Because better detections are still possible: Multi-aspect object detection with Boosted Hough Forest

Carolina Redondo-Cabrera
carolina.redondoc@edu.uah.es
Roberto López-Sastre
robertoj.lopez@uah.es

University of Alcalá
GRAM
Alcalá de Henares, ES

Supplementary Material

1 PASCAL3D+ Extended Results

We include here the experimental validation of the models for all the object classes included in the *PASCAL3D+* [5] dataset. We report the performance of the HF [2], the ARF [4], and the novel BHF for the problem of object detection and pose estimation.

Table 1: Detection Results on PASCAL3D+ [5]. Results reported in terms of average precision (AP).

Models	aeroplane	bicycle	boat	bottle	bus	car	chair	diningtable	motorbike	sofa	train	tvmonitor	Average
HF	0.290	0.211	0.015	0.070	0.195	0.163	0.021	0.110	0.107	**0.148**	0.126	0.212	0.139
ARF	0.299	0.222	0.023	0.099	0.142	0.182	**0.099**	**0.113**	0.112	0.105	0.145	0.168	0.142
BHF	**0.319**	**0.283**	**0.097**	**0.110**	**0.189**	**0.189**	0.097	0.061	**0.171**	0.134	**0.161**	**0.248**	**0.157**
DPM	0.180	0.360	0.091	0.091	0.207	**0.266**	0.091	0.091	**0.359**	0.091	0.182	0.333	0.195
BHF + verif.	**0.364**	**0.387**	**0.098**	**0.179**	**0.224**	0.265	**0.111**	**0.113**	0.321	**0.159**	**0.251**	**0.352**	**0.235**

Table 1 shows the results achieved by the BHF with and without the verification step described in the paper. Our BHF *without* verification (see row 3 in Table 1) obtains a better AP than the HF and ARF approaches. When we incorporate the verification step to the BHF (Table 1 row 5), we are not loosing accuracy compared to the DPM[1]. Furthermore, we get a gain in terms of AP, with respect to the BHF without verification, of 11.9%. Figure 1 shows the precision/recall curves for all the models and four classes (aeroplane, bicycle, car and motorbike). The BHF casts the highest AP for the class bicycle. The DPM reports better results than the BHF *without* verification only for 7 classes.

For the pose estimation quantitative results, we use the following two metrics: the Average Orientation Similarity (AOS) introduced in [3] (see Table 2) and the Average Viewpoint Precision (AVP) [5] (see Table 3). For the AVP, we fixed a threshold of 15°. This value was suggested by the authors of the PASCAL3D+ dataset. Again, the result reveal that our BHF outperforms the pose estimation performance reported by the ARF and HF (considering both evaluation metrics). The BHF achieves a gain of 5.6% and 4.4% in terms of AOS over the HF and the ARF, respectively. And a gain of 4.8% and 3.1% in terms of AVP.

Figure 2 shows more qualitative results for this dataset.

Table 2: Pose estimation results PASCAL3D+ [5]. Results are reported by the Average Orientation Similarity (AOS).

Models	aeroplane	bycicle	boat	bottle	bus	car	chair	diningtable	motorbike	sofa	train	tvmonitor	Average
HF	0.172	0.160	0.010	0.070	0.117	0.113	0.013	0.081	0.059	**0.139**	0.110	0.105	0.096
ARF	0.212	0.108	0.013	0.099	0.101	0.118	**0.096**	**0.108**	0.065	0.094	0.119	0.161	0.108
BHF	**0.261**	**0.227**	**0.083**	**0.110**	**0.172**	**0.154**	0.095	0.058	**0.150**	0.130	**0.149**	**0.239**	**0.152**

Table 3: Pose estimation results PASCAL3D+ [5]. Results are reported by the Average Viewpoint Precision (AVP).

Models	aeroplane	bycicle	boat	bottle	bus	car	chair	diningtable	motorbike	sofa	train	tvmonitor	Average
HF	0.065	0.021	0.004	0.070	0.041	0.090	0.010	0.010	0.023	0.109	0.080	0.147	0.055
ARF	**0.119**	0.035	0.003	0.099	0.043	0.088	**0.093**	**0.095**	0.020	0.070	0.079	0.122	0.072
BHF	0.105	**0.134**	**0.016**	**0.110**	**0.105**	**0.112**	0.093	0.049	**0.115**	**0.112**	**0.139**	**0.149**	**0.103**

References

[1] P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan. Object detection with discriminatively trained part based models. *PAMI*, 32(9):1627–1645, 2010.

[2] J. Gall, A. Yao, N. Razavi, L. van Gool, and V. Lempitsky. Hough forests for object detection, tracking, and action recognition. In *PAMI*, 2011.

[3] A. Geiger, P. Lenz, and R. Urtasun. Are we ready for autonomous driving? the KITTI vision benchmark suite. In *CVPR*, 2012.

[4] S. Schulter, C. Leistner, P. Wohlhart, P. M. Roth, and H. Bischof. Alternating regression forests for object detection and pose estimation. In *ICCV*, 2013.

[5] Y. Xiang, R. Mottaghi, and S. Savarese. Beyond PASCAL: A benchmark for 3D object detection in the wild. In *WACV*, 2014.

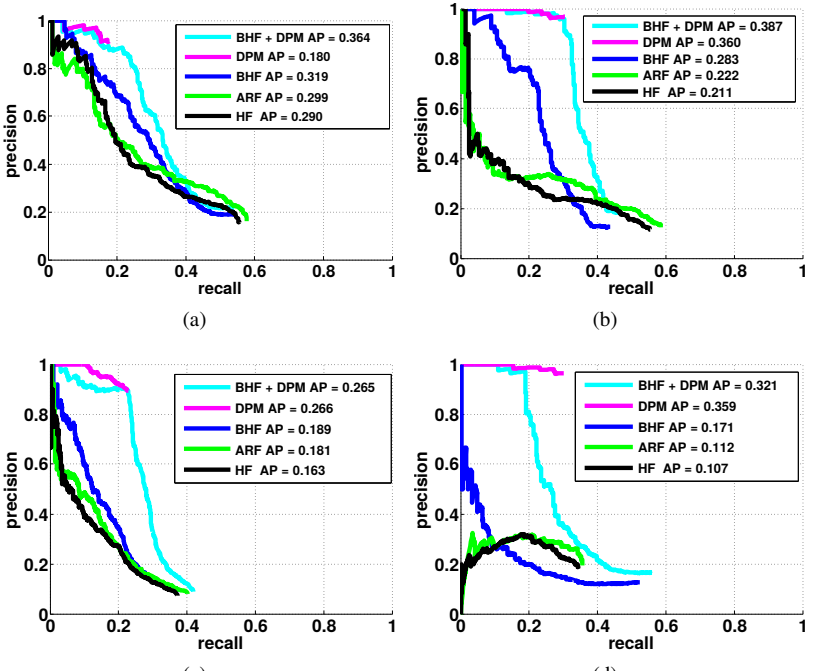

Figure 1: Precision/recall curves obtained by our BHF with and without the verification step, a DPM [1], an ARF [4] and a HF [2] on the PASCAL3D+ [5] dataset for the clases: (a) *aeroplane*, (b) *bicycle*, (c) *car* and (d) *motorbike*.

(a)

(b)

Figure 2: (a) Qualitative results on the PASCAL3D+ dataset [9] for the classes (a) aeroplane and (b) car. Columns 2,4 and 6 show the training images selected to estimate the azimuth and zenith. Ground truth in yellow, estimations in green and wrong detections in red.