LOOSECUT: INTERACTIVE IMAGE SEGMENTATION WITH LOOSELY BOUNDED BOXES

Hongkai Yu[1], Youjie Zhou[1], Hui Qian[2], Min Xian[3], and Song Wang[1]
[1] University of South Carolina [2] Zhejiang University [3] Utah State University

ICIP 2017
IEEE International Conference on Image Processing
September 17-20, 2017, Beijing, China

Interactive image segmentation

- Interactive image segmentation has wide applications in computer vision community.
- However, the existing algorithms for interactive image segmentation prefer a bounding box that tightly encloses the object.
- Tight bounding box increases the annotation burden, and prevents these algorithms from utilizing automatically detected bounding boxes.

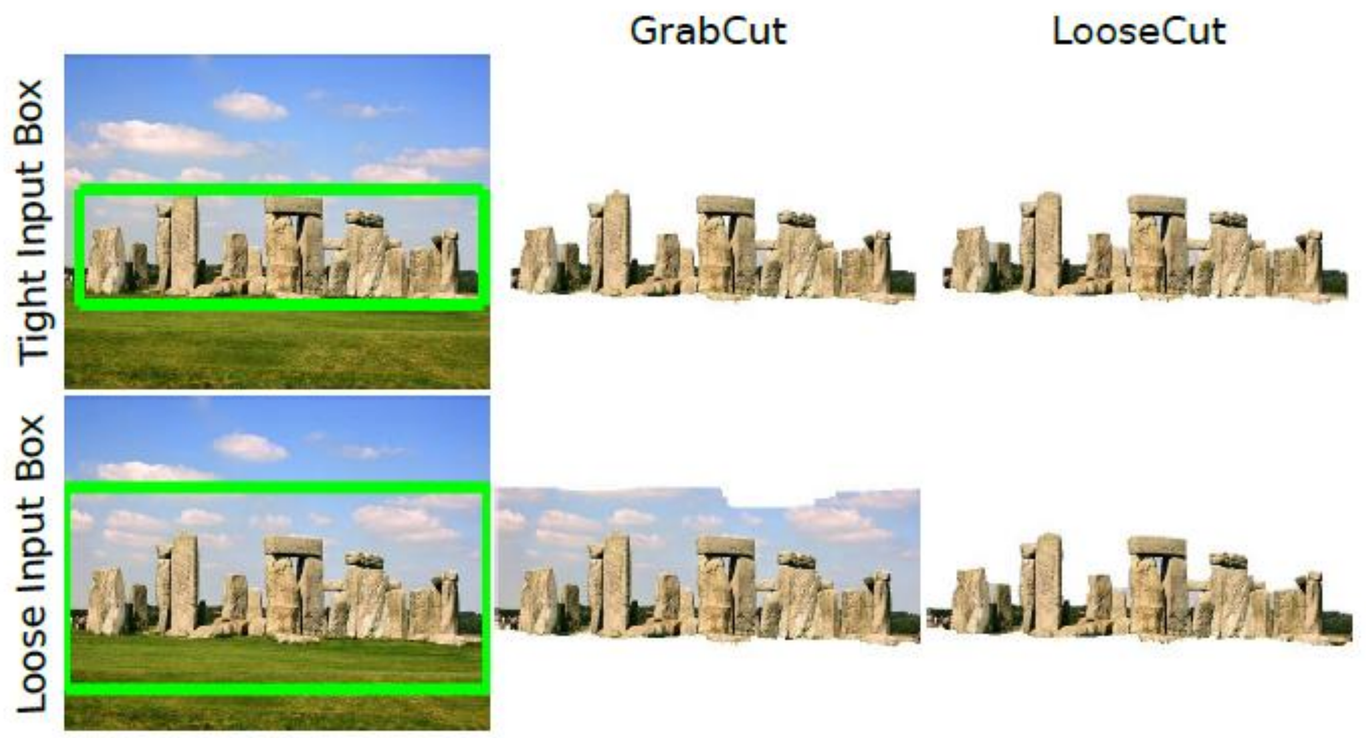


Fig. 1. Sample results from GrabCut and the proposed LooseCut with tightly and loosely bounded boxes.

