Piecewise Flat Embedding for Image Segmentation

Yizhou Yu[†] , Chaowei Fang [†‡], Zicheng Liao[‡]

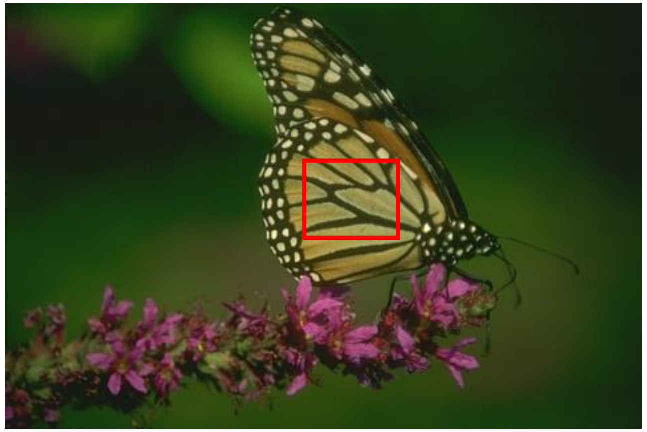
† Department of Computer Science, The University of Hong Kong

‡ College of Computer Science and Technology, Zhejiang University

Challenges in Image Segmentation

1. Small differences between neighboring pixels may accumulate to significant differences between distant pixels on the same object.
2. Textures exist everywhere.

Solution: embedding pixels into a different feature space where distances are more consistent with visual pixel grouping results


Related Work

- Normalized Cuts [Shi & Malik 2000]

Approximate solution: Rayleigh quotient problem,

$$\min_{y} \frac{y^T (D-W) y}{y^T D y} \quad s.t.\ y^T D 1 = 0$$

- Laplacian Eigenmaps [Belkin & Niyogi 2001]

Dimensionality data reduction, $L = D - W$

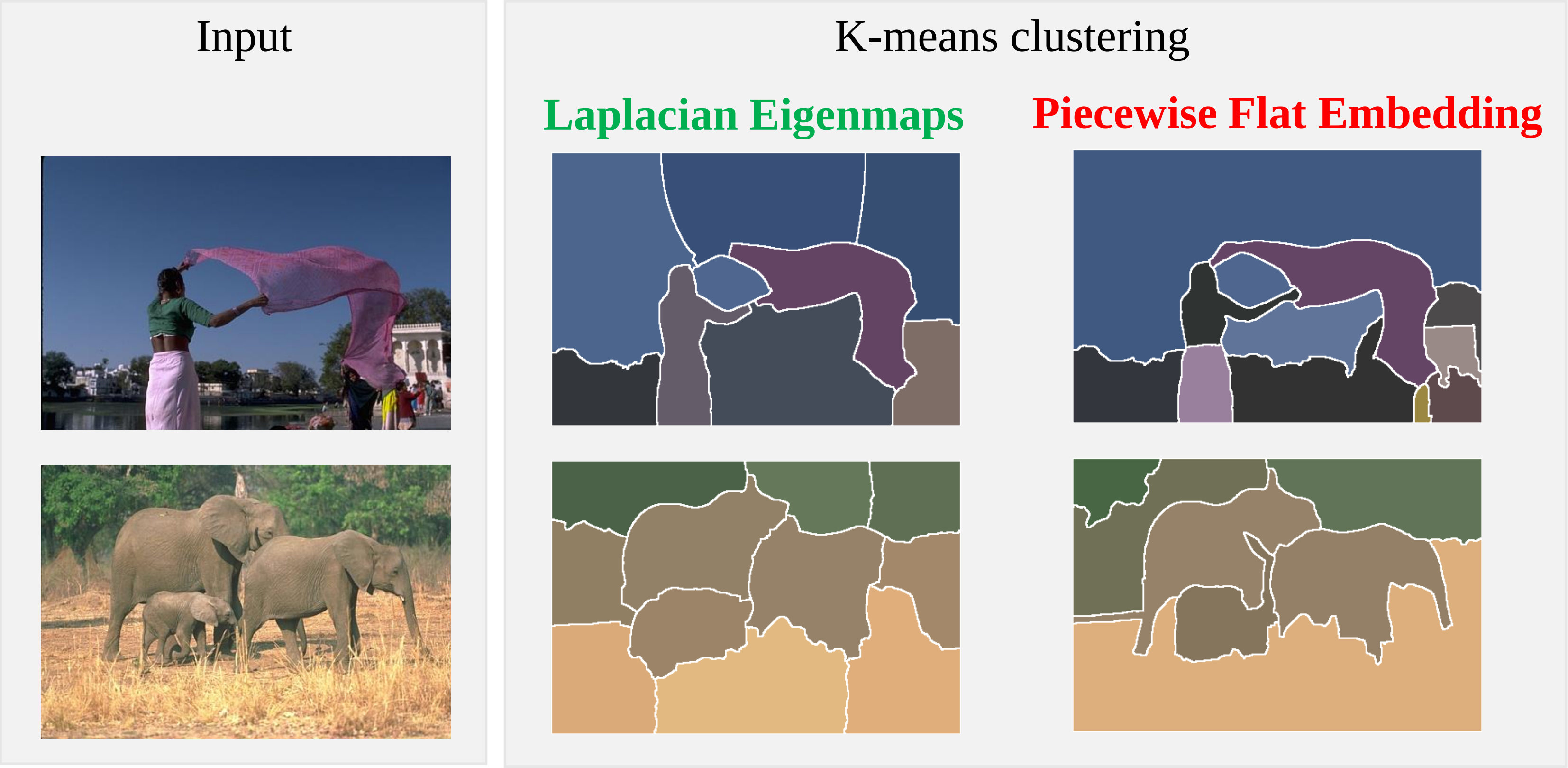
$$\min_{Y} \sum_{i,j} W_{ij} \left\| Y_i - Y_j \right\|^2 \quad s.t.\ Y^T D Y = I$$

