# Outlining Objects for Interactive Segmentation on Touch Devices

Matthieu Pizenberg
Axel Carlier
Emmanuel Faure
Vincent Charvillat

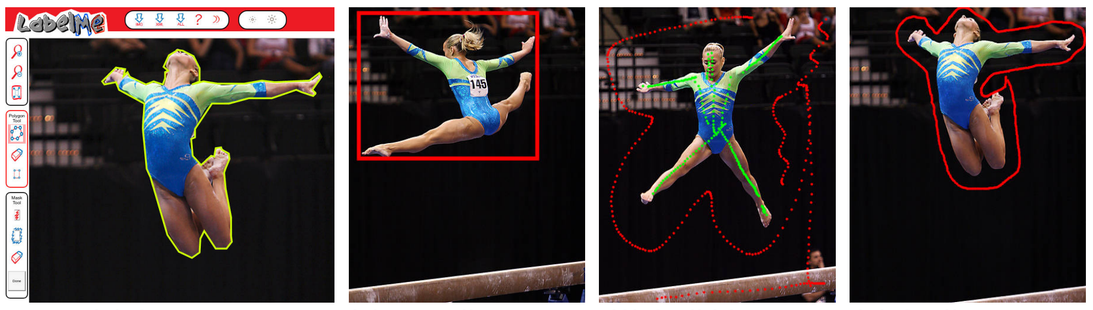
ACM multimedia 25
October 23 - 27 2017
Mountain View - CA USA

github.com/mpizenberg/otis

user study demo (online)
mm17-otis.pizenberg.fr

## Introduction

There is a need for training datasets to support recent advances in machine learning for segmentation [1,2,7]
-> Segmentation datasets are tedious to create with traditional interactions
-> We propose to use outlining on touch devices




(A) Contour (B) Bounding Box (C) Scribbles (D) Outline

Figure 1: Different approaches for the task of interactive segmentation. Images are from the iCoseg dataset [3].

## Experiments

* Twenty users (10 male / 10 female)
* Three tested interactions (bounding box, outline, scribbles)
* 11 images per interaction (from iCoseg dataset [3])
* Regular 8" android touch tablet
  -> source code available at github.com/mpizenberg/otis
  -> online demo at mm17-otis.pizenberg.fr

| Method | Bounding box | Outline | Scribble |
|---|---|---|---|
| Ease | 2.1 ± 0.62 | 2.65 ± 0.74 | 2.1 ± 0.61 |
| Time | 2.35 ± 0.69 | 2.5 ± 0.67 | 2.6 ± 0.70 |
| Rank | 1.95 ± 0.43 | 1.90 ± 0.32 | 2.15 ± 0.37 |

Table 1: Results of the questionnaire with a 95% confidence intervals. "Ease" and "Time" are measured on a scale from 1 (better) to 7.

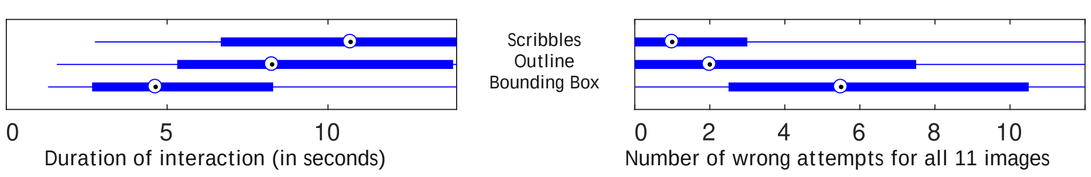



Figure 3: Measures of duration and errors for each interaction.

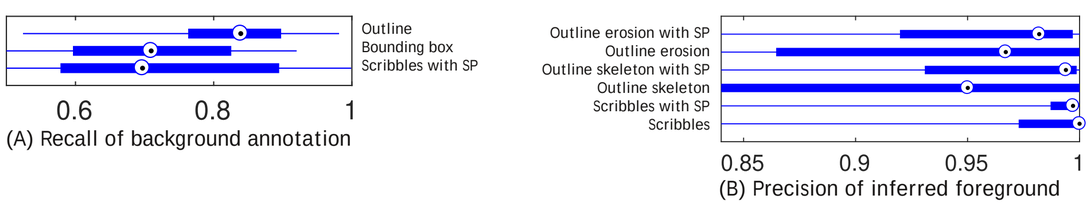



Figure 4: Evaluation of the quality of the participants annotations.