Contributions

- In this paper, we develop a new LooseCut algorithm that can handle cases where the bounding box only loosely covers the object.
- We propose a new Markov Random Fields (MRF) model for segmentation with loosely bounded boxes, including an additional energy term to encourage consistent labeling of similar-appearance pixels and a global similarity constraint to better distinguish the foreground and background.

MRF Energy function

$$E(X,\theta) = E_{GC}(X,\theta) + \beta E_{LC}(X)$$

X **: binary labels**

θ **: appearance f/b GMM models** $\theta = (M_f, M_b)$

$E_{GC}(X,\theta)$**: GrabCut Energy Term**

$E_{LC}(\mathbf{X})$ **: Label Consistency Energy Term**

$$E_{LC}(\mathbf{X}) = \sum_k \sum_{i \in C_k} \phi(x_i \neq x_{C_k})$$

$\phi(\cdot)$ is an indicator function

Idea: pixels in same cluster are encouraged to be set the same label.

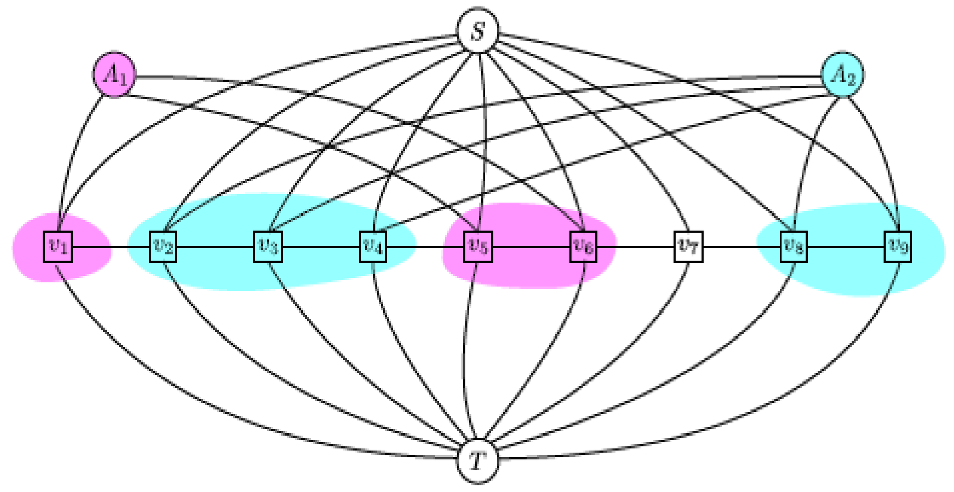


Fig. 2. Graph construction for the step of optimizing over X with a fixed θ. v_i's are the nodes for pixels and A_i's are the auxiliary nodes. S and T are the source and sink nodes. Same color nodes represent a cluster.

Global Similarity Constraint

Let M_f have K_f Gaussian components M_f^i with means μ_f^i, $i = 1, 2, \cdots, K_f$ and M_b have K_b Gaussian components M_b^j with means μ_b^j, $j = 1, 2, \cdots, K_b$.

$$Sim(M_f, M_b) = \sum_{i=1}^{K_f} S\left(M_f^i, M_b\right)$$

$$j(i) = \arg \min_{j \in \{1,\ldots,K_b\}} \left|\mu_f^i - \mu_b^j\right|$$

$$S\left(M_f^i, M_b\right) = \frac{1}{\left|\mu_f^i - \mu_b^{j(i)}\right|}$$

Idea: foreground and background similarity is encouraged to be small.

