A Graphical Model Framework for Image Segmentation

Rui Huang, Vladimir Pavlovic and Dimitris N. Metaxas

Summary. Graphical models are probabilistic models defined in terms of graphs. The intuitive and compact graph representation and its ability to model complex probabilistic systems make graphical models a powerful modeling tool in various research areas. In this paper we introduce a graphical model framework for image segmentation based on the integration of Markov random fields (MRFs) and deformable models. A graphical model is constructed to represent the relationship of the observed image pixels, the true region labels and the underlying object contour. We then formulate the problem of image segmentation as the one of joint region-contour inference and learning in the graphical model. The graphical model representation allows us to use an approximate structured variational inference technique to solve this otherwise intractable joint inference problem. Using this technique, the MAP solution to the original model is obtained by finding the MAP solutions of two simpler models, an extended MRF model and a probabilistic deformable model, iteratively and incrementally. In the extended MRF model, the true region labels are estimated using the BP algorithm in a band area around the estimated contour from the probabilistic deformable model, and the result in turn guides the probabilistic deformable model to an improved estimation of the contour. Finally, we generalize our method from 2D to 3D. Experimental results on both synthetic and real images, in both 2D and 3D, show that our new hybrid method outperforms both the MRF-based and the deformable model-based methods using only homogeneous constraints.

1 Introduction

Graphical models are a marriage between probability theory and graph theory [1]. A graphical model is a probabilistic model defined in terms of a graph in which the nodes represent random variables and the edges describe the probabilistic relationships among these variables. In particular, these probabilistic relationships are usually defined by conditional probabilities among the related variables or potential functions on the cliques of the graph, depend on whether the graph is directed or undirected. The joint probability distribution of a set of variables or the whole system can then be computed by taking products over the functions defined on relevant nodes. The graph theoretic side of graphical models provides an intuitive and compact representation for the complex probabilistic system, as well as well-defined data structures and efficient general-purpose algorithms. Probability theory, on the other

R. Huang et al.: *A Graphical Model Framework for Image Segmentation*, Studies in Computational Intelligence (SCI) **52**, 43–63 (2007)
www.springerlink.com

hand, ensures the consistency of the whole system, and provides various statistical inference and learning methods to analyze the data.

Graphical models have recently received extensive attention from many different research communities, including artificial intelligence, machine learning, computer vision, etc. In this paper, we apply graphical models to the image segmentation problem, one of the most important and difficult tasks in computer vision area. We are able to integrate two fundamentally different traditional segmentation methods and take advantage of both using graphical models. Furthermore, the graphical model theory allows us to employ an approximate, computationally efficient solution to the otherwise intractable inference problem. We will focus on the graphical model representation and inference (mainly approximate inference) techniques for image segmentation. See [2] for a more comprehensive introduction to graphical models and [1] for more advanced topics.

The rest of this paper is organized as follows: Section 2 defines the segmentation problem and reviews the previous work; Sect. 3 introduces a new integrated model and its decoupled approximation using the variational inference method; detailed inferences on the decoupled models are described in Sect. 4; the 2D model is then generalized to 3D in Sect. 5; Sect. 6 shows the experimental results on both synthetic and real 2D images and 3D volumes; and finally Sect. 7 summarizes the paper.

2 Previous Work

Image segmentation is one of the most important and difficult preliminary processes for high-level computer vision and pattern recognition problems. The main goal of image segmentation is to divide an image into its constituent parts that have a strong correlation with objects or areas of the real world depicted by the image.

Region-based and edge-based segmentations are the two major classes of segmentation methods. Though one can label regions according to edges or detect edges from regions, these two kinds of methods are naturally different and have respective advantages and disadvantages.

Region-based methods assign image pixels to a region according to some image property (e.g., region homogeneity). These methods work well in noisy images, where edges are usually difficult to detect while the region homogeneity is preserved. The disadvantages of region-based methods are that they may generate rough edges and holes inside the objects, and they do not take account of object shape.

On the other hand, edge-based methods generate boundaries of the segmented objects. A prior knowledge of object shape and topology can be easily incorporated to constrain the segmentation result. While this often leads to sufficiently smooth boundaries, the oversmoothing may be excessive. Because edge-based methods rely on edge detecting operators, they are sensitive to image noise and need to be initialized close to the actual region boundaries.

Most segmentation methods are either region-based or edge-based. Among region-based methods, besides the classical region growing method [3], the Markov random field (MRF) model has been extensively used. Because the exact MAP

inference in MRF models is computationally infeasible, various techniques for approximating the MAP estimation have been proposed, such as Markov Chain Monte Carlo (MCMC) [4], iterated conditional modes (ICM) [5], maximizer of posterior marginals (MPM) [6], etc. [7] presents a comparative analysis of some of these methods. Two of the more recent algorithms, Belief Propagation (BP) [8,9] and Graph Cuts [10] are compared in [11]. The estimation of the MRF model parameters is another related problem, often solved using the EM algorithm [12].

In edge-based methods, since Kass et al. introduced Snakes [13], deformable models have attracted much attention. Variants of deformable models have been proposed to address different problems. For instance, Balloons [14] and Gradient Vector Flow (GVF) Snakes [15] introduces different external forces, and Topologically Adaptable Snakes [16] allow changes in the model's topology. See [17] for a review of deformable models and [3] for other edge-based methods and some basic edge detecting operators.

Hybrid approaches [18–20] attempt to combine region-based and edge-based segmentations to alleviate deficiencies of the individual methods and improve the segmentation results. There are different choices of the combination. For instance, [20] proposes a way of integrating MRFs and deformable models. MRFs are used to initially estimate the boundary of objects in noisy images. Balloons are then fitted to the estimated boundary. The result of the fitting is in turn used to update the MRF parameters. Final segmentation is achieved by iteratively integrating these processes. While this hybrid method attempted to take advantage of both MRFs and deformable models, the model coupling was loose. This may cause failure of deformable models if the initial estimation of the boundary by MRF is not closed, and it may also yield oversmoothed boundaries.

We propose a new framework to combine the MRF-based and the deformable model-based segmentation methods. To tightly couple the two models, we construct a graphical model to represent the relationship of the observed image pixels, the true region labels and the underlying object contour. Exact inference in the graphical model is intractable because of the large state spaces and the couplings of model variables. To tackle this problem we use a variational inference method to seemingly decouple the graphical model into two simpler models: one extended MRF model and one probabilistic deformable model. Then we obtain the MAP solution in the original model by solving the MAP problems of the two simpler models iteratively and incrementally. In the extended MRF model, the true region labels are estimated using the BP algorithm in a band area around the estimated contour from the probabilistic deformable model, and the result in turn guides the probabilistic deformable model to an improved estimation of the contour.