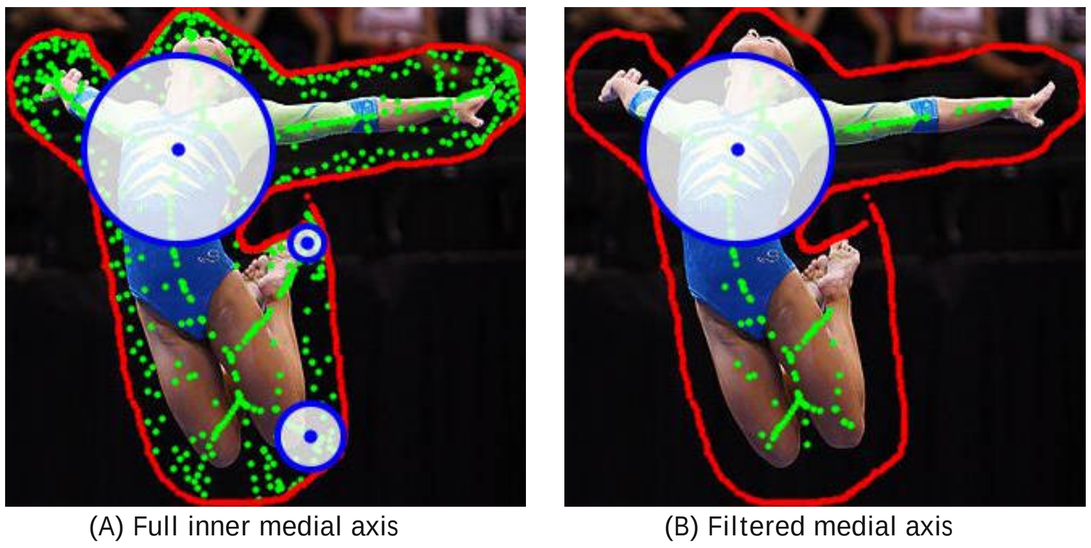
Best segmentation results (Table 2) obtained with:
Outline interaction + inferred foreground + superpixels enhancement

| Method | Scrib. | B. Box | Outl. | Outl. + er. | Outl. + BMA |
|---|---|---|---|---|---|
| Mean Jaccard | 0.79 | 0.82 | 0.86 | 0.88 | 0.89 |

Table 2: Mean Jaccard index obtained on all images for all users, for each interaction.

## Outlining Interaction

GrabCut [4] alteration with initial foreground inferred using:
* morphologic erosion
* shape skeleton [5] (Fig. 2)
-> enhanced by superpixels [6]



(A) Full inner medial axis (B) Filtered medial axis

(C) Skeleton (D) Refined through superpixels

Figure 2: Inferring an initial foreground using the Blum medial axis [5], refined thanks to mean-shift superpixels [6].

## Conclusion

Outlining is a simple and natural interaction on touch devices.
Simplicity + swiftness + accuracy -> good candidate for datasets creation.

Figure 5: Some segmentation results using bounding box and outline interactions.

## References

[1] J. Long et al. CVPR 2015. Fully convolutional networks for semantic segmentation.
[2] Mark Everingham et al. IJCV 2010. The pascal visual object classes (voc) challenge.
[3] D. Batra et al. CVPR 2010. iCoseg: Interactive co-segmentation with intelligent scribble guidance.
[4] C. Rother et al. ACM SIG 2004. "GrabCut": Interactive Foreground Extraction Using Iterated Graph Cuts.
[5] Harry Blum et al. Pattern recognition (1978). Shape description using weighted symmetric axis features.
[6] Dorin Comaniciu et al. PAMI 2002. Mean shift: A robust approach toward feature space analysis.
[7] Bryan C Russell et al. IJCV 2008. LabelMe: a database and web-based tool for image annotation.

INP - ENSEEIHT, Toulouse, France
University of Toulouse, France
CNRS - IRIT, Toulouse, France