Combined Object Categorization and Segmentation with an Implicit Shape Model

B. Leibe, A. Leonardis, and B. Schiele.

Marco Andreetto
marco@caltech.edu
Addressed Problems

Category detection and localization (cars)
Addressed Problems

“Soft” segmentation
Addressed Problems
Multiple-Object Scene Analysis

Recognition of Articulated Object (???)
Proposed solution

• Learning an \( I_{SM}(C) = (I_C, P_{I,C}) \)

  – \( I_C \) codebook of local patches

  – \( P_{I,C} \) Non parametric prob. distribution of a patch position (w.r.t. the center)

Similar to the solution of S. Agarwal and D. Roth 2002
Rec./Seg. procedure

Original Image → Interest Points → Matched Codebook Entries → Probabilistic Voting

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

Voting Space (continuous)
Construction of the codebook

\[
\text{similarity}(C_1, C_2) = \frac{\sum_{p \in C_1, q \in C_2} NGC(p, q)}{|C_1| \times |C_2|} > t
\]

\[
NGC(p, q) = \frac{\sum_i (p_i - \overline{p_i})(q_i - \overline{q_i})}{\sqrt{\sum_i (p_i - \overline{p_i})^2 \sum_i (q_i - \overline{q_i})^2}}
\]
Construction of prob. distribution

Non parametric density estimation: kernel implicitly defined

Segmented training data
Probabilistic framework (1/2)

\[ p(o_n, x | e, \ell) = \sum_{i} p(o_n, x | e, I_i, \ell) p(I_i | e, \ell). \]

\[ p(o_n, x | e, \ell) = \sum_{i} p(o_n, x | I_i, \ell) p(I_i | e). \]

\[ = \sum_{i} p(x | o_n, I_i, \ell)p(o_n | I_i, \ell)p(I_i | e). \]
Probabilistic framework (2/2)

\[ p(o_n, x) = \sum_k p(o_n, x \mid e_k, \ell_k) p(e_k, \ell_k) \]

\[ \text{score}(o_n, x) = \sum_k \sum_{x_j \in W(x)} p(o_n, x_j \mid e_k, \ell_k). \]

Mean Shift Algorithm for local maxima
Experimental results: cars

50x2 training examples
Experimental results: cars
Experimental results: cows
Object Segmentation (1/3)

• Given \( p(o_n, x|e, \ell) = \sum_i p(o_n, x|e, I_i, \ell)p(I_i|e, \ell). \)

\[
p(e, \ell|o_n, x) = \frac{p(o_n, x|e, \ell)p(e, \ell)}{p(o_n, x)} = \sum_I p(o_n, x|I, \ell)p(I|e)p(e, \ell)p(o_n, x)
\]

\[
p(p = figure|o_n, x) = \sum_{p \in (e, \ell)} p(p = figure|o_n, x, e, \ell)p(e, \ell|o_n, x)
\]
Object Segmentation (2/3)

\[
p(p = \text{figure}|o_n, x) = \sum_{p \in (e, \ell)} \sum_{I} p(p = \text{fig.}|o_n, x, e, I, \ell) p(e, I, \ell|o_n, x)
\]

\[
= \sum_{p \in (e, \ell)} \sum_{I} p(p = \text{fig.}|o_n, x, I, \ell) \frac{p(o_n, x|I, \ell) p(I|e) p(e, \ell)}{p(o_n, x)}
\]

\[
L = \frac{p(p = \text{figure}|o_n, x)}{p(p = \text{ground}|o_n, x)}.
\]
Segmentation results (3/3)
Hypothesis verification (1/2)

MDL principle (Occam's razor)
Hypothesis verification (2/2)

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

\[ S_{error} = \sum_{p \in \text{Seg}(h)} (1 - p(p = \text{figure}|h)) \]

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

Combinatorial complexity of space \( \text{H} \)
Experimental results: cars
## Comparison

<table>
<thead>
<tr>
<th>Method</th>
<th>EER</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agarwal and Roth (2002)</td>
<td>~79%</td>
</tr>
<tr>
<td>Garg et al. (2002)</td>
<td>~88%</td>
</tr>
<tr>
<td>Fergus et al. (2003)</td>
<td>88.5%</td>
</tr>
<tr>
<td>Our algorithm</td>
<td>97.5%</td>
</tr>
</tbody>
</table>

UIICI car database
Open Questions

• Large training set for non-parametric est.
• Segmented training set
• Shape partially modeled
• Not too articulated objects
• Multiple classes in one image?
• Star model?
References

