EFFICIENT TOPOLOGY-CONTROLLED SAMPLING OF IMPLICIT SHAPES

Jason Chang and John W. Fisher III

Massachusetts Institute of Technology
Department of Electrical Engineering and Computer Science
32 Vassar St. Cambridge, MA 02139

ABSTRACT

Sampling from distributions of implicitly defined shapes enables analysis of various energy functionals used for image segmentation. Recent work [1] describes a computationally efficient Metropolis-Hastings method for accomplishing this task. Here, we extend that framework so that samples are accepted at every iteration of the sampler, achieving an order of magnitude speed up in convergence. Additionally, we show how to incorporate topological constraints.

***Index Terms*—** Markov chain Monte Carlo, MCMC, Metropolis-Hastings, level sets, segmentation

1. INTRODUCTION

For many Bayesian inference tasks, evaluating marginal event probabilities may be more robust than computing point estimates (*e.g.* the MAP estimate). Image segmentation, particularly when the signal-to-noise ratio (SNR) is low, is one such task. However, because the space of shapes is infinitely large, direct inference or sampling is often difficult, if not infeasible. In these cases, Markov chain Monte Carlo (MCMC) sampling approaches can be used to compute empirical estimates of marginal probabilities. Recently, the work of [1] derived an efficient MCMC method for sampling from distributions of implicit shapes (*i.e.* level sets). We improve upon that algorithm in two ways. First, we improve the convergence rate by defining a Gibbs-like iteration in which *every* sample is accepted and, second, we demonstrate how to efficiently incorporate both local and global topological constraints on sample shapes.

Many approaches formulate image segmentation as an energy optimization. One can derive a related Bayesian inference procedure by viewing the energy functional $E(\ell; x)$ as the log of a probability function

$$p_{L|X}(\ell|x) \propto \exp\left[\pm E(\ell; x)\right], \tag{1}$$

where ℓ is the labeling associated with some segmentation, x is the observed image, and the $\pm$ in the exponent depends on whether one is maximizing or minimizing. In PDE-based level-set methods, $\ell = \text{sign}(\varphi)$, where φ is the level-set function. In graph-based segmentation algorithms, such as ST-Mincut [2] [3] or Normalized Cuts [4], ℓ is the label assignment. Often, the energy functional decomposes into a data fidelity term and a regularization of the segmentation (Normalized Cuts being an exception). Bayesian formulations treat the former as the data likelihood and the latter as a prior on segmentations:

$$p_{L|X}(\ell|x) \propto p_{X|L}\left(x|\ell\right) p_L\left(\ell\right). \tag{2}$$

This research was partially supported by the Air Force Office of Scientific Research under Award No. FA9550-06-1-0324. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the views of the Air Force.

While both the data likelihood and prior terms are user-defined, the form of the prior varies considerably depending on the optimization method; a curve-length penalty is typically used in level-set methods while a neighborhood affinity is typically used in graph-based methods. We refer to these two forms as the PDE-based energy and the graph-based energy, respectively.

Another form of prior information incorporates the topology of the segmented object. Han *et al.* [5] first showed this using topology-preserving level-set methods, where the topology of the segmented object was not allowed to change. This methodology was later extended in [6] to topology-controlled level-set methods, where the topology was allowed to change but only within an allowable set of topologies. To our knowledge, topology-constrained sampling methods have not previously been considered.

While [7], [8], and [1] all present Metropolis-Hastings MCMC sampling methods to generate implicitly represented shapes, the first two have been shown to converge very slowly and cannot accommodate any topological changes. Recently, however, [1] presented an algorithm called BFPS that acts directly on the level-set function. BFPS generates proposals by sampling a sparse set of delta functions and smoothing them with an FIR filter. Delta locations and heights were biased by the gradient of the energy to increase the Hastings ratio. Empirical results demonstrated that BFPS was orders of magnitude faster than [7] and [8], and representations, unlike the previous two methods, could change topology. Here, we show how to efficiently sample from a distribution over segmentations in both the PDE-based and graph-based energies. In contrast to previous MCMC samplers, proposals at each iteration are accepted with certainty, achieving an order of magnitude speed up in convergence. Additionally, we incorporate topological control to exploit prior knowledge of the shape topology.

