# Histograms of oriented gradients

2014/08/22

ked

# Authors

- Navneet Dalal and Bill Triggs

# Authors

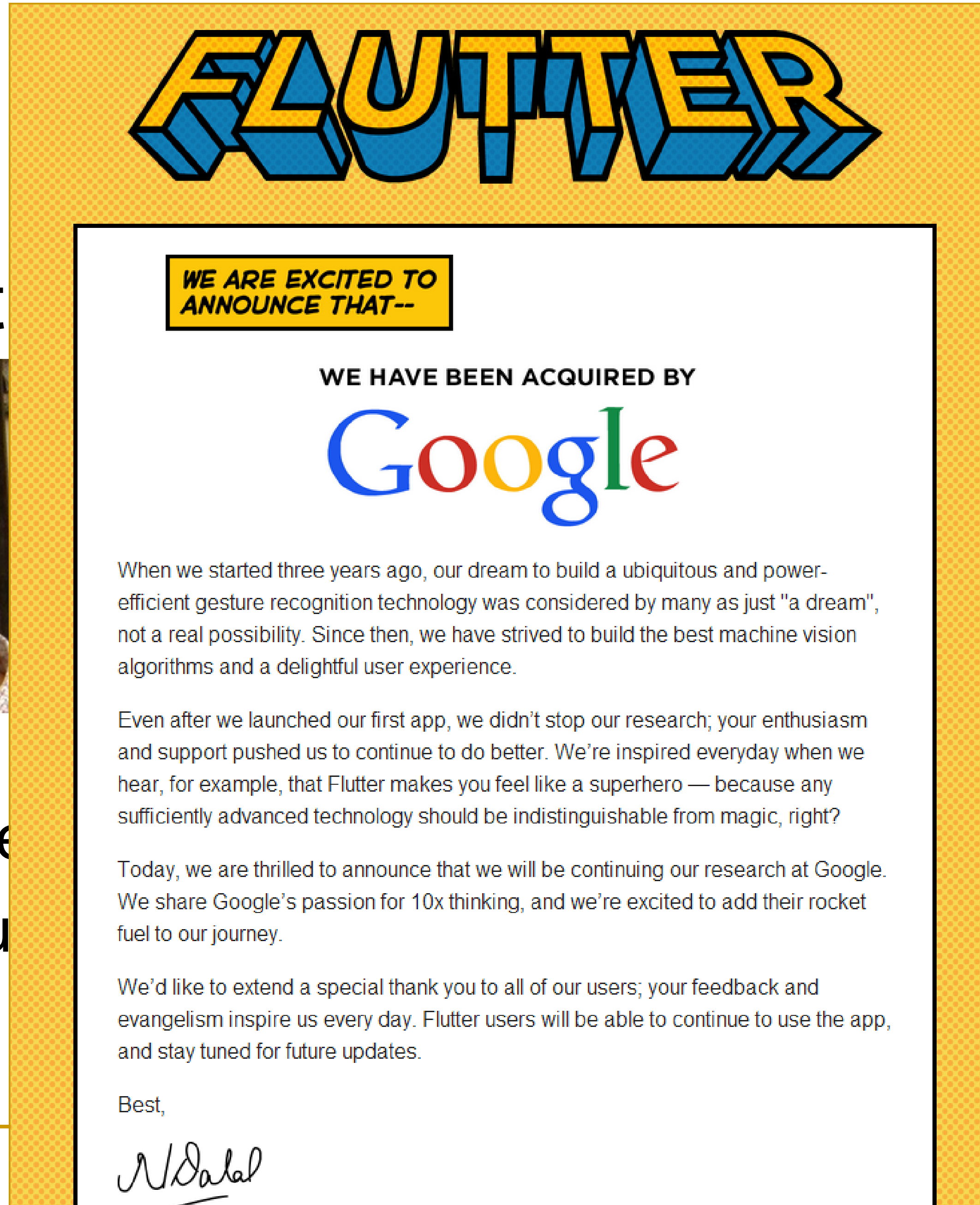
- Navneet Dalal

  - A founder of Flutter
  - (A gesture recognition startup company created in 2010)

