

Marius Leordeanu
Email: leordeanu@gmail.com
Telefon: +40746 033 711

Memoriu de Activitate Stiintifica

Domeniul nostru de cercetare este **Computer Vision**, o ramura din Inteligenta Artificiala de o importanta crescanda in stiinta, tehnologie si industrie. Computer Vision isi propune sa *inteleaga problema vederii (a interpretarii lumii vizuale)* si sa faca computerele sa *vada* si sa inteleaga lumea din imagini si secvente video, intr-un mod inteligent similar cu cel uman.

Metodele dezvoltate in computer vision fac posibila interpretarea automata a imaginilor si secventelor video: computerele de azi “vad”, de exemplu, daca intr-o imagine exista o fata umana, si sunt capabile sa o gaseasca si sa o recunoasca. Ele pot sa alinieze automat scene tri-dimensionale urbane pentru a crea automat harti 3D fotorealiste de tip Google Streetview. Metode recente din vision fac posibila lumea jocurilor pe calculator interactive (Microsoft Kinect) si a masinilor cu pilot automat (Google Driverless Car). Computer Vision leaga rezultate si cunostinte stiintifice din mai multe domenii precum Matematica, Statistica, Informatica, Invatarea Automata (Machine Learning), Inteligenta Artificiala, Robotica, dar si Stiinte Neuro-cognitive. Aceasta stiinta va contribui esential si profund la dezvoltarea viitoare a lumii in care traim.

2010-2013: Detectarea Generalizata a Marginilor (Generalized Boundary Detection)

Problema de *Boundary detection* este fundamentala in computer vision si e esentiala pentru o varietate de sarcini cum ar fi gasirea automata a marginilor si conturilor de obiecte, segmentarea de regiuni semantice, detectarea simetriei, si recunoasterea de obiecte si categorii de obiecte. In [1,2] am introdus o formulare generalizata de boundary detection, cu solutie *closed-form*, aplicabila la localizarea diverselor tipuri de margini (boundaries), cum ar fi: margini in imagini naturale, si margini cauzate de ocluzie in secvente video. Metoda noastra de *Generalized Boundary Detection (Gb)* combina simultan interpretari ale imaginilor la mai diferite nivele de interpretare intr-o singura problema de optimizare (gasirea vectorului si a valorii proprii a unei matrici special construite) care

gaseste in mod optimal atat orientarea (vectorul propriu) cat si puterea raspunsului de margine (radicalul valorii proprii). Aceassta solutie complet noua face ca metoda noastra sa obtina rezultate state-of-the-art la un cost computational si de memorie mult mai scazut (40x mai rapid) decat metodele competitive curente. De asemenea, propunem doua componente care se combina natural cu Gb: mai intai, introducem o procedura foarte rapida de soft-segmentare care creeaza noi straturi de input (nivele de interpretare) pentru o imbunatarire semnificativa a performantei, si doi, propunem o metoda eficienta de grupare si clasificare pe contouri care, la un cost computational foarte scazut, imbunatateste si mai mult calitatea marginilor produse.

Rezultatele noastre, in forma lor initiala, au fost acceptate si publicate [1] mai intai la European Conference on Computer Vision (ECCV), Florenta, Italia, 2012 – una din primele trei conferinte din domeniu la nivel mondial. O extensie a acestei lucrari [2], care include metoda de interpretare inteligenta a conturilor a fost submisă la IEEE-Transactions on Pattern Analysis and Machine Intelligence (impact factor 6.68) si a trecut de prima faza de reviewing (o revizie a fost deja submisă si ne asteptam ca va fi acceptata). Cercetarea pe problema de boundary detection face parte dintr-o colaborare stransa si continua cu Google Research (Dr. Rahul Sukthankar).

La nici un an de la publicare, Gb a fost folosit deja [3] pentru imbunatarirea descriptorilor care folosesc informatii de segmentare semantica a obiectelor si regiunilor. Noii descriptori au fost publicati in IEEE-Computer Vision and Pattern Recognition, USA, 2013 [3].

