Parsing Images into Region and Curve Processes

Zhuowen Tu and Song-Chun Zhu

Department of Computer and Information Science
The Ohio State University
{ztu,szhu}@cis.ohio-state.edu

Abstract. Natural scenes consist of a wide variety of stochastic patterns. While many patterns are represented well by statistical models in two dimensional regions as most image segmentation work assume, some other patterns are fundamentally one dimensional and thus cause major problems in segmentation. We call the former *region processes* and the latter *curve processes*. In this paper, we propose a stochastic algorithm for parsing an image into a number of region and curve processes. The paper makes the following contributions to the literature. Firstly, it presents a generative *rope* model for curve processes in the form of Hidden Markov Model (HMM). The hidden layer is a Markov chain with each element being an image base selected from an over-complete basis, such as Difference of Gaussians (DOG) or Difference of Offset Gaussians (DOOG) at various scales and orientations. The rope model accounts for the geometric smoothness and photometric coherence of the curve processes. Secondly, it integrates both 2D region models, such as textures, splines etc with 1D curve models under the Bayes framework. Because both region and curve models are generative, they compete to explain input images in a layered representation. Thirdly, it achieves global optimization by effective Markov chain Monte Carlo methods in the sense of maximizing a posterior probability. The Markov chain consists of reversible jumps and diffusions driven by bottom up information. The algorithm is applied to real images with satisfactory results. We verify the results through random synthesis and compare them against segmentations with region processes only.

1 Introduction

Natural images consist of a wide variety of stochastic visual patterns. As an example, Figure 1 illustrates how an image is decomposed into point, line, curve processes, regions of coherent color and textures, and objects. Parsing an image into its constituent components is a fundamental problem in image understanding. It augments the current pixel-based image representation to a semantic object-based description, and thus has broad impacts on many applications, including image compression, photo editing, image database retrieval, object recognition in vision, and non-photo-realistic rendering (NPR) in graphics. Solving the image parsing problem needs 1). a number of types of probabilistic models which can characterize various visual patterns and are compatible (or comparable) with each other, and 2). inference algorithms which can effectively handle
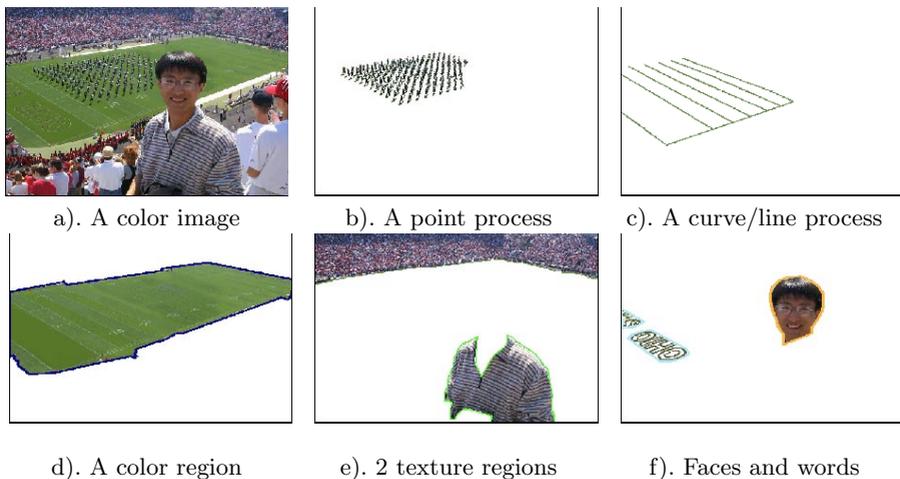
a). A color image   b). A point process   c). A curve/line process

d). A color region   e). 2 texture regions   f). Faces and words

Fig. 1. An example of image parsing and decomposition.

different types of probabilistic models and achieve globally optimal solution. These are general topics of this paper.

In the literature, image segmentation work usually assume that images contain regions of homogeneous properties, such as textures, textons (attributed points), colors, and shading. While many visual patterns are represented well by such 2D models, there are many other visual patterns which are fundamentally one dimensional and thus cause major problems in segmentation. For example, see those lines, trees, grids in Figures 6,8, and 10. We call the 2D patterns *region processes* and 1D patterns *curve processes*.

In this paper, we propose a stochastic algorithm for parsing an image into a number of region and curve processes. The paper makes contributions in the following aspects.

Firstly, it presents a generative rope *model for curve patterns.* The rope model is in the form of hidden Markov model. The hidden layer is a chain of connected *knots*. Each knot has 1-3 image bases selected from an over-complete basis[9], such as DOG (difference of Gaussians) or DOOG (difference of offset Gaussians) at various scales and orientations. The Markov model accounts for not only the geometric smoothness as in the SNAKE[10] or Elastica models[8], but also the photometric coherence along the curve.

Secondly, it integrates both 2D region and 1D curve models for image parsing. We adopt a generative image model which represents an image $\mathbf{I}$ as a superposition of two layers $\mathbf{I} = \mathbf{I}^r + \mathbf{I}^c$. $\mathbf{I}^r$ is partitioned into disjoint regions and $\mathbf{I}^c$ is a linear sum of some image bases which are grouped into curves. The algorithm consists of two parts. Part I is image segmentation on $\mathbf{I}^r$, and this is done by a recent data driven Markov chain Monte Carlo (DDMCMC) algorithm[11]. Part II infers the curves patterns from $\mathbf{I}^c$, and is the main focus of this paper.

