An Object-Based Approach for Detecting Small Brain Lesions: Application to Virchow-Robin Spaces

Xavier Descombes*, Frithjof Kruggel, Gert Wollny, and Hermann Josef Gertz

***Abstract*—This paper is concerned with the detection of multiple small brain lesions from magnetic resonance imaging (MRI) data. A model based on the marked point process framework is designed to detect Virchow-Robin spaces (VRSs). These tubular shaped spaces are due to retraction of the brain parenchyma from its supplying arteries. VRS are described by simple geometrical objects that are introduced as small tubular structures. Their radiometric properties are embedded in a data term. A prior model includes interactions describing the clustering property of VRS. A Reversible Jump Markov Chain Monte Carlo algorithm (RJMCMC) optimizes the proposed model, obtained by multiplying the prior and the data model. Example results are shown on T_1-weighted MRI datasets of elderly subjects.**

***Index Terms*—Features extraction, marked point processes, reversible jump MCMC, Virchow Robin spaces.**

I. Introduction

DETECTING focal lesions in magnetic resonance imaging (MRI) data sets of the human head is considered a nontrivial segmentation task. Segmentation approaches require prior knowledge about the lesion charactistics (e.g., their expected compartment, size and shape, their signal statistics in relation to the embedding tissue), and thus, are generally targeted for detecting a specific lesion type (e.g., [6], [18], [24], and [29]). For historical reasons and their medical significance, most approaches focused on white matter (WM) lesions due to Multiple Sclerosis (e.g., [15], [19], [26]), while methods for segmenting large focal lesions appeared only recently (e.g., [13]).

Different types of small focal lesions are typically found in MRI scans of elderly subjects. Their neuropathological substrate and their influence on cognitive abilities is still under debate [10]. The neurobiological background of this work is to aid in discriminating healthy from pathological aging as revealed by MRI brain data sets.

While a trained human observer is still outperforming automatic approaches for lesion detection and discrimination, estimating the lesion count and describing their position is tedious due to their multitude. So typically, these lesions are evaluated visually in the acquired data sets and rated by semiquantitative scales [22]. One such lesion type is called "enlarged Virchow-Robin space" (VRS) that corresponds to a small gap around a deep penetrating artery supplying the WM. Such lesions appear as small tubular structures filled with cerebro-spinal fluid (CSF) that are (ideally) radial to the brain surface. Typically, they are close to the spatial resolution limit of current MRI methods (1–3 mm in diameter and 3–15 mm long), and a single brain may contain hundreds of such lesions.

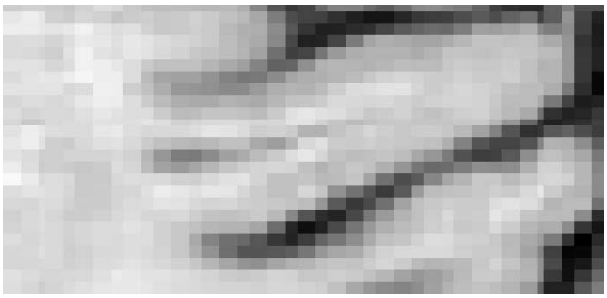


Fig. 1. An example of the tubular shape of VRS.

Herein, we consider an approach to detect small lesions that are nonuniformly distributed in space, have some shape prior and some constraints for their relative positions. Our framework is based on a marked point process and general under these hypotheses.

Structure and distribution of VRS have motivated the following approach. First, a VRS typically consists of several voxels forming a tubular shape (see Fig. 1). Thus, an approach based on a geometrical object appears suited for this problem. Second, VRS are not uniformly localized and some regionally higher occurence can be observed (see Fig. 2). Interactions between objects are included in the model to favor clustering of VRS while suppressing an overlap of different VRS. A certain shape variability (length and diameter of the tubular structures in our case) and data noise have to be taken into account introducing local minima of the model-related cost function. To overcome this problem, we consider a stochastic framework.

The template theory introduced by Grenander allows modeling random shapes [9]. Within this pattern theory framework, several kinds of objects can be considered and complex scenes containing several objects can be represented by a graph [17]. In our context, the object's shape model is simple and represented by a small segment. The complexity of the model lies in the different object interactions we will introduce. Therefore, we have preferred the marked point process framework here [1], [2]. This framework has been successfully applied in image analysis, e.g.,

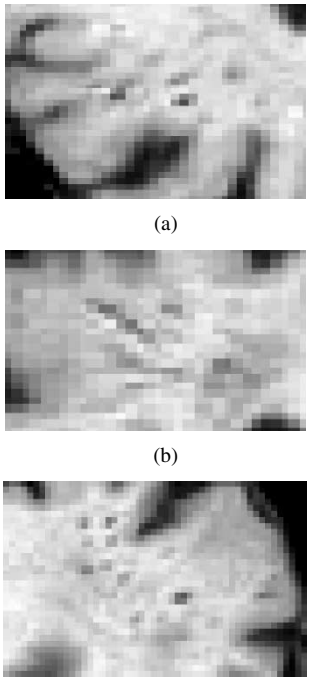
Manuscript received October 7, 2002; revised October 27, 2003. The work of X. Descombes was supported in part by the Max-Planck Institute (MPI) through a travel grant during summer 2001. The Associate Editor responsible for coordinating the review of this paper and recommending its publication was W. Niessen. *Asterisk indicates corresponding author.*

*X. Descombes is with Ariana, common project CNRS/INRIA/UNSA, INRIA, BP93, 2004 route des Lucioles, 06902 Sophia Antipolis Cedex, France (e-mail: xdescomb@sophia.inria.fr).

F. Kruggel and G. Wollny are with the Max-Planck Institute of Cognitive Neuroscience, 04103 Leipzig, Germany.

H. J. Gertz is with the Department of Psychiatry, University Clinic Leipzig, 04103 Leipzig, Germany.

Digital Object Identifier 10.1109/TMI.2003.823061



(a) (b) (c)

Fig. 2. An example of the clustering property of VRS: (a) axial, (b) coronal, and (c) sagittal sections.

in the context of cell segmentation [14], [21] or road network extraction [23]. A point process with a Poisson density measure models the number of objects and their localization in the scene. This process may include interactions between points such as explicit relations or clustering properties. Shape parameters are associated with the points to define the object geometry (e.g., length and orientation of an object).

Model optimization can be achieved by a birth and death process [2], a jump diffusion [17] or a Reversible Jump Markov Chain Monte Carlo algorithm (RJMCMC) [5], [8]. Birth and death processes require the computation of a birth (or death) rate map which leads to costly algorithms on a three–dimensional (3-D) volume. The main advantage of the RJMCMC approach lies in the flexibility of the derived algorithms. Within this framework, we can define dynamics adapted to the target distribution. Therefore, we prefer a RJMCMC algorithm embedded in a simulated annealing scheme. We define a proposition kernel mixing different subkernels allowing us to move or to add and suppress objects. Some specific subkernels allow us to speed up the convergence rate of the algorithm.

In the Section II, we describe a data model for VRS. Filters designed as VRS indicators define the space where the first points of the objects are located. The subsequent section introduces the marked point process that consists of the prior model (i.e., the geometry of the object and their interactions) and the data term (i.e., the filter output). The RJMCMC algorithm is described in Section IV. Section V is devoted to results. Finally, some conclusions are drawn in Section VI.

