Learning to Detect Objects in Images via a Sparse, Part-Based Representation

Shivani Agarwal, Aatif Awan, and Dan Roth

Presented by:
Piotr Dollar
Nov. 30, 2004

Problem:
Object Detection

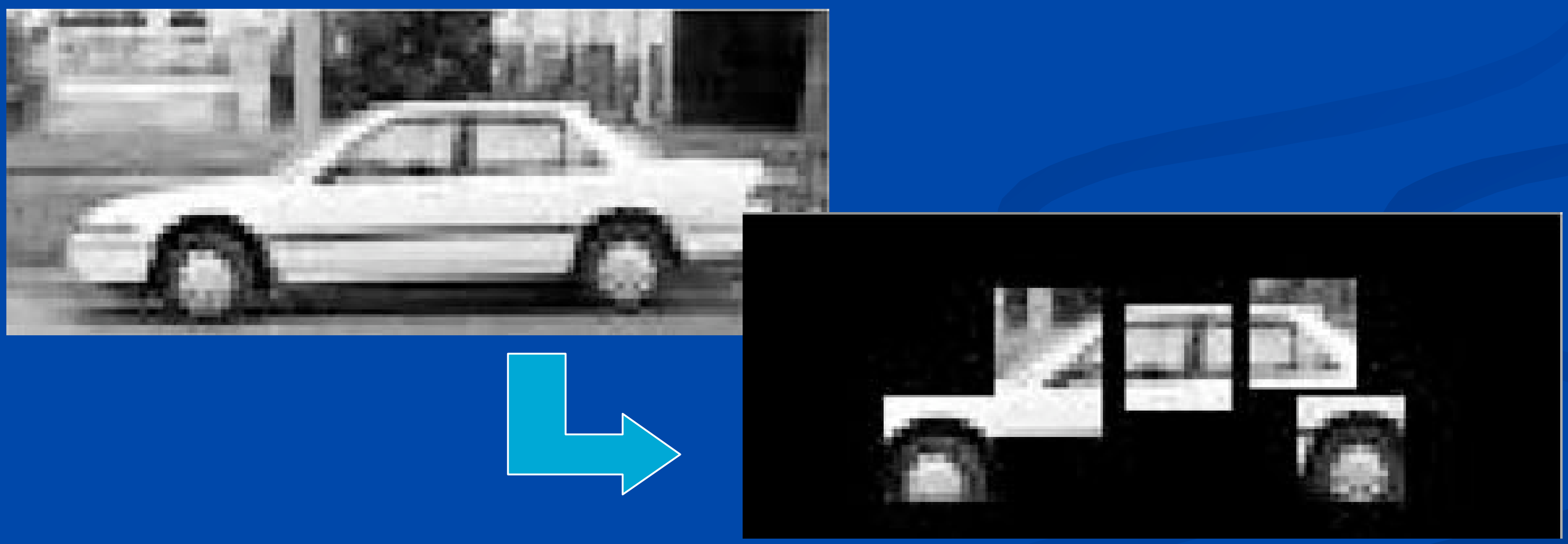
Car

Not a Car

Difficulties in Object Detection

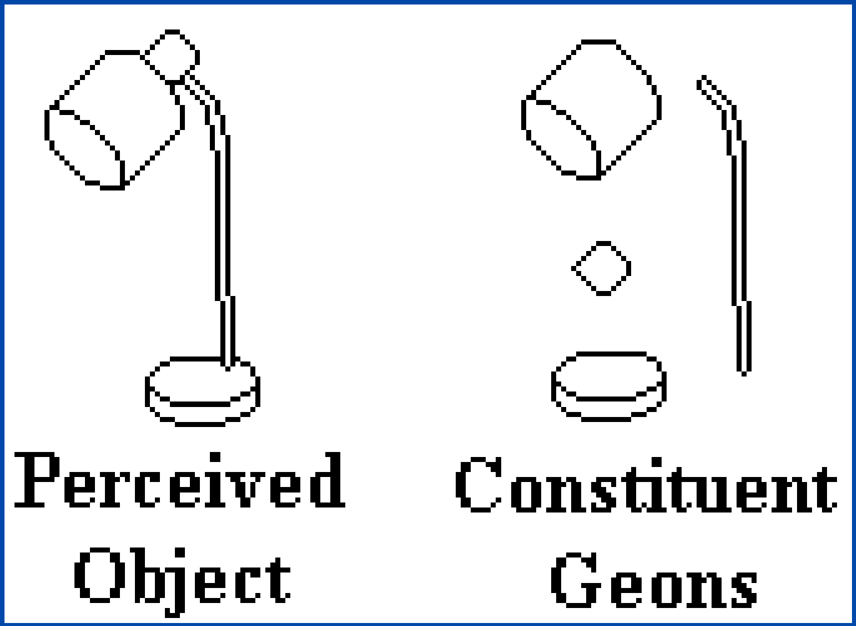
- Image clutter
- Occlusion
- Pose variation
- Object variation

Proposed Method: Represent an Object by a constellation of Parts



What is a Part?