By virtue of the generative image models, the cooperation of the two parts is governed in the posterior probability.

   Thirdly, it achieves global optimization by effective Markov chain Monte Carlo. Due to the use of different families of models, the solution space consists of many subspaces of varying dimensions. Therefore the Markov chain consists of reversible jumps[4,5] and stochastic diffusions to explore the complex solution space. The jump dynamics realize the death/birth, split and merge of regions and curves, the switching of models, and so on, while the diffusion process realizes region growing/competition[12], and curve deformation[10].

   The algorithm is applied to a set real images and some results are shown in Figures 8 and 10 . We verify the results through random synthesis from the computed solution and compare the results against segmentations with region process only.

   We organize this paper as follows. We first formulate the problem in Section 2. The Section 3 briefly overviews the region models used by region processes. Section 4 discusses the rope model for curve patterns. Section 5 discusses the structure of the solution space. Then we present the integrated algorithm in Section 6. Some experiments are shown in section 7. We conclude the paper with a discussion in section 8.

2   Problem Formulation in Bayes Statistics

Let $\Lambda = \{(i,j) : 1 \leq i \leq L, 1 \leq j \leq H\}$ be an image lattice, and $\mathbf{I}$ be an intensity image defined on Λ. We assume that the image $\mathbf{I}$ is a superposition of two layers,

$$\mathbf{I} = \mathbf{I}^r + \mathbf{I}^c, \tag{1}$$

where $\mathbf{I}^r$ and $\mathbf{I}^c$ are called *region layer* and *curve layer* respectively.

   We assume that the region layer $\mathbf{I}^r$ consists of a number of K^r disjoint regions which form a partition of the lattice Λ.

$$\cup_{i=1}^{K^r} R_i = \Lambda, \quad R_i \cap R_j = \emptyset, \quad \forall i \neq j.$$

Let $\mathbf{I}_R^r$ denote the image intensities in a region R, and a region is said to be coherent in the sense that $\mathbf{I}_R^r$ is a realization from a probabilistic model $p_\ell(\mathbf{I}_R; \Theta)$. ℓ indexes the model or a stochastic process, and Θ is a vector valued parameter of the model. We should discuss the families of region models shortly. Therefore the region processes are represented by a vector W^r of unknown dimension.

$$W^r = (K^r, \{(R_i, \ell_i, \Theta_i); i = 1, 2, ..., K^r\}). \tag{2}$$

   We assume that the curve layer $\mathbf{I}^c$ is a linear sum of a number of N image bases, following the literature of image coding[7,9].

$$\mathbf{I}^c = \sum_{j=1}^{N} \alpha_j B_j, \qquad B_j \in \Delta, \tag{3}$$

α_j is the coefficient and base B_j is selected from a over-complete basis or dictionary Δ, for example, Δ includes Difference of Gaussains (DoG) and Difference of Offset Gaussians (DOOG) over a group of transforms (scaling, rotating, and translation). These image bases are grouped into a number of $K^c \leq N$ curves $C_i, i = 1, ..., K^c$ based on a probabilistic curve model that we should deliberate in section (4.2). C_i is a list of bases with certain geometric and photometric regularities. Thus the curve processes are denoted by a vector W^c and $\mathbf{I}^c$ is a deterministic function of W^c,

$$W^c = (K^c, \{C_i; \;\; i = 1, 2, ..., K^c\}), \qquad \mathbf{I}^c = \mathbf{I}^c(W^c). \qquad (4)$$

In a Bayesian framework, our objective is to make inference about $W = (W^r, W^c)$ from $\mathbf{I}$ that maximizes a posterior probability,

$$W^* = (W^r, W^c)^* = \arg \max_{\Omega \ni W} p(W^r, W^c | \mathbf{I}) = \arg \max_{\Omega \ni W} p(\mathbf{I} - \mathbf{I}^c(W^c) | W^r) p(W^r) p(W^c).$$

In the computation, the region and curve processes W^r, W^c are coordinated by the generative models expressed in equations (1) and (3), i.e.

$$\mathbf{I}^r = \mathbf{I} - \mathbf{I}^c(W^c).$$

In a language of neuroscience, the two layers W^r, W^c (supposed they are represented by two cortical areas) have some *mutual inhibition* as they compete to explain the observed image, like the lateral inhibition between adjacent neurons (which are bases B_j in our representation). This enables the use of multiple families of image models either 2D or 1D , and distinguishes the generative methods from the discriminative methods for image segmentation (e.g. [6]).

In the following two sections, we discuss the mathematical models for the region and curve processes.

3   Probabilistic Models for Region Processes

In this section, we briefly overview the probabilistic models for region processes, following the DDMCMC work in [11].

1. The likelihood for the region layer $\mathbf{I}^r$. This is a product of individual region models,

$$p(\mathbf{I}^r | W^r) = \prod_{i=1}^{K^r} p_{\ell_i}(\mathbf{I}^r_{R_i}; \Theta_i).$$

$\ell_i \in \{1, 2, 3, 4\}$ indexes the following four families of intensity models.

Family 1: ϖ_1. This is a simple Gaussian model for flat regions. It assumes that pixel intensities in a region R are constant subject to an independently and identically distributed (i.i.d.) Gaussian noise.

