La segmentation d'image et de vidéo dans le cadre de la vision cognitive

Vincent Martin, INRIA Sophia Antipolis
Equipe projet PULSAR

GDR-ISIS, réunion SCATI
ENST, Paris
27 Mars 2008
Introduction
Image and Video Segmentation in Vision Systems

• Motivations
  • Image segmentation is a crucial task in vision systems
  • Issue 1: How to choose the right segmentation algorithm?
  • Issue 2: How to tune its parameters?
Introduction

Issue 1: to choose the right segmentation algorithm

- **Segmentation task:**
  - Extract the white flies from the background (rose leaves)

- **Problems:**
  - No unique solution
  - Rely on the knowledge of the expert in image segmentation
Introduction
Issue 2: to tune a segmentation algorithm

- **Algorithm:**
  Hysteresis thresholding

- **Low threshold setting:**
  A: 0.33, B: 0.39, C: 0.55
  Best value ≈ C

- **Problems:**
  - Number of parameters
  - Interaction between parameters
  - Subjective and time-consuming task
  - No semantic results with low-level algorithms
Introduction

Goal: to propose an approach for

- Extracting automatically optimal segmentation parameters
- Learning the selection and the parameter tuning of a segmentation algorithm
- Performing an adaptive segmentation of new images w.r.t. the image content and the application needs.

A cognitive vision approach

- Understanding of the environment
- Learning from real world examples
- Adaptability to its environment
- Reasoning/interpretation of knowledge
- Re-usability by generalization of operations
- Convenience on usage
Related Work
Image Segmentation

• Different approaches
  • Feature space based approaches: clustering, histograming
  • Image domain based approaches: edge detectors, watershed, snakes, region growing, split-and-merge
  • Object-based approaches: top-down + bottom-up

<table>
<thead>
<tr>
<th>Approaches</th>
<th>Strengths</th>
<th>Weakness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feature space based</td>
<td>▪ Detection of homogeneity in a global context</td>
<td>▪ Spatial relationship between pixels is ignored</td>
</tr>
<tr>
<td>Image domain based</td>
<td>▪ Smooth and accurate results</td>
<td>▪ Edge linking problem</td>
</tr>
<tr>
<td></td>
<td>▪ Robust to noise</td>
<td>▪ Need of prior information on objects</td>
</tr>
<tr>
<td>Object based</td>
<td>▪ Combination of several methods</td>
<td>▪ Computational cost</td>
</tr>
<tr>
<td></td>
<td>▪ Meaningful results</td>
<td>▪ Restricted applicability</td>
</tr>
</tbody>
</table>
Related Work
Video Segmentation

• Different approaches
  • Optical flow: based on motion vector fields
  • Background modelling techniques: running average, MoG, Kernel Density Estimator, Codebooks
  • Spatio-temporal approaches: region + pixel based approaches

<table>
<thead>
<tr>
<th>Approaches</th>
<th>Strengths</th>
<th>Weakness</th>
<th>Computational complexity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Optical Flow</td>
<td>▪ Motion detection accuracy</td>
<td>▪ Aperture problem</td>
<td>High</td>
</tr>
<tr>
<td>Background Modelling</td>
<td>▪ Robustness</td>
<td>▪ Parameter tuning</td>
<td>low/medium</td>
</tr>
<tr>
<td></td>
<td>▪ Fast</td>
<td>▪ Fail to model both slow and fast variations</td>
<td></td>
</tr>
<tr>
<td>Spatio-temporal</td>
<td>▪ Meaningful results</td>
<td>▪ Slow, Parameter tuning</td>
<td>High</td>
</tr>
</tbody>
</table>
Related Work
Segmentation in Vision Systems

- Knowledge-based systems with program supervision [Clouard99, Thonnat99]
- Learning the algorithm selection based on visual assessment [Zhang00, Xia05]
- Bottom-up segmentation: adaptive selection of key parameters [Bahnu95, Pignalberi03, Abdul-Karim05]
- Top-down segmentation: learning template-based models [Borenstein06]
- Cognitive vision systems [ECVision roadmap05]
Approach Overview

1. Adaptive Image Segmentation
   - Selection of the proper algorithm and its parameters
   - Semantic segmentation
   - Manual segmentation easily available

2. Adaptive Video Segmentation
   - Focus on two learning-based algorithms (codebook and MoG)
   - Huge data set available (without manual segmentation)
Approach Overview
Context Analysis and Clustering

• Context Analysis
  • Context described by an image histogram sensitive to local and global image variations as Coherence Color Feature Vector [Pass96]

