă FPGA [4]. Au raportat o rată de procesare (frame rate) de 27Hz pentru imagini de 340x200 (scalate de la 680x400), o disparitate maximă de 64 și cu o verificare stânga-dreapta completă, a.i. SGM e calculat de două ori pentru fiecare pereche de imagini. Pentru o regiune de interes cât rezoluția imaginilor originale, este capabil să calculeze o disparitate de până la 128 pixeli.

Într-o versiune mai recentă Gehrig și Rabe [5] arată cum se paralelizează un algoritm SGM pe procesoarele desktop. Pentru ca algoritmul să ruleze în timp real, imaginile sunt scalate la o rezoluție înjumătățită și la un sfert din rezoluția inițială. Pentru imagini 640x320 și o disparitate maximă de 128, implementarea lor rulează la o frecvență de 14 Hz pe un procesor Intel Core i7 rulând la 3.3 GHz.

Ernst și Hirschmüller [6] prezintă o implementare GPU al algoritmului SGM care utilizează MI (mutual information) ca și metrică. Soluția rulează pe NVIDIA GeForce 8800 ULTRA la 8.8 fps pentru imagini 470x375 și o disparitate de până la 64 de pixeli. Pe imagini cu rezoluții mai mici, metoda rulează în timp real.

În ceea ce privește algoritmul original al lui Hirschmüller am constatat două probleme. Prima problemă vizează lipsa implementării în timp real a algoritmului pentru rezoluții medii și mari. A doua problemă vizează acuratețea metodei, care în unele situații reale, în scene de trafic poate fi destul de scăzută.

Metoda originală proprie propune ca un număr de minim de 8 direcții să fie folosite pentru optimizare în vreme ce 16 sunt necesare pentru calitate maximă. Pentru a reduce costul computațional propunem folosirea a 4 direcții de optimizare. Această idee a fost evaluată și în [7] autorii raportând o creștere a numărului de corelații eronate de la 12.8%, în cazul a 8 direcții de optimizare, la 14% în cazul optimizării de-a lungul a 4 direcții. Pentru sisteme unde constrângerile de timp sunt importante, această creștere relativ mică a numărului de erori poate fi considerată neglijabilă datorită câștigului în timp de execuție. Cea de-a doua problemă legată de algoritmul original a fost descoperită în timpul evaluărilor preciziei la nivel sub-pixel. S-a dovedit că având o valoare mică pentru penalizarea $P1$, se obține o reconstrucție mai dispersată a punctelor 3D. Dispersia aceasta este vizibilă și în zonele mai bine texturate. Folosirea unei valori mari pentru penalizarea $P1$ nu ar fi corectă din punct de vedere matematic deoarece există constrângerea ca $P1$ trebuie să fie mai mic decât cea mai mică valoare pe care o poate lua $P2$.

A1.2.1 Dezvoltarea fundamentelor matematice și proiectarea soluției originale

În acest proiect, propunem îmbunătățiri a costului de potrivire, în optimizarea disparităților și în rafinarea acestora, în așa fel încât să rezulte o creștere a acurateței algoritmului SGM păstrând în același timp o execuție în timp real pe GPU. S-a propus o implementare care să păstreze doar penalizarea $P2$, funcția de energie devenind:

$$E(D) = \sum_p (C(p, D_p) + \sum_{q \in N_p} P2 * T \left[|D_p - D_q| > 1 \right]).$$

Rezultate comparative pentru folosirea a 4 sau 8 direcții de optimizare sunt foarte apropiate, justificându-se astfel alegerea primei variante.

S-a constatat ca exista o legătura strânsă între algoritmul de reconstrucție și funcția de interpolare la precizie de sub-pixel. În consecință s-au dezvoltat doua metode originale de proiectare a funcțiilor de interpolare una bazata pe metoda egalizării histogramelor și a doua bazata pe căutarea funcției optime de interpolare pornind de la un set de date sintetice cu comportare cunoscuta. Metodele au fost prezentate și publicate în [8], [9].

SO1.2 Flux optic dens de înalta acuratețe în timp real

A1.1.2 Studiu comparativ, teoretic și experimental, al stării actuale în domeniu

În [10] Horn & Schunk (HS) au pus bazele modelului variațional de estimare a fluxului optic. Problema variațională conține doi termeni: primul reprezintă termenul de potrivire a datelor, iar al doilea este termenul de netezime sau regularizare. Termenul de potrivire a datelor măsoară gradul de similitudine în intensitate dintre doi pixeli (sau regiuni). El este eficient la marginile obiectelor deoarece acestea reprezintă trăsături importante și cea mai buna potrivire este unica sau sunt puține potriviri posibile. Pentru zone omogene mari, posibilitățile de potrivire sunt mai multe, iar fluxul corect este dedus din fluxul zonelor de margine. Metodele variaționale utilizează scheme piramidale pentru a detecta deplasamente mari, deoarece ele se bazează pe o aproximare în serie Taylor de grad 1, care este valida doar pentru deplasamente mici.