Family 2: ϖ_2. This is a non-parametric model for cluttered regions. It assumes that pixel intensities are iid distributed according to a histogram which is discretized as a step function expressed by a vector $\Theta = (h_1, h_2, ..., h_G)$.

Family 3: ϖ_3. This is Markov random field model (FRAME) for textured regions. We choose a set of 8 filters in FRAME and formulate the model in pseudo-likelihood form.

Family 4: ϖ_4. This is a spline model for regions with gradually changing intensities, such as lighting areas.

We refer to [11] for detailed specification of these models. In summary, the algorithm can switch between the families to search for a good fit in terms of high likelihood function for each region $\mathbf{I}_R^r$.

2. The prior model for region process W^r. The prior model $p(W^r)$ penalizes model complexity and ensure boundary smoothness. Let $A_i = |R_i|$ be the area (pixel number) of R_i, and $\Gamma_i = \partial R_i$ be the region boundary, then

$$p(W^r) = p(K^r) \prod_{i=1}^{K^r} p(R_i)p(\ell_i)p(\Theta_i|\ell_i) = p(K^r) \prod_{i=1}^{K^r} p(A_i)p(\Gamma_i)p(\ell_i)p(\Theta_i|\ell_i) \quad (5)$$

$$p(K^r) \propto e^{-\lambda_0 K^r}, \ \ p(A_i) \propto e^{-\gamma A_i^c}, \ \ p(\Gamma_i) \propto e^{-\mu \oint_{\Gamma_i} ds}, \ \ p(\Theta_i) \propto e^{-\nu \mathrm{len}(\Theta_i)}$$

In these probabilities, γ, c, ν, μ are some constants and $\mathrm{len}(\Theta_i)$ is the number of parameters in Θ_i.

A DDMCMC algorithm using this set of priors and image models are tested in a large set of images with satisfactory results[11]. It was also tested in a benchmark dataset of 50 natural images by the Berkeley group, and achieved the best results among the algorithms that have been tested. See http://www.cs.berkeley.edu /~dmartin/segbench/BSDS100/html/benchmark. However, it is evident in these experiments that such region models are not suitable for 1D curve processes. This motivates our curve models below.

4   Probabilistic Models for Curve Processes

In this section, we present a rope model for curve processes, and we start with a brief review of existing curve models.
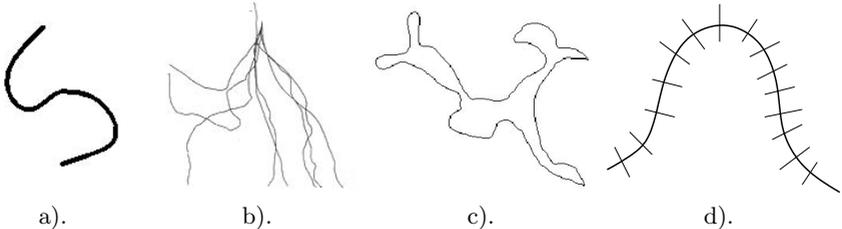

a).   b).   c).   d).

Fig. 2. Four existing curve models. a). SNAKE (Kass et al 1988), b). Random curves sampled from the Elastica model (Mumford, 1994). c). A random curve sampled from a MRF curve model (Zhu, 1999), d). A profile model for modeling faces (Cootes et al, 1995) and contour tracking (Isard and Blake, 1996).

4.1   Previous Curve Models

There are four interesting curve models in the literature as Figure 2 shows. The first one is the SNAKE model[10]. Let $C(s)$, $s \in [a, b]$ be a continuous curve, a SNAKE model has a smooth term plus an image gradient term for intensity discontinuity.

$$p(C) \sim \exp\{-\int_a^b [-|\nabla \mathbf{I}|^2 + \alpha \dot{C}^2(s) + \beta \ddot{C}^2(s)]ds\}.$$

Mumford studied an Elastica model in [8] motivated by a Ulenbeck process for a moving particle with friction, let $\kappa(s)$ be the curvature, then

$$p(C) \sim \exp\{-\int_a^b [\beta + \alpha \kappa^2(s)]ds\}.$$

Figure 2.b shows some typical curves sampled from the above model. The third model was proposed by Zhu in [13] which integrates model Gestalt properties and symmetry in a Markov random field model for closed contours, and is learned by a maximal entropy principle. Figure 2.c shows a typical sample of curve from this model. The fourth model is used by Cootes et al. [3] for face contour and are also used in tracking[2]. It is a smooth contour but also measures image profiles along lines perpendicular to the contour as Figure 2.d shows. This profile is used to local edges nearby. Recently, August and Zucker [1] constituted a so called *curve indicator random field* model.

   As we can see, the gradient term in the SNAKE model or the profile measure are both feature detectors, and they are not generative models for image intensities.

4.2   A Generative Rope Model of Curve Processes

In the real images, the curve patterns are generated by elongated objects, such as trees, stems, cables, strings, rails, and so on. Instead of being curves of constant intensities, such objects often have interesting intensity profiles due to lightings and are blended with the background regions through image formation.
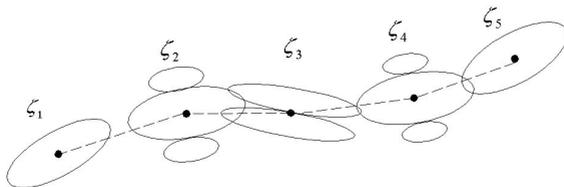

Fig. 3. A rope model of curves consists of a chain of knots. A knot has 1-3 image bases shown by the ellipses.

