Journal of Zhejiang University SCIENCE B
ISSN 1673-1581 (Print); ISSN 1862-1783 (Online)
www.zju.edu.cn/jzus; www.springerlink.com
E-mail: jzus@zju.edu.cn

SVM for density estimation and application to medical image segmentation*

ZHANG Zhao†, ZHANG Su, ZHANG Chen-xi, CHEN Ya-zhu
(*Biomedical Instrument Institute, Shanghai Jiao Tong University, Shanghai 200030, China*)
†E-mail: z_ball@sjtu.edu.cn
Received Sept. 29, 2005; revision accepted Feb. 27, 2006

Abstract: A method of medical image segmentation based on support vector machine (SVM) for density estimation is presented. We used this estimator to construct a prior model of the image intensity and curvature profile of the structure from training images. When segmenting a novel image similar to the training images, the technique of narrow level set method is used. The higher dimensional surface evolution metric is defined by the prior model instead of by energy minimization function. This method offers several advantages. First, SVM for density estimation is consistent and its solution is sparse. Second, compared to the traditional level set methods, this method incorporates shape information on the object to be segmented into the segmentation process. Segmentation results are demonstrated on synthetic images, MR images and ultrasonic images.

Key words: Support vector machine (SVM), Density estimation, Medical image segmentation, Level set method
doi:10.1631/jzus.2006.B0365 **Document code:** A **CLC number:** TN 911.73

INTRODUCTION

Image segmentation problem is an important task in image understanding and computer vision. Especially, medical image processing applications such as surgical planning, navigation, simulation, diagnosis, and therapy evaluation all benefit from segmentation of anatomical structures from medical images. Image segmentation is usually based on four different philosophical perspectives (Hampton *et al*., 1998): region approach where each voxels is assigned to a particular object or region, boundary approach where the location of the boundaries existing between regions is desired, edge approach where identification of the edge voxels is followed by a linking process of these edges which will form the required boundaries, and hybrid methods.

For medical images, the segmented regions often do not correspond well to true objects because of the noise, large "internal" intensity variations in the structures to be identified and the overlapping of the distribution of intensity values corresponding to one structure with those of another structure. An effective way to solve these problems is to incorporate the known shape information on the desired objects into the segmentation process. The support vector machine (SVM) approach is considered as a good candidate for utilizing the prior knowledge because of its high generalization performance and sparse solution. In this paper, we use SVM density estimator to construct a prior knowledge model of the structure based on previously segmented training data. To segment an object from a novel image, we improve the level set method by using the evolution metric defined by the prior knowledge model instead of by minimizing an energy function over a curve. For efficiency, narrow band method is used during the segmentation process.

The rest of the paper is organized as follows. First, the theoretical concept of density estimation based on SVM is presented. Then the implementation

* Project (No. 2003CB716103) supported by the National Basic Research Program (973) of China and the Key Lab for Image Processing and Intelligent Control of National Education Ministry, China

of the segmentation method is described. In the end, the results of segmentation and discussion are provided.

DENSITY ESTIMATION MODEL BASED ON SVM

SVM is a new type of learning machines based on statistical learning theory. It takes both empirical risk and confidence interval into account. SVM is mainly developed to solve the classification problem, regression problem and density estimation problem.

SVM method for density estimation is a non-parametric approach. This method is consistent and results in a sparse solution. The SVM and Parzen method have similar quality solutions, but the SVM solution is sparse. The SVM and the Gaussian mixture model (GMM) approaches are both sparse but the SVM approach is consistent and more accurate than GMM. The idea of SVM method is that using the technique of solving linear operator equations by SVM to estimate the density directly based on the definition of probability density (Vapnik and Mukherjee, 2000).

