



A Tour of Image Segmentation

Dr. Scott Cohen | Adobe Research



Image Segmentation

- Segment image pixels into different classes



- 2 Classes : Boy (Foreground), Not Boy (Background)

- Scott Cohen, Gregg Wilensky, Jeff Chien (Adobe)

Segmentation Variations

- How many classes?
- How are the classes defined?
- What features are used to compute the segmentation?
- Hard Segmentation or Soft Segmentation?
 - Hard: a pixel is assigned to exactly one class
 - Soft: a pixel may be assigned to more than one class
- Automatic or Interactive computation?
 - What user input is provided?
- How many images are segmented?

Interactive Binary Segmentation

- User Strokes (Scribbles) in **Foreground** and **Background**



User Intent

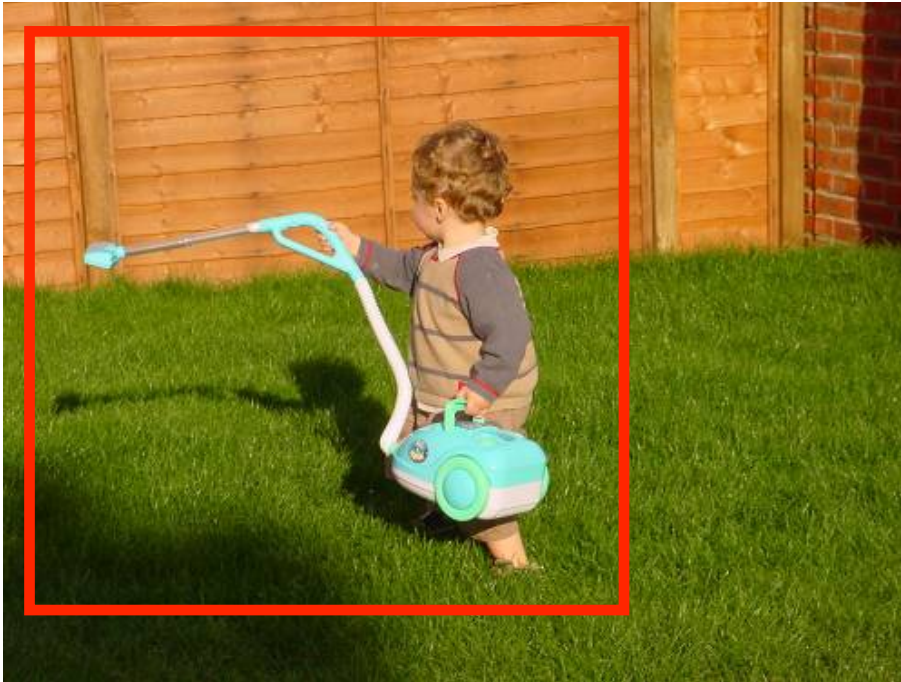


Desired Segmentation

Interactive Graph Cuts for Optimal Boundary & Region Segmentation
Boykov, Jolly (ICCV 2001)

Interactive Binary Segmentation

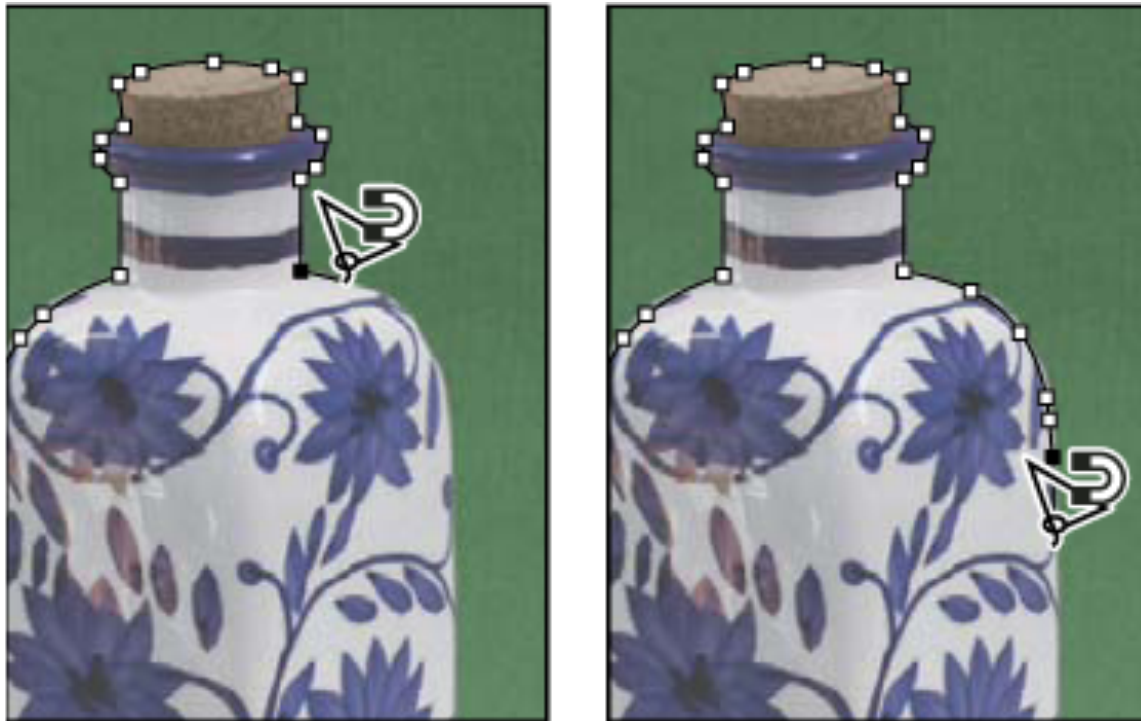
- GrabCut: Draw a rectangle around the object to select



GrabCut: Interactive Foreground Extraction using Iterated Graph Cuts
Rother, Kolmogorov, Blake (Siggraph 2004)

Interactive Binary Segmentation

- Magnetic Lasso: Trace around the object to select



Intelligent Scissors for Image Composition
Mortensen, Barrett (Siggraph 1995)