The curve processes W^c have a number of curves $\{C_i;\ i = 1, 2, ..., K^c\}$, and each curve C is a chain of n knots as shown in Figure 3.

$$C = (n, \zeta_1, \zeta_2, ..., \zeta_n)$$

Each knot ζ consists of 1-3 image bases. One major base shown by a large ellipse is often associated with 0-2 small minor bases to achieve smooth blending with background regions, thus

$$\zeta = (\ (\alpha_1, B_1),\ ...,\ (\alpha_k, B_k)\ ),\quad k \le 3.$$

$\alpha_1, ..., \alpha_k$ are the coefficients of these bases and thus specifies the photometric properties and $B_1, ..., B_k$ are bases selected from an over-complete basis or dictionary Δ. We utilize 134 base functions shown in figure 4 which are Gaussians, DoG, and DOOG bases at various orientation and scales, and can be translated to any location.
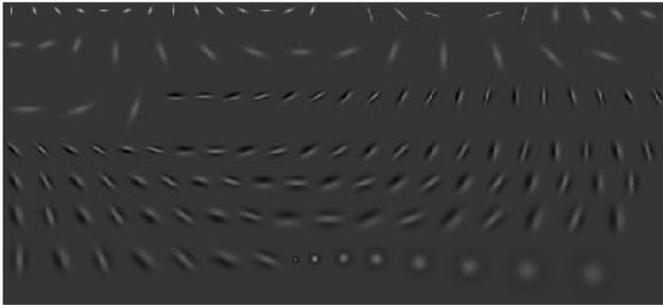

Fig. 4. Base functions used for the disctionary Δ.

For each curve C we have a 2nd order Markov chain model for the knots

$$p(C) = p(n)p(\zeta_1)p(\zeta_2|\zeta_1) \prod_{i=3}^{n} p(\zeta_i|\zeta_{i-1}, \zeta_{i-2}),$$

where $p(n)$ controls the length of the curve. The conditional probabilities $p(\zeta_2|\zeta_1)$ and $p(\zeta_i|\zeta_{i-1}, \zeta_{i-2})$ are defined so that the curve is smooth in geometry (by energy terms for co-linearity and co-circularity of adjacent bases as in the SNAKE/Elastica models) and in appearance (by energy term for the change of base types and coefficients). This model can be learned by a minimax entropy method as in [13].

We draw a set of random curves (ropes) from the model $p(C)$ in Figure 5. Figure 5.a shows the geometric curves for the ropes, and the knots of these chains are shown in Figure 5.b. Then Figure 5.c is the curve layer $\mathbf{I}^c$ which is a linear sum of all bases in the rope, as equation (3) defines.

To summarize, the curve model for $p(W_C)$ is,

$$p(W^c) = p(K^c) \prod_{i}^{K^c} p(C_i), \qquad \mathbf{I}^c = \mathbf{I}^c(W^c).$$
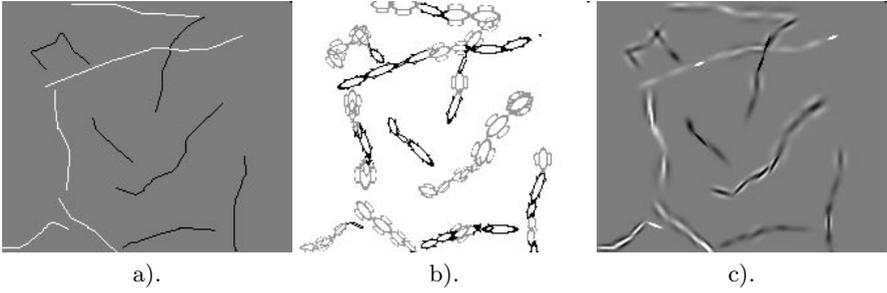
Fig. 5. Random ropes sampled from the prior model $p(C)$. a). The geometric curves for the sampled ropes. b). A symbolic representation of the sampled ropes. c). Curve layer I^c by the ropes.

5   The Solution Space Ω

Given the image models in the previous sections, we now briefly analyze the solution space Ω for $W = (W^r, W^c)$. As K^r, K^c are unknown numbers, the solution space Ω is

$$\Omega = [\cup_{K^r} \Omega^r_{K^r}] \times [\cup_{K^c} \Omega^c_{K^c}],$$

where $\Omega^r_{K^r}$ is the space for W^r with exactly K^r regions and $\Omega^c_{K^c}$ is the space with exactly K^c curves.

