ICME 2016, Paper #254

Salient Object Detection for RGB-D Image via Saliency Evolution

Jingfan Guo, Tongwei Ren, Jia Bei

Nanjing University

July 12, 2016

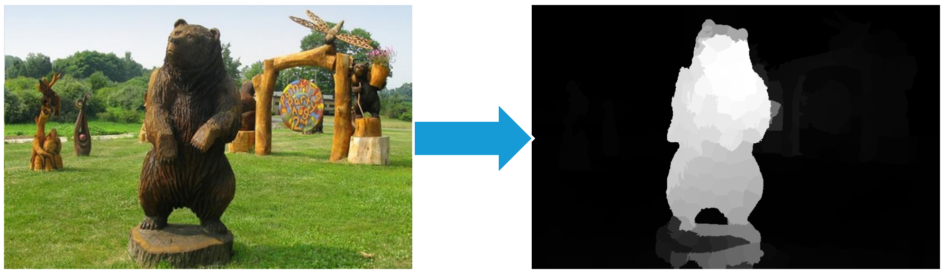
MAGUS

Multimedia AnalyzinG and UnderStanding

Salient object detection

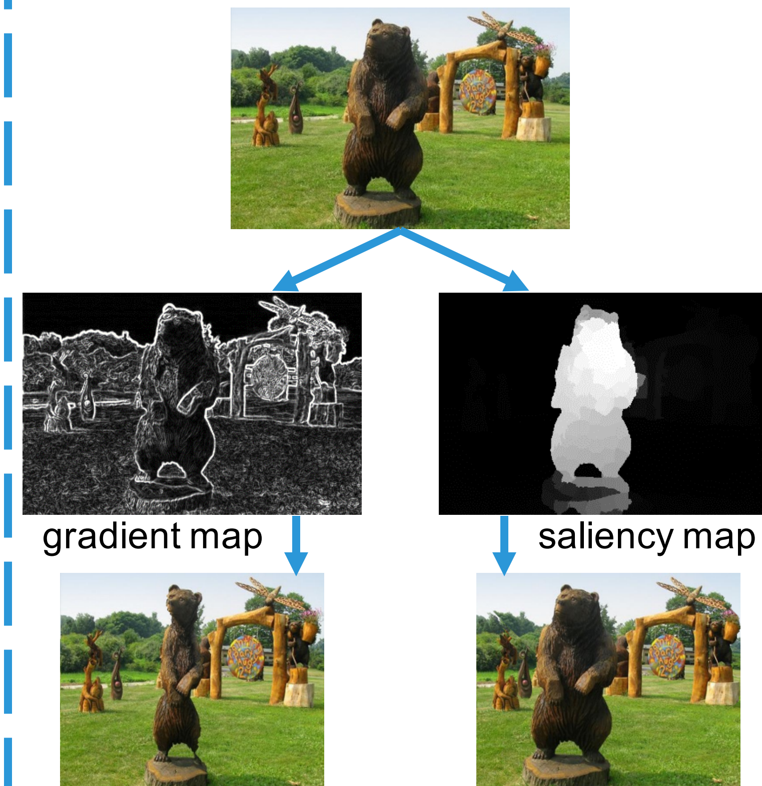
MAGUS
Multimedia AnalyzinG and UnderStanding

- Find visual attractive objects



- Multimedia applications
  - Image and video compression
  - Image editing and manipulating
  - Video summarization
  - Content-based image retrieval
  - ……

Content-aware image resizing



[1] S. Avidan & A. Shamir. Seam Carving for Content-Aware Image Resizing. TOG 2007

