Interactive Image Composition through Draggable Objects

Yuichiro Yamaguchi* Takuya Saito* Yosuke Bando*,† Bing-Yu Chen‡ Tomoyuki Nishita*
*The University of Tokyo †TOSHIBA Corporation ‡National Taiwan University

1 Introduction

In traditional image composition methods for cutting out a source object from a source image and pasting it onto a target image, users have to segment a foreground object in a target image when they want to partially hide a source object behind it. While recent image editing tools greatly facilitate segmentation operations, it can be tedious to segment each object if users try to place a source object in various positions in a target image before obtaining a satisfying composition. We propose a method which allows users to drag a source object and slip it behind a target object as shown in Fig. 1, so that users can move a source object around without manually segmenting each part of a target image.

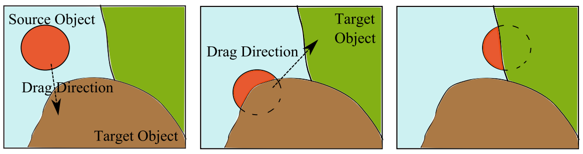


Figure 1: *Image composition by dragging objects.*

2 Proposed Image Composition Method

Our system performs local segmentation of the region which is newly covered with the source object when it is dragged, which we call *newly covered area* (shown in yellow in Fig. 2). For segmentation we need local probability estimations of foreground and background colors, and we propose to obtain them by looking ahead into color distributions of the target image in the drag direction.

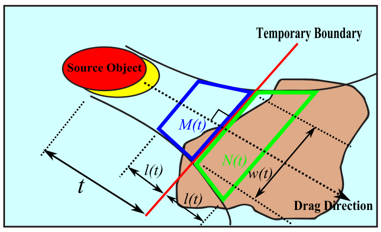


Figure 2: *Color distributions look-ahead.*

Step 1: We consider two regions $M(t)$ and $N(t)$, divided by the temporary boundary t, as containing candidate color distributions for background and foreground, respectively. The width $w(t)$ and length $l(t)$ of the regions are set to be comparable to the source object width W and length L as $w(t) = W\exp(s_w t)$ and $l(t) = L\exp(s_l t)$, but we introduced exponential scaling to enlarge the region sizes to account for the uncertainty of the drag direction look-ahead in farther regions (we set $s_w = 0.03$ and $s_l = 0.04$, but results seem largely insensitive to the choise of values up to 0.05).

Step 2: We compute distances $D(t)$ between the candidate foreground and background color distributions while varying t, i.e., while moving the temporary boundary along the drag direction, as:

$$D(t) = \frac{\lambda(t)}{|M(t)||N(t)|} \sum_{m\in M(t)} \sum_{n\in N(t)} |c_m - c_n|^2,$$

where c_m and c_n are the RGB color values of the pixels m and n, and $\lambda(t) = \exp(-s_\lambda t)$ is a weighting function to put more confidence on estimates for nearer regions ($s_\lambda \in [0.01, 0.07]$ performed equally well, and we used $s_\lambda = 0.03$ for the results shown below).

Step 3: We regard the two regions having the largest distance between them as providing the foreground and background color distributions; we determine the boundary as $t_{max} = \arg\max D(t)$.

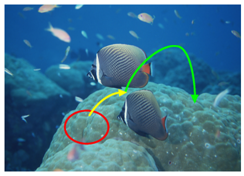
Step 4: We segment the newly covered area into foreground and background with these color distributions $M(t_{max})$ and $N(t_{max})$ by the graph-cut method [Boykov and Jolly 2001].

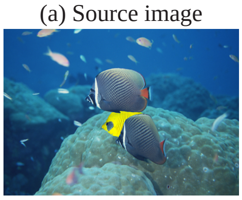
3 Results and Conclusion

We have presented a method to composite new objects into images behind existing objects using a dragging interface, which are automatically detected by color distribution look-ahead.

Fig. 3 shows an example of composing a yellow fish in a source image (a) into a target image (b) of another underwater scene. We put the source fish on the left side of the target image as indicated by the red circle, and dragged it to the right (yellow arrow) to compose it behind the other fishes (c). We subsequently dragged it as indicated by the green arrow to compose it behind the rock (d). Fig. 4 shows our system can also compose a source object (balloon) behind moving target objects (swaying flowers) in a video by applying the history of user interaction for one frame to the other frames.


(a) Source image


(b) Target image


(c) Behind the fishes

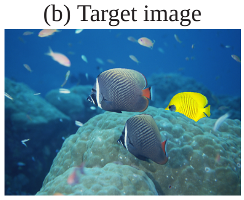

(d) Behind the rock

Figure 3: *Composition results of our method.*

Figure 4: *Three frames from a video composition result.*

References

BOYKOV, Y., AND JOLLY, M. P. 2001. Interactive graph cuts for optimal boundary & region segmentation of objects in N-D images. In *ICCV*.