Implicit Shape Model

[Leibe, Schiele 04]

Mario Fritz
Object Categorization in Real-World Scenes

- How to recognize ANY car
- How to recognize ANY cow
Object Categorization and Segmentation
Object Categorization and Segmentation
Object Categorization and Segmentation
Object Categorization and Segmentation
Object Categorization and Segmentation
Object Categorization and Segmentation
Overview

- Implicit Shape Model:
  - Hough transform idea
  - Non-parametric object model
  - Voting scheme for detection
  - Detection and segmentation
  - Limitations and outlook
Hough Transform

- Simple Example: find lines in image
Hough Transform

- Simple Example: find lines in image

```
+---------------------------+    +---------------------+
| distance from center      |    | angle               |
+---------------------------+    +---------------------+
```

- Image of circle with lines detected.

- Diagram shows detected lines: distance from center and angle.
Hough Transform

- Simple Example: find lines in image
Hough Transform

- Simple Example: find lines in image
Hough Transform

• example:
Codebook Representation

- Extraction of local object patches
  - scale-invariant interest points (difference of gaussian)

- Collect patches from whole training set

- Example:
Codebook Representation

- 50 car images
- only side views were used
Implicit Shape Model (ISM)

- For every codebook entry, store possible “occurrences”
Implicit Shape Model (ISM)

- For every codebook entry, store possible “occurrences”
Implicit Shape Model (ISM)

- For every codebook entry, store possible "occurrences"

- For new image, let the matched patches vote for possible object positions
Implicit Shape Model (ISM)

- For every codebook entry, store possible “occurrences”

- For new image, let the matched patches vote for possible object positions
Object Categorization Procedure

Interest Points
Object Categorization Procedure

Interest Points

Matched Codebook Entries
Object Categorization Procedure

**Interest Points**

**Matched Codebook Entries**

**Probabilistic Voting**

**Image Patch**

**Interpretation (Codebook match)**

**Object Position**

\[
p(o_n, x | e) = \sum_j p(o_n, x | I_j) p(I_j | e)
\]
Object Categorization Procedure

Interest Points → Matched Codebook Entries → Probabilistic Voting
Object Categorization Procedure

1. **Interest Points**
2. **Matched Codebook Entries**
3. **Probabilistic Voting**

Voting Space (continuous)
Object Categorization Procedure

Interest Points → Matched Codebook Entries → Probabilistic Voting

Voting Space (continuous)

Backprojection of Maximum
Object Categorization Procedure

**Interest Points** → **Matched Codebook Entries** → **Probabilistic Voting**

- **Voting Space** (continuous)
- **Backprojected Hypothesis** → **Backprojection of Maximum**
Object Categorization Procedure

1. **Interest Points**
   - Image showing detected interest points on a car image.

2. **Matched Codebook Entries**
   - Diagram showing matched codebook entries in a voting space.

3. **Probabilistic Voting**
   - Shows the backprojection of the maximum refined hypothesis.

   - Voting Space (continuous)

4. **Refined Hypothesis (uniform sampling)**
   - Image showing the refined hypothesis of the car.

5. **Backprojected Hypothesis**
   - Diagram showing backprojected hypothesis.

6. **Backprojection of Maximum**
   - Diagram illustrating backprojection of the maximum hypothesis.
Results on Cows
Results on Cows
Results on Cows
Results on Cows
Results on Cows
Results on Cows
Results on Cows
Object Categorization Procedure

Interest Points → Matched Codebook Entries → Probabilistic Voting

Voting Space (continuous)

Refined Hypothesis (uniform sampling) → Backprojected Hypothesis → Backprojection of Maximum
Object Categorization Procedure

Interest Points → Matched Codebook Entries → Probabilistic Voting

Segmentation → Refined Hypothesis (uniform sampling) → Backprojected Hypothesis → Backprojection of Maximum

Voting Space (continuous)
Object Categorization Procedure

Interest Points

Matched Codebook Entries

Probabilistic Voting

Segmentation

Refined Hypothesis (uniform sampling)

Backprojected Hypothesis

Backprojection of Maximum
Motorbikes: Detection/Segmentation Results
Results on New Sequences

