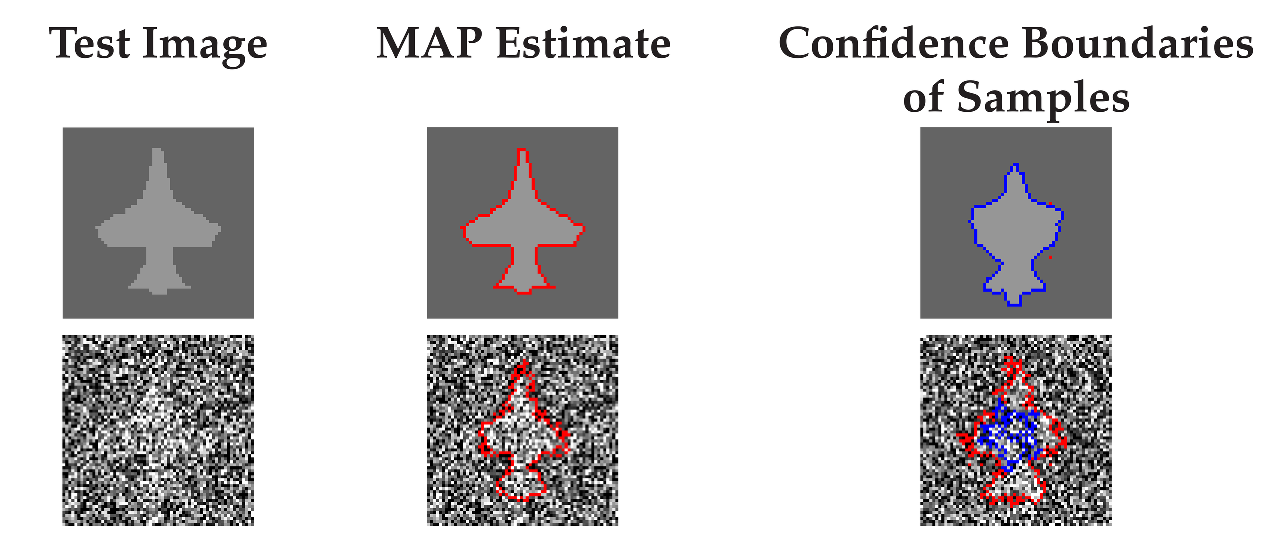
Image Segmentation with Pseudo-Marginal MCMC Sampling and Nonparametric Shape Priors

Ertunc Erdil[1], Sinan Yildirim[1], Tolga Tasdizen[2], Mujdat Cetin[1,3]
[1]Sabanci University, Istanbul, Turkey [2]University of Utah, Salt Lake City, UT, USA [3]University of Rochester, Rochester, NY, USA
{ertuncerdil, sinanyildirim, mcetin}@sabanciuniv.edu tolga@sci.utah.edu

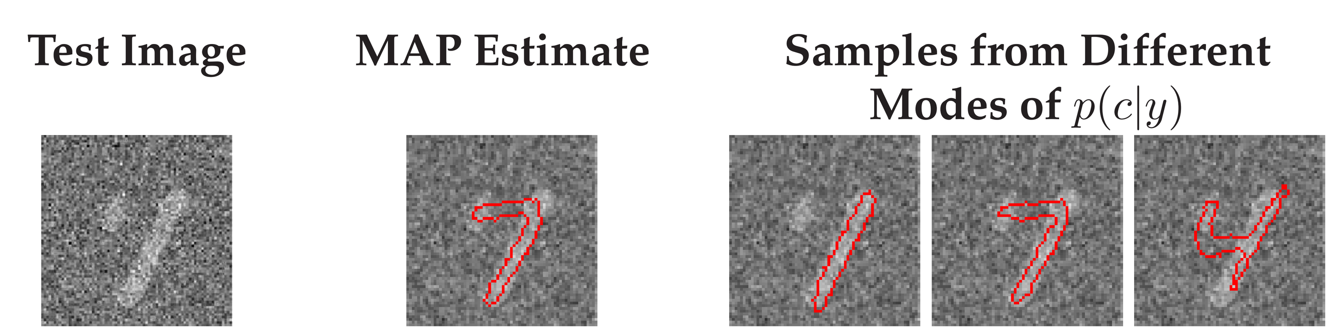
Motivation

- Segmentation of images having low quality or missing data requires prior shape information [1]
- Segmentation is often formulated as maximum *a posteriori* (MAP) estimation: $\hat{c} = \underset{c}{\operatorname{argmax}}\, p(c|y)$
- MAP estimation does not provide

(1) a measure of the degree of confidence in that result,



(2) a picture of other probable solutions based on the data and the priors.



Approach:

- Instead of MAP estimation, characterize the entire posterior distribution through its samples, potentially generating samples from different modes of a **multimodal** probability density.

Challenges:

- Generating samples from $p(c|y)$ is non-trivial?
- Conventional MCMC sampling become inefficient when the training set size increases.