Solution: generalized eigendecomposition $LY = \lambda DY$

Piecewise Flat Embedding

- L_1-regularized energy function adapted from Laplacian Eigenmaps

$$\min_{Y} \sum_{i,j} W_{ij} \left\| Y_i - Y_j \right\|^2 \quad s.t.\, Y^T D Y = I \;\Rightarrow\; \min_{Y} \sum_{i,j} W_{ij} \left\| Y_i - Y_j \right\|_1 \quad s.t.\, Y^T D Y = I$$



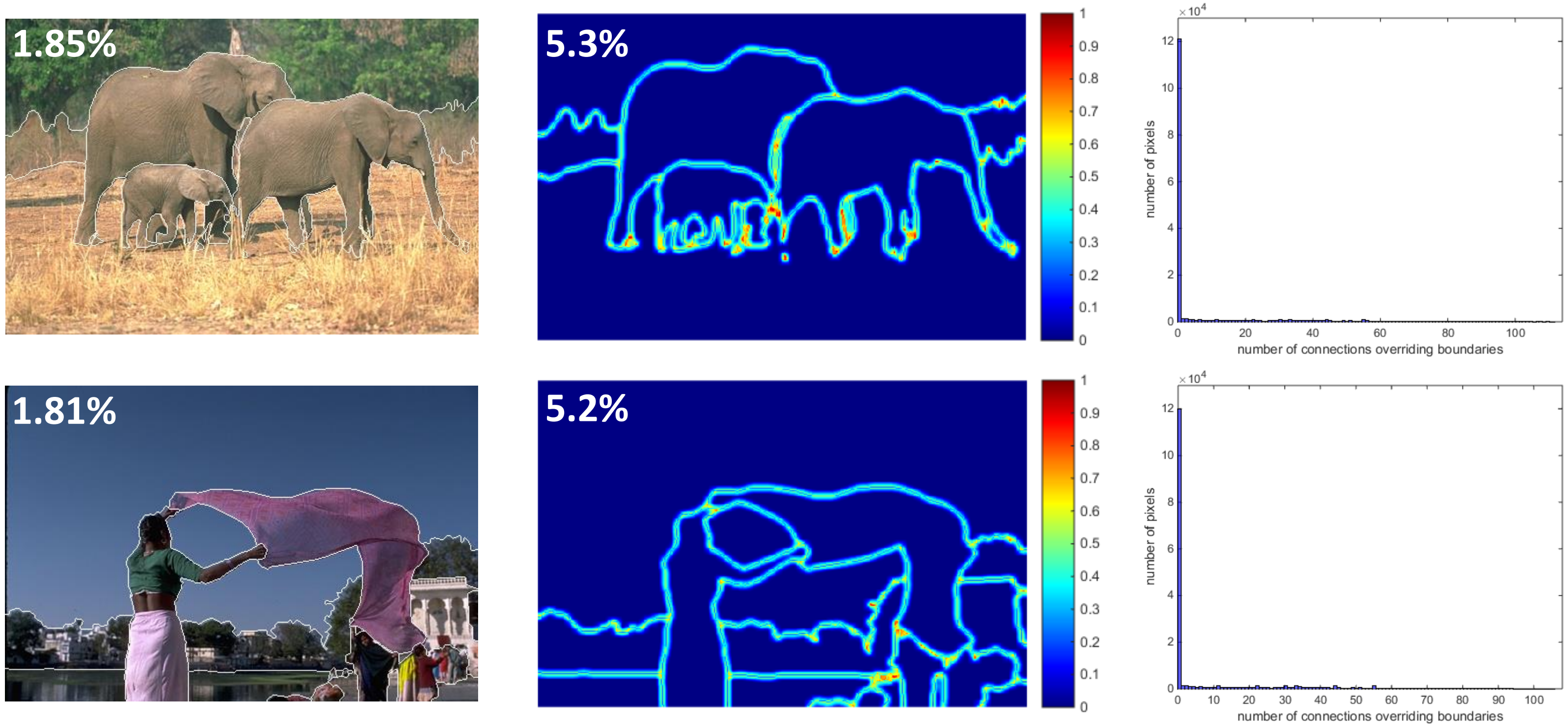
Motivations I

- Boundary Sparsity in Locally Connected Graphs

a) 1D boundaries occupy few pixels in a 2D image plane.

b) Percentage of pairwise connections crossing boundaries is very small.