Automatic Binary Segmentation

- Segment the “Salient” Region



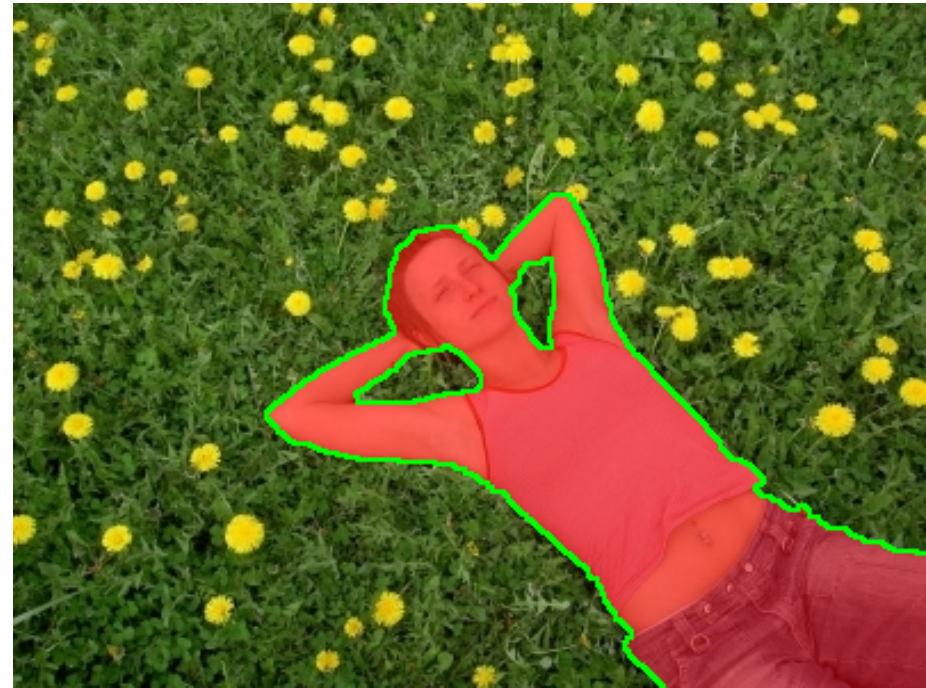
Automatic Binary Segmentation

- Segment the “Salient” Region



Automatic Binary Segmentation

- Segment the “Salient” Region



Automatic Binary Segmentation

- Segment In-focus Regions



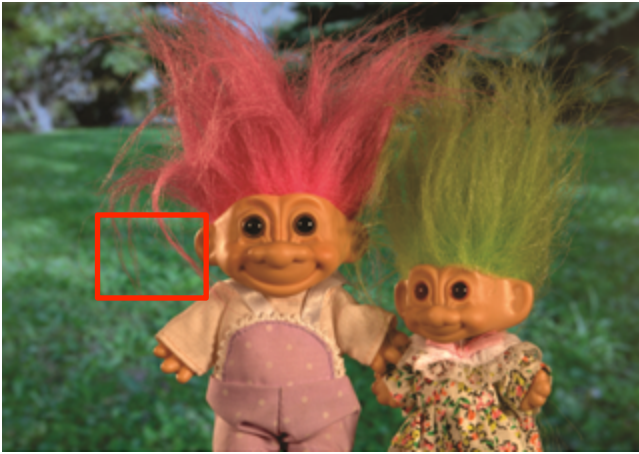
Input Image



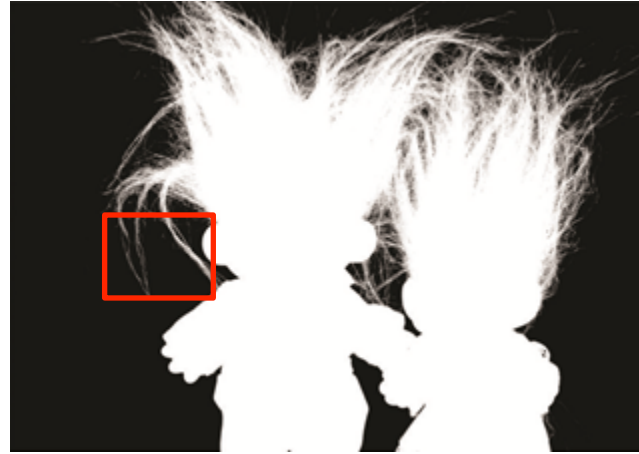
Output Segmentation

Image Matting

- Soft Binary Segmentation $I_p = \alpha_p F_p + (1 - \alpha_p) B_p$



Input Image I



Output Segmentation $\alpha \in [0, 1]$



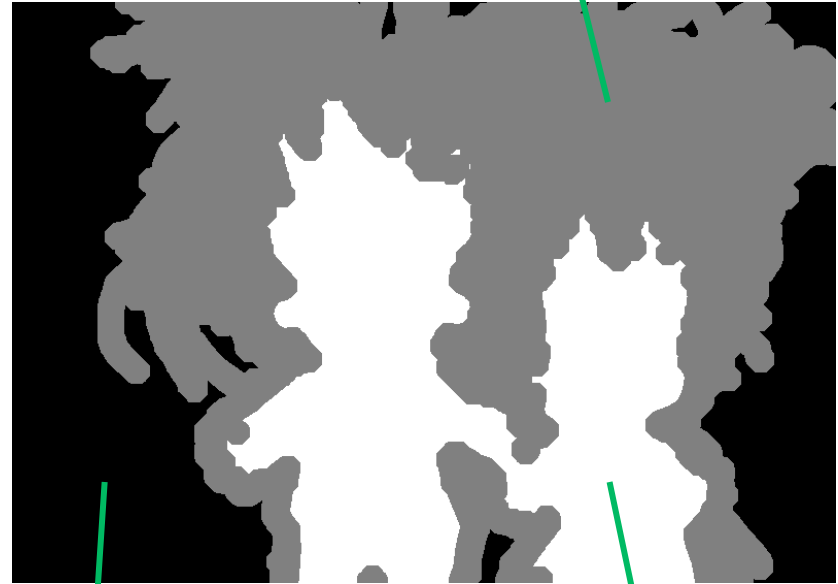
Image Matting : Compositing Application

$$I_p = \alpha_p F_p + (1 - \alpha_p) B_p \quad \hat{I}_p = \alpha_p^{01} H_p + (1 - \alpha_p^{01}) \hat{B}_{pp}$$



Image Matting : Trimap Input

$$I_p = \alpha_p F_p + (1 - \alpha_p) B_p$$



Unknown Region

Known Background

Known Foreground

Demo: Interactive Matting

- Brian Price, Scott Cohen (Adobe)

Co-Segmentation

- Segment the object in common in multiple images



Input Image Pair



Cosegmentation



Input Image pair



Cosegmentation

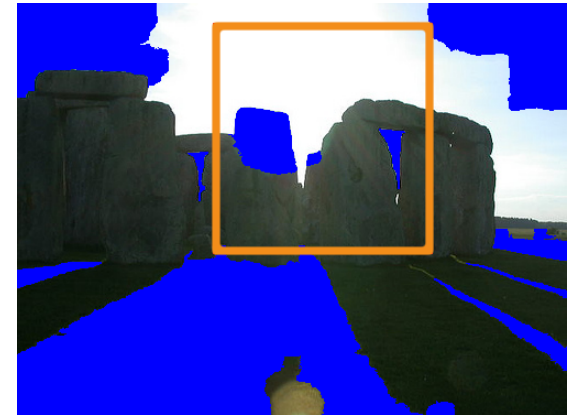
Cosegmentation of Image Pairs by Histogram Matching
Rother, Kolmogorov, Minka, Blake (CVPR 2006)

Co-Segmentation Methods

Method	Foregrounds	Backgrounds	Automatic or Interactive?	# of Images
Histogram Matching (CVPR06)	Same	Different	Automatic	2

Co-Segmentation : Similar BGs, Interactive, Many Images

iCoseg: Interactive Co-segmentation with Intelligent Scribble Guidance
Batra, Kowdle, Parikh, Luo, Chen (CVPR 2010)

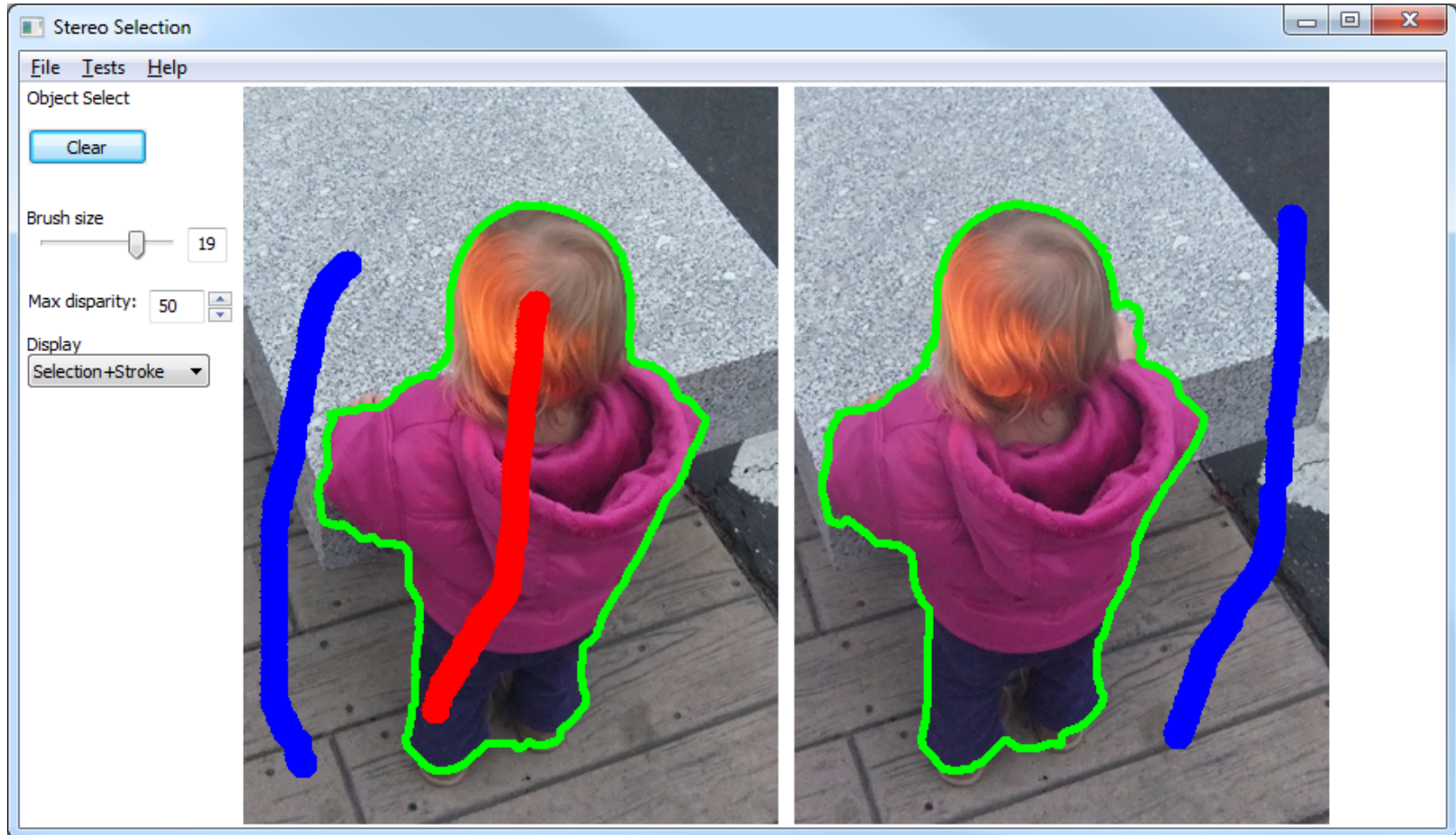


Co-Segmentation Methods

Method	Foregrounds	Backgrounds	Automatic or Interactive?	# of Images
Histogram Matching (CVPR06)	Same	Different	Automatic	2
iCoseg (CVPR10)	Similar	Similar	Interactive	Many

