

MCMC Shape Sampling for Image Segmentation with Nonparametric Shape Priors

Ertunc Erdil, Sinan Yıldırım, Müjdat Çetin
Sabanci University,
Tuzla 34956, Istanbul

{ertuncerdil, sinanyildirim, mcetin}@sabanciuniv.edu

Tolga Tasdizen
Utah University
Salt Lake City, UT, USA
tolga@sci.utah.edu

Abstract

Segmenting images of low quality or with missing data is a challenging problem. Integrating statistical prior information about the shapes to be segmented can improve the segmentation results significantly. Most shape-based segmentation algorithms optimize an energy functional and find a point estimate for the object to be segmented. This does not provide a measure of the degree of confidence in that result, neither does it provide a picture of other probable solutions based on the data and the priors. With a statistical view, addressing these issues would involve the problem of characterizing the posterior densities of the shapes of the objects to be segmented. For such characterization, we propose a Markov chain Monte Carlo (MCMC) sampling-based image segmentation algorithm that uses statistical shape priors. In addition to better characterization of the statistical structure of the problem, such an approach would also have the potential to address issues with getting stuck at local optima, suffered by existing shape-based segmentation methods. Our approach is able to characterize the posterior probability density in the space of shapes through its samples, and to return multiple solutions, potentially from different modes of a multimodal probability density, which would be encountered, e.g., in segmenting objects from multiple shape classes. We present promising results on a variety of data sets. We also provide an extension for segmenting shapes of objects with parts that can go through independent shape variations. This extension involves the use of local shape priors on object parts and provides robustness to limitations in shape training data size.

1. Introduction

Prior knowledge about the shapes to be segmented is required for segmentation of images involving limited and low quality data. In many applications, object shapes come from multiple classes (i.e., the prior shape density is “mul-

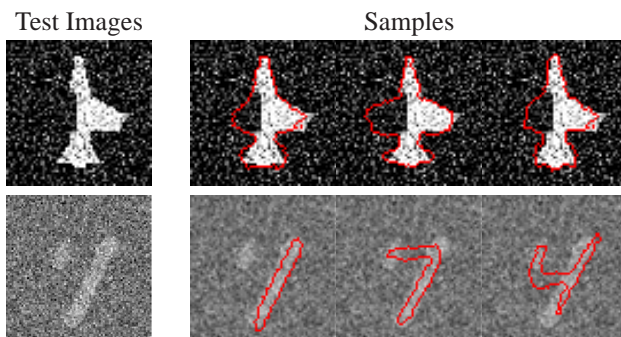


Figure 1. Examples of MCMC sampling. First row: on an object with a unimodal shape density. Second row: on an object with a multimodal shape density.

timodal”) and the algorithm does not know the class of the object in the scene. For example, in the problem of segmenting objects in a natural scene (e.g., cars, planes, trees, etc.), a segmentation algorithm should contain a training set of objects from different classes. Another example of a multimodal density is the shape density of multiple handwritten digits, e.g., in an optical character segmentation and recognition problem. In this paper, we consider segmentation problems that involve limited and challenging image data together with complex and potentially multimodal shape prior densities.

The shape-based segmentation approach of Tsai et al. [20] uses a parametric model for an implicit representation of the segmenting curve by applying principal component analysis to a training data set. Such techniques can handle only unimodal, Gaussian-like shape densities and cannot deal with “multimodal” shape densities. Kim et al. [12] and Cremers et al. [6] propose nonparametric methods for handling multimodal shape densities by extending Parzen density estimation to the space of shapes. These methods minimize an energy function containing both data fidelity and shape terms, and find a solution at a local optimum. Having such a point estimate does not provide any measure of the degree of confidence/uncertainty in that result

or any information about the characteristics of the posterior density, especially if the prior shape density is multimodal. Such a more detailed characterization might be beneficial for further higher-level inference goals such as object recognition and scene interpretation.

Our contributions in this paper are twofold. First, as the major contribution, we present a Markov chain Monte Carlo (MCMC) sampling approach that uses nonparametric shape priors for image segmentation. Our MCMC sampling approach is able to characterize the posterior shape density by returning multiple probable solutions and avoids the problem of getting stuck at a single local optimum. To the best of our knowledge, this is the first approach that performs MCMC shape sampling-based image segmentation through an energy functional that uses nonparametric shape priors and level sets. We present experimental results on several data sets containing low quality images and occluded objects involving both unimodal and multimodal shape densities. As a second contribution, we provide an extension within our MCMC framework, that involves a local shape prior approach for scenarios in which objects consist of parts that can exhibit independent shape variations. This extension allows learning shapes of object parts independently and then merging them together. This leads to more effective use of potentially limited training data. We demonstrate the effectiveness of this approach on a challenging segmentation problem as well.

Some exemplary results of our MCMC shape sampling approach that uses nonparametric shape priors are illustrated in Figure 1. The first row of Figure 1 shows three different samples obtained by our approach given the partially occluded aircraft test image in the first column of the corresponding row. In this experiment, the training set contains only examples of aircraft shapes, i.e., the shape density is unimodal, meaning that there are no well-defined subclasses. The second row of the figure contains an MCMC shape sampling example on handwritten digits. In this example, the training set consists of examples from ten different digit classes. Here, our approach is able to find multiple probable solutions from different modes of the shape density.