# Authors

- Navneet



FLUTTER

WE ARE EXCITED TO ANNOUNCE THAT--

WE HAVE BEEN ACQUIRED BY

Google

When we started three years ago, our dream to build a ubiquitous and power-efficient gesture recognition technology was considered by many as just "a dream", not a real possibility. Since then, we have strived to build the best machine vision algorithms and a delightful user experience.

Even after we launched our first app, we didn't stop our research; your enthusiasm and support pushed us to continue to do better. We're inspired everyday when we hear, for example, that Flutter makes you feel like a superhero — because any sufficiently advanced technology should be indistinguishable from magic, right?

Today, we are thrilled to announce that we will be continuing our research at Google. We share Google's passion for 10x thinking, and we're excited to add their rocket fuel to our journey.

We'd like to extend a special thank you to all of our users; your feedback and evangelism inspire us every day. Flutter users will be able to continue to use the app, and stay tuned for future updates.

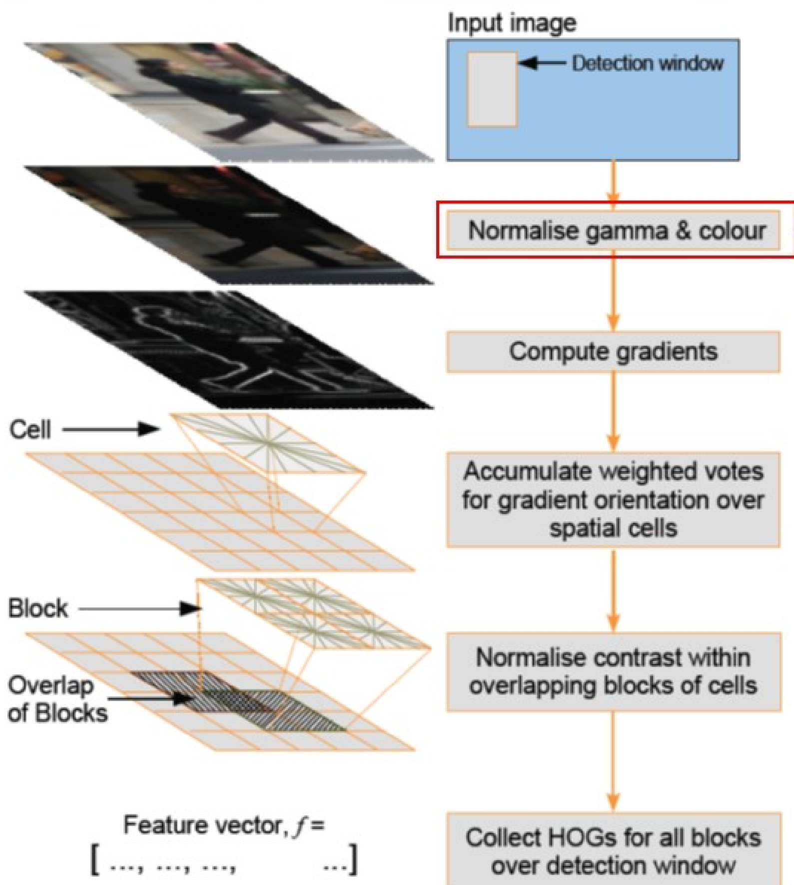
Best,

NDalal

- A founde
- (A gestu ... 10)

# Overview – HOG

## Typical person detection scheme using SVM



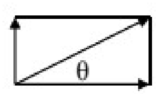

In practice, effect is very small (about 1%) while some computational time is required*

*Navneet Dalal and Bill Triggs. Histograms of Oriented Gradients for Human Detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, SanDiego, USA, June 2005. Vol. II, pp. 886-893.