2. METROPOLIS-HASTINGS MCMC SAMPLING

We begin with a brief discussion of two MCMC sampling algorithms (*c.f.* [9] for details). For notational convenience, we drop the dependence on x in distributions. We denote $p_{L|X}(\ell)$ as the *target distribution*, *i.e.* the distribution from which we would like to sample. MCMC methods are applicable when one can compute a value proportional to the target distribution, but not sample from it directly. Distributions defined over the infinite-dimensional space of shapes fall into this category. MCMC methods address this problem by constructing a first-order Markov chain such that the stationary distribution of that chain is exactly the target distribution. For this condition to hold, it is sufficient to show that the chain is *ergodic* and satisfies *detailed balance*. The Metropolis-Hastings sampler (MH-MCMC) [9] is one such approach. At time t in the chain, the algorithm samples from some user-defined proposal distribution,

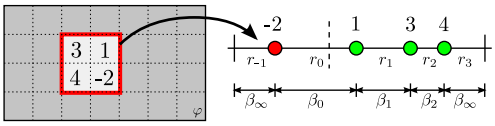

Fig. 1. Mapping the masked level-set function (left) onto the real line (right). Possible ranges, r, and widths, β, are shown on the right.

$\hat{\varphi}^{(t)} \sim q(\hat{\varphi}^{(t)}|\varphi^{(t-1)}, x)$, and assigns the transition probability:

$$\varphi^{(t)} = \begin{cases} \hat{\varphi}^{(t)} & \text{w/ prob} \quad \min(\alpha^{(t)}, 1) \\ \varphi^{(t-1)} & \text{w/ prob} \quad 1 - \min(\alpha^{(t)}, 1) \end{cases}, \tag{3}$$

$$\alpha^{(t)} = \frac{p_{L|X}(\hat{\ell}^{(t)})}{p_{L|X}(\ell^{(t-1)})} \frac{q(\varphi^{(t-1)}|\hat{\varphi}^{(t)}, x)}{q(\hat{\varphi}^{(t)}|\varphi^{(t-1)}, x)}, \tag{4}$$

where $\alpha^{(t)}$ is the Hasting's ratio, and the segmentation labels are $\hat{\ell}^{(t)} = \text{sign}(\hat{\varphi}^{(t)})$ and $\ell^{(t-1)} = \text{sign}(\varphi^{(t-1)})$. As this transition probability satisfies detailed balance, the correct stationary distribution is ensured as long as the chain is ergodic. A single sample from the target distribution is then generated by repeating this proposal/transition step until the chain converges.

Gibbs sampling is a special case of MH-MCMC, where $\alpha^{(t)} = 1$. Gibbs proposals typically select a random dimension and sample from the target distribution conditioned on all other dimensions. Empirical evidence indicates that when learning appearance models (e.g. in [10]), Gibbs sampling exhibits slow convergence times. Block Gibbs sampling, where a group of dimensions are sampled conditioned on all other dimensions, allows for larger moves. The proposed algorithm is related to this type of Gibbs sampling.

3. GIBBS-INSPIRED METROPOLIS-HASTINGS

BFPS [1] achieves a high acceptance ratio at each iteration by biasing the proposal by the gradient of the energy functional. Here, we design a similar proposal such that samples are accepted with certainty. Given some previous level-set function, $\varphi^{(t-1)}$, we generate a proposal, $\hat{\varphi}^{(t)}$ with the following steps: (1) sample a mask, m, that selects a subset of pixels and (2) add a constant value, f, to all pixels within this mask. We can express this as

$$\hat{\varphi}^{(t)}(m, f) = \varphi^{(t-1)} + f \cdot m, \tag{5}$$

where the mask, m, is a set of indicator variables with the same size as φ. The support of the mask can be of any shape and size, though in practice we use circles of random center and radius. Deferring the choice of f, we write the proposal likelihood as

$$q(\hat{\varphi}^{(t)}|\varphi^{(t-1)}, x) = p_M(m)\, p_{F|XM\Phi}(f|m, \varphi^{(t-1)}). \tag{6}$$

Figure 1 shows a notional mask overlaid on a level-set function, φ, with the height of φ plotted on the real line for pixels in the support of the mask. The dotted line marks the zero-height which splits pixels into foreground and background. As all pixels in the support of the mask are summed with the same constant, f, choosing an f is equivalent to shifting the zero-height by $-f$. The resulting proposed zero-height can fall into one of $(|m| + 1)$ different *ranges*, where $|m|$ counts the pixels in the support of the mask.