2. Related Work

Most of the sampling-based segmentation methods in the literature use an energy functional that include only a data fidelity term [2, 3, 9] which means that they are only capable of segmenting objects whose boundaries are mostly visible. Among these approaches, Fan et al. [9] have developed a method that utilizes both implicit (level set) and explicit (marker-based) representations of shape. The proposal distribution generates a candidate sample by randomly perturbing a set of marker points selected on the closed curve. Due to the use of marker points in perturbation, this ap-

proach is only applicable to segmentation of simply connected shapes, i.e., it cannot handle topological changes. Later, Chang et al. [2] have proposed an efficient MCMC sampling approach on a level set-based curve representation that can handle topological changes. Random curve perturbation is performed through an addition operator on the level set representation of the curve. Additive perturbation is generated by sampling from a Gaussian distribution. They also introduce some bias to the additive perturbation with the gradient of the energy function to achieve faster convergence. The method is further extended in [3] to achieve order of magnitude speed up in convergence by developing a sampler whose samples at every iteration are accepted. Additionally, they incorporate topological constraints to exploit prior knowledge of the shape topology.

Chen et al. [5] use the shape prior term suggested by Kim et al. [12] and Cremers et al. [6] together with a data fidelity term in the energy functional. Samples are generated by constructing a smooth normal perturbation at a single point on the curve which preserves the signed distance property of the level set. The method is restricted to segmentation of simply connected shapes due to its inability to handle topological changes. Therefore, the approach is not applicable to shapes with complex boundaries.

De Bruijne et al. [7] propose a particle filter-based segmentation approach that exploits both shape and appearance priors. The method assumes that the underlying shape distribution is unimodal. Therefore, it cannot handle cases when the shapes in the training set comes from a multimodal density.

Eslami et al. [8] propose a shape model that learns binary shape distributions using a type of deep Boltzmann machine [19] and generates samples using block-Gibbs sampling. The model is able to learn multimodal shape densities, however, samples generated from the distribution come only from a particular class closest to the observed data.

3. Metropolis-Hastings Sampling in the Space of Spaces

With a Bayesian perspective, segmentation can be viewed as the problem of estimating the boundary C based on image data:

$$p(C|data) \propto \exp(-E(C)) \quad (1)$$

where,

$$E(C) = E_{data}(C) + E_{shape}(C) = -\log p(data|C) - \log p_C(C) \quad (2)$$

In this paper, we present an algorithm to draw samples from $p(C|data)$ which is, in general, a complex distribution and is not possible to sample from directly.

MCMC methods were developed to draw samples from a probability distribution when direct sampling is non-trivial. We use Metropolis-Hastings sampling [16] which

has been previously used for image segmentation [9, 5, 2]. In Metropolis-Hastings sampling, instead of directly sampling from p , a proposal distribution q is defined and samples from q are accepted in such a way that samples from p are generated asymptotically. The Metropolis-Hastings acceptance probability is defined as

$$Pr[C^{(t+1)} = C^{(t+1)} | C^{(t)}] = \min \left[\underbrace{\frac{\pi(C^{(t+1)})}{\pi(C^{(t)})} \cdot \frac{q(C^{(t)} | C^{(t+1)})}{q(C^{(t+1)} | C^{(t)})}}_{\text{Metropolis-Hastings ratio}}, 1 \right]. \quad (3)$$

The Metropolis-Hastings threshold, η , is randomly generated from the uniform distribution in $[0, 1]$. The candidate (proposed) sample $C^{(t+1)}$ is accepted if $Pr[C^{(t+1)} = C^{(t+1)} | C^{(t)}]$ is greater than η . Otherwise, $C^{(t+1)} = C^{(t)}$. In Equation 3, $C^{(t)}$ and $C^{(t+1)}$ represent the current sample and proposed sample, respectively. The superscripts (t) and $(t+1)$ denote the sampling iteration count, and $\pi(C) \propto \exp(-E(C))$. After a sufficient number of iterations (i.e., the mixing time) a single sample from the posterior is produced by converging to the stationary distribution. Evaluating the acceptance probability is a key point in MCMC methods. Correct evaluation of the acceptance probability satisfies the sufficient conditions for convergence to the desired posterior distribution: detailed balance and ergodicity. Therefore, the problem turns into the correct computation of forward $q(C^{(t+1)} | C^{(t)})$ and reverse $q(C^{(t)} | C^{(t+1)})$ transition probabilities of the proposal distribution.

4. MCMC Shape Sampling using Nonparametric Shape Priors

We assume that the curve at the 0^{th} sampling iteration, $C^{(0)}$, is the curve that is found by minimizing only the data fidelity term, $E_{data}(C)$. We use piecewise-constant version of the Mumford-Shah functional [18, 1] for data driven segmentation. One can consider optimizing more sophisticated energy functions such as mutual information [13], J-Divergence [11], and Bhattacharya Distance [17] to obtain $C^{(0)}$. Also, using an MCMC sampling based approach for data driven segmentation can enrich the sampling space since it would allow subsequent MCMC shape sampling to use several initial curves to start from. After the curve finds all the portions of the object boundary identifiable based on the image data only (e.g., for a high SNR image with an occluded object, one would expect this stage to capture the non-occluded portions of the object reasonably well), we activate the process of generating samples from the underlying space of shapes using nonparametric shape priors.

The overall proposed MCMC shape sampling algorithm is given in Algorithm 1. The steps of the algorithm are explained in the following three subsections.