From RGB to RGB-D

MAGUS
Multimedia AnalyzinG and UnderStanding

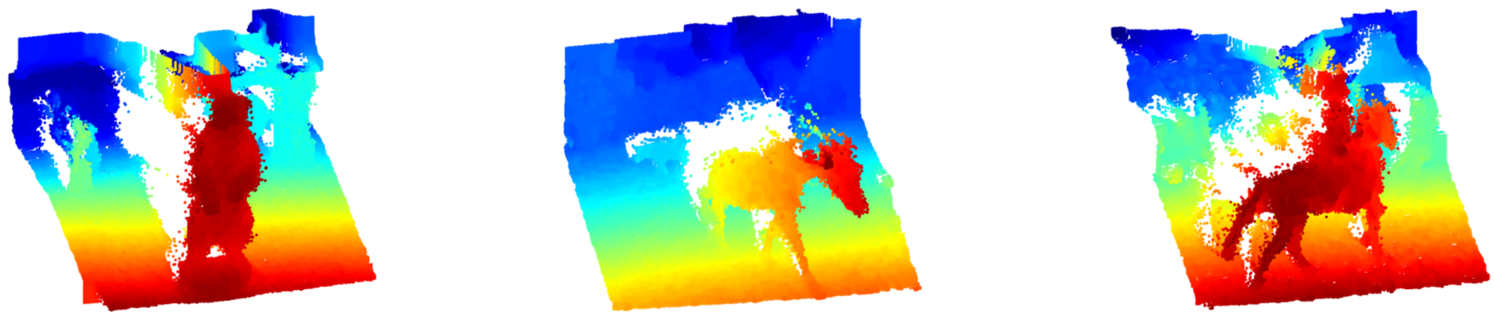
- Low-cost depth sensors

- Depth data provides important spatial information
  - complementary to color channels
  - potential to improve salient object detection

Challenge

MAGUS
Multimedia AnalyzinG
and UnderStanding

- How to manipulate depth data?
  - always noisy



  - may conflict with color cue

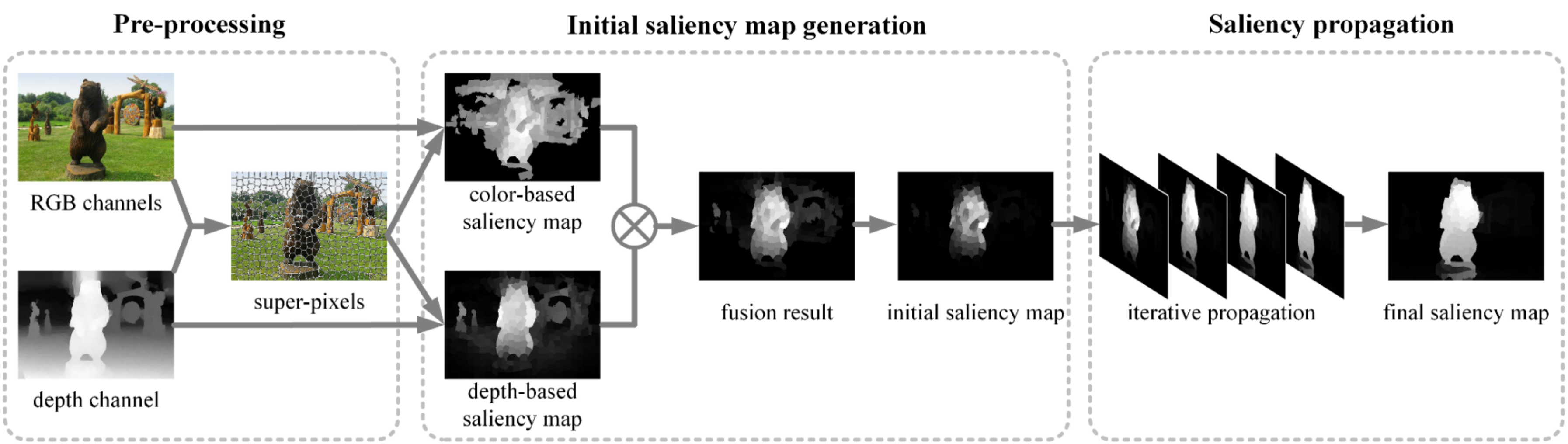
Color saliency: characters on the wall

Depth saliency: bonsai

Overview of our method

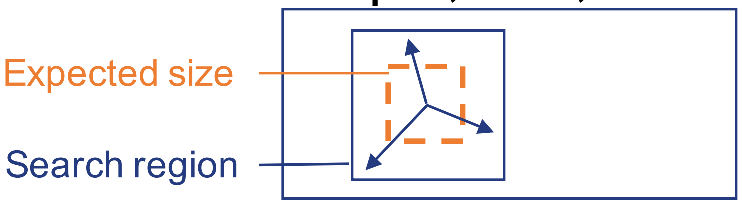
MAGUS
Multimedia AnalyzinG and UnderStanding