Sampling a perturbation, f, can then be decomposed into sampling a range followed by a value within that range. Most energy functionals only depend on the sign of pixels, which conveniently is determined solely by the range. If we choose $p_{F|R}(f|r) = \frac{1}{\beta_r}$ (i.e. a uniform distribution), we can write the perturbation likelihood as

$$p_{FR|XM\Phi}(f|m, \varphi^{(t-1)}) = \frac{1}{\beta_r} p_{R|XM\Phi}(r|m, \varphi^{(t-1)}), \tag{7}$$

where β_r is the width or range r (note that r is deterministic conditioned on f). Because the value of f within a range r does not affect the sign of the resulting level-set function, the proposed labeling can be expressed as

$$\hat{\ell}^{(t)}(m, r) = \text{sign}(\hat{\varphi}^{(t)}(m, f \sim p_{F|R}(f|r)). \tag{8}$$

By choosing the following Gibbs-like proposal for the ranges,

$$p_{R|XM\Phi}(r|m, \varphi^{(t-1)}) \propto \beta_r\, p_{L|X}(\hat{\ell}^{(t)}(m, r)), \tag{9}$$

one can verify via Equations 4, 6, and 7 that the Hastings ratio evaluates to one at every iteration.

One subtlety exists in the endpoints of the first and last range. Since these ranges extend to $\pm\infty$, their corresponding β is also ∞. We therefore restrict both ranges to be of some finite length β_∞. While any finite value will suffice, in practice, we choose $\beta_\infty = 1$.

The resulting algorithm, called the Gibbs-Inspired Metropolis Hastings Shape Sampler (GIMH-SS), is described by the following steps: (1) Sample a circular mask, m, with a random center and radius; (2) Sample a range, r, according to Equation 9; (3) Sample a perturbation, f, uniformly in this range; (4) Compute $\hat{\varphi}^{(t)}(m, f)$ using Equation 5; (5) Repeat from Step 1 until convergence. Preliminary results (*c.f.* [11]) show that GIMH-SS draw a single sample from a 500x300 pixel image in approximately 10 seconds. This convergence speed is approximately 10x faster than BFPS, 10^5x faster than [8], and 10^6x faster than [7].

3.1. Relation to Block Gibbs Sampling

Block Gibbs sampling first selects a mask of pixels (or dimensions) to changes, and then samples from the target distribution conditioned on all other pixels. In binary segmentation, a block of size $|m|$ would require one to evaluate an exponential number ($2^{|m|}$) of different configurations, which is intractable for a reasonably sized mask.

The GIMH-SS algorithm similarly selects a mask of pixels. The level-set function orders the masked pixels, in that, if a pixel of height $+h$ changes sign, all pixels of height $0 < \varphi_i < h$ must as well. Consequently, this algorithm samples from a subset of the conditional distribution, resulting in a linear number ($|m| + 1$) of different configurations. We note that ergodicity is ensured because the ordering of pixels by the level-set function changes over time.

4. TOPOLOGY-CONTROLLED SAMPLING

In this section, we extend GIMH-SS to a topology-controlled sampler. The topology of a continuous, compact surface is often described by its genus (i.e. the number of "handles"). Digital topology [12] is the discrete counterpart of continuous topology, where regions are represented via binary variables on a grid.

In digital topology, *connectiveness* must be defined for the foreground (FG) and background (BG) describing how pixels in a local neighborhood are connected. For example, in 2D, a 4-connected region corresponds to a pixel being connected to its neighbors above, below, left, and right. An 8-connected region corresponds to being connected to the eight pixels in a 3x3 neighborhood. Connectiveness must be jointly defined for the foreground (n) and background ($\overline{n}$) to

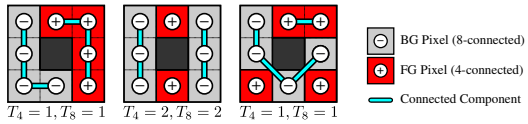

Fig. 2. Examples of topological numbers with $(n, \overline{n}) = (4, 8)$.