• Context Clustering
  • Problem: unknown number of classes, clusters of arbitrary shapes, moderately dimensional space (~$10^2$), numerous samples with noise (~$10^4$)
  • Algorithm: Density-based Spatial Clustering for Applications with Noise (DBSCAN) [Ester96]
  • Based on two main concepts: density reachability and density connectability
  • Need to set manually the $\epsilon$-neighbourhood Parameter
Approach Overview
Adaptive Image Segmentation

Training Stage

- Training Image Set
- Segmentation Algorithms
- Ground Truth Data
- Segmentation Knowledge Base
  - Learnt Clusters of Training Images
  - Learnt Algo Parameters
  - Trained Region Classifiers
Approach Overview
Adaptive Image Segmentation

Online segmentation

New Image → Learnt Clusters of Training Images → Context Identification

Algorithm Selection → Learnt Parameters

Bottom-Up Segmentation → Trained Region Classifiers

Semantic Segmentation → Segmented Image
Approach Overview
Adaptive Video Segmentation

Training stage
Approach Overview
Adaptive Video Segmentation

Real-time segmentation
Adaptive Image Segmentation
Algorithm Parameter Optimization

Diagram:

1. Training Image
2. Segmentation Algorithm
3. Segmentation
5. Stop?
6. Optimal Parameters + Final segmentation assessment value
7. Updated Segmentation Parameters
8. Global Optimization Algorithm
9. Algorithm Parameter Space

Flowchart:

- Segmentation Algorithm
- Manual Segmentation Evaluation
- Stop?
- Updated Segmentation Parameters
- Global Optimization Algorithm
- Algorithm Parameter Space
- Optimal Parameters + Final segmentation assessment value
Adaptive Image Segmentation
Evaluation Metric of the Segmentation Performance

- Ground truth pixel $x \in B^G$
- Segmented pixel $x \in B^A$
- $x \in (B^G \cap B^A)$

$$T_1 = \{x \mid (x \in B^G) \wedge (x \notin B^A)\}$$
$$T_2 = \{x \mid (x \in B^A) \wedge (x \notin B^G)\}$$

Multi-objective optimization

$$\text{obj.a: } e^B_{\text{missed}} = \frac{|T_1|}{|B^G|}, \quad \text{obj.b: } e^B_{\text{false}} = \frac{|T_2|}{|B^A|}$$

Segmentation error $E^A = \frac{1}{2}(w_m \cdot e^B_m + w_f \cdot e^B_f)$
Adaptive Image Segmentation
Algorithm Selection Strategy
Adaptive Image Segmentation
Region Classifier Training

Training Image Set

Manual Segmentation → Manual Region Labelling

Segmentation with Learned Parameters → Automatic Region Labelling

Labelled ground truth regions

Feature Extraction → Region Classifier Training

Trained Region Classifiers
Adaptive Image Segmentation
Region Classifier Training

• **Input:**
  - Low-level Region Features:
    - Color histograms in different color spaces $S$ and quantification levels $q$
    - Texture features (Quadrature Mirror Filter Bank) based on the energy property of pixels
  - Region semantic labels (e.g., background, object class #1,...)

• **Learning algorithm:**
  - Support Vector Machine (OVA-SVM) with a Radial Basis Function (RBF) kernel ($C, \gamma$)
  - $S, q, C,$ and $\gamma$ optimized using cross-validations.

• **Output:** trained region classifiers
Adaptive Image Segmentation
Experiment on a Biological Application

- Biological issue: the early detection of pest (e.g., insects) on rose leaves (cooperation with P. Boissard, INRA Sophia)
- 100 scanned images of rose leaves (training/test = 50/50)
- 5 region-based segmentation algorithms:

<table>
<thead>
<tr>
<th>Segmentation Algorithm</th>
<th>No. of Free Parameters</th>
<th>Parameter Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Color Structure Code (CSC)</td>
<td>1</td>
<td>Region merging threshold</td>
</tr>
<tr>
<td>Statistical Region Merging (SRM)</td>
<td>1</td>
<td>Coarse-to-fine scale control</td>
</tr>
<tr>
<td>Hysteresis thresholding (THRESH)</td>
<td>2</td>
<td>Low and high thresholds</td>
</tr>
<tr>
<td>Efficient Graph-Based Image Segmentation (EGBIS)</td>
<td>2</td>
<td>Smooth control and color threshold</td>
</tr>
<tr>
<td>Color Watershed Adjacency-Graph Merge (CWAGM)</td>
<td>3</td>
<td>Merging threshold for the graph representation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Min number of regions</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Min prob. For watershed threshold</td>
</tr>
</tbody>
</table>
Adaptive Image Segmentation
Experiment on a Biological Application

Examples of the problem complexity (different training/test images)
Adaptive Image Segmentation
Examples of Optimal Parameter Extraction

training image

manual segmentation
Adaptive Image Segmentation
Examples of Evaluation Profiles

CSC
SRM
Hysteresis Thresholding
EGBIS

Good Performance Value
↓
↑ Algo Param value
Adaptive Image Segmentation
Examples of Optimal Parameter Extraction