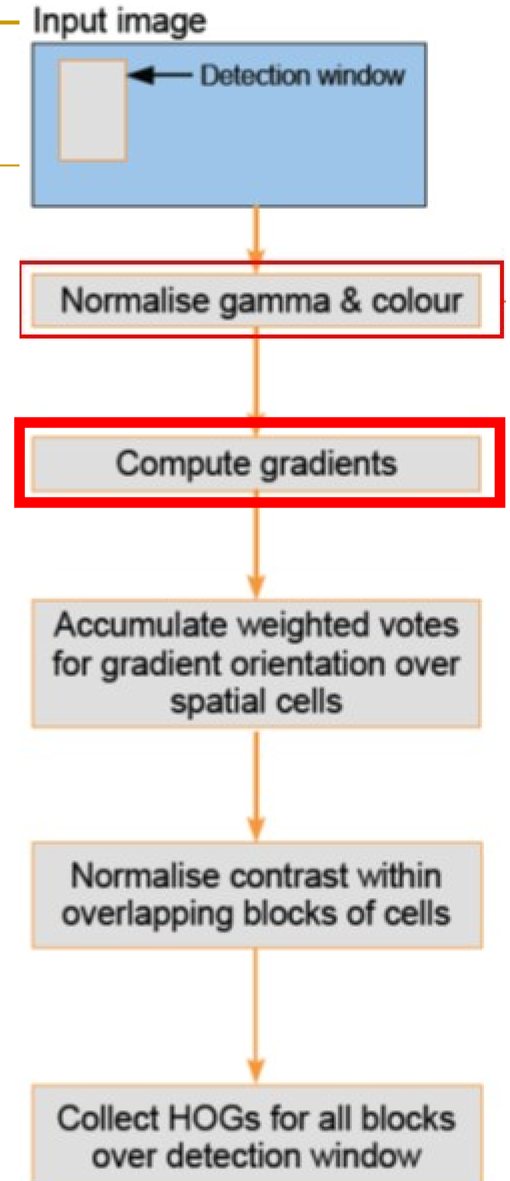
# Gradient

## Computing Gradients

- Centered: $f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x-h)}{2h}$
- Filter masks in x and y directions
  - Centered: $\begin{bmatrix} -1 & 0 & 1 \end{bmatrix}$ $\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$
- Gradient
  - Magnitude: $s = \sqrt{s_x^2 + s_y^2}$

