

198:500 Light Seminar: Readings on Generic Object Recognition

Organizer: Ahmed Elgammal

Description

The field of computer vision witnesses recently a great interest focused on solving the generic object recognition problem. Traditionally, object recognition has been at the center of the computer vision field. The explosion of digital imaging and digital video that we are witnessing makes a huge urgent demand for systems and applications that are able to understand images and videos at the semantic level. The goal of this reading seminar is to catch up with the state of the art in this field and understand the achievements and challenges towards solving the generic object recognition problem.

Class Time: Thursday 2-3pm

Class Location: Hill 254

Extended Reading List

- Sept 19: Holistic appearance-based approaches: Murase, Hiroshi, and Nayar [1995]
Presented by: Chan Su Lee
- Sept 27: Part-based Models: Schmid and Mohr [1997], Agarwal, Awan, and Roth [2004]
Presented by: Afzal Mazhar and Marwan A. Torki
- Oct 5: Probabilistic Part-based Models: Burl, Leung, and Perona [1995], Weber, Welling, and Perona [2000], Fergus, Perona, and Zisserman [2003, 2005]
Presented by: Xiaoru Wang and Atul Kanaujia
- Oct 12: What Parts, What Features: Interest point detectors, invariant features, SIFT features PCA-SIFT etc.: Harris and Stephens [1988], Kadir and Brady [2001], Lowe [1999, 2001], D.Lowe [2004]

- Oct 12: Part Selection: Dorkó and Schmid [2003]
Presented by: Zhipeng Zhao
- Oct 19: Bag of Features: Willamowski et al. [2004], Sivic et al. [2005].
Presented by: Jintao Feng
- Nov 2: Beyond Bag of Features: Lazebnik, Schmid, and Ponce [2006], Marszaek and Schmid [2006]
Presented by:
- Nov 2: Boosting for part based recognition: Opelt, Fussenegger, Pinz, and Auer [2004]
Presented by: Junzhou Huang
- Nov 9: Contour-Based Approaches: contour fragments: Ferrari, Tuytelaars, and Van Gool [2006], Opelt, Pinz, , and Zisserman [2006]
Presented by:
- Nov 16: Models for recognition: Generative vs. Discriminative models: Ulusoy and Bishop [2005]
Presented by: Nicholas Michael
- Nov 23: Conditional Random Field for Recognition: Quattoni, Collins, and Darrell [2004], Ren, Fowlkes, and Malik [2005], Levin and Weiss, Shotton, Winn, Rother, and Criminisi [2006]
Presented by: Zhiguo Li
- Dec 7: Biologically Motivated Models: Knoblich, Riesenhuber, Freedman, Miller, and Poggio, Riesenhuber and Poggio [1999, 2003], Serre, Wolf, and Poggio [2005]
Presented by: Ishani Chakraborty
- Search-Based object detection and recognition: Schneiderman and Kanade [2000], Viola and Jones [2002]
 - Rigid 3D Objects: Wolfson and Rigoutsos [1997]
 - Pictorial Structures: Fischler and Elschlager [1973], Felzenszwalb and Huttenlocher [2005], Borenstein and Ullman [2002]
 - Feature Sharing: Scalability to many classes Torralba, Murphy, and Freeman [2004], Murphy-Chutorian and Triesch [2005]
 - Role of Correspondences: Belongie, Malik, and Puzicha [2000], Belongie, Malik, and Puzicha, Berg, Berg, and Malik [2005]
 - Deformable and Articulated Objects: Ramanan, Forsyth, and Zisserman [2005], Fanti, Zelnik-Manor, and Perona [2005]

Resources

International Workshop on Object Recognition 2004

<http://vasc.ri.cmu.edu/~hebert/04workshop/workshop.html>

Data sets: PASCAL database of labeled objects

<http://www.robots.ox.ac.uk/~manik/PASCAL>

References

- Murase, Hiroshi, and Shree K. Nayar. Visual learning and recognition of 3d objects from appearance. *IJCV*, 14(1):5–24, /1995.
- Cordelia Schmid and Roger Mohr. Local grayvalue invariants for image retrieval. *IEEE PAMI*, 19(5):530–535, 1997.
- Shivani Agarwal, Aatif Awan, and Dan Roth. Learning to detect objects in images via a sparse, part-based representation. 26(11):1475–1490, 2004. URL <http://12r.cs.uiuc.edu/cogcomp/Data/Car/>.
- M. Burl, T. Leung, and P. Perona. Face localization via shape statistics. In *In Intl. W. on Automatic Face and Gesture Recognition*, 1995. URL citeseer.ist.psu.edu/burl195face.html.
- Markus Weber, Max Welling, and Pietro Perona. Unsupervised learning of models for recognition. In *ECCV (1)*, pages 18–32, 2000. URL citeseer.ist.psu.edu/weber00unsupervised.html.
- R. Fergus, P. Perona, and A. Zisserman. Object class recognition by unsupervised scale-invariant learning. In *CVPR (2)*, pages 264–271, 2003.
- R. Fergus, P. Perona, and A. Zisserman. A sparse object category model for efficient learning and exhaustive recognition. In *CVPR*, 2005.
- C.J. Harris and M. Stephens. A combined corner and edge detector. 1988. URL www.csse.uwa.edu.au/pk/Research/MatlabFns/Spatial/Docs/Harris/.
- T. Kadir and M. Brady. Scale, saliency and image description. *IJCV*, 2001.
- David G. Lowe. Object recognition from local scale-invariant features. In *Proc. of the International Conference on Computer Vision ICCV, Corfu*, pages 1150–1157, 1999. URL citeseer.ist.psu.edu/lowe99object.html.
- David G. Lowe. Local feature view clustering for 3d object recognition. In *CVPR (1)*, pages 682–688, 2001.
- D.Lowe. Distinctive image features from scale invariant keypoints. *International Journal of Computer Vision*, 60:91–110, 2004.