II. Data Model

In this section, we derive filters which can be interpreted as VRS indicators. These filters extract the information provided by the radiometry (grey level of voxels). However, this information is not sufficient to characterize the lesion, so a prior is defined to include knowledge about the relative localization and geometric properties of VRS. Filter results serve to define the data term and the volume where the point of the process defined in Section III are living. In a first approach, they can be modeled as small tubular structures, with a diameter close to the data resolution (typically 1–3 mm, [3]). Their length varies with an average of 3–4 mm but can reach up to 15 mm in exceptional cases. The data representation of these structures is discrete due to the discrete nature of the image lattice. The induced discretization is far from being isotropic. Therefore, we can only consider the three main directions (u, v, w) defined by the data without loosing significant information. For these three directions we design filters that take the diameter variability and the partial volume effect into account. A filter lies in the plane perpendicular to the considered direction, say u. It consists of a central voxel s, the eight surrounding neighbors $t_1 \in \mathcal{N}_u^1(s)$, and the next 16 neighbors $t_2 \in \mathcal{N}_u^2(s)$ (see Fig. 4). A VRS is characterized by three properties: 1) VRS contain CSF that should appear as low-intensity voxels in T_1-weighted datasets. 2) Surrounding tissue (i.e., WM or GM voxels) may increase lesion intensities due to the partial volume effect. Thus, neighboring voxels will have higher intensity and 3) are contrasted with the VRS voxels. We define three filters corresponding to these three properties $F^{\text{black}}(i_s)$, $F^{\text{white}}(\min(i_t, t \in \mathcal{N}_u^2(s)))$ and $F^{\text{contrast}}(\sum_{t\in\mathcal{N}_u^2(s)} i_t/16)$, where i_s represents the grey level of voxel s in the data. Note that these three properties are not equivalent. Considering only one of them leads to false alarms in the CSF compartment for property 1), in the WM for property 2), and in the thin GM structures for property 3). The three defined functions depend on the data statistics and are defined on Fig. 3. We combine these three properties as shown in (1) at the bottom of the page. To avoid multiple detections of the same VRS, the final filter is written as follows:

$$F_u(s) = \begin{cases} f_u(s), & \text{if } \forall t \in \mathcal{N}_u^1(s), i_s \leq i_t \\ \min(0, f_u(s)), & \text{otherwise} \end{cases}. \quad (2)$$

To compute the mean and standard deviation of the CSF, GM, and WM, we used a segmentation based on a region growing

$$f_u(s) = \min\left(F^{\text{black}}(i_s), F^{\text{white}}(\min(i_t, t \in \mathcal{N}_u^2(s)), F^{\text{contrast}}\left(\frac{\sum_{t\in\mathcal{N}_u^2(s)} i_t}{16}\right)\right). \quad (1)$$

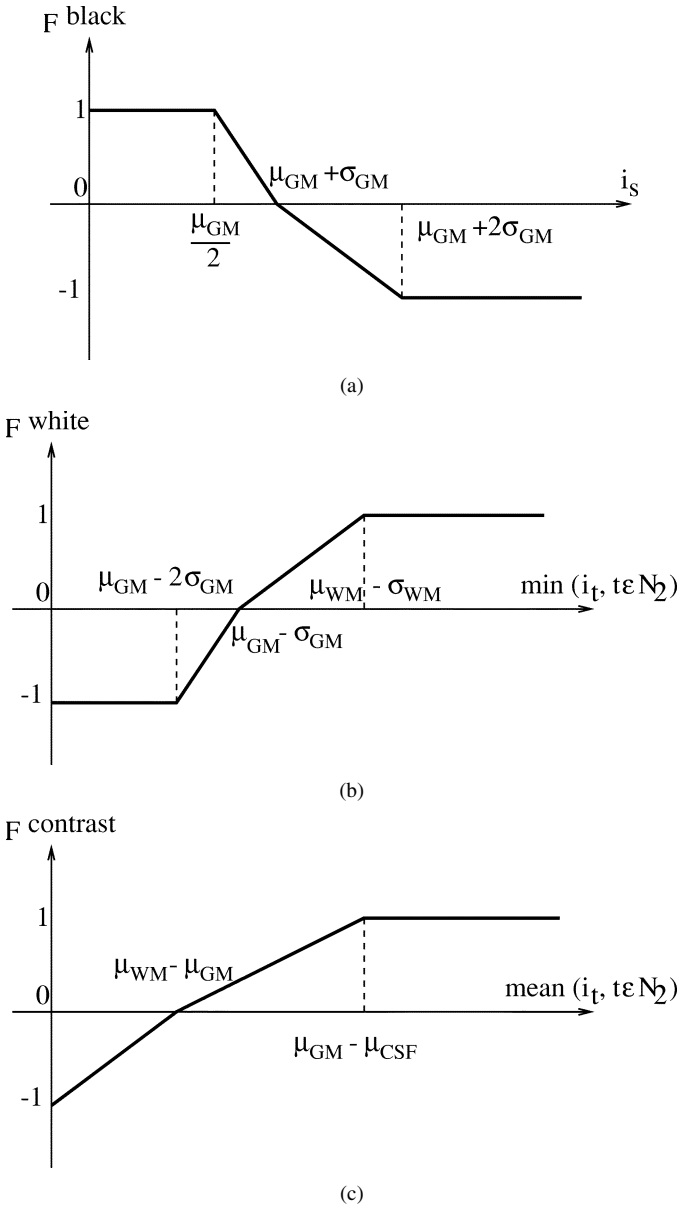


Fig. 3. The three filters characterizing VRS data: μ_{CSF}, μ_{GM}, and μ_{WM} represent the mean of CSF, GM, and WM, σ_x being the associated standard deviations.

approach [12]. The presence of lesions leads to local segmentation errors. Since their total volume is small [13], we neglect this influence on the grey level statistics of the different classes.

III. The Marked Point Process

To increase the sensitivity and reduce the false alarms rate, we have to consider geometric information about VRS. Besides, VRS are not distributed uniformly in the brain. The information on their location obtained from medical studies has to be considered. The marked point process framework is adapted here because it can embed such properties. The number of objects is random (we do not know *a priori* the number of VRS in a dataset), some prior knowledge concerning the geometry is modeled by the object definition and by a density on the underlying parameters. Besides, some interactions between objects

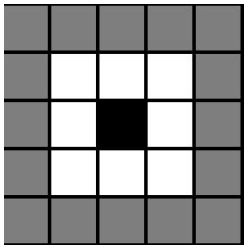


Fig. 4. VRS indicator filter: central voxel s in black, first neighborhood $\mathcal{N}_u^1(s)$ in white, and second neighborhood $\mathcal{N}_u^2(s)$ in grey.

allow us to take into account the information concerning the localization and the relations between the different VRS.

The full model $h(x)$ for the marked point process consists of the prior model $f(x)$ and the data term $g(x)$

$$h(x) = f(x)g(x). \tag{3}$$

Both terms are now discussed in detail.

A. Prior Model

The prior model takes the geometry of a detected object and its interactions into account. Each object is represented by a point with attributes (or marks) that define its geometry. The resulting configurations is a set of marked points $x = \{\vec{s(1)} = (x(1), \vec{l(1)}), \ldots, \vec{s(n)} = (x(n), \vec{l(n)})\}$ where $x(i) \in V$ and $\vec{l(i)}$ is a vector in $\mathcal{R}^3$. We restrict the volume V to the space of points that satisfy $\max(F_u(x(i)), F_v(x(i)), F_w(x(i))) \geq 0$, where F_u, F_v and F_w are the filters defined in Section II.

