GrabCut
Interactive Foreground Extraction using Iterated Graph Cuts

Carsten Rother
Vladimir Kolmogorov
Andrew Blake

Microsoft Research Cambridge-UK
Photomontage
Problem

GrabCut – Interactive Foreground Extraction
What GrabCut does

User Input

Magic Wand (198?)

Intelligent Scissors
Mortensen and Barrett (1995)

Result

Regions

Boundary

Regions & Boundary

GrabCut – Interactive Foreground Extraction
Framework

- **Input:** Image $x \in \{R, G, B\}^n$
- **Output:** Segmentation $S \in \{0, 1\}^n$
- **Parameters:** Colour $\Theta$, Coherence $\lambda$
- **Energy:** $E(\Theta, S, x, \lambda) = E_{Col} + E_{Coh}$
- **Optimization:** $\arg\min_{S, \Theta} E(S, \Theta, x, \lambda)$
Graph Cuts

Boykov and Jolly (2001)

**Image**

**Cut:** separating source and sink; **Energy:** collection of edges

**Min Cut:** Global minimal energy in polynomial time

**Graph Cuts**

*Boykov and Jolly (2001)*

**GrabCut – Interactive Foreground Extraction**
Iterated Graph Cut

User Initialisation

\[ \arg\min_{\Theta} E(S, \Theta, x, \lambda) \]

K-means for learning colour distributions

Graph cuts to infer the segmentation

\[ \arg\min_{S} E(S, \Theta, x, \lambda) \]
Iterated Graph Cuts

Guaranteed to converge

Result

Energy after each Iteration

GrabCut – Interactive Foreground Extraction
Colour Model

Gaussian Mixture Model (typically 5-8 components)

$$E_{Col}(\Theta, S, x) = \sum_n D(S_n, \Theta, x_n)$$
Coherence Model

An object is a coherent set of pixels:

\[ E_{coh}(S, x, \lambda) = \lambda \sum_{i,j \text{ adj.}} (S_i \neq S_j) \exp\left\{-\frac{1}{2\sigma^2}||x_i - x_j||^2\right\} \]

\[ \lambda = 0 \quad \lambda = 50 \quad \lambda = 1000 \]

Blake et al. (2004): Learn \( \Theta, \lambda \) jointly
Moderately straightforward examples

... GrabCut completes automatically
Difficult Examples

Camouflage & Low Contrast

Initial Rectangle

Fine structure

Initial Result

No telepathy

GrabCut – Interactive Foreground Extraction
Evaluation – Labelled Database

Available online: http://research.microsoft.com/vision/cambridge/segmentation/
Comparison

Boykov and Jolly (2001)

GrabCut

User Input

Result

Error Rate: 0.72%

Error Rate: 0.72%

Error Rate: 1.87%

Error Rate: 1.32%

Error Rate: 1.25%

GrabCut – Interactive Foreground Extraction
Summary

Magic Wand (198?)
Intelligent Scissors Mortensen and Barrett (1995)
Graph Cuts Boykov and Jolly (2001)
LazySnapping Li et al. (2004)
GrabCut Rother et al. (2004)

GrabCut – Interactive Foreground Extraction