CSC
Region Merging Color Threshold = 88.33

SRM
Coarse-to-Fine Scale Control = 338.67

Hysteresis Thresholding
Low Threshold = 0.33 High Threshold = 0.37
Adaptive Image Segmentation
Evaluation of the Segmentation Optimization

On the training image set using the performance evaluation metric (‘0’ = no segmentation error)

<table>
<thead>
<tr>
<th>Segmentation Algorithm</th>
<th>Systematic Search</th>
<th>Simplex Algorithm</th>
<th>Genetic Algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td>CSC</td>
<td>0.30</td>
<td>0.32</td>
<td>0.30</td>
</tr>
<tr>
<td>SRM</td>
<td>0.29</td>
<td>0.42</td>
<td>0.31</td>
</tr>
<tr>
<td>THRESH</td>
<td>0.26</td>
<td>0.26</td>
<td>0.26</td>
</tr>
<tr>
<td>EGBIS</td>
<td>0.51</td>
<td>0.53</td>
<td>0.52</td>
</tr>
<tr>
<td>CWAGM</td>
<td>0.36</td>
<td>0.32</td>
<td>0.30</td>
</tr>
<tr>
<td>mean</td>
<td>0.344</td>
<td>0.370</td>
<td>0.338</td>
</tr>
</tbody>
</table>
# Adaptive Image Segmentation

## Evaluation of the Segmentation Optimization

<table>
<thead>
<tr>
<th>Segmentation Algorithm</th>
<th>Systematic search</th>
<th>GA</th>
<th>Simplex Algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td>CSC</td>
<td>1000</td>
<td>733</td>
<td>83</td>
</tr>
<tr>
<td>SRM</td>
<td>1000</td>
<td>734</td>
<td>82</td>
</tr>
<tr>
<td>THRESH</td>
<td>10000</td>
<td>840</td>
<td>704</td>
</tr>
<tr>
<td>EGBIS</td>
<td>2550</td>
<td>840</td>
<td>497</td>
</tr>
<tr>
<td>CWAGM</td>
<td>1250</td>
<td>840</td>
<td>1821</td>
</tr>
</tbody>
</table>

Up to two free parameters: prefer **Simplex Algorithm**

More than two: prefer **Genetic Algorithm**
Adaptive Image Segmentation
Evaluation of the Segmented Test Images

Performance evaluation scores using: (i) default parameters, (ii) learnt parameters, (iii) learnt parameters + region labelling (0 = no segmentation error)

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>(i)</th>
<th>(ii)</th>
<th>(iii)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CSC</td>
<td>0.57</td>
<td>0.21</td>
<td>0.22</td>
</tr>
<tr>
<td>SRM</td>
<td>0.29</td>
<td>0.22</td>
<td>0.13</td>
</tr>
<tr>
<td>THRESH</td>
<td>-</td>
<td>0.68</td>
<td>0.19</td>
</tr>
<tr>
<td>CWAGM</td>
<td>0.72</td>
<td>0.35</td>
<td>0.20</td>
</tr>
<tr>
<td>mean</td>
<td>0.51</td>
<td>0.37</td>
<td>0.19</td>
</tr>
</tbody>
</table>
Adaptive Image Segmentation
Algorithm Selection

- Two identified contexts: back sides and front sides of leaves
- One selected algorithm (SRM) with two parameter settings

Context: 1/2
- Algo: SRM
- Param: $Q = 70.7$

Context: 2/2
- Algo: SRM
- Param: $Q = 78.1$
Adaptive Image Segmentation with the SRM Algorithm

<table>
<thead>
<tr>
<th>Original image</th>
<th>Manual segmentation</th>
<th>Segmentation with default parameters</th>
<th>Segmentation with learnt parameters</th>
<th>+ region labelling</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1" alt="Original image" /></td>
<td><img src="image2" alt="Segmentation" /></td>
<td><img src="image3" alt="Segmentation default parameters" /></td>
<td><img src="image4" alt="Segmentation learnt parameters" /></td>
<td><img src="image5" alt="Segmentation + region labelling" /></td>
</tr>
</tbody>
</table>
Adaptive Image Segmentation with Different Segmentation Algorithms