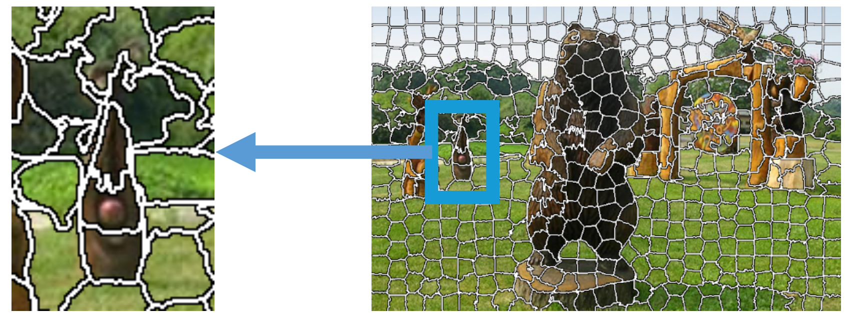
- Evolution strategy for RGB-D saliency
  1. initiate by different cues separately
  2. fuse to roughly locate the salient regions
  3. propagate to capture the whole salient object



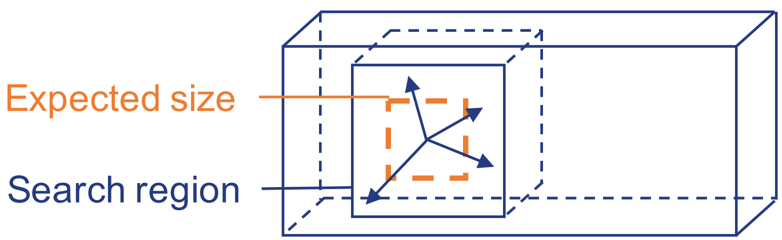
Super-pixel segmentation

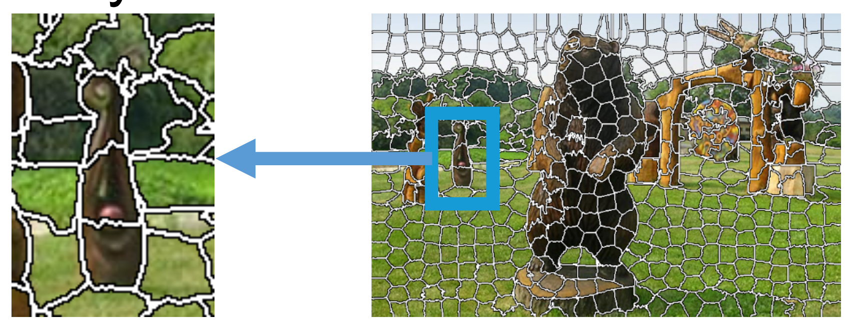
- SLIC - simple linear iterative clustering [1]
  - simple, fast, adhering to boundaries





- Extend to RGB-D
  - bring depth in spatial proximity term



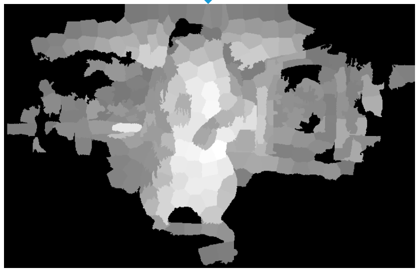


[1] R. Achanta et al. SLIC Superpixels Compared to State-of-the-Art Superpixel Methods. TPAMI 2012

Color-based saliency

MAGUS
Multimedia AnalyzinG and UnderStanding

- Weighted global color contrast
  - rarer colors tend to be salient
  - weighted by background probability and spatial distance



$$S_i^c = \sum_{j=1}^{N} \omega_j^b \cdot \omega_{i,j}^s \cdot \widetilde{D}_{i,j}^c$$

background weight
background probability of super-pixel j computed by boundary connectivity

spatial weight
Euclidean distance between centers of super-pixel i and j

color contrast
Euclidean distance between mean color in Lab color space

Depth-based saliency

MAGUS
Multimedia AnalyzinG
and UnderStanding

- Local depth contrast
  - a salient object tends to outstand from its surroundings in depth space