3 Our Method

The goal of our segmentation method is to find one specific region with a smooth and closed boundary. A seed point is manually specified and the region containing it is then segmented automatically. Thus, without significant loss of modeling generality,

we simplify the MRF model and avoid possible problems caused by segmenting multiple regions simultaneously.

In this section, we first briefly review MRFs and deformable models, define the notation, and then introduce our hybrid framework.

3.1 MRF-Based Segmentation

MRF models are a special case of undirected graphical models. They are often used for image analysis, because of their ability to capture the context of an image (i.e., dependencies among neighboring image pixels) and deal with the noise.

A typical MRF model for image segmentation, as shown in Fig. 1, is a graph with two types of nodes: observable nodes (shaded nodes in Fig. 1, representing image pixel values) and hidden nodes (clear nodes in Fig. 1, representing region labels). The edges in the graph depict the relationships among the nodes.

Let n be the number of the hidden/observable states (i.e., the number of pixels in the image). A configuration of the hidden layer is:

$$\mathbf{x} = (x_1, ..., x_n), x_i \in L, i = 1, ..., n$$

where L is a set of region labels, such as $L = \{inside, outside\}$.

Similarly, a configuration of the observable layer is:

$$\mathbf{y} = (y_1, ..., y_n), y_i \in D, i = 1, ..., n$$

where D is a set of pixel values, e.g., gray values 0–255.

The relationship between the hidden states and the observable states (also known as local evidence) can be described by the potential (or compatibility) function: $\phi(x_i, y_i)$, which is often a conditional Gaussian to handle the image noise; the relationship between the neighboring hidden states is described by the second potential function: $\psi(x_i, x_j)$, which usually penalizes differences between the states to keep region smoothness. The detailed definitions will be discussed later.

Now the segmentation problem can be viewed as a problem of estimating the MAP solution of the MRF model:

$$\mathbf{x}_{\mathrm{MAP}} = \arg\max_{\mathbf{x}} P(\mathbf{x}|\mathbf{y}) \tag{1}$$

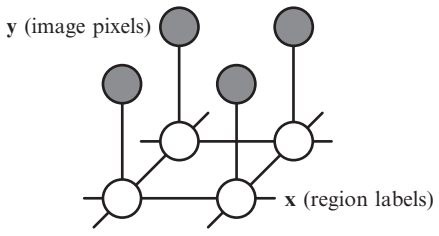


Fig. 1. MRF model

where

$$P(\mathbf{x}|\mathbf{y}) \propto P(\mathbf{y}|\mathbf{x})P(\mathbf{x}) \propto \prod_i \phi(x_i, y_i) \prod_{(i,j)} \psi(x_i, x_j)$$

As mentioned previously, the exact MAP inference in MRFs is computationally infeasible, and various techniques have been used for approximating the MAP estimation. In our method, we use the BP algorithm. The MRF model parameters (i.e., the parameters in the potential functions) are learned using the EM algorithm. However, in the presence of multiple regions in the image, the automatic determination of the number of regions and the initial guess of the parameters could be difficult. More importantly, like other region-based methods, MRFs do not take account of object shape and may generate rough edges and even holes inside the objects.

3.2 Deformable Model-Based Segmentation

Many deformable model-based methods have also been used in image segmentation. A deformable model is usually a parameterized geometric primitive, whose deformation is determined by geometry, kinematics, dynamics, and other constraints (e.g., material properties, etc.) [21]. Snakes [13], a special case of deformable models, are a parametric contour:

$$\begin{aligned} \Omega = [0,1] &\rightarrow \Re^2 \\ s &\rightarrow \mathbf{c}(s) = (x(s), y(s)) \end{aligned}$$

where s is the parametric domain and x and y are the coordinate functions. The energy of the contour:

$$E(\mathbf{c}) = E_{\text{int}}(\mathbf{c}) + E_{\text{ext}}(\mathbf{c}) = \int_\Omega \omega_1(s) \left|\frac{\partial \mathbf{c}}{\partial s}\right|^2 + \omega_2(s) \left|\frac{\partial^2 \mathbf{c}}{\partial s^2}\right|^2 + F(\mathbf{c}(s))ds$$

where $\omega_1(s)$ and $\omega_2(s)$ control the "elasticity" and "rigidity" of the contour, and F is the potential associated to the external forces. The final shape of the contour corresponds to the minimum of this energy.

To minimize the above energy term, one can use the discretized first-order Lagrangian dynamics equation:

$$\dot{\mathbf{d}} + \mathbf{K}\mathbf{d} = \mathbf{f}$$

where $\mathbf{d}$ is discretized version of $\mathbf{c}$, $\mathbf{K}$ is the stiffness matrix calculated from $\omega_1(s)$ and $\omega_2(s)$, and $\mathbf{f}$ is the generalized force vector.

Image gradient forces are usually used to attract a deformable model to edges. However, when far from the true boundary, the model often gets attracted to spurious image edges. Balloon forces have been introduced to solve this problem [14]. Namely, the deformable model is considered a balloon, which is inflated by an additional force and stopped by strong edges. The initial contour need no longer be close to the true boundary. Mathematically, a force along the normal direction to the curve at point $\mathbf{c}(s)$ with some appropriate amplitude k is added to the original forces.

$$\mathbf{f}' = \mathbf{f} + k\overrightarrow{\mathbf{n}}(s)$$

Deformable models can also be viewed in a probabilistic framework [17]. The internal energy $E_{\text{int}}(\mathbf{c})$ leads to a Gibbs prior distribution of the form:

$$P(\mathbf{c}) = \frac{1}{Z_i}\exp(-E_{\text{int}}(\mathbf{c})) \tag{2}$$

while the external energy $E_{\text{ext}}(\mathbf{c})$ can be converted to a sensor model with conditional probability:

$$P(\mathbf{I}|\mathbf{c}) = \frac{1}{Z_e}\exp(-E_{\text{ext}}(\mathbf{c})) \tag{3}$$

where $\mathbf{I}$ denotes the image, and $E_{\text{ext}}(\mathbf{c})$ is a function of the image $\mathbf{I}$.

The deformable models can now be fitted by solving the MAP problem:

$$\mathbf{c}_{\text{MAP}} = \arg\max_{\mathbf{c}} P(\mathbf{c}|\mathbf{I}) \tag{4}$$

where

$$P(\mathbf{c}|\mathbf{I}) \propto P(\mathbf{c})P(\mathbf{I}|\mathbf{c})$$

One limitation of the deformable model-based method is its sensitivity to image noise, a common drawback of edge-based methods. This may result in the deformable model being "stuck" in a local energy minimum of a noisy image.

