Random Walks based Multi-Image Segmentation: Quasiconvexity Results and GPU-based Solutions

Maxwell D. Collins[1] Jia Xu[1] Leo Grady[2] Vikas Singh[1]
{mcollins,jiaxu}@cs.wisc.edu lgrady@heartflow.com vsingh@biostat.wisc.edu
[1]University of Wisconsin-Madison [2]HeartFlow

Cosegmentation

Cosegmentation is the task of segmenting a common salient foreground object from two or more images.

We approach cosegmentation as an optimization problem of the form

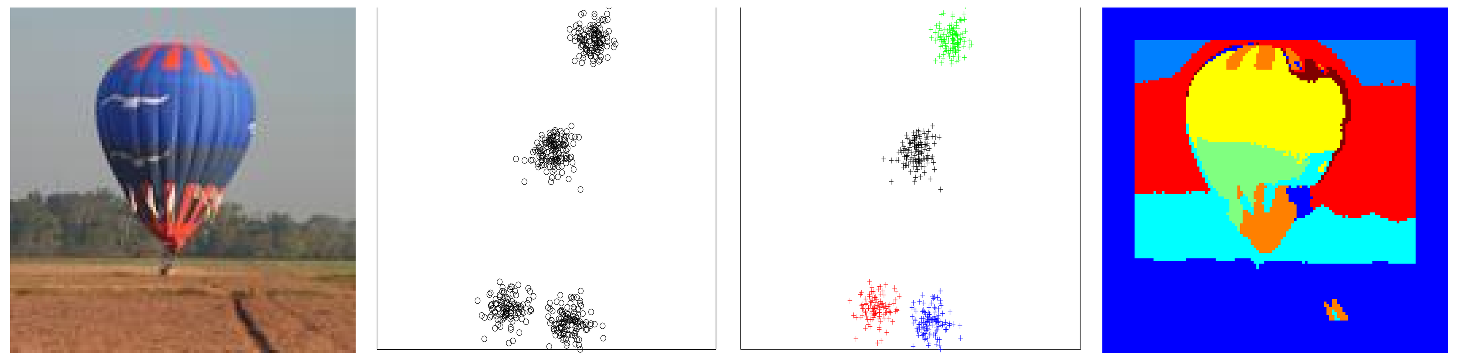
$$\min_{\mathbf{x}} \sum_{i \in \text{images}} E_{\text{segmentation}}(x_i) + E_{\text{cosegmentation}}(x_1, x_2, \ldots)$$

Random Walks based $E_{segmentation}$ provides fast and effective nonparametric cosegmentation without entropy assumptions.

Histogram Distance Penalties

A *histogram* gives a global description by binning/counting features.

Histograms bin similar pixels, and so provide a similarity measure in cosegmentation by comparing the histograms of subregions.



Use $E_{\text{cosegmentation}}$ based on a distance measure between histograms.

Optimization Model

$E_{\text{segmentation}}(x_i) = x_i^T L_i x_i$: Random walker energy with Laplacian matrix L_i.

Cosegmentation potential $\|h_i - \bar{h}\|_2^2$, the distance between foreground histogram h_i and common model histogram $\bar{h}$.

Given histogram matrices H_i and seeds s with values m_i, solve

$$\min_{x_i, h_i, \bar{h}} \sum_i x_i^T L_i x_i + \lambda \|h_i - \bar{h}\|_2^2$$
$$\text{s.t.} \quad x_i \in [0,1]^{n_i}, \quad x_i^{(s)} = m_i^{(s)}, \quad H_i x_i = h_i, \quad i = 1 \ldots m.$$

to output segmentation potentials x_i for each image i.

Box-QP

The model can be expressed as a QP with box constraints:

$$\min_{x_1, \ldots, x_m, \bar{h}} \begin{bmatrix} x_1 \\ \vdots \\ x_m \\ \bar{h} \end{bmatrix}^T \begin{bmatrix} L_1 + \lambda H_1^T H_1 & & & -\lambda H_1 \\ & \ddots & & \vdots \\ & & L_m + \lambda H_m^T H_m & -\lambda H_m \\ -\lambda H_1^T & \ldots & -\lambda H_m^T & \lambda m I \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_m \\ \bar{h} \end{bmatrix}$$
$$\text{s.t.} \quad l_i \le x_i \le u_i \quad x_i \text{ is of size } [0,1]^{n_i} \quad i = 1, \ldots, m$$

where (l_i, u_i) is $(1,1)$ for foreground seeds, $(0,0)$ for background seeds, and $(0,1)$ otherwise.