center of super-pixel i

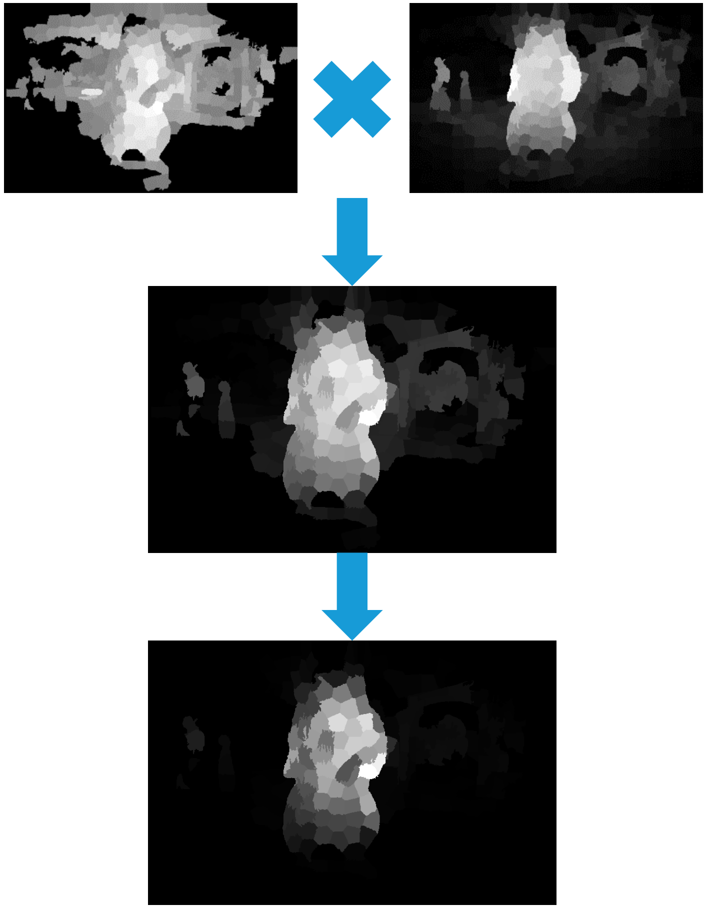
$$S_i^d = \sum_{\theta} \widetilde{D}^d\left(\hat{p}_i, P_i(\theta)\right)$$

set of pixels on the scanning radius emitting from $\hat{p}_i$ with angle θ

Manhattan distance between the depth of $\hat{p}_i$ and the minimum depth of $P_i(\theta)$

Fusion and refinement

- Element-wise product
  - common salient regions
- Depth-biased weighting
  - regions closer to us
- Emphasize precision

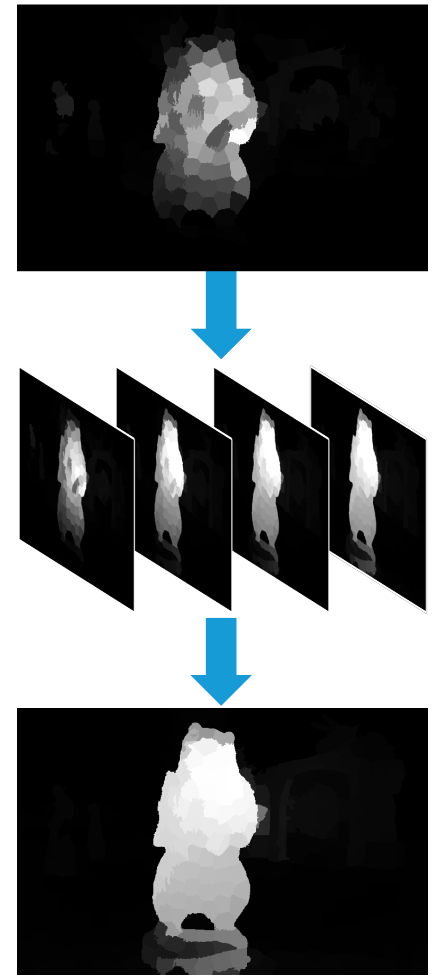


Saliency propagation

- Cellular Automata
  - cell: super-pixel, state: saliency value
  - propagate based on current state, neighbors' states, and similarity
  - Salient regions share similar features

$$S_i^* = \underbrace{\alpha_i S_i}_{\text{state coherence term}} + \overbrace{(1-\alpha_i)\sum_{j=1}^{N} \omega_{i,j}^{F} S_j}^{\text{neighbor coherence term}}$$

weigh the influences of adjacent super-pixels, in which the impact factor is based on both color cue and depth cue