We consider objects as small oriented segments $\vec{s}$ with ending points x and $x + \vec{l}$ where $x \in V$ and $\| \vec{l} \| \in [\rho_{\min}, \rho_{\max}]$, $\rho_{\min}$ ($\rho_{\max}$) corresponds to the minimum (maximum) segment length. We define a distribution on the configuration space whose density with respect to a Poisson measure is written as follows:

$$f(x) \propto \beta^n \prod_{k \in \langle 1,n \rangle} q_1(\| \vec{l(k)} \|) \prod_{k,k' \in \langle 1,n \rangle : k \sim k'} q_2(\vec{s(k)}, \vec{s(k')}) \tag{4}$$

where $\sim$ defines a neighborhood relation. β is a density parameter which drives the number of objects in the configuration (note that the density $f(x) \propto \beta^n$ corresponds to a Poisson process with intensity β). The term $q_1(\| \vec{l(k)} \|)$ represents a prior on the segment length and is defined as follows:

$$q_1(\| \vec{l(k)} \|) = \exp - \left[A \left(\frac{\rho_{\max} - \| \vec{l(k)} \|}{\rho_{\max} - \rho_{\min}} \right)^2 \right]. \tag{5}$$

In our experiments, we have chosen $\rho_{\min} = 2$ and $\rho_{\max} = 15$, according to the length of VRS observed in our datasets and reported in the literature [10], [11]. The term $q_2(\vec{s(k)}, \vec{s(k')})$ defines interactions between neighboring segments. We consider three kinds of interactions (see Fig. 5): 1) an explicit term which penalizes intersecting segments; 2) a clustering term which favors neighboring segments with similar orientation; and 3) a repulsive term which penalizes neighboring segments

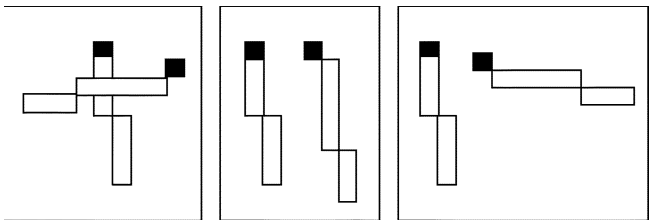


Fig. 5. Explicit (left), clustering (middle) and repulsive (right) interactions.

with different orientations. To define the different interactions, we consider the silhouette $S(\vec{s})$ of the segment $\vec{s}$ as the discrete projection of the segment onto the lattice. We use the following explicit interaction to penalize intersecting segments (hard core interaction):

$$q_2^{hc}(s(\vec{k}), s(\vec{k'})) = \exp[-B] \text{ if } S(s(\vec{k})) \cap S(s(\vec{k'})) \neq \emptyset. \quad (6)$$

Denote the three coordinates of $\vec{s}$ by (s_u, s_v, s_w). Two segments $\vec{s}$ and $\vec{s'}$ have a similar u-direction if and only if

$$\begin{cases} |s_u| & > & \max(|s_v|, |s_w|) \\ |s'_u| & > & \max(|s_v|, |s'_w|) \\ s_u \times s'_u & > & 0 \end{cases}. \quad (7)$$

If $s(\vec{k})$ and $s(\vec{k'})$ have a similar u-direction, we have a clustering interaction

$$q_2^{clu}(s(\vec{k}), s(\vec{k'})) = \exp\left[+C \min\left(|s_u(k)|, |s_u(k')|\right)\right], \quad \text{if } |x_u(k) - x_u(k')| < c. \quad (8)$$

This interaction is easily extended to segments having similar v or w directions. Finally if two segments have no similar directions, we define a repulsive interaction as follows:

$$q_2^{\text{rep}}(s(\vec{k}), s(\vec{k'})) = \exp[-D] \text{ if } \|x(k) - x(k')\| < d. \quad (9)$$

Note that $q_2^{clu}(s(\vec{k}), s(\vec{k'}))$ and $q_2^{\text{rep}}(s(\vec{k}), s(\vec{k'}))$ are mutually excluding. Since we found VRS pairs with no similar direction and $d = 4$, we have chosen $c = 5$ and $d = 3$ in our experiments.

B. Data Term

We now define the data term $g(x)$. Assuming that data are independent conditionally to the segments we can write

$$g(x) \propto \prod_{k \in \langle 1,n \rangle} s(\vec{k}). \quad (10)$$

The data term corresponding to a segment $\vec{s}$ consists of two parts. The first part is proportional to the filter values along the segment as defined in Section II. In the second part, we assume

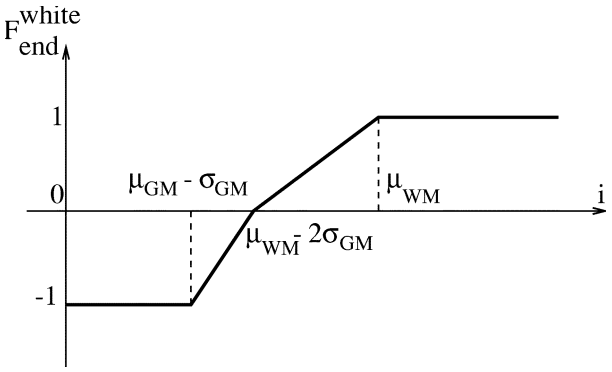


Fig. 6. The prolongating filter: μ_{CFS}, μ_{GM}, and μ_{WM} represent the mean of CSF, GM, and WM, σ_x correspond to the associated standard deviations.

that voxels at the segment end have a high intensity as they do not belong to the VRS. The data term is then written as follows:

$$r(\vec{s}) = \exp\left[E\left(\sum_{t \in S(\vec{s})} F_u(t) + \frac{n_{\vec{s}}}{2} F_u^{\text{end}}(\vec{s})\right)\right] \text{ if } |s_u| \geq \max(|s_v|, |s_w|)$$

$$r(\vec{s}) = \exp\left[E\left(\sum_{t \in S(\vec{s})} F_v(t) + \frac{n_{\vec{s}}}{2} F_v^{\text{end}}(\vec{s})\right)\right] \text{ if } |s_v| \geq \max(|s_u|, |s_w|)$$

$$r(\vec{s}) = \exp\left[E\left(\sum_{t \in S(\vec{s})} F_w(t) + \frac{n_{\vec{s}}}{2} F_w^{\text{end}}(\vec{s})\right)\right] \text{ if } |s_w| \geq \max(|s_u|, |s_v|) \quad (11)$$

where $n_{\vec{s}}$ is the number of voxel in $S(\vec{s})$ and (12), shown at the bottom of the page, holds, where $F_{\text{end}}^{\text{white}}(\cdot)$ is defined on Fig. 6. Similar definitions are used for $F_v^{\text{end}}(\vec{s})$ and $F_w^{\text{end}}(\vec{s})$.

IV. RJMCMC Algorithm

In this section, we derive an RJMCMC algorithm to optimize the model $h(x)$. For details about the theoretical background of this algorithm we refer to [8]. The optimal solution is defined by the configuration maximizing the density distribution $h(x)$. We consider a simulating annealing scheme by iteratively sampling the distributions $h_T(x) = (h(x))^{1/T}$, while the temperature parameter T is decreasing. Contrary to more classical models such as Markov random fields, the number of random variables is not fixed because we do not know the number of segments beforehand. The Metropolis-Hastings algorithm must

$$F_u^{\text{end}}\left(\vec{s} = (x, \vec{l})\right) = \begin{cases} F_{\text{end}}^{\text{white}}\left(\min\left(i_{x+\vec{l}+(1,0,0)}, i_t, t \in \mathcal{N}_u^1(i_{x+\vec{l}+(1,0,0)})\right)\right) + F_{\text{end}}^{\text{white}}\left(\min\left(i_{x+(-1,0,0)}, i_t, t \in \mathcal{N}_u^1(i_{x+(-1,0,0)})\right)\right), & \text{if } \vec{l}\,_u > 0 \\ F_{\text{end}}^{\text{white}}\left(\min\left(i_{x+(1,0,0)}, i_t, t \in \mathcal{N}_u^1(i_{x+(1,0,0)})\right)\right) + F_{\text{end}}^{\text{white}}\left(\min\left(i_{x+\vec{l}+(-1,0,0)}, i_t, t \in \mathcal{N}_u^1(i_{x+\vec{l}+(-1,0,0)})\right)\right), & \text{if } \vec{l}\,_u \leq 0 \end{cases} \quad (12)$$

be extended to allow jumps between configurations of different dimension, which can be achieved by the RJMCMC algorithm proposed by Green [8].