3.3 Integrated Model

As shown in (1) and (4), both the MRF-based and the deformable model-based segmentations can be viewed as the MAP estimation problems. In previous work [20], these two models were loosely coupled. Our new framework uses the graphical model theory to tightly couple the two models. This is achieved, as depicted in Fig. 2, by adding a new hidden state to the traditional MRF model to represent the underlying contour.

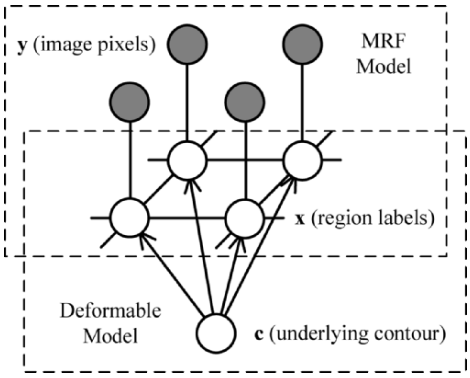


Fig. 2. Integrated model

In the new model, the segmentation problem can also be viewed as a *joint* MAP estimation problem:

$$(\mathbf{c}, \mathbf{x})_{\mathrm{MAP}} = \arg\max_{\mathbf{c},\mathbf{x}} P(\mathbf{c}, \mathbf{x}|\mathbf{y})$$

where

$$P(\mathbf{c}, \mathbf{x}|\mathbf{y}) \propto P(\mathbf{y}|\mathbf{x})P(\mathbf{x}|\mathbf{c})P(\mathbf{c})$$

To define the joint distribution of the integrated model, we model the image likelihood term $P(\mathbf{y}|\mathbf{x})$ as:

$$P(\mathbf{y}|\mathbf{x}) = \prod_i \phi(x_i, y_i)$$

identical to the traditional MRF model. The second term $P(\mathbf{x}|\mathbf{c})$, modeling the distribution of the region labels conditioned on the contour, is defined as:

$$P(\mathbf{x}|\mathbf{c}) = \prod_{(i,j)} \psi(x_i, x_j) \prod_i P(x_i|\mathbf{c})$$

where we incorporated a shape prior $\mathbf{c}$ to constrain the region labels $\mathbf{x}$, in addition to the original Gibbs distribution.

Since we only segment one specific region at one time, we need only consider the pixels near the contour, and label them either *inside* or *outside* the contour.

We model the dependency between the contour $\mathbf{c}$ and the region labels $\mathbf{x}$ using the softmax function:

$$P(x_i = inside|\mathbf{c}) = \frac{1}{1 + \exp(-dist(i, \mathbf{c}))} \tag{5}$$

$$P(x_i = outside|\mathbf{c}) = 1 - P(x_i = inside|\mathbf{c}) \tag{6}$$

induced by the signed distance of pixel i from the contour $\mathbf{c}$:

$$dist(i, \mathbf{c}) = sign(i) \min_{s \in \Omega} \|loc(i) - \mathbf{c}(s)\| \tag{7}$$

where $sign(i) = 1$ if pixel i is inside contour $\mathbf{c}$, $sign(i) = -1$ when it is outside, and $loc(i)$ denotes the spatial coordinates of pixel i.

Lastly, the prior term $P(\mathbf{c})$, as in (2), can be represented as a Gibbs distribution when the shape prior is given by a parametric contour $\mathbf{c}$.

Despite the compact graphical representation of the integrated model, the exact inference in the model is computationally intractable. One reason for this is the large state space size and the complex dependency structure introduced by the Gibbs distribution of the prior $P(\mathbf{c})$. The second reason is the existence of loops in the graphical model, which preclude polynomial-time inference. To deal with these problems we propose an approximate, yet tractable, solution based on structured variational inference.

3.4 Approximate Inference Using Structured Variational Inference

Structured variational inference techniques [22, 23] consider parameterized distribution which is in some sense close to the desired posterior distribution, but is easier to compute. Namely, for a given image $\mathbf{y}$, a distribution $Q(\mathbf{c}, \mathbf{x}|\mathbf{y}, \theta)$ with an additional set of *variational parameters* θ is defined such that the Kullback–Leibler (KL) divergence between $Q(\mathbf{c}, \mathbf{x}|\mathbf{y}, \theta)$ and $P(\mathbf{c}, \mathbf{x}|\mathbf{y})$ is minimized with respect to θ:

$$\theta^* = \arg\min_{\theta} \sum_{\mathbf{c},\mathbf{x}} Q(\mathbf{c}, \mathbf{x}|\mathbf{y}, \theta) \log \frac{P(\mathbf{c}, \mathbf{x}|\mathbf{y})}{Q(\mathbf{c}, \mathbf{x}|\mathbf{y}, \theta)}$$

The dependency structure of Q is chosen such that it closely resembles the dependency structure of the original distribution P. However, unlike P the dependency structure of Q *must* allow a computationally efficient inference.

In our case we define Q by decoupling the MRF model and the deformable model components of the original integrated model in Fig. 2. The original distribution is factorized into two independent distributions: an extended MRF model Q_M with variational parameter **a** and a probabilistic deformable model Q_D with variational parameter **b** (Fig. 3). The *extended* MRF model means we have an additional layer to the traditional MRF model to deal with the shape prior, and the *probabilistic* deformable model means the contour is not fitted to the image directly, but to the probabilistic label image.

Because Q_M and Q_D are independent,

$$Q(\mathbf{c}, \mathbf{x}|\mathbf{y}, \mathbf{a}, \mathbf{b}) = Q_M(\mathbf{x}|\mathbf{y}, \mathbf{a}) Q_D(\mathbf{c}|\mathbf{b})$$

According to the extended MRF model, we have:

$$Q_M(\mathbf{x}|\mathbf{y}, \mathbf{a}) \propto Q_M(\mathbf{y}|\mathbf{x}) Q_M(\mathbf{x}|\mathbf{a})$$

$$Q_M(\mathbf{y}|\mathbf{x}) = \prod_i \phi(x_i, y_i)$$

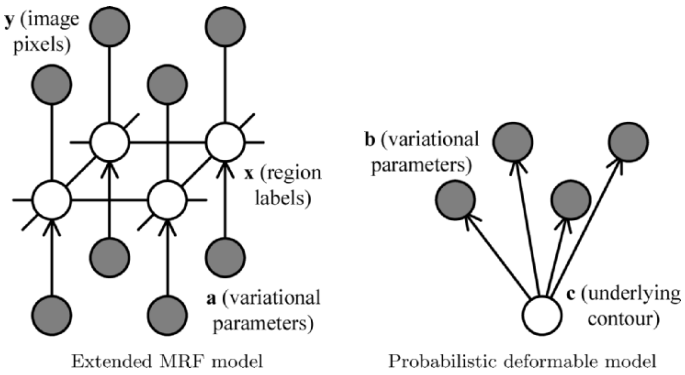


Fig. 3. Decoupled models

$$Q_M(\mathbf{x}|\mathbf{a}) = \prod_{(i,j)} \psi(x_i, x_j) \prod_i P(x_i|a_i)$$

Hence,

$$Q_M(\mathbf{x}|\mathbf{y}, \mathbf{a}) \propto \prod_i \phi(x_i, y_i) \prod_{(i,j)} \psi(x_i, x_j) \prod_i P(x_i|a_i) \tag{8}$$