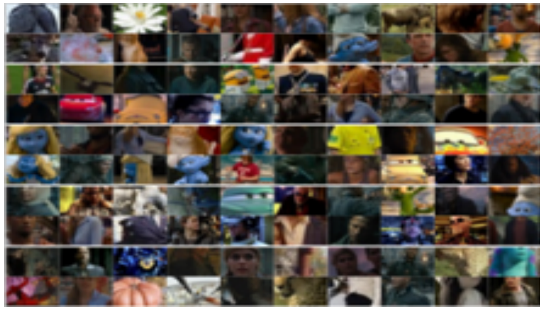


Datasets

MAGUS
Multimedia AnalyzinG and UnderStanding

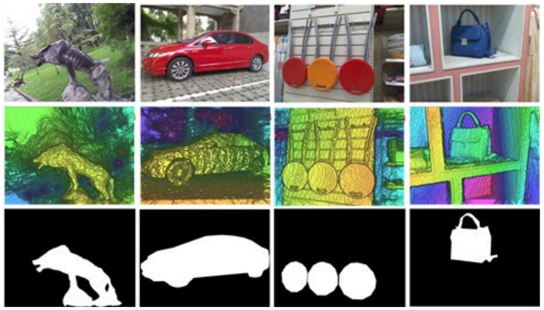
NJU2000 [1]

- 2000 stereo images
  - disparity recovered by optical flow
- 3D movies / Fuji W3



RGBD1000 [2]

- 1000 RGB-D images
- Microsoft Kinect

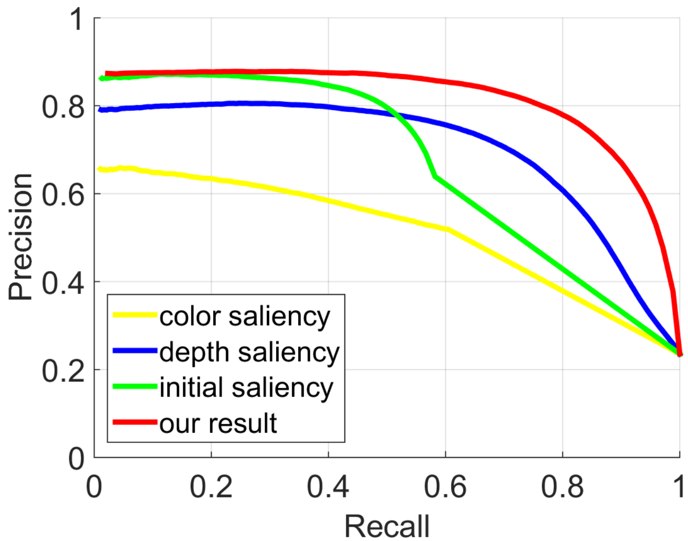


[1] R. Ju et al. Depth-Aware Salient Object Detection Using Anisotropic Center-Surround Difference. SPIC 2015

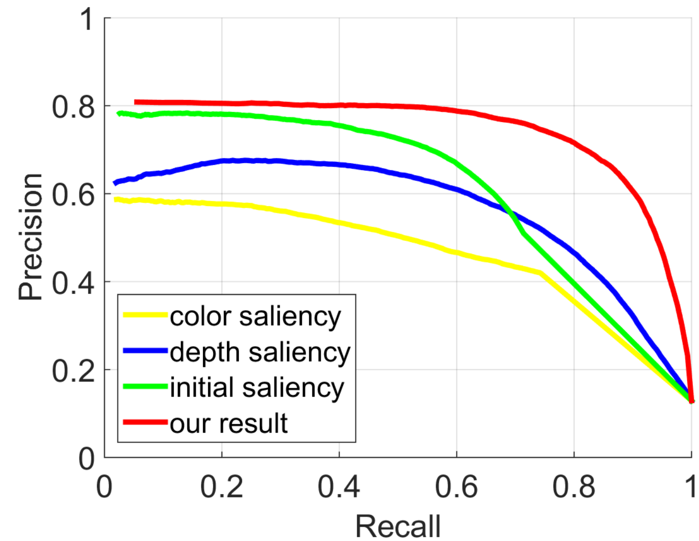
[2] H. Peng et al. RGBD Salient Object Detection: A Benchmark and Algorithms. ECCV 2014