Modelul HS are dezavantajul ca nu păstrează discontinuitățile mișcării și propaga fluxul în toate direcțiile, fără a ține cont de proprietățile locale prezente în imagini.

Utilizarea funcțiilor robuste de penalizare [11-24] definește o problema de optimizare ne-neteda, permițând soluții netede și astfel păstrând discontinuitățile care apar în mișcare. O aproximare des întâlnita [11-14, 19, 23, 24] în aceste modele este: $|x| \approx \sqrt{x^2 + \epsilon}$. Norma L1 nu este diferențiabilă, dar aproximarea ei epsilon este diferențiabilă, permițând o manipulare matematică mai ușoară. În modelele amintite aceasta aproximare este cel mai adesea însoțită de o schema SOR (successive-over-relaxation), sacrificant paralelismul. Modelul TV-L1 [6] nu utilizează aceasta aproximare și încorporează algoritmul paralelizabil de optimizare convexă prezentat în [25].

În aplicațiile reale schimbările de luminozitate au loc destul de frecvent. Pentru a rezolva schimbările de luminozitate imaginile pot fi pre-procesate [17, 22, 26], iar o alta variantă este de a utiliza o funcție de cost care este invarianta la aceste schimbări. Pe de alta parte, acuratețea impune o metrica foarte discriminativă. Metrica ZNCC (Zero-mean Normalized Cross-Correlation) îndeplinește ambele condiții. Din păcate, ZNCC are o formă mai complicată și literatura nu furnizează o soluție pe care să ne putem baza în totalitate.

În cadrul acestui proiect, propunem un model rapid și sigur ce utilizează ZNCC ca funcție de cost. De asemenea, pentru o rulare în timp real, modelul propus trebuie să fie paralelizabil.

A1.2.2 Dezvoltarea fundamentelor matematice și proiectarea soluției originale

Fie f și T două matrici (două regiuni) de aceeași dimensiune. O metrica comună pentru măsurarea similitudinii dintre f și T este SSD (sum of squared distances):

$$SSD(f, T) := \sum_{s \in \mathcal{N}} (f_s - T_s)^2,$$

unde \mathcal{N} conține indicii s a fiecărei locații din matrici.

Zero-Mean Normalized Cross-Correlation (ZNCC) se definește ca:

$$ZNCC(f, T) = \frac{1}{|\mathcal{N}|} \frac{\langle f - \mu_f, T - \mu_T \rangle}{\sigma_f \cdot \sigma_T},$$

unde $|\mathcal{N}|$ reprezintă dimensiunea regiunii, μ_X și σ_X sunt media și respectiv deviația standard ale lui $X \in \{f, T\}$. Dacă f și T sunt identice valoarea acestei măsuri este 1. Cea mai bună potrivire minimizează expresia $\{1 - ZNCC(f, T)\}$.

Fie $I_1, I_2 : \Omega (\subseteq \mathbb{N}^2) \rightarrow \mathbb{R}$ două imagini consecutive ale aceleiași secvențe și fie $w = (u, v) : \Omega \rightarrow \mathbb{R}^2$ fluxul optic dintre ele, care este necunoscut și urmează a fi determinat. $\Omega = \{(x, y) | 1 \leq x \leq N, 1 \leq y \leq M\}$ este lătecea bidimensională carteziană standard iar (N, M) reprezintă dimensiunea celor două imagini. Fiecare pixel are un index bidimensional $i \in \Omega$ și o vecinătate notată cu \mathcal{N}_i , care conține indicii din jurul lui i . Imaginile pot lua valori reale ca și argumente, dar valorile în aceste locații sunt calculate prin interpolare. Notăm cu w_i sau $w(i)$ fluxul optic în locația i , și aceleași notații sunt utilizate pentru toate celelalte variabile. Presupunem că la un moment dat un flux inițial $w_0 = (u_0, v_0)$ este disponibil și notăm cu

$\tilde{I}_2(i) = I_2(i + w_0(i))$ imaginea translatată cu acest flux. Aceasta imagine trebuie în continuare rafinată cu deplasamentul $dw = w - w_0$ în așa fel încât să se potrivească cât mai bine cu imaginea fixată I_1 , în sensul distanței ZNCC dintre blocuri:

$$E_d(w) := \sum_{i \in \Omega} \left\{ 1 - \frac{1}{|\mathcal{N}_i|} \sum_{s \in \mathcal{N}_i} \frac{(\tilde{I}_2(s + dw_i) - \tilde{\mu}_2(i + dw_i)) \cdot (I_1(s) - \mu_1(i))}{\tilde{\sigma}_2(i + dw_i) \cdot \sigma_1(i)} \right\}$$