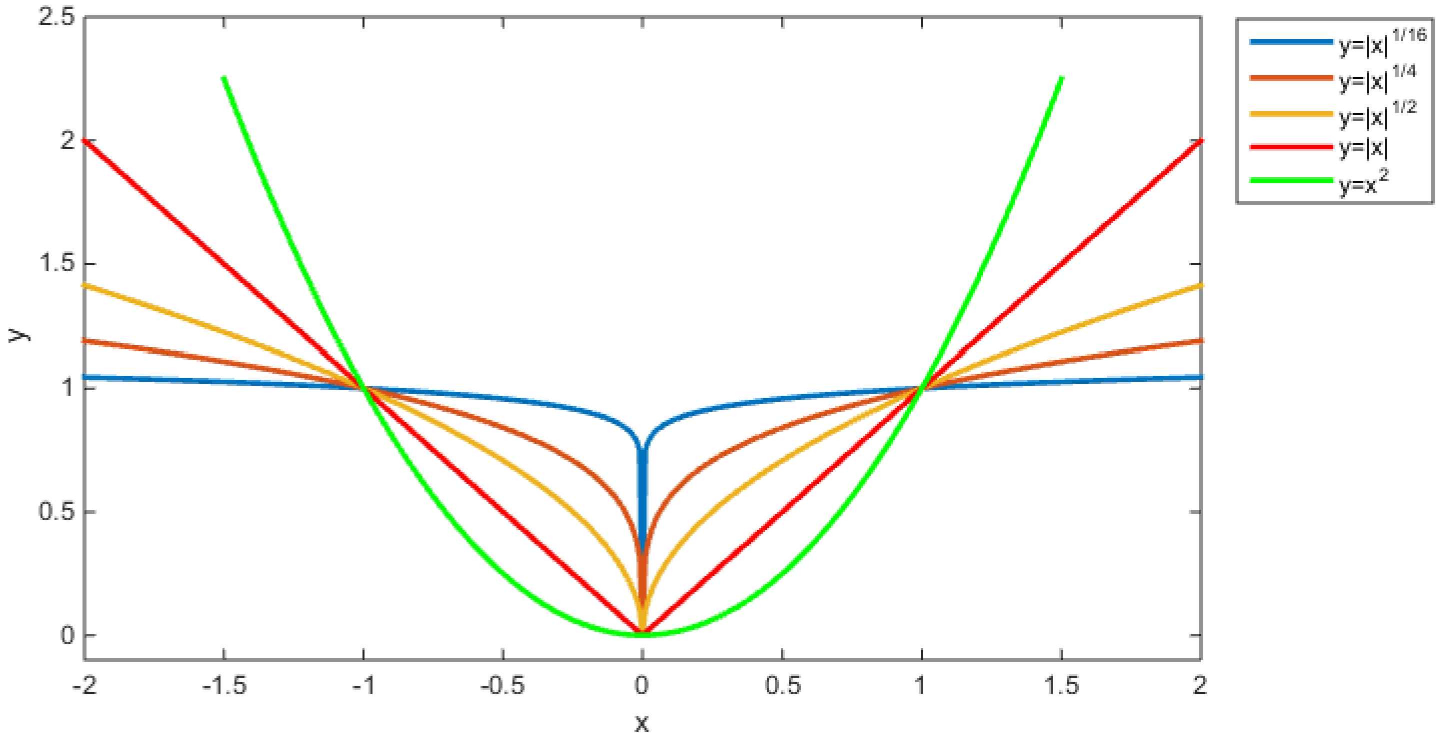
Boundary Pixels **Connections Overriding Boundaries**



Neighborhood Size 11x11

Motivations II

- L_1-norm promotes sparser solutions with few nonzero entries than L_2
- L_1-norm gives rise to a convex energy function closest to the one from L_0



Justifications

- L_1-norm in Piecewise Flat Embedding

$$\min_{Y} \sum_{i,j} W_{ij} \left\| Y_i - Y_j \right\|_1 \quad s.t. \, Y^T D Y = I$$