Component evaluation

Multimedia AnalyzinG and UnderStanding



NJU2000



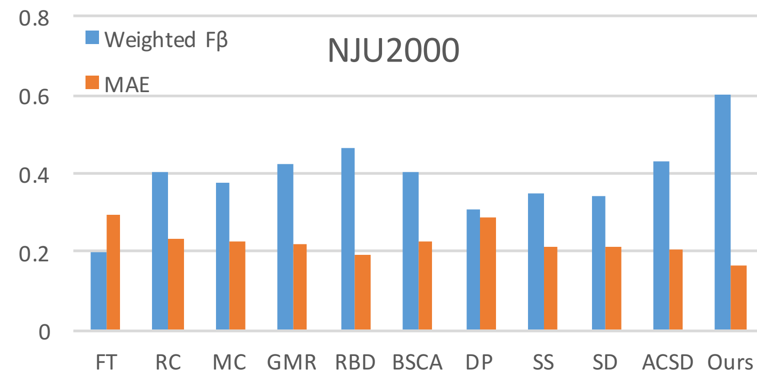
RGBD1000

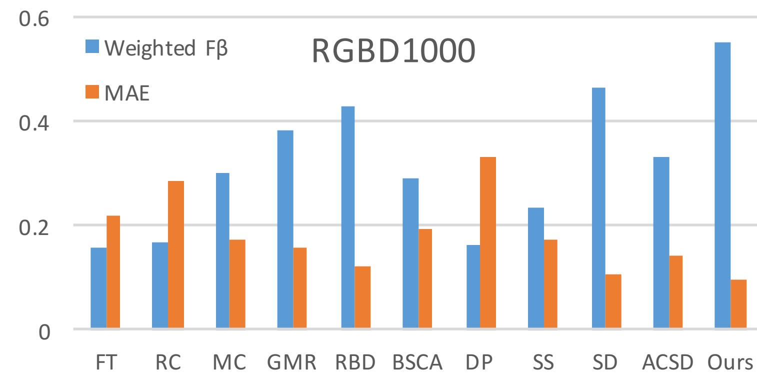
- Fusion: improve precision
- Propagation: improve recall

Performance evaluation

MAGUS
Multimedia AnalyzinG and UnderStanding

		NJU2000		RGBD1000	
		F_β^ω	MAE	F_β^ω	MAE
2D methods	FT	0.2009	0.2973	0.1583	0.2175
	RC	0.4025	0.2306	0.1689	0.2856
	MC	0.3749	0.2278	0.3018	0.1735
	GMR	0.4265	0.2174	0.3838	0.1593
	RBD	0.4678	0.1939	0.4300	0.1222
	BSCA	0.4040	0.2270	0.2886	0.1961
RGB-D methods	DP	0.3062	0.2896	0.1654	0.3305
	SS	0.3507	0.2102	0.2323	0.1750
	SD	0.3430	0.2144	0.4647	0.1091
	ACSD	0.4318	0.2031	0.3310	0.1452
	Ours	**0.6009**	**0.1634**	**0.5487**	**0.0974**