Stereo Co-Segmentation : Same BGs, Interactive, 2 Images

StereoCut: Consistent Interactive Object Selection in Stereo Image Pairs
Price, Cohen (ICCV 2011)

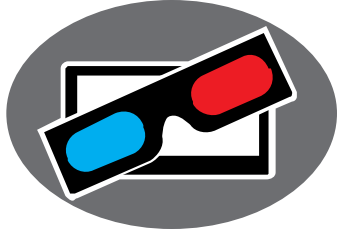
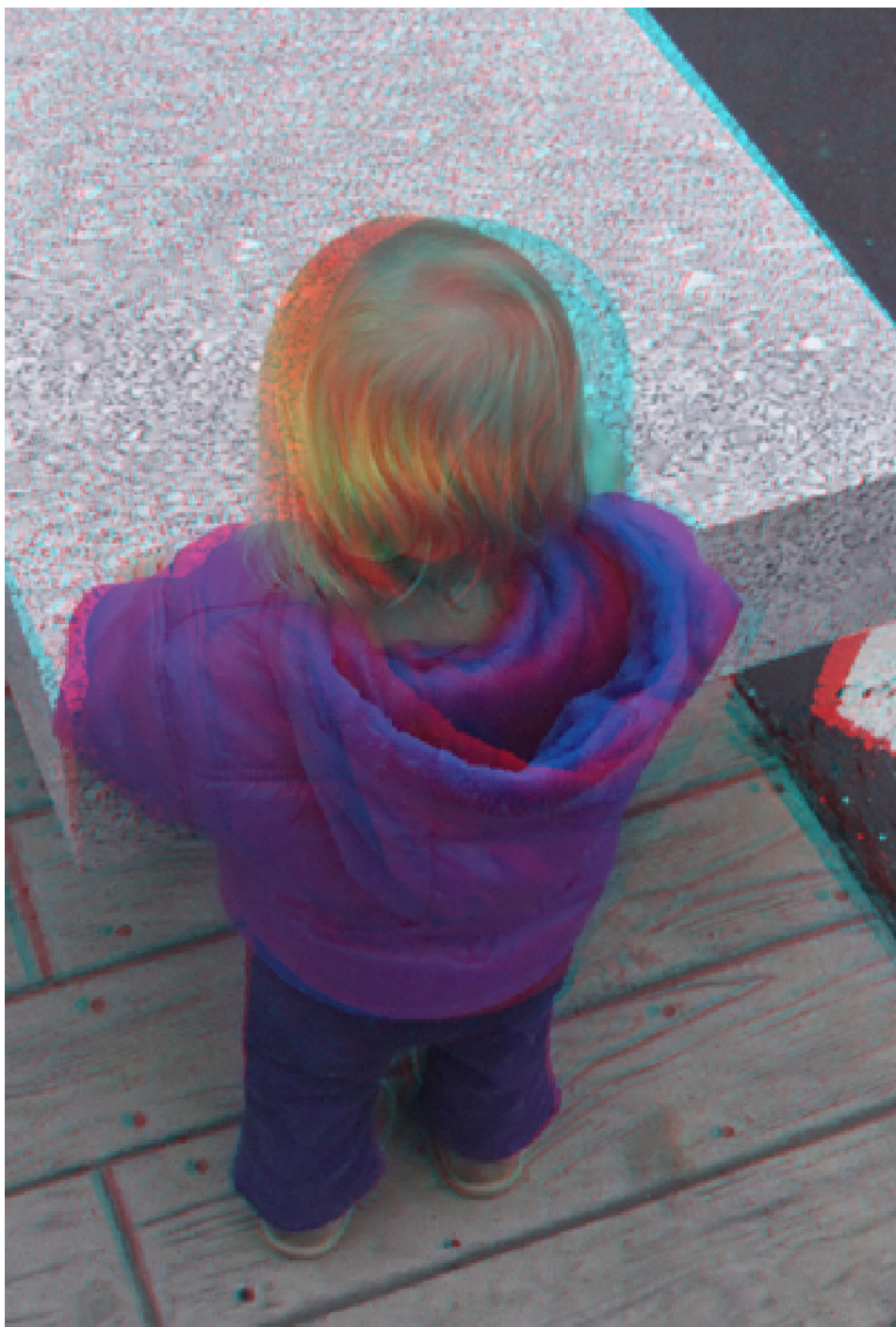


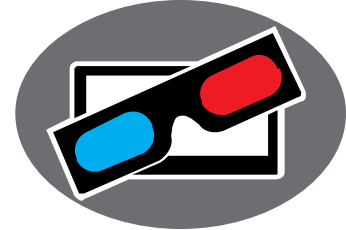
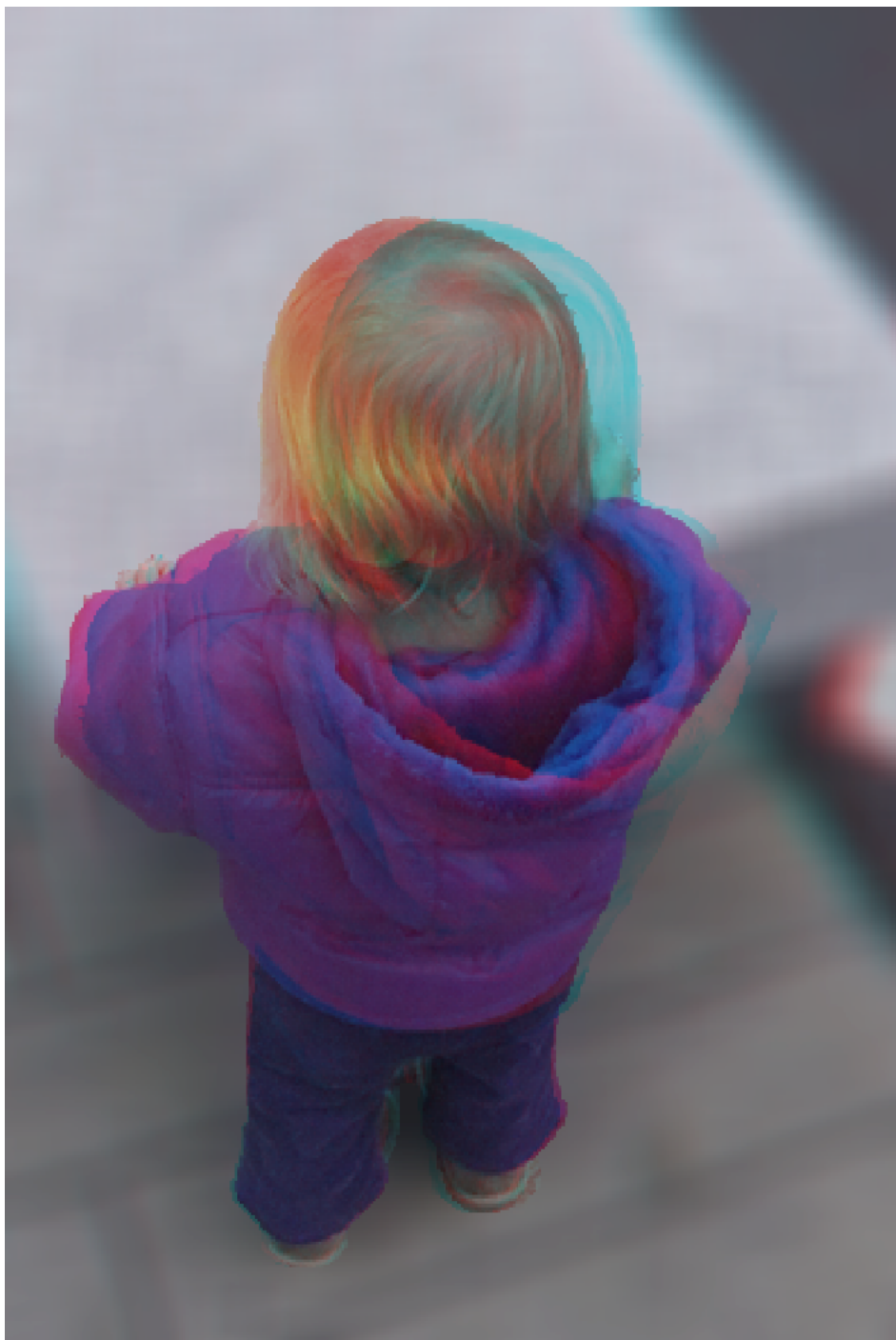
Stereo Co-Segmentation Applications