În formula de mai sus $\mu_1(i)$ și $\sigma_1(i)$ sunt media și deviația standard ale regiunii \mathcal{N}_i din jurul locației i în imaginea I_1 . Definițiile lui $\tilde{\mu}_2$ și $\tilde{\sigma}_2$ sunt similare. $|\mathcal{N}_i|$ reprezintă dimensiunea regiunii i (dimensiunile sunt egale). Fiecare termen al sumei globale reprezintă eroarea ZNCC dintre regiunea imaginii rafinate $\tilde{I}_2(i + dw_i)$ și regiunea lui $I_1(i)$ în locația dată. Am presupus că deplasamentul local dw_i este constant în regiunea din jurul locației i . Energia de mai sus este destul de complicată pentru abordările variaționale tradiționale pentru că trebuie să o liniarizăm și să o transformăm într-o funcție convexă.

SO1.3 Estimarea de înaltă acuratețe a mișcării proprii și extragerea vectorilor 3D de mișcare

A1.1.3 Studiu comparativ, teoretic și experimental, al stării actuale în domeniu

În general metodele bazate pe stereoviziune sunt mai puțin sensibile la erorile de potrivire, nu prezintă problema ambiguității scalării, și oferă rezultate mai precise. În general, o funcție de cost se calculează pe baza potrivirii punctelor de trasaturi dintre două perechi de cadre stereo, și este folosită la calcularea mișcării sistemului stereo. Recenzii a unor metode de calculare a mișcării proprii bazate pe stereoviziune sunt realizate în [27] și [28].

În [29] și [30] minimizarea funcției de cost este formulată ca o problemă de calculare a verosimilității maxime în coordonate 3D ale lumii. În abordarea prezentată în [31] parametrii de mișcare sunt calculați pe baza unui număr dublu de „quaternions” și funcția de cost este exprimată în coordonate 3D. Metodele prezentate în [32] și [33] sunt bazate pe algoritmul P3P prezentat în [27]. În [34], [35] și [36] funcția de cost este exprimată în planul imaginii, pe baza transferului tensorului trifocal. În [37], [38] și [39] funcția de cost este exprimată în planul imaginii, pe baza reconstrucției stereo a punctelor de trăsături din primul cadru considerat și re-proiecția lor în cadrul al doilea. Metodele [40], [41] și [42] lucrează în spațiul disparității pentru a reprezenta funcția de cost.

Minimul funcției de cost poate fi calculat folosind metode non-lineare de minimizare, dacă rotația este reprezentată folosind unghiurile Euler ([35], [37], [34], [38]), sau poate fi formulată ca o problemă de pătrat minim, dacă rotația este aproximată folosind extinderea lui Taylor de ordin I ([40], [41] și [42]).

A1.2.3 Dezvoltarea fundamentelor matematice și proiectarea soluției originale

Algoritmul de estimare a mișcării proprii operează pe două cadre stereo: cadrul stâng anterior, cadrul drept anterior, cadrul stâng, cadrul drept. Pașii algoritmului sunt următorii: determinarea corespondențelor între cele patru cadruri; calcularea unui set maxim de date care respectă modelul („inliers”) folosind algoritmul RANSAC (RANDOM SAMPLE CONSENSUS) [43]; calcularea valorii finale a parametrilor de mișcare pe baza setului de date care respectă modelul. Detaliile metodei sunt prezentate în [44].

SO2.1 Fuziunea multi-sensor temporală a informației

A2.1.1 Studiu comparativ, teoretic și experimental, al stării actuale în domeniu; specificații

Metodele de fuziune temporală se pot clasifica după tipul senzorilor utilizați în procesul de achiziție a datelor. Majoritatea algoritmilor se bazează pe date provenite de la senzori laser [45-47], ultrasunet [48] sau sisteme bazate pe stereo-viziune [49].

Soluțiile existente se pot clasifica și după tipul trăsăturilor utilizate. Anumite metode pot aplica fuziunea temporală direct pe întreg norul de puncte 3D. Fiecare punct fiind descris printr-o poziție și printr-un vector de mișcare. Alte abordări se bazează pe utilizarea unor atribute de nivel înalt cum ar fi voxelii [54], stixelii [52], boxuri 2D, cuboide 3D [55], fronturi de diferență [56], modele poligonale [57] sau contururi ale obiectelor [58][59]. Anumite metode existente constau în utilizarea informației de intensitate [50]. Alte soluții se bazează doar pe informația de geometrie a obiectelor [46][54].