A. General Algorithm

In this section, we briefly recall the RJMCMC algorithm. First, we define a simulated annealing scheme by considering the following procedure.

1. Set T to $T^{(0)}$, n to 0.
2. Compute I iterations of the RJMCMC algorithm to simulate $h_{T^{(n)}}(x)$.
3. If n = *stop* exit, else set n to $n+1$, $T^{(n+1)} = aT^{(n)}$ (with $0 < a < 1$) and go to 2.

To simulate a distribution $\pi(x)$, we consider several transition kernels $Q_m(x \to dx')$ corresponding to different moves (birth/death of a segment, changing the mark of a segment, etc.) in the configuration space. The global transition kernel is then given by $Q(x \to dx') = \sum_m p_m Q_m(x \to dx')$ with $\sum_m p_m = 1$. An iteration of the RJMCMC algorithm is then written as follows.

1. Select a move m with probability p_m.
2. Select x' according to $Q_m(x \to dx')$.
3. Compute the acceptance ratio

$$\alpha_m = \frac{\pi(x')}{\pi(x)} \frac{Q_m(x' \to dx)}{Q_m(x \to dx')} \tag{13}$$

where $Q_m(x' \to dx)$ is the reverse move associated with $Q_m(x \to dx')$
4. Accept the transition $x \to dx'$ with probability $p = \min(1, \alpha_m)$

For some move (e.g., a birth), the dimension of x' may differ from the dimension of x, say $\dim(x') > \dim(x)$. In that case, we have to match the dimension by considering auxilliary random variables u, which are sampled using a distribution $q(u)$, and such that there exists a bijective transformation $\mathcal{T}$ with $x' = \mathcal{T}(x, u)$. Then, we have

$$Q_m(x \to dx') = q(u)|\mathcal{J}(\mathcal{T})|P_m \tag{14}$$

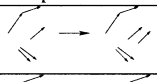
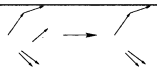
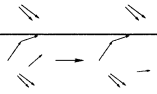
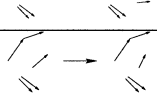
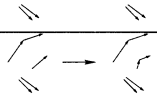
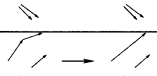
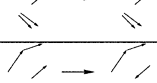
where $\mathcal{J}$ is the Jacobian and P_m is the probability to propose the forward jump [in the birth example $P_m = P(\text{birth})/P(\text{birth}) + P(\text{death})$].

B. The Different Moves

Hereafter, we describe the different moves. These moves and their associated acceptance ratio are summarized in Table I. The derivations of the acceptance ratios are given in the Appendix .

We first consider a birth and death move. This move is essential to establish the real number of segments in the configuration. Besides, it guaranties the irreductibility property. This move is chosen with probability p_1. It consists in adding a new segment with probability P_b or removing a segment from the configuration with probability $P_d = 1 - P_b$. In the case of a death move, the segment is chosen uniformly among the segments of the current configuration. For the birth move, we propose a new segment uniformly in the parameter space. To overcome the low acceptance ratio of the death move when the current configuration contains only a few segments, we propose to move a segment. We choose a segment uniformly within the configuration and select a new location and new marks uniformly within the parameter space. The move is made with probability p_2. With this move, we can change the segments having a low probability. To allow a local improvement of the segments localization with respect to the data we consider the move of end points. This move is made with probability p_3. We choose a new location of the end point uniformly in a neighborhood of its current position. This move may also extend segments in partially detected VRS. A VRS can be covered by several nonoverlapping segments during optimization. Extending segments has a very low acceptance ratio in this configuration because it may induce overlapping, which is penalized by the explicit interaction term. A death move followed by an extension has also a low acceptance ratio because all the segments within this local configuration fit the data well. Therefore, we have introduced a move which merges neighboring segments. The reverse move corresponds to splitting a segment in two parts. This move is made with probability p_4. A merging is proposed with probability P_m and a splitting with probability $P_s = 1 - P_m$. Segments having an opposite direction cannot be merged because the constraint imposing that the first point of each segment should belong to V leads to difficulties in the computation of the acceptance ratio. To merge such configurations we define a new move which returns a segment. It is proposed with probability $p_5 = 1 - (p_1 + p_2 + p_3 + p_4)$.

TABLE I
DIFFERENT MOVES DEFINING THE RJMCMC ALGORITHM: x IS THE CURRENT CONFIGURATION AND x' IS THE PROPOSED ONE (SEE THE APPENDIX FOR DETAILS AND NOTATIONS)

Move	Representation	Proba.	Acceptance ratio
Birth		$p_1 \times P_b$	$\frac{P_d}{P_b} \frac{h_T(x')}{h_T(x)} \frac{4\pi^2(\rho_{max}-\rho_{min})V}{n+1}$
Death		$p_1 \times P_d$	$\frac{P_b}{P_d} \frac{h_T(x')}{h_T(x)} \frac{n}{4\pi^2(\rho_{max}-\rho_{min})V}$
Seg. mv		p_2	$\frac{h_T(x')}{h_T(x)}$
End pt mv		p_3	$\frac{h_T(x')}{h_T(x)}$
Split		$p_4 \times P_s$	$\frac{P_m}{P_s} \frac{h_T(x')}{h_T(x)} \frac{(2D_m+1)^5(\lvert I_u\rvert-2*\rho_{min})}{n_m+1}$
Merge		$p_4 \times P_m$	$\frac{P_s}{P_m} \frac{h_T(x')}{h_T(x)} \frac{n_m+1}{(2D_m+1)^5(\lvert I_u\rvert-2*\rho_{min})}$
Return		p_5	$\frac{h_T(x')}{h_T(x)}$

C. Convergence

The optimization is performed by a MCMC sampler embedded into a simulated annealing scheme. If the considered point process is stable, which is the case here, the sampler converges to a stationary measure, which is the defined process. Convergence is reached for any initial configuration, e.g., the empty configuration chosen here. Convergence is achieved

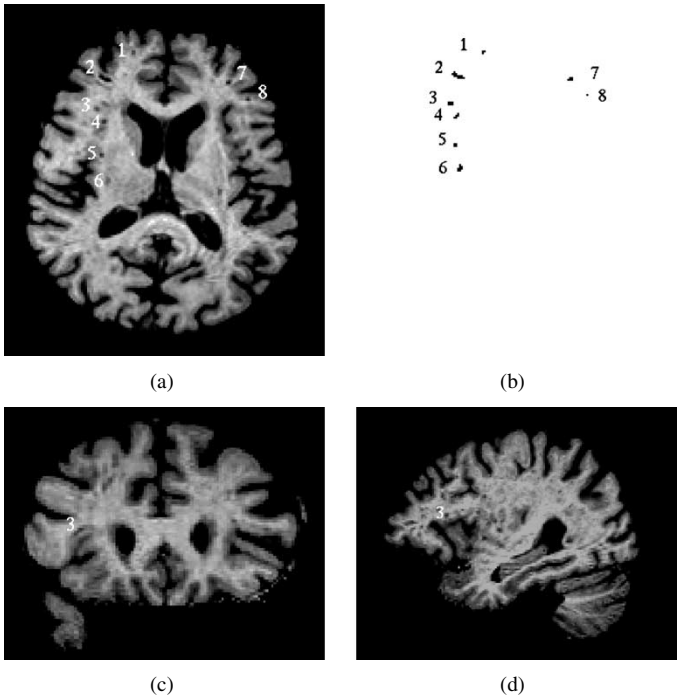


Fig. 7. An example of VRS detection: (a) Original axial slice, (b) detected VRS, (c) Coronal slice corresponding to VRS 3, and (d) sagittal slice corresponding to VRS 3.

in theory with only birth and death moves. However, adding other moves speeds up the optimization process. The weighting between the different moves defining the kernel has been set empirically to optimize the convergence rate. The simulating annealing scheme provides a convergence to a Dirac measure on the configuration maximizing the defined density. This convergence property has been proved for a logarithmic decrease of the temperature. In practice, to speed up the process, we consider a geometric decrease of the temperature.

