PhD and Internship proposals in École Centrale Paris-INRIA

Contact: Iasonas Kokkinos, iasonas.kokkinos@ecp.fr **Center for Visual Computing**
http://cvn.ecp.fr/personnel/iasonas/ http://cvn.ecp.fr/

The Center for Visual Computing of ECP-INRIA will have two PhD position openings at the intersection of computer vision, machine learning and optimization, starting as early as March 2015, if possible.

The candidate is expected to be highly motivated, have a solid background in electrical engineering and/or computer science, and experience with computer vision and/or machine learning. Good coding skills are necessary, while prior experience with object detection or GPU programming will be positively appreciated.

The candidate should have a Masters in Computer Science, or an equivalent Diploma of Engineering degree. It is possible, and preferable, to start with an internship of six months or more and then continue for a PhD. All positions have secured funding until 2018.

Applicants should upload a copy of their CV, a motivation letter, a record of their grades and names of at least two references (ideally, also their letters) at the following link:
https://dbinbox.com/iasonas
Please do not send files larger than 5MB in total. Applications will start being processed on the 20th of March and shortlisted candidates will be contacted before the 30th of April. Preference will be given to applications received before the 20th of March.

Research topics

The two PhD positions will be on (i) deep learning and (ii) deformable part models, and directly relate to our on-going research in ECP-INRIA:
http://cvn.ecp.fr/personnel/iasonas/publications.html
http://cvn.ecp.fr/personnel/iasonas/deeplearning.html
http://cvn.ecp.fr/personnel/iasonas/dpms.html
An indicative, but not exhaustive, description of proposed research topics follows below.

Deep Learning

Topic 1: Joint Semantic Segmentation, Object Detection and Scene Recognition

Deep Learning has pushed the envelope of high-level vision in problems such as image classification, semantic segmentation and object detection; e.g. computers now outperform humans in image classification [1,2].

Our goal will be to explore the interplay between these problems: how can semantic segmentation help and be helped by detection? For instance, if we know that there is a car at some position, can we segment it more easily, and vice versa? How can object detection be used together with image classification? How can we train a system that performs all of these tasks in an integrated, end-to-end manner?

For development we will be using a caffe-based hybrid of Matlab/C++/cuda and systems that have already been developped in our group [3,4,5].

References

[1] Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification, K. He, X. Zhang, S. Ren, J. Sun, http://arxiv.org/abs/1502.01852, 2015.

[2] Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift, S. Ioffe, C. Szegedy, http://arxiv.org/abs/1502.03167, 2015.

[3] Semantic Image Segmentation with Deep Convolutional Nets and Fully Connected CRFs, L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy and A. Yuille, http://arxiv.org/abs/1412.7062, 2014.

[4] Untangling Local and Global Deformations in Deep Convolutional Networks, G. Papandreou, I. Kokkinos, P.-A. Savalle, http://arxiv.org/abs/1412.0296, 2014.

[5] Deformable Part Models with CNN Features, P.-A. Savalle, S. Tsogkas, G. Papandreou and I. Kokkinos, 3rd Parts and Attributes Workshop, ECCV 2014.

Topic 2: Efficient Deep Learning for ImageNet Classification

Current optimization algorithms for deep learning can come up with good solutions, while being particularly simple, using e.g. Stochastic Gradient Descent with momentum. This simplicity comes at the cost of computation time: training one of the best performing networks for the last ImageNet competition took more than 2 weeks on 4 high-end GPUs. But better optimization algorithms could reduce training time.

Recent works in this direction [1],[2] have considered well-established optimization algorithms, including Conjugate Gradients, L-BFGS and Truncated Newton methods [3]. These have been shown to yield accelerations on smaller datasets, but their performance on ImageNet has not yet been tested. Our main objective is to combine such techniques with adaptations particular to the kinds of classifiers used in recent architectures; in particular we will consider efficient implementations of the Dropout [4,5], and explore how these can be combined with the most efficient current algorithms for deep network training [6].

References

[1] On Optimization Methods for Deep Learning, Q. V. Le, et. al., ICML 2011.

[2] Hessian-Free Optimization for Deep Learning, J. Martens, ICML 2010.

[3] Introduction to Numerical Optimization, S. Nocedal and J. Wright, Springer, 2006.

[4] Fast dropout training, S. Wang and C. Manning, ICML 2013.

[5] Dropout Training as Adaptive Regularization, S. Wager, S. Wang and P. Liang, NIPS 2012.

[6] Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift, S. Ioffe, C. Szegedy, http://arxiv.org/abs/1502.03167, 2015.

Topic 3: Learning Invariances for Dense Image and Surface Descriptors

Image and surface descriptors are a main ingredient of success in applications as diverse as recognition, registration, retrieval and segmentation. One challenging aspect of descriptor design is the tradeoff between invariance to signal transformations (e.g. image scaling, rotation, and/or affine transformations) and the discriminative power of the descriptor. Descriptor invariances are typically hand-crafted in descriptor construction, by relying on either front-end operations, e.g. [1], or ideas from image processing [2].