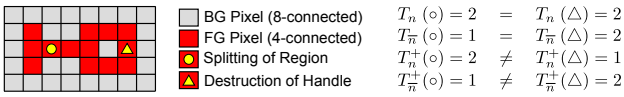

Fig. 3. Splitting a region vs. destroying a handle while a pixel is added to BG. Topologal numbers are shown on right.

avoid topological paradoxes. As shown in [12], valid connectivities for 2D are $(n, \overline{n}) \in \{(4, 8), (8, 4)\}$. Given a pair of connectiveness, the topological numbers [13] at a particular pixel, T_n (for the FG) and $T_{\overline{n}}$ (for the BG) count the number of connected components a pixel is connected to in a 3x3 neighborhood. Figure 2 shows a few neighborhoods with their corresponding topological numbers. While these topological numbers reflect topological changes, they do not distinguish splitting or merging of regions from the creation or destruction of a handle. Segonne [6] defines two additional topological numbers, T_n^+ and $T_{\overline{n}}^+$, which count the number of *unique* connected components a pixel is connected to over the entire image domain. T_n^+ and $T_{\overline{n}}^+$ depend on how pixels are connected outside of the 3x3 region and allow one to distinguish all topological changes.

By labeling each connected component in the foreground and background, Segonne shows that T_n^+ can be computed efficiently when a pixel is added to the foreground and $T_{\overline{n}}^+$ when a pixel is added to the background. In 3D, one *must* calculate $T_{\overline{n}}^+$ when adding a pixel to the foreground, which can be computationally expensive. We show here that this calculation is not necessary in 2D.

Consider the two pixels marked by $\circ$ and $\triangle$ in Figure 3. Removing the $\circ$ pixel from the foreground splits the region and removing the $\triangle$ pixel destroys a handle. We emphasize that in this 2D case, both $T_n^+(\circ) \neq T_n^+(\triangle)$ *and* $T_{\overline{n}}^+(\circ) \neq T_{\overline{n}}^+(\triangle)$, which generally is not true in 3D. In fact, in 2D, the destruction of a handle in the foreground corresponds directly to a merging of regions in the background. Likewise, the splitting of regions in the foreground corresponds directly to a creation of a handle in the background. Because of this one-to-one mapping, we do not need to compute the expensive $T_{\overline{n}}^+$ when adding a pixel to the foreground. Table 1 summarizes the topological changes of the foreground and background in 2D as a function of the four topological numbers.

				Add to FG		Add to BG	
T_n	T_n^+	$T_{\overline{n}}$	$T_{\overline{n}}^+$	FG	BG	FG	BG
0	0	1	1	CR	CH	DR	DH
1	1	0	0	DH	DR	CH	CR
1	1	1	1	-	-	-	-
≥ 2	$< T_n$	X	X	CH	SR	X	X
≥ 2	≥ 2	X	X	MR	DH	X	X
X	X	≥ 2	$< T_{\overline{n}}$	X	X	SR	CH
X	X	≥ 2	≥ 2	X	X	DH	MR

Table 1. Topological changes as a function of topological numbers. 'C' - Create; 'D' - Destroy; 'S' - Split; 'M' - Merge; 'H' - Handle(s); 'R' - Region(s); 'X' - any value; '-' - no topological change. Omitted configurations are impossible in 2D.

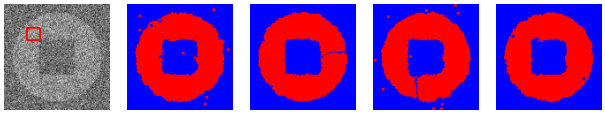

Fig. 4. Example samples imposing different topology constraints.