V. Results

The target application of the algorithm described above is the detection of VRS in MRI brain data sets. Since VRS are a typical sign of an aged brain, we selected a random subsample of datasets acquired in a clinical study of minimal cognitive dysfunction in a population of elderly subjects [27].

All 3-D datasets were acquired on a Siemens Vision 1.5-T scanner using a T_1-weighted MPRAGE protocol (TR 11.4 ms, TE 4.4 ms, 128 sagittal slices, matrix 256×256, voxel size $0.9 \times 0.9 \times 1.5$ mm). An affine transformation using b-spline interpolation [25] was applied to yield a dataset aligned with the stereotactical coordinate system [16] at an isotropical resolution of 1 mm.

For all experiments, we used the same parameters: $\beta = 1$, $A = 10$, $B = 1e10$, $C = 20$, $D = 60$, and $E = 80$. The huge value of parameter B is used to forbid overlapping segments in the result. Other parameters were set by trials on a single dataset.

A. Detail Analysis on One Example

First, we present results on a particular dataset. This dataset has been selected among cases with a high prevalence of VRS. Fig. 7 shows the result on a sample axial slice and the detected lesions on Fig. 7(b). The coronal and sagittal slices corresponding to one particular VRS are, respectively, shown in Fig. 7(c) and (d). A 3-D vizualization of the result is shown on Fig. 8. We performed a shape analysis of the different connected components. As seen on Fig. 9, the anisotropy of the VRS indicates that the tubular structures are not very elongated. A possible explanation is that we are detecting only a part of the VRS which is mostly within the WM compartment. Indeed, the detected lesions do not reach the cortical surface.

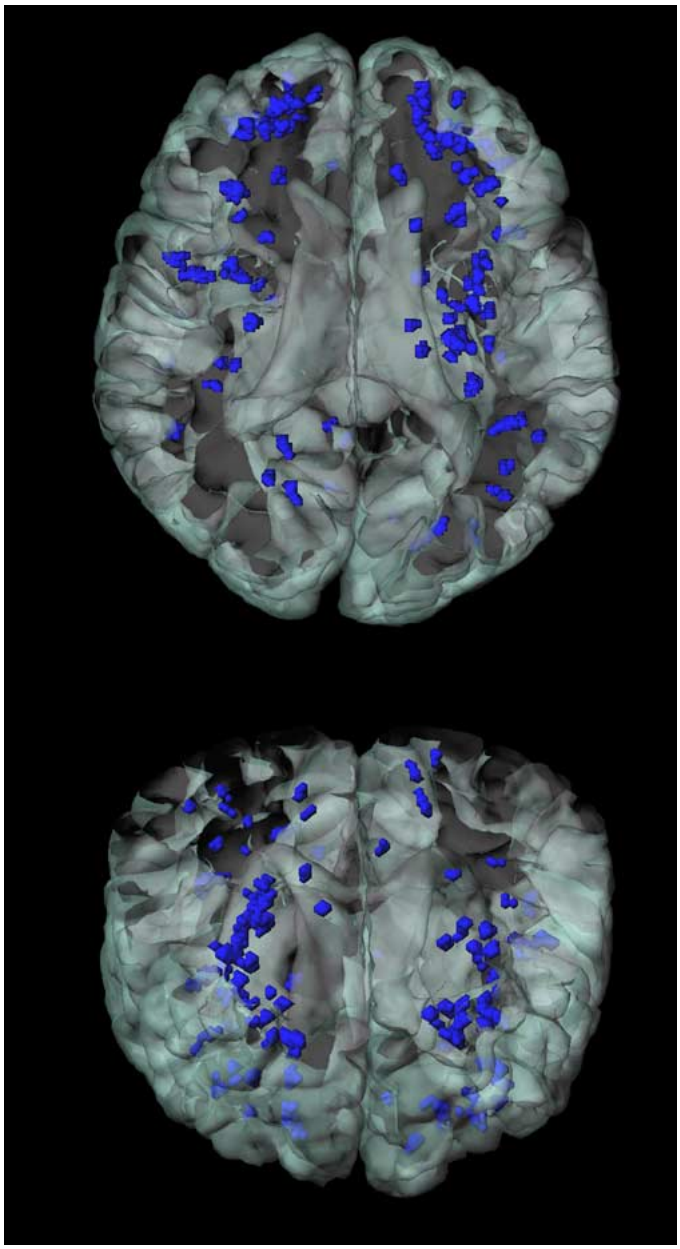


Fig. 8. Detected VRS in relation to the brain surface in the example dataset of Fig. 7, views from top and top-frontal.

B. Evaluation on a Database

Now, we report about results obtained on a database of 37 subjects. To compare with current practice in neurobiology, datasets were rated by an independent expert (HJG) for the

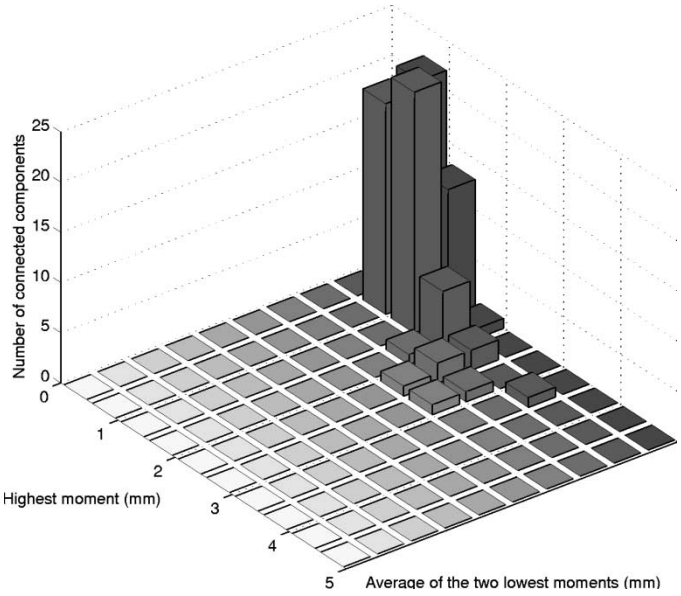


Fig. 9. Moments of the connected components. X axis: average value of the two lowest moments; y axis: Highest moment; and z axis: Number of connected components.