References:

- [1] M. Leordeanu, R. Sukthanakar and C. Sminchisescu, „Efficient Closed-Form Solution to Generalized Boundary Detection”, Florence, Italy, ECCV, 2012
- [2] M. Leordeanu, R. Sukthankar and C. Sminchisescu, „Generalized Boundaries from Multiple Image Interpretations”, (revision submitted), PAMI 2013
- [3] E. Trulls, I. Kokkinos, A. Sanfeliu and F. Moreno-Noguer, „Dense Segmentation-aware Descriptors”, CVPR, Ohio, USA, 2013

2010-2013: Sparse-to-dense Matching pentru estimarea automata de ocluzii si optical flow

Estimarea unui camp dens de corespondente intre frame succesive in video (optical flow estimation) e o problema esentiala in computer vision, cu aplicatii in multe probleme de invatare automata si recunoastere. Noi am dezvoltat o metoda noua de Sparse-to-Dense Matching pentru estimarea de optical flow si a zonelor de ocluzie. Aceast algoritm a fost acceptat pentru publicare la International Conference on Computer Vision (ICCV), Sydney, Australia, 2013 [4], (prima conferinta in lume din domeniu, , <http://www.iccv2013.org/>).

Ca o alternativa la abordarile piramidale curente de tip coarse-to-fine din literatura de optical flow, algoritmul nostru porneste cautarea de la nivelul mai inalt de interpretare folosind *Sparse Matching* cu descriptori de aparenta sofisticati si relatii geometrice complexe, folosind un model nou local afin care ia in considerare si zonele de ocluzie. Apoi, la pasii ulteriori, cautarea se rafineaza si se indreapta catre un model de flow mai clasic, folosind o metoda noua de interpolare sparse-to-dense, care ofera o tranzitie natural intre campurile sparse si dense. De asemenea, demonstram experimental ca aspectele noastre de aparenta bogate, impreuna cu modelul de constrangeri geometrice complexe permit estimarea corecta a miscarii chiar si in cazurile de deplasari mari si schimbari de aparenta vizuala semnificativa. De asemenea, introducem si o metoda noua de clasificare pentru detectarea de ocluzii, care functioneaza in conjunctie cu metoda noastra de sparse-to-dense matching. Algoritmul introdus a fost validat experimental pe cea mai recenta, mai mare si mai dificila baza de date (Sintel) publica, pe care am obtinut rezultate state-of-the-art [5].

References

- [4] M. Leordeanu, A. Zanfır and C. Sminchisescu, „Locally Affine Sparse-to-Dense Matching for Motion and Occlusion Estimation”, Sydney, Australia, 2013
- [5] **Rezultatele noastre State-of-the-Art** oficiale (metoda noastra apare S2D-Matching in tabelul oficial): <http://sintel.is.tue.mpg.de/results>

Important: metoda noastra (S2D-Matching) depaseste in performanta toate celelalte metode publicate pana acum, de la grupuri de cercetare de top din lume (Berkeley, Japan, Microsoft Research printre altele) pe baza de date Sintel – cel mai nou, mai extins si mai dificil set de testare

existent. In medie, metoda nepublicata inca Deep-Matching-Flow ne-a intrecut (dar nu pe toate tipurile de eroare) – aceasta metoda va apare (va fi publicata) impreuna cu a noastra la ICCV, Sydney, Australia, 2013. De notat e si faptul ca metoda Complex-Flow (care apare in tabel) nu a fost acceptata pentru publicare.

2012-2013: Metode eficiente pentru recunoastere in real-time de actiuni in RGB-D video

Recunoasterea de actiuni umane cu latentă observatională scăzută se bucură de un interes crescând în computer vision datorită dezvoltării rapide de tehnologii de interacțiune om-computer, jocuri interactive pe calculator și aplicații de supraveghere. Noi am dezvoltat o metodă rapidă, simplă și eficientă (numită **Moving Pose Framework**) pentru recunoastere cu latentă scăzută (Real-Time) de activități și acțiuni, care a fost acceptată pentru publicare la International Conference on Computer Vision [6] (ICCV), Sydney, Australia 2013 (prima conferință în lume din domeniu: <http://www.iccv2013.org/>).

Central metodologiei noastre este descriptorul **Moving Pose** care consideră atât informații despre poze (poziția corpului, ale membrilor și încheieturilor) cât și cantități diferențiale (viteză și accelerația) ale încheieturilor corpului în diferite momente în timp. Descriptorul propus, centrat temporal în jurul unei anumite frame de referință, este folosit în conjuncție cu un clasificator modificat de tip kNN - care consideră atât localizarea globală a framei în cadrul secvenței de acțiune, cât și puterea de discriminativitate a descriptorului Moving Pose relativ la alte frame din setul de antrenare. Metoda rezultată e simplă, non-parametrică și face posibilă recunoasterea eficientă și cu latentă scăzută, învățarea *one-shot*, și detectia de acțiuni în secvențe nesegmentate dificile. De asemenea, framework-ul nostru e real-time, scalabil și depășește ca performanță abordările competitive pe baze de date precum Microsoft Research - Action3D și Microsoft Research - DailyActivities3D.