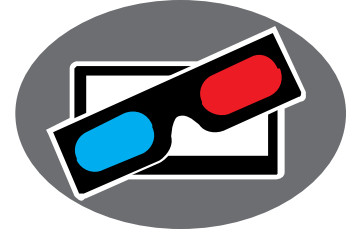
- Localized Stereo Editing
- Stereo Inpainting: remove Co-Segmented object from the stereo picture

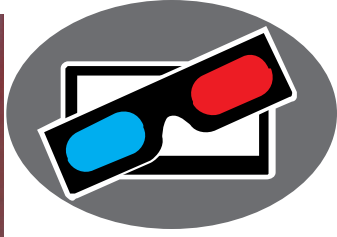


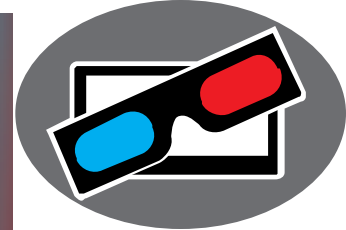
PatchMatch-based Content Completion of Stereo Image Pairs
Morse, Howard, Cohen, Price (3DimPVT 2012)

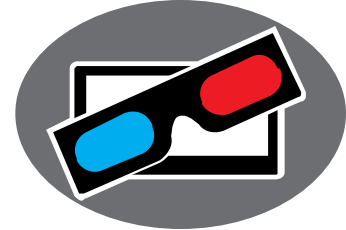


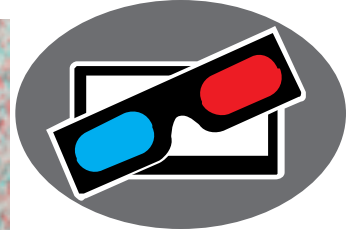












"On my first day, I
got hit badly by a
girl and her brother."
— Lella



WHAT'S YOUR IDENTITY?

How do you define yourself? By the clothes you wear and the products you buy. By the people you spend your time with. By the things you do. By the way you express yourself and make yourself heard.

SOUTH CAMDEN COMMUNITY SCHOOL

At the South Camden Community School, we believe in the power of education. We are committed to providing a high-quality, culturally relevant education for all students. Our school is a place where every student is valued and every voice is heard. We are proud to be a part of the South Camden community and to work together to create a bright future for all.

South Camden Community School
1000 South Camden Ave
Camden, NJ 08102
609.291.1234
www.southcamden.k12.nj.us

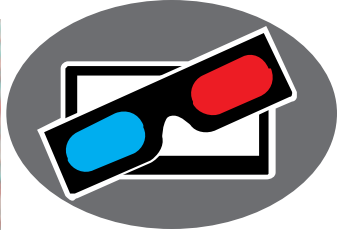
ST SEATTLE ELEMENTARY SCHOOL

St. Seattle Elementary School is a place where every student is valued and every voice is heard. We are committed to providing a high-quality, culturally relevant education for all students. Our school is a place where every student is valued and every voice is heard. We are proud to be a part of the St. Seattle community and to work together to create a bright future for all.



St. Seattle Elementary School
1000 St. Seattle Ave
St. Seattle, WA 98102
206.462.1234
www.stseattle.k12.wa.us





"On my first day, I got hit badly by a girl and her brother."
 — Lella

WHAT IS YOUR IDENTITY?

How do you define yourself? By the clothes you wear and the products you buy. By the people you spend your time with and make yourself heard.

SOUTH CAMDEN COMMUNITY SCHOOL

At the South Camden Community School, we believe in the power of education to transform lives. Our students are the future, and we are committed to providing them with the best possible education. We offer a variety of programs and services to support our students' learning and personal growth.

South Camden
 2007-2008

ST SEATTLE ELEMENTARY SCHOOL

At St. Seattle Elementary School, we believe in the power of education to transform lives. Our students are the future, and we are committed to providing them with the best possible education. We offer a variety of programs and services to support our students' learning and personal growth.

St. Seattle
 2007-2008















Co-Segmentation Methods

Method	Foregrounds	Backgrounds	Automatic or Interactive?	# of Images
Histogram Matching (CVPR06)	Same	Different	Automatic	2
iCoseg (CVPR10)	Similar	Similar	Interactive	Many
Stereo (ICCV11)	Same	Same	Interactive	2

Co-Segmentation : Similar BGs, Automatic, Many Images

Unsupervised Joint Object Discovery and Segmentation in Internet Images
Rubinstein, Joulin, Kopf, Liu (CVPR 2013)

“Car” Internet Search



Input

Output

Co-Segmentation Methods

Method	Foregrounds	Backgrounds	Automatic or Interactive?	# of Images
Histogram Matching (CVPR06)	Same	Different	Automatic	2
iCoseg (CVPR10)	Similar	Similar	Interactive	Many
Stereo (ICCV11)	Same	Same	Interactive	2
Object Discovery (CVPR13)	Similar, but more variation than iCoseg	Similar and Different	Automatic	Many

Scene Parsing | Semantic Segmentation

- Label each pixel in an image with its semantic category



Input Image



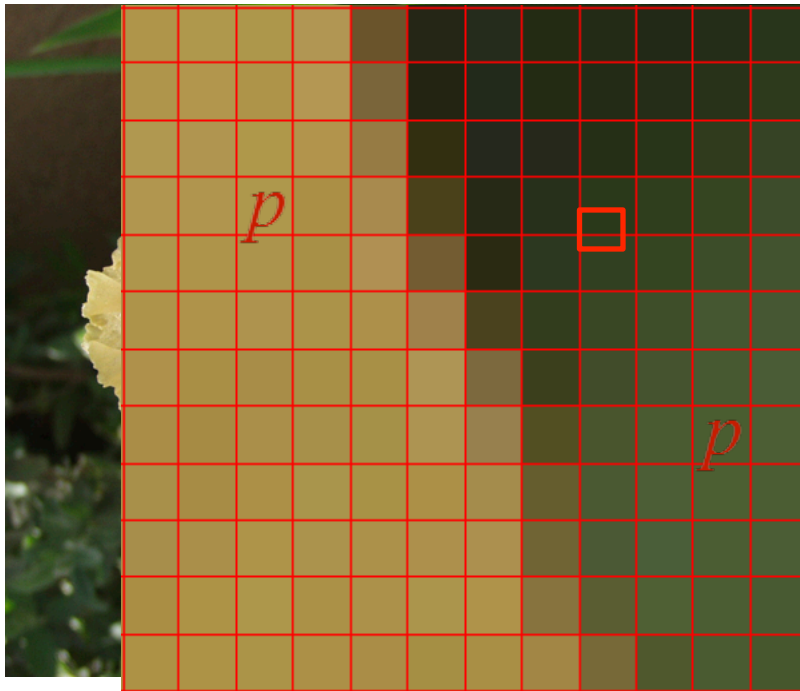
Desired Output

black	unlabeled
red	building
purple	car
cyan	person
light purple	pole
grey	road
brown	sidewalk
orange	sign
purple	trash can
green	tree

