September 11, 2013

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

Research Activity

Our field of research, Computer Vision, is an increasingly important part of Artificial Intelligence, which tries to make computers see and understand the visual world in an intelligent manner, in the same way as humans do. Methods and models from Computer Vision make it possible for computers to automatically interpret and understand images and video sequences. For example, today's computers are able to detect and recognize human faces with almost perfect accuracy. They can also reconize human actions from video and object categories. The list of working computer vision applications is increasing at an exponential rate and the field is starting to mature, with a visible impact on industry and human life. Driveless cars strongly relying on computer vision applications have been developed (Google, Mercedes etc), computer games recognizing human actions are already available and best selling on the market (Microsoft Kinect), programs that automatically detect and recognize human faces are available on virtually every photo camera and image processing software. Even commonly used online maps with 3D street views (such as Google or Bing maps) are made possible by computer vision methods that automatically align 3D laser scanned urban, large scale scenes.

Computer vision connects theoretical and experimental results from many disciplines, such as Mathematics, Statistics, Computer Science, Machine Learning, AI and Neuroscience (Cognitive/Brain Science). Below I briefly enumerate my contributions to this field, in which I have been continuously working for more than eleven years (since January 2002). Please note that each section below has its own **references,** for easier access.

2010-2013: Generalized Boundary Detection and Contour Grouping and Reasoning

Boundary detection is a fundamental computer vision problem that is essential for a variety of tasks, such as contour and region segmentation, symmetry detection and object recognition and categorization. We introduced [1,2] a generalized formulation for boundary detection, with closed-form solution, applicable to the localization of different types of boundaries, such as object edges in natural images and occlusion boundaries from video. Our generalized boundary detection method (Gb) simultaneously combines low-level and mid-level image representations in a single eigenvalue problem and solves for the optimal continuous boundary orientation and strength. The closed-form solution to boundary detection enables our algorithm to achieve state of the art results at a significantly (40x) lower computational cost than current methods. We also propose two complementary novel components that can seamlessly be combined with Gb: first, we introduce a soft-segmentation procedure that provides region input layers to our boundary detection algorithm for a significant improvement in accuracy, at negligible computational cost; second, we present an efficient method for contour grouping and reasoning, which when applied as a final post-processing stage, further increases the boundary detection performance.

The first version of our work was published at the European Confernce on Computer Vision (ECCV), Florence, Italy, 2012, which is one of the top three computer vision conferences in the world [1]. An extension of our approach [2], with the addition of contour grouping and reasoning was submitted to IEEE-Transactions on Pattern Analysis and Machine Intelligence (impact factor 6.68) and passed the first phase of reviewing (a revision has already been submitted, we expect it will eventually be accepted). This work is part of an ongoing collaboration with **Google Research** (Dr. Rahul Sukthankar).

Since its publication, our method was already used in the creation of segmentation-aware descriptors for matching and recognition [3], published in the International Conference on Computer Vision and Pattern Recognition (CVPR) 2013, also one of the first top three computer vision conferences in the world.

Refercences:

[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

<u>2010-2013: Sparse to dense matching for motion (optical flow) and occlusion estimation</u>

Estimating a dense correspondence field (the optical flow) between subsequent video frames is an important task in computer vision, with applications to many visual learning and recognition tasks. We developed a novel sparse-to-dense matching method for motion field estimation with occlusion detection, which was accepted for publication at the International Conference on Computer Vision [4] (ICCV), Sydney, Australia, 2013 (the premier Computer Vision event in the world, http://www.iccv2013.org/).

As an alternative to the current coarse-to-fine approaches from the optical flow literature, we start from the higher level of sparse matching with rich appearance and geometric constraints, using a novel, occlusion aware, locally affine model. Then, we move towards the simpler, but denser classic flow field model, using a novel sparse-to-dense matching interpolation procedure, which offers a natural transition between the sparse and dense correspondence fields. We demonstrate experimentally that our appearance features and complex geometric constraints permit the correct motion estimation even in difficult cases of large displacements and significant appearance changes. We also introduced a novel classification method for occlusion detection,
which works in conjunction with our sparse-to-dense matching. We validated our approach on the newly released Sintel dataset and benchmark, on which we obtain state-of-the-art results [5].

References

[4] M. Leordeanu, A. Zanfir and C. Sminchisescu, „Locally Afine Sparse-to-Dense Matching for Motion and Occlusion Estimation", Sydney, Australia, 2013

[5] **Our State-of-the-Art official restuls (our method is S2D-Matching)** http://sintel.is.tue.mpg.de/results

Important Note: our method (S2D-Matching) is better than all previously published methods, from top research groups in the world (Berkeley, Japan, Microsoft Research among others) on Sintel Dataset and Benchmark, which is the most recent, difficult and extensive optical flow (motion flow) dataset curently available. We are surpassed (on average) by Deep-Matching-Flow (which is not published yet, it will also appear in ICCV). Also note that the method Complex-Flow appearing in the list was NOT accepted for publication.

2012-2013: Efficient methods for real time action reconition from RGB-D video

Human action recognition with low observational latency is receiving a growing interest in computer vision due to the rapidly developing technologies in human-robot interaction, computer gaming and surveillance. We developed a fast, simple, yet powerful non-parametric **Moving Pose (MP)** framework to low-latency human action and activity recognition, with **State-of-the-Art** performance on the latest RGB-D (captured with Microsoft Kinect) datasets, which was accepted for publication at the International Conference on Computer Vision [6] (ICCV), Sydney, Australia, 2013 (the premier Computer Vision event in the world, http://www.iccv2013.org/).