Previous Work on MCMC-based Segmentation

- [2], [3], [4] can generate samples from $p(c|y) \propto p(y|c)$ by assuming that $p(c)$ is uniform.
- [2] uses explicit (marker-based) shape representation
- [5] is restricted to simple closed curve
- [6] assumes that the underlying shape distribution is unimodal
- [7] becomes inefficient for large training sets
- All above-mentioned methods approximately satisfies necessary conditions to implement MCMC sampling

References

[1] J. Kim, M. Cetin and A. S. Willsky. Nonparametric shape priors for active contour-based image segmentation In *Signal Processing '07.*

[2] A. C. Fan, J. W. Fisher, W. M. Wells III, J. J. Levitt and A. S. Willsky MCMC curve sampling for image segmentation In *MICCAI '07.*

[3] C. Chang and J. W. Fisher Efficient MCMC sampling with implicit shape representation In *CVPR '11.*

[4] C. Chang and J. W. Fisher Efficient topology-controlled sampling of implicit shapes In *ICIP '11.*

[5] S. Chen and R. J. Radke Markov chain monte carlo shape sampling using level sets In *ICCV Workshops '09.*

[6] M. D. Bruijne and M. Nielsen Image segmentation by shape particle filtering In *ICPR '04.*

[7] E. Erdil, S. Yildirim, M. Cetin and T. Tasdizen MCMC shape sampling for image segmentation with nonparametric shape priors In *CVPR '16.*

Contributions

- We propose a Markov chain Monte Carlo (MCMC) sampling-based image segmentation approach that exploits nonparametric shape priors.
- The proposed approach achieves scalability as the training set size grows by exploiting pseudo-marginal sampling. To the best of our knowledge, pseudo-marginal sampling has not been used for image segmentation.
- The proposed pseudo-marginal MCMC sampling approach perfectly satisfies necessary conditions to implement MCMC. To the best of our knowledge, existing methods in the literature approximately satisfy these conditions.

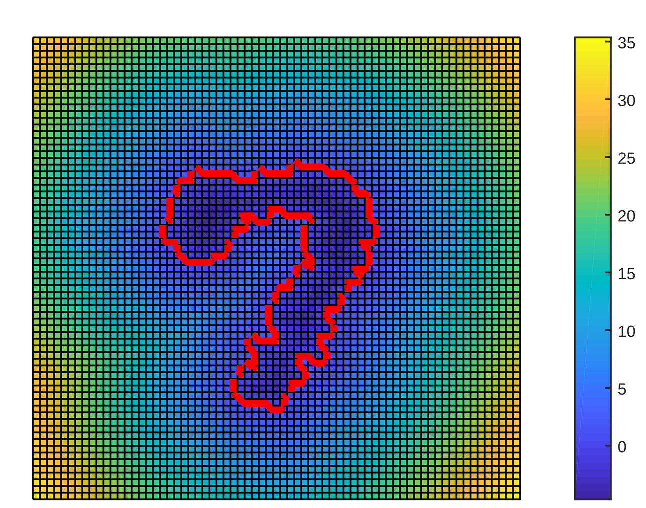
Model and Problem Definition

- Bayesian image segmentation can be formulated as

$$p(x|y) \propto p(y|x)p(x) = p(y|x)\sum_{s=1}^{n} p(s)p(x|s)$$

where x is the level set representation of the segmenting curve c, y is the input image and $s \sim \mathcal{U}(1,n)$ ($p(s) = 1/n$) is the class of the shape to be segmented as we are interested in segmentation problems where $p(x)$ is multimodal.

- **Level set representation, x, of a curve, c, shown by red.**



- Given level set representations of training shapes $\mathbf{X} = \{\mathcal{X}_1, \ldots, \mathcal{X}_n\}$ where each $\mathcal{X}_i = \{x_{i,1}, \ldots, x_{i,m_i}\}$ is the collection of $m_i \geq 1$ level set representation of segmented curves for class i, we estimate $p(x|s)$ as $p(x|s) = \frac{1}{m_s}\sum_{i=1}^{m_s} \mathcal{N}(x; x_{s,i}, \sigma^2 I)$ where $\mathcal{N}(x;\mu,\Sigma)$ is a Gaussian kernel with mean μ and covariance matrix Σ.
- Estimating $p(x|y)$ can be difficult since the summation over classes makes the distribution hard to infer, e.g., using Monte Carlo sampling methods.
- Alternatively, we aim for the joint posterior distribution of s and x given y ($p(s,x|y) \propto p(s,x,y)$) whose marginal is still the desired posterior $p(x|y)$.