References

[6] M. Zanfir, M. Leordeanu and C. Sminchisescu, „The *Moving Pose*: An Efficient 3D Kinematics Descriptor for Low-Latency Action Recognition and Detection”, ICCV, Sydney, Australia 2013.

2003:2013: Graph Matching, Invatare Automata si Optimizare, Modele Grafice Probabilistice

O parte importantă din cercetarea mea in computer vision include dezvoltarea de metode computationale eficiente pentru recunoasterea automata de obiecte si categorii de obiecte din imagini si secvente video. Algoritmii pe care i-am dezvoltat iau in considerare atat caracteristici vizuale locale dar si relatii geometrice de ordinul doi sau superior. Ele sunt legate de problema cunoscuta sub numele de *graph matching*: obiectele sunt reprezentate ca si graf-uri cu caracteristic vizuale continute in nodurile grafului si cu relatiile geometrice modelate de legaturile dintre noduri. Folosind graph matching obiectele pot fi gasite si recunoscute automat in imagini prin potrivirea automata dintre un graf-model (invatat automat in prealabil) si graful extras automat din imagine sau secventa video.

Metodele de graph matching pe care le-am publicat sunt considerate state-of-the-art in domeniu si, datorita eficientei lor, au fost preluate, utilizate si extinse mai departe in universitati si centre de cercetare de prestigiu ale lumii. Acum, dupa zece ani de munca de cercetare pe aceasta tema, as putea spune ca rezultatele noastre in graph matching (majoritatea incluse in teza mea de doctorat [27]) au deschis poarta pentru graph matching-ul modern in computer vision, bazat pe formularea QAP (Quadratic Assignment Problem) si au contribuit fundamental la formarea unui drum nou pe problema de recunoastere automata de obiecte in vision.

Problemele teoretice pe care le-am abordat pana acum sunt: 1) Graph și hiper-graph matching [7,16]; 2) Inferenta Maximum A Posteriori (MAP) în modele grafice probabilistice Markov Random Fields (MRFs) [14]; 3) Invățare automata supervizata, semi-supervizata si complet nesupervizata pentru graph si hypergraph matching si modele grafice probabilistice Markov Random Fields [11,16]. 4) Metode eficiente de hypergraph clustering [26] și alte metode de optimizare, cum ar fi optimizarea generală a funcțiilor non-negative [15].

Pe lângă contribuțiile mele pentru problema de Graph Matching, am dezvoltat, de asemenea, metode noi pentru alte probleme din computer vision cum ar fi: object tracking in secvente de rezolutie scazuta [35], recunoasterea de categorii de obiecte[8], alinierea / gasirea de corespondente intre scene urbane scanate 3D [36]. Mai in detaliu, rezultatele mele stiintifice includ:

1. Recunoașterea automata a obiectelor si categoriilor de obiecte

- Am combinat detectarea de obiecte cu graph matching folosind fragmente de contururi, astfel depășind cele mai bune rezultate oficiale din 2007 Pascal pentru patru clase de obiecte (în teza de doctorat). Am combinat, de asemenea, recunoașterea de categorii de obiecte bazata pe forma cu detectie de contururi specifice clasei de obiecte si am demonstrat o imbunatatire semnificativa fata de stat-of-the-art (ECCV 2008 [9]).
- Am proiectat si implementat un sistem care foloseste fragmente de contururi si metoda noastra de Spectral Graph Matching, rezultand intr-o metoda noua de invatare semi-supervizata de modele de clase de obiecte (publicate in CVPR 2007 [8] – prezentare orala). Această metodă a fost folosită ulterior pentru recunoașterea de acțiuni din video, la Universitatea din Florida (Yan et al., CVPR 2008 [25]). Ea a devenit subiectul unui capitol de carte (Dickinson et al., 2009 [21]) și a influențat mai multe abordări recente pentru descoperirea automata de obiecte din secvente video.
- Am dezvoltat o metodă de învățare nesupervizata de obiecte din secvente video, cu rezoluție redusă si în prezența ocluziilor (publicata in CVPR 2005 [13]).