4.1. Topology-Controlled GIMH-SS

In this section, we summarize how to sample from the space of segmentations while enforcing topology constraints (*c.f.* [11] for details). While the goal is similar to [6], a simple alteration of the level-set velocity in an MCMC framework will *not* preserve detailed balance. A naïve approach generates proposals using GIMH-SS and rejects samples that violate topology constraints. Such an approach wastes significant computation generating samples that are rejected due to their topology. We take a different approach: only generate proposed samples from the set of allowable topological changes. Recalling the discussion of ranges in GIMH-SS, one can determine which ranges correspond to allowable topologies and which do not. Ranges corresponding to restricted topologies have their likelihood set to zero in a topology-controlled version of GIMH-SS (TC-GIMH-SS). An efficient implementation for finding allowable ranges can be designed by considering the masked pixels in ascending absolute height while maintaining the labeling of connected components. This methodology treats the topology as a hard constraint. A distribution over topologies is implemented by weighting ranges based on topologies rather than eliminating restricted ones.

5. RESULTS

The TC-GIMH-SS algorithm can be applied to any PDE-based or graph-based energy functional that depends only on the labeling and that can be evaluated. For deomnstration purposes, we make a specific choice on the energy functional. Similar to [10], we learn nonparametric probability densities over intensities and combine mutual information with a curve length penalty. We impose four different topology constraints on the foreground: unconstrained (GIMH-SS), topology-preserving (TP), genus-preserving (GP), and connected-component-preserving (CCP). The TP sampler does not allow any topology changes, the GP sampler only allows the splitting and merging of regions, and the CCP sampler only allows the creation and destruction of handles. Typical samples from each of these constraints are shown in Figure 4. When the topology constraint is incorrect for the object (e.g. using TP or GP), the resulting sample may be undesirable (e.g. creating a strait connecting the two connected components of the background). When the topology is correct, however, more robust results can be obtained. For example, the CCP constraint removes some noise in the background.

The usefulness of the topology constraint relies on a valid initialization. Histogram images [7] display the marginal probability that a pixel belongs to the foreground or background where lighter colors equate to higher foreground likelihood. The computed histogram images are shown in Figure 5 for each of the topology constraints. We initialize the samples either using a random circle containing the foreground (FG Init), or a random circle placed anywhere in the image (Random Init). While not always the case, marginal statistics are often robust to incorrect topology constraints. For example, in the FG initialization case, the strait in the TP and GP constraints is no longer visible. Additionally, if the initialization only captures one connected component of the background (which is possible with ran-

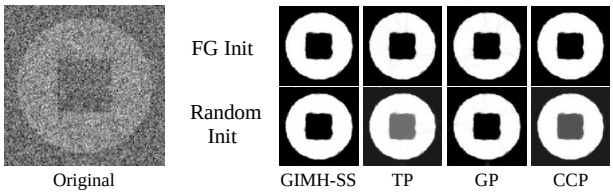

Fig. 5. Histogram images using different initializations.

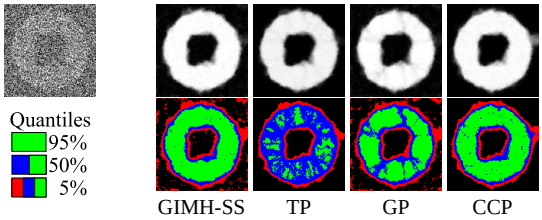

Fig. 6. Low SNR image using FG Init. Histograms are shown above, and quantiles (thresholded histograms) are shown below.

dom initialization), certain samplers that prohibit region splits (TP and CCP) will not be able to capture the entire region. This is reflected in the histogram image with the gray center.

Consider the low SNR image of Figure 6. The work in [1] demonstrated that sampling improves results when compared to optimization based methods in low SNR cases. When the problem is less likelihood dominated, as in this case, the prior has greater impact. The top row of Figure 6 shows the histogram images obtained using each of the topological constraints. One can see remnants of the straits in the TP and GP cases. Thresholding the normalized histogram image at a value t reveals the tth quantile of the segmentation. For example, if $t = 0.9$, the resulting foreground of the thresholded histogram contains pixels that are in the foreground for at least 90% of the samples. We show the 95%, 50%, and 5% quantile segmentations in Figure 6. Since reducing the threshold never shrinks the foreground segmentation, we can overlay these quantiles on top of each other. The 95% quantile shows the straits in the TP and GP cases which results from the wrong prior topology. However, using a correct CCP constraint improves results as compared to the unconstrained case by removing a lot of the background noise.