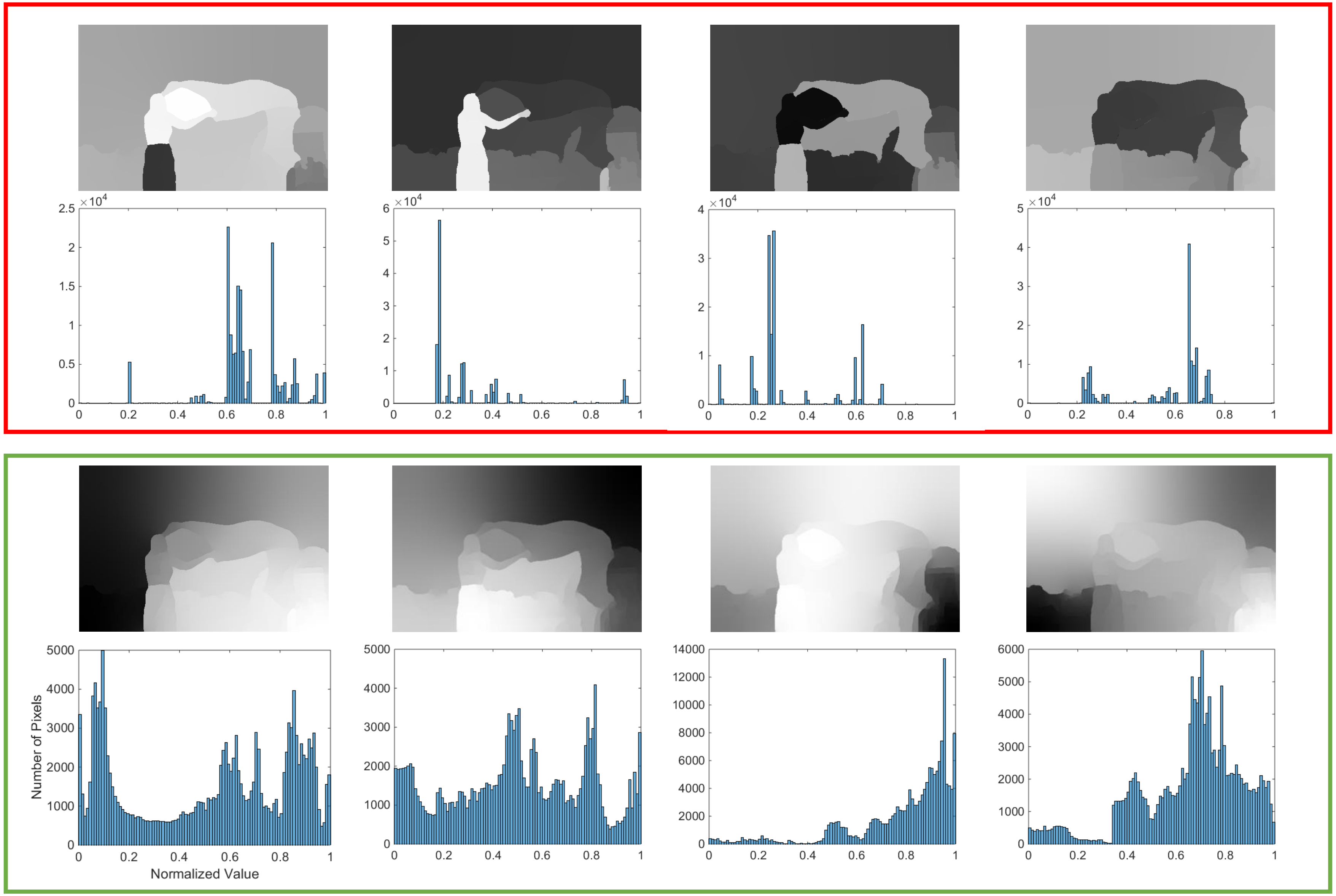
Ideally, $d_{ij} = \left| Y_i - Y_j \right| \begin{cases} > 0 & \text{if pixel } i \text{ and } j \text{ belong to different objects} \\ = 0 & \text{if pixel } i \text{ and } j \text{ belong to same object} \end{cases}$

a) Optimal Y would make nonzero entries in d_{ij}'s sparse. Thus L_1-based regularization produces solutions consistent with boundary sparsity.
b) Ideally, pixels within the same segment would have constant embedding coordinates, giving rise to piecewise flat embedding results.

An Example of Sparsity Analysis

PFE

LE



Numerical Solution I

The numerical solver for piecewise flat embedding consists of a double loop

a) Outer loop: enforce the orthogonality constraint using Splitting Orthogonality Constraint (SOC) [Lai & Osher 2014]

b) Inner loop: minimize an L_1-regularized energy using Split Bregman Iterations [Goldstein & Osher 2009]

- **Outer Loop** $\min_{Y} \sum_{i,j} W_{ij} \left\| Y_i - Y_j \right\|_1 \quad s.t.\ D^{1/2}Y = P,\ P^T P = I$

Outer Loop (SOC)

$$Y^{(k+1)} = \arg\min_{Y} \sum_{ij} W_{ij} \left\| Y_i - Y_j \right\|_1 + \frac{r}{2} \left\| D^{1/2}Y - P^{(k)} + B^{(k)} \right\|_2^2$$

Solved by an Inner Loop

$$P^{(k+1)} = \arg\min_{P} \left\| P - (D^{1/2}Y^{(k+1)} + B^{(k)}) \right\|_2^2 \quad s.t.\ P^T P = I$$

$$B^{(k+1)} = B^{(k)} + D^{1/2}Y^{(k+1)} - P^{(k+1)}$$

$$D^{1/2}Y^{(k+1)} + B^{(k)} = U\Sigma_{n\times d}V^T$$

$$P^{(k+1)} = UI_{n\times d}V^T$$

Numerical Solution II

- **Inner Loop** solves $Y^{(k+1)} = \arg\min_{Y} \sum_{ij} W_{ij} \|Y_i - Y_j\|_1 + \frac{r}{2} \|D^{1/2}Y - P^{(k)} + B^{(k)}\|_2^2$

1. Concatenate columns of $Y^{(k)}$, $P^{(k)}$ and $B^{(k)}$, forming $Y_v^{(k)}$ $P_v^{(k)}$ and $B_v^{(k)}$

$$Y^{(k+1)} = \arg\min_{Y} \sum_{ij} \|LY_v\|_1 + \frac{r}{2} \|\tilde{D}^{1/2}Y_v - P_v^{(k)} + B_v^{(k)}\|_2^2$$

2. Split Bregman Iterations

Inner Loop

$$Y_v^{(k,l+1)} = \arg\min_{Y_v} \frac{\lambda}{2} \|LY_v + b^l - d^l\|_2^2 + \frac{r}{2} \|\tilde{D}^{1/2}Y_v - P_v^{(k)} + B_v^{(k)}\|_2^2$$