On the other hand, the probabilistic deformable model yields:

$$Q_D(\mathbf{c}|\mathbf{b}) \propto Q_D(\mathbf{b}|\mathbf{c}) Q_D(\mathbf{c})$$

$$Q_D(\mathbf{b}|\mathbf{c}) = \prod_i P(b_i|\mathbf{c})$$

leading to

$$Q_D(\mathbf{c}|\mathbf{b}) \propto \prod_i P(b_i|\mathbf{c}) Q_D(\mathbf{c}) \tag{9}$$

The optimal values of the variational parameters $\theta = (\mathbf{a}, \mathbf{b})$ are obtained by minimizing the KL-divergence. It can be shown, using e.g., [24], that the optimal parameters $\theta^* = (\mathbf{a}^*, \mathbf{b}^*)$ should satisfy the following equations:

$$\log P(x_i|a_i^*) = \sum_{\mathbf{c}} Q_D(\mathbf{c}|\mathbf{b}^*) \log P(x_i|\mathbf{c}) \tag{10}$$

$$\log P(b_i^*|\mathbf{c}) = \sum_{x_i \in L} Q_M(x_i|\mathbf{y}, \mathbf{a}^*) \log P(x_i|\mathbf{c}) \tag{11}$$

Notice that the inference solutions, (8) and (9), together with the parameter optimizations, (10) and (11), form a set of *fixed-point equations*. Solution of this fixed-point set yields a tractable approximation to the intractable original posterior.

Since the state space of $\mathbf{c}$ (all possible contour configurations in the image plane) is too large, (10) is still intractable. We simply use the winner-take-all strategy and approximate $Q_D(\mathbf{c}|\mathbf{b})$ as a delta function:

$$Q'_D(\mathbf{c}|\mathbf{b}) = \begin{cases} 1 \text{ if } \mathbf{c} = \arg\max_{\mathbf{c}} Q_D(\mathbf{c}|\mathbf{b}) \\ 0 \text{ else} \end{cases}$$

and (10) can be simplified to:

$$P(x_i|a_i) = P(x_i|\mathbf{c}) \tag{12}$$

where $\mathbf{c} = \arg\max_{\mathbf{c}} Q_D(\mathbf{c}|\mathbf{b})$.

3.5 Algorithm Description

The variational inference algorithm for the hybrid MRF-DM model can now be summarized as:

Initialize contour **c**;
while $(error > maxError)$ {
1. Calculate a band area B around **c**. Perform remaining steps inside B;
2. Calculate $P(x_i|a_i)$ based on (12) using **c**;
3. Estimate the MRF-MAP solution $Q_M(x_i|\mathbf{y}, \mathbf{a})$ based on (8) using $P(x_i|a_i)$;
4. Calculate $\log P(b_i|\mathbf{c})$ based on (11) using $Q_M(x_i|\mathbf{y}, \mathbf{a})$;
5. Estimate the DM-MAP solution $Q_D(\mathbf{c}|\mathbf{b})$ based on (9) using $\log P(b_i|\mathbf{c})$;
}

Steps 2 and 4 follow directly from (12) and (11). The details of steps 1, 3, and 5 are discussed in Sect. 4.

4 Implementation Issues

4.1 Solve MRF-MAP with EM and BP

Step 3 of our algorithm solves the MAP problem in the extended MRF model. The EM algorithm is used to estimate both the MAP solution of region labels **x** and the parameters of the model (i.e., the parameters in the potential functions).

Particularly, in E step, the MAP solution of region labels **x** is estimated based on current parameters. Unlike most of the previous work mentioned in Sect. 2, we solve this MRF-MAP estimation problem using the BP algorithm. BP is an inference method proposed by Pearl [8] to efficiently estimate Bayesian beliefs in the network by the way of iteratively passing messages between neighbors. It is an exact inference method in the network without loops. Even in the network with loops, the method often leads to good approximate and tractable solutions [25].

There are two variants of the BP algorithm: sum–product and max–product. The sum–product message passing rule can be written as:

$$m_{ij}(x_j) = \sum_{x_i} \Psi_{ij}(x_i, x_j) \Phi_i(x_i) \prod_{k \in \aleph(i) \setminus j} m_{ki}(x_i)$$

The max–product has analogous formula, with the marginalization replaced by the maximum operator. At convergence:

$$x_{i\mathrm{MAP}} = \arg\max_{x_i} \Phi_i(x_i) \prod_{j \in \aleph(i)} m_{ji}(x_i)$$

According to our extended MRF model the compatibility functions are:

$$\Phi_i(x_i) = \phi(x_i, y_i) P(x_i|a_i)$$

$$\Psi_{ij}(x_i, x_j) = \psi(x_i, x_j)$$

We again note the difference from a traditional MRF model, due to the incorporated shape prior. $P(x_i|a_i)$ is calculated in step 2 of the algorithm. $\phi(x_i, y_i)$ and $\psi(x_i, x_j)$ can be calculated using current MRF parameters.

In this model we assume the image pixels are corrupted by white Gaussian noise:

$$\phi(x_i, y_i) = \frac{1}{\sqrt{2\pi\sigma_{x_i}^2}} \exp\left(-\frac{(y_i - \mu_{x_i})^2}{2\sigma_{x_i}^2}\right)$$

On the other hand, to penalize differences between the neighboring labels (i.e., to keep local region smoothness),

$$\psi(x_i, x_j) = \frac{1}{Z} \exp\left(\frac{\delta(x_i - x_j)}{\sigma^2}\right)$$

where $\delta(x) = 1$ if $x = 0$; $\delta(x) = 0$ if $x \neq 0$, σ controls the similarity of neighboring hidden states, and Z is a normalization constant.

As shown in step 1 in our algorithm, belief propagation is restricted to a single band of model variables around the current contour estimates. A reason for this is that, in practice, we only need to care about the statistics of pixels near the boundary. More importantly, the banded inference significantly speeds up the whole algorithm. Although convergence of the banded algorithm is not guaranteed, in our experiments, the BP algorithm does converge, usually in only one or two iterations.