Pseudo-marginal Sampling

- Computing $p(x|s)$ requires evaluation of m_s multivariate Gaussian densities of dimension MN.
- Assume that we have a non-negative random variable z such that given x and s, its conditional density $g_{s,x}(z)$ satisfies

$$\int_0^\infty g_{s,x}(z)zdz = p(x|s).$$

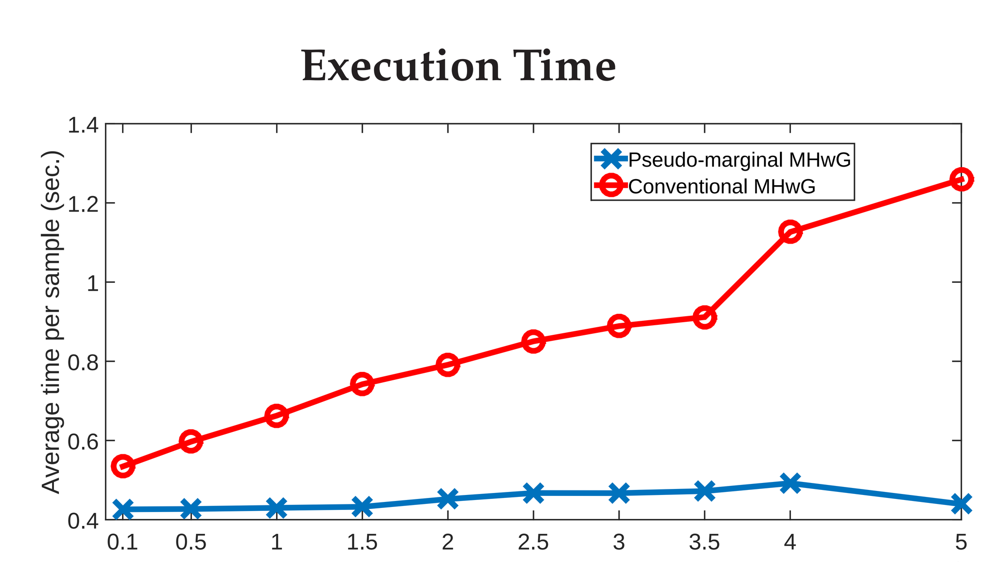
We choose z as an approximation to $p(x|s)$, in particular $z = \widehat{p}(x|s) = \frac{1}{\hat{m}_s}\sum_{j=1}^{\hat{m}_s} \mathcal{N}(x; x_{s,u_j}, \sigma^2 I)$ where $\{u_1, u_2, \ldots, u_{\hat{m}_s}\} \subset \{1, 2, \ldots, m_s\}$ is a set of subsamples generated via sampling without replacement and $\hat{m}_s \ll m_s$.

- We define the extended posterior density that we generate sample as $p(x,s,z|y) \propto p(s)zg_{s,x}(z)p(y|x)$

Proposed Method

1: Initialize $s^{(0)}, x^{(0)}$ and generate an approximation of the prior $z^{(0)} = \widehat{p}(x^{(0)}|s^{(0)})$ using subsampling.
2: **for** $t = 1 \to N$ **do**
3: Set $s = s^{(t-1)}$, $x = x^{(t-1)}$, and $z = z^{(t-1)}$.
$\triangleright$ Steps $4-13$ generate samples of $(s\prime, z\prime)$ from $p(s,z|y,x)$.
4: Sample $s' \sim q(s'|s)$.
5: Generate an approximation $z' = \widehat{p}(x|s')$ of $p(x|s')$ using a subsample of size $\hat{m}_{s'}$ where $\hat{m}_{s'} \ll m_{s'}$.
6: Compute $\alpha = \min\left\{1, \frac{p(s')z'q(s|s')}{p(s)zq(s'|s)}\right\}$,
7: Sample $\eta \sim \mathcal{U}(0,1)$.
8: **if** $\alpha > \eta$ **then**
9: $s^{(t)} = s'$; $z^{(t)} = z'$; $\triangleright$ Accept the candidate
10: **else**
11: $s^{(t)} = s$; $z^{(t)} = z$; $\triangleright$ Reject the candidate
12: **end if**
13: Set $s = s^{(t)}$, and $z = z^{(t)}$.
$\triangleright$ Steps $14-23$ generate samples of $(x\prime, z\prime)$ from $p(x,z|y,s)$.
14: Uniformly draw a random number j in $\{1, \ldots, m_s\}$.
15: Sample x' from $q_{s,j}(x'|x,y) = \mathcal{N}\left(x'; x - \widehat{\nabla E}_{s,j}(x), \Sigma\right)$.
16: Generate an approximation $z' = \widehat{p}(x'|s)$ of $p(x'|s)$ using a subsample of size $\hat{m}_s$.
17: Compute $\alpha = \min\left\{1, \frac{z'p(y|x')q_{s,j}(x|x',y)}{zp(y|x)q_{s,j}(x'|x,y)}\right\}$,
18: Sample $\eta \sim \mathcal{U}(0,1)$.
19: **if** $\alpha > \eta$ **then**
20: $x^{(t)} = x'$; $z^{(t)} = z'$; $\triangleright$ Accept the candidate
21: **else**
22: $x^{(t)} = x$; $z^{(t)} = z$; $\triangleright$ Reject the candidate
23: **end if**
24: **end for**

Experimental Results



Dice Score

Training Set Size	Pseudo-marginal Shape Sampling	Conventional MCMC Sampling
1000	0.7736	0.7758
5000	0.7735	0.7818
10000	0.7744	0.7765
15000	0.7756	0.7745
20000	0.7706	0.7791
25000	0.7689	0.7776
30000	0.7703	0.7767
35000	0.7700	0.7769
40000	0.7691	0.7801
50000	0.7719	0.7776
Average	0.7718	0.7777