2. Invatare automata si inferenta in modele grafice probabilistice

- Am dezvoltat un algoritm de învățare semi-supervizata pentru hypergraph matching si i-am demonstrat eficacitatea pe seturi dificile de date, cu performanța state-of-the-art (publicat in ICCV 2011 [16]). Acesta este primul algoritm publicat pentru invatare in Hypergraph Matching.
- Am propus metoda Integer Projected Fixed Point (IPFP) (NIPS 2009 [12]) pentru graph matching si inferenta MAP in Markov Random Fields. Aceasta a depasit semnificativ metodele state-of-the-art existente. Algoritmul este acum considerat a fi state of the art și a fost deja folosit si extins de către alți cercetatori din străinătate pentru diferite probleme din computer vision: de exemplu, metoda IPFP a fost folosită pentru segmentarea imaginilor la Universitatea de Stat din Oregon - USA (Brendel et al, NIPS 2010 [17].) și Universitatea din Maryland - SUA (Chen și colab, CVPR 2011 [20]). Acesta a fost, de asemenea, extinsă la Universitatea din New South Wales - Australia pentru inferență in MRFs de ordin superior (Semenovich et al, ICPR 2010 [28]).
- Am propus un algoritm eficient de învățare nesupervizata pentru graph matching (IJCV 2012 [11]). Este primul algoritm de invatare nesupervizata pentru graph matching. Am propus de asemenea o metodă similară de învățare pentru probleme de inferenta in modele grafice probabilistice (IJCV 2012 [11]).

- Am creat o metoda de inferenta pentru MRF's inspirata din metoda noastra de Spectral Graph Matching (publicat in ICML 2006 [14]). Metoda noastra a fost extinsa ulterior la Universitatea din Pennsylvania - SUA (Cour et al, AISTATS 2007 [18]). Acesta a fost, de asemenea, folosita recent si de cercetatorii de la Universitatea Stanford și Google pentru segmentarea si clasificare de imagini (Huang et al., CVPR 2011 [29]).
- Am creat algoritmul Spectral Graph Matching (ICCV 2005 [7]) – o metoda eficienta de gasire de corespondente intre doua imagini (sau intre o imagine si un model) care foloseste atat caracteristici vizuale locale cat si relatii geometrice. Această metodă a fost modificata ulterior și extinsa de mai multe grupuri de cercetare, cum ar fi printre altele: Universitatea din Pennsylvania - SUA si INRIA-Paris (Cour et al, NIPS 2007 [19]; Duchenne et al, CVPR 2009 [22]). Acest algoritm este utilizat de diverse grupuri de cercetare pentru: descoperire de obiecte (University of Central Florida [24], Intel Research Labs [33]), captarea miscarilor umane (Max Plank Institutul de Informatică [30]), potrivire de aspecte articulate pentru obiecte in imagini (Institutul Tehnologic Toyota din Chicago [32]) si modelare de scene 3D (Universitatea Tübingen [31]). O discutie mai ampla despre alte abordari care au folosit Spectral Graph Matching [7] se gaseste in teza mea de doctorat [27], in capitolul introductiv.

3. Optimizare

- Am proiectat o nouă metodă de Hypergraph Clustering, folosind relații de ordin doi si superior, care surclaseaza alti algoritmi publicati recent (publicata in AISTATS 2012 [26]).
- Am dezvoltat un algoritm de optimizare globala, care surclasează pe experimentele noastre metode bine-cunoscute, cum ar fi Markov Chain Monte Carlo și Simulated Annealing (publicat in CVPR 2008 [15]). Un nou rezultat teoretic care leagă metoda noastră de metoda de inferenta Belief Propagation a fost publicat la Penn State University (M. Park, teza de doctorat, 2010 [34]).

References

- [7] M. Leordeanu and M. Hebert, A Spectral Technique for Correspondence Problems Using Pairwise Constraints, International Conference on Computer Vision (ICCV), 2005. **Citations: 331.**
- [8]. M Leordeanu, M. Hebert, R. Sukthankar, Beyond Local Appearance: Category Recognition from Pairwise Interactions of Simple Features CVPR 2007. **Citations: 102.**

