Universitatea Politehnica București

Facultatea de Automatică și Calculatoare

Catedra de Calculatoare

Detecting Cars in Images Using Histogram of Oriented Gradients

Research Area, Fall 2013 - 2014

Autor

Vlad Păunescu
vlad.paunescu@cti.pub.ro

Conducători științifici

Prof. Dr. Ing. Marius Leordeanu

Agenda

- Introduction
- Motivation
- Theory
- Results
- Conclusion and Future Work
- Questions

Computer Vision

- Field of Artificial Intelligence
- Acquiring, processing;
- Analyzing and understanding of images
- Extracting information from images
- Reconstructing the model that generated the image – inverse of Computer Graphics

Object Category Recognition

- Detecting objects in images
- Classifying detected objects
- Objects at different sizes(scale)
- Different view angles
- Different classes of objects

Motivation

- Humans detect a multitude of objects
- No effort
- Computer Vision aims to ahieeve and exceed human vision
- Still a challange in computer vision
- Many categories (classes) of objects

A Machine Learning Approach

- Idea – capture somehow the features of one category of objects from many labeled images;
- Construct a model;
- Being able to identify new instances in unseen images

HOG

- Histogram of Oriented Gradients
- The HOG person detector was introduced by Dalal and Triggs at the CVPR conference in 2005.

Oriented Gradients

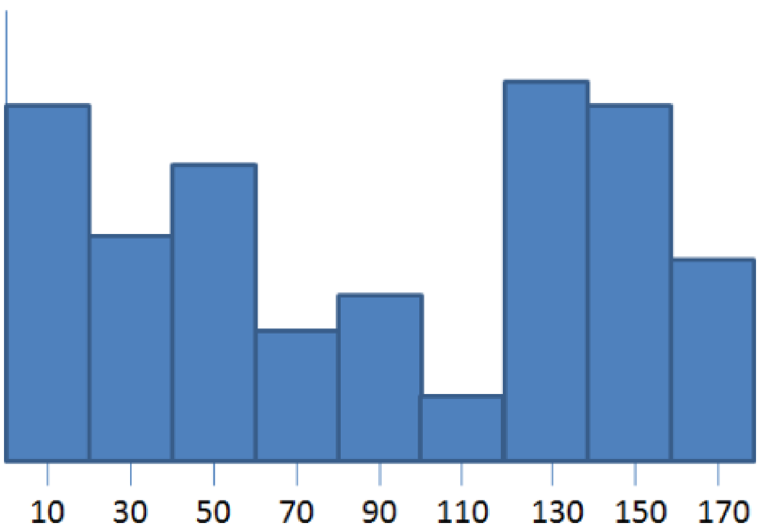
- At every point compute Gradient magnitude and orientation over x and over y. (Convolutuion with [-1, 0, 1]).
- [Gx , Gy]
- Magnitude = $\sqrt{Gx^2 + Gy^2}$
- Angle = atan2(Gy,Gx)