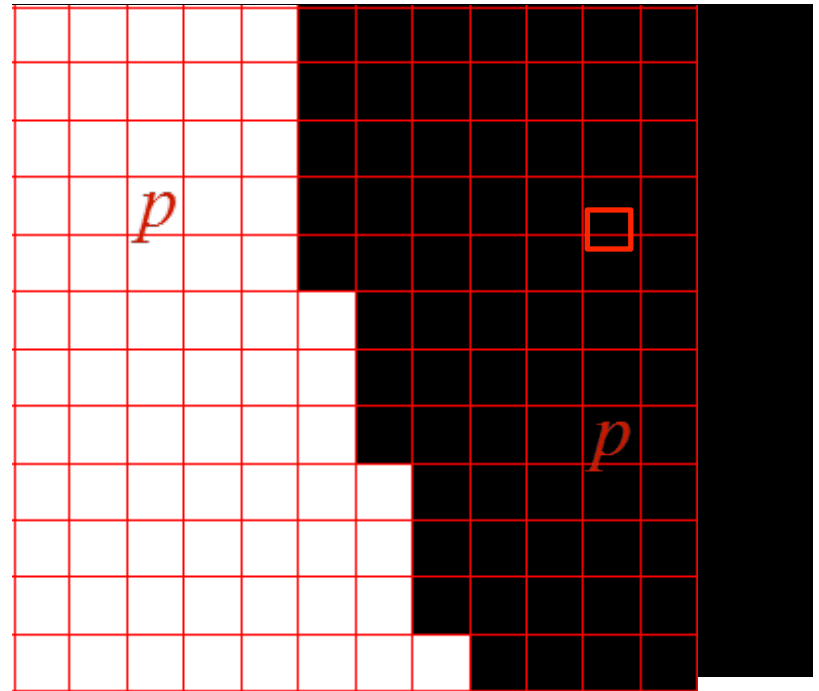
Details About Some Segmentation Methods

- Common Framework
- Notation: Segmentation $X = \{x_p\}$

$$I = \{I_p\}$$



$$x_p = 1 \quad X = \{x_p\} \quad x_p = 0$$

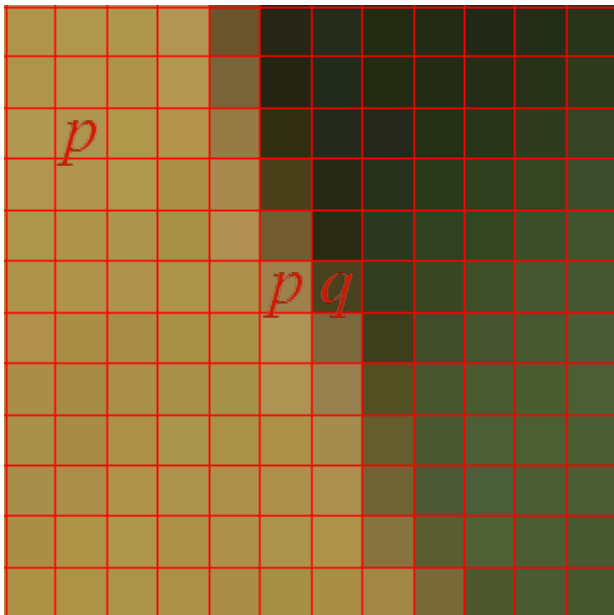


Common Energy Minimization Framework

- Energy function to measure quality of segmentation $X = \{x_p\}$

$$E(X) = \sum_p D_p(x_p) + \lambda \sum_{p,q \in N} V_{pq}(x_p, x_q)$$

- * Global minimum found by min graphcut / maxflow algorithms



Data Term

$D_p(0)$ = cost of labeling pixel p as BG

$D_p(1)$ = cost of labeling pixel p as FG

Smoothness Term

$V_{pq}(1,0)$ = cost of p as FG, q as BG

$V_{pq}(0,1)$ = cost of p as BG, q as FG

$V_{pq}(0,0)$ = cost of p as BG, q as BG = 0

$V_{pq}(1,1)$ = cost of p as FG, q as FG = 0

Energy E for Stroke-based Binary Segmentation using Color

- K-means on FG Strokes



- K-means on BG Strokes



- \rightarrow Probability $P_p(FG | I_p)$

$$\begin{aligned} D_p(FG) &= 0 \text{ if } p \in F \\ &= \infty \text{ if } p \in B \\ &= -\log P_p(FG | I_p) \text{ otherwise} \end{aligned}$$

$$\begin{aligned} D_p(BG) &= 0 \text{ if } p \in B \\ &= \infty \text{ if } p \in F \\ &= -\log P_p(BG | I_p) \text{ otherwise} \end{aligned}$$



Energy E for Binary Segmentation: Smoothness

$$V_{pq}(BG, BG) = V_{pq}(FG, FG) = 0$$

- Encourage segmentation boundaries to occur at image edges

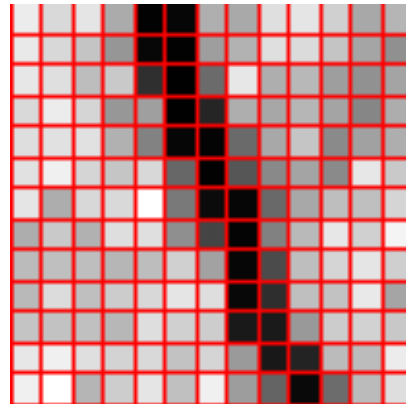
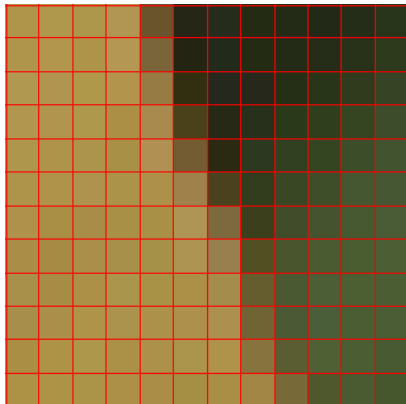
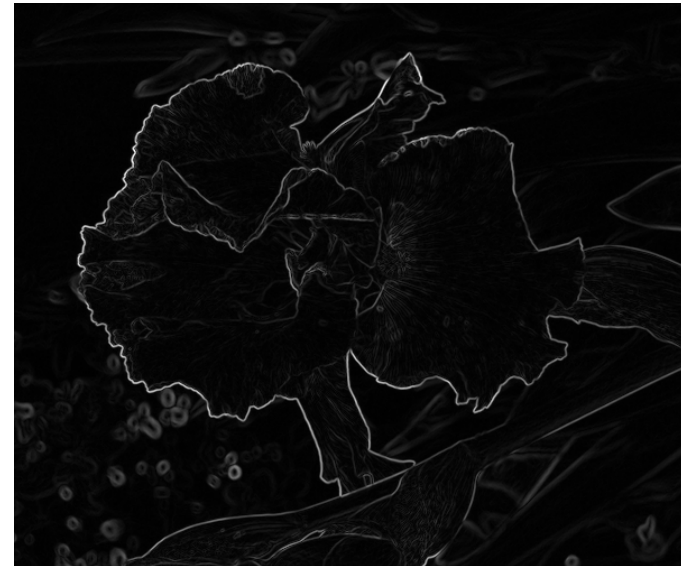
$$V_{pq}(FG, BG) = \exp\left(-\frac{\|I_p - I_q\|_2^2}{2\sigma^2}\right)$$

$$V_{pq}(BG, FG) = V_{pq}(FG, BG)$$

I



$\|\nabla I\|$



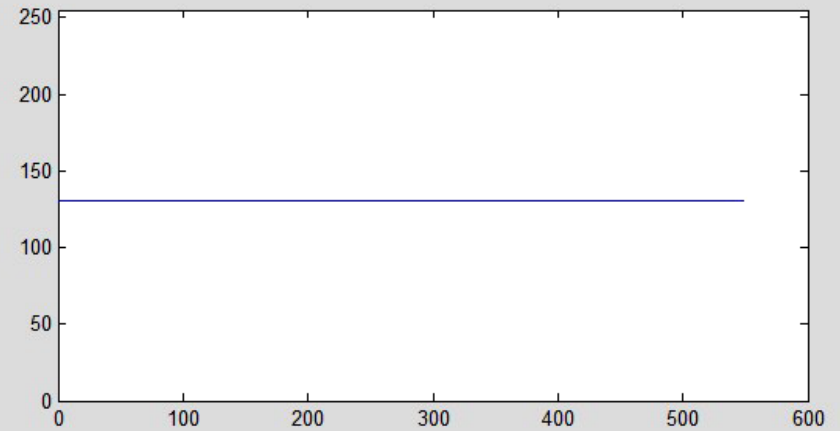
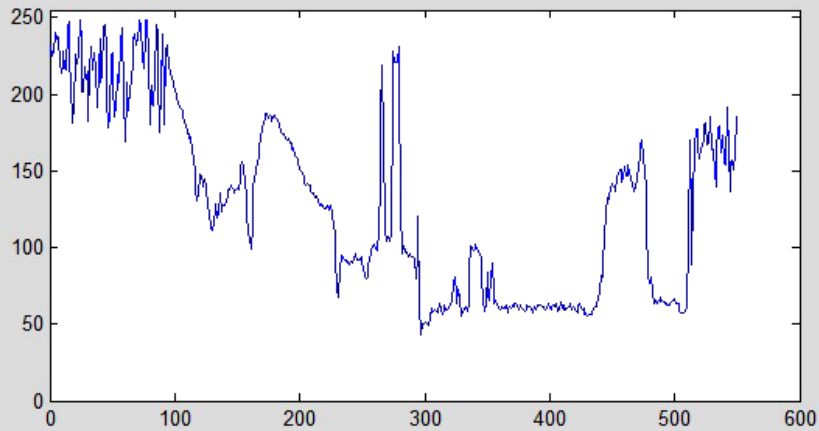
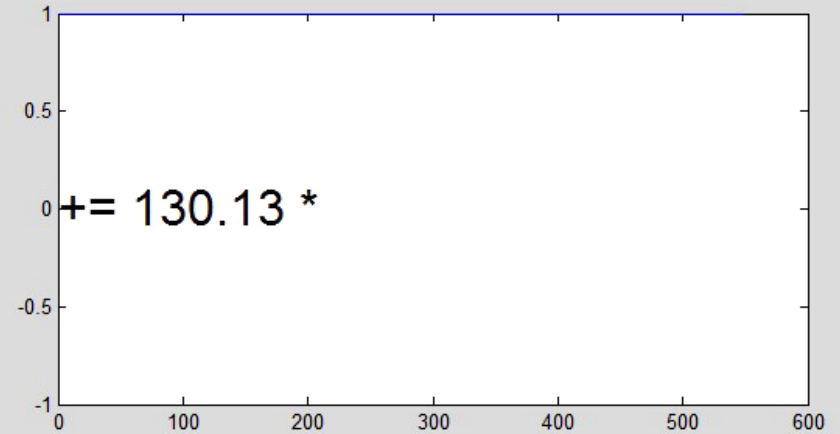
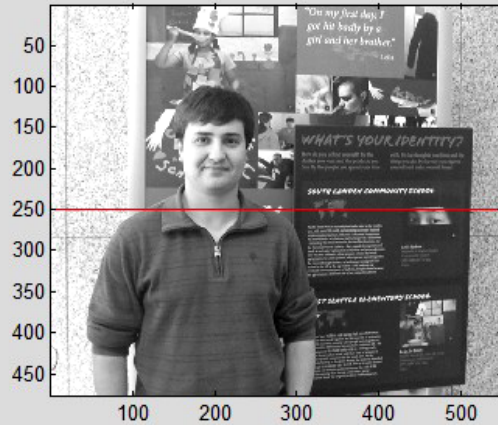
Why is V called the smoothness term?

