From Hierarchies of Regions to Image Understanding

Prof. Sinisa Todorovic
Acknowledgment

**UIUC:**
Prof. Narendra Ahuja

**Oregon State University:**
William Brendel
Nadia Payet
Muhamed Amer
Prof. Eugene Zhang
Goals: Object Recognition

input image set

discover and learn all objects present

detect
segment
explain
all occurrences
of the learned objects

new image
Goal: Video Painterly Rendering

video sequence enhanced with multiple painting styles
-- one per each object

flower petals = van Gogh
stamens = expressionism
background = pointilism
Many applications require unsupervised partitioning of the image into textured and non-textured subimages.
### Prior Work: Object Recognition

#### Prior Work
- High degree of supervision
- Predominance of keypoint features
- Ignoring the spatial info
- Limited goals

#### Our Extensions
- Relaxing supervision requirements
- Using richer features: regions
- Accounting for multiscale spatial info
- Unified framework for many goals
## Prior Work: Painterly Rendering

<table>
<thead>
<tr>
<th>PRIOR WORK</th>
<th>OUR EXTENSIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uses only a single style</td>
<td>Object-based multiple styles</td>
</tr>
<tr>
<td>Unrealistic, poor artistic expression</td>
<td>Rich artistic expression</td>
</tr>
</tbody>
</table>

![Prior Work Comparison](image1.png)

![Our Extensions Comparison](image2.png)
## Prior Work: Texture Segmentation

<table>
<thead>
<tr>
<th>PRIOR WORK</th>
<th>OUR EXTENSIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uses a pre-specified bank of filters</td>
<td>Extraction of texels</td>
</tr>
<tr>
<td>Assumptions: smoothness, scale</td>
<td>Relaxing the assumptions</td>
</tr>
</tbody>
</table>

### PRIOR WORK
- **meanshift**
- **active contours**
- **our results**
WHAT IS AN OBJECT?
### Properties of Objects

**3D objects in the scene** → **2D objects in the image**

<table>
<thead>
<tr>
<th>cohesive</th>
<th>occupy regions</th>
</tr>
</thead>
<tbody>
<tr>
<td>form characteristic spatial configurations with other objects</td>
<td>context</td>
</tr>
<tr>
<td>have parts</td>
<td>subregions</td>
</tr>
<tr>
<td>parts have characteristic spatial layout</td>
<td>spatial layout of subregions</td>
</tr>
</tbody>
</table>
Rationale for Learning – Like a Small Child

It is likely to be meaningful:

• If some parts repeat in the set of images
• If some configurations of the learned parts repeat in the set
It is likely to be meaningful:

- If some parts repeat in the set of images
- If some configurations of the learned parts repeat in the set
Rationale for Learning – Like a Small Child

It is likely to be meaningful:

- If some parts repeat in the set of images
- If some configurations of the learned parts repeat in the set
Rationale for Learning – Like a Small Child

It is likely to be meaningful:

• If some parts repeat in the set of images

• If some configurations of the learned parts repeat in the set
It is likely to be meaningful:

- If some parts repeat in the set of images
- If some configurations of the learned parts repeat in the set
Category = Set of Recurring Similar 2D Objects

(1) Photometric (e.g., color)

(2) Geometric (e.g., area, shape)

(3) Structural:

spatial layout of subcategories

containment of subcategories
Category = Set of Recurring Similar 2D Objects

(1) Photometric (e.g., color)

(2) Geometric (e.g., area, shape)

(3) Structural:

spatial layout of subcategories

containment of subcategories

recursive definition
Problem Statement

- Given a set of images
Problem Statement

• Given a set of images

• Discover frequently occurring 2D objects
  • Under illumination and scale changes
  • Amidst background clutter
  • Under partial occlusion
Problem Statement

• Given a set of images
• Discover frequently occurring 2D objects
  • Under illumination and scale changes
  • Amidst background clutter
  • Under partial occlusion
• Learn their generative, statistical models
Problem Statement

• Given a set of images

• Discover frequently occurring 2D objects
  • Under illumination and scale changes
  • Amidst background clutter
  • Under partial occlusion

• Learn their generative, statistical models

• Use the models for
  • Object recognition
  • Object-based painterly rendering and synthesis
  • Texel-based texture segmentation
1. Image representation = Hierarchy of regions