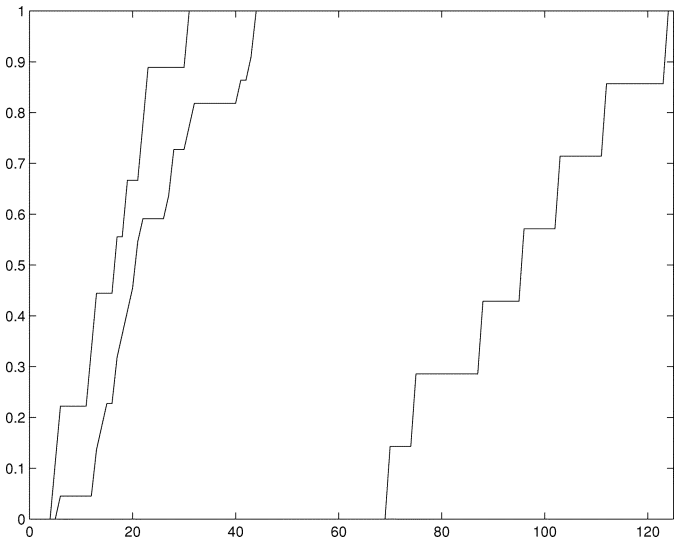


Fig. 10. Cumulative distribution functions of detected tubes (x axis) versus expert rating level 1 (left curve), 2 (middle curve), and 3 (right curve).

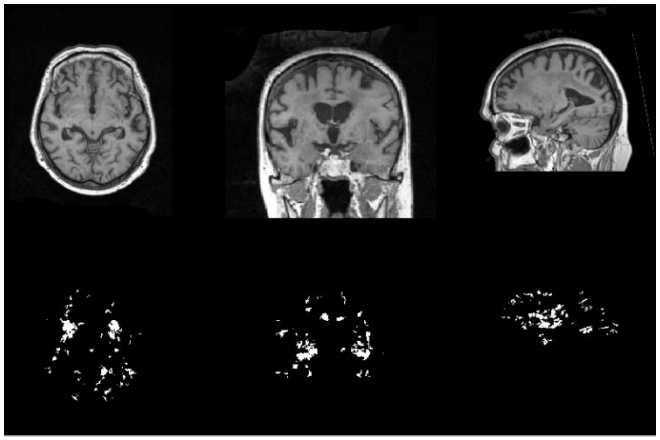


Fig. 11. Cumulative spatial distribution of VRS collected from 37 datasets, together with the reference brain. A higher prevalence of VRS is found in the area supplied by the medial and lateral striate arteries (e.g., putamen and pallidum) and in the area supplied by the long penetrating arteries, especially in the frontal WM compartment.

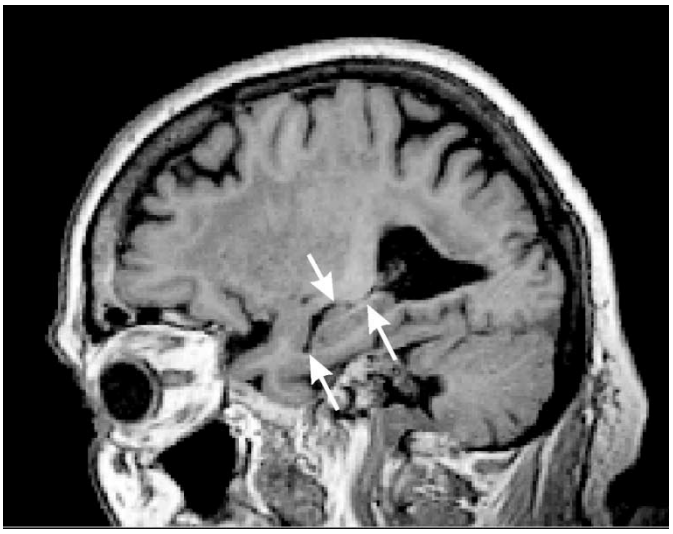


Fig. 12. Sagittal slice through the temporal horn of the lateral ventricle. Arrows mark locations where false positive detections may occur.

presence of VRS using a semiquantitative rating scale [22]: 9 cases as 1 (corresponding to a few VRS), 22 cases as 2, and 8 cases as 3 (corresponding to many VRS). The number of connected components as determined by our approach was considered for comparison. We used the same parameters (as defined in Section V-A) for all datasets. Fig. 10 shows a plot of the cumulative distributions of detected VRS count for the datasets rated as 1 (left curve), 2 (middle curve), and 3 (right curve). The subset rated as 3 is clearly discriminated. Indeed, the minimum obtained value is 70 connected components whereas the maximum value of the two other subsets is 44. Although there is a reasonable shift between the cumulative distributions corresponding to the subsets 1 and 2, we observe an overlap between the distributions. A visual inspection of the results indicates that the overlap between level 1 and 2 is also due to a certain amount of rater bias. A linear regression between the ratings and the VRS count yields a correlation coefficient of 0.77 ($p \sim 1e-08$).

All 37 brain datasets were registered to the same brain using a nonlinear procedure based on fluid dynamics [4], [28]. The obtained displacement fields were used to transform the detected VRS into a common space. Fig. 11 shows the cumulative spatial distribution of VRS collected from all datasets, together with the reference brain. A higher prevalence of VRS is found in the area supplied by the medial and lateral striate arteries (e.g., putamen and pallidum) and in the area supplied by the long penetrating arteries, especially in the frontal WM compartment.

Since VRS lesions are at the resolution limit of the scanner and often hard to discriminate from other lesion types (e.g., lacunes), a formal validation requires the comparison of high-resolution MRI data (at 0.25 mm resolution or better) with histological slices. Such analyses are currently performed at MPI. As an informal evaluation of the method's performance, VRS were marked manually in the 7 datasets rated as 3 by two experts in two separate runs. An average of 597 lesions were found

by the experts, compared with 668 automatically detected VRS. A type I error of 13.0% (false positives) and a type II error of 2.7% (false negatives) was determined. The intraclass correlation coefficient (ICC [20]) between both experts was 0.96, within experts 0.97, the ICC between experts and the automatic detection was 0.87. About half of the false positive detections (40) were found at locations that resemble VRS but belong to healthy structures (e.g., at the anterior tip of the temporal horn of the lateral ventricle, see Fig. 12). Even with normal datasets (rated as 0), false positive detections at these sites may occur. It is suggested that atlas information may be used as an additional prior to improve the specificity. The other false positive detections (31) were addressed to other pathologic features of these datasets (e.g., lacunar infarctions).

VI. Conclusion

We applied a marked point process framework to derive a model for multiple small brain lesions detection that was optimized by a RJMCMC algorithm. A special feature of this approach is the object based modeling. In our application example, VRS are not considered at a voxel level but as geometrical shapes. This induces robustness with respect to noise and, thus, a good detection sensitivity. The sensitivity is further improved by the prior model which embeds specific interactions. The clustering interaction appears to be well adapted to VRS. Future work will consist of improving the specificity further by removing false positives.

The proposed model involves several parameters. The parameters defining the data term are derived from the intensity statistics of the dataset provided that a segmentation into the compartments CSF, WM, and GM is available. The prior model parameters and the hyper-parameters weighting the prior model and the data term were calibrated. Estimating these weights would be an improvement toward a fully automatic approach. However, using the same parameter values for all the datasets yielded satisfactory results which proves the robustness of the approach with respect to the parameters involved. Parameters of the prior model are relatively data independent. Therefore, spending extra computation time for parameter estimation appears unnecessary here.

The usual drawback of MCMC algorithms is the computation time necessary for model optimization, which may become critical in this 3-D case. Therefore, we have reduced the volume by using the filters defining the data term. Depending on the number of detected VRS, the computation time varies between 10 min and 2 hours on a personal computer (Intel 750-MHz processor). The optimal choice of the weighting between the different moves in the RJMCMC algorithm is considered an open issue.