Franke *et al* [50] propune o soluție de urmărire a unui set de puncte din imagine și elaborează o fuziune spațială (pozițiile punctelor 3D) și temporală (mișcarea 3D). Pentru efectuarea acestei fuziuni se utilizează un procedeu de extragere a trăsăturilor optime cu ajutorul abordării KLT (Kanade-Lucas-Tomasi) [60] precum și o filtrare temporală folosind un filtru Kalman extins.

Majoritatea tehnicilor existente încearcă să asigure o complexitate redusă a algoritmilor de procesare și în același timp să mențină un nivel de acuratețe ridicat a rezultatelor. Din acest motiv, procesarea datelor și fuziunea lor la nivel intermediar sunt cel mai des întâlnite în literatură. Astfel, informația primară poate fi transformată în reprezentări intermediare mai

compacte sub forma unor hărți de ocupare, structuri de evidență arborescente (octrees) [53][54], Stixel World [51], hărți de elevație multinivel [52] hărți de gaussiani multirezoluție [53] sau hărți de diferență [56].

Modelele de fuziune descrise în acest capitol pornesc de la premisa că reprezentarea lumii este dată sub forma unei hărți de ocupare probabilistică. Fiecare celulă din hartă este descrisă de o probabilitate de ocupare. S-au specificat următoarele elemente ale proceselor de fuziune temporala: modelul probabilistic generic, modelul de fuziune prin acumulare, politica de update a măsurătorilor, alinierea sistemelor de coordonate, asocierea datelor. S-a experimentat o soluție de fuziune a informației folosind Hărți de Diferență Temporală pentru reprezentarea mediilor dinamice [56].

SO2.2 Percepția și urmărirea multi-modală și multi-scalară a mediului

A2.1.1 Studiu comparativ, teoretic și experimental, al stării actuale în domeniu; specificații

O reprezentare suficient de generală, dar totuși compactă, este cea bazată pe „stixeli” dinamici [61]. Stixelii sunt coloane verticale, ce pornesc de pe suprafața drumului, au o grosime fixă dar suficient de mică pentru a reda detaliile fine, și înălțime variabilă. Părțile vizibile ale obstacolelor din trafic pot fi modelate pe baza acestor coloane 3D, care se vor constitui într-o înfășurătoare a acestora. Stixelii au un vector viteză în planul drumului, vector ce se actualizează prin urmărirea folosind filtrul Kalman.

Hărțile digitale de înălțime (Digital Elevation Maps, DEM) sunt o altă abordare compactă pentru reprezentarea mediilor 3D complexe, fără a face presupuneri asupra compoziției acestora. În hărțile de înălțime mediul este reprezentat sub forma unei rețele 2D de celule, fiecare celulă fiind caracterizată de o valoare a înălțimii. Hărțile digitale de înălțime pot fi structuri enorme, folosite pentru caracterizarea terenului pe suprafețe mari [62], ceea ce le face foarte utile în domeniul aplicațiilor spațiale de explorare robotizată a suprafețelor planetare [63], dar pot fi și structuri locale, folosite pentru navigarea roboților autonomi [64], pentru reprezentarea mediului în aplicații de asistență a conducerii auto [65], sau chiar și pentru aplicații în medii închise, controlate, precum urmărirea pietonilor în interiorul unei clădiri [66]. Hărțile digitale de înălțime pot fi actualizate în timp real, folosind mai multe tipuri de senzori 3D, cele mai populare fiind din familia laserscannerelor sau a sistemelor de stereoviziune [65] [66]. Odată înălțimile calculate, harta poate fi analizată pentru a putea identifica celulele traversabile, celulele de tip obstacol, sau de altă natură [64] [65]. În [52] găsim o extensie mai avansată a hărților de înălțime, extensie numită Hartă de Suprafețe Multinivel (Multi Level Surface Map), hartă ce poate modela suprafețele suspendate. O altă extensie, mai generală, este harta de ocupare multi-volum, prezentată în [67]. O soluție similară, ce combină elevația și ocuparea folosind un element de incertitudine (sub numele de „credibilitate”) este prezentată în [68].

Hărțile de ocupare pot fi considerate a fi o extensie sau o alternativă la hărțile de înălțime. Primele hărți de ocupare au fost prezentate de Elfes în [69]. Sistemul prezentat de Coue și echipa în [70] folosește o hartă de ocupare cu patru dimensiuni, unde fiecare celulă are poziție și două componente de viteză pe fiecare axă. O soluție mai naturală pentru reprezentarea distribuției de probabilitate a vitezelor pe celulă de hartă de ocupare este prezentată de Chen în [71].