- Gy. Dorkó and C. Schmid. Selection of scale-invariant parts for object class recognition. In *ICCV*, pages 634–640, 2003.
- Jutta Willamowski, Damian Arregui, Gabriella Csurka, Christopher R. Dance, and Lixin Fan. Categorizing nine visual classes using local appearance descriptors. In *IWLAVS*, 2004.
- Josef Sivic, Bryan C. Russell, Alexei A. Efros, Andrew Zisserman, and William T. Freeman. Discovering objects and their location in images. In *ICCV*, 2005.
- S. Lazebnik, C. Schmid, and J. Ponce. Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories. pages II: 2169–2178, 2006.
- M. Marszaek and C. Schmid. Spatial weighting for bag-of-features. pages II: 2118–2125, 2006.
- Andreas Opelt, Michael Fussenegger, Axel Pinz, and Peter Auer. Weak hypotheses and boosting for generic object detection and recognition. In *ECCV (2)*, pages 71–84, 2004.
- V. Ferrari, T. Tuytelaars, and L.J. Van Gool. Object detection by contour segment networks. pages III: 14–28, 2006.
- Andreas Opelt, Axel Pinz, , and Andrew Zisserman. A boundary-fragment-model for object detection. In *ECCV*, 2006.
- I. Ulusoy and C.M. Bishop. Generative versus discriminative methods for object recognition. In *CVPR*, 2005.
- Ariadna Quattoni, Michael Collins, and Trevor Darrell. Conditional random fields for object recognition. In *NIPS*, 2004.
- X. Ren, C. Fowlkes, and J. Malik. Scale-invariant contour completion using conditional random fields, 2005. URL citeseer.ist.psu.edu/745408.html.
- Anat Levin and Yair Weiss. Learning to combine bottom-up and top-down segmentation. In *ECCV06*.
- J. Shotton, J. Winn, C. Rother, and A. Criminisi. Textonboost: Joint appearance, shape and context modeling for multi-class object recognition and segmentation. In *ECCV*, 2006.
- Ulf Knoblich, Maximilian Riesenhuber, David J. Freedman, Earl K. Miller, and Tomaso Poggio. Visual categorization: How the monkey brain does it. URL citeseer.ist.psu.edu/569436.html.
- M. Riesenhuber and T. Poggio. Hierarchical models of object recognition in cortex, 1999. URL citeseer.ist.psu.edu/riesenhuber99hierarchical.html.

- M. Riesenhuber and T. Poggio. How visual cortex recognizes objects: The tale of the standard model, 2003.
- Thomas Serre, Lior Wolf, and Tomaso Poggio. Object recognition with features inspired by visual cortex. In *CVPR*, 2005.
- H. Schneiderman and T. Kanade. A statistical method for 3d object detection applied to faces and cars. pages 45–51, 2000.
- Paul Viola and Michael Jones. Robust real-time object detection. *International Journal of Computer Vision* 2002, 2002.
- Haim J. Wolfson and Isidore Rigoutsos. Geometric hashing: An overview. *IEEE Computational Science & Engineering*, 4(4):10–21, /1997. URL citeseer.ist.psu.edu/wolfson97geometric.html.
- M. Fischler and R. Elschlager. The representation and matching of pictorial structures, 1973. *IEEE Transaction on Computer c-22(1)*: 67-92.
- Pedro F. Felzenszwalb and Daniel P. Huttenlocher. Pictorial structures for object recognition. *IJCV*, 61(1):55–79, 2005.
- E. Borenstein and S. Ullman. Class-specific, top-down segmentation. In *ECCV*, pages 109–124, 2002.
- Antonio B. Torralba, Kevin P. Murphy, and William T. Freeman. Sharing visual features for multiclass and multiview object detection. In *CVPR*, 2004.
- Erik Murphy-Chutorian and Jochen Triesch. Shared features for scalable appearance-based object recognition. In *WACV/MOTION*, pages 16–21, 2005.
- Serge Belongie, Jitendra Malik, and Jan Puzicha. Shape context: A new descriptor for shape matching and object recognition. In *NIPS*, pages 831–837, 2000. URL citeseer.nj.nec.com/belongie00shape.html.
- Serge Belongie, Jitendra Malik, and Jan Puzicha. Matching shapes. In *ICCV 2001*, pages 454–463. URL citeseer.nj.nec.com/belongie01matching.html.
- A.C. Berg, T.L. Berg, and J. Malik. Shape matching and object recognition using low distortion correspondences. In *CVPR*, 2005.
- D. Ramanan, D.A. Forsyth, and A. Zisserman. Strike a pose: Tracking people by finding stylized poses. In *CVPR*, 2005.
- C. Fanti, L. Zelnik-Manor, and P. Perona. Hybrid models for human motion recognition. In *CVPR*, 2005.