Future work is twofold: we will study the correlation of VRS with clinical variables to address their significance for the process of aging. Second, this model may be modified for similar problems, such as detecting small lesions in Multiple Sclerosis. Our approach is quite general and suitable for detecting multiple lesions. However, for other lesion types, the object geometry and object relations in space must be revised and adapted to the target application.

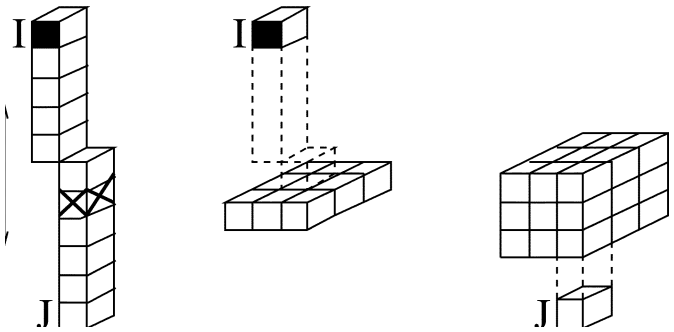


Fig. 13. Splitting move. (a) Initial segment $\rightarrow IJ$, the arrow indicates candidate points for splitting ($\rho_{\min} = 3$). (b) First resulting segment, the initial point I and the 3×3 candidates for the final point. (c) Second resulting segment, the final point J and the candidates for the initial point ($D_m = 1$).

Appendix

In this appendix, we derive the acceptance ratio for the different moves involved in the proposed RJMCMC algorithm. We denote the current configuration $x = \{\vec{s(1)} = (x(1), \vec{l(1)}), \ldots, \vec{s(n)} = (x(n), \vec{l(n)})\}$.

A. Birth and Death

This move consists in adding a new segment η with probability P_b or removing a segment ξ from the configuration with probability $P_d = 1 - P_b$. In the case of a death move, the segment is chosen uniformly among the segments of the current configuration and uniformly on the parameter space for the a birth. The birth kernel is given by $Q(x \rightarrow dx' = x \cup d\eta)$. Denote $\vec{s^{\text{new}}} = (x_{n+1}, \vec{l(n+1)})$, we then have

$$q(\vec{s^{\text{new}}}) = \frac{1}{V} \frac{1}{4\pi^2(\rho_{\max} - \rho_{\min})}. \tag{15}$$

The transformation from (x, s^{new}) to x' is the identity, so its Jacobian is equal to 1. The inverse transition consists in removing the segment s^{new} from the $n+1$ segments contained in x'. The acceptance ratio is, therefore, given by

$$\alpha_{\text{birth}} = \frac{P_d}{P_b} \frac{h_T(x')}{h_T(x)} \frac{4\pi^2(\rho_{\max} - \rho_{\min})V}{n+1}. \tag{16}$$

Applying the same considerations, the death kernel $Q(x \rightarrow x' = x/\xi)$ has the acceptance ratio

$$\alpha_{\text{death}} = \frac{P_b}{P_d} \frac{h_T(x')}{h_T(x)} \frac{n}{4\pi^2(\rho_{\max} - \rho_{\min})V}. \tag{17}$$

B. Move a Segment

We choose a segment uniformly within the configuration and select a new location and new marks uniformly within the parameter space. This move is identical to its reverse. Therefore, we have

$$\alpha_{\text{mvseg}} = \frac{h_T(x')}{h_T(x)}. \tag{18}$$

$$\max(|x_u(2)-x_u(1)-l_u(1)|,|x_v(2)-x_v(1)-l_v(1)|,|x_w(2)-x_w(1)-l_w(1)|)\leq D_m$$

$$\max(|x_u(2)+l_u(2)-x_u(1)|,|x_v(2)+l_v(2)-x_v(1)|,|x_w(2)+l_w(2)-x_w(1)|)\leq D_m$$

C. Move the End Point

We choose a new location of the end point uniformly in a neighborhood of its current position. As for the previous move, the acceptance ratio is given by

$$\alpha_{\text{mvend}}=\frac{h_T(x')}{h_T(x)}. \tag{19}$$

D. Split and Merge

This move merges neighboring segments. The reverse move corresponds to splitting a segment in two parts. A merging is proposed with probability P_m and a splitting with probability $P_s=1-P_m$.

The dimension matching is not easily achieved here. We say that two segments $\vec{s(1)}=(x(1),\vec{l(1)})$ and $\vec{s(2)}=(x(2),\vec{l(2)})$ are candidates for merging if the first or second equation shown at the top of the page holds. The merged segment is given by $(x^{\text{new}}=x(1),\vec{l^{\text{new}}}=(x(1),x(2)+\vec{l(2)}))$ in the former case and $(x^{\text{new}}=x(1),\vec{l^{\text{new}}}=(x(2),x(1)+\vec{l(1)}))$ in the latter case.

To define splitting we have to consider the discrete nature of the data (see Fig. 13). We first define the voxel where the segment $\vec{s}=(x,\vec{l})=\vec{xy}$, where $y=x+\vec{l}$, will be split. To obtain the discrete segment $S(\vec{s})$ we compute the direction corresponding to the maximum value of $(|l_u|,|l_v|,|l_w|)$. Let us assume that the maximum is obtained for u. Then, we have $E(|l_u|)$ voxel in $S(\vec{s})$ (E represents the integer part). Candidate split points p are such that $\|\vec{xp}\|\geq\rho_{\min}$ and $\|\vec{py}\|\geq\rho_{\min}$. We first uniformly select a point p within the candidate points and denote $\nu=\|\vec{xp}\|/\|\vec{l}\|$. In practice, p is chosen among point of $S(\vec{s})$ for which $|p_u-x_u|\geq\rho_{\min}$ and $|p_u-y_u|\geq\rho_{\min}$, so that $\nu\in[\rho_{\min}/E(|l_u|),1-\rho_{\min}/E(|l_u|)]$. The segment $\vec{s}$ is then split into $\vec{xy'}$ and $\vec{x'y}$ with $y'=p+q$, $x'=y'+r$, where $q_u=0$ and q_v,q_w,r_u,r_v,r_w are uniformly chosen in $\{-D_m,-D_m+1,\ldots,0,\ldots D_m-1,D_m\}$. Therefore, the splitting transformation $\mathcal{T}$ can be summarized as follows (omitting the identity terms):

$$\begin{aligned}
y'_u&=\nu y_u+(1-\nu)x_u\\
y'_v&=\nu y_v+(1-\nu)x_v+q_v\\
y'_w&=\nu y_w+(1-\nu)x_w+q_w\\
x'_u&=\nu y_u+(1-\nu)x_u+r_u\\
x'_v&=\nu y_v+(1-\nu)x_v+q_v+r_v\\
x'_w&=\nu y_w+(1-\nu)x_w+q_w+r_w.
\end{aligned} \tag{20}$$

The Jacobian of $\mathcal{T}$ is then given by

$$|\mathcal{J}(\mathcal{T})|=|l_u|. \tag{21}$$

The density of the auxilliary variables are given by

$$q(\nu)=\frac{|l_u|-2*\rho_{\min}}{|l_u|} \tag{22}$$

and

$$q(q_v)=q(q_w)=q(r_u)=q(r_v)=q(r_w)=\frac{1}{2D_m+1}. \tag{23}$$

The acceptance ratio for the split move is written as follows:

$$\alpha_{\text{split}}=\frac{P_m}{P_s}\frac{h_T(x')}{h_T(x)}\frac{(2D_m+1)^5\,(|l_u|-2*\rho_{\min})}{n_m+1} \tag{24}$$

where n_m is the number of candidate pairs for merging. The merging acceptance ratio is the reverse of the splitting acceptance ratio. In practice, we have chosen $D_m=1$.