$$E(X) = \sum_p D_p(x_p) + \lambda \sum_{p,q \in N} V_{pq}(x_p, x_q)$$

$$P_p(x_p) + \lambda \sum_{p,q \in N} V_{pq}(x_p, x_q)$$

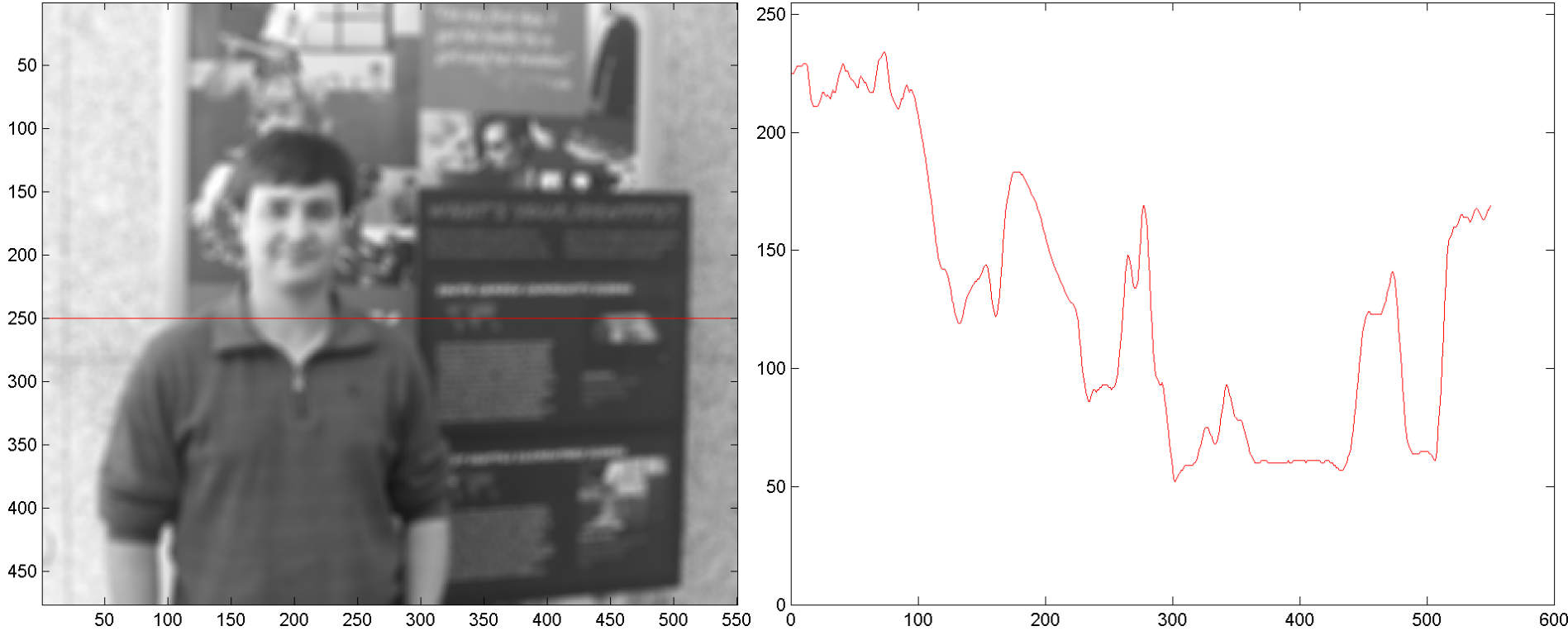


In-Focus Segmentation: Frequency Decomposition

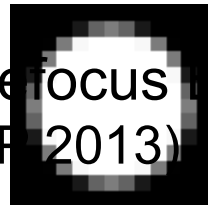


In-Focus Segmentation: Modeling Defocus Blur

- Larger blurs remove higher frequencies



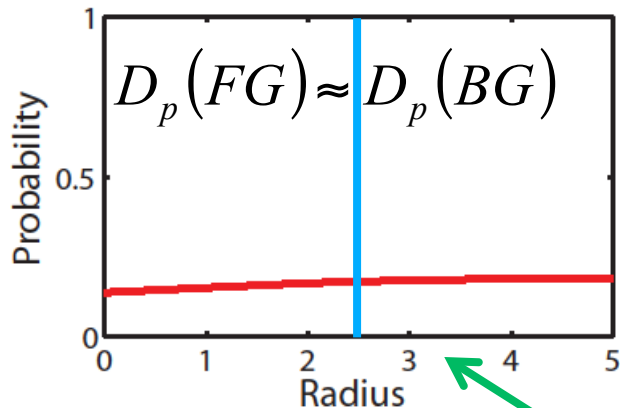
Estimating Spatially Varying Defocus Blur from A Single Image
Zhu, Cohen, Schiller, Milanfar (TIP 2013)