The CCP constraint is particularly useful when an unknown number of handles exist (e.g. deformable objects). Objects with a known number of handles in 3D projected onto a 2D plane can have any number of handles. We show two example images of a human and the resulting thresholded histogram image in Figure 7. In the first image, the handles formed by the arms are not captured well with TP and GP. In the second image, the vignetting causes the unconstrained and the GP to group some background with foreground.

6. CONCLUSION

We have presented a new MCMC sampler for implicit shapes. We have shown how to design a proposal such that every proposed sample is accepted. Unlike previous methods, GIMH-SS efficiently samples PDE-based and graph-based energy functionals and does not require the evaluation of the gradient of the energy functional (which may be hard to compute). Additionally, GIMH-SS was extended to include hard topological constraints by only proposing samples that are topologically consistent with some prior. Publicly available source code for the TC-GIMH-SS algorithm can be found at

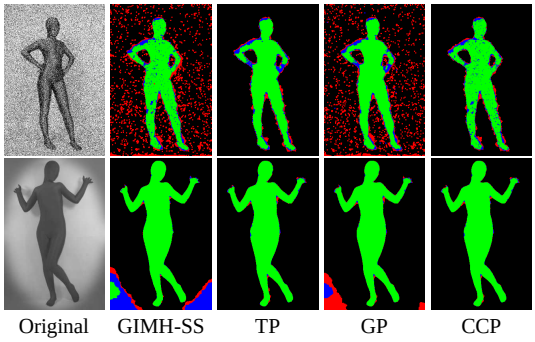

Fig. 7. Example images illustrating the utility of topology priors.

`http://people.csail.mit.edu/jchang7/`.

7. REFERENCES

[1] J. Chang and J. W. Fisher III, "Efficient MCMC Sampling with Implicit Shape Representations," in *IEEE CVPR*, 2011, pp. 2081–2088.

[2] D. M. Greig, B. T. Porteous, and A. H. Seheult, "Exact Maximum A Posteriori Estimation for Binary Images," *Jrnl. of the Royal Statistics Society*, vol. 51, pp. 271–279, 1989.

[3] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *IEEE Trans. PAMI*, vol. 23, pp. 1222–1239, November 2001.

[4] J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Trans. PAMI*, vol. 22, no. 8, pp. 888–905, 2000.

[5] X. Han, C. Xu, and J. L. Prince, "A topology preserving level set method for geometric deformable models," *IEEE Trans. PAMI*, vol. 25, no. 6, pp. 755–768, 2003.

[6] F. Ségonne, "Active contours under topology control–genus preserving level sets," *Int. J. Comput. Vision*, vol. 79, pp. 107–117, August 2008.

[7] A. C. Fan, J. W. Fisher III, W. M. Wells, J. J. Levitt, and A. S. Willsky, "MCMC curve sampling for image segmentation.," *MICCAI*, vol. 10, no. Pt 2, pp. 477–485, 2007.

[8] S Chen and R. J. Radke, "Markov chain Monte Carlo shape sampling using level sets," *IEEE ICCV Workshops*, pp. 296–303, 2009.

[9] W. K. Hastings, "Monte Carlo sampling methods using Markov chains and their applications," *Biometrika*, vol. 57, no. 1, pp. 97–109, 1970.

[10] J. Kim, J. W. Fisher III, A. Yezzi, M. Cetin, and A. S. Willsky, "A nonparametric statistical method for image segmentation using information theory and curve evolution.," *IEEE Trans. Image Proc.*, vol. 14, no. 10, pp. 1486–1502, 2005.

[11] J. Chang and J. W. Fisher III, "Efficient topology-controlled sampling of implicit shapes," May 2012, arXiv:1205.3766v1 [cs.CV].

[12] T. Y. Kong and A. Rosenfeld, "Digital topology: introduction and survey," *Comput. Vision Graph. Image Process.*, vol. 48, pp. 357–393, December 1989.

[13] G. Bertrand, "Simple points, topological numbers and geodesic neighborhoods in cubic grids," *Pattern Recogn. Lett.*, vol. 15, pp. 1003–1011, October 1994.