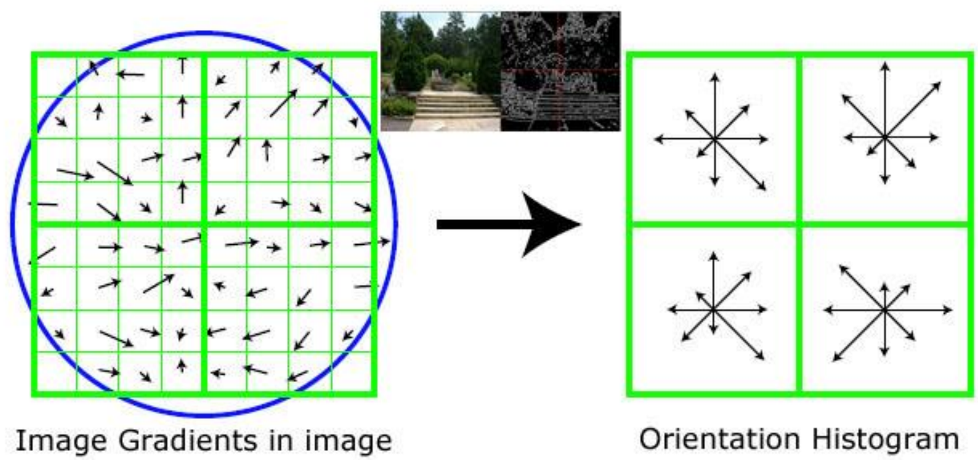
Change in x-direction

Change in y-direction

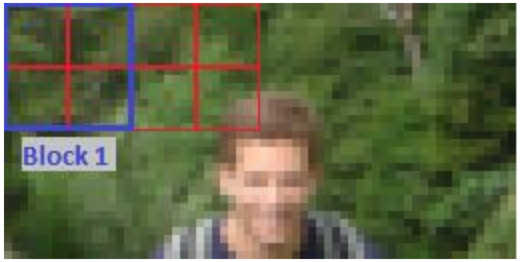
Binning (Histogram)

- Split image in cells. Compute gradients.
- For each cell build a histogram – binning.
- Split interval 0 – 180 into bins.

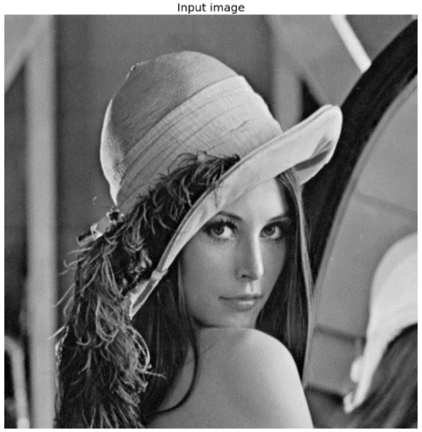




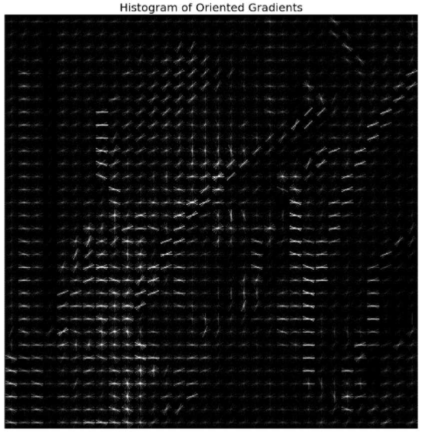
HOG



Input image



Histogram of Oriented Gradients



HOG Training set

- Example, split [0...180] into bins of 20 degs => b = 9 bins.
- Each HOG h_i has bin size b.
- There is some degree ov overlapping.
- Linearize $h_i = [a_1, a_2 \dots a_b]$
- Concatenate each hog in the image vector X_i
- $X_i = [h_1, h_2, \dots, h_n]$
- N = # vertical cells x # horizontal cells

HOG Training set

- $X_i = [h_1, h_2, \dots, h_n]$
- N = # vertical cells x # horizontal cells
- M images. Positive and Negative
- Training set = $\begin{bmatrix} X_1 & 1 \\ \vdots & \vdots \\ X_m & 0 \end{bmatrix}$
- Usually a linear model SVM is trained – best results.

Results and Experiments

- Used Caltech Cars (Rear) Dataset
- 126 images of cars from the rear
- Background dataset for negative examples
- 2 Tasks:
  - Classification
  - Detection

Classification

- Classification
- Worked only on cropped images, and on background
- Image taken as a whole.
- Is it a car or not?
- K – fold cross validation (K = 10)
- 60 % Train set. 40 % Test set