Cholmod

$$d^{l+1} = Shrink(LY_v^{(k,l+1)} + b^l, \frac{1}{\lambda})$$

$$b^{l+1} = b^l + LY_v^{(k,l+1)} - d^{l+1}$$

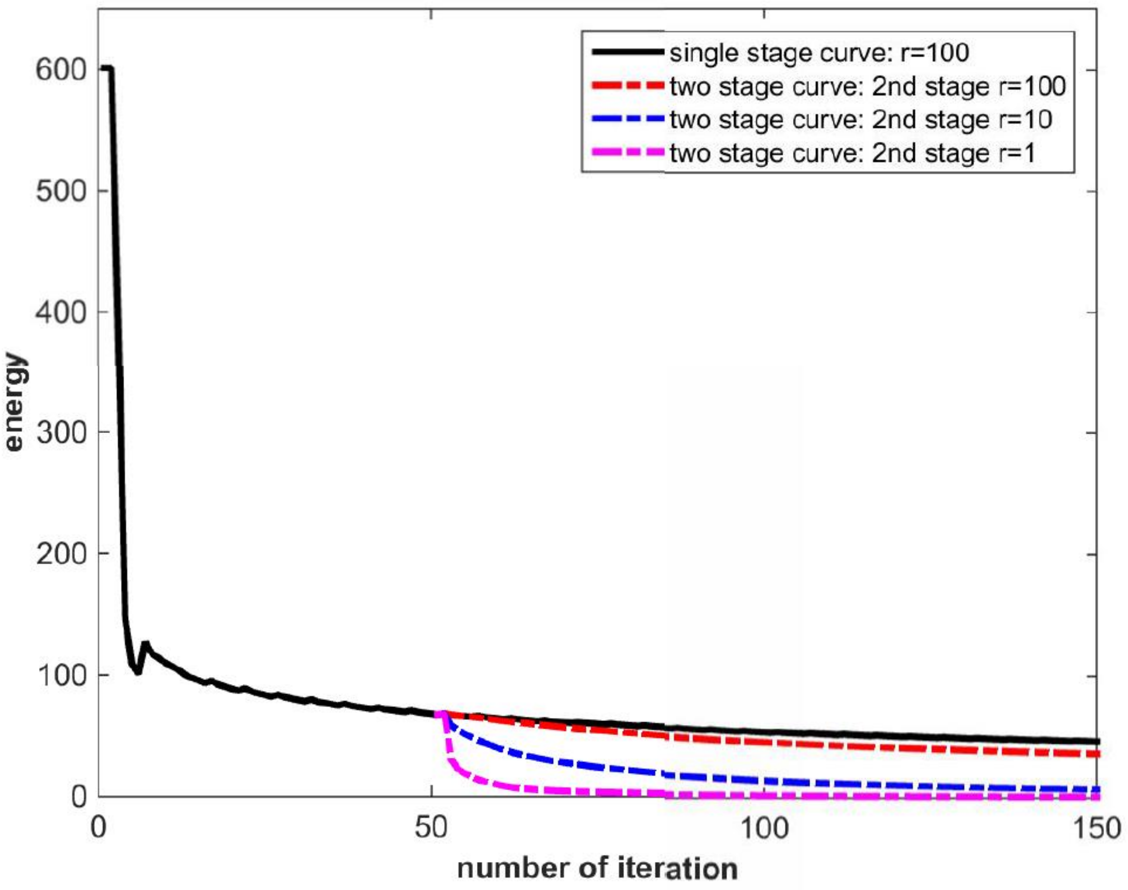
Numerical Solution III

- Two-stage Implementation

Stage I: run the full numerical solver with nested Bregman iterations.

Stage II: only minimize the L_1-regularized energy function in the inner loop without strictly enforcing orthogonality.