In M step, the MRF parameters are updated based on the MAP solution of the region labels **x** using the following equations:

$$\mu_l = \frac{\sum_i Q_M(x_i = l|y_i, a_i) y_i}{\sum_i Q_M(x_i = l|y_i, a_i)}$$

$$\sigma_l^2 = \frac{\sum_i Q_M(x_i = l|y_i, a_i)(y_i - \mu_l)^2}{\sum_i Q_M(x_i = l|y_i, a_i)}$$

where $l \in \{inside, outside\}$.

4.2 Probabilistic Deformable Model

In step 5, according to (3), we use the negative log term, $-\log P(\mathbf{b}|\mathbf{c})$, as the external energy in the deformable model. Given this "label image" energy landscape, the image force is simply $\nabla(\log P(\mathbf{b}|\mathbf{c}))$. With the additional balloon forces, this leads to the discretized first-order Lagrangian dynamics equation:

$$\dot{\mathbf{d}} + \mathbf{K}\mathbf{d} = \nabla(\log P(\mathbf{b}|\mathbf{c})) + k\vec{\mathbf{n}}(s)$$

We note that this formulation is different from that of [20] where the deformable model is fitted to a *binary* label image obtained from the MAP configuration of **x**. In our method, we use a *probabilistic* measurement of the label of each pixel as specified in (11).

Finally, following the definition in (5)–(7), we note that the gradient of the coupling energy at pixel i, $\nabla(\log P(\mathbf{b}|\mathbf{c}))$, can be shown to be:

$$\frac{\partial \log P(\mathbf{b}|\mathbf{c})}{\partial \mathbf{c}} = -\frac{\partial \log P(\mathbf{b}|\mathbf{c})}{\partial loc(i)}$$

5 3D Image Segmentation

Increasing availability of high-resolution 3D image data using modalities such as magnetic resonance (MR) and computed tomography (CT) has prompted the need for 3D segmentation approaches. However, 3D image segmentation remains an extremely difficult problem, due to the complex topology of 3D objects, the massive data, and demanding computational algorithms. Many 3D approaches are often 2D in nature (i.e., applying the 2D algorithm slice by slice to the 3D volume data [26]). The lack of interaction among individual slice solutions, however, leads to results that are inferior to true 3D-based solutions [27].

In this section, we generalize our framework to 3D image segmentation based on the integration of 3D MRFs and deformable surface models. The proposed method is a true 3D method that fully exploits the structure of the 3D data, resulting in improved object segmentation. The generalization is straightforward using the graphical model representation, and the variational inference in the graphical model also leads to computationally more efficient solutions, which, in the 3D case, is still of main concern.

A 3D MRF model is shown in Fig. 4. The hidden nodes are positioned at the vertices of a regular 3D grid of the same size as the volume data (Fig. 4 left). Each hidden node x_i is connected to six neighboring hidden nodes (more neighbors can be connected by adding diagonals in the grid) and one observable node y_i (Fig. 4 right). Again, the observable nodes represent the voxel values of the 3D volume data and the hidden nodes represent the region labels of corresponding voxels.

As to the deformable models, Finite-Element Method (FEM)-based balloon models [27] and Polygonal Geometrically Deformed Model (GDM) [28] are commonly used for representation of 3D surfaces and segmentation of volume data.

Similar to the 2D case, a new hidden node representing the underlying boundary surface is added to the 3D MRF model (Fig. 5 left, only one pair of voxel/label nodes is drawn for simplicity). We again use the structured variational inference technique to seemingly decouple the integrated model into two simpler models (Fig. 5 right): one extended 3D MRF model with shape prior constraints and one probabilistic deformable surface model.

The 3D algorithm is similar to the 2D one. However, the expansion process of the 3D balloon model far away from the true boundary can be time-consuming and needs frequent reparametrization, and often suffers from local energy minima in

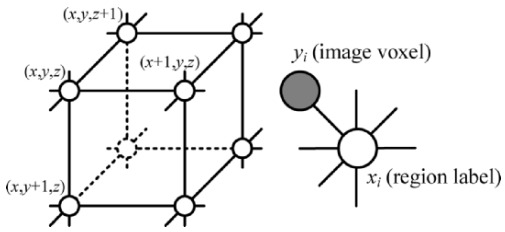


Fig. 4. 3D MRF model

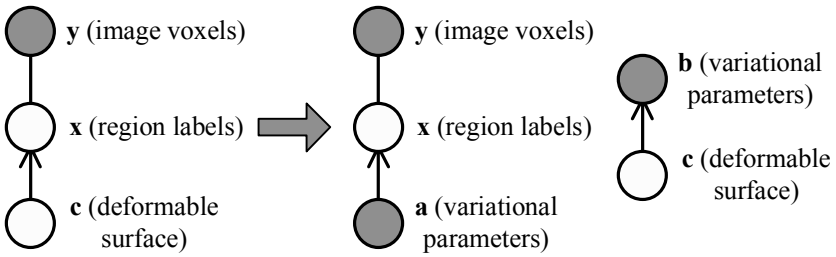


Fig. 5. Integrated and decoupled models

noisy images. An interactive initialization procedure or a learned shape prior would be helpful. When there is no shape prior, one can use the 3D MRF segmentation algorithm alone to generate an initial region segmentation and apply the Marching Cubes algorithm [29] to the 3D belief image to generate an initial surface. Marching Cubes is an algorithm for constructing triangle models of constant density surfaces from discrete volume data. The resulting surface representation is suitable for the FEM-based balloon model. The rest of the 3D algorithm is a straightforward generalization of the 2D one.

6 Experiments

6.1 2D Synthetic Images

The initial study of properties and utility of our method was conducted on a set of synthetic images. The images were synthesized in a way similar to [7]. In [7], the 64×64 perfect images contain only two gray levels representing the *object* (gray level is 160) and the *background* (gray level is 100) respectively. In our experiments, we made the background more complicated by introducing a gray level gradient. The gray levels of the background are increasing from 100 to 160, along the normal direction of the object contour (Fig. 6a). Figure 6b shows the result of a traditional MRF-based method. The object is segmented correctly, however some regions in the background are misclassified. On the other hand, the deformable model fails because of the leaking from the high-curvature part of the object contour, where the gradient in the normal direction is too weak (Fig. 6c). Our hybrid method, shown in Fig. 6d, results in a significantly improved segmentation.