Solved via GPCG, alternating gradient projection/conjugate gradient.

After distributing the multiplication

$$(L_i + H_i^T H_i) x_i = L x_i + H_i^T (H_i x_i).$$

we can calculate gradient with sparse matrix computations on GPU.

Scale-Free Similarities

A metric distance between histogram vectors will be dependent on *scale* and only accurate segment foregrounds of the same size.
Cosegmentation can be robust to distances in scale by choosing an energy E such that

$$E(h, \bar{h}) = E(sh, \bar{h}) \quad \forall s \in \mathbb{R} \setminus \{0\}$$

This condition is met by a *normalized* histogram similarity, such as

$$-\frac{\langle h, \bar{h} \rangle}{\|h\|_2} = -\|\bar{h}\|_2 \cos(\angle h \bar{h})$$

Quasiconvexity

A scale-free distance is not *convex* and used in E_{coseg} cannot efficiently be solved through traditional methods. It is however *quasiconvex* in h.

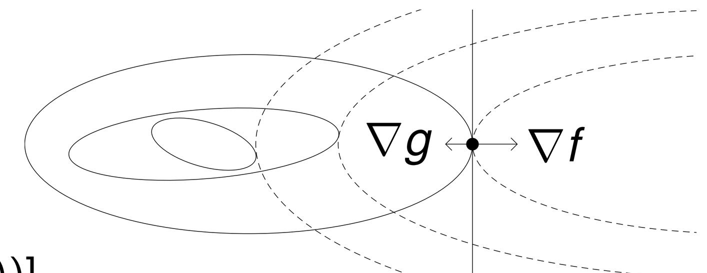
Quasiconvexity is the property of a function f such that its *sublevel sets* are convex:

$$f((1-\lambda)x + \lambda x') \le \max(f(x), f(x')) \quad \forall x, x' \in X, \lambda \in [0,1]$$

Optimization Scheme

Given a pair of quasiconvex functions f, g, the solution to $\min_x f(x) + g(x)$ is a solution to

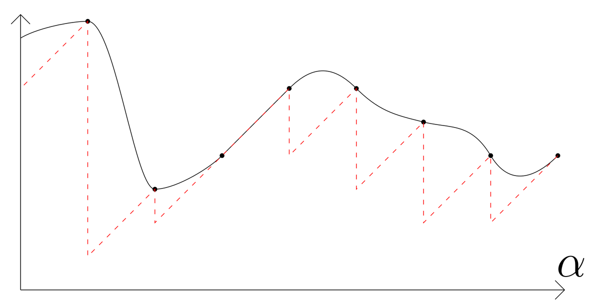
$$\min_x \; f(x) \quad \text{s.t} \; g(x) \le \alpha$$



for $\alpha \in [\min_x g(x), g(\operatorname{argmin}_x f(x))]$

In the epitome-based RWCoseg setting, we solve a series of QCQPs,

$$\min_{x,h} \; x^T L x \quad \text{s.t.} \; Hx = h, \; 0 \le x \le 1, \quad h^T \left(\bar{h}\bar{h}^T - \alpha I\right) h \le 0$$