Out-of-Focus Blur Estimation

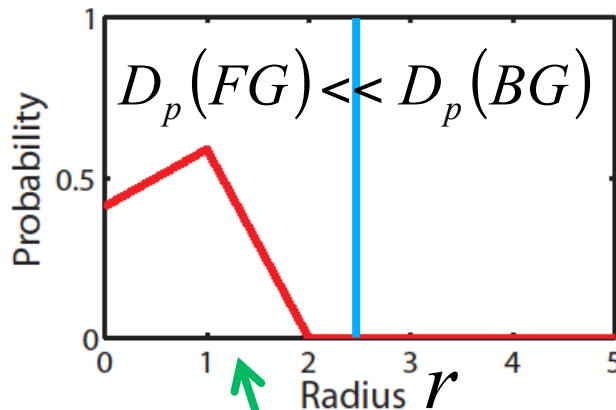
$(x,y) = (54,52)$

most probable radius = 5



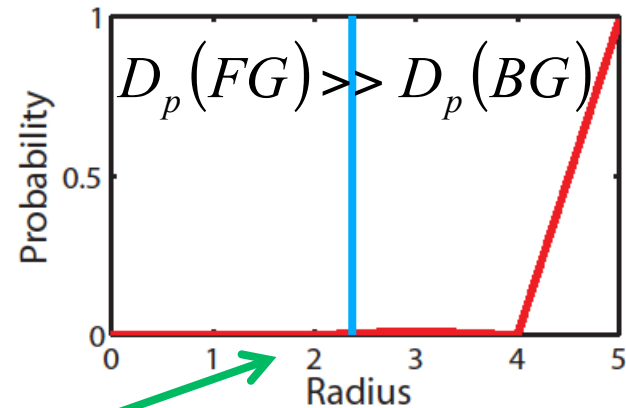
$p = (x,y) = (193,223)$

most probable radius = 1

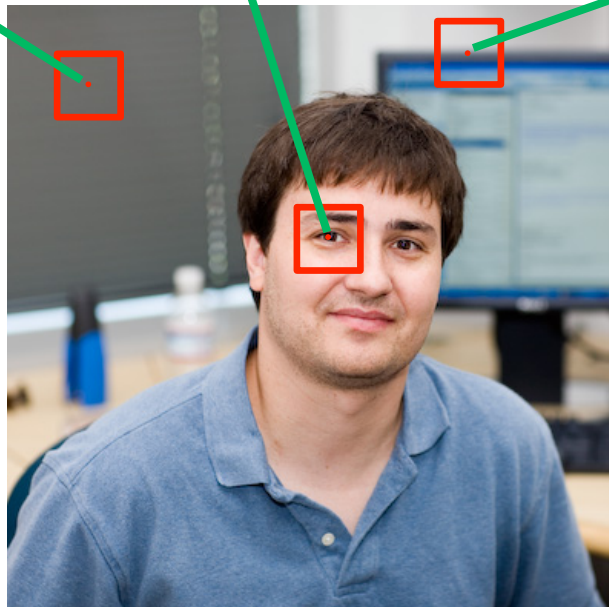


$(x,y) = (302,31)$

most probable radius = 5



- Measure power across frequencies



Automatic In-Focus Segmentation Results

$$E(X) = \sum_p D_p(x_p) + \lambda \sum_{p,q \in N} V_{pq}(x_p, x_q)$$



Automatic In-Focus Segmentation Results



Automatic In-Focus Segmentation Results

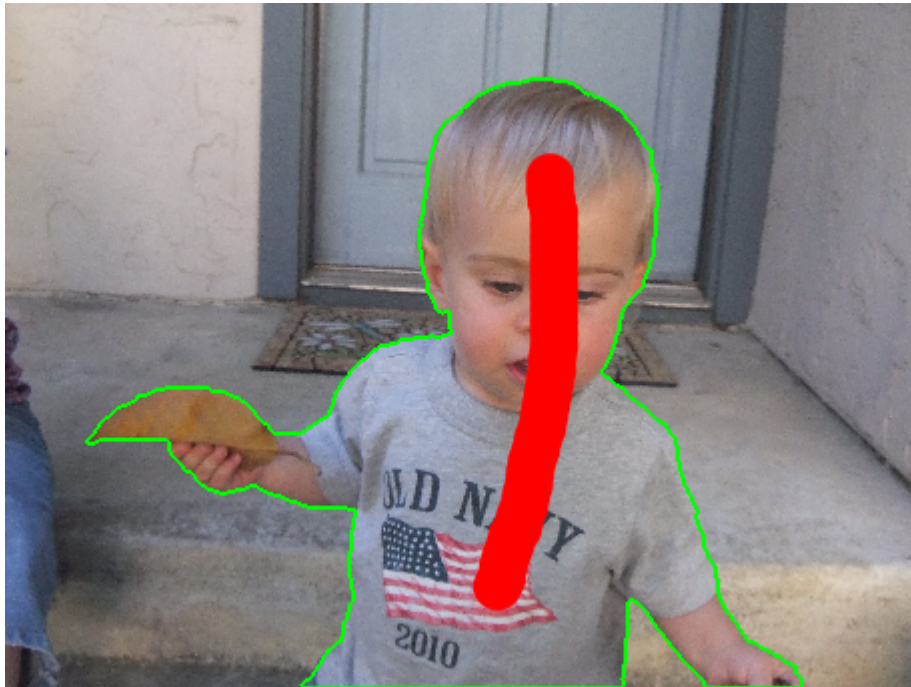
- Works when foreground and background colors are similar



Automatic In-Focus Segmentation Results



Interactive Stereo Co-Segmentation



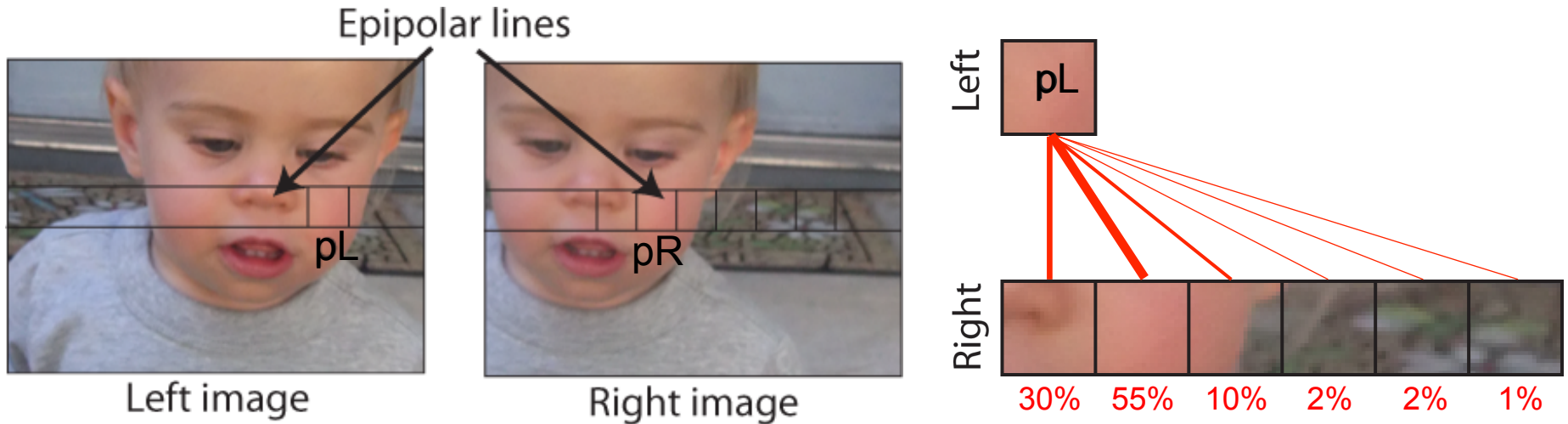
Left



Right

Energy E for Stereo Segmentation

- Add an energy term to include correspondence information



- If (pL, pR) likely correspond, then their labels should be same

$$E(X) = \sum_p D_p(x_p) + \lambda \sum_{p, q \in N} V_{pq}(x_p, x_q) + \mu \sum_{pL, pR} P_{match}(pL, pR) |x_{pL} - x_{pR}|$$

Stereo Segmentation Results

File Tests Edit Help

Calc Disparity Edit Disparity **Object Select** Discontinuity

Object Select

Clear

Brush size

23

Display

Selection+Stroke



Scene Parsing | Semantic Segmentation

- Label each pixel in an image with its semantic category



Input Image

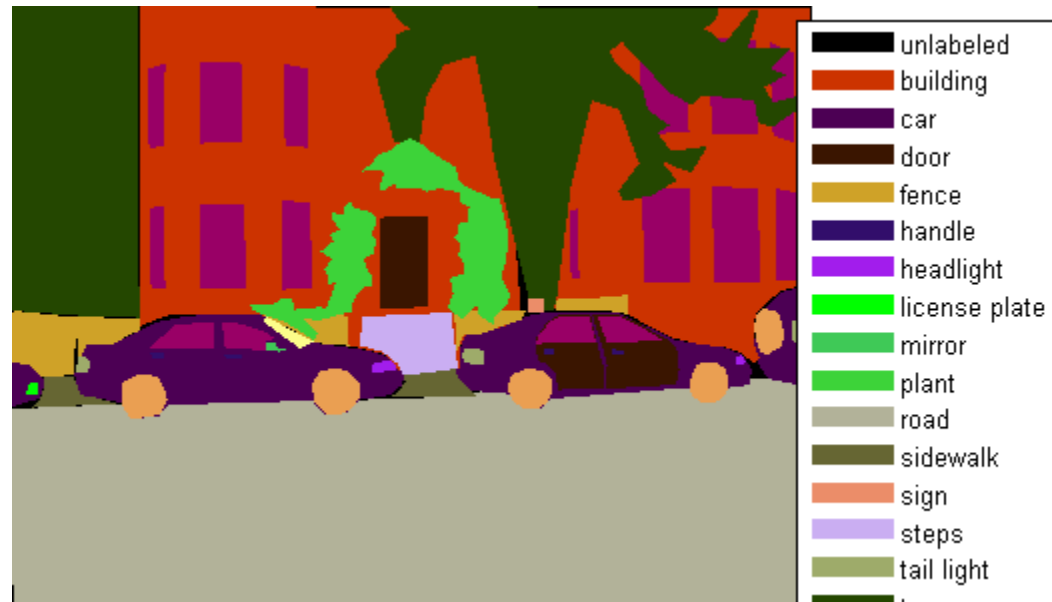


Desired Output

black	unlabeled
red	building
purple	car
cyan	person
light purple	pole
grey	road
brown	sidewalk
orange	sign
purple	trash can
green	tree

