Connected Segmentation Tree -- A Joint Representation of Region Layout and Hierarchy

Narendra Ahuja and Sinisa Todorovic
{n-ahuja, sintod}@uiuc.edu

ILLINOIS
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

BECKMAN INSTITUTE
AT THE UNIVERSITY OF ILLINOIS

CVPR 2008

MOTIVATION

Objects in 3D world: **Their image projections:**

cohesive ⟶ regions of contiguous pixels = 2D objects
have distinct locations & finite volume ⟶ geometric properties: shape, area
different materials on their surfaces ⟶ photometric properties: color, texture
structure of surface materials ⟶ layout and embedding of subregions within regions
comprised of other objects = parts ⟶ spatial layout and hierarchy of 2D objects

PROBLEM STATEMENT

GIVEN images, each containing >= 0 occurrences of an object category
IDENTIFY all regions occupied by the category in the image set
LEARN a model of the category that JOINTLY captures:
- Geometric properties (e.g., shape, area, relative displacements)
- Photometric properties (e.g., intensity contrast along the boundary)
- Embedding (or containment) relationships
- Neighbor relationships and their strength

of the regions identified to represent category occurrences.

GIVEN a new image, use the learned category model to
DETECT, RECOGNIZE, SEGMENT any occurrences of the category.

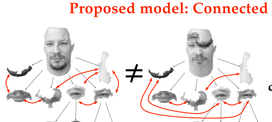
PRIOR WORK vs. PROPOSED OBJECT REPRESENTATION

bag of keypoints ↓ no spatial information

planar-graph models ↓ preselected same-scale parts

hierarchical models
e.g., segmentation tree [1, 2]
↓
the same segmentation tree for different layouts

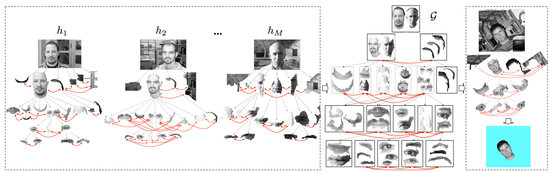
Proposed model: Connected Segmentation Tree



hierarchy of adjacency graphs ↓ different connected segmentation trees for different layouts

modeling properties of regions occupied by 2D objects	photometric	geometric	spatial layout	embedding
bag of keypoints	✔	✘	✘	✘
planar-graph models	✔	✔	✔	✘
hierarchical models	✔	✔	✘	✔
Connected Segmentation Tree	✔	✔	✔	✔

OVERVIEW OF OUR APPROACH



1) Images = Connected segmentation trees (CST) ⇒ Similar objects = Similar subgraphs
2) Similarity defined in terms of region properties:
   - Geometric and photometric
   - Strengths of region neighbor relationships

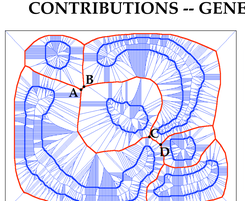
   and recursively the same properties of embedded subregions
3) Similar subgraphs found by using max-clique based graph matching
4) Maximally matching subgraphs are fused into a graph-union = CST category model
5) Matching the model with the CST of a new image:
   - Simultaneous recognition and segmentation all category occurrences
   - Explanation of recognition in terms of object parts and their neighbor relationships

CONTRIBUTIONS -- DEFINITION OF REGION NEIGHBORLINESS

1) Regions are exposed to one other through their nearby boundary segments
2) If boundary parts of two regions are:
   - Visible to each other
   - Nearby
   - Sufficiently far from other region boundaries

   } ⇒ The two regions are called neighbors
3) Relative degrees of boundary exposure to neighbors = Strengths of a region's neighborliness
4) Region neighborliness is asymmetric

CONTRIBUTIONS -- GENERALIZED VORONOI DIAGRAM



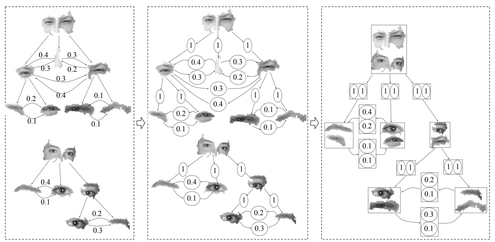
Boundary pixels and their corresponding standard Voronoi polygons for points

Generalized Voronoi polygons for regions
⇕
Union of Voronoi polygons of boundary pixels

- Regions are neighbors if their generalized Voronoi polygons touch
- Strength of neighborliness = Percentage of a Voronoi polygon's perimeter that is shared

CONTRIBUTIONS -- GENERALIZED MAX-CLIQUE GRAPH MATCHING

Problem: How to match graphs whose both edges and nodes are weighted?