According to the definition of the density, $P(\boldsymbol{x})$, is the solution of the following linear operator equation:

$$F(\boldsymbol{x}) == \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_N} P(\boldsymbol{t}) \mathrm{d}t_1 \cdots \mathrm{d}t_N, \tag{1}$$

where $\boldsymbol{x}=(x_1, \ldots, x_N)\in R^n$, and it satisfies the following properties:

$$P(\boldsymbol{x}) \geq 0, \tag{2}$$

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} P(\boldsymbol{t}) \cdot \cdot \mathrm{d}t_1 \cdots \mathrm{d}t_N = 1. \tag{3}$$

Since $F(\boldsymbol{x})$ is unknown and a random independent sample is given as $x_1, \ldots, x_l$, the empirical distribution function can be evaluated as:

$$F_l(\boldsymbol{x}) = \frac{1}{l}\sum_{i=1}^{l} \theta(x_1 - x_{i,1}) \cdots \theta(x_N - x_{i,N}), \tag{4}$$

where

$$\theta(x) = \begin{cases} 1, & x > 0, \\ 0, & \text{otherwise}. \end{cases} \tag{5}$$

Constructing the data

$$(\boldsymbol{x}_1, F_l(\boldsymbol{x}_1)), \ldots, (\boldsymbol{x}_l, F_l(\boldsymbol{x}_l)), \tag{6}$$

to solve the density estimation problem we use the SVM for solving linear equations

$$AP(\boldsymbol{x})=F(\boldsymbol{x}), \tag{7}$$

where operator A is a linear mapping from a Hilbert space of function $P(\boldsymbol{x})$ to a Hilbert space of function $F(\boldsymbol{x})$. We solve a regression problem in the image space $P(\boldsymbol{x}, \boldsymbol{w})$ based on Eq.(6) and this solution, which is an expansion on the support vectors, can be used to describe the solution in the pre-image space $P(\boldsymbol{x}, \boldsymbol{w})$. Namely, the problem of finding the solution of linear operator equations is equivalent to the regression problem in image space using SVM method. In pre-image space, $P(\boldsymbol{x})$ can be expressed as:

$$P(\boldsymbol{x}, \boldsymbol{w}) = \sum_{r=0}^{\infty} w_r \varphi_r(\boldsymbol{x}) = \boldsymbol{w} \cdot \Phi(\boldsymbol{x}). \tag{8}$$

The result of the mapping from the pre-image to the image space can be expressed as:

$$F(\boldsymbol{x}, \boldsymbol{w}) = AP(\boldsymbol{x}, \boldsymbol{w}) = \sum_{r=0}^{\infty} w_r \psi_r(\boldsymbol{x}) = \boldsymbol{w} \cdot \psi(\boldsymbol{x}). \tag{9}$$

Solving the regression problem in image space used SVM method:

$$\boldsymbol{w} = \sum_{i=1}^{l} \alpha_i \Psi(x_i), \tag{10}$$

so the solution of Eq.(1) is:

$$P(\boldsymbol{x}) = \sum_{i=1}^{l} \alpha_i \Psi(x_i) \Phi(\boldsymbol{x}). \tag{11}$$

We define the kernel in image space:

$$K(\boldsymbol{x}_i, \boldsymbol{x}_j) = \sum_{r=0}^{\infty} \psi_r(\boldsymbol{x}_i) \psi_r(\boldsymbol{x}_j), \tag{12}$$

and cross kernel:

$$\kappa(\boldsymbol{x}_i, \boldsymbol{x}) = \sum_{r=0}^{\infty} \psi_r(\boldsymbol{x}_i)\varphi_r(\boldsymbol{x}), \tag{13}$$

so the density can be calculated as:

$$P(\boldsymbol{x}) = \sum_{i=1}^{l} \alpha_i \kappa(\boldsymbol{x}_i, \boldsymbol{x}). \tag{14}$$

In classical density estimation theory, in order to satisfy Eq.(2) and Eq.(3), the weight α_i should be as follows:

$$\sum_{i=1}^{l} \alpha_i = 1, \quad \alpha_i \geq 0, \ i = 1, \ \ldots, \ l. \tag{15}$$

Then, when we use the technique of linear SVM and ε-insensitive loss function, the problem of density estimation is equivalent to the linear programming problem as follows:

$$\min\left(\sum_{i=1}^{l} \delta_i \alpha_i + C\sum_{i=1}^{l} \xi_i + C\sum_{i=1}^{l} \xi_i^*\right), \tag{16}$$

subject to the constraints Eq.(15) and

$$F_l(\boldsymbol{x}_i) - \varepsilon_i - \xi_i \leq \sum_{j=1}^{l} \alpha_j K(\boldsymbol{x}_i, \boldsymbol{x}_j) \leq F_l(\boldsymbol{x}_i) + \varepsilon_i + \xi_i^*, \quad i=1, \ldots, l \tag{17}$$

$$\xi_i \geq 0, \ \xi_i^* \geq 0, \quad i=1, \ldots, l \tag{18}$$

where

$$\varepsilon_i = \lambda \sqrt{\frac{1}{l} F_l(\boldsymbol{x}_i)(1 - F_l(\boldsymbol{x}_i))}, \tag{19}$$

$$\delta_i = \frac{1}{l}\sum_{j=1}^{l} \| \boldsymbol{x}_i - \boldsymbol{x}_j \|^2. \tag{20}$$