In image segmentation, there are two ways of defining a region. One represents a region boundary $\Gamma(s) = \partial R$ as a continuous contour parameterized by s and treats the image domain as a 2D plane. The other considers a segmentation as a label map ψ_Λ, and a region R_i is a set of pixels sharing the same label, say n,

$$R_i = \{(x, y): \ \psi(x, y) = n, (x, y) \in \Lambda\}, \quad \text{for } i = 1, 2, ..., K.$$

In this paper, we adopt the label map notation and define the solution space $\Omega^r_{K^r}$ as a product of a K^r-partition space and K^r spaces for the image models as

$$\Omega^r_{K^r} = [\, \varpi_{\pi_{K^r}} \times \underbrace{\varpi_\Theta \times \cdots \times \varpi_\Theta}_{K^r} \,],$$

where $\varpi_\Theta = \cup^4_{i=1} \varpi_{gi}$. The set of all K^r-partitions, $\varpi_{\pi_{K^r}}$, is a quotient space of the set of all possible K^r-labelings divided by a permutation group, $\mathcal{PG}$, for the labels. It can be denoted as

$$\varpi_{\pi_{K^r}} = \{(R_1, R_2, ..., R_{K^r}) = \pi_{K^r}; \quad |R_i| > 0, \ \forall i = 1, 2, ..., K^r.\}/\mathcal{PG},$$

where

$$\pi_{K^r} = (R_1, R_2, ..., R_{K^r}), \quad \cup^{K^r}_{i=1} R_i = \Lambda, R_i \neq R_j, \ \forall i \neq j$$

is a set of K^r regions of Λ.

2. Solution space for $\Omega^c_{K^c}$

Recall the definition of the curve layer in section 4.2. The curve layer is composed of an unknown number of curves; Each curve consists of an unknown number of knots which are made up of 1-3 bases. In correspondence to this definition, the solution space for the curve layer is defined as

$$\Omega^c_{K^c} = \underbrace{\varpi_C \times \ldots \times \varpi_C}_{K^c}.$$

The space for each curve ϖ_C is of the form

$$\varpi_C = \cup^{|A|}_{n=0} [\underbrace{\varpi_\varsigma \times \ldots \times \varpi_\varsigma}_{n}],$$

where ϖ_ς is the space for each knot and is defined as

$$\varpi_\varsigma = \varpi_B \times \varpi_B \times \varpi_B.$$

In the curve layer, the space for the bases is defined as

$$\varpi_B = \{\phi\} \cup \{(b_i, B_i), i = 1, 2, ..., G_b \times N_B \times |A|\},$$

where G_b is the number of possible values for coefficient b_i, N_B is the number of base functions in the over-complete basis and equals to 134 in this paper, and $|A|$ is the size of the lattice.

6   Integration of Region and Curve Processes by MCMC

Now we turn to the design of the algorithm which forms ergodic Markov chain with reversible jumps[4,5] and diffusion to explore the solution space Ω.

6.1   Dynamics Design

We use the Metroplis-Hasting algorithm in realizing the jump processes. To achieve the detailed balance equation the acceptance rate is computed as

$$\alpha(W \rightarrow dW') = \min(1, \frac{G(W' \rightarrow dW)p(W'|\mathbf{I})dW'}{G(W \rightarrow dW')p(W|\mathbf{I})dW}), \tag{6}$$

where $G(W \rightarrow dW')$ and $G(W' \rightarrow dW)$ are the two proposal probabilities for the Markov chain to jump between solution W and W'. The data-driven techniques are used to propose important proposals to guide the Markov chain in traveling in the solution space Ω more efficiently. The diffusion dynamics are realized by steepest ascent algorithms.

In the following, we briefly discuss different types of these dynamics for region and curve processes.

Jump and Diffusion Dynamics for Region Processes

I. Boundary diffusion This is a diffusion process used to adjust the region boundaries., which are represented by continuous curves evolving to maximize the posterior probability through a region competition equation [12].

II. Model adaptation. This is simply to fit the parameters of a region by steepest descent equation,

$$\frac{d\Theta_i}{dt} = \frac{\partial \log p(\mathbf{I}^r_{R_i}; \Theta_i)}{\partial \Theta_i}.$$

III and IV. Split a region into two and Merge two regions into one. These are a pair of reversible jumps. Suppose at a certain time step, a region R_k with model Θ_k is split into two regions R_i and R_j with models Θ_i, Θ_j, or vice verse, and this realizes a jump between two states W to W'.

$$W = (K, \ (R_k, \ell_k, \Theta_k), \ W_-) \longleftrightarrow (K+1, \ (R_i, \ell_i, \Theta_i), (R_j, \ell_j, \Theta_j), \ W_-) = W',$$

where W_- is the remaining variables that are unchanged during the move.

V. Switch image models. This switches the image model within the four families for a region R_i. For example, from texture description to a spline surface etc.

$$W = (\ell_i, \Theta_i, \ W_-) \longleftrightarrow (\ell'_i, \Theta'_i, \ W_-) = W'.$$

Jump and Diffusion Dynamics for Curve Processes

VI Curve diffusion. This is a diffusion process by curve deformation [10].

VII and VIII. Create a new curve and Delete a curve. They are a pair of reversible jumps. At a time step, a new curve could be created with 1-3 knots or a curve with less than three knots could be killed. The two solution states before and after the change are denoted as

$$W = (K_C, \ W_-) \longleftrightarrow (K_C + 1, \ C_{K_C+1}, \ W_-) = W'. \tag{7}$$

IX and X. Split a curve into two and Merge two curves into one. They are a pair of reversible jumps as well. At a step, two curves could be merged into a new curve or a curve having at least two knots could be split into two curves. We have the two states defined as

$$W = (K_C, \ C_k, W_-) \longleftrightarrow (K_C + 1, \ C_i, \ C_j \ W_-) = W'. \tag{8}$$

There have been a great deal of work done in the area of perceptual organization about how to group 1D elements, such as straight line segments, which potentially belong to the same object together. These grouping methods are considered as bottom-up techniques and could be utilized in designing proposals to help to achieve a fast convergence rate.