Algorithm 1 MCMC Shape Sampling

```

1: for  $i = 1 \rightarrow M$  do ▷  $M$  : # of samples to be generated
2:   Randomly select class of  $C^{(0)}$  as introduced in Section 4.1.
3:   for  $t = 0 \rightarrow (N - 1)$  do ▷  $N$  : # of sampling iterations
4:     Generate candidate sample  $\tilde{C}^{(t+1)}$  from curve  $\tilde{C}^{(t)}$  as introduced in Section 4.2.
▷ The steps between 5 - 10 are introduced in Section 4.3
5:     Calculate Metropolis-Hastings ratio,  $Pr$ 
6:      $\eta = \mathcal{U}_{[0,1]}$ 
7:     if  $(t + 1) = 1$  OR  $\eta < Pr$  then
8:        $\tilde{C}^{(t+1)} = \tilde{C}^{(t+1)}$  ▷ Accept the candidate
9:     else
10:       $\tilde{C}^{(t+1)} = \tilde{C}^{(t)}$  ▷ Reject the candidate
11:    end if
12:  end for
13: end for

```

4.1. Random class decision

Suppose that we have a training set $\mathbf{C} = \{C_1, \dots, C_n\}$ consisting of shapes from n different classes where each class $C_i = \{C_{ij} | j \in [1, m_i] \in \mathbb{Z}\}$ contains m_i different example shapes. We align training shapes C_{ij} into \tilde{C}_{ij} using the alignment approach presented in Tsai et al. [20] in order to remove the artifacts due to pose differences such as translation, rotation, and scaling.

We exploit the shape prior term $p_C(C)$ proposed by Kim et al. [12] to select the class of the curve $\tilde{C}^{(0)}$. The prior probability density function of the curve evaluated at sampling iteration zero is

$$p_C(\tilde{C}^{(0)}) = \frac{1}{n} \sum_{i=1}^n \frac{1}{m_i} \sum_{j=1}^{m_i} k(d_{L_2}(\phi_{\tilde{C}^{(0)}}, \phi_{\tilde{C}_{ij}}), \sigma) \quad (4)$$

where $k(\cdot, \sigma)$ is a 1D Gaussian kernel with kernel size σ , $d_{L_2}(\cdot, \cdot)$ is the L_2 distance metric and ϕ denotes the level set representation of a curve. Also, note that $\tilde{C}^{(0)}$ is the aligned version of $C^{(0)}$ with the training set. By exploiting Equation 4, we can compute the prior probability density of the shapes in C_i evaluated at $\tilde{C}^{(0)}$, $p'_{C_i}(\tilde{C}^{(0)})$, as follows

$$p'_{C_i}(\tilde{C}^{(0)}) \propto \frac{1}{m_i} \sum_{j=1}^{m_i} k(d_{L_2}(\phi_{\tilde{C}^{(0)}}, \phi_{\tilde{C}_{ij}}), \sigma). \quad (5)$$

We randomly select a class for shape $\tilde{C}^{(0)}$ where the probability of selecting a class is proportional to the value of $p'_{C_i}(\tilde{C}^{(0)})$ computed in Equation 5. When we generate multiple samples, the random class selection step helps us obtain more samples from the classes having higher probabilities.

4.2. Generating a candidate sample

In this section, we explain how to generate a candidate sample from the proposal distribution q . Once the class of $\tilde{C}^{(0)}$ is randomly selected, we perform curve perturbation exploiting the training samples in this class. Let \tilde{C}_r be the set that contains the training shapes from the selected class r . We randomly choose γ training shapes from \tilde{C}_r where the probability of selecting each shape is proportional to its similarity with $\tilde{C}^{(t)}$. We compute the similarity between a training shape \tilde{C}_{rj} and $\tilde{C}^{(t)}$ as the value of the probability density function, s , at \tilde{C}_{rj} where,

$$s_{\tilde{C}^{(t)}}(\tilde{C}_{rj}) \propto k(d_{L_2}(\phi_{\tilde{C}^{(t)}}, \phi_{\tilde{C}_{rj}}), \sigma). \quad (6)$$

Note that a training shape can be selected multiple times and random training shape selection is repeated in each sampling iteration. We represent the set composed of randomly selected γ training shapes at sampling iteration t by $\tilde{\mathbf{C}}_{\mathbf{R}}^{(t)}$.

Finally, we define the forward perturbation for the curve $\tilde{C}^{(t)}$ with level sets as follows:

$$\phi_{\tilde{C}^{(t+1)}} = \phi_{\tilde{C}^{(t)}} + \alpha \mathbf{f}^{(t)} \quad (7)$$

We choose $\mathbf{f}^{(t)}$ as the negative gradient of the energy function given in Equation 2 in order to move towards a more probable configuration in each perturbation. Here, α indicates the step size for gradient descent. Note that we use randomly selected training samples, $\tilde{C}_{Rj} \in \tilde{\mathbf{C}}_{\mathbf{R}}^{(t)}$, for curve perturbation. Mathematically this is expressed as

$$\begin{aligned} \mathbf{f}^{(t)} = & -\frac{\partial E(\phi_{\tilde{C}^{(t)}})}{\phi_{\tilde{C}^{(t)}}} = \frac{\partial \log p(\text{data}|\tilde{C}^{(t)})}{\partial t} \\ & + \frac{1}{p_{\tilde{C}^{(t)}}(\tilde{C}^{(t)})} \frac{1}{\gamma} \sum_{j=1}^{\gamma} k(d_{L_2}(\phi_{\tilde{C}^{(t)}}, \phi_{\tilde{C}_{Rj}}), \sigma)(\phi_{\tilde{C}_{Rj}} - \phi_{\tilde{C}^{(t)}}) \end{aligned} \quad (8)$$

In other words, updating the curve $\tilde{C}^{(t)}$ toward the negative gradient direction of the energy functional produces the candidate curve $\tilde{C}^{(t+1)}$.

