Another Descriptor

Histograms of Oriented Gradients for Human Detection

Navneet Dalal and Bill Triggs

CVPR 2005

Overview

1. Compute gradients in the region to be described
2. Put them in bins according to orientation
3. Group the cells into large blocks
4. Normalize each block
5. Train classifiers to decide if these are parts of a human

Details

- Gradients

  $[-1\ 0\ 1]$ and $[-1\ 0\ 1]^T$ were good enough.

- Cell Histograms

  Each pixel within the cell casts a weighted vote for an orientation-based histogram channel based on the values found in the gradient computation. (9 channels worked)

- Blocks

  Group the cells together into larger blocks, either R-HOG blocks (rectangular) or C-HOG blocks (circular).

More Details

- Block Normalization

They tried 4 different kinds of normalization.
Let υ be the block to be normalized and e be a small constant.

L2-norm: $f = \frac{v}{\sqrt{\|v\|_2^2 + e^2}}$

L2-hys: L2-norm followed by clipping (limiting the maximum values of v to 0.2) and renormalizing,

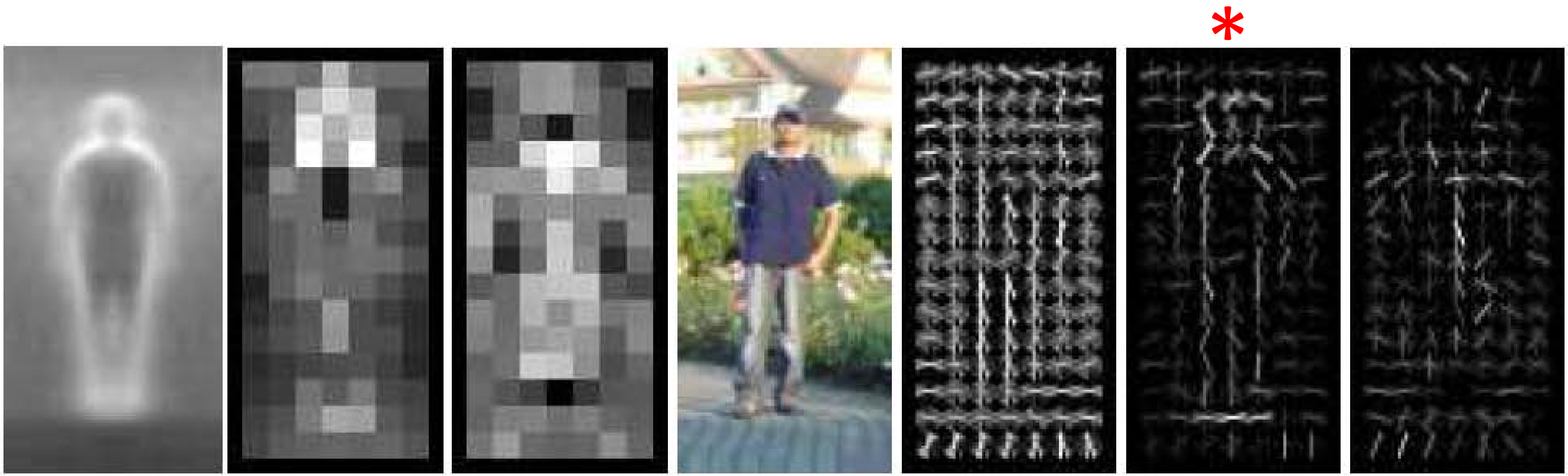
L1-norm: $f = \frac{v}{(\|v\|_1 + e)}$

L1-sqrt: $f = \sqrt{\frac{v}{(\|v\|_1 + e)}}$

R-HOG compared to SIFT Descriptor

- R-HOG blocks appear quite similar to the SIFT descriptors.
- But, R-HOG blocks are computed in dense grids at some single scale without orientation alignment.
- SIFT descriptors are computed at sparse, scale-invariant key image points and are rotated to align orientation.

Pictorial Example



(a) (b) (c) (d) (e) (f) (g)

(a) average gradient image over training examples
(b) each “pixel” shows max positive SVM weight in the block centered on that pixel
(c) same as (b) for negative SVM weights
(d) test image
(e) its R-HOG descriptor
(f) R-HOG descriptor weighted by positive SVM weights
(g) R-HOG descriptor weighted by negative SVM weights