O unificare a celor două metode de reprezentare și de inferență asupra mediului 3D, o combinație dintre harta de înălțimi și harta de ocupare, pare a fi o soluție naturală pentru un model mai complet, și un mecanism de inferență mai performant. În [72] găsim o soluție bazată pe stereoviziune ce menține două hărți, una pentru ocupare și alta pentru înălțime, iar în [73] harta de înălțimi este folosită ca pas intermediar spre obținerea reprezentării de tip hartă de ocupare.

Soluția propusă pentru modelarea și urmărirea mediului 3D se va baza pe rezultatele noastre anterioare, descrise în [74].

Referințe

- [1] D. Scharstein, R. Szeliski, "A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms," *Int. J. of Computer Vision*, vol. 47, pp. 7-42, 2002.
- [2] M. Z. Brown, D. Burschka, G. D. Hager, "Advances in Computational Stereo," *Trans. PAMI*, vol. 25, pp. 993-1008, 2003.
- [3] I. J. Woodfill, G. Gordon, R. Buck, "Tyxz DeepSea High Speed Stereo Vision System," *CVPRW 2004*, pp. 41-45.
- [4] S. K. Gehrig, F. Eberli, T. Meyer, "A Real-Time Low-Power Stereo Vision Engine Using Semi-Global Matching," *ICVS 2009*, pp. 134-143.
- [5] S. K. Gehrig, C. Rabe, "Real-time Semi-Global Matching on the CPU," *CVPRW*, 2010, pp. 85-92.
- [6] I. Ernst, H. Hirschmuller, "Mutual Information Based Semi-Global Stereo Matching on the GPU," *Advances in Visual Computing*, vol. 5358, pp. 228-239, 2008.
- [7] S. Hermann, R. Klette, E. Destefanis, "Inclusion of a Second-Order Prior into Semi-Global Matching," *Advances in Image and Video Technology*, vol. 5414, pp. 633-644, 2009.
- [8] I. Haller, S. Nedeveschi, "Design of Interpolation Functions for Subpixel-Accuracy Stereo-Vision Systems," *IEEE Trans. on Image Processing*, vol. 21, pp. 889-898, 2012.
- [9] C. D. Pantilie, S. Nedeveschi, "SORT-SGM: Subpixel Optimized Real-Time Semiglobal Matching for Intelligent Vehicles," *IEEE Trans. on Vehicular Technology*, vol. 61, pp. 1032-1042, 2012.
- [10] B. Horn, B. Schunck, "Determining optical flow," *Artificial Intelligence*, vol. 17, pp. 185-203, 1981.
- [11] T. Brox, A. Bruhn, N. Papenberg, J. Weickert, "High accuracy optical flow estimation based on a theory for warping," *European Conference and Computer Vision*, 2004, pp. 25-36.
- [12] A. Bruhn, J. Weickert, "Towards Ultimate Motion Estimation: Combining Highest Accuracy with Real-Time Performance," *ICCV 2005*, pp. 749 - 755.
- [13] A. Bruhn, J. Weickert, C. Schnörr, "Lucas/Kanade Meets Horn/Schunck: Combining Local and Global Optic Flow Methods," *Int. J. of Computer Vision*, vol. 61, pp. 211-231, 2005.
- [14] J. Weickert, A. Bruhn, N. Papenberg, T. Brox, "Variational Optic Flow Computation: From Continuous Models to Algorithms," *IWCIVIA 2003*, pp. 1-6.
- [15] C. Zach, T. Pock, H. Bischof, "A duality based approach for realtime TV-L1 optical flow," *DAGM Conference on Pattern Recognition*, 2007, pp. 214-223.
- [16] M. Werlberger, T. Pock, H. Bischof, "Motion estimation with non-local total variation regularization," *CVPR*, 2010, pp. 2464-2471.
- [17] M. Werlberger, W. Trobin, T. Pock, A. Wedel, D. Cremers, H. Bischof, "Anisotropic Huber-L1 Optical Flow," *BMVC*, 2009.
- [18] M. Drulea, S. Nedeveschi, "Total variation regularization of local-global optical flow," *IEEE Int. Transp. Sys. Conf.*, 2011, pp. 318-323.