4.3. Evaluating the Metropolis-Hastings ratio

Computation of the first fraction in the Metropolis-Hastings ratio in Equation 3 is straightforward since $\pi(C') \propto \exp(-E(C'))$. Recall that the candidate curve $\tilde{C}^{(t+1)}$ is dependent on the forward perturbation $\mathbf{f}^{(t)}$. Therefore, we compute the forward perturbation probability by considering the value of the probability density function, s , for each randomly selected training shape $\tilde{C}_{Rj} \in \tilde{\mathbf{C}}_{\mathbf{R}}^{(t)}$ as follows:

$$q(\tilde{C}^{(t+1)}|C^{(t)}) = \prod_{\tilde{C}_{Rj} \in \tilde{\mathbf{C}}_{\mathbf{R}}^{(t)}} s(\tilde{C}_{Rj}) \quad (9)$$

Similarly, the reverse perturbation probability in sampling iteration $(t+1)$ is computed as the probability of selecting random shapes in $\tilde{\mathbf{C}}_{\mathbf{R}}^{(t-1)}$ which have been used to

produce the curve $\tilde{C}^{(t)}$:

$$q(\tilde{C}^{(t)}|\tilde{C}^{(t+1)}) = \prod_{\tilde{C}_{Rj} \in \tilde{\mathbf{C}}_{\mathbf{R}}^{(t-1)}} s(\tilde{C}_{Rj}) \quad (10)$$

Note that, given the above formulations, computation of the reverse perturbation probability is not possible for candidate curve $\tilde{C}^{(1)}$, the curve at sampling iteration 1, since we have to use information from sampling iteration -1 for evaluation of Equation 10, which is not available. Therefore, we accept the candidate sample $\tilde{C}^{(1)}$ without evaluating the Metropolis-Hastings ratio and consider the above-mentioned steps for generating samples after sampling iteration 1.

4.4. Discussion on sufficient conditions for MCMC sampling

Convergence to the correct stationary distribution is crucial in MCMC methods. Convergence is guaranteed with two sufficient conditions: (1) that the chain is ergodic, and (2) that detailed balance is satisfied in each sampling iteration. Ergodicity is satisfied when the Markov chain is aperiodic and irreducible. Aperiodicity of a complicated Markov chain is a property that is hard to prove as attested in the literature [10].

Detailed balance is satisfied as long as the Metropolis-Hastings ratio in Equation 3 is calculated correctly. We have already described how we compute the Metropolis-Hastings ratio in the previous section. Empirical results show that a stationary distribution is most likely reached since our samples converge. Related pieces of work in [9], [2], and [5] argue that the Markov chain is unlikely to be periodic because the space of segmentations is so large. Similarly, we can also assert that our Markov chain is unlikely to be periodic. Even if the chain is periodic in exceptional cases, the average sample path converges to the stationary distribution as long as the chain is irreducible. Irreducibility of a Markov chain requires showing that transitioning from any state to any other state has finite probability. Chen et al. [5] and Chang et al. [2] provide valid arguments that the Markov chain is irreducible whereas Fan et al. [9] does not discuss this property. As explained in the previous section, curve perturbation in our framework is performed with randomly selected training samples $\tilde{\mathbf{C}}_{\mathbf{R}}^{(t)}$ and each shape has finite probability to be selected at any sampling iteration. With this perspective, we can also argue that each move between shapes has finite probability in our approach.

5. Extension to MCMC Sampling using Local Shape Priors

In this section, we consider the problem of segmenting objects with parts that can go through independent shape variations. We propose to use local shape priors on object

parts to provide robustness to limitations in shape training size [4, 15]. Let us consider the motivating example shown in Figure 2. In this example, there are three images of walking silhouettes: two for training and one for testing. Note that the left leg together with the right arm of the test silhouette involves missing regions. When segmenting the test image using nonparametric shape priors [12] based on global training shapes¹, the result may not be satisfactory (see the rightmost image in the first row of Figure 2), because the shapes in the training set do not closely resemble the test image. This motivates us to represent shapes with local priors such that resulting segmentation will mix and match information from independent object parts (e.g., by taking information about the right arm from the first training shape and about the left leg from the second training shape).

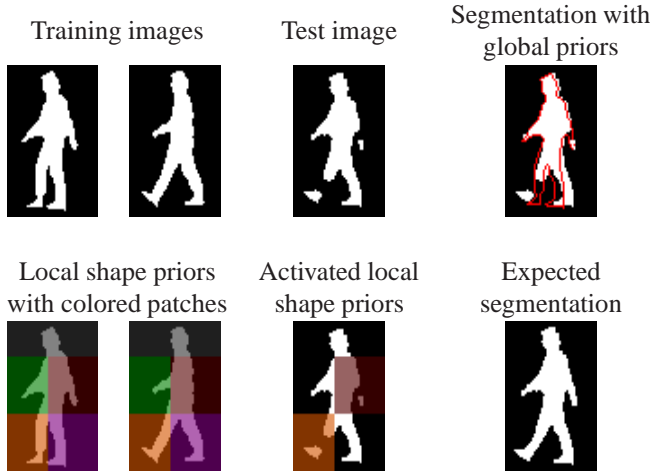


Figure 2. Motivating example for using local shape priors in walking silhouettes data set.