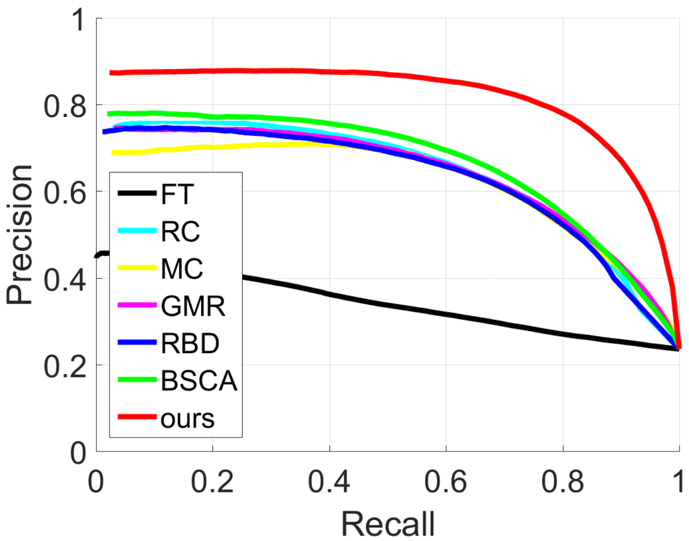




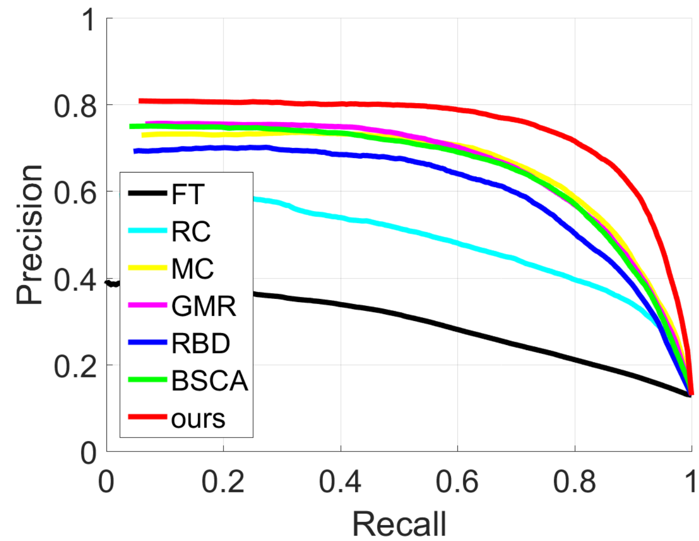
- State-of-the-art performance
- RGB-D methods aren't always better than 2D methods
  - How to manipulate depth data is crucial

