

Real-time Human Detection based on Personness Estimation

Supplementary Material

Kyuwon Kim^{1,2}
q1.kim@samsung.com

¹ School of Electrical and Electronic
Engineering
Yonsei University
Republic of Korea

Kwanghoon Sohn¹
khsohn@yonsei.ac.kr

² Mobile Division
Samsung Electronics
Republic of Korea

In this supplementary material, we show experiment results for grid search over the patch size (*psize*) and the aspect-ratio threshold (*arthres*) on VOC 2012 dataset [1], and discuss their meaning. All recall-time curves are obtained from the DPM detection results based on personness estimation.

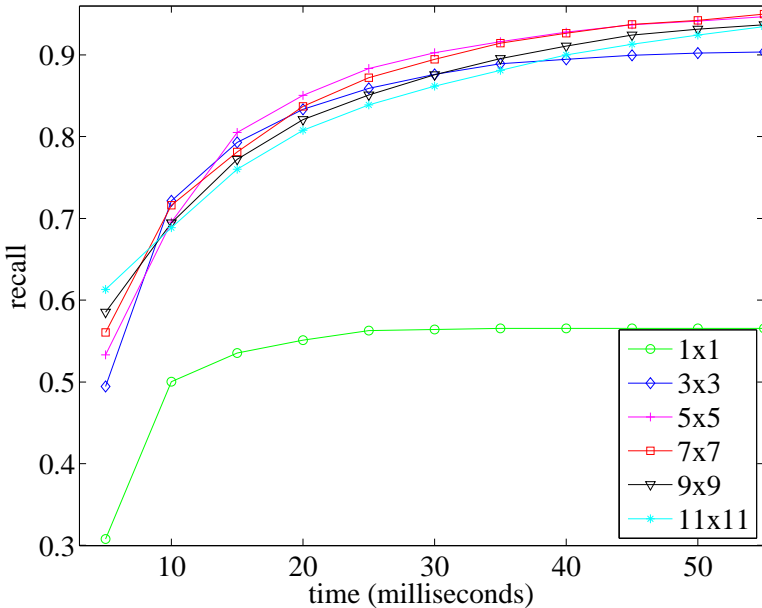


Figure 1: Recall curves according to the changes of the patch size (arthres = 1.54)

Patch size for pinpoint Patch size 1×1 poorly performs as shown in Fig. 1. At this size,

the DPM detector consumes all candidate windows at about 25ms since the DPM detector performs convolutions only for pinpoints. As soon as size 3×3 is adopted, the performance greatly increases. And as we increase the patch size further, detection rate slightly increases. But, from size 9×9 , detection rate decreases. The performances of size 5×5 and size 7×7 are similar. At 5ms and 10ms, size 7×7 is slightly better than size 5×5 . In other time limits, size 5×5 is better.

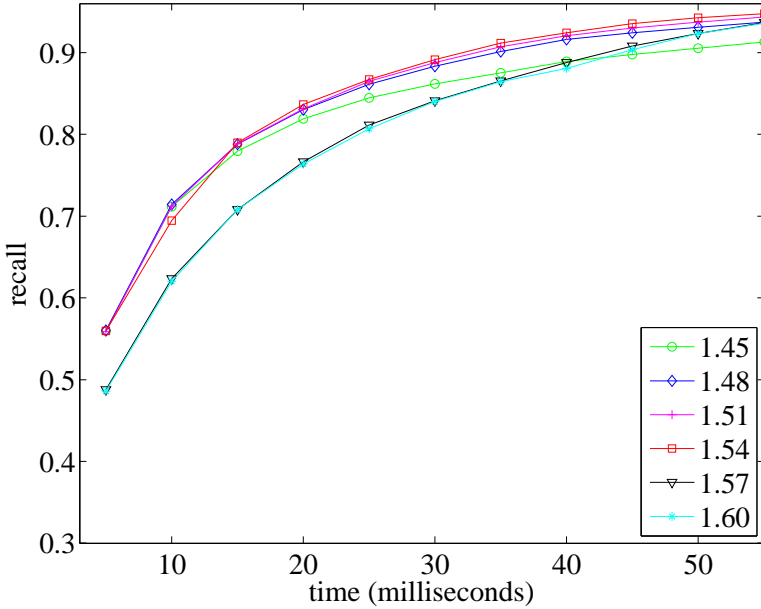


Figure 2: Recall curves according to the changes of the aspect-ratio threshold (psize = 7×7)

Aspect-ratio threshold By decreasing the aspect-ratio threshold, we can ignore the irrelevant candidate windows. As can be seen in Fig. 2, the detection performances decrease if the aspect-ratio threshold is greater than 1.57. The reason of the degradation is that the DPM detector considers too many irrelevant candidate windows in its initial stage. Threshold 1.54 shows the best performance in general. Threshold 1.51 also shows good detection performance in a very short period of time, especially at the 10ms time limit. However, decreasing threshold further degrades the detection performance since too many windows are discarded.

References

- [1] M. Everingham, V. Gool, L. Williams, C. K. I., J. Winn, and A. Zisserman. The pascal visual object classes (VOC) challenge. *International Journal of Computer Vision*, 88: 303–338, Jun. 2010.