<table>
<thead>
<tr>
<th>Original image</th>
<th>CSC</th>
<th>SRM</th>
<th>THRESH</th>
<th>CWAGM</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Original image" /></td>
<td><img src="image2.png" alt="CSC" /></td>
<td><img src="image3.png" alt="SRM" /></td>
<td><img src="image4.png" alt="THRESH" /></td>
<td><img src="image5.png" alt="CWAGM" /></td>
</tr>
<tr>
<td><img src="image2.png" alt="Original image" /></td>
<td><img src="image3.png" alt="CSC" /></td>
<td><img src="image4.png" alt="SRM" /></td>
<td><img src="image5.png" alt="THRESH" /></td>
<td><img src="image6.png" alt="CWAGM" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Score</th>
<th>Score</th>
<th>Score</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.2592</td>
<td>0.1645</td>
<td>0.4254</td>
<td>0.3502</td>
</tr>
<tr>
<td>0.1277</td>
<td>0.1500</td>
<td>0.7251</td>
<td>0.9106</td>
</tr>
</tbody>
</table>
Adaptive Image Segmentation
Training/Test Comparison

Original image

Manual segmentation

Segmentation with default parameters

Segmentation with learned parameters

+ region labelling
Adaptive Image Segmentation
Integration in a Cognitive Vision Platform [Hudelot05]

- Based on two Knowledge-Based Systems (KBS):
  - Image interpretation layer using a visual concept ontology [Maillot05]
  - Image processing layer using program supervision

- Integration of our adaptive segmentation to complement the program supervision KBS
## Adaptive Image Segmentation

Cognitive Vision Platform: Counting Results

Results for early detection of mature white flies

<table>
<thead>
<tr>
<th></th>
<th>With <em>ad hoc</em> segmentation</th>
<th>With a cognitive vision based segmentation</th>
</tr>
</thead>
<tbody>
<tr>
<td>False Negative Rate</td>
<td>9.1%</td>
<td>11.0%</td>
</tr>
<tr>
<td>False Positive Rate</td>
<td>9.4%</td>
<td>2.7%</td>
</tr>
</tbody>
</table>
Adaptive Video Segmentation

Figure-ground segmentation

- Learning-based algorithms
  - Codebook model [Kim05] based on color distortion and brightness bounds → compressed models
  - Mixture of Gaussians [Stauffer99] based on multimodal background distributions
  - Both Trained on background samples

- Limits
  - Unable to cope with both quick and long-term changes
Adaptive Video Segmentation
Video Sequence
Adaptive Video Segmentation
Image Context Analysis

Color Coherence Feature Vector in HSV Color Space
H=[1-20], S=[21-34], V=[35-52]
Adaptive Video Segmentation
Cluster identification

Night context

Morning context

Afternoon context
Adaptive Video Segmentation
Context Adaptation and Filtering Effect

No Context adaptation

Context Adaptation + Filtering

Context adaptation
Adaptive Video Segmentation
Codebook vs MoG

MoG

Codebooks
Adaptive Video Segmentation
Quantitative Evaluation

- **Training set:** 11277 frames taken along 3 days (one every 4 minutes)
- **Test set:** next three days (84236 frames, 248 with manual segmentation)

Results of False Positive Rate (FPR) and False Negative Rate (FNR) on the 248 test frames:

<table>
<thead>
<tr>
<th></th>
<th>without context adaptation</th>
<th>with filtered context adaptation</th>
</tr>
</thead>
<tbody>
<tr>
<td>FPR</td>
<td>2.26%</td>
<td>4.65%</td>
</tr>
<tr>
<td>FNR</td>
<td>73.70%</td>
<td>36.38%</td>
</tr>
</tbody>
</table>
Conclusion and Future Work
Summary for Image Segmentation

• Contributions:
  • Image processing: goal-oriented method for segmentation algorithm ranking and parameter tuning
  • End-users: easy set up of the segmentation task w.r.t. their needs

• Advantages:
  • Automatic evaluation of the segmentation performance
  • Method independent of the set of segmentation algorithms

• Limitations:
  • Supervised learning (manual ground truth)
  • Availability of representative training data
Conclusion and Future Work
Summary for Video Segmentation

• Contributions:
  • An adaptive method for background model selection w.r.t. the context variations

• Advantages:
  • Weak supervision
  • Adapted to very long sequences

• Limitations:
  • Unable to cope with unforeseen situations
Conclusion and Future Work

Summary for the Cognitive Vision Approach

• Contributions in cognitive vision systems:
  • For a given segmentation problem
    – Automatically learns/finds the suitable algorithm and its parameters
    – Adaptability for the image/video context
  • Re-usability (generic cognitive vision system)
  • Convenience (user-friendly)

• limitations:
  • No feedback mechanisms between high and low-level layers
  • Knowledge gained during learning process is not used completely
Conclusion and Future Work

Future Work

• Short term:
  • Use incremental learning techniques to learn on-the-fly unforeseen situations
  • Apply the parameter tuning method to the video segmentation algorithms

• Long term:
  • Image segmentation: Extend the experiments to (>2)-class problems
  • Construct and share video databases with ground truth for video segmentation benchmarking