- [9] J. Mairal, M. Leordeanu, F. Bach, M. Hebert and J. Ponce, Discriminative Sparse Image Models for Class-Specific Edge Detection and Image Interpretation, ECCV, 2008. **Citations: 84.**
- [10] J. Hays, M. Leordeanu, A. Efros and Y. Liu, Discovering Texture Regularity as a Higher-order Correspondence Problem, ECCV, 2006. **Citations: 83.**
- [11] M. Leordeanu, R. Sukthankar and M. Hebert, Unsupervised Learning for Graph Matching, International Journal of Computer Vision (IJCV), 2012. Impact factor: 5.76. **Citations: 32.**
- [12] M. Leordeanu, M. Hebert and R. Sukthankar, An Integer Projected Fixed Point Method for Graph Matching and MAP Inference, Neural Information Processing Systems, 2009. **Citations : 33.**
- [13]. M. Leordeanu and R. Collins, Unsupervised Learning of Object Features from Video Sequences, IEEE Computer Vision and Pattern Recognition (CVPR), 2005. **Citations: 28.**
- [14] M. Leordeanu and M. Hebert, Efficient MAP Approximation for Dense Energy Functions, International Conference on Machine Learning, USA, 2006. **Citations: 13.**
- [15] M. Leordeanu and M. Hebert, Smoothing-based Optimization, Computer Vision and Pattern Recognition, USA, 2008. **Citations: 6.**
- [16] M. Leordeanu, A. Zanfir and C. Sminchisescu, Semi-supervised Learning and Optimization for Hypergraph Matching, International Conference on Computer Vision, Barcelona, 2011. **Citations 6**
- [17] Brendel W, Todorovic S, "Segmentation as maximum weight independent set", Advances in Neural Information Processing (NIPS), Canada, 2010. **Citations: 12.**
- [18] Cour et al., Solving markov random fields with spectral relaxation, Artificial Intelligence and Statistics (AISTATS), USA, 2007. **Citations: 27.**
- [19] Cour T, Srinivasan P, Shi J, Balanced Graph Matching, Advances in Neural Information Processing Systems (NIPS), Canada, 2007. **Citations: 120.**
- [20] Chen, Xi, et al. "Piecing together the segmentation jigsaw using context." Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on. IEEE, 2011. **Citations: 4.**
- [21]. Dickinson et al, "Object Categorization: Computer and Human Vision Perspectives", Cambridge University Press, 2009
- [22] Duchenne O, Bach F, Kweon I, Ponce J, "A tensor-based algorithm for high-order graph matching", Prize paper, CVPR 2009. **Citations: 115.**
- [23] Gaur, Utkarsh, et al. "A "string of feature graphs" model for recognition of complex activities in natural videos.", ICCV, 2011. **Citations 22.**

- [24] Liu, Jingen, Jiebo Luo, and Mubarak Shah. "Recognizing realistic actions from videos "in the wild"." IEEE Computer Vision and Pattern Recognition, USA, 2009. **Citations: 331.**
- [25] Yan, Pingkun, Saad M. Khan, and Mubarak Shah. "Learning 4d action feature models for arbitrary view action recognition." Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on. IEEE, 2008. **Citations 74.**
- [26] M. Leordeanu and C. Sminchisescu, Efficient Hypergraph Clustering, International Conference on Artificial Intelligence and Statistics (AISTATS), La Palma, 2012.
- [27] M. Leordeanu, Spectral Graph Matching, Learning and Inference for Computer Vision, PhD Thesis, Carnegie Mellon University, 2009.
- [28] Semenivich and Sowmya, Tensor power method for efficient MAP inference in higher-order MRF's, International Conference on Pattern Recognition (ICPR), 2010.
- [29] Huan et al., A hierarchical conditional random field model for labeling and segmenting images of street scenes, IEEE-Computer Vision and Pattern Recognition, 2011. **Citations: 8.**
- [30] E De Aguitar et al. Performance capture from sparse multi-view video, ACM Transactions on Computer Graphics SIGGRAPH, 2008, **Citations: 242.**
- [31]. B. Huhle et al., On-the-fly scene acquisition with a handy multi-sensor system, Int. Journal of Intelligent Systems Technologies and Applications, 2008. **Citations: 43**
- [32]. X. Ren, Learning and Matching Line Aspects for Articulated Objects, IEEE-Computer Vision and Pattern Recognition, 2007. **Citations: 14.**
- [33]. D. Parikh and T. Chen, Hierarchical semantics of objects (hSOs), International Conference on Computer Vision, 2007. **Citations: 29.**
- [34]. Minwoo Park, Efficient Mean-Shift Belief Propagation and Its Applications in Computer Vision, Doctoral Dissertation, USA, 2010.
- [35]. R. Collins, Y. Liu and M. Leordeanu, Online Selection of Discriminative Tracking Features, IEEE-TPAMI, 2005. **Citations: 927.**
- [36]. I. Stamos and M. Leordeanu, Automated Feature-based Range Registration of Urban Scence of Large Scale, IEEE-Computer Vision and Pattern Recogniton (CVPR), 2003. **Citations: 118.**