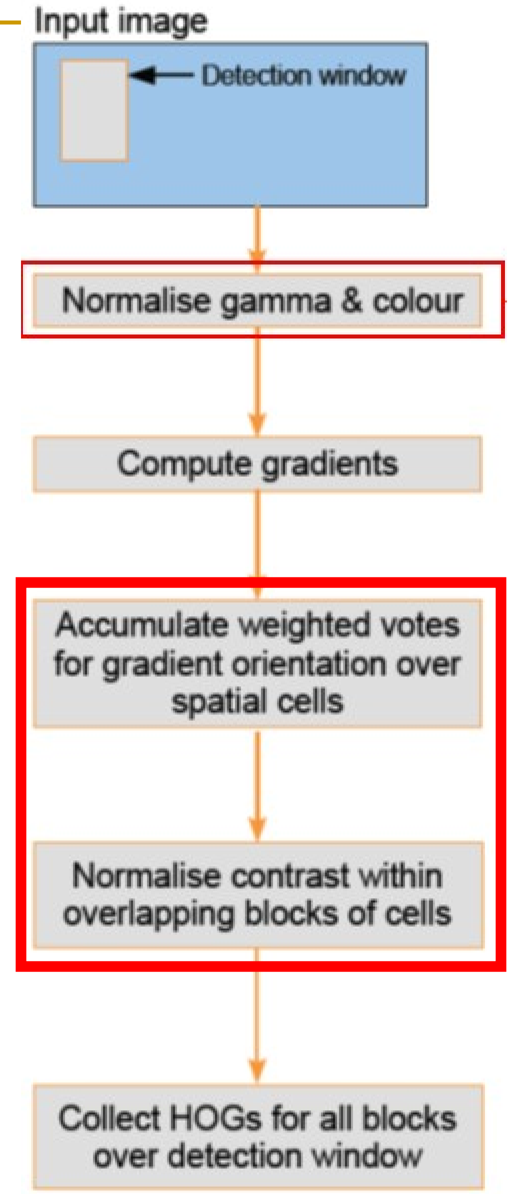



  - Orientation: $\theta = \arctan(\frac{s_y}{s_x})$



# Cell & block

- Cell
  - 8 x 8 pixels
- Block
  - 2 x 2 cells
  - 50% overlap for neibhbors

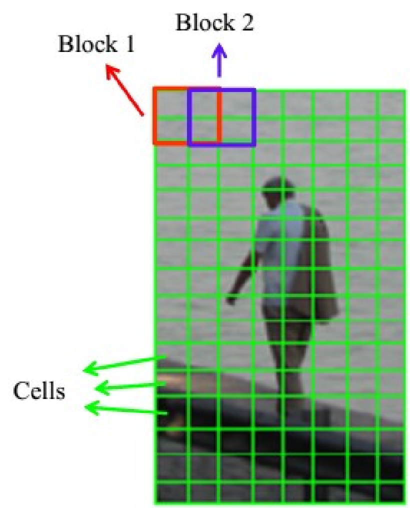


# Example

- For a 64 x 128 image

## Blocks, Cells

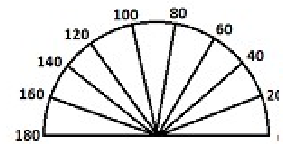
- 16x16 blocks of 50% overlap.
  - 7x15=105 blocks in total
- Each block should consist of 2x2 cells with size 8x8.



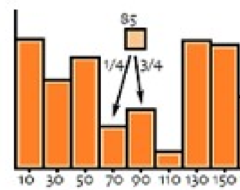
# Quantization

## Tri-linear Interpolation

- Each block consists of 2x2 cells with size 8x8
- Quantize the gradient orientation into 9 bins (0-180)