We can get the solution of density estimation by substituting α_i into Eq.(14). Usually most of the α_i values in the SVM estimate will be zero, namely, the solution is sparse. The $\boldsymbol{x}_i$ values corresponding to the nonzero α_i values are called support vectors. For the Parzen method, the solution is not sparse. So we see that the computational labor of estimating the density will be reduced by SVM method than by Parzen method.

In this paper, we choose the kernel and cross kernel function as follows:

$$K(\boldsymbol{x}, \boldsymbol{x}') = \prod_{i=1}^{N} \frac{1}{1 + e^{-\gamma(x_i - x_i')}}, \tag{21}$$

$$\kappa(\boldsymbol{x}, \boldsymbol{x}') = \prod_{i=1}^{N} \frac{\gamma}{2 + e^{\gamma(x_i - x_i')} + e^{-\gamma(x_i - x_i')}}. \tag{22}$$

SEGMENTATION ALGORITHM BASED ON SVM

The basic idea of level set approaches is the evolution of a curve C that is embedded as a zero level set of a higher dimensional surface (Sethian, 1999). The surface is usually expressed as a function ϕ of time. So given an image $I(\boldsymbol{x})$, the curve C can be defined as follows:

$$\begin{cases} C_0 = \{\boldsymbol{x} \mid \phi(\boldsymbol{x}, 0) = 0\}, \\ C_t = \{\boldsymbol{x} \mid \phi(\boldsymbol{x}, t) = 0\}. \end{cases} \tag{23}$$

Commonly, we define the function as follows:

$$\phi(\boldsymbol{x}, 0) = \pm d. \tag{24}$$

It is called singed distance function, where d is the distance from the point $\boldsymbol{x}$ to the nearest point in the curve C at t=0, and the plus (minus) sign is chosen while $\boldsymbol{x}$ is outside (inside) the curve. With time progressing, the surface can change to take on the desired shape. The motion for this evolving function ϕ is determined from a partial differential equation in the higher dimension which permits cusps, sharp corners, and changes in topology in the zero level set describing the interface. This process can be described as follows:

Step 1: Initialize $\phi(\boldsymbol{x}, 0)$ with an initial closed curve.

Step 2: Update $\phi(\boldsymbol{x}, t)$ based on different metrics according to different level set approaches.

Step 3: Extract the boundary of the desired objects by $C_t=\{\boldsymbol{x}|\phi(\boldsymbol{x}, t)=0\}$ if the stop conditions are satisfied, otherwise return to Step 2.

Traditionally, the metric in Step 2 is the energy-based active contours minimization problem by the computation of geodesics or minimal distance curves. To incorporate the prior information into the process of segmenting desired objects in an image, we choose $P(\phi(\boldsymbol{x})|I(\boldsymbol{x}),\phi(N(\boldsymbol{x})))$ as the metric, where $N(\boldsymbol{x})$ is the neighborhood of the point $\boldsymbol{x}$. It expresses the probability of the value of the surface at the point $\boldsymbol{x}$, given the intensity value of the image at the same point and the neighboring values of the surface. This conditioned distribution can be approximated as follows (Leventon *et al*., 2002):

$$\begin{aligned}P(\phi(\boldsymbol{x})\,|\,I(\boldsymbol{x}),\phi(N(\boldsymbol{x}))) \propto P(I(\boldsymbol{x}),\phi(\boldsymbol{x}))\\ P(\phi(\hat{t}+),\phi(\hat{t}-)\,|\,\phi(\boldsymbol{x}))P(\phi(\hat{n}+),\phi(\hat{n}-)\,|\,\phi(\boldsymbol{x})).\end{aligned} \tag{25}$$

Here, we use four neighbors in the direction of the local normal ($\hat{n}+$, $\hat{n}-$) and the tangent ($\hat{t}+$, $\hat{t}-$) to the embedded curve as shown in Fig.1. The first term on right hand side of Eq.(25) is called the image term. The middle term is called curvature term. The last term is called linearity term.