Classification

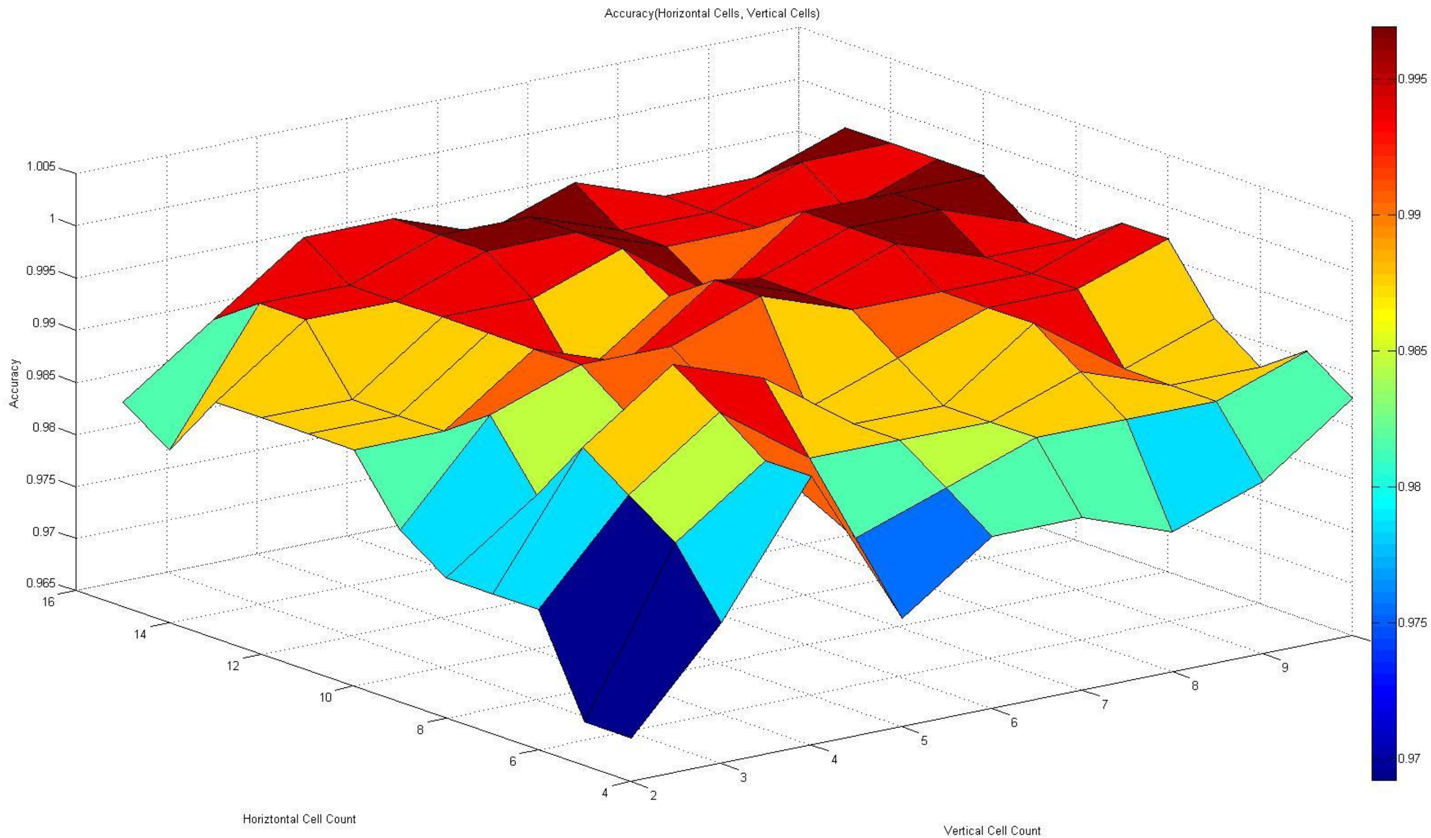
Classification

- K – fold cross validation for 4 height cells and 8 horizontal cells: 97 %.
- Varied
  - – Height Cells count [2… 10]
  - – Horiztontal Cells count [4-14]
- On 60% / 40 % Train / Test obtained high > 97%) on all height/horiz. Cell counts.