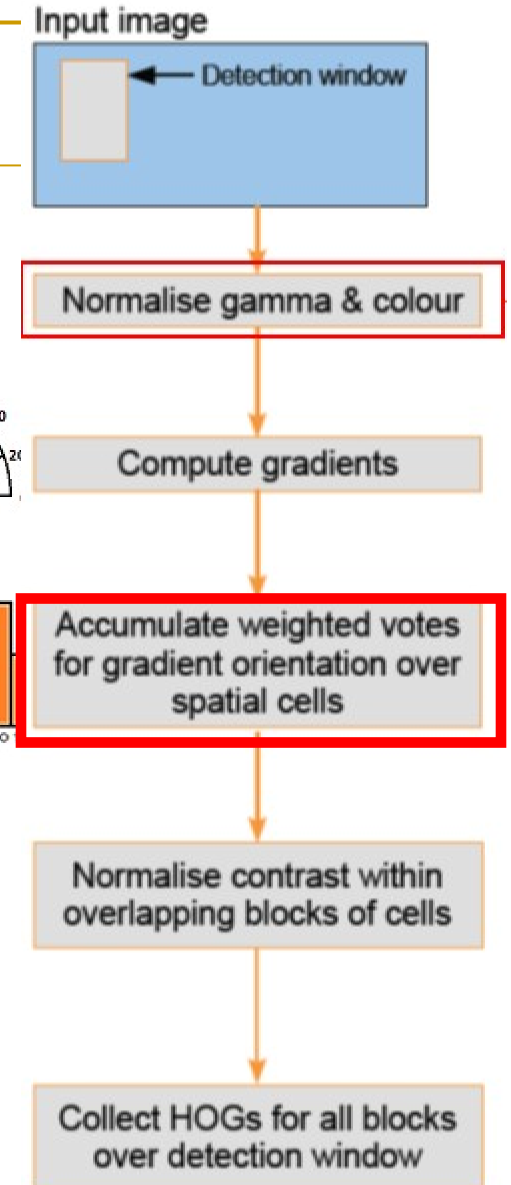



9 Bins



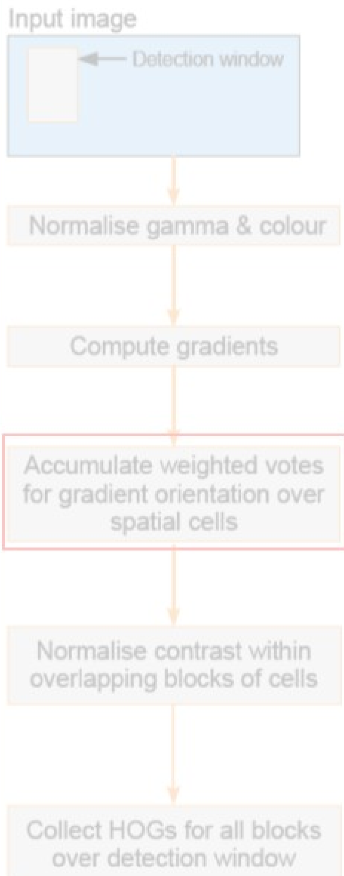

Bin centers

  - The vote is the gradient magnitude
  - Interpolate votes linearly between neighboring bin centers.
    - Example: if θ=85 degrees.
    - Distance to the bin cente Bin 70 and Bin 90 are 15 and 5 degrees, respectively.
    - Hence, ratios are 5/20=1/4, 15/20=3/4.
  - The vote can also be weighted with Gaussian to downweight the pixels near the edges of the block.