XI and XII. Engage a new knot to a curve and Kill a knot from a curve. They are two complementary moves. At a step, a new knot is proposed to engage to one of a curve's two extremes or one of a proposed curve's two extreme knots could be killed. The two states can be written as

$$W = ((\zeta_{i1}, ..., \zeta_{in}), \ W_-) \longleftrightarrow ((\zeta_{i1}, ..., \zeta_{in}, \zeta_{i(n+1)}), \ W_-) = W'. \tag{9}$$

In the following, we give an example of how to compute the proposal probabilities $G(W \to dW')$ and $G(W' \to dW)$ for computing the acceptance rate $\alpha(W \to dW')$ in equation (6). The readers are referred to [11] for a detailed discussion of computing proposal probabilities for region processes.

The two proposal probabilities for move type IX (split a curve into two curves) can be computed as

$$G(W \to dW') = q(IX)q(C_k)q(C_i|C_k)q(C_j|C_k)dW',$$

and

$$G(W' \to dW) = q(X)(q(C_i)q(C_j|C_i) + q(C_j)q(C_i|C_j))dW,$$

where $q(IX)$ is the probability for choosing dynamic type IX, $q(C_k)$ is the probability for proposing curve C_k to split. The two new curves, C_i and C_j, are created with the probabilities $q(C_i|C_k)$ and $q(C_j|C_k)$ respectively. Similarly, $q(X)$ is the probability for choosing dynamic type X, $q(C_i)$ is the probability for choosing curve C_i, and $q(C_j|C_i)$ is the probability for proposing curve C_j to merge with C_i. Likewise, we could compute the values for $q(C_j)$ and $q(C_i|C_j)$.

6.2   Initialization by Image Pyramid and Matching Pursuit

To initialize the algorithm, we decompose an image into two layers $\mathbf{I} = \mathbf{I}_o^r + \mathbf{I}_o^c$ by a simply Laplacian pyramid method, as Figure 6 shows. $\mathbf{I}_o^r$ is a lowpass filtered version of the input image $\mathbf{I}$, and $\mathbf{I}_o^c = \mathbf{I} - \mathbf{I}_o^r$ is the residual image. In general, $\mathbf{I}_o^r$ contains regions, and $\mathbf{I}_o^c$ contains mostly high frequency components for the curves as well as region boundaries.

Then we adopt a match pursuit method by Mallat and Zhang[7] to quickly compute an initial set of image bases from image $\mathbf{I}_o^c$. In particular, we select the ridge type of bases with coefficients larger than a certain significant threshold τ.

$$\mathbf{I}_o^c = \sum_{\alpha_j \geq \tau} \alpha_j B_j + \mathbf{I}_o'.$$

These selected bases are often good seeds for the curves. The reconstruction residual image $\mathbf{I}_o'$ is added to the region layer $\mathbf{I}_o^r \leftarrow \mathbf{I}_o^r + \mathbf{I}_o'$.

For the region layer $\mathbf{I}_o^r$, we conduct some edge detection and clustering as the DDMCMC algorithm did in [11].
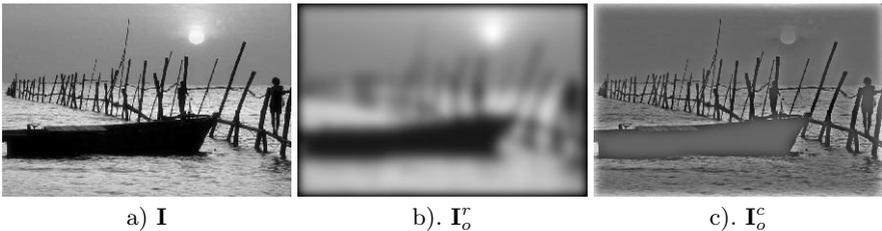

a) $\mathbf{I}$ b). $\mathbf{I}_o^r$ c). $\mathbf{I}_o^c$

Fig. 6. An input image $\mathbf{I}$ (a) is initially decomposed into two components: $\mathbf{I}_o^r$ (b) is a lowpass filtered version, and $\mathbf{I}_o^c$ is the residue for high frequency components.

6.3   Summary of the Algorithm

To summarize, we list the algorithm below.

---

1. Initialize $\mathbf{I}^c = \mathbf{I}^c_o$ and $\mathbf{I}^r = \mathbf{I}^r_o$ by the Laplacian pyramid method.
2. Initial W^c by match pursuit, and W^r by clustering.
3. Select a diffusion dynamic or a jump dynamic at random.
4. If a diffusion dynamic is selected, run the following dynamics at random:
   - For type I, a region is randomly proposed to execute region competition for its boundaries.
   - For type II, a region is randomly selected and its corresponding model parameter is adapted by a steepest descent method.
   - For type VI, a curve is randomly selected and a deformation process is executed as in [10].
5. If a jump dynamic is selected,
   - a new solution W' is randomly sampled according to the dynamic picked at random below.
     - For type III, a region is randomly picked to be split into two new regions which are randomly proposed by one of the partition maps.
     - For type IV, two neighboring regions are randomly chosen to form a new region.
     - For type V, a different model is randomly picked for a randomly chosen region.
     - For type VII, 1-3 knots are randomly proposed to form a new curve.
     - For type VIII, a curve is randomly chosen and is then killed.
     - For type IX, a curve is randomly picked and is then split into two new curves.
     - For type X, two neighboring curves are randomly proposed to form a new curve.
     - For type XI, a new knot is proposed to engage to one of a randomly picked curve's two extreme knots.
     - For type XII, one of a randomly picked curve's two extreme knots is proposed to be killed.
   - The overall posterior probability, $p(W'|\mathbf{I})dW'$, is computed.
   - The proposal probabilities, $G(W' \rightarrow dW)$ and $G(W \rightarrow dW')$ are then computed.
   - A decision of acceptance is made according the acceptance probability $\alpha(W \rightarrow dW')$ in equation 6.
6. Repeat the above steps to draw samples W from $p(W|\mathbf{I})$.

