Multimedia Systems and Applications

Precise Object Cutout From Images

James Wang

Based on ACM Multimedia 2008 paper:

Liu et al., Precise Object Cutout From Images, in Proceedings of ACM Multimedia 2008, pp. 623 – 626, October 26–31, 2008

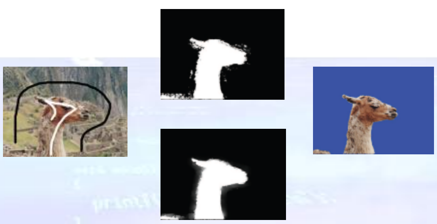
Problem and Existing Approaches

- **Problem:**
  - Interactive segmentation, whose task is to extract objects from an image using as little interaction as possible, is of great importance in many applications such as photo and movie editing.
- **Approaches:**
  - Boundary-based:
    - requires the user to draw along the object boundary and then adjusts the curve to snap to the real boundary piecewise.
  - Region-based:
    - requires the user to specify two small sets of pixels belonging to the foreground and the background.
    - It is more convenient.

2

Approach in this paper

- **It is a region-based approach:**
  - Several strokes (seeds) are provided by the user to indicate the foreground and the background
  - the algorithm does the segmentation automatically.



3

Methods

- **Two Gaussian mixture models (GMMs) for modeling the foreground and background are trained using the initial seeds.**
- **Two probability quantities are proposed to measure the initial likelihoods of each pixel belonging to the foreground and the background based on the GMMs.**
- **The segmentation is a problem of labeling the non-seed pixels as 1 (the foreground) and 0 (the background).**
  - Thus the labels can be regarded as the probabilities of the pixels belonging to the foreground.

4

Methods (cont.)

- **The inconsistency between the label and the quantity of the same pixel constructs the data cost in an optimization function, in which the boundary discontinuity and coherent region information are also incorporated.**
- **To solve the optimization problem, a soft segmentation that regards the label of each pixel as its probability of belonging to the foreground in the continuous domain.**
  - a closed form global optimal solution can be achieved.
- **With the global optimal label configuration in the continuous domain, it is straight-forward to use a threshold (0.5 in this paper) to obtain the final binary segmentation result.**

5

GMMs

- **Two GMMs $G_o(c)$ and $G_b(c)$, each with K components (K = 10 in this paper) are used to model the color distributions of the foreground (object) and the background, where c denotes the vector consisting of the R, G and B components of a pixel.**
  - The parameters in the two GMMs are estimated from the seeds provided by the user.
  - Here $G_o(c)$ ($G_b(c)$) can be considered as the likelihood of c belonging to the foreground (background).
  - Then, for each pixel i with its color vector c_i, we define two probability quantities as

$$p_{oi} = \frac{G_o(c_i)}{G_o(c_i) + G_b(c_i)}, \quad p_{bi} = \frac{G_b(c_i)}{G_o(c_i) + G_b(c_i)}. \qquad (1)$$

  - p_{oi} (p_{bi}) the initial probability of pixel i belonging to the foreground (background).
  - From the figures in slide 3, It is not difficult to see that p_{oi} reflects the main object region very well. However, in many small regions close to the boundary of the object, p_{oi} gives wrong results.

6

Data Cost Function

- **Let the pixel labels of an image be represented by $L = [l_1, l_2, ..., l_n]^T$, where n is the number of pixels of the image and l_i is the label of pixel i.**
- **A data cost function $D(l_i, p_{oi}, p_{bi})$ is used to measure the inconsistency between the assigned label l_i and the initial probabilities, where**

$$D(l_i = 1, p_{oi}, p_{bi}) = (1 - p_{oi})^2, \qquad (2)$$
$$D(l_i = 0, p_{oi}, p_{bi}) = (1 - p_{bi})^2. \qquad (3)$$

- **Since $p_{oi} = 1-p_{bi}$, the above two equations can be combined as the following one:**

$$D(l_i, p_{oi}, p_{bi}) = (l_i - p_{oi})^2. \qquad (4)$$

7

Objective Function

- **Anatural image usually has the property of pairwise smoothness.**
  - a smoothness penalty term $w_{ij} (l_i - l_j)^2$ to impose this constraint.
- **The objective function is defined as**

$$f(L) = \sum_{i=1}^{n} (l_i - p_{oi})^2 + \sum_{i=1}^{n} \sum_{j \in \mathcal{N}(i)} w_{ij}(l_i - l_j)^2, \qquad (5)$$

where N(i) is a neighborhood of pixel i, and

$$w_{ij} = a \cdot k(i, j) \cdot \exp(-\frac{(p_{oi} - p_{oj})^2}{\sigma_i^2 + \sigma_j^2}), \qquad (6)$$

8

Optimization Process

- **With the objective function f(L), we relax the binary labeling to continuous labeling, ranging from 0 to 1, then obtain a closed form global optimal solution to this continuous labeling problem.**
  - Construct an undirected weighted graph G =(V, E), where V is the set of vertices denoting all the image pixels and E is the set of weighted edges.
  - The elements of the adjacent matrix $W = [W_{ij}]_{n\times n}$ of G are obtained by

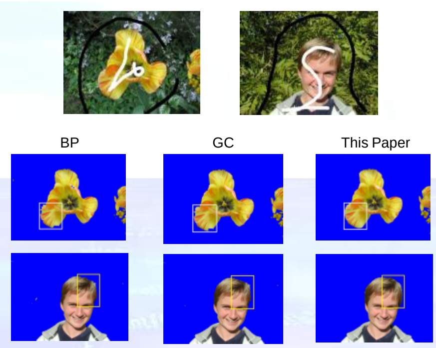
$$W_{ij} = \begin{cases} w_{ij}, & \text{if } i \neq j, j \in \mathcal{N}(i) \\ 0, & \text{if } i \neq j, j \notin \mathcal{N}(i) \\ c, & \text{if } i = j, \end{cases} \qquad (7)$$

  - With the objective function f(L) and the corresponding graph G, a closed form solution can be achieved to the following problem:

$$\min f(L) = \min \left(\sum_{i=1}^{n} (l_i - p_{oi})^2 + \sum_{i,j=1}^{n} W_{ij}(l_i - l_j)^2 \right). \qquad (8)$$

9

Experiments



10

More Experiments

11

Conclusions

- **In this paper, a novel interactive segmentation algorithm has been proposed.**
- **The experimental results demonstrate the excellent performance of this algorithm.**
- **This approach can also be generalized to 3D images easily.**
- **This framework can be applied to other applications such as image denoising and stereo correspondence, where continuous labels are preferred.**

12