Our idea of constructing local shape priors is straightforward. Once the training shapes are aligned, we divide the shapes into patches, such that each patch contains a different local shape region. Each patch is indicated by a different color in the second row of Figure 2. Note that the patches representing the same local shape have identical size. For MCMC shape sampling using local shape priors, it is straightforward to adapt the formulation in the previous sections to consider local priors. In particular, instead of choosing random global shapes using the values computed by Equation 6, we compute these values for each patch (local shape) and select random patches among all training images. Note that evaluation of forward and reverse perturbation probabilities should also be modified accordingly.

¹Unless otherwise stated, the shape priors we use are global. We explicitly refer to global shape priors when we need to distinguish them from local shape priors.

6. Experimental Results

In this section, we present empirical results of our MCMC shape sampling algorithm on segmentation of potentially occluded objects in low-quality images. Note that, when dealing with segmentation of objects with unknown occlusions, $E_{data}(C)$ increases when the shape term delineates the boundaries in the occluded region. This can lead to overall increasing effect on $E(C^{(t)})$ for a candidate curve and to the rejection of the candidate sample. In order to increase the acceptance rate of our approach, we use $\pi(C) \propto \exp(-E_{shape}(C))$ instead of $\pi(C) \propto \exp(-E(C))$ in our experiments involving occluded objects (see supplementary material for experiments involving missing data in which we use $\pi(C) \propto \exp(-E(C))$). This does not cause any problem in practice since the data fidelity term (together with the shape prior term) is involved in the curve perturbation step, enforcing consistency with the data.

We perform experiments on several data sets: aircraft [12], MNIST handwritten digits [14], and walking silhouettes [6]. In the following subsections, we present quantitative and visual results together with discussions of the experiments for each data set.

6.1. Experiments on the aircraft data set

The aircraft data set [12] contains 11 synthetically generated binary aircraft images as shown in the top row of Figure 3. We construct the test images shown in the middle and the bottom rows of the same figure by cropping the left wings from the binary images to simulate occlusion and by adding different amounts of noise. Note that the test images shown in the middle row of Figure 3 (test image set - 1) have higher SNR than the ones shown in the bottom row (test image set - 2). In our experiments, we use this data set in leave-one-out fashion, i.e., we use one image as test and the remaining 10 binary images for training.

In Figure 4(a), we present some visual and quantitative results on the first three images from the test image set - 1 shown in Figure 3. In this experiment, we generate 500 samples using our shape sampling approach for each test image. We also obtain segmentations using the optimization-based segmentation approach of Kim et al. [12] (see the second column of Figure 4(a)). We compare each sample and the result of Kim et al. [12] with the corresponding ground truth image using precision - recall values and the F-measure. The samples with the best F-measure value are shown in the third column of Figure 4(a). Finally, we plot the precision - recall values (PR plots) for each sample and for the result of Kim et al. [12] in the fourth column of Figure 4(a). Here, the data fidelity term keeps the curve at the object boundaries and shape prior term helps to complete the shape in the occluded part. In our approach, since we select the most probable subset of training images and evolve the curve with the weighted average of these im-

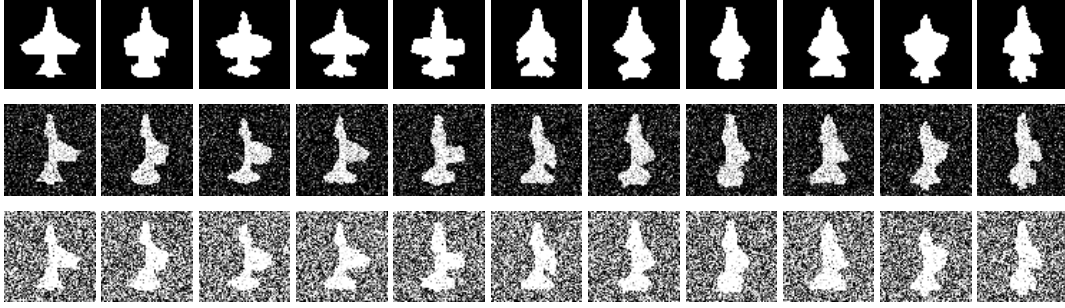


Figure 3. The aircraft data set. Top row: Training set, middle row: test image set - 1 and bottom row: test image set - 2. Note that we remove the corresponding training image from the training set for each test image in our experiments.

ages, the results of our approach are more likely to produce better fits for the occluded part. In the experiments shown in Figure 4(a), our approach can generate better samples than the result of Kim et al. [12] in all test images. Moreover, our algorithm is able to generate many different samples in the solution space. By looking at these samples, one can also have more information about the confidence in a particular solution.

We also perform experiments on the aircraft test image set - 2 shown in Figure 3 and present results on the first three images in Figure 4(b). The segmentation problem in this image set is more challenging than the previous case because of lower SNR. We perform experiments with the same settings as in test image set - 1 and present the results in the same way in Figure 4(b). In this case, we have to give more weight to the shape prior term during evolution to complete the occluded part because of the high amount of noise. Because of the limited role of the data fidelity term, the curve starts losing some part of the boundary after the shape term is turned on since the role of the data term is limited. Therefore, in this case, not only the occluded part but also the other parts of the aircraft shape approach a weighted average of the objects in the training set during curve evolution. Note from Figure 4(b) that the results of Kim et al. [12] on different test images are very similar to one another. However, our sampling approach produces more diverse samples including better ones than the result of Kim et al. [12] in terms of F-measure in most cases. Additional results on all remaining test images shown in Figure 3 can be found in the supplementary material.