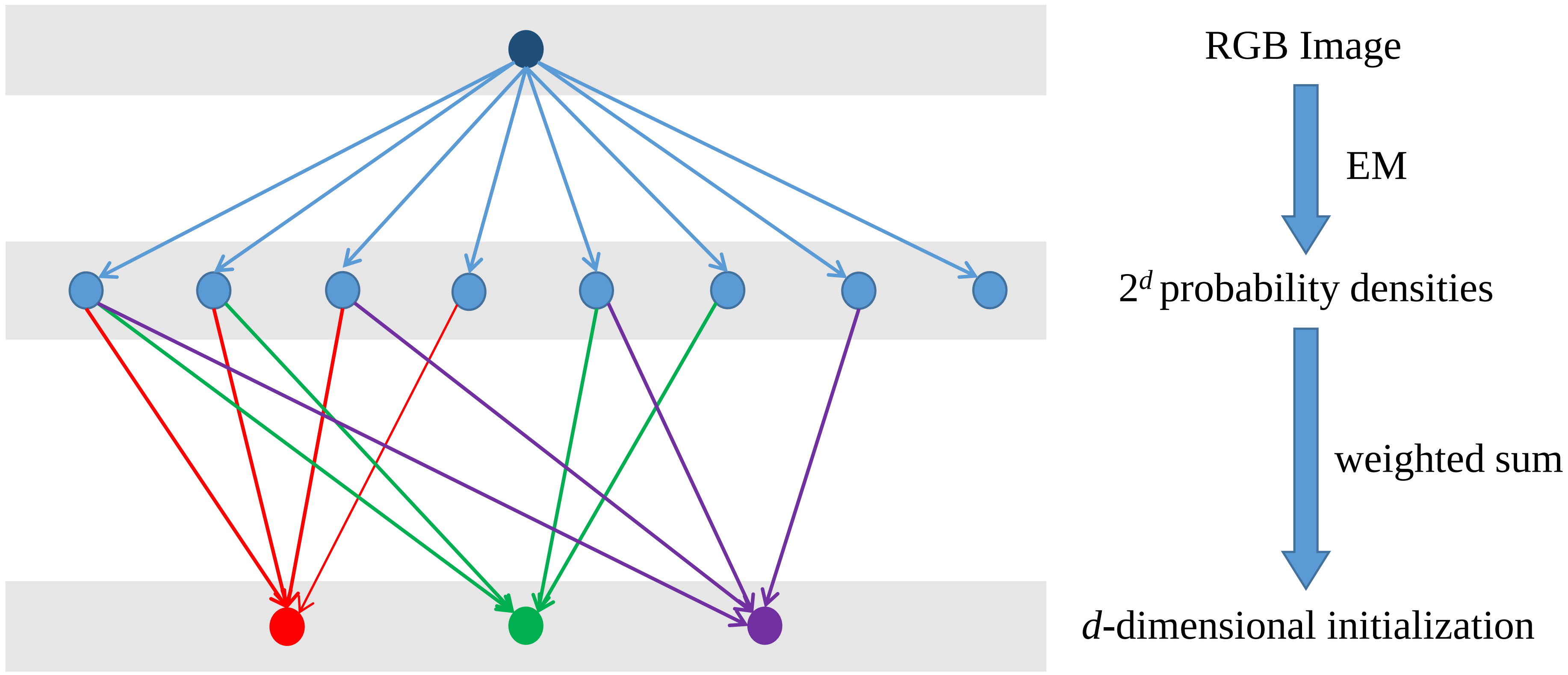
$$Y^{(k+1)} = \arg\min_{Y} \sum_{ij} W_{ij} \left\| Y_i - Y_j \right\|_1 + \frac{r}{2} \left\| D^{1/2} Y - P^{(k)} + B^{(k)} \right\|_2^2$$



Initialization

Input: original image.