E. Return a Segment

The segment $\vec{xy}$ becomes $\vec{yx}$ provided that y belongs to V. The reverse move is the identical and we have

$$\alpha_{ret}=\frac{h_T(x')}{h_T(x)}. \tag{25}$$

Acknowledgment

The authors would like to acknowledge that the acquisition of the datasets was supported by the project C-08 of the Interdisciplinary Centre for Clinical Research at the University Clinic, Leipzig, Germany.

References

[1] A. J. Baddeley and M. N. M. van Lieshout, "ICM for object recognition," *Computational Statist.*, vol. 2, pp. 271–286, 1992.

[2] ——, "Stochastic geometry models in high-level vision," in *Statist. Images*, K. V. Mardia and G. K. Kanji, Eds., 1993, vol. 1, J. Appl. Stat., pp. 231–256.

[3] H. Bokura, S. Kobayashi, and S. Yamaguchi, "Distinguishing silent lacular infarction from enlarged Virchow-Robin spaces: A magnetic resonance imaging and pathological study," *J. Neurol.*, vol. 245, pp. 116–122, 1998.

[4] G. Christensen, "Deformable Shape Models for Anatomy," Ph.D., Washington University, St. Louis, 1994.

[5] X. Descombes, R. Stoica, L. Garcin, and J. Zerubia, "A RJMCMC algorithm for object processes in image processing," *Monte Carlo Meth. Applicat.*, vol. 7, pp. 149–156, 2001.

[6] J. A. Fiez, H. Damasio, and T. J. Grabowski, "Lesion segmentation and manual warping to a reference brain: Intra- and interobserver reliability," *Human Brain Mapping*, vol. 9, pp. 191–211, 2000.

[7] G. Gerig, D. Welti, C. R. G. Guttmann, A. C. F. Colchester, and G. Szekely, "Exploring the discrimination power of time domain for segmentation and characterization of active lesions in serial MR data," *Med. Imag. Anal.*, vol. 4, pp. 31–42, 2000.

[8] P. Green, "Reversible jump MCMC computation and Bayesian model determination," *Biometrika*, vol. 82, pp. 711–732, 1995.

[9] U. Grenander, *General Pattern Theory*. Oxford, U.K.: Oxford Univ. Press, 1993.

[10] L. A. Heier, C. J. Bauer, L. Schwartz, R. D. Zimmerman, S. Morgelli, and M. D. Deck, "Large Virchow-Robin spaces: MR-clinical correlation," *Amer. J. Neurorad.*, vol. 10, pp. 929–936, 1989.

[11] N. Hirabuki, N. Fujita, K. Fujii, T. Hashimoto, and T. Kozuka, "MR appearance of virchow robin spaces along lenticulostriate arteries spin echo and 2 dimensional fast low angle shot imaging," *Amer. J. Neurorad.*, vol. 15, pp. 277–281, 1994.

[12] S. A. Hojjatoleslami and J. Kittler, "Region growing: A new approach," *IEEE Trans. Imag. Proc.*, vol. 7, pp. 1079–1083, July 1998.

[13] S. A. Hojjatoleslami and F. Kruggel, "Segmentation of large brain lesions," *IEEE Trans. Med. Imag.*, vol. 20, pp. 666–669, July 2001.

[14] M. A. Hurn, "Confocal fluorescence microscopy of leaf cells: An application of Bayesian image analysis," *J. Roy. Statist. Soc.*, ser. C, vol. 47, pp. 361–377, 1998.

[15] M. Kamber, R. Shinghai, D. L. Collins, G. S. Francis, and A. C. Evans, "Model-based 3-D segmentation of multiple sclerosis lesion in magnetic resonance brain images," *IEEE Trans. Med. Imag.*, vol. 14, pp. 442–453, Sept. 1995.

[16] F. Kruggel and D. Y. von Cramon, "Alignment of magnetic-resonance brain datasets with the stereotactical coordinate system," *Med. Imag. Anal.*, vol. 3, pp. 1–11, 1999.

[17] A. D. Lanterman, M. I. Miller, and D. L. Snyder, "Implementation of jump-diffusion algorithms for understanding FLIR scenes," *Proc. SPIE*, vol. 2485, pp. 309–320, 1995.

[18] S. Loncaric, A. P. Dhawan, D. Cosic, D. Kovacevic, J. Broderick, and T. Brott, "Quantitative intracerebral brain hemorrhage analysis," *Proc. SPIE*, vol. 3661, pp. 886–894, 1999.

[19] C. Pachai, Y. M. Zhu, J. Grimmaud, M. Hermier, A. Dromigny-Badin, A. Boudraa, G. Gimenez, C. Confavreux, and J. C. Froment, "A pyramidal approach for automatic segmentation of multiple sclerosis lesions in brain MRI," *Comp. Med. Imag. Graphics*, vol. 22, pp. 399–408, 1998.

[20] L. G. Portney and M. P. Watkins, *Foundations of Clinical Research: Applications and Practice*. Norwalk, CT: Appleton & Lange, 1993.

[21] H. Rue and A. R. Syverseen, "Bayesian object recognition with Baddeley's delta loss," *Adv. Appl. Prob.*, vol. 30, pp. 64–84, 1998.

[22] P. Scheltens, T. Erkinjunti, D. Leys, L. O. Wahlund, D. Inzitari, T. del Ser, F. Pasquier, F. Barkhof, R. Mäntylä, J. Bowler, A. Wallin, J. Ghika, F. Fazekas, and L. Pantoni, "White matter changes on CT and MRI: An overview of visual rating scales," *Eur. Neurol.*, vol. 39, pp. 80–89, 1998.

[23] R. Stoica, X. Descombes, and J. Zerubia, "Road extraction in remote sensed images using a stochastic geometry framework," in *Proc. Int. Workshop Bayesian Inference and Maximum Entropy Methods*, Gif-sur-Yvette, France, 2000.

[24] Y. Tao, W. I. Grosky, L. Zamorano, Z. Jiang, and J. Gong, "Segmentation and representation of lesions in the MRI brain images," *Proc. SPIE*, vol. 3661, pp. 930–939, 1999.

[25] P. Thevenaz, T. Blu, and M. Unser, "Image interpolation and resampling," in *Handbook of Medical Image Processing*. Los Alamitos, CA: IEEE Press, 1999.

[26] J. K. Udupa, L. Wei, S. Samaraseka, Y. Miki, M. A. van Buchem, and R. I. Grossman, "Multiple sclerosis lesion quantification using fuzzy-connectedness principles," *IEEE Trans. Med. Imag.*, vol. 16, pp. 598–608, Oct. 1997.

[27] H. Wolf, M. Grunwald, F. Kruggel, S. G. Riedel-Heller, S. Angerhöfer, S. A. Hojjatoleslami, A. Hensel, T. Arendt, and H. J. Gertz, "Hippocampal volume discriminates between normal condition, questionable and mild dementia in the elderly," *Neurobiol. Aging*, vol. 22, pp. 177–186, 2001.

[28] G. Wollny and F. Kruggel, "Computational cost of nonrigid registration algorithms based on fluid dynamics," *IEEE Trans. Med. Imag.*, vol. 21, pp. 946–952, Aug. 2002.

[29] A. P. Zijdenbos, B. M. Dawant, R. A. Margolin, and A. C. Palmer, "Morphometric analysis of white matter lesions in MR images: Method and validation," *IEEE Trans. Med. Imag.*, vol. 13, pp. 716–724, Dec. 1994.