Scene Parsing | Semantic Segmentation

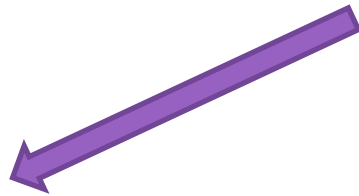
- Training Data : Labeled Images (Input)
- SIFTflow : 2488 Labeled Training Images, 33 Classes
- LMSun : 45176 Labeled Training Images, 232 Classes



Input Image



Image Retrieval



Sea, sand, sky, mountain,
field, tree, rock, plant, road,
grass, boat, river, person

$P_p(x_p = \text{sea}), P_p(x_p = \text{sand}), \dots$

Similar Training Images



Scene Parsing | Semantic Segmentation

- Tighe, Lazebnik. Finding things: Image parsing with regions and per-exemplar detectors. CVPR13

Cars from Similar Training Images

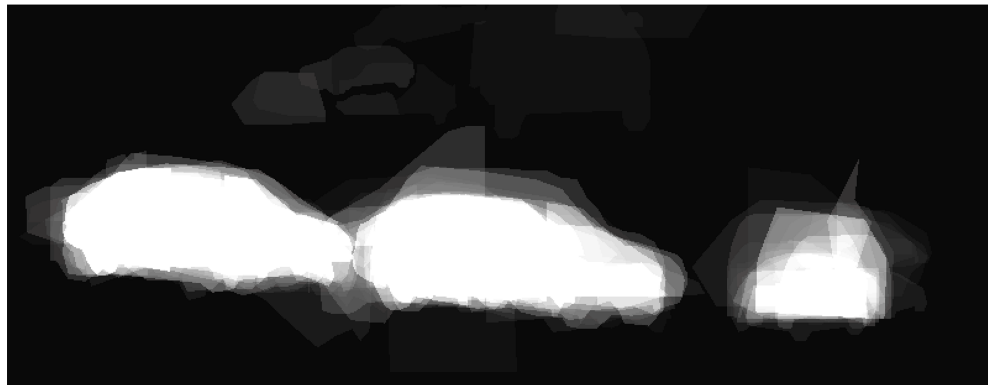


Test Image



Likelihood of Car

$$P_p(x_p = \text{car})$$

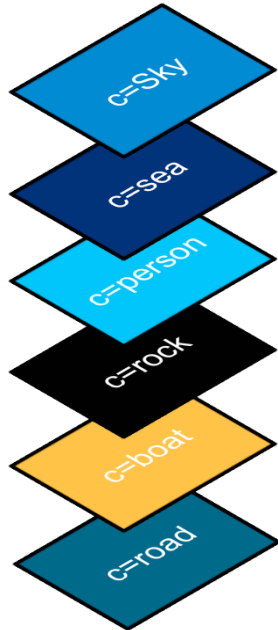


Energy E for Scene Parsing | Semantic Segmentation

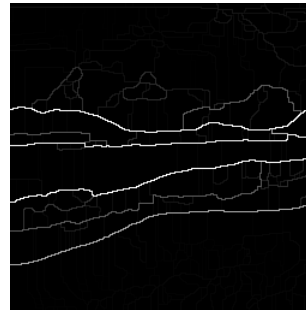
$$E(X) = \sum_p D_p(x_p) + \lambda \sum_{p,q \in N} V_{pq}(x_p, x_q)$$

$$P_p(x_p = c)$$

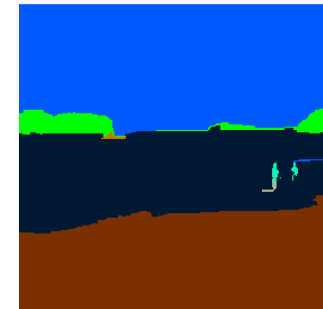
- $V_{pq}(\text{sea, sand})$
- $V_{pq}(\text{sky, rock})$
- \vdots
- $V_{pq}(\text{sofa, sea})$
- \vdots



+



→



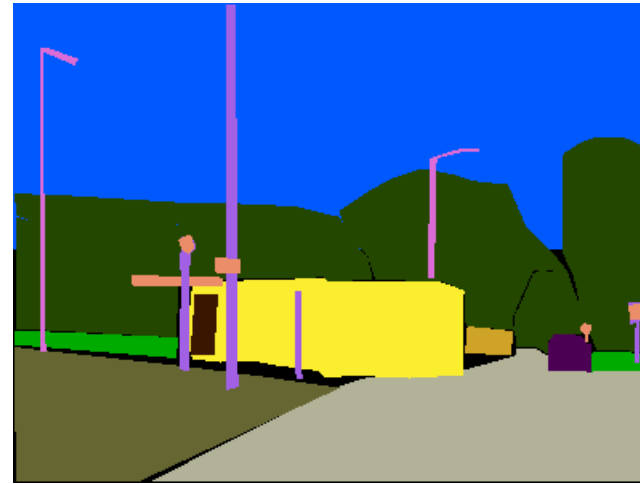
- boat
- person
- road
- rock
- sand
- sea
- sky

2 p=0,91517

Scene Parsing | Semantic Segmentation Results



Image



- unlabeled
- bus
- car
- door
- fence
- grass
- pole
- road
- sidewalk
- sign
- sky
- streetlight
- tree

Human Annotation



- building
- bus
- car
- grass
- pole
- road
- sky
- tree
- van

Tighe CVPR13

Scene Parsing | Semantic Segmentation Results



Image



Human Annotation



Tighe CVPR13

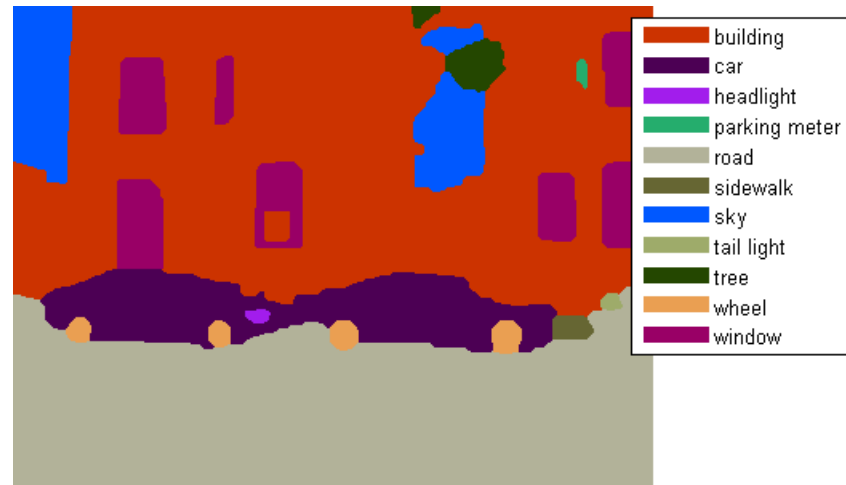
Scene Parsing | Semantic Segmentation Results



Image



Human Annotation

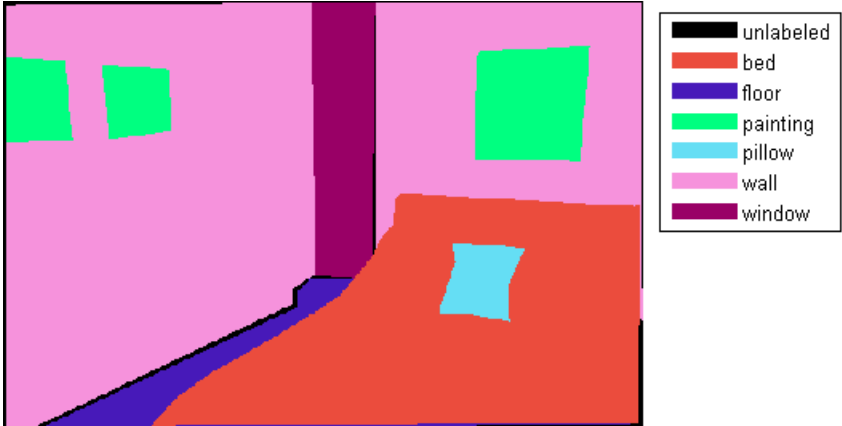


Tighe CVPR13

Scene Parsing | Semantic Segmentation Results



Image



Human Annotation



Tighe CVPR13

Summary and Conclusion

- There are many variations of segmentation problems
 - How many classes? What are the classes?
 - What features are used?
 - Hard or Soft Segmentation?
 - Automatic or Interactive? What User Input?
 - How many images are segmented?

- Semantic Segmentation: still a lot of work to be done



Adobe