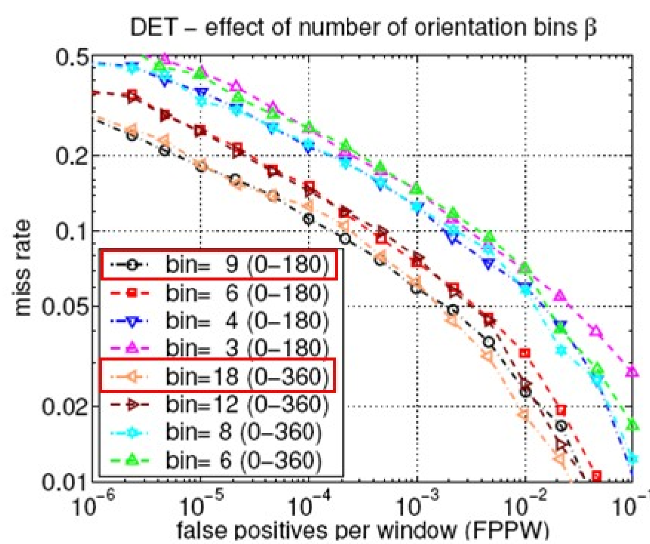


# Parameters – bins

## Accumulate weight votes over spatial cells

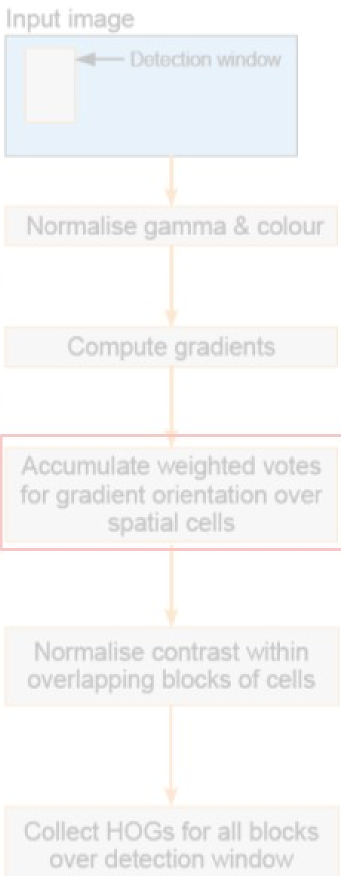



- How many bins should be in histogram?
- Should we use oriented or non-oriented gradients?
- How to select weights?
- Should we use overlapped blocks or not? If yes, then how big should be the overlap?
- What block size should we use?

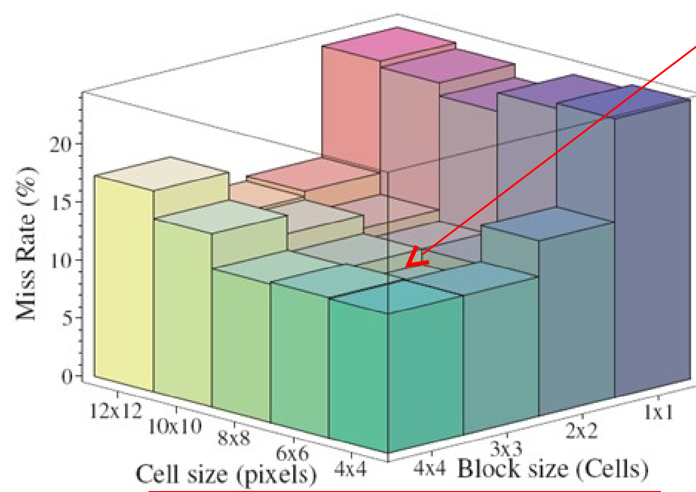


# Parameters – cell & blcok

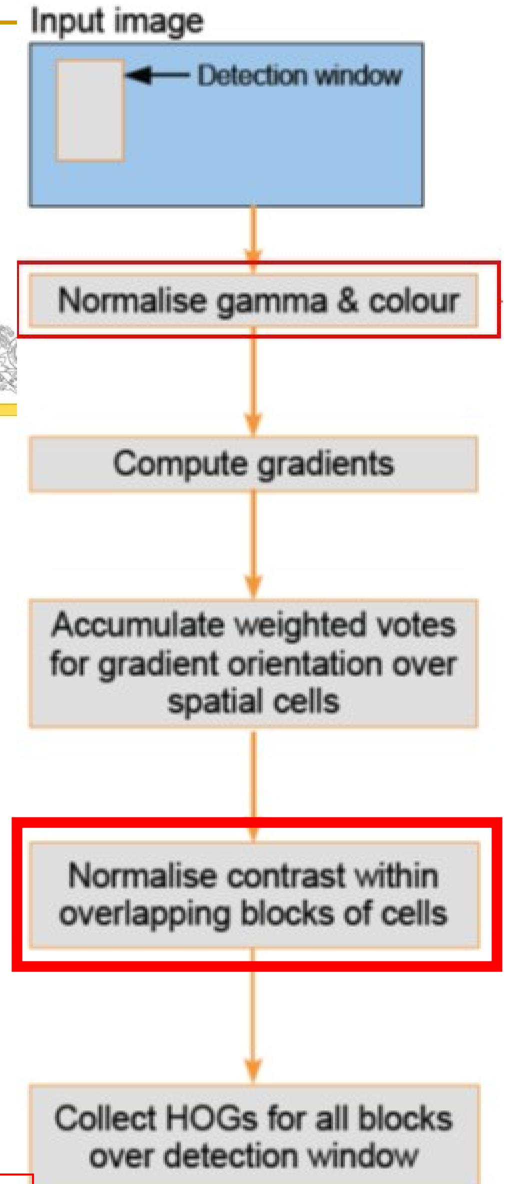
## Accumulate weight votes over spatial cells