- [19] M. Drulea, I. R. Peter, S. Nedeveschi, "Optical flow. A combined local-global approach using L1 norm," ICCV, 2010, pp. 217-222.
- [20] L. Xu, J. Chen, J. Jia, "A Segmentation Based Variational Model for Accurate Optical Flow Estimation," European Conf. on Computer Vision, 2008, pp. 671-684.
- [21] L. Xu, J. Jia, Y. Matsushita, "Motion detail preserving optical flow estimation," CVPR, 2010, pp. 1293-1300.
- [22] D. Sun, S. Roth, M. Black, "Secrets of optical flow estimation and their principles," CVPR, 2010, pp. 2432-2439.
- [23] T. Brox, J. Malik, "Large Displacement Optical Flow: Descriptor Matching in Variational Motion Estimation," Trans. PAMI, vol. 33, pp. 500-513, 2011.
- [24] T. Brox, C. Begler, J. Malik, "Large displacement optical flow," CVPR, 2009, pp. 41-48.
- [25] A. Chambolle, "An Algorithm for Total Variation Minimization and Applications," J. of Mathematical Imaging and Vision, vol. 20, pp. 89-97, 2004.
- [26] A. Wedel, T. Pock, C. Zach, H. Bischof, and D. Cremers, "An Improved Algorithm for TV-L1 Optical Flow," SGAVMA, 2009, pp. 23-45.
- [27] H. Alismail, B. Browning, M. B. Dias, "Evaluating Pose Estimation Method for Stereo Visual Odometry on Robots," IAS, 2010.
- [28] N. Sünderhauf, P. Protzel, "Stereo Odometry-A Review of Approaches", Technical Report 3/07, Institute of Automation, Dep. of Electrical Engineering and Information Technology, Chemnitz University of Technology, 2007.
- [29] C. F. Olson, L. H. Matthies, M. Schoppers, Mark W. Maimone, "Stereo Ego-motion Improvements for Robust Rover Navigation", ICRA, 2001, pp. 1099-1104.
- [30] C.F. Olson, L.H. Matthies, H. Schoppers, M.W. Maimone, "Robust stereo ego-motion for long distance navigation", CVPR 2000, pp. 453-458.
- [31] A. Milella, R. Siegwart, "Stereo-Based Ego Motion Estimation Using Pixel Tracking and Iterative Closest Point", ICCV, 2006, pp. 21-27.
- [32] D. Nistér, O. Naroditsky, J. Bergen, "Visual odometry", CVPR, 2004, pp. 652-659.
- [33] T.Y. Tian, C. Tomasi, D.J. Heeger, "Comparison of approaches to egomotion computation", CVPR, 1996, pp. 315-320
- [34] S. A. Rodríguez, V. Frémont, P. Bonnfait, "An Experiment of a 3D Real-Time Robust Visual Odometry for Intelligent Vehicles", ITSC, 2009, pp. 1-6.
- [35] S. A. Rodríguez, V. Frémont, P. Bonnfait, V. Cherfaoui, "An embedded multi-modal system for object localization and tracking", IV, 2010, pp. 211-216.
- [36] A.I. Comport, E. Malis, P. Rives, "Accurate Quadrifocal Tracking for Robust 3D Visual Odometry", ICRA, 2007, pp. 40-45.
- [37] A. Howard, "Real-time stereo visual odometry for autonomous ground vehicles", IROS, 2008, pp. 3946 - 3952.
- [38] C. Golban, I. Golban, S. Nedeveschi, "Vision based three-dimensional vehicle motion detection by minimizing nonlinear functions", ICCP 2010, pp.239-242.
- [39] A. Geiger, J. Ziegler, C. Stiller, "StereoScan: Dense 3d Reconstruction in Real-time", IV, 2011, pp. 963-968.
- [40] A. Talukder, L. Matthies, "Real-time detection of moving objects from moving vehicle using dense stereo and optical flow", IROS 2004, pp. 3718-3725.
- [41] A. Bak, S. Bouchafa, D. Aubert, "Detection of Independently Moving Objects Through Stereo Vision and Ego-Motion Extraction", IV, 2010, pp. 863 - 870.
- [42] D. Demirdjian, T. Darrell, "Motion estimation from disparity images", ICCV , 2001, pp. 213 - 218.
- [43] C. V. Stewart, "Robust Parameter Estimation in Computer Vision", SIAM REVIEW, 1999, vol. 41, no. 3, pp. 513-537.
- [44] C. Golban, I. Szakacs, S. Nedeveschi, "Stereo based visual odometry in difficult traffic scenes", IV 2012, pp. 736-741.
- [45] R. Madhavan, "Terrain aided localization of autonomous vehicles", Symposium on Automation and Robotics in Construction, 2002. pp. 1-8.
- [46] A. Petrovskaya, S. Thrun, "Model based vehicle tracking for autonomous driving in urban environments", RSS, 2008.
- [47] T. de Candia, "3d tracking of dynamic objects with a laser and a camera", Technical report, Autonomous System Lab, ETH Zurich, 2010.
- [48] N. Fairfield, G.A. Kantor, D. Wettergreen, "Real-Time SLAM with Octree Evidence Grids for Exploration in Underwater Tunnels," J. of Field Robotics, vol.24, no.1/2, pp. 3-21, 2007.