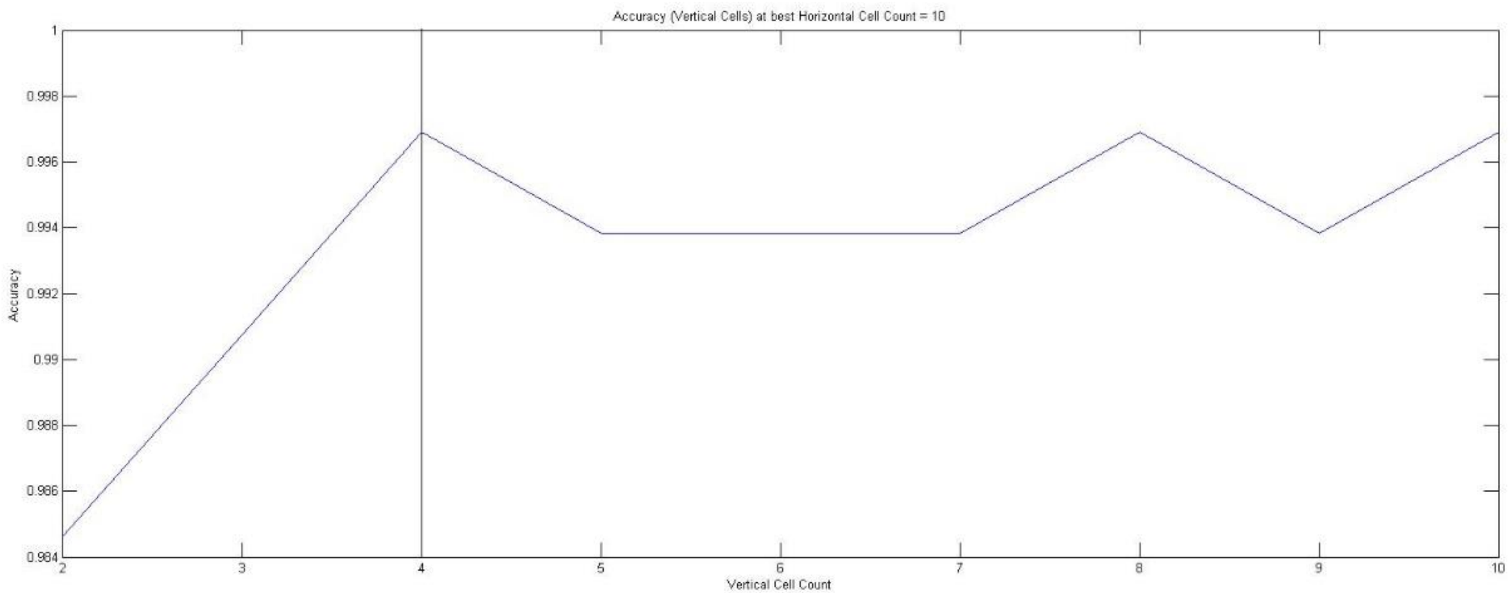
Classification

- Best accuracy is achieved at 4 height cells (vertical cells), and 10 horizontal cells (width cells). Accuracy = 0.9969
- Note: Very high accuracy