Given two CST image representations: $G = (V, E)$ and $G' = (V', E')$
find structure-preserving bijection: $f = \{(v, v')\} \subset V \times V'$
which maximizes their similarity measure defined as

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

unary potential = function of region intrinsic properties

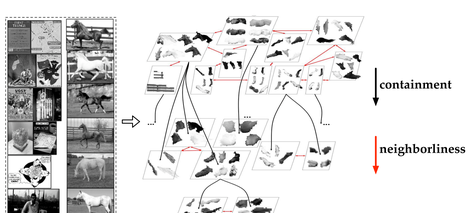
pairwise potential = function of region neighborliness and embedding

Properties of the proposed algorithm:

- It matches regions with regions, and separately region spatial relationships with corresponding relationships
- The maximum common subgraph, found by the algorithm, preserves the original structure of input graphs
- It seeks legal region-region and relationship-relationship matches whose unary and pairwise potentials are large

Theorem: Maximum subgraph isomorphism between two graphs with weighted edges and nodes
⇕
Maximum weight clique of their generalized association graph

EXAMPLE: CST MODEL OF WEIZMANN HORSES

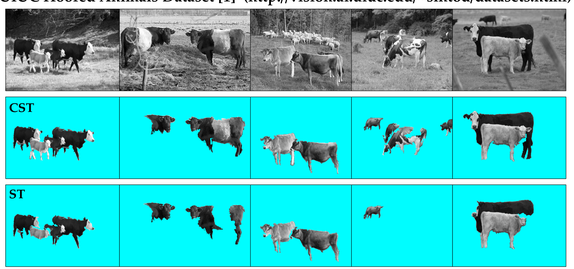


training set with positive and negative examples

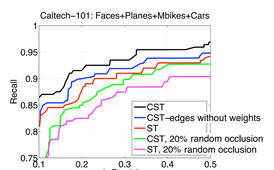
sample parts and their spatial relations in the learned CST model

RESULTS

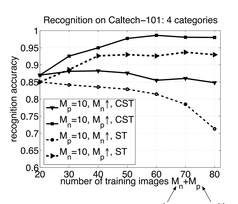
UIUC Hoofed Animals Dataset [1] (http://vision.ai.uiuc.edu/~sintod/datasets.html)



Simultaneous recognition and segmentation of category "cow" represented by the CST model (middle row), and the segmentation-tree (ST) model [1, 2] (bottom row). CSTs outperform STs.



Simultaneous recognition and segmentation of category "car"



Characteristics of recognition and segmentation:

1) Invariant to translation and in-plane rotation
2) Robust against a certain degree of:
   - Scale variations
   - Illumination changes
   - Object articulation
   - Partial occlusion
   - Background clutter
3) Segmentation is good when object boundaries are:
   - Jagged
   - Blurred
   - Form complex topologies

	LabelMe Trees	LabelMe Buildings	LabelMe Cars	Weizmann Horses	Horses	Cows	Deer	Sheep	Goats	Camels
Recall	47.6±6.9	92.6±6.9	67.6±6.9	91.9±5.2	81.2±10.3	78.4±4.2	88.1±6.9	81.2±5.3	78.2±8.6	89.9±7.2
Seg. error	41.6±7.9	34.6±13.4	32.5±8.2	7.2±2.5	15.9±5.3	17.1±4.6	11.1±8.4	24.8±7.2	20.1±8.1	11.5±5.1
Rec. error	19.7±3.8	11.6±2.9	12.9±4.8	7.9±4.1	7.8±4.2	6.5±6.2	7.7±3.4	7.8±4.1	12.2±5.4	3.2±3.9

ACKNOWLEDGMENT

The support of the National Science Foundation under grant NSF IIS 07-43014 is gratefully acknowledged

REFERENCES

[1] N. Ahuja and S. Todorovic, "Learning the taxonomy and models of categories present in arbitrary images," in ICCV 2007
[2] S. Todorovic and N. Ahuja, "Unsupervised category modeling, recognition and segmentation in images," in IEEE TPAMI, 2008