We next generated a test image (Fig. 6e) by adding Gaussian noise with mean 0 and standard deviation 60 to Fig. 6a. The result of the MRF-based method on the noisy image (Fig. 6f) is somewhat similar to that in Fig. 6b, which shows the MRF can deal with image noise to some extent. But significant misclassification occurred because of the complicated background and noise levels. The deformable model either sticks to spurious edges caused by image noise or leaks (Fig. 6g) because of the weakness of the true edges. Unlike the two independent methods, our hybrid algorithm, depicted in Fig. 6h, correctly identifies the object boundaries despite the

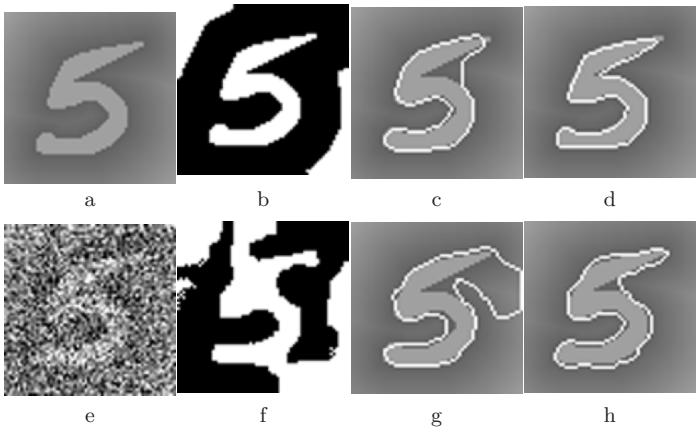


Fig. 6. Experiments on 2D synthetic images

excessive image noise. For visualization purposes we superimpose the contour on the original image (Fig. 6a) to show the quality of the result in Fig. 6g and h.

6.2 2D Medical Images

Experiments with synthetic images in Sect. 6.1 outlined some of the benefits of our hybrid method. The real world images usually have significant, often nonwhite noise and contain multiple regions and objects, rendering the segmentation task a great deal more difficult. In this section we show results of applying our method to real medical images on which we can hardly get satisfying results with either the MRF-based or the deformable model-based methods alone.

In the following comparisons, we manually specified the inside/outside regions to get an initial guess of the parameters for the MRF-only method. For the deformable model method, we started the balloon model at several different initial positions and use the best results for the comparison. On the other hand, our hybrid method is significantly less sensitive to the initialization of the parameters and the initial seed position.

Figure 7a shows a 2D MR image of the left ventricle of the human heart. Figure 7b is the result of the MRF-based method. While it is promising, the result still exhibits rough edges and holes. Figure 7c depicts the result of the deformable model-based method. Although we carefully chose the magnitude of the balloon forces, parts of the contour begin to leak others stick to spurious edges. Our hybrid method, started from the initial contour shown in Fig. 7e, generated better result (Fig. 7d). One of the intermediate iterations is shown in Fig. 7f. The corresponding external energy in the band area is depicted in Fig. 7g (image intensities are proportional to the magnitude of the energy), showing a more useful profile than the traditional edge energy $-|\nabla(G_\sigma * I)|^2$ shown in Fig. 7h.

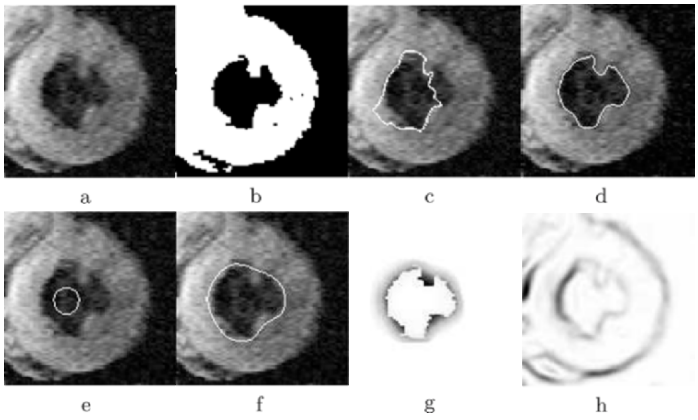


Fig. 7. Experiments on 2D medical images (1)

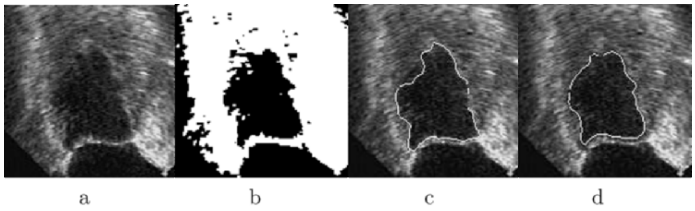


Fig. 8. Experiments on 2D medical images (2)

Figure 8a is an ultrasound image. The MRF gets rough edges and holes in the objects (Fig. 8b) while the deformable model cannot escape a local minimum (Fig. 8c). Our hybrid method eliminates the rough edges and holes caused by the MRF while outlining the region more accurately than the deformable model (Fig. 8d).

Finally, Figs. 9a and 10a are both examples of difficult images with complicated global properties, requiring the MRF-based method to automatically determine the number of regions and the initial values of the parameters. Figure 9b is obtained by manually initializing the MRF model. Our method avoids this problem by creating and updating an MRF model locally and incrementally. The images are also difficult for deformable models because the boundaries of the objects to be segmented are either high-curvature (Fig. 9a) or low-gradient (Fig. 10a). Figure 9c exemplifies the oversmoothed deformable models. Our method's results, shown in Figs. 9d and 10b, do not suffer from either of the problems.

6.3 3D Synthetic Images

Our 3D method was also first experimented on a set of synthetic images. The perfect image contains two gray levels representing the object (gray level is 160) and the background (gray level is 100), respectively. Gaussian noise with mean 0 and standard deviation 60 is added to the whole image to generate the test image.

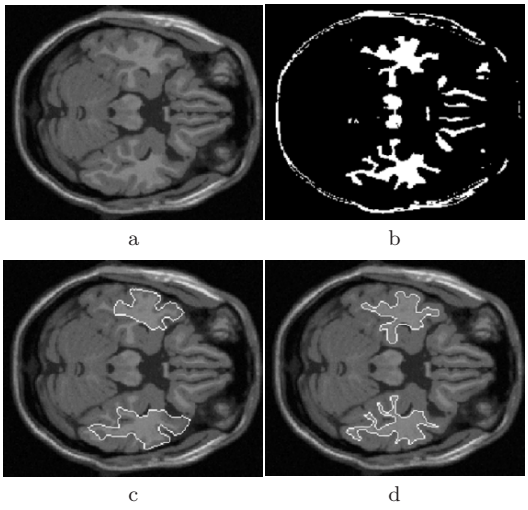


Fig. 9. Experiments on 2D medical images (3)

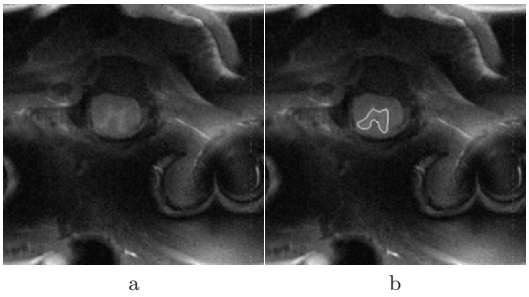
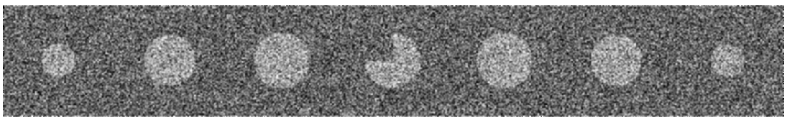