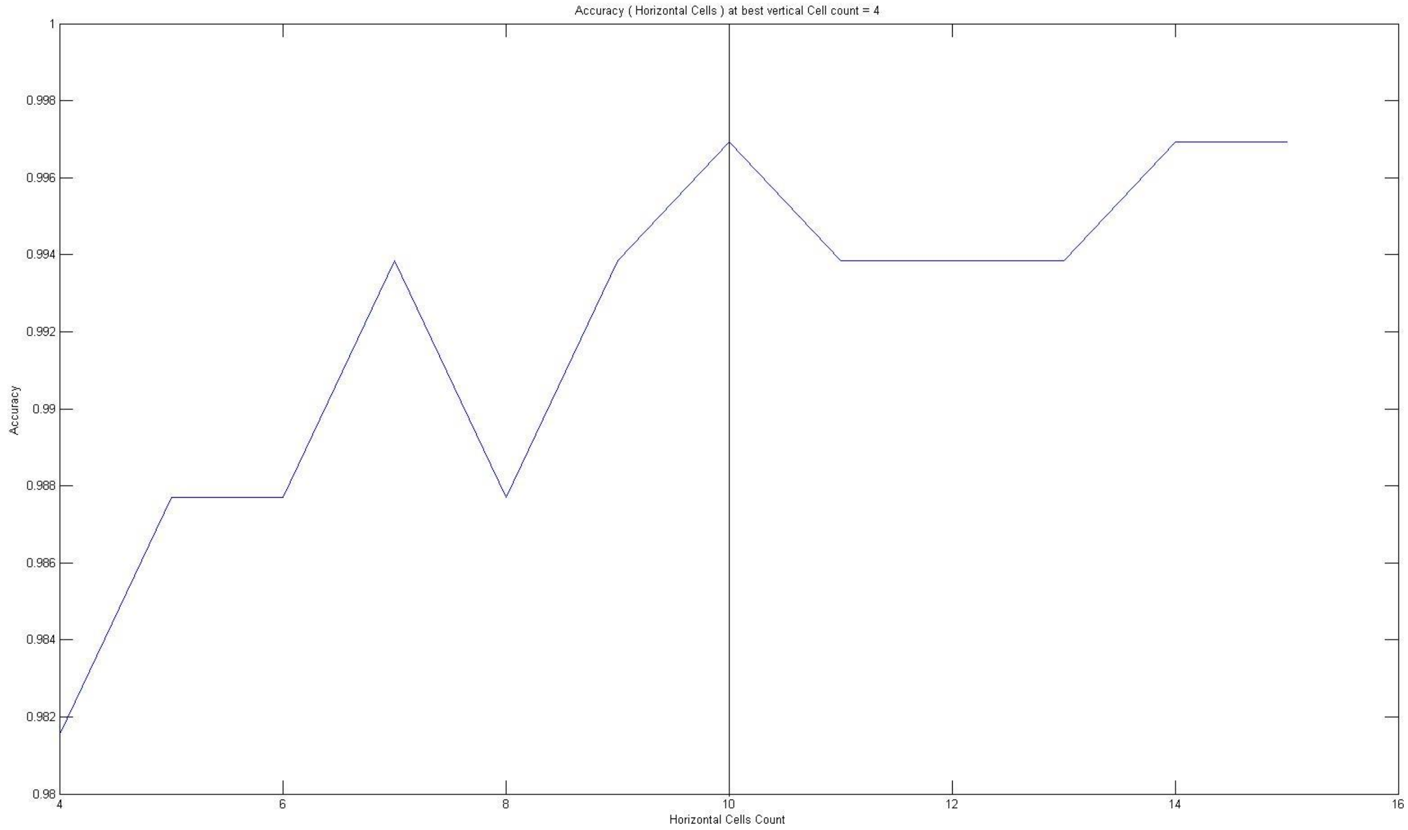
Classification



Classification



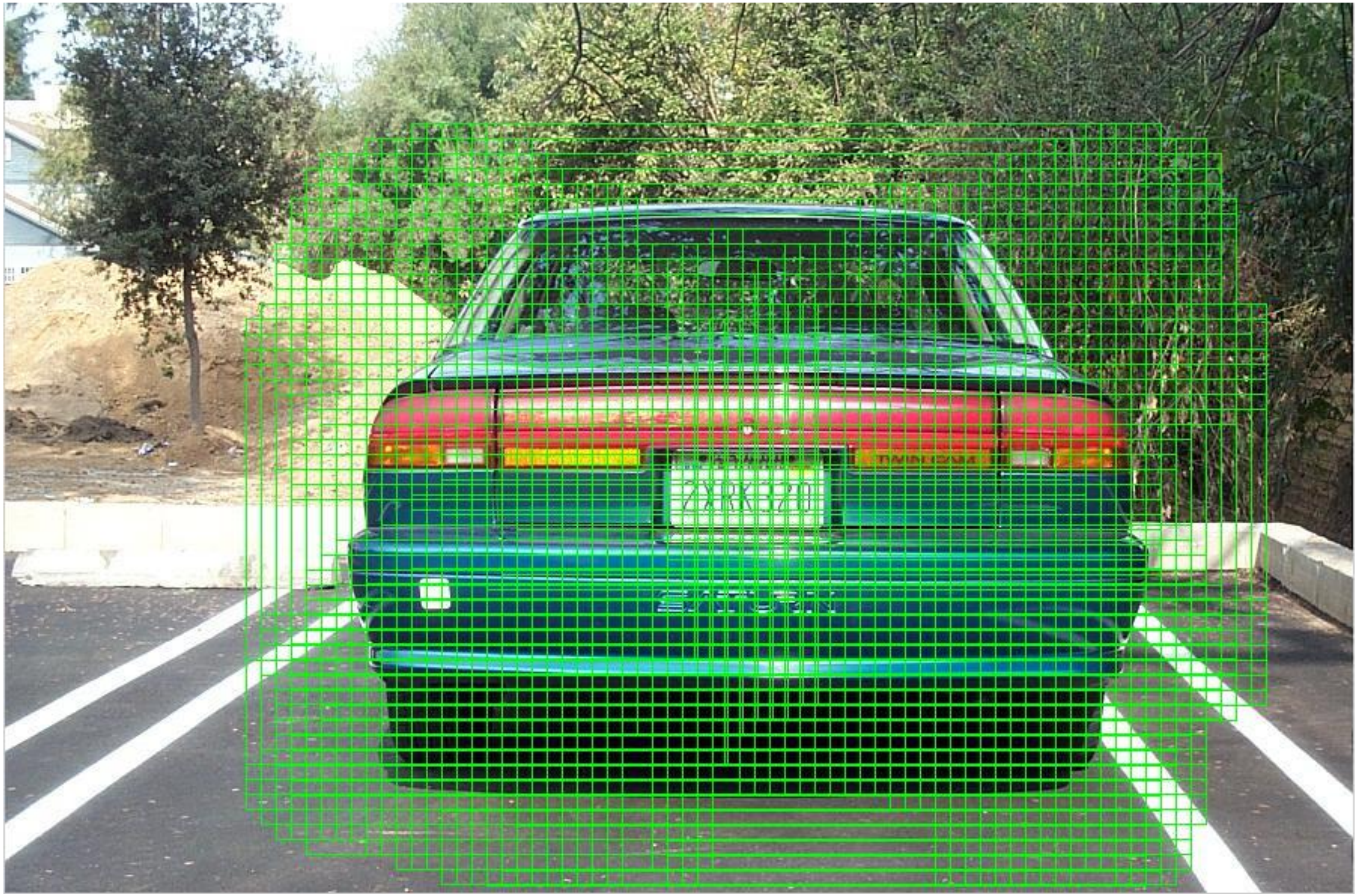
Classification



Detection

- Scan the image with a moving window, and detect the bounding box of the car.
- Method: Scanned at multiple scales
  - 3: 1, ½ 1/3.
  - Used 10 px increments for the moving window.
- Whenever the SVM says there is a car in given window, the window "votes" for all pixels.
- The pixel with highest number of votes is chosen. (max number of votes)

Detection



Detection

- Expand the bounding box until we reach pixels with votes < 0.3 max number of votes.
- The rectangle is said to be the detected car in the image.
- A detection is correct if the overlap percent with the ground truth > some threshold – typically 50 %.

Detection

- Overlap = intersection / union of detection rectangle and real rectangle

Conclusion

- We achieve both good recognition (classification) rate, and detection rate.
- However, we have only one class – cars.
- Also, cars appear only from back position.
- Initially used for pedestrian detection, HOG proves a very robust method for car detection.

Future Work

- Expand detection for cars – side view.
- Add multiple object categories.
- Investigate other classifiers, other than SVMs.

Final

- Thank you!
- Questions?