Output: initial embedding with *d* channels

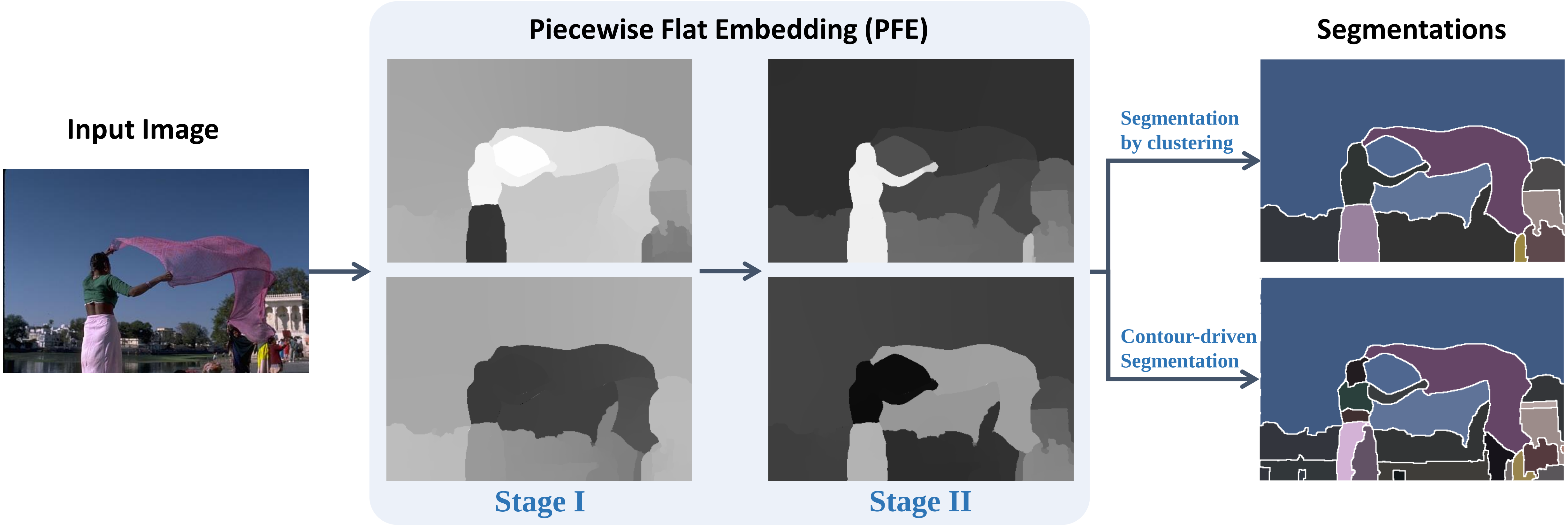


Initialization ---- An Example

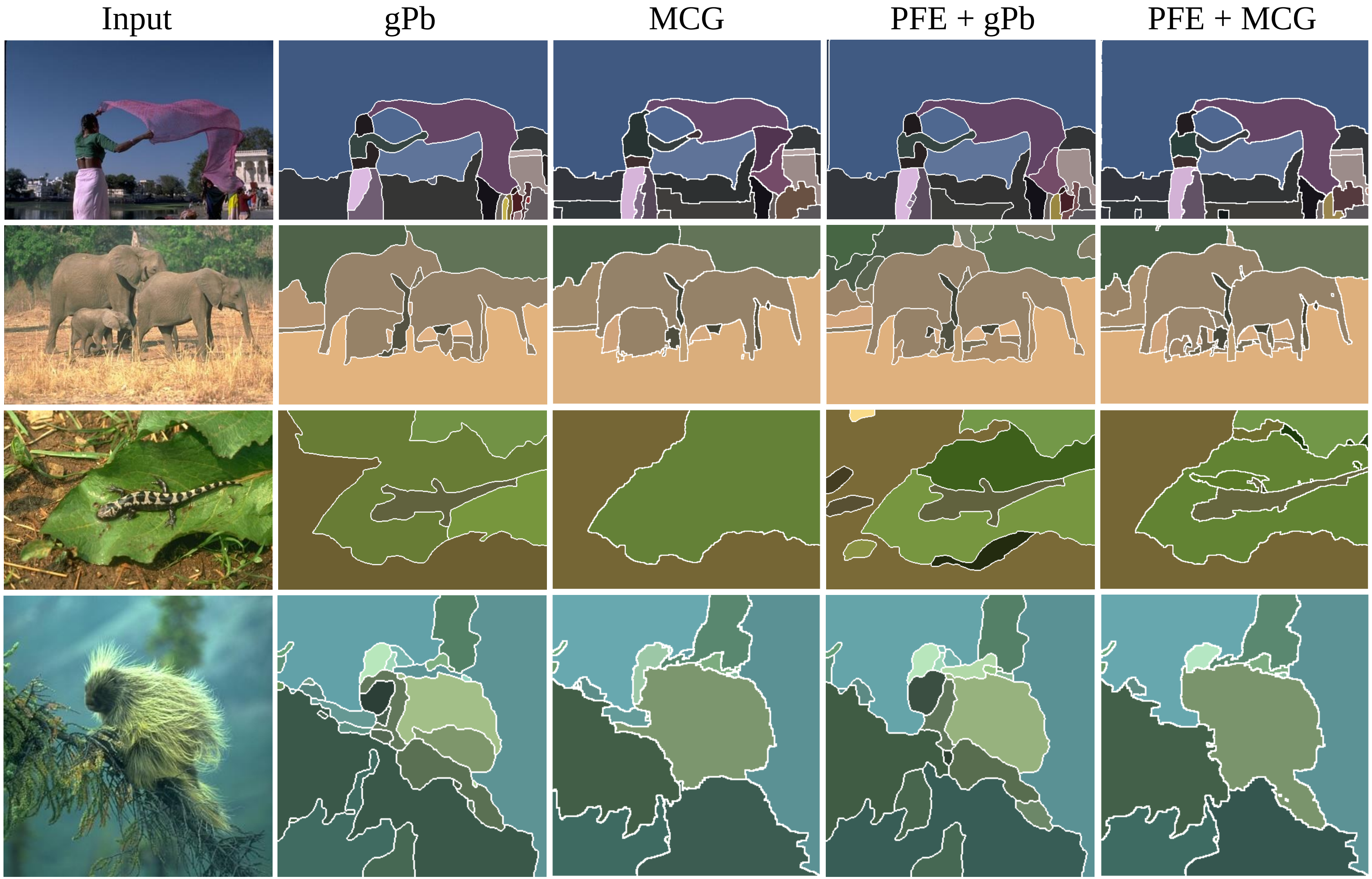
The top two rows show pixelwise density defined by the 16 components of the Gaussian Mixture Model (GMM) learned from an input image. Images in the bottom row show a GMM-based initialization of the 4 dimensions of our embedding. Each image in the bottom represents the mixed density of 8 components of the GMM.

