

GrabCut

Interactive Foreground Extraction using Iterated Graph Cuts

Carsten Rother
Vladimir Kolmogorov
Andrew Blake



Microsoft Research Cambridge-UK

Photomontage



Problem



→
Fast &
Accurate ?



What GrabCut does

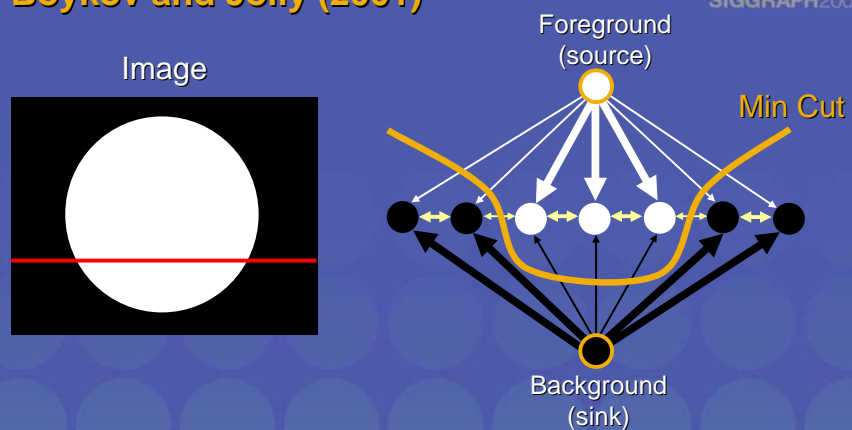
	Magic Wand (198?)	Intelligent Scissors Mortensen and Barrett (1995)	GrabCut
User Input			
Result			
	Regions	Boundary	Regions & Boundary

Framework

- **Input:** Image $x \in \{R, G, B\}^n$
- **Output:** Segmentation $S \in \{0, 1\}^n$
- **Parameters:** Colour Θ , Coherence λ
- **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)



Cut: separating source and sink; Energy: collection of edges
Min Cut: Global minimal energy in polynomial time

Iterated Graph Cut



User Initialisation

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

$$\arg \min_S E(S, \Theta, x, \lambda)$$

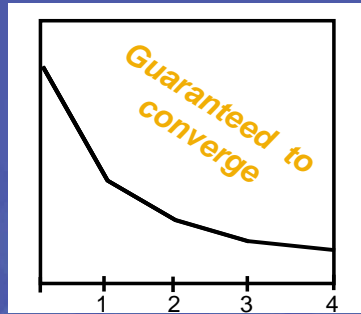
K-means for learning colour distributions

Graph cuts to infer the segmentation

Iterated Graph Cuts

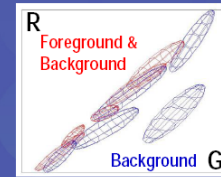


Result

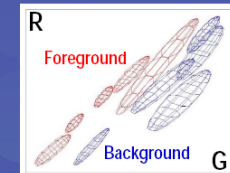


Energy after each Iteration

Colour Model



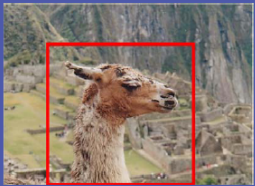
Iterated graph cut



Gaussian Mixture Model (typically 5-8 components)

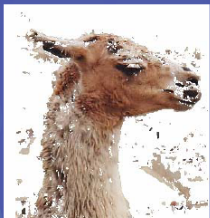
$$E_{Col}(\Theta, S, x) = \sum_n D(S_n, \Theta, x_n)$$

Coherence Model

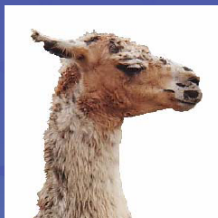


An object is a coherent set of pixels:

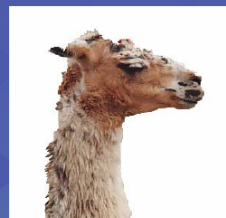
$$E_{coh}(S, x, \lambda) = \lambda \sum_{i,j \text{ adj.}} (S_i \neq S_j) \exp\left\{-\frac{1}{2\sigma^2} \|x_i - x_j\|^2\right\}$$



$\lambda = 0$



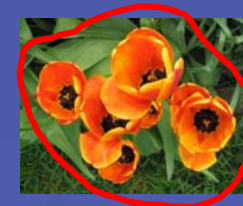
$\lambda = 50$



$\lambda = 1000$

Blake et al. (2004): Learn Θ, λ jointly

Moderately straightforward examples



... GrabCut completes automatically

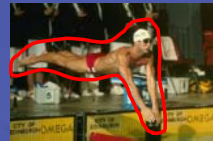
Difficult Examples

Camouflage & Low Contrast

Fine structure

No telepathy

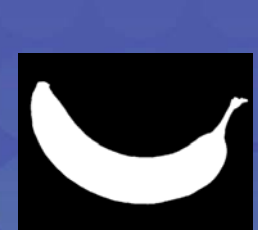
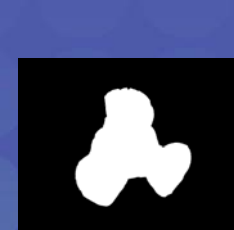
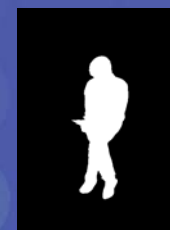
Initial Rectangle



Initial Result



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%

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)