2. Region matching

3. Applications and results
Image = Tree \Rightarrow Object = Subtree

Ahuja PAMI96, Tobb & Ahuja TIP97, Arora&Ahuja ICPR06
Connected Segmentation Trees

Lateral links = Region neighbor relations
Hierarchical links = Region embedding

Ahuja& Todorovic CVPR08
Region Properties Associated with Each Node

- Gray-level contrast with surround
- Boundary shape
- Displacement of centroids
- Orientation

Properties relative wrt parent ⇒ Scale and in-plane rotation invariance
1. Image representation = Hierarchy of regions

2. Region matching

3. Applications and results
How to Discover Repeating Image Parts?

Object category is present = Many similar subgraphs

Discovering objects = Graph matching
Graph Matching = Subgraph Isomorphism

Match two regions

- If their immediate properties are similar
- AND the same holds for their subregions
- AND the same holds for their neighbors
Given two graphs: $G = (V, E)$ and $G' = (V', E')$

Find the mapping $f = \{(v, v')\} \subset V \times V'$

which minimizes their cost of matching:

$$COST_{GG'} = \min_f \left[ \sum_{(v,v') \in f} \psi_{vv'} + \sum_{(v,v',u,u') \in f \times f} \phi_{vv'uu'} \right]$$

unary potential function of region properties

pairwise potential function of spatial relationships
Graph Matching: Formulation

Linearization by introducing an indicator vector

\[ X = [0 \ 0 \ 1 \ 0 \ 0 \ 1 \ ... \ 0 \ 1 \ 0]^T \]

matched pair \((v, v')\)  
unmatched pair \((u, u')\)

\[ \downarrow \]

Discrete problem

\[ \min_X [\Psi^T X + X^T \Phi X] \]

s.t. \(x_{vv'} \in \{0, 1\}\)
Relaxation of the discrete problem

\[
\min_X \left[ \Psi^T X + X^T \Phi X \right]
\]

s.t. \( \forall x_{\nu\nu'} \geq 0, \sum_{\nu} x_{\nu\nu'} = 1, \sum_{\nu'} x_{\nu\nu'} = 1 \)

Todorovic&Ahuja IJCV08, PAMI08, CVPR06-08, ICCV07, ICPR06-08
1. Image representation = Hierarchy of regions

2. Region matching

3. Applications and results
   a. Object recognition
   b. Painterly rendering
   c. Texture segmentation
Discovering Objects = Matching + Clustering

training images
Discovering Objects = Matching + Clustering

training images

Each cluster = Distinct Object
Learning a Model of Each Cluster = Structural EM

matched subgraphs

hierarchical object model

model structure ?

model parameters ?
Category Model = Bayesian Net

object part (hidden)

region properties

number of children
Category Model = Bayesian Net

\[ P(X, Y, N | T, \Omega) = \prod_{j \in T} P(N_j | x_j) \prod_{i=1}^{N_j} P(x_i | x_j) P(y_{ij} | x_i x_j) \]

- object part (hidden)
- region properties
- number of children
- structure + parameters
- Markovian chain
- Exponential
- Gaussian
Learning a Model = Structural EM

Given $\mathcal{T}$
Belief propagation $\Rightarrow \Omega$

Given $\Omega$
Graph matching $\Rightarrow \mathcal{T}$

Category model
$\mathcal{T} \Omega$

$\mathcal{T}$ model structure
$\Omega$ model parameters

Todorovic&Ahuja ICCV07
Results: Weizmann Horses

Training images

Category model

Todorovic & Ahuja PAMI08
Results: Weizmann Horses

- Object segmentation is good on contours that are:
  - Jagged
  - Blurred
  - Form complex patterns

- Low-contrast regions merge with background

Todorovic&Ahuja  PAMI09
UIUC Hoofed Animals Dataset

http://vision.ai.uiuc.edu/~sintod/HoofedAnimalsDataset.html

training images
Multi-Object Recognition

1. TREE MATCHING

$\dot{t}_1$ $\dot{t}_2$ $\dot{t}_3$ $\ldots$ $\dot{t}_M$
Multi-Object Recognition