Fig. 7. The algorithm that integrates the region and curve processes.

7   Experiments

We test the algorithm on many natural images and report some results in Figures 8, 9 and 10. Figure 8 shows the results for a sea shore image where regions and curves are parsed. Figure 8.b shows the regions, and we sample a synthesized image from the likelihood $\mathbf{I}^r_{\mathrm{syn}} \sim p(\mathbf{I}|W^r)$ given the computed W^r, this illustrates the computed region models. Clearly the shading effects in the water

a). input image   ⇓   b). region process   ⇓   c). curve process

d).synthesized   e). Synthesis by region process   f). synthesis by curve process
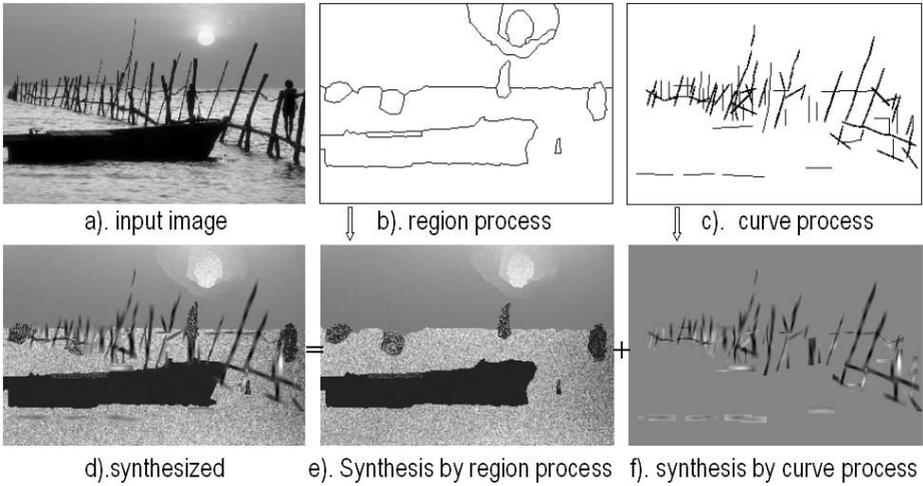
Fig. 8. Result: parsing a sea shore image into regions and curves.

and the sky are captured by the spline model in family 4. Figure 8.c is the sketch for the computed curves W^c, and Figure 8.f is the image $\mathbf{I}^c = \mathbf{I}^c(W^c)$. Thus we have an overall synthesis $\mathbf{I}_{syn} = \mathbf{I}^r_{syn} + \mathbf{I}^c$ shown in Figure 8.d. By comparing $\mathbf{I}^r_{syn}$ and $\mathbf{I}^c$ to the initial decomposition $\mathbf{I}^r_o$, $\mathbf{I}^c_o$ in Figure 6, we clearly see that it is the more sophisticated region and curve models that refine the decomposition. Figure 9 shows the segmented regions W^r and $\mathbf{I}^r_{syn}$ by the previous DDMCMC algorithms which assumes region processes only.