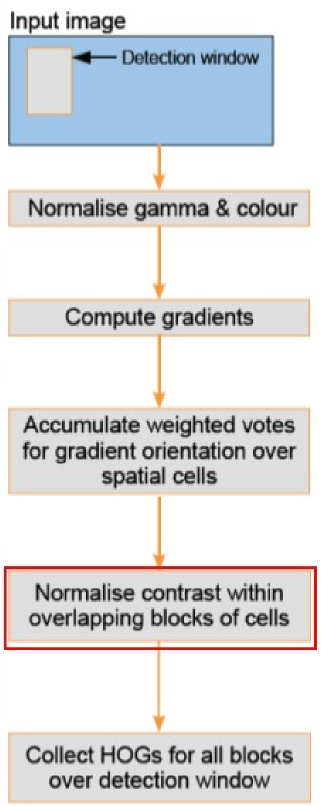

- How many bins should be in histogram?
- Should we use oriented or non-oriented gradients?
- How to select weights?
- Should we use overlapped blocks or not? If yes, then how big should be the overlap?
- What block size should we use?

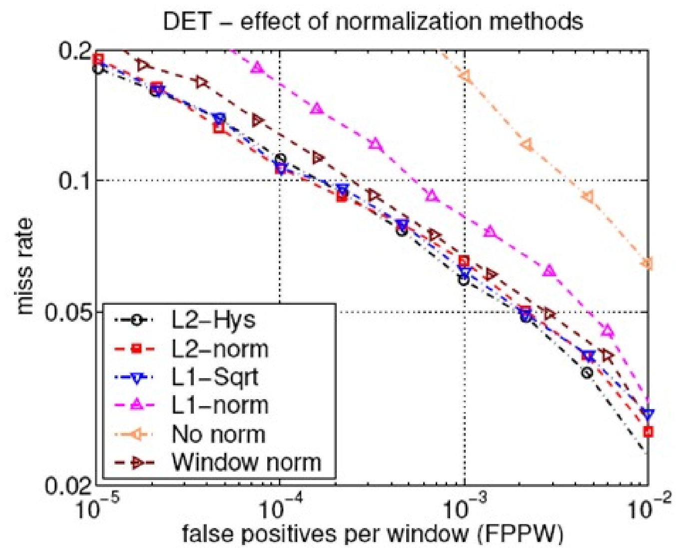


# Histogram normalization




## Contrast normalization

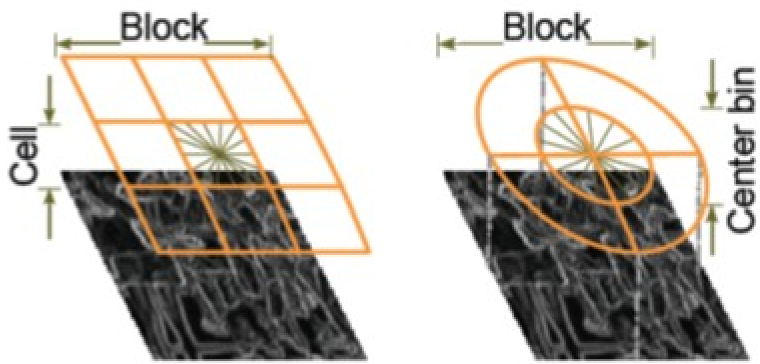


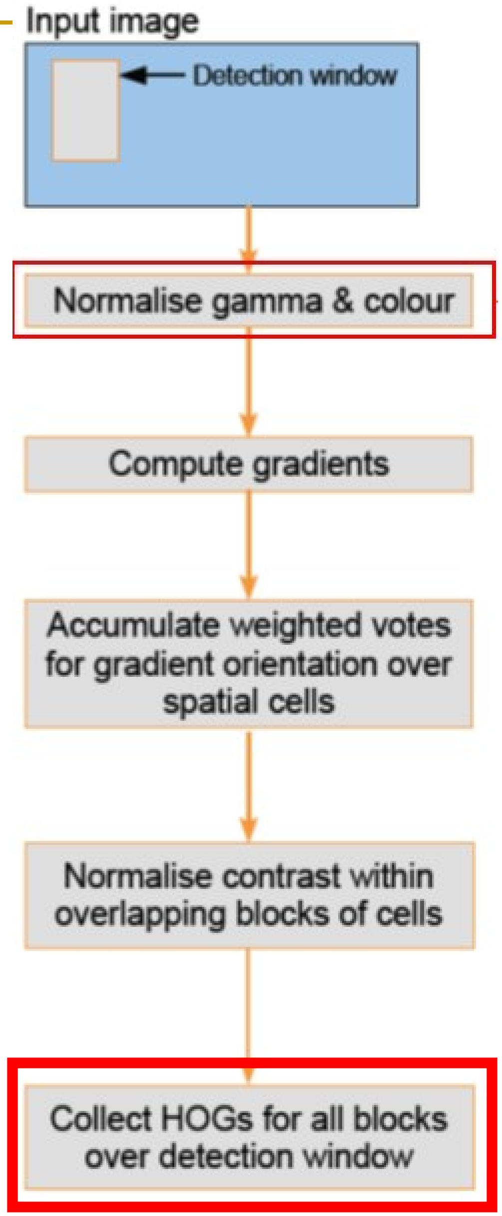