- [49] A. Barth, U. Franke, "Where will the oncoming vehicle be the next second?," IV 2008, pp.1068-1073.
- [50] U. Franke, C. Rabe, H. Badino, S. Gehrig, "6d-vision: Fusion of stereo and motion for robust environment perception," DAGM Conference on Pattern Recognition, 2005, pp. 216-223.
- [51] D. Pfeiffer, U. Franke, "Efficient Representation of Traffic Scenes by Means of Dynamic Stixels", IV, 2010, pp. 217-224.
- [52] R. Triebel, "Multi-level surface maps for outdoor terrain mapping and loop closing", IROS, 2006, pp. 2276-2282.
- [53] R. Danescu, F. Oniga, S. Nedeveschi, "Particle Grid Tracking System for Stereovision Based Environment Perception", IV 2010, pp. 987 - 992.
- [54] A. Azim, O. Aycard, "Detection, classification and tracking of moving objects in a 3D environment," IV, 2012, pp. 802-807
- [55] R. Danescu, S. Nedeveschi, M.M. Meinecke, T. Graf, "Stereovision Based Vehicle Tracking in Urban Traffic Environments", IEEE Int. Transp. Sys. Conf., 2007, pp. 859-865
- [56] A. Vatavu, R. Danescu, S. Nedeveschi; "Real-time dynamic environment perception in driving scenarios using difference fronts," IV, 2012, pp.717-722.
- [57] C.C. Wang, C. Thorpe, M. Hebert, S. Thrun, and H. Durrant-Whyte. "Simultaneous localization, mapping and moving object tracking", Int. J. of Robotics Research, vol. 26 no. 9, pp. 889-916, 2007.
- [58] S. Prakash, C. Thomas, "Contour tracking with condensation/stochastic search", Dept. of CSE, IIT Kanpur, 2007.
- [59] M. Yokoyama and T. Poggio, "A Contour-Based Moving Object Detection and Tracking", ICCCN, 2005, pp. 271-276.
- [60] B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," IJCAI, 1981, pp. 674-679.
- [61] D. Pfeiffer, U. Franke, "Efficient representation of traffic scenes by means of dynamic stixels," IV, 2010, pp. 217-224.
- [62] D. F. Huber, M. Hebert, "A new approach to 3-D terrain mapping," IROS, 1999, pp. 1121-1127.
- [63] M. Vergauwen, M. Pollefeys, L. Van Gool, "A stereo-vision system for support of planetary surface exploration," MVA, vol. 14, pp. 5-14, 2003.
- [64] S. Lacroix, A. Mallet, D. Bonnafous, G. Bauzil, S. Fleury, M. Herrb, R. Chatila, "Autonomous Rover Navigation on Unknown Terrains: Functions and Integration," Int. J. of Robotics Research, vol. 21, pp. 917-942, 2002.
- [65] F. Oniga S. Nedeveschi, "Processing Dense Stereo Data Using Elevation Maps: Road Surface, Traffic Isle, and Obstacle Detection," IEEE Trans. on Vehicular Technology, vol. 59, pp. 1172-1182, 2010.
- [66] M. Harville, "Stereo person tracking with adaptive plan-view templates of height and occupancy statistics," J. of Image and Vision Computing, vol. 22, pp. 127-142, 2004.
- [67] I. Dryanovski, W. Morris, X. Jizhong, "Multi-volume occupancy grids: An efficient probabilistic 3D mapping model for micro aerial vehicles," IROS 2010, pp. 1553-1559.
- [68] B. Gassmann, L. Frommberger, R. Dillmann, K. Berns, "Real-time 3D map building for local navigation of a walking robot in unstructured terrain," IROS, 2003, pp. 2185-2190.
- [69] A. Elfes, "Using Occupancy Grids for Mobile Robot Perception and Navigation," Computer, vol. 22, pp. 46-57, 1989.
- [70] C. Coué, C. Pradalier, C. Laugier, T. Fraichard, and P. Bessière, "Bayesian Occupancy Filtering for Multitarget Tracking: An Automotive Application," Int. J. of Robotics Research, vol. 25, pp. 19-30, 2006.
- [71] C. Chen, C. Tay, C. Laugier, K. Mekhnacha, "Dynamic Environment Modeling with Gridmap: A Multiple-Object Tracking Application," ICARC 2006, pp. 1-6.
- [72] M. Kumar, D.P. Garg, "Three-Dimensional Occupancy Grids With the Use of Vision and Proximity Sensors in a Robotic Workcell," ASME 2004, pp. 1029-1036.
- [73] H. Lategahn, W. Derendarz, T. Graf, B. Kitt, and J. Effertz, "Occupancy grid computation from dense stereo and sparse structure and motion points for automotive applications," IV, 2010, pp. 819-824.
- [74] R. Danescu, F. Oniga, and S. Nedeveschi, "Modeling and Tracking the Driving Environment With a Particle-Based Occupancy Grid," IEEE Trans. on Intelligent Transportation Systems, vol. 12, pp. 1331-1342, 2011.