- Object Detections
Results on New Sequences

- Segmentation
Secondary hypotheses

- Desired property of algorithm! ⇒ robustness to occlusion
- Standard solution: reject based on bounding box
  ⇒ Problematic - may lead to missing detections!
  ⇒ Use segmentations to resolve ambiguities instead defining costs and savings for acceptance of hypotheses
• Secondary hypotheses
  - Desired property of algorithm! ⇒ robustness to occlusion
  - Standard solution: reject based on bounding box
    ⇒ Problematic - may lead to missing detections!
    ⇒ Use segmentations to resolve ambiguities instead defining costs and savings for acceptance of hypotheses
• Secondary hypotheses
  ➢ Desired property of algorithm!  ⇒ robustness to occlusion
  ➢ Standard solution: reject based on bounding box
    ⇒ Problematic - may lead to missing detections!
    ⇒ Use segmentations to resolve ambiguities instead defining costs and savings for acceptance of hypotheses
Secondary hypotheses

- Desired property of algorithm! \(\Rightarrow\) robustness to occlusion
- Standard solution: reject based on bounding box
  \(\Rightarrow\) Problematic - may lead to missing detections!
  \(\Rightarrow\) Use segmentations to resolve ambiguities instead defining costs and savings for acceptance of hypotheses
• Secondary hypotheses
  ➢ Desired property of algorithm! ⇒ robustness to occlusion
  ➢ Standard solution: reject based on bounding box
  ⇒ Problematic - may lead to missing detections!
  ⇒ Use segmentations to resolve ambiguities instead defining costs and savings for acceptance of hypotheses
Secondary hypotheses
- Desired property of algorithm! ⇒ robustness to occlusion
- Standard solution: reject based on bounding box
  ⇒ Problematic - may lead to missing detections!
  ⇒ Use segmentations to resolve ambiguities instead defining costs and savings for acceptance of hypotheses
Secondary hypotheses
- Desired property of algorithm! \(\Rightarrow\) robustness to occlusion
- Standard solution: reject based on bounding box
  \(\Rightarrow\) Problematic - may lead to missing detections!
  \(\Rightarrow\) Use segmentations to resolve ambiguities instead defining costs and savings for acceptance of hypotheses
Secondary hypotheses

- Desired property of algorithm! $\Rightarrow$ robustness to occlusion
- Standard solution: reject based on bounding box
  $\Rightarrow$ Problematic - may lead to missing detections!
  $\Rightarrow$ Use segmentations to resolve ambiguities instead defining costs and savings for acceptance of hypotheses
Secondary hypotheses

- Desired property of algorithm! \(\Rightarrow\) robustness to occlusion
- Standard solution: reject based on bounding box
  \(\Rightarrow\) Problematic - may lead to missing detections!
  \(\Rightarrow\) Use segmentations to resolve ambiguities instead defining costs and savings for acceptance of hypotheses
Secondary hypotheses

- Desired property of algorithm! ⇒ robustness to occlusion
- Standard solution: reject based on bounding box
  ⇒ Problematic - may lead to missing detections!
  ⇒ Use segmentations to resolve ambiguities instead defining costs and savings for acceptance of hypotheses
Formalization in MDL Framework

- Savings of a hypothesis

\[ S_h = K_0 S_{\text{area}} - K_1 S_{\text{model}} - K_2 S_{\text{error}} \]

- with
  - \( S_{\text{area}} \): \#pixels \( N \) in segmentation
  - \( S_{\text{model}} \): model cost, assumed constant
  - \( S_{\text{error}} \): estimate of error, according to
    \[ S_{\text{error}} = \sum_{p \in \text{seg}(h)} (1 - p(p = \text{figure}|h)) \]

- Savings of *combined* hypothesis

\[ S_{h_1 \cup h_2} = S_{h_1} + S_{h_2} - S_{\text{area}} (h_1 \cap h_2) + S_{\text{error}} (h_1 \cap h_2) \]

- \( \rightarrow \) greedy optimization of total savings
Extension to Scale Invariance
Extension to Scale Invariance
Extension to Scale Invariance

$\text{scale} = 3.71$

$\text{scale} = 0.75$
Extensions to Scale Invariance

• Generate scale votes
  ➢ Scale as 3rd dimension in voting space
    \[
    \begin{align*}
    x_{vote} &= x_{img} - x_{occ}(s_{img}/s_{occ}) \\
    y_{vote} &= x_{img} - y_{occ}(s_{img}/s_{occ}) \\
    s_{vote} &= (s_{img}/s_{occ})
    \end{align*}
    \]
  ➢ Search for maxima in 3D voting space
Extension to Rotation Invariance
[Mikolajczyk06]
Complexity of Recognition: Local Voting vs. Global Consistency
Complexity of Recognition:
Local Voting vs. Global Consistency
Complexity of Recognition: Local vs. Global
Outlook to Lecture on 3rd March

- Recovering global consistency
- Adding discriminance to the model