Comparison with 2D methods

Multimedia AnalyzinG and UnderStanding



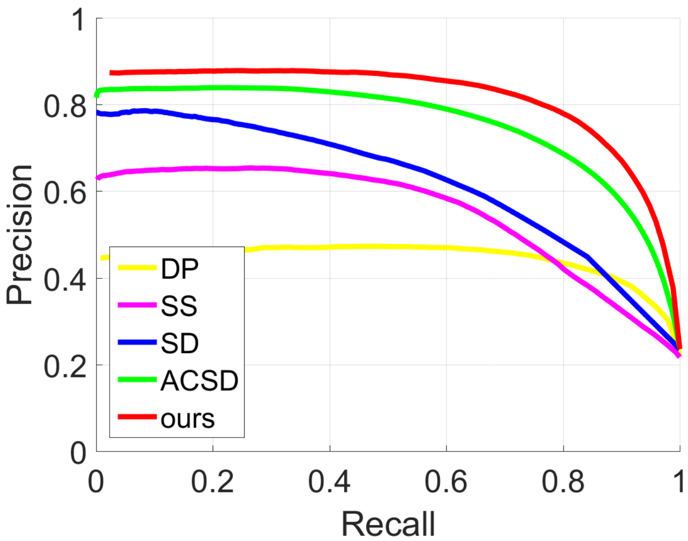
NJU2000



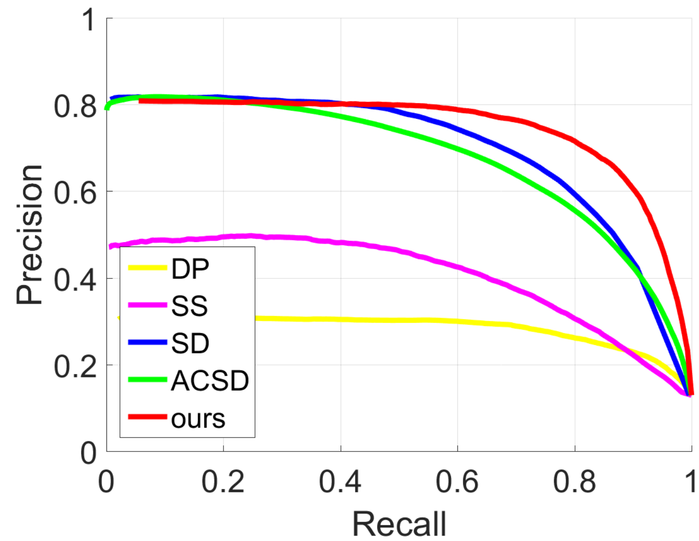
RGBD1000

- Depth cue takes positive effect in our method

Comparison with RGB-D methods



NJU2000



RGBD1000

- State-of-the-art performance
  - better integration of color cue and depth cue

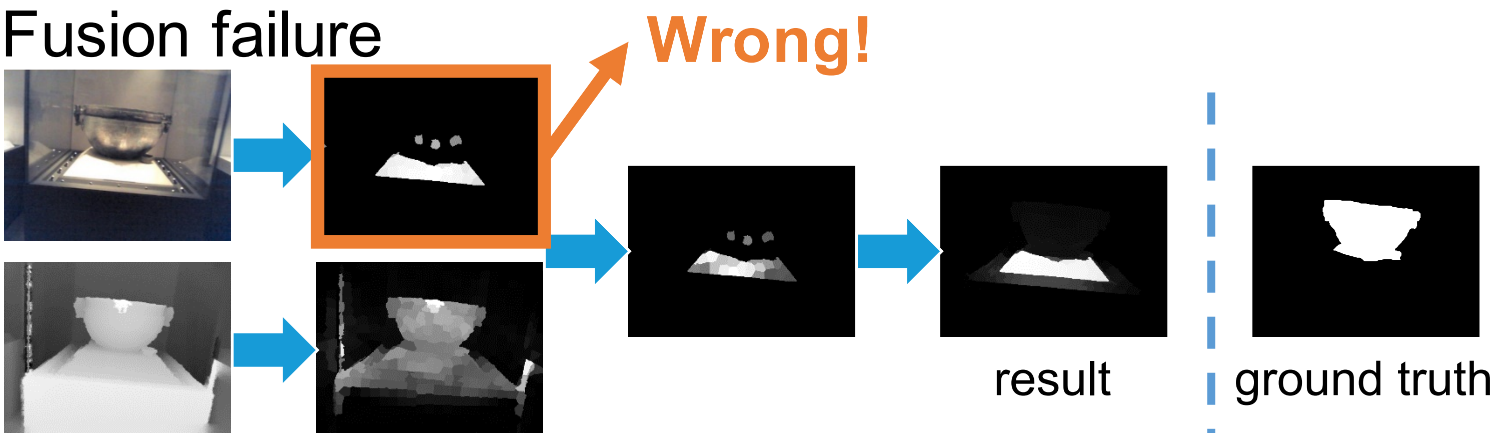
Comparison with other methods

MAGUS
Multimedia AnalyzinG and UnderStanding

Failure examples

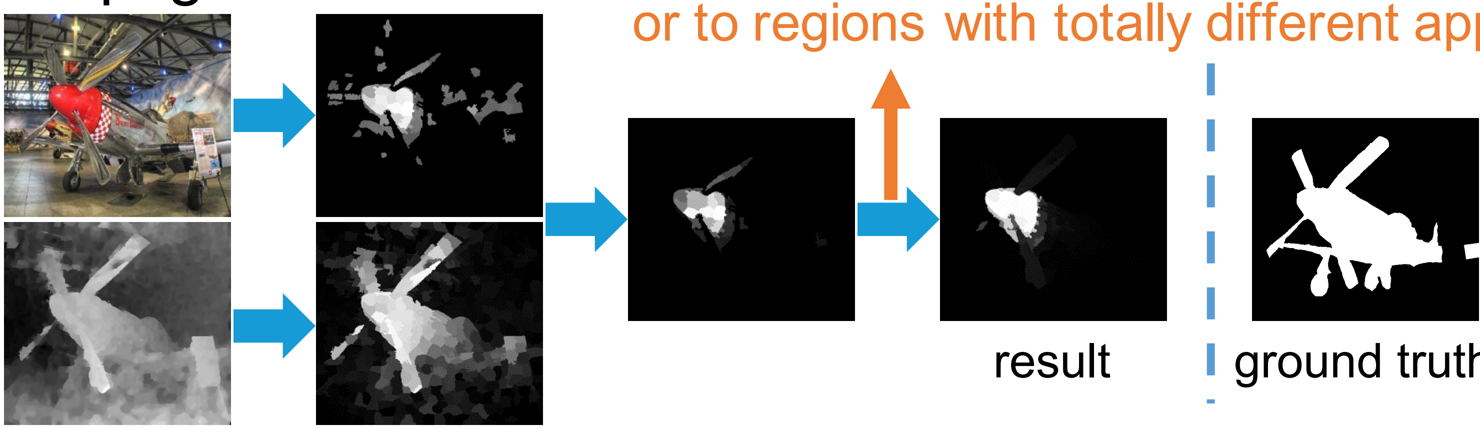
MAGUS
Multimedia AnalyzinG
and UnderStanding

- Fusion failure



- Propagation failure

cannot propagate to the whole depth space
or to regions with totally different appearance



Conclusion

MAGUS
Multimedia AnalyzinG and UnderStanding

- RGB-D saliency by evolution strategy
  1. initiate by different cues separately
  2. fuse to roughly locate the salient regions
  3. propagate to capture the whole salient object

- Integration of color cue and depth cue
- State-of-the-art performance
- Each component could be improved individually

Thank You

guojf@smail.nju.edu.cn