6.2. Experiments on the MNIST data set

In this section, we present empirical results on the MNIST handwritten digits [14] data set which includes a multimodal shape density (i.e. training set contains shapes from multiple classes corresponding to different modes of the shape density). The MNIST handwritten digits data set contains 60,000 training examples and 10,000 test images from 10 different digit classes. In our experiments, we take a subset of 100 images for training such that each class con-

tains 10 training examples. Test images, none of which are contained in the training set, are obtained by cropping some parts of the digits and adding noise. The test images that we use in our experiments are shown in Figure 5.

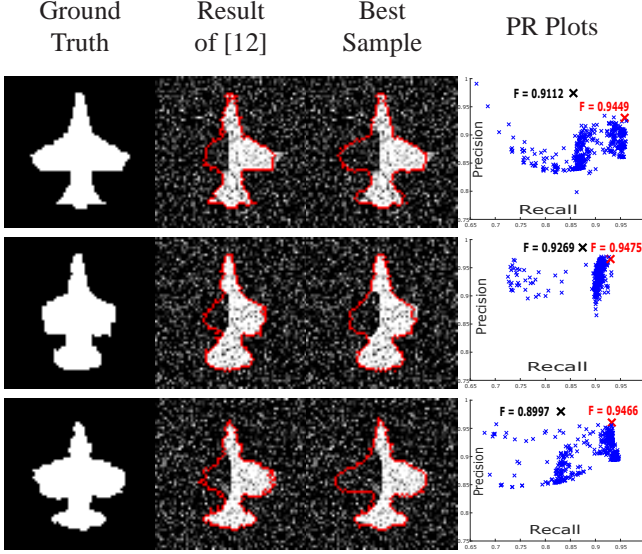
In our experiments on the MNIST data set, we generate 1000 samples using our shape sampling approach. In order to interpret our results, we use three methodologies: (1) Compute the average energy for each class by considering the samples generated in that class. Choose the best three classes with respect to average energy values. Display the best three samples from each class in terms of energy. These samples are most likely good representatives of the modes of the target distribution, (2) Compute the histogram images $H(x)$ which indicate in what percentage of the samples a particular pixel is inside the boundary. This can be simply computed by summing up all the binary samples and dividing by the number of samples [9]. $H(x)$ can be computed for each class for problems involving multimodal shape densities. We draw the marginal confidence bounds, the bounds where $H(x) = 0.1$ and $H(x) = 0.9$, over the test image for each class, (3) Count the number of samples obtained from each class. This can allow a probabilistic interpretation of the results.

Figure 6 demonstrates the average shape energy for each class, $E_{shape}(C)$, as a function of sampling iterations for test image MNIST - 1. We note that while the average energy appears to be smoothly converging, the energy for each sample path can sharply increase and decrease. The plot of class 9 in Figure 6 exhibits such a pattern because there is only one sample generated from this class. As the number of samples generated in each class increases, the average sample path converges to a stationary distribution.

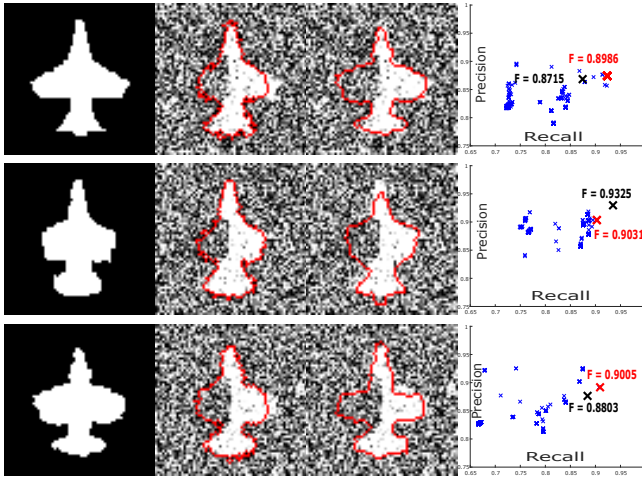
Test Image	Digit Class									
	0	1	2	3	4	5	6	7	8	9
MNIST - 1	336	433	6	18	29	38	115	16	8	1
MNIST - 2	4	691	8	3	96	9	0	120	3	66
MNIST - 3	119	661	8	1	2	11	154	14	28	2

Table 1. Number of samples generated for each digit class in test images from the MNIST data set.

Number of samples generated from each digit class for



(a) Test image set - 1



(b) Test image set - 2

Figure 4. Experiments on aircraft data set. Note that each row contains the results for a different test image. In the PR plots, ‘ \times ’ and ‘ \times ’ mark the samples produced by our approach where ‘ \times ’ indicates the sample with the best F-measure value, and ‘ \times ’ marks that of segmentation of Kim et al. [12].

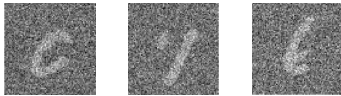


Figure 5. Test images from the MNIST data set. From left to right: MNIST - 1, MNIST - 2, and MNIST - 3.

all the three test images is shown in Table 1. This allows us to make a probabilistic interpretation of the segmentation results. One can evaluate the confidence level of the results by analyzing the number of samples generated from a class over all samples.

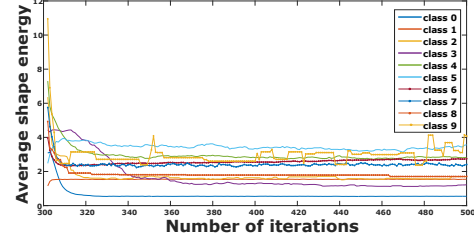


Figure 6. Average shape energy ($E_{shape}(C)$) across all sampling iterations for all digit classes for test image MNIST - 1. Note that the number of iterations start from 300 in x -axis because the previous iterations involve segmentation with the data term only.