Our goal is to avoid hard-wiring these invariances in descriptor construction, but rather let them be learned from data. We will rely on deep learning techniques developped for equivariant recognition [3,4] and combine them with recent works on metric learning for descriptor construction [5,6]. We will originally explore applications in dense image and surface correspondence [7,8] and then move on to recognition.

References

[1] Object recognition from local scale-invariant features, D. Lowe, ICCV 1999.

[2] Scale invariance without scale selection, I. Kokkinos and A. L. Yuille, CVPR 2008.

[3] Learning to Relate Images, R. Memisevic, PAMI 2013.

[4] Modeling the joint density of two images under transformations J. M. Susskind, G. Hinton, R. Memisevic and M. Pollefeys, CVPR 2011.

[5] Fracking Deep Convolutional Image Descriptors, E. Simo-Serra, E. Trulls, L. Ferraz, I. Kokkinos and F. Moreno-Noguer, arXiv:1412.7062v1, 2014.

[6] ShapeNet: Convolutional Neural Networks on Non-Euclidean Manifolds, J. Masci, D. Boscaini, M. Bronstein and P. Vandergheynst `http://arxiv.org/abs/1501.06297`, 2015.

[7] Dense Segmentation-Aware Descriptors, E. Trulls, I. Kokkinos, A. Sanfeliu, and F. Moreno, CVPR 2013.

[8] Intrinsic Shape Context Descriptors for Deformable Shapes, I. Kokkinos, M. Bronstein, R. Littman and A. Bronstein, CVPR 2012.

Deformable Part Models

The Deformable Part Model (DPM) paradigm [1] has been established as a robust framework to tackle a broad range of problems in object recognition, including e.g. face localization [2], human pose estimation [3] and general object detection [1]. But a main impediment to their broader application is their computation time. Over the last four years ECP-INRIA has developped optimization approaches that tackle several of the underlying problems: fast techniques for part score computation [4,5], combinatorial optimization for star-shaped models [6] and coordination-decomposition algorithms for loopy graphs [7].

Having tackled the optimization problems underlying DPMs in 2D pose spaces, we now want to explore the use of DPMs in higher-dimensional spaces along the following directions.

Topic I: 3D surface fitting and tracking from RGB-D data

Our goal is to use Deformable Part Models for real-time surface registration and tracking using point cloud data: parameterizing a 3D surface as an enseemble of ‘points’ in 3D, we want to rephrase problems such as surface matching and registration in terms of combinatorial optimization. Our goal is to recover globally optimal solutions efficiently, by employing combinatorial optimization, rather than the brute-force techniques currently in use.

Topic II: 3D DPMs for Medical Image Segmentation and Registration

Our recent results [7] have demonstrated that loopy Deformable Part Models yield state-of-the-art results in shape-based segmentation of 2D X-ray images. By increasing the dimensionality of the pose space we intend to address problems such as 3D medical image segmentation, registration and tracking. Our group has extensive collaborations with clinical partners, and a successful application of these techniques can lead to software prototypes used by physicians.

Topic III: 3D category pose from RGB images

3D pose estimation from RGB images has been extensively studied for tasks such as human pose estimation, and is becoming increasingly relevant for 3D object recognition [8-11]. Part-based models are increasingly popular for such tasks, most noteably in the context of object detection. Our goal will be to accelerate such algorithms using discrete optimization, and push further their performance by leveraging upon recent advances in deep learning.

References

[1] Object Detection with Discriminatively Trained Part-Based Models, P. Felzenszwalb, R. Girshick, D. McAllester, D. Ramanan, PAMI 2010

[2] Discriminative appearance models for pictorial structures. M. Andriluka et. al. IJCV 2012.

[3] Face detection, pose estimation, and landmark localization in the wild. X. Zhu and D. Ramanan, CVPR 2012.

[4] Bounding Part Scores for Rapid Detection with Deformable Part Models, I. Kokkinos, PnA-ECCV 2012.

[5] Shufflets: Shared mid-level parts for fast object detection, I. Kokkinos, ICCV, 2013.

[6] Rapid deformable object detection using dual-tree branch-and-bound. I. Kokkinos, NIPS 2011.

[7] Fast and Exact: ADMM-Based Discriminative Shape Segmentation with Loopy Part Models, H. Boussaid and I. Kokkinos, CVPR 2014

[8] Detailed 3d representations for object recognition and modeling, M. Z. Zia, M. Stark, B. Schiele, and K. Schindler, PAMI, 2013

[9] 3d2pm3d deformable part models. B. Pepik, P. Gehler, M. Stark, and B. Schiele., ECCV 2012

[10] Parsing ikea objects: Fine pose estimation. J. J. Lim, H. Pirsiavash, and A. Torralba, ICCV, 2013

[11] Seeing 3d chairs: exemplar part-based 2d-3d alignment using a large dataset of cad models, M. Aubry, D. Maturana, A. Efros, B. Russell, and J. Sivic. CVPR 2014