Central to our methodology is a moving pose descriptor that considers both pose information as well as differential quantities (speed and acceleration) of the human body joints at different temporal locations. The proposed frame-based descriptor is used in conjunction with a modified kNN classifier that considers both the global temporal location of a particular frame within the action sequence as well as the discrimination power of its moving pose descriptor compared to other frames in the training set. The resulting method is simple, non-parametric, and enables low-latency recognition, one-shot learning, and action detection in difficult unsegmented sequences. Moreover, the framework is real-time, scalable, and outperforms more sophisticated approaches on challenging benchmarks like MSR-Action3D or MSR-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, Learning and Optimization, Probabilistic Graphical Models

An important part of my research includes the development of efficient methods for object matching and recognition using second or higher-order geometric relationships between local visual features. The theoretical problems I have addressed are: 1) Graph and hyper-graph matching [7,16]; 2) Maximum A Posteriori (MAP) inference in Markov Random Fields (MRFs) [14]; 3) Semi-supervised learning for graph matching, hyper-graph matching, and Markov Random Fields [11,16]; 4) Efficient hyper-graph clustering [26] and other optimization methods, such as general optimization of non-negative functions [15]. My work has been published in top international conferences and journals, receiving considerable attention from the international computer vision community. Besides my theoretical contributions, I have also developed novel methods for computer vision applications such as: finding correspondences between images, object matching, category discovery and recognition from images and video, object tracking in low resolution video, and registration of large scale urban scenes. My algorithms for graph matching (Spectral Graph Matching [1]) and the Integer Projected Fixed Point Method [7]) are currently among the most efficient in the literature, being generally accepted as **State-of-the-Art**. They are being used and extended by research groups from several prestigious universities and research labs. In more detail, I have been working on:

1. Object Recognition and Segmentation:

- Combined detection and matching using contour fragments for object category classification, outperforming the best official results from the Pascal 2007 challenge on four object classes (in the PhD thesis). We have also combined object class recognition using shape with object category specific contours and demonstrated improvement over state-of-the-art (ECCV 2008 [9]).
- Designed and implemented a system for both object specific and category recognition using contour fragments and our spectral matching method, and developed a new method for learning object category models in a semi-supervised fashion (CVPR 2007 [8]). This method was later used for action recognition at the University of Central Florida (Yan et al., CVPR 2008 [25]). It became the subject of a book chapter (Dickinson et al., 2009 [21]) and influenced other recent approaches to object discovery and recognition using shape (please see Intro of my PhD Thesis [27]).
- Developed a method for learning object models from image sequences in an unsupervised manner, while dealing with low resolution data, occlusions and change in pose (CVPR 2005 [13]).

2. Inference and Learning for Graph Matching and Markov Random Fields:

- Developed a semi-supervised learning algorithm for hyper-graph matching and demonstrated its effectiveness on difficult datasets (ICCV 2011 [16]). It is the first method published on learning for hyper-graph matching.
- Designed a novel hyper-graph matching method that outperforms all current state-of-the-art algorithms and can also improve on their performance if used in combination (ICCV 2011 [16]).
- Designed the Integer Projected Fixed Point (IPFP) (NIPS 2009 [12]) algorithm for graph matching and MAP inference using pair-wise constraints, which significantly outperformed previous state-of-the-art methods. It is now generally considered to be a state-of-the-art algorithm for graph matching and it has already been used and extended by other researchers around the world for different computer vision tasks: IPFP was used for image segmentation at Oregon State University - USA (Brendel et al., NIPS 2010 [17]) and University of Maryland – USA (Chen et al., CVPR 2011, [20]). It was also extended at University of South Wales - Australia for MAP inference in higher-order MRF (Semenovich et al., ICPR 2010, [28]).
- Developed and implemented an efficient algorithm for unsupervised learning for matching using pair-wise constraints (IJCV 2012 [11]); it is the first algorithm published for learning graph matching in an unsupervised fashion; developed a similar learning method for MAP problems (also in IJCV 2012 [11]).
- Designed a novel inference method for MRF inspired from our spectral matching algorithm (ICML 2006 [14]). Our method was later extended at the University of Pennsylvania – USA (Cour et al., AISTATS 2007 [18]). It has also been recently used by researchers from Stanford University and Google Research on image segmentation and labeling (Huang et al., CVPR 2011 [29]).

- Developed the spectral matching method and implemented a system for object specific recognition (ICCV 2005 [7]). This method was later modified and extended by research groups from: University of Pennsylvania – USA and INRIA-Paris (Cour et al., NIPS 2007 [19]; Duchenne et al., CVPR 2009 [22]). It has been used by other research groups on: object discovery (University of Central Florida [24], Intel Research Labs [33]), capturing human performance (Max Plank Institute of Informatics [30]), matching articulated 2D object aspects (Toyota Technological Institute of Chicago [32]) and 3D scene modeling (Tubingen University [31]).

3. Optimization:

- Designed a novel hyper-graph clustering method, using relationships beyond second-order, which outperforms recently published algorithms (AISTATS 20012 [26]).
- Developed a novel algorithm for global optimization that outperforms on our experiments well-known methods such as Markov Chain Monte Carlo and Simulated Annealing (CVPR 2008 [15]). A novel theoretical result linking our method to Mean-shift Belief Propagation and Gaussian Belief Propagation was established at Penn State University (M. Park, PhD thesis, 2010 [34]).

References

[7] M. Leordeanu and M. Hebert, A Spectral Technique for Correspondence Problems Using Pair-wise 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 Confernence 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, Learing 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 hierarchicala 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.