Robust Face Detection by Simple Means

Martin Köstinger, Paul Wohlhart, Peter M. Roth, and Horst Bischof

Institute for Computer Graphics and Vision
Graz University of Technology, Austria
{*koestinger,wohlhart,pmroth,bischof*}*@icg.tugraz.at*

1 Motivation

Face detection is still one of the core problems in computer vision, especially in unconstrained real-world situations where variations in face pose or bad imaging conditions have to be handled. These problems are covered by recent benchmarks such as *Face Detection Dataset and Benchmark* (FDDB) [2], which reveals that established methods, e.g, Viola and Jones [8] suffer a drop in performance. More effective approaches exist, but are closed source and not publicly available. Thus, we propose a simple but effective detector that would be available to the public. It combines Histograms of Orientated Gradient (HOG) [1] features with linear Support Vector Machine (SVM) classification.

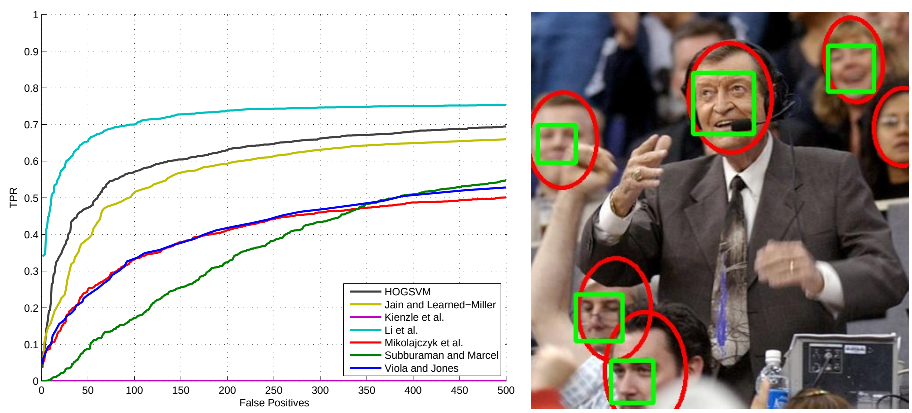
2 Technical Details

One important aspect in the training of our face detector is bootstrapping. Thus, we rely on iterative training. In particular, each iteration consists of first describing the face patches by HOGs [1] and then learning a linear SVM. At the end of each iteration we bootstrap with the preliminary detector hard examples to enrich the training set. We perform several bootstrapping rounds to improve the detector until the desired false positive per window rate is reached. Interestingly, we found out that picking up false positives at multiple scales in a sliding window fashion yields better results than just at a single scale. Testing several patch sizes and HOG layouts revealed that a patch size of 36 by 36 delivers the best results. For the HOG descriptor we ended up with a block size of 12x12, 4x4 for the cells. Prior to the actual training we gathered face crops of the *Annotated facial landmarks in the wild* (AFLW) dataset [4]. As AFLW includes the coarse face pose we are able to retrieve about 28k frontal faces by limiting the yaw angle between $\pm\frac{\pi}{6}$ and mirroring them. For each face we crop a square region between forehead and chin. The non-face patches are obtained by randomly sampling at multiple scales of the PASCAL VOC 2007 dataset, excluding the persons subset.

3 Results

In Figure 1 we report the performance of our final detector on the challenging FDDB benchmark compared to state-of-the-art methods. Despite the simplicity of our detector it is able to improve considerably over the boosted classifier cascade of Viola and Jones [8] and even outperforms the recent work of Jain and

Learned-Miller [3], which adapts a pre-trained classifier by reclassifying hard examples near the decision boundary at test time. Only the work of Li [5], which uses a boosted classifier cascade and SURF features, improves over our results. Moreover, we successfully applied our detector in several applications ranging from PTZ surveillance to the processing of news broadcasts. Even though the detector is ready to be used on a GPU, future work is concerned with speed issues. In particular, we aim at reducing the number of needed feature computations, e.g., by using approximated responses at nearby scales.



(a) (b)

Fig. 1: FDDB benchmark. In (a) we report ROC curves for [8, 5, 3, 7, 6] and our method (HOG/SVM). In (b) we provide an illustrative detection example. The red ellipses denote the FDDB ground truth, whereas the green rectangles are the respective detector outputs.

References

1. Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: Proc. CVPR (2005)
2. Jain, V., Learned-Miller, E.: FDDB: A benchmark for face detection in unconstrained settings. Tech. Rep. UM-CS-2010-009, UMASS, Amherst (2010)
3. Jain, V., Learned-Miller, E.G.: Online domain adaptation of a pre-trained cascade of classifiers. In: Proc. CVPR (2011)
4. Köstinger, M., Wohlhart, P., Roth, P.M., Bischof, H.: Annotated facial landmarks in the wild: A large-scale, real-world database for facial landmark localization. In: Benchmarking Facial Image Analysis Technologies (ICCV Workshop) (2011)
5. Li, J., Wang, T., Zhang, Y.: Face detection using surf cascade. In: Benchmarking Facial Image Analysis Technologies (ICCV Workshop) (2011)
6. Mikolajczyk, K., Schmid, C., Zisserman, A.: Human detection based on a probabilistic assembly of robust part detectors. In: Proc. ECCV (2004)
7. Subburaman, V.B., Marcel, S.: Fast Bounding Box Estimation based Face Detection. In: ECCV, Workshop on Face Detection (2010)
8. Viola, P., Jones, M.J.: Rapid object detection using a boosted cascade of simple features. In: Proc. CVPR (2001)