Object Segmentation as a Concurrent Grouping Problem

Stella X. Yu

Computer Science Division
University of California, Berkeley
Goal of Object Segmentation
Intuition

An object pops out from background if it is familiar and salient.
Intuition

An object pops out from background if it is familiar and salient.

- It has parts laid out in familiar configurations.
- It has boundaries well defined by intensity or texture.
- Each part has its corresponding image support.
Intuition

An object pops out from background if it is familiar and salient.

- It has parts laid out in familiar configurations.
- It has boundaries well defined by intensity or texture.
- Each part has its corresponding image support.

Object segmentation as a concurrent grouping problem:

- good grouping at object-part level
- good grouping at pixel level
- both subject to part-pixel ownership constraints
Grouping by Graph Cuts: 1. Representation
Grouping by Graph Cuts: 1. Representation

Step 1: Build a weighted graph $W$

$$W(p, q) = e^{-\frac{\text{intensity difference}(p, q)^2}{\sigma}}$$
Grouping by Graph Cuts: 1. Representation

Step 1: Build a weighted graph $W$

$$W(p, q) = e^{-\frac{\text{intensity difference}(p, q)^2}{\sigma}}$$
Grouping by Graph Cuts: 2. Criterion

\[
\min \text{ normalized cuts} = \frac{\min \text{ between-group affinity}}{\max \text{ within-group affinity}} \quad (\text{ShiMalik97})
\]

Step 2: Seek an optimal node partitioning
Grouping by Graph Cuts: 3. Solution

eigenvectors of $W$

Step 3: Find near-global optima to this NP-complete problem
Step 3: Find near-global optima to this NP-complete problem

eigenvectors of $W \implies$ discretization (YuShi03)
Constrained Concurrent Grouping Model

Patch Detection
Constrained Concurrent Grouping Model

Patch Detection

Edge Detection
Constrained Concurrent Grouping Model

Patch Detection

Expectation

Constraints

Evidence

Constraints

Edge Detection
Constrained Concurrent Grouping Model
Object Parts for a Nearest Neighbour Detector

Parts that maximize discrimination among classes are selected in training images.

Patch Affinity Encodes Spatial Compatibility \( B \)

Patches with mis-aligned object silhouettes have low affinity.
Pixel Affinity Encodes Intensity Similarity $A$

Pixels intersected by edges have low affinity.

$A(1, 3) \approx 1$

$A(1, 2) \approx 0$
Each patch claims a set of pixels based on training images.
Patch-Pixel Ownership Constraints $C$

Refined patch-pixel ownerships based on both patch and pixel affinity.
An Algebraic Description of the Model

Build a weighted graph:

\[ W = \begin{bmatrix} A \\ B \end{bmatrix} \]
An Algebraic Description of the Model

Build a weighted graph:

$$W = \begin{bmatrix} A & \cdot \\ \cdot & B \end{bmatrix}$$

Seek a joint partitioning:

$$\begin{bmatrix} X \\ Y \end{bmatrix}$$
An Algebraic Description of the Model

Build a weighted graph:

\[ W = \begin{bmatrix} A & B \end{bmatrix} \]

Seek a joint partitioning:

\[ \begin{bmatrix} X \\ Y \end{bmatrix} \]

Subject to constraints:

\[ Y = CX \]
How Object Knowledge Helps Segmentation

pixel grouping alone

pixel-patch grouping
How Segmentation Helps Object Detection

image | patch density | segmentation
Equally Applicable to Multiple Objects
When Does Our Method Fail

image  patch density  segmentation

image  patch density  segmentation
Retrospective: Boundaries + Alignment = Object

Perceiving Shapes through Region and Boundary Interaction
Stella X. Yu and Jianbo Shi
Technical Report CMU-RI-TR-01-21,
Robotics Institute, Carnegie Mellon University, July 2001

\( 41 \times 30 \) image

\( \lambda_2 = 1.00 \) \hspace{1cm} \( \lambda_1 = 1.00 \) \hspace{1cm} \( \lambda_1 = 1.08 \) \hspace{1cm} \( \lambda_1 = 1.20 \) \hspace{1cm} \( \lambda_1 = 1.33 \)

\( t = 0 \) \hspace{1cm} \( t = 0.25 \) \hspace{1cm} \( t = 0.5 \) \hspace{1cm} \( t = 0.75 \) \hspace{1cm} \( t = 1 \)

Appeal in theory: joint grouping of pixels and edgels
Difficulty in reality: curvilinearity, texture, specific shapes
Retrospective: Parts $+$ Alignment $=$ Object

Concurrent Object Recognition and Segmentation by Graph Partitioning
Stella X. Yu, Ralph Gross and Jianbo Shi
Neural Information Processing Systems, Dec 2002

Appeal in theory: joint grouping of pixels and body parts
Difficulty in reality: appearance variation in clothed body parts
Retrospective: Parts + Alignment = Object

Object-Specific Figure-Ground Segregation
Stella X. Yu and Jianbo Shi
IEEE Conference on Computer Vision and Pattern Recognition, June 2003

Appeal in theory: concurrent grouping of pixels and parts
Difficulty in reality: hallucination, incompleteness, insensitivity to pose
Acknowledgements

- Jianbo Shi
  Ralph Gross
  Shyjan Mahamud

- CMU HumanID Lab
  CMU Hebert Lab

- ONR N00014-00-1-0915
  NSF IRI-9817496
  NSF LIS 9720350
  NSF CAREER 9984706
  NIH EY 08098