Bibliografie – lucrari publicate de colectivul proiectului

1. S. Nedevschi, V. Popescu, D. Radu, M. Tiberiu, F. Oniga, "Accurate Ego-Vehicle Global Localization at Intersections through Alignment of Visual Data with Digital Map", accepted in IEEE Transactions on Intelligent Transportation Systems, 2012.
2. C. Pantilie, S. Nedevschi, "SORT-SGM: Sub-pixel Optimized Real-Time Semi-Global Matching for Intelligent Vehicles", in IEEE Transactions on Vehicular Technology, vol. 61, no. 3, 2012, pp. 1032-1042.
3. I. Haller, S. Nedevschi, "Design of Interpolation Functions for Sub-Pixel Accuracy Stereo-Vision Systems", IEEE Transactions on Image Processing, vol. 21, no. 2, 2012, pp. 889-898.
4. R. Danescu, F. Oniga, S. Nedevschi, "Modeling and Tracking the Driving Environment with a Particle Based Occupancy Grid", IEEE Transactions on Intelligent Transportation Systems, Vol.12, Issue 4, 2011, pp. 1331 – 1342.
5. A. Ciurte, S. Nedevschi, I. Rasa, "An algorithm for solving some nonlinear systems with applications to extremum problems", in Taiwanese Journal of Mathematics, vol. 16, no. 3, 2012, pp. 1137-1150.
6. Popescu, V.; Danescu, R.; Nedevschi, S.; , "On-road position estimation by probabilistic integration of visual cues," Intelligent Vehicles Symposium (IV), 2012 IEEE , vol., no., pp.583-589, 3-7 June 2012
7. Vatavu, A.; Danescu, R.; Nedevschi, S.; , "Real-time dynamic environment perception in driving scenarios using difference fronts," Intelligent Vehicles Symposium (IV), 2012 IEEE , vol., no., pp.717-722, 3-7 June 2012
8. Golban, C.; Istvan, S.; Nedevschi, S.; , "Stereo based visual odometry in difficult traffic scenes," Intelligent Vehicles Symposium (IV), 2012 IEEE , vol., no., pp.736-741, 3-7 June 2012
9. Vatavu, A.; Nedevschi, S., "Real-time modeling of dynamic environments in traffic scenarios using a stereo-vision system," Intelligent Transportation Systems (ITSC), 2012 15th International IEEE Conference on , vol., no., pp.722-727, 16-19 Sept. 2012
10. Vatavu, A.; Nedevschi, S., "Vision based Real-time Modeling of Dynamic Unstructured Environments in Driving Scenarios", in proceedings of ICINCO 2012, pp.141-149
11. Pantilie, C.; Nedevschi, S.; , "Optimizing the Census Transform on CUDA enabled GPUs," Intelligent Computer Communication and Processing (ICCP), 2012 IEEE International Conference on , vol., no., pp.201-207, Aug. 30 2012-Sept. 1 2012
12. Popescu, Voichita; Nedevschi, Sergiu; , "Cut-in maneuver recognition and behavior generation using Bayesian networks and fuzzy logic," Intelligent Computer Communication and Processing (ICCP), 2012 IEEE International Conference on , vol., no., pp.123-130, Aug. 30 2012-Sept. 1 2012
13. Giosan, Ion; Nedevschi, Sergiu; , "A solution for probabilistic inference and tracking of obstacles classification in urban traffic scenarios," Intelligent Computer Communication and Processing (ICCP), 2012 IEEE International Conference on , vol., no., pp.221-227, Aug. 30 2012-Sept. 1 2012
14. Cosma, Andrei Claudiu; Brehar, Raluca; Nedevschi, Sergiu; , "Part-based pedestrian detection using HoG features and vertical symmetry," Intelligent Computer Communication and Processing (ICCP), 2012 IEEE International Conference on , vol., no., pp.229-236, Aug. 30 2012-Sept. 1 2012
15. Pojar, Dan; Jeong, Pangyu; Nedevschi, Sergiu; , "Robust visual odometry using stereo reconstruction error model," Intelligent Computer Communication and Processing (ICCP), 2012 IEEE International Conference on , vol., no., pp.149-154, Aug. 30 2012-Sept. 1 2012
16. C. Pantilie, "Advanced Techniques for Real-Time Stereovision based Perception", PhD Thesis, TUCN, 01.09.2012.
17. C. Lemnaru, "Strategies for Dealing with Real World Classification Problems", PhD Thesis, TUCN, 01.09.2012
18. V. Popescu, "Contributions to Localization and Situation Assessment Based on Stereovision for Driving Assistance Systems", TUCN, 19.11.2012.

Director proiect,
Sergiu Nedevschi _____