In different segmentation applications, one can investigate solutions obtained from different parts of the posterior probability density. Especially, in the case of multimodal shape densities, segmentation results obtained from multiple modes might be interesting and might offer reasonable solutions. Figure 7 shows some visual results obtained from the experiments on the MNIST data set. For each test image, we display the results from the best three digit classes where, the quality of each class is computed as the average energy, $E(C)$, of the samples in that class. Also, for each class, we show three samples having the best energy values. These results show that our algorithm is able to find reasonable solutions from different modes of the posterior density. In Figure 7, we also present marginal confidence bounds (MCB images) obtained from the samples in each class. The figure demonstrates the marginal confidence bounds at different levels of the histogram image, $H(x)$, for the best classes in all test images. $H(x) = 0.1$ and $H(x) = 0.9$ indicate the low probability and the high probability regions, respectively.

6.3. Experiments on the walking silhouettes data set

In this experiment, we test the performance of local shape priors extension of our MCMC shape sampling approach and compare it with the one that uses global shape priors, as well as with the method of Kim et al. [12]. We choose a subset of 30 binary images of a walking person from the walking silhouettes data set [6]. A subset of 16 images shown in Figure 8 among these 30 binary images are used for training. The remaining 14 binary images are used to construct test images by adding a high amount of noise.

For the sake of brevity, we present results on 3 test images in Figure 9. Additional results can be found in the supplementary material. Similar to the evaluations performed for the aircraft data set, we plot the PR values for each sample obtained by our approaches (with global and local priors) and by the approach of Kim et al. [12]. According to the results, our proposed approach with global shape priors

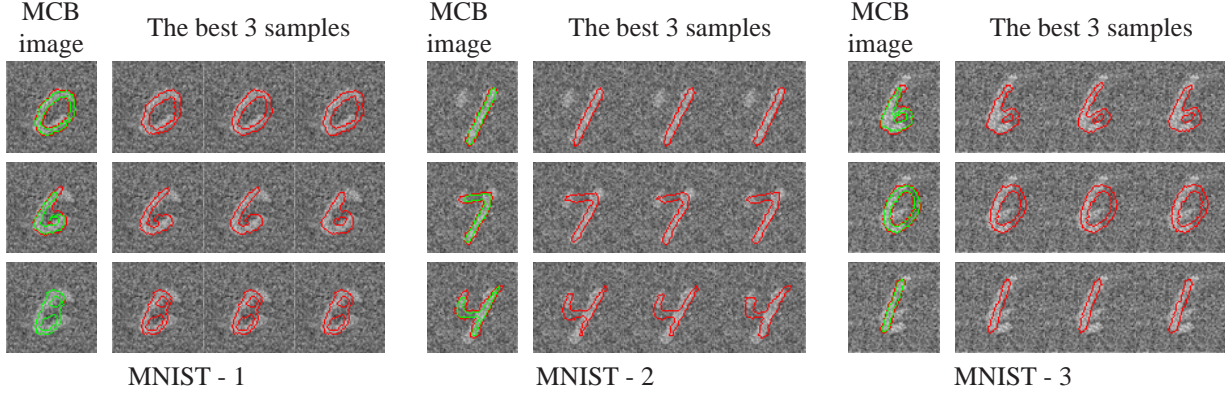


Figure 7. Experiments on the MNIST data set. Note that in MCB images, red and green contours are the marginal confidence bounds at $H(x) = 0.1$ and $H(x) = 0.9$, respectively.

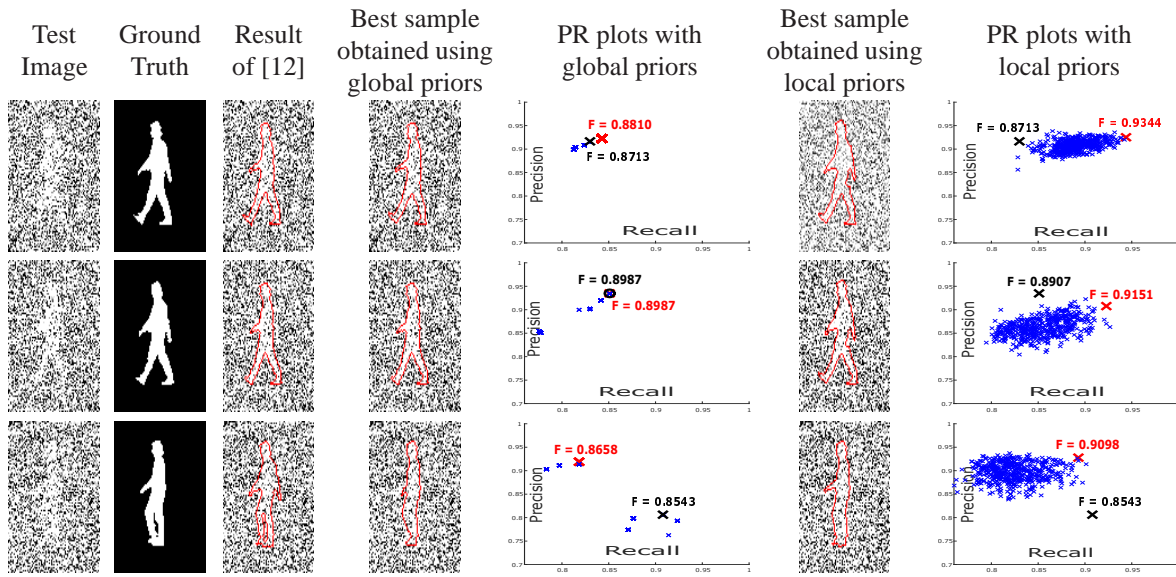


Figure 9. Experiments on walking silhouettes data set. In the PR curves, the 'x' marks the sample having the best F-measure value obtained using the proposed approach (with either global or local shape priors), and the 'x' marks that of segmentation of Kim et al. [12].

