Integral Channel Features

**Idea:** Generate and compute features efficiently using integral images over multiple registered image channels.

**Key properties:**
- Naturally integrate heterogeneous information
- Few parameters and insensitive to exact settings
- State of the art performance
- Fast to compute
- More accurate spatial localization

**Channels and Feature Types:**
- A channel $C$ is a registered image map of $I$: $C = \Omega(I)$
- Restrict features to local rect. sums over $C$: $f(C) = \sum_x \sum_y C(x, y)$
- Require translation invariance of $C$: $f(\Omega(I)) = f'(\Omega(I'))$
- Grayscale: original Viola framework; Haar-like features [Viola04] ...
- Color and linear filters: RGB, LUV, DoG, Gabor [Tu05, Dollar06] ...
- Nonlinear transforms: Canny edges, thresholding, [Dollar06] ...
- Integral/Gradient Histograms: [Porikli05, Zhu06, Laptev06, Zhang08], ...
- Haar-like features [Viola04], generalized Haar features [Dollar07], ...
- Point-wise transforms: log, power; allow for local product, max ...

**Scale:**
- Pre-smoothing (local scale)
- Post-smoothing (integration scale)
- Binomial filtering “Gaussian with $\sigma = \sqrt{\frac{1}{4}}$”

**Computation time:**
- LUV color channels: 135 fps
- Gradient Magnitude: 60 fps
- Gradient histogram channels (6): 60 fps
- LUV + grad-mag + grad-hist + pre-smoothing: 30-34 fps

**Learning Framework**

**Random Features:**
- Random combinations of weighted rectangles [can span multiple channels]
- Large pool of candidate features [O(n^k) possibilities for k rectangles in nxn image]
- Random features work well in practice [Dollar07]

**Boosting:**
- Standard method: soft cascades [Zhang&Viola07]
- Fast: 20k examples x 5k features: 5-10m training
- Weak classifier: depth 2 decision tree

**Details:**
- Channels: LUV, grad-mag, grad-hist (6), pre-smoothing r=1
- Two rounds of bootstrapping [15k negatives + 5k positives total]
- Boosted 2000 weak classifiers from 30,000 random features
- Full-Image Detection: Spatial Stride: 4 pixels; Scale Stride: $2^{1/6}$
- Non-maximal Suppression: pairwise-max w thr=0.6

**Future Work**

- Merge with more sophisticated learning method
- Multiple Component Learning [Dollar08]
- Motion and other channels (e.g. [Wojek09])
- Automatic channel selection / discovery
- Demonstrate on other domains:
  - mouse tracking by detection, learned edge detection [Dollar06]

**Effectiveness for Pedestrian Detection**

**Framework Evaluation**

**Integral Channel Features**

1. Electrical Engineering, California Institute of Technology
   pdollar.perona@caltech.edu
2. Lab of Neuro Imaging, University of California, Los Angeles
   zhuxu@loni.ucla.edu
3. Computer Science and Engineering, University of California, San Diego
   sjb@cs.ucsd.edu