HOGgles
Visualizing Object Detection Features

C. Vondrick, A. Khosla, T. Malisiewicz, A. Torralba
ICCV, 2013.

presented by
Ezgi Mercan
Object Detection Failures

- Why do our detectors think water looks like a car?
  - Data set?
  - Machine learning method?
  - Features?

A high scoring car detection from DPM.
Object Detection Failures

Object Categories

- Aeroplane
- Bicycle
- Bird
- Boat
- Bottle
- Bus
- Car
- Cat
- Chair
- Cow
- Table
- Dog
- Horse
- Motorbike
- Person
- Potted plant
- Sheep
- Sofa
- Train
- TV/monitor
Reconstructing an Image

Original image → copied patches → elliptic region of interest → affine normalization to square patch → local descriptor

calculate SIFT

Histogram of Oriented Gradients

Histograms of oriented gradients for human detection, Dalal, N., Triggs, B., CVPR’05.
Algorithms

Image: \( x \in \mathbb{R}^D \)

HOG descriptor: \( y = \emptyset(x) \in \mathbb{R}^d \)

HOG inverse: \( \emptyset^{-1}(y) \)

- Exemplar LDA

\[
\emptyset_A^{-1}(y) = \frac{1}{K} \sum_{i=1}^{K} z_i
\]

- Ridge Regression

\[
\emptyset_B^{-1}(y) = \Sigma_{XY} \Sigma_{YX}^{-1}(y - \mu_Y) + \mu_X
\]
Algorithms

• Direct Optimization

Image: \( x \in \mathbb{R}^D \)
HOG descriptor: \( y \in \mathbb{R}^d \)
Image basis: \( U \in \mathbb{R}^{D \times K} \)
Coefficients: \( \rho \in \mathbb{R}^K \); \( x = U \rho \)

\[ \phi_c^{-1}(y) = U \rho^* \]

\[ \rho^* = \arg \min_{\rho \in \mathbb{R}^K} \| \phi(U \rho) - y \|_2^2 \]
Algorithms

• Paired Dictionary Learning

Image: \( x \in \mathbb{R}^D \)
HOG descriptor: \( y \in \mathbb{R}^d \)
Image basis: \( U \in \mathbb{R}^{D \times K} \)
HOG basis: \( V \in \mathbb{R}^{d \times K} \)
Coefficients: \( \alpha \in \mathbb{R}^K \); \( x = U \alpha \) and \( y = V \alpha \)

\[
\phi_D^{-1}(y) = U \alpha^*
\]

\[
\alpha^* = \arg\min_{\alpha \in \mathbb{R}^K} \| V\alpha - y \|_2^2 \quad \text{s.t.} \quad \| \alpha \|_1 \leq \lambda
\]
Paired Dictionary Learning

\[ y = \phi_D^{-1}(y) \]

\[ y = \alpha_1 + \alpha_2 + \ldots + \alpha_k \]

HOG feature

HOG Basis

Image Basis

HOG Inversion
Paired Dictionary Learning

• Solving paired dictionary learning problem:

$$\arg\min_{U,V,\alpha} \sum_{i=1}^{N} (\|x_i - U a_i\|_2^2 + \|\phi(x_i) - V a_i\|_2^2)$$

s.t.  \(\|a_i\|_1 \leq \lambda \ \forall i, \quad \|U\|_2^2 \leq \gamma_1, \quad \|V\|_2^2 \leq \gamma_2\)

some learned pairs of dictionaries for \(U\) and \(V\)
Feature Visualization
Evaluation of Performance

- ELDA
- Ridge
- Direct
- PairDict

10/11/2013
CSE590V
Evaluation of Visualization Performance

- ELDA
- Ridge
- Direct
- PairDict
- Glyph
- Expert

Mean: 0.68

10/11/2013
CSE590V
Deformable Parts Model

Does HOG capture color?

Potted Plant

Chair

Computers can see better than us
Questions?

Visit [web.mit.edu/vondrick/ihog/](web.mit.edu/vondrick/ihog/) for more cool stuff.