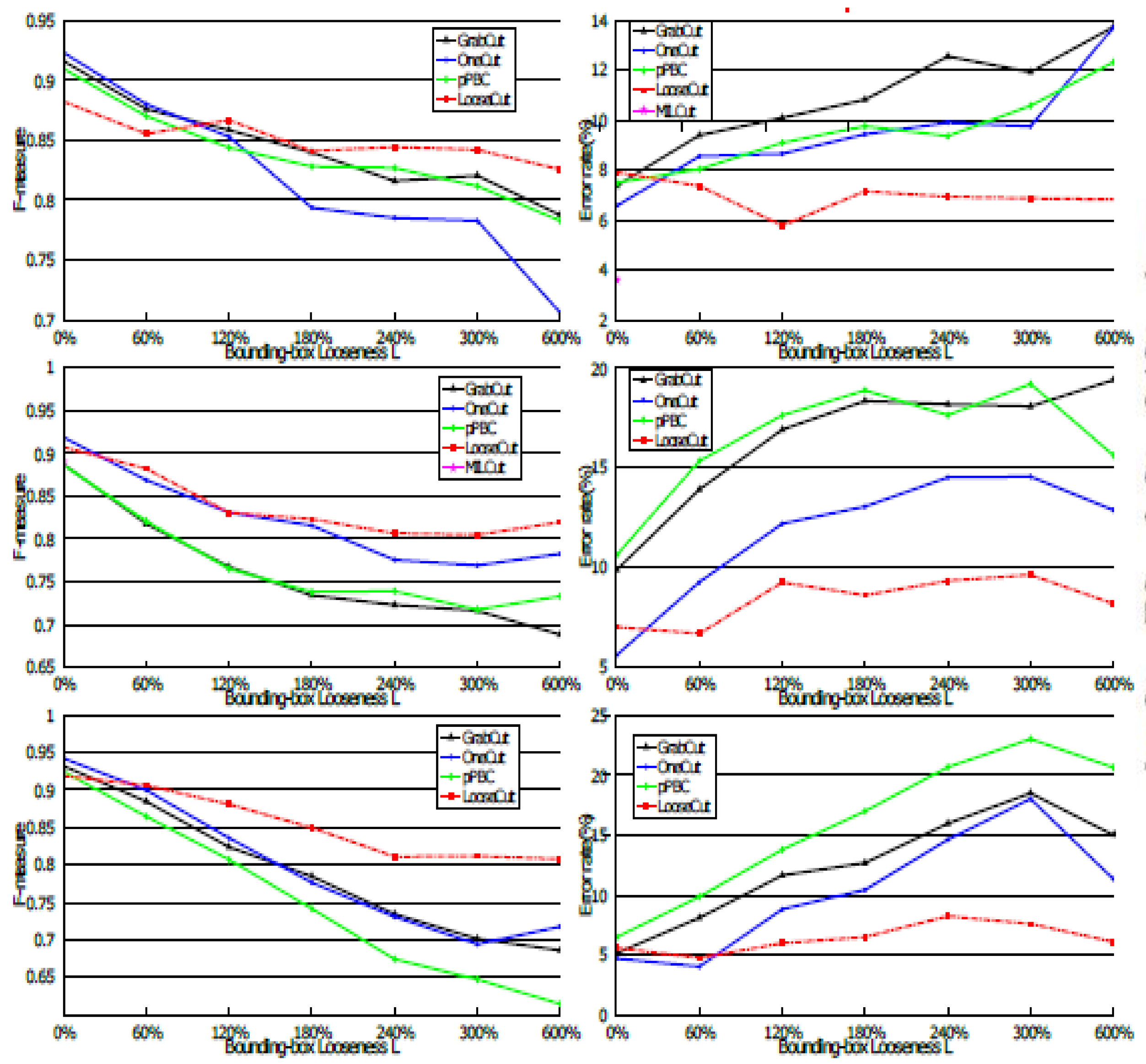
Algorithm

Algorithm 1 LooseCut

Input: Image I, bounding box B, # of clusters N
Output: Binary labeling X to pixels in I

1: Construct N superpixel-based clusters.
2: Create initial labeling X using box B.
3: **repeat**
4: Based on the current labeling X, estimate and update θ by enforcing $Sim(M_f, M_b) \leq \delta$.
5: Construct the graph using the updated θ with N auxiliary nodes as shown in Fig. 2.
6: Apply the max-flow algorithm to update labeling X by minimizing $E(X, \theta)$.
7: **until** Convergence or maximum iterations reached

Experimental Results on datasets of GrabCut, Weizmann, iCoseg



Sample Results

Code will be released soon at:
https://cvl.cse.sc.edu/research.html