Cosegmentation of object categories

Jean Ponce: Jean.Ponce@ens.fr
Cordelia Schmid: Cordelia.Schmid@inria.fr
Armand Joulin: ajoulin@fb.com

Location of the internship: The internship will be jointly between the WILLOW team at INRIA Paris[1] and the LEAR team at INRIAGrenoble[2], and will be co-supervised by Jean Ponce and Cordelia Schmid, a ENS professor and a INRIA research director respectively. Both teams are specialized in computer vision, in particular visual recognition. Armand Joulin, a researcher at FaceBook, will be an external advisor to the project. There exists the possibility to continue the research as a PhD student.

Topic: With the amount of on-line available images increasing daily, weakly supervised approaches to object recognition become increasingly important. In many cases, annotations may be noisy (user tags accompanying personal photos on Flickr for example), or missing altogether (in family videos for example). The problem is, then, three-fold: (a) how to deal with limited amounts of noisy supervisory data, (b) how to localize the corresponding objects automatically in the image, and (c) how to segment the objects precisely. A solution to these problems will allow learning object detectors without manual intervention.

Objective: Initial solutions have been developed in both teams [2, 3, 6, 5, 1], using the fact that a set of images or videos share the same labels to find the regions corresponding to these labels — this is the cosegmentation problem [4, 7, 8]. However, the methods we have developed are purely discriminative and do not impose the constraint that regions with the same label in different images should have similar visual feature distributions. The goal of the project is to enrich our current discriminative approach to cosegmentation [3] to model the distribution of the object relevant features with a generative model [7, 8], and to exploit noisy labels. Furthermore, the project will explore combining state-of-the-art approaches for co-localization [1] with co-segmentation approaches [2, 3]. Most of the previous methods are computationally expensive and could not scale to large dataset. In this project, we will explore how to leverage the specificity of our problem to tailor efficient optimization scheme.

Skills and profile: The student must have solid programming skills (the project involves programming in matlab and C) as well as solid mathematics knowledge (especially linear algebra and statistics).

References

[1] R. Cinbis, J. Verbeek, and C. Schmid. Multi-fold mil training for weakly supervised object localization. In *CVPR*, 2014.

[1] http://www.di.ens.fr/willow
[2] http://lear.inrialpes.fr

[2] A. Joulin, F. Bach, and J. Ponce. Discriminative clustering for image co-segmentation. In *CVPR*, 2010.

[3] A. Joulin, F. Bach, and J. Ponce. Multi-class cosegmentation. In *CVPR*, 2012.

[4] G. Kim, E.P. Xing, L. Fei-Fei, and T. Kanade. Distributed cosegmentation vis submodular optimization on anisotropic diffusion. In *ICCV*, 2011.

[5] A. Prest, C. Leistner, J. Civera, C. Schmid, and V. Ferrari. Learning object class detectors from weakly annotated video. In *CVPR*, 2012.

[6] A. Prest, C. Schmid, and V. Ferrari. Weakly supervised learning of interactions between humans and objects. *IEEE PAMI*, 2012.

[7] C. Rother, V. Kolmogorov, T. Minka, and A. Blake. Cosegmentation of image pairs by histogram matching - incorporating a global constraint into MRFs. In *CVPR*, 2006.

[8] S. Vicente, C. Rother, and V. Kolmogorov. Object cosegmentation. In *CVPR*, 2011.