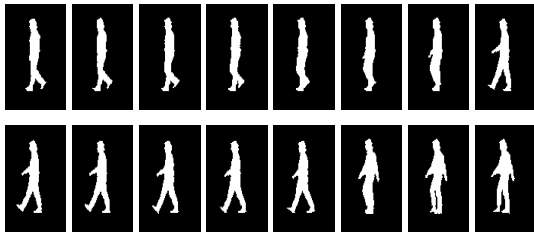


Figure 8. The training set for the walking silhouettes data set.

produces samples that have F-measure values better than or equal to the result of Kim et al. [12] in all test images. By using local shape priors, we can generate even better samples than both Kim et al. [12] and the approach with global shape priors. Moreover, it seems that our approach based on local shape priors is able to sample the space more effectively than the approach with global shape priors.

7. Conclusion

We have presented a MCMC shape sampling approach for image segmentation that exploits prior information about the shape to be segmented. Unlike existing MCMC sampling methods for image segmentation, our approach can segment objects with occlusion and suffering from severe noise, using nonparametric shape priors. We also provide an extension of our method for segmenting shapes of objects with parts that can go through independent shape variations by using local shape priors on object parts. Empirical results on various data sets demonstrate the potential of our approach in MCMC shape sampling. The implementation of the proposed method is available at spis.sabanciuniv.edu/data_code.

References

- [1] T. F. Chan, L. Vese, et al. Active contours without edges. *IEEE transactions on Image processing*, 10(2):266–277, 2001.
- [2] J. Chang and J. Fisher. Efficient mcmc sampling with implicit shape representations. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2081–2088. IEEE, 2011.
- [3] J. Chang and J. W. Fisher III. Efficient topology-controlled sampling of implicit shapes. In *Proceedings of the IEEE International Conference on Image Processing (ICIP)*, Sept 2012.
- [4] F. Chen, H. Yu, R. Hu, and X. Zeng. Deep learning shape priors for object segmentation. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1870–1877. IEEE, 2013.
- [5] S. Chen and R. J. Radke. Markov chain monte carlo shape sampling using level sets. In *IEEE 12th International Conference on Computer Vision Workshops (ICCV Workshops)*, pages 296–303. IEEE, 2009.
- [6] D. Cremers, S. J. Osher, and S. Soatto. Kernel density estimation and intrinsic alignment for shape priors in level set segmentation. *International Journal of Computer Vision*, 69(3):335–351, 2006.
- [7] M. De Bruijne and M. Nielsen. Image segmentation by shape particle filtering. In *Proceedings of the 17th International Conference on Pattern Recognition (ICPR)*, volume 3, pages 722–725. IEEE, 2004.
- [8] S. A. Eslami, N. Heess, C. K. Williams, and J. Winn. The shape boltzmann machine: a strong model of object shape. *International Journal of Computer Vision*, 107(2):155–176, 2014.
- [9] A. C. Fan, J. W. Fisher III, W. M. Wells III, J. J. Levitt, and A. S. Willsky. Mcmc curve sampling for image segmentation. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pages 477–485. Springer, 2007.
- [10] W. R. Gilks, S. Richardson, and D. J. Spiegelhalter. (1996). markov chain monte carlo in practice.
- [11] N. Houhou, J.-P. Thiran, and X. Bresson. Fast texture segmentation model based on the shape operator and active contour. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1–8. IEEE, 2008.
- [12] J. Kim, M. Çetin, and A. S. Willsky. Nonparametric shape priors for active contour-based image segmentation. *Signal Processing*, 87(12):3021–3044, 2007.
- [13] J. Kim, J. W. Fisher III, A. Yezzi, M. Çetin, and A. S. Willsky. A nonparametric statistical method for image segmentation using information theory and curve evolution. *IEEE Transactions on Image Processing*, 14(10):1486–1502, 2005.
- [14] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.
- [15] F. Mesadi, M. Cetin, and T. Tasdizen. Disjunctive normal shape and appearance priors with applications to image segmentation. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pages 703–710. Springer, 2015.
- [16] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. Equation of state calculations by fast computing machines. *The journal of chemical physics*, 21(6):1087–1092, 1953.
- [17] O. Michailovich, Y. Rathi, and A. Tannenbaum. Image segmentation using active contours driven by the bhat-tacharyya gradient flow. *IEEE Transactions on Image Processing*, 16(11):2787–2801, 2007.
- [18] D. Mumford and J. Shah. Optimal approximations by piecewise smooth functions and associated variational problems. *Communications on pure and applied mathematics*, 42(5):577–685, 1989.
- [19] R. Salakhutdinov and G. E. Hinton. Deep boltzmann machines. In *International conference on artificial intelligence and statistics*, pages 448–455, 2009.
- [20] A. Tsai, A. Yezzi Jr, W. Wells, C. Tempny, D. Tucker, A. Fan, W. E. Grimson, and A. Willsky. A shape-based approach to the segmentation of medical imagery using level sets. *IEEE Transactions on Medical Imaging*, 22(2):137–154, 2003.