1. TREE MATCHING

2. CLUSTERING
Multi-Object Recognition

1. TREE MATCHING

2. CLUSTERING
Multi-Object Recognition

1. TREE MATCHING

2. CLUSTERING
Multi-Object Recognition

1. TREE MATCHING

2. CLUSTERING
Multi-Object Recognition

1. TREE MATCHING

2. CLUSTERING
Multi-Object Recognition

1. TREE MATCHING

2. CLUSTERING
Multi-Object Recognition

1. TREE MATCHING

2. CLUSTERING
Multi-Object Recognition

1. TREE MATCHING

2. CLUSTERING
Multi-Object Recognition

1. TREE MATCHING

2. CLUSTERING
1. TREE MATCHING

2. CLUSTERING
Multi-Object Recognition

1. TREE MATCHING

2. CLUSTERING
Overview of Multi-Category Recognition

1. TREE MATCHING

2. CLUSTERING
Overview of Multi-Category Recognition

1. TREE MATCHING

2. CLUSTERING
Overview of Multi-Category Recognition

1. TREE MATCHING

2. CLUSTERING
Overview of Multi-Category Recognition

1. TREE MATCHING

2. CLUSTERING

3. TAXONOMY OF ALL DISCOVERED CATEGORIES WITH DIFFERENT COMPLEXITIES
Overview of Multi-Category Recognition

1. TREE MATCHING

2. CLUSTERING

3. TAXONOMY OF ALL DISCOVERED CATEGORIES WITH DIFFERENT COMPLEXITIES

4. RECOGNIZE SEGMENT EXPLAIN
Simultaneous Recognition and Segmentation
Simultaneous Recognition and Segmentation
Discriminative Learning of Object Parts

Discovery of subcategories in segmentation trees

Trees = points in the feature space of subcategories

Linear classifier

Rescaled axes

Todorovic & Ahuja CVPR08
Published, best categorizations on Caltech-101 and Caltech-256

Caltech-101:
- Ours
- Bosch-ICCV07
- Frome-ICCV07
- CaltechTechRep07
- Zhang-CVPR06
- Lazebnik-CVPR06
- Todorovic-CVPR06

Caltech-256:
- Ours
- Bosch-ICCV07
- CaltechTechRep07
- Todorovic-CVPR06

Average recognition rate versus number of training images per category.
1. Image representation = Hierarchy of regions

2. Region matching under unstable segmentations

3. Applications and results
   a. Object recognition
   b. Painterly rendering
   c. Texture segmentation
Multi-style Painterly Rendering

multistyle painting

Collaboration with Prof. Eugene Zhang at Oregon State University
Results: Video Object Segmentation

Brendel&Todorovic ICCV09
Results: Multi-style Painterly Rendering

multistyle without field

Collaboration with Prof. Eugene Zhang at Oregon State University
1. Image representation = Hierarchy of regions

2. Region matching under unstable segmentations

3. Applications and results
   a. Object recognition
   b. Video object segmentation
   c. Painterly rendering
   d. Texture segmentation
What is image texture?

...Repeated occurrence of image texture elements (or texels)...

[Beck ‘82]
Texture = Spatial Repetition of Texels

- Image texels = Images of physical texture elements
- Texels are not identical, only statistically similar
- Texel placement is not regular
Results: UIUC Texture Dataset

original image

extracted texels

Ahuja&Todorovic ICCV07
Results: UIUC Texture Dataset

original image  texel segmentation

Ahuja&Todorovic ICCV07
Results: Texture Segmentation

Todorovic&Ahuja ICCV09
Results: Texture Segmentation

original image

textel-based
Todorovic&Ahuja ICCV09

filter-based
Galun et al ICCV03
Summary

- Hierarchical region-based image representation
Summary

• Hierarchical region-based image representation

• Robust matching of regions
Summary

• Hierarchical region-based image representation

• Robust matching of regions

• Operative definition of an object category
Summary

- Hierarchical region-based image representation
- Robust matching of regions
- Operative definition of an object category
- Hierarchical taxonomy of shared categories
Summary

• Hierarchical region-based image representation
• Robust matching of regions
• Operative definition of an object category
• Hierarchical taxonomy of shared categories
• The framework allows:
  • Simultaneous recognition and segmentation
  • Semantic basis of recognition
  • Space-time coherent video object segmentation
  • Texel-based texture analysis
Thank you!