Piecewise Flat Embedding Results

Segmentation Pipeline Using Piecewise Flat Embedding



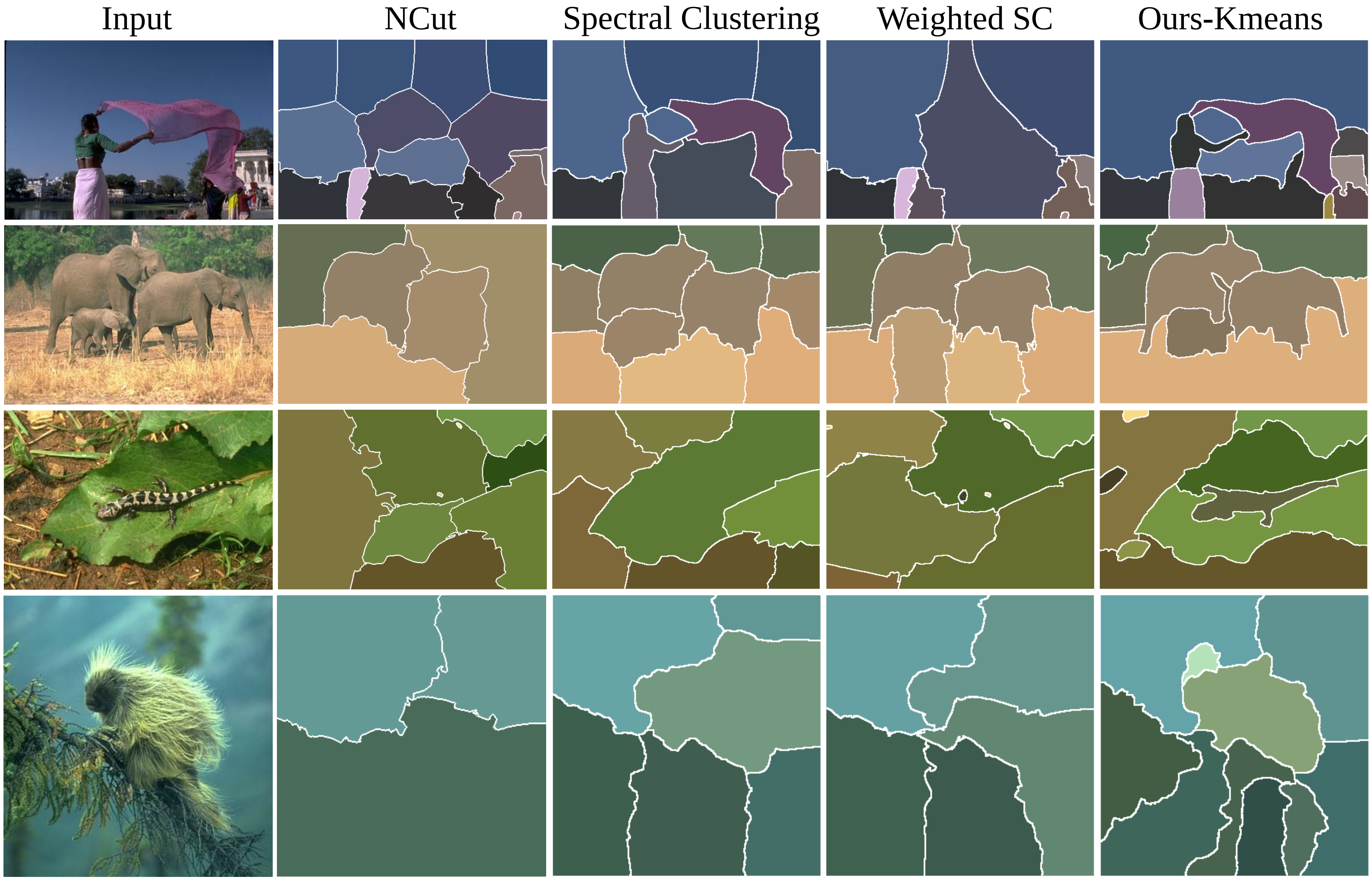
Comparison of Contour-Driven Segmentation Results



Performance of Contour-Driven Segmentation on BSDS500

method	Covering			PRI		VI	
	ODS	OIS	Best	ODS	OIS	ODS	OIS
MS-Ncut	0.45	0.53	0.67	0.78	0.80	2.23	1.89
Felz-Hutt	0.52	0.57	0.69	0.80	0.82	2.21	1.87
SCG-owt-ucm	0.51	0.56	0.66	0.78	0.83	1.98	1.84
Mean Shift	0.54	0.58	0.66	0.79	0.81	1.85	1.64
Hoiem *et al.*	0.56	0.60	-	0.80	0.77	1.78	1.66
gPb-owt-ucm	0.59	0.65	0.74	0.83	0.86	1.69	1.48
ISCRA	0.59	0.66	-	0.82	0.85	1.60	1.42
MCG	0.61	0.66	0.76	0.83	0.86	1.57	1.39
PFE+mPb	**0.62**	0.67	0.76	0.84	0.86	1.61	1.43
PFE+MCG	**0.62**	**0.68**	**0.77**	**0.84**	**0.87**	**1.56**	**1.36**

Comparison of Clustering-Based Segmentation Results



Performance of Clustering-Based Segmentation on BSDS500

- Affinity matrix in the original Normalized Cut

method	Covering		PRI		VI	
	fixed	dynamic	fixed	dynamic	fixed	dynamic
Ncut	0.33	0.40	0.75	0.76	2.77	2.39
SC	0.36	0.44	0.75	0.77	2.68	2.24
WSC	0.36	0.44	0.75	0.77	2.63	2.21
Ours	**0.46**	**0.52**	**0.77**	**0.79**	**2.21**	**1.91**

- Affinity matrix in gPb-owt-ucm

method	Covering		PRI		VI	
	fixed	Dynamic	Fixed	dynamic	fixed	dynamic
SC	0.35	0.45	0.76	0.77	2.66	2.17
WSC	0.35	0.44	0.76	0.77	2.67	2.20
Ours	**0.45**	**0.56**	**0.78**	**0.81**	**2.26**	**1.77**

An L_1 Image Transform for Edge-Preserving Smoothing and Scene-Level Intrinsic Decomposition [SIGGRAPH 2015]

Sai Bi, Xiaoguang Han, Yizhou Yu

Input Images

Shading Image

Our Smoothing Results

Reflectance Image

Conclusions

- We propose piecewise flat embedding which adopts an L_1-regularized energy term to promote sparse solutions.
- We devise an effective two-stage numerical algorithm based on Bregman iterations to solve the proposed embedding.
- Experiments indicate that segmentation algorithms incorporating piecewise flat embedding achieve much improved results.

Acknowledgments

- Hong Kong Research Grants Council under General Research Funds (HKU719313)
- Zhejiang Natural Science Foundation (ZJNSF Q15F020006)