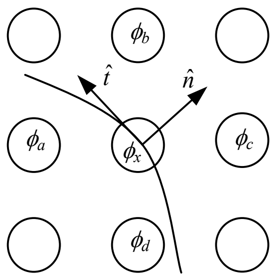


Fig.1 The directions of the normal and tangent of ϕ at the point x

We model this conditioned distribution based on a set of segmented images that have the same modality containing the same anatomical structure of various subjects. The distribution of the novel images to be segmented should be similar to that of the training data. So to segment a novel image, we maximize the probability to estimate each surface point $\phi(\boldsymbol{x})$ independently while assuming the rest of the surface is constant:

$$\phi(\boldsymbol{x}) = \max_{\phi(\boldsymbol{x})} \log P(\phi(\boldsymbol{x})\,|\,I(\boldsymbol{x}),\phi(N(\boldsymbol{x}))). \tag{26}$$

So in the Step 2, in order to update $\phi(\boldsymbol{x})$, we use the function as follows:

$$\begin{aligned}&\phi(\boldsymbol{x},\ t+1)=\phi(\boldsymbol{x},\ t)\\ &+\lambda(\frac{\mathrm{d}}{\mathrm{d}\phi(\boldsymbol{x},\ t)}\log P(\phi(\boldsymbol{x},\ t)\,|\,I(\boldsymbol{x}),\ \phi(N(\boldsymbol{x}),\ t))).\end{aligned} \tag{27}$$

Namely:

$$\begin{aligned}\phi(\boldsymbol{x},\ t+1) = \phi(\boldsymbol{x},\ t) + \lambda(&\frac{\mathrm{d}}{\mathrm{d}\phi(\boldsymbol{x},\ t)}\log P(I(\boldsymbol{x}),\ \phi(\boldsymbol{x},\ t))\\ &+\frac{\mathrm{d}}{\mathrm{d}\phi(\boldsymbol{x},\ t)}\log P(\phi(\hat{t}+),\ \phi(\hat{t}-)\,|\,\phi(\boldsymbol{x},\ t))\\ &+\frac{\mathrm{d}}{\mathrm{d}\phi(\boldsymbol{x},\ t)}\log P(\phi(\hat{n}+),\ \phi(\hat{n}-)\,|\,\phi(\boldsymbol{x},\ t))).\end{aligned} \tag{28}$$

This update is repeated until there is little change in the surface. Then the boundary of the objects can be extracted from the zero level set of ϕ.

To construct the prior distribution, we model image term, curvature term and linearity term independently.

Image term: It relates the intensity and the surface at $\boldsymbol{x}$. Let $\{I_1, I_2, \ldots, I_n\}$ be the training images and $\{\phi_1, \phi_2, \ldots, \phi_n\}$ be the corresponding singed distance functions. So training set of the intensity term is $T=\{<I_1,\ \phi_1>,\ \ldots,\ <I_n,\ \phi_n>\}$. It is a joint distribution. The image term can be derived from the training set using the method of SVM described in the last section. Then the derivative of the log image term in Eq.(28) is computed by taking the gradient of the sampled probability in the direction of $\phi(\boldsymbol{x})$.

Curvature term: It reflects the curvature profile of the training data. When the surface is the singed distance function, the second partial derivatives of ϕ in the tangent and normal directions are as follows (Leventon *et al*., 2002):

$$\frac{\partial^2\phi}{\partial\hat{t}^2} = k,\ \frac{\partial^2\phi}{\partial\hat{n}^2} = 0. \tag{29}$$

Considering the formulas for finite differences of an arbitrary 1D function:

$$\frac{\mathrm{d}^2 f(x)}{\mathrm{d}t^2} \approx \frac{f(x+t)+f(x-t)-2f(x)}{t^2}. \tag{30}$$