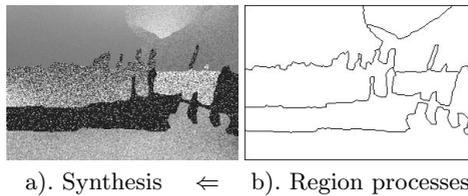a). Synthesis   ⇐   b). Region processes

Fig. 9. The segmentation and synthesis with region processes only for fig. 8

Figure 10 displays more examples in a similar way. The improvement of synthesized images for all the examples demonstrate the the advantages of engaging curve models. The parameters for the two methods are set to be the same to have a fair comparison. Although Figure 10.a by the region processes only successfully segmented many tree trunks out, it disconnects the background. The synthesis by region and curve processes for Figure 10.b is more similar to the original image than that by the region processes, thus, the new algorithm acheives a better visual effect.
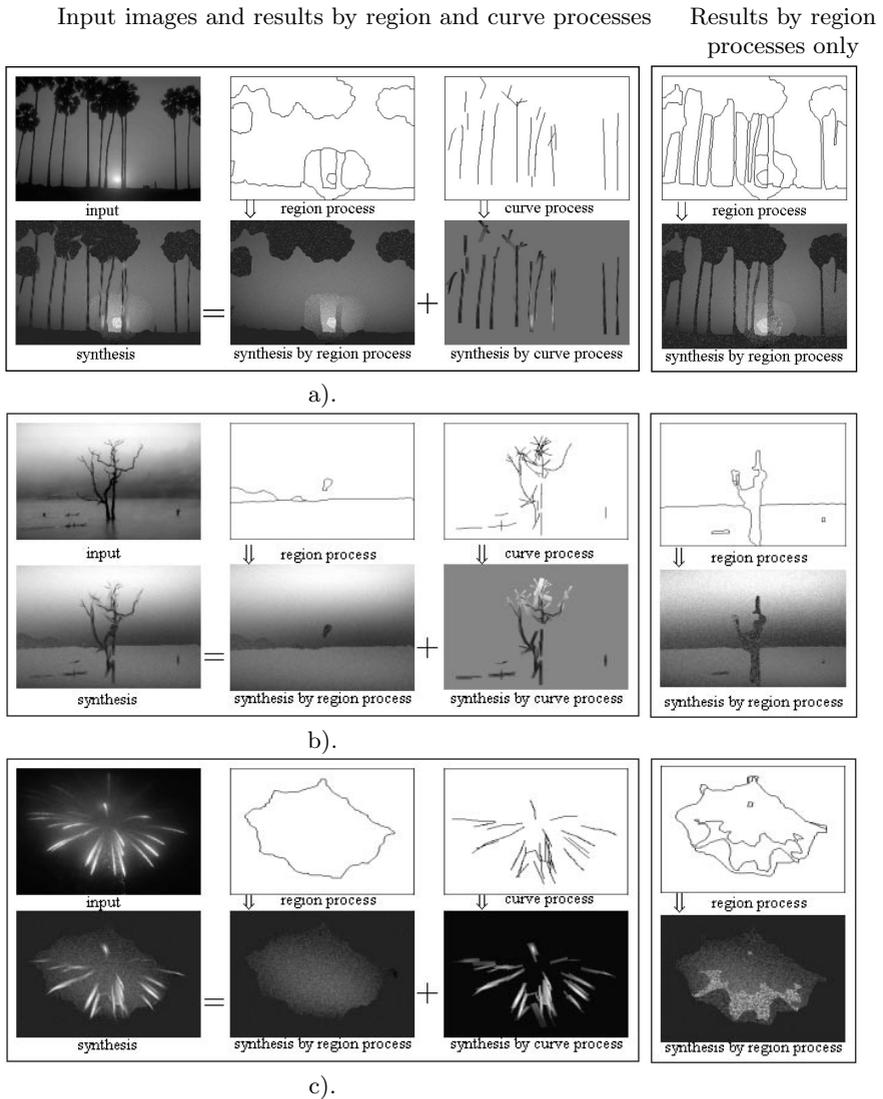
Input images and results by region and curve processes   Results by region processes only

a).

b).

c).

Fig. 10. More image parsing results by the algorithms with the integration of region and curve processes and by the previous DDMCMC algorithm with region processes only.

The semantical representation W by region and curve models can largely reduce the coding length of image. The table below lists the number of bytes used by W for the synthesized images in comparison with the jpeg algorithm for the original images.

Table 1. A comparion of coding lengths in bytes for the example images by the proposed algorithm and the jpeg algorithm.

	The image in fig. 8	Image a in fig. 10	Image b in fig. 10	Image c in fig. 10
Region and Curve processes	2,387	2,266	1,347	922
jpeg	17,001	10,620	10,620	6,656

8   Future Work

In future study, we plan to engage generative models for faces and texton process to realize the image parsing as Figure 1 demonstrates and apply the algorithm for non-photo-realistic rendering to render stylish paintings.

References

1. J. August and S.W. Zucker. A generative model of curve images with a completely-characterized non-gaussian joint distribution. In *Proc. of Workshop on Stat. and Comp. Theories of Vis.*, July, 2001.
2. M. Isard and A. Blake. Contour tracking by stochastic propagation of conditional density. *Proc. ECCV*, 1996.
3. T.F. Cootes, D. Cooper, C.J. Taylor and J. Graham. Active shape models - their training and application. *Computer Vision and Image Understanding*, vol.61, no. 1, Jan. pages 38-59,1995.
4. U. Grenander and M. I. Miller. Representation of knowledge in complex systems. *J. R. Stat. Soc., B*, vol 56, 549-603, 1994
5. P. J. Green. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika*, vol. 82, 711-732, 1995.
6. J. Malik, S. Belongie, T. Leung, and J. Shi. Contour and texture analysis for image segmentation. *IJCV*, vo. 43, no.1, June 2001.
7. S.G. Mallat and Z. Zhang. Matching pursuits with time-frequency dictionaries. *IEEE Tranc. on Signal Processing*, vol. 41, no. 12, Dec. 1993.
8. D. Mumford. Algebraic geometry and its applications, chaper elastic and computer vision, pp. 491-506. Springer-Verlag, 1994.
9. B. A. Olshausen and D. J. Field. Sparse coding with an overcomplete basis set: a strategy employed by v1?. *Vision Res.*, vol. 37, no. 23 3311-25 1997.
10. M. Kass, A. Witkin and D. Terzopoulos. Snakes: active contour models. *IJCV*, 1, 1988.
11. Z. Tu, S.C. Zhu., and H.Y. Shum. Image segmentation by data driven markov chain monte carlo. *Proc. ICCV*, 2001.
12. S. C. Zhu and A. L. Yuille. Region Competition: Unifying Snakes, Region Growing, and Bayes/MDL for Multiband Image Segmentation. *IEEE Trans. PAMI* vol.18, No.9, 1996.
13. S.C. Zhu. Embedding gestalt laws in markov random fields- a theory for shape modeling and perceptual organization. *PAMI*, vol. 21, no.11, Nov. 1999.