Geon



Problems:

- Difficulty in detection
- lack of representational power

First: Feature Detection

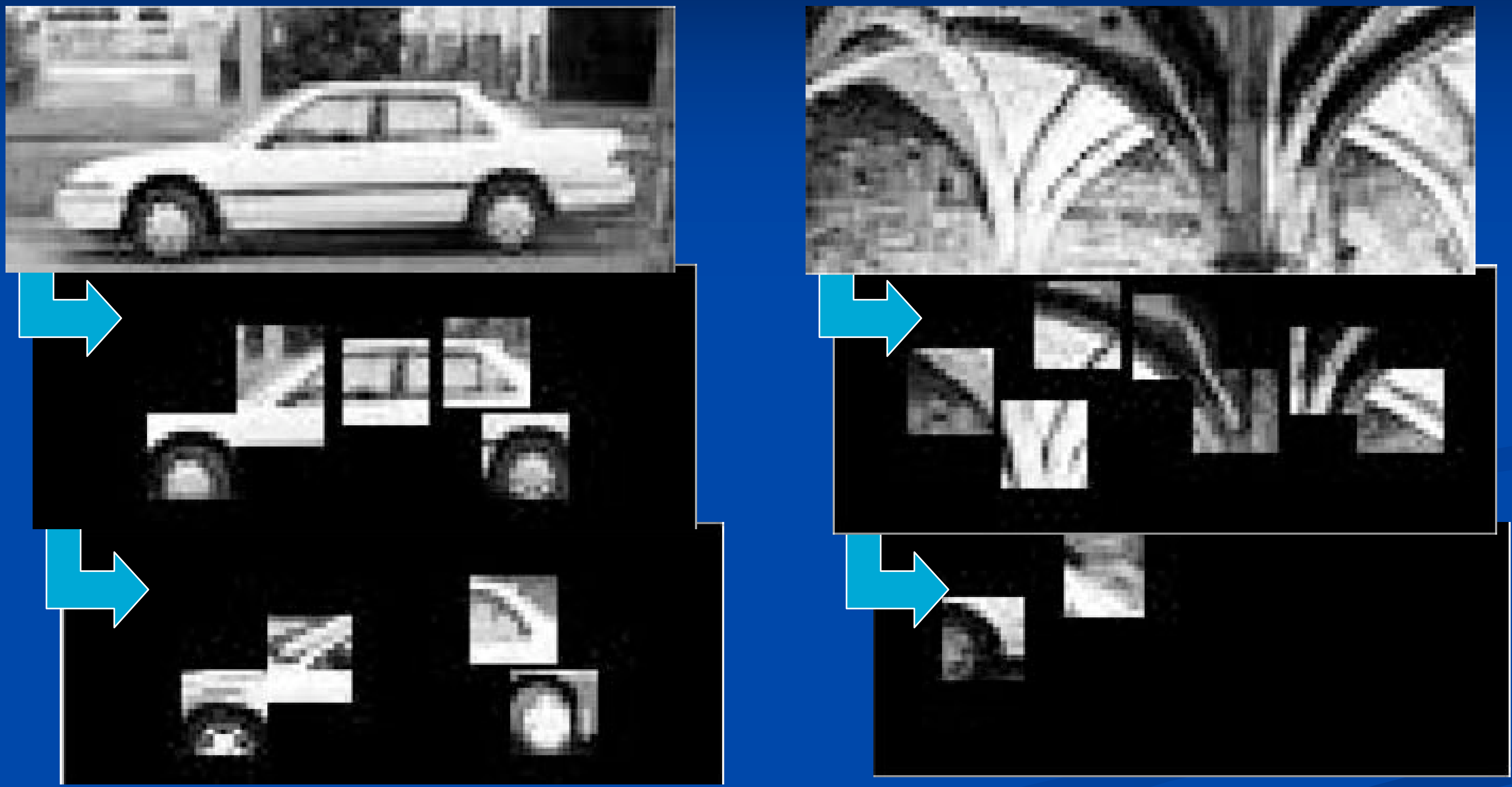
- Corner Detection
  - Low Level Operation
  - Looks for regions with large change in gradient in both directions
  - (Förstner, Harris)

Defining Parts

- Step 1: Detect a large number of features over training images:

- Step 2: Cluster patches:

Example



Feature patches as Parts:

Advantages

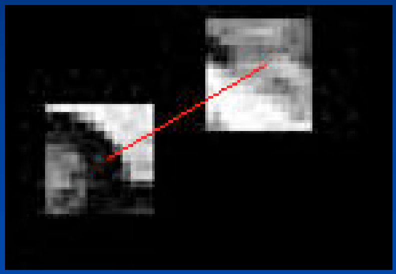
- No manual intervention needed to define a part
- Captures inherent variation
- Low level operations where any step can be quite noisy

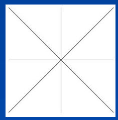
Disadvantages

- Tied to feature detector
- Redundancy in representation
- Does not necessarily capture intuitive notion of what a part is

However: It works!

Relation over Detected Parts



- Discretize angle into 8 canonical directions 
- Discretize distance into 5 bins

Any two parts have 8*5=40 possible relations between them.

Sparse Binary Features

- One feature for presence of each part
  - n total features (where n = number of parts)
- One feature for presence of a particular relation among a pair of features
  - n*n*40 total features (order 10^6 features!)
- Additional features if any of original features occur additional times
  - Theoretically no limit on number of such features, rare in practice.

Learning a Classifier

- Feature set very large, but in any given instance number of active features small
- Winnow: learns a linear decision boundary
  - Like a perceptron but with a multiplicative rather than additive update rule
  - **Sample complexity grows linearly with number of relevant features and logarithmically with total number of features**

Learning a Classifier II

- SNoW: Sparse Network of Winnows
- Winnow : SNoW ←→ Perceptron : Neural Nets
- Theory for SNoW not well developed

Windows vs Images

- All training data is in the form of windows that either have the object or don't
- When testing, use a scanning window over entire image
- Interesting problem: how to create a final classifier?
- Also, raises questions on evaluation methodology that authors address at length

Relation to Previous Work

(that we've studied)

- AdaBoost
  - Not the only algorithm for Object Recognition!
- Ivan Laptev:
  - Local Spatiotemporal Features for Behavior Recognition

Conclusion

Conclusion

I. Journal papers are long

Conclusion

II. Robust method; it works on real, difficult data

Conclusion

III. Active, promising area of research