We see that curvature k reflects the distribution of $\phi(\hat{t}+)$, $\phi(\hat{t}-)$ and $\phi(\boldsymbol{x})$. So we use $\{k_1, \ldots, k_n\}$ as the training set of curvature term. We derived the curvature term from the training set using the method of SVM described in the last section. Then the derivative of the log curvature term in Eq.(28) can be computed by taking the gradient of the sampled probability in the direction of $\phi(\boldsymbol{x})$. The curvature of the underlying level set is calculated as follows (Sethian, 1999):

$$k = \frac{\phi_{xx}\phi_y^2 - 2\phi_x\phi_y\phi_{xy} + \phi_{yy}\phi_x^2}{(\phi_x^2 + \phi_y^2)^{3/2}}. \tag{31}$$

Linearity term: It is a regularization term for preventing the surface from evolving arbitrarily. From Eq.(29) and Eq.(30), we see that the second derivative in the normal direction should be zero. So the derivative of the log linearity term in Eq.(28) is defined as follows:

$$\begin{aligned} &\frac{\mathrm{d}}{\mathrm{d}\phi(\boldsymbol{x},\ t)}\log P(\phi(\hat{n}+),\ \phi(\hat{n}-)\,|\,\phi(\boldsymbol{x},\ t)) \\ &\quad = \alpha(\phi(\hat{n}+) + \phi(\hat{n}-) - 2\phi(\boldsymbol{x},\ t)). \end{aligned} \tag{32}$$

For efficiency, instead of evolving the surface at every point over the image, we just update the surface in a narrow band around the boundary during the evolution (Adalsteinsson and Sethian, 1995). So the segmentation algorithm in this paper is summarized as follows:

Step 1: Use SVM method to model image term and curvature term from their training set independently.

Step 2: To segment a novel image, give an initial closed curve, and calculate the signed distance function $\phi(\boldsymbol{x}, 0)$ on a tube around the curve.

Step 3: Update $\phi(\boldsymbol{x}, t)$ on the tubular domain according to Eq.(28).

Step 4: If the curve is within a set distance of the tube boundary or instability is developing on the tube boundary, reinitialize the surface by going to Step 2 and resize the rectangular square where we do the calculations.

Step 5: If there is little change in the surface, exit iteration and extract the boundary of the desired objects by $C_t=\{\boldsymbol{x}|\phi(\boldsymbol{x}, t)=0\}$, otherwise go to Step 3.