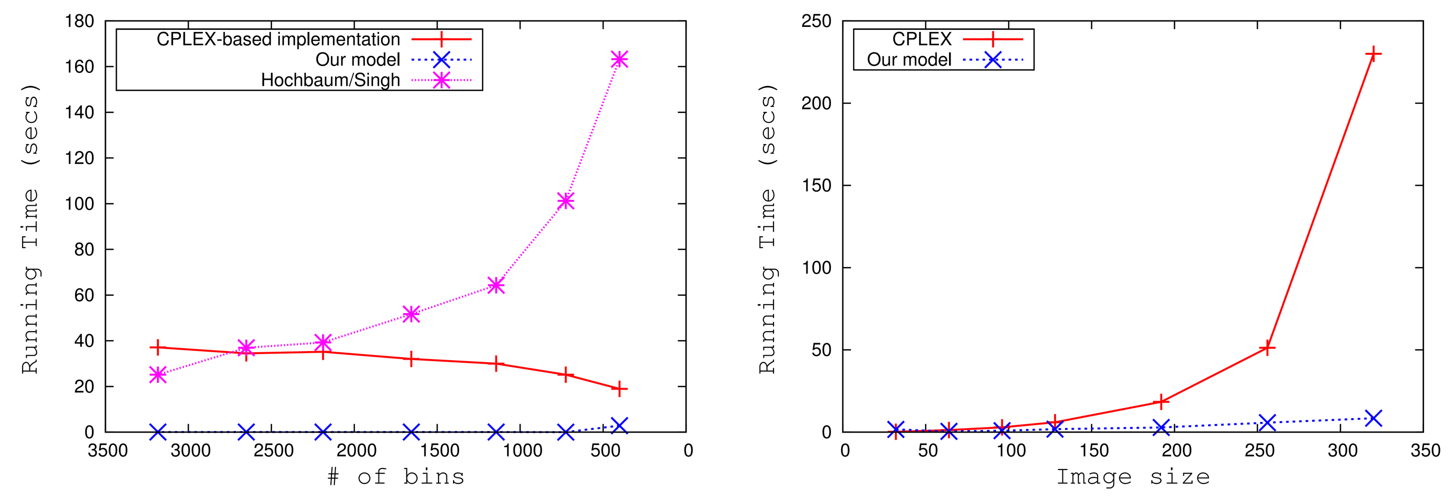


performing a grid search over α to find the optimal segmentation.

$f + g$ restricted to the α solutions will be *one-sided Lipschitz*.

Can therefore provide a lower bound on true minimum from solutions given a set of α's.

Running Time

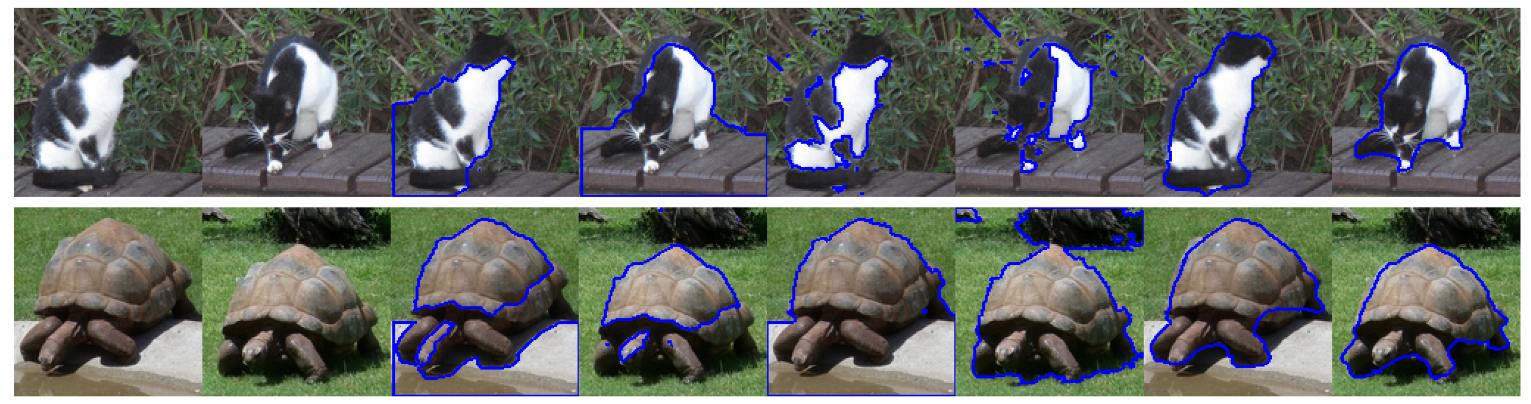


Running time varies with histogram complexity and image size.

Comparison to Independent RW

Segmentation potentials comparing Random Walks cosegmentation with independent runs on the same input.

Cosegmentation Comparison



iCoseg Dataset

	RWCoseg	Joulin
Airshow (8/21)	99.7	80.5
Alaskan bear (5/18)	92.2	76.6
Christ (6/14)	95.6	97.3
Ferrari (4/10)	96.4	59.4
Goose (13/31)	97.3	97.5
Helicopter (9/11)	96.9	98.1
Kendo (10/31)	91.4	93.4
Lobster Kite (4/11)	97.0	88.5
Monk (4/17)	83.3	87.1
Soccer (5/36)	93.9	79.9
Speed Skater (9/13)	80.8	77.3
Liberty (11/41)	93.2	86.1

Segmentation accuracy for some iCoseg image sets. Subsets were chosen which have similar appearance under a histogram model.

Cosegmentation for Video

May be applied to video, with "histograms" based on optical flow and frame-by-frame penalties $\|h_i - h_{i+1}\|_2^2$