Fig. 10. Experiments on 2D medical images (4)

The first experiment intended to show the advantages of the true 3D method over the 2D slice-based method. In this experiment, we generated a $100 \times 100 \times 100$ 3D image containing a ball-like object. Figure 11a shows several slices of the perfect image. Our test image is generated by cutting out a quarter of the pie-like object from the 50^{th} frame and adding the Gaussian noise (Fig. 11b). The segmentation results by 2D MRFs and 3D MRFs are shown in Fig. 11c, d. Both models handled noise successfully. The 3D MRF model obviously recovered the pie-like object in the 50^{th} frame by retaining region smoothness in the direction perpendicular to the frame. The 2D MRF model cannot achieve this due to the lack of interaction between neighboring frames. The boundary of the results from 3D MRFs also look smoother.

The second experiment was performed on a $64 \times 64 \times 64$ volume containing a "5"-like object similar to Fig. 6a. The thickness of the object is 8 (i.e., frames 29 to 36 contain the object). Besides the zero mean Gaussian noise, extra noise with mean 160 is also added to a part of the two successive frames 32 and 33. The test

a. Perfect image slices (Frame 30, 40, 49, 50, 51, 60, 70)

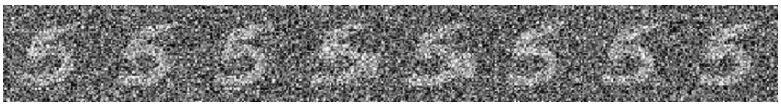


b. Corrupted Image slices

c. 2D MRFs segmented slices

d. 3D MRFs segmented slices

Fig. 11. Experiments on 3D synthetic images (1)



a. noisy slices (Frame 29 36)

b. 2D MRF segmentation slices (error = 3.98%)

c. 3D MRF segmentation slices (error = 2.79%)

d. 3D MRF + DM segmentation slices (error = 1.62%)

Fig. 12. Experiments on synthetic 3D images (2)

image slices are shown in Fig. 12a. The results of 2D MRFs are shown in Fig. 12b. Each slice looks different from others, especially for the two frames with extra noise. The slices in Fig. 12c (results of 3D MRFs), however, are smoother and similar to

their neighbors, except for the first and last frames, which suffered more interference from the background. These two outermost frames are improved by coupling the DM with the 3D MRF model, and other frames are also slightly smoother (Fig. 12d). The average error rates of the three methods are 3.98%, 2.79%, and 1.62%.

6.4 3D Medical Images

Experiments with synthetic images in Sect. 6.3 outlined the advantages of both the 3D method over the 2D method and the hybrid method over the MRF-only method. In this section, we show experimental results of applying our methods to 3D medical images. We do not show the results of the slice-based method with 2D MRFs as in previous experiments mainly because this method is sensitive to initialization and we cannot get satisfying results on these medical images. While our 3D method also needs manual initialization when the shape prior is not given, the slice-based method requires manual initialization for almost each single slice.

We first test our algorithms on simulated brain MRI data from BrainWeb [30]. The database contains simulated brain MRI data based on two anatomical models: normal and multiple sclerosis. For both of these, full 3D data volumes have been simulated using three sequences (T1-, T2-, and proton-density (PD)-weighted) and a variety of slice thicknesses, noise levels, and levels of intensity nonuniformity. We segmented the white matter from three different normal brain data volumes using the hybrid method. Figure 13a shows a slice from the ground truth data of the white matter. Figure 13d is the result from our hybrid method. The second column of Fig. 13 shows the segmentation results on T1 image without noise and intensity nonuniformity (RF inhomogeneity). The segmented white matter is slightly thicker than the results from the ground truth, because some of the grey matter is misclassified due to its similar grey value to the white matter. Same misclassification can be observed in

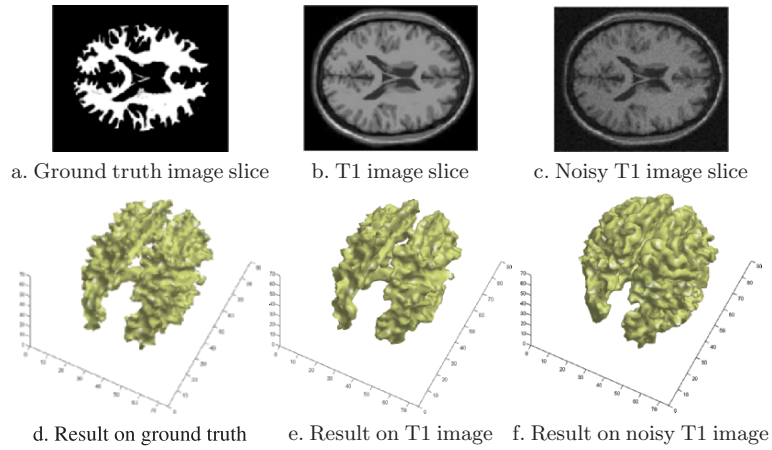


Fig. 13. Experiments on 3D medical images (1)

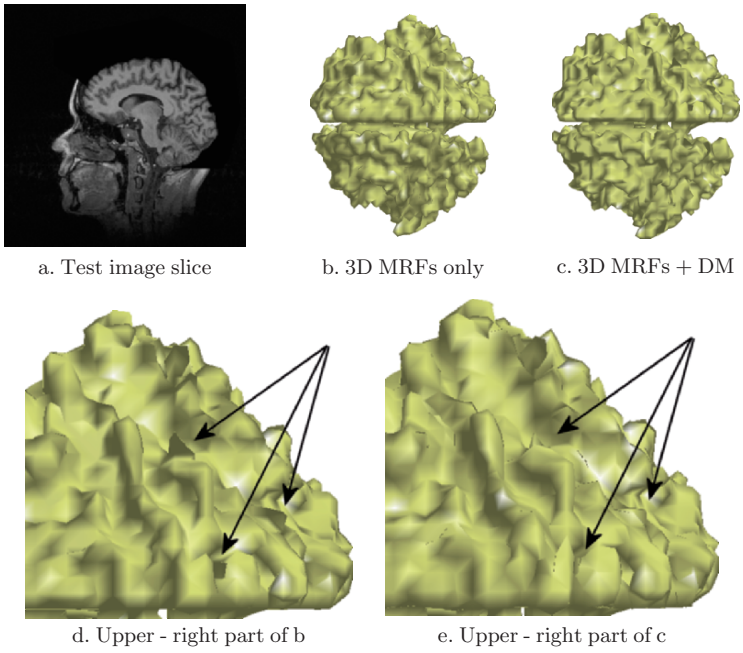


Fig. 14. Experiments on 3D medical images (1)

the third column, which is the segmentation result on T1 image with 9% noise and 40% intensity nonuniformity. One possible solution to the misclassification problem is using the 3D MRF-only algorithm to do a multiregion segmentation first.