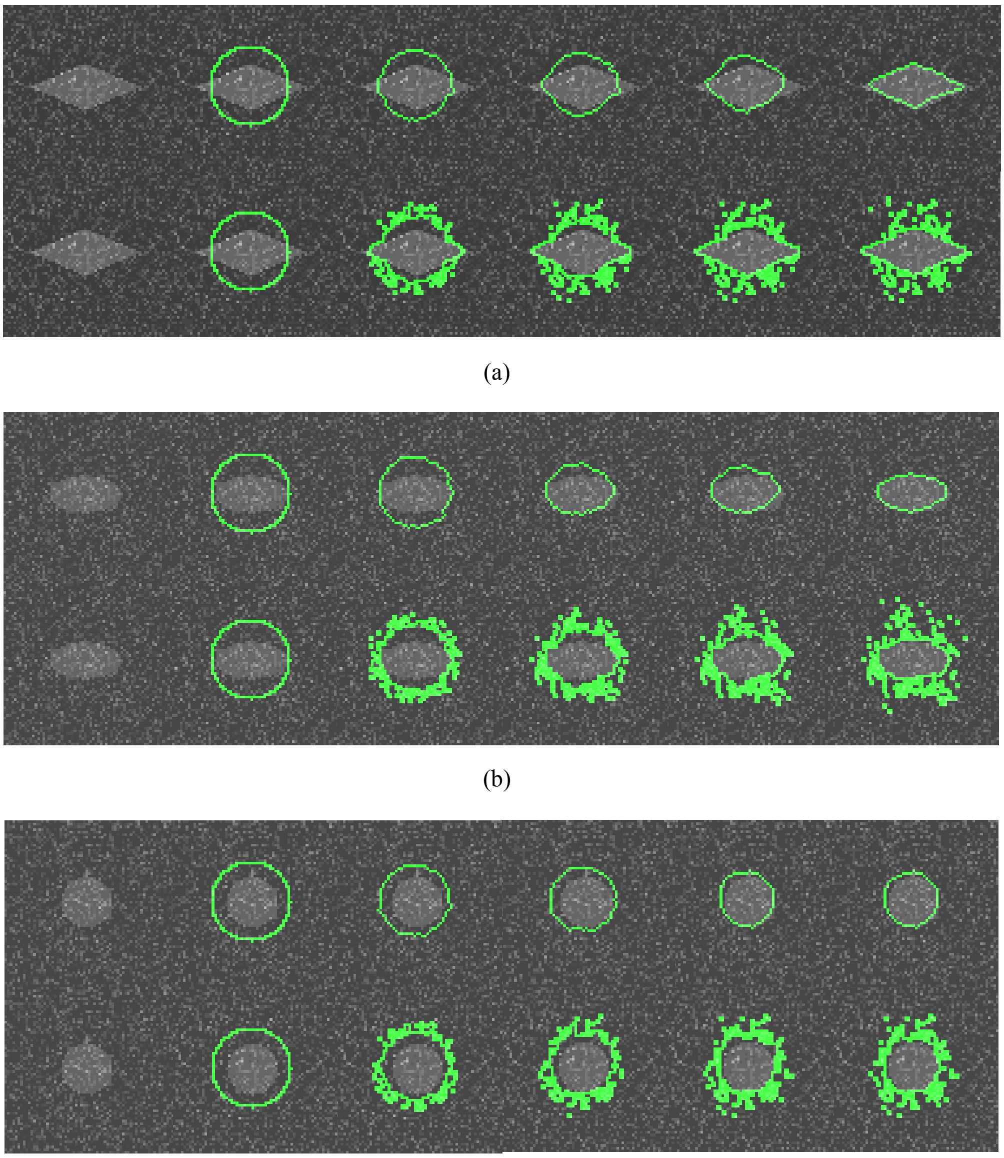
EXPERIMENTAL RESULTS AND DISCUSSION

Let us present some experiments to illustrate our segmentation method. There are six sets of images including three sets of synthetic images, one set of brain in sagittal MR images, one set of femur in MR images of the knee and one set of egg in ultrasonic images. For each set of images, both the segmentation method described in this paper and the level set method based on the C-V model (Chan and Vese, 2001) are used. The experimental results are shown in Figs.2~5. The top row in each Figure shows the results of the first method. The bottom row shows the results of the second method. The figures illustrate the initial, middle, and final steps in the evolution process of segmentation. The synthetic images were constructed by embedding the shapes of random rhombi, ellipses and circularities with the addition of random noise. In each set of synthetic images, 20 images with similar shape were generated. Nineteen images of them were used to construct the training set. The other one was used as the test set. The image term and curvature term were modelled based on the training set. Fig.2 shows the results of the segmentation. There are raw images to be segmented in the first column from left. The images in the second column from left have the initial curves. Fig.3 shows the segmentation of femur in MR image of the knee. Fig.4 shows the segmentation of the brain in sagittal MR image. The segmentation of egg in ultrasonic image is shown in Fig.5. In each of their cases, the training set consisted of the ten neighboring slices of the test slice. The training sets were segmented manually. Then the image term and curvature term were modelled based on the training set.

In our segmentation method, SVM for density estimation were used to model the image term and curvature term. The experiments showed that the number of support vectors only amounts to 2%~7% of the number of training samples. Because of the sparse solution, SVM method to model the image term and the curvature term reduced the computational labor of updating the surface, compared to Parzen method. By using the prior model, this segmentation method provides effective solution to incorporate shape information on the desired objects into the segmentation process. Furthermore, because of adding image intensity information into the model, this segmentation method can detect objects whose boundaries do not necessarily have a strong gradient.

From the first three experiments, we can see that the method mentioned in this paper is robust under the condition of noise. And it can segment different shapes according to the training set. From the last three experiments, we see that the method works well with different modality medical images and different anatomical structures.

Compared to the level set method based on the C-V model, our method has faster convergence speed and higher quality segmentation results. In the six experiments, to achieve the final curve from the initial curve, our method separately iterated 5, 5, 5, 30, 40 and 45 times, while the level set method based on the C-V model needs 5, 5, 5, 150, 50, 60 iterations. The segmentation did not completely extract the correct boundary due to noise (Fig.2 and Fig.5), the large "internal" intensity variations in the structures to be identified (Fig.3), and the overlapping distribution of intensity values corresponding to one structure with those of another structure (Fig.4). