Learning SURF Cascade for Fast and Accurate Object Detection Supplementary materials

Jianguo Li, Yimin Zhang
Intel Labs China

1. Spatial Cell Configuration

Figure 1 depicts spatial configuration of SURF patches.

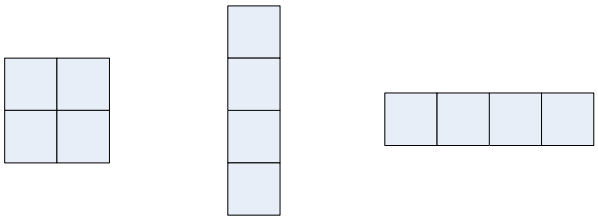


Figure 1. Possible spatial cell configuration of local patches. Each local patch has 4 same-size cells, but these cells may be in different forms. From left to right, 2×2 cells, 4×1 cells, 1×4 cells.

2. L_2HYS Feature Normalization

Given a feature vector $\mathbf{v} = (v_1, \ldots, v_d)$, the L_2Hys normalization works like below

(1) L_2-normalization: $u_i = v_i/\sqrt{\|\mathbf{v}\|_2^2 + \epsilon}$, where ϵ is small positive value to avoid dividing by zero;

(2) Clipping u_i with the following rule

$$u_i = \begin{cases} \theta & \text{if } u_i > \theta \\ -\theta & \text{if } u_i < -\theta \\ u_i & \text{otherwise,} \end{cases}$$

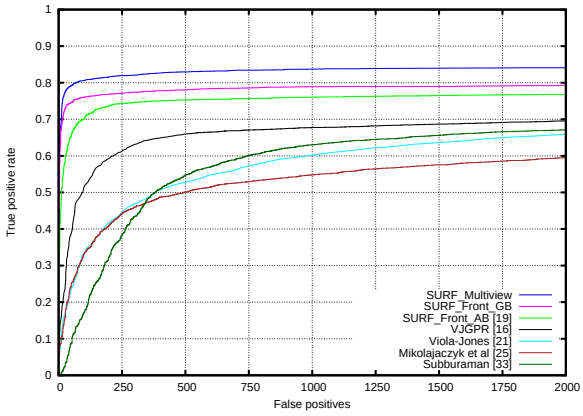
where $\theta = 2/\sqrt{d}$ empirically [1].

(3) Re-normalization: $v_i' = u_i/\sqrt{\|\mathbf{u}\|_2^2 + \epsilon}$, and $\mathbf{v}' = (v_1', \ldots, v_d')$ is the L_2Hys normalization result.

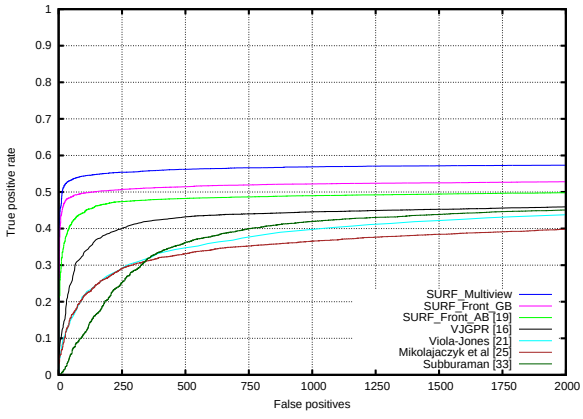
3. Results on UMass FDDB

Figure 2 depicts ROC curve from both discrete score and continuous score on FDDB benchmark by SURF cascade.

[1] After normalization, assuming $|u_i| < \theta$, thus $\sum u_i^2 < d\theta^2$. u_i can be viewed as samples for a Gaussian variable, the variance $\sigma^2 = (\sum u_i^2)/d = 1/d < \theta^2$. As is known, about 95% samples from the Gaussian distribution fall within the range $[-2\sigma, 2\sigma]$. Therefore, we define $\theta = 2\sigma = 2/\sqrt{d}$.



(a)



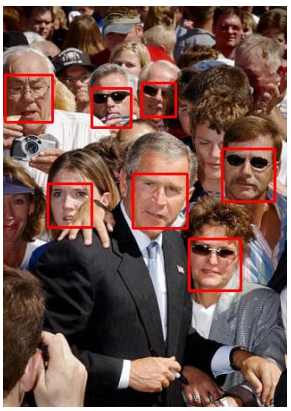
(b)

Figure 2. (a) Discrete score ROC curves and (b) Continuous score ROC curves for different methods on UMass FDDB dataset.

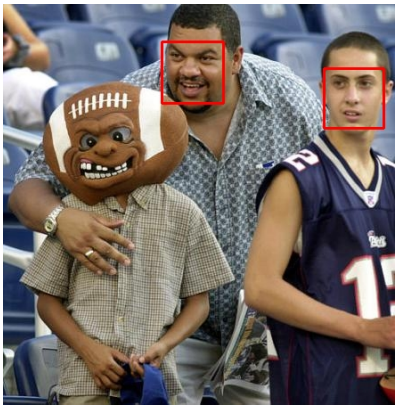
4. Example Results of Face and Car Detection

Figure 3 depicts some face detection results on FDDB and CMU+MIT datasets. Figure 4 depicts some car detection results on TUGRAZ dataset.

(a)



(b)

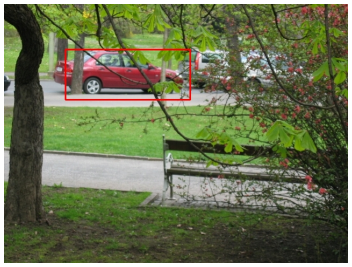


(c)

(d)

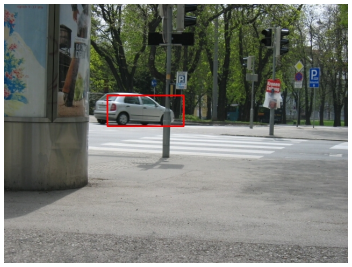
Figure 3. Example detection results on CMU+MIT (a) and UMass FDDB datasets (b,c,d).

(a)



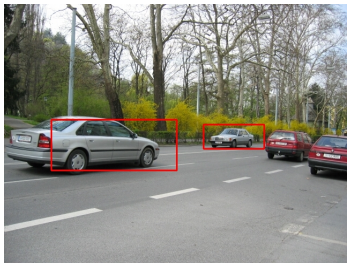
(b)

(c)



(d)

(e)



(f)

Figure 4. Some car detection results on TUGRAZ dataset.