Finally, we show some results on a real medical image [31], which is an MR image of a head with the skull partially removed to reveal the brain. Figure 14a is one of the slices from the volume. The results of our methods are shown in Fig. 14b, c. To show the difference between the two algorithms (i.e., the effect of adding deformable models), the upper-right parts of Fig. 14b, c are magnified in Fig. 14d, e. The arrows show that some incorrect patches are eliminated by the deformable fitting process. Surface smoothness can be easily controlled by tuning the parameters in the stiffness matrix. Because the white matter itself is a complicated object with high curvature, the parameters are usually chosen according to experts' opinion.

7 Conclusions

We proposed a new framework to combine the MRF-based and the deformable model-based segmentation methods. The framework was developed under the auspices of the graphical model theory allowing us to employ a well-founded set of statistical inference and learning techniques. In particular, we employed the variational inference method, an approximate, computationally efficient solution, to otherwise intractable inference of region boundaries. Experimental results on both synthetic

and real 2D images and 3D volumes show that the hybrid methods outperforms both the MRF-based and the deformable model-based methods using only homogeneous constraints.

References

1. Jordan, M.I. ed.: Learning in Graphical Models. The MIT, Cambridge, MA (1998)
2. Jordan, M.I.: Graphical models. Statistical Science (Special Issue on Bayesian Statistics) **19**(1) (2004) 140–155
3. Sonka, M., Hlavac, V., Boyle, R.: Image Processing, Analysis and Machine Vision, Second Edition. Thomson Learning (1998)
4. Geman, S., Geman, D.: Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. IEEE Transaction on Pattern Analysis and Machine Intelligence **6**(6) (1984) 721–741
5. Besag, J.E.: On the statistical analysis of dirty pictures. Journal of the Royal Statistical Society Series B **48**(3) (1986) 259–302
6. Marroquin, J., Mitter, S., Poggio, T.: Probabilistic solution of ill-posed problems in computational vision. Journal of the American Statistical Association **82**(397) (1987) 76–89
7. Dubes, R., Jain, A., Nadabar, S., Chen, C.: MRF model-based algorithm for image segmentation. In: Proceedings of International Conference on Pattern Recognition. Volume 1. (1990) 808–814
8. Pearl, J.: Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Morgan Kaufmann, Stanford, CA (1988)
9. Yedidia, J., Freeman, W.T., Weiss, Y.: Understanding belief propagation and its generalizations. In: International Joint Conference on Artificial Intelligence, Distinguished Presentations Track. (2001)
10. Boykov, Y., Veksler, O., Zabih, R.: Fast approximate energy minimization via graph cuts. IEEE Transactions on Pattern Analysis and Machine Intelligence **23**(11) (2001) 1222–1239
11. Tappen, M., Freeman, W.: Comparison of graph cuts with belief propagation for stereo, using identical MRF parameters. In: Proceedings of IEEE International Conference on Computer Vision. Volume 2. (2003) 900–907
12. Zhang, Y., Brady, M., Smith, S.: Segmentation of brain MR images through a hidden Markov random field model and the expectation–maximization algorithm. IEEE Transaction on Medical Imaging **20**(1) (2001) 45–57
13. Kass, M., Witkin, A., Terzopoulos, D.: Snakes: Active contour models. International Journal of Computer Vision **1**(4) (1988) 321–331
14. Cohen, L.: On active contour models and balloons. Computer Vision, Graphics, and Image Processing: Image Understanding **53**(2) (1991) 211–218
15. Xu, C., Prince, J.: Gradient vector flow: A new external force for snakes. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. (1997) 66–71
16. McInerney, T., Terzopoulos, D.: Topologically adaptable snakes. In: Proceedings of IEEE International Conference on Computer Vision. (1995) 840–845
17. McInerney, T., Terzopoulos, D.: Deformable models in medical image analysis: A survey. Medical Image Analysis **1**(2) (1996) 91–108
18. Ronfard, R.: Region-based strategies for active contour models. International Journal of Computer Vision **13**(2) (1994) 229–251

19. Jones, T., Metaxas, D.: Image segmentation based on the integration of pixel affinity and deformable models. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. (1998) 330–337
20. Chen, T., Metaxas, D.: Image segmentation based on the integration of Markov random fields and deformable models. In: Proceedings of International Conference on Medical Image Computing and Computer-Assisted Intervention. (2000) 256–265
21. Metaxas, D.: Physics-based Deformable Models: Applications to Computer Vision, Graphics and Medical Imaging. Kluwer, Dordecht (1997)
22. Jordan, M., Ghahramani, Z., Jaakkola, T., Saul, L.: An introduction to variational methods for graphical models. Machine Learning **37**(2) (1999) 183–233
23. Pavlovic, V., Frey, B., Huang, T.: Variational learning in mixed-state dynamic graphical models. In: Proceedings of Conference on Uncertainty in Artificial Intelligence. (1999) 522–530
24. Ghahramani, Z.: On structured variational approximations. Technical Report CRG-TR-97-1 (1997)
25. Weiss, Y.: Belief propagation and revision in networks with loops. Technical Report MIT A.I. Memo 1616 (1998)
26. Choi, S., Lee, J., Kim, J., Kim, M.: Volumetric object reconstruction using the 3D-MRF model-based segmentation. IEEE Transaction on Medical Imaging **16**(6) (1997) 887–892
27. Cohen, L., Cohen, I.: Finite-element methods for active contour models and balloons for 2-D and 3-D images. IEEE Transactions on Pattern Analysis and Machine Intelligence **15**(11) (1993) 1131–1147
28. Miller, J., Breen, D., Lorensen, W., O'Bara, R., Wozny, M.: Geometrically deformed models: A method for extracting closed geometric models from volume data. Computer Graphics (Proceedings of SIGGRAPH) **25**(4) (1991) 217–226
29. Lorensen, W., Cline, H.: Marching cubes: A high resolution 3D surface construction algorithm. Computer Graphics (Proceedings of SIGGRAPH) **21**(4) (1987) 163–169
30. BrainWeb: Simulated Brain Database. (http://www.bic.mni.mcgill.ca/brainweb/)
31. Stanford: Volume Data Archive. (http://graphics.stanford.edu/data/voldata/)