$$L1-norm=\frac{v}{\|v\|_1+\varepsilon} \qquad L1-sqrt=\sqrt{\frac{v}{\|v\|_1+\varepsilon}} \qquad L2-norm=\frac{v}{\sqrt{\|v\|_2^2+\varepsilon}}$$

$L2-Hys$ - L2-norm followed by clipping (limiting the maximum values of v to 0.2) and renormalising

# Feature description




(a) R-HOG/SIFT (b) C-HOG (c) Single centre C-HOG



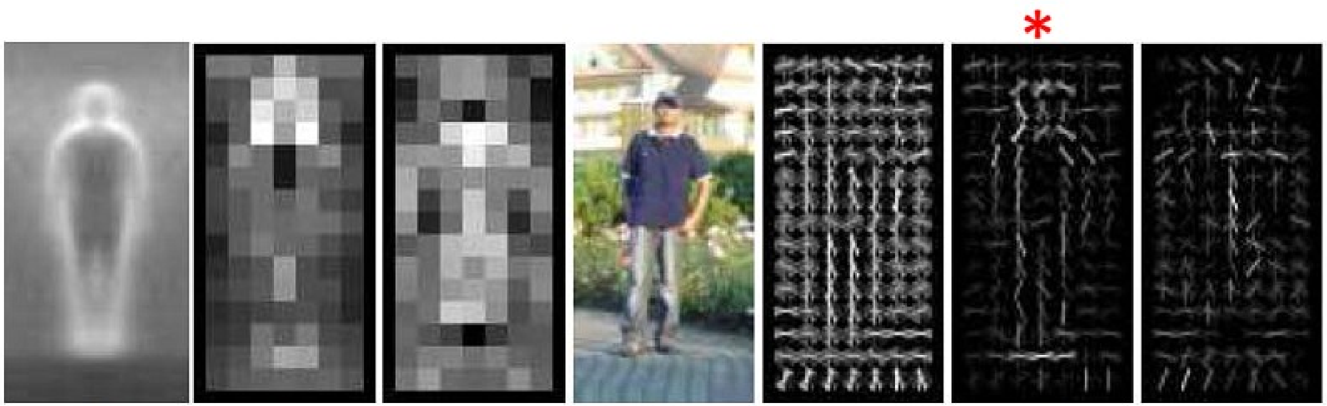
# Compare to SIFT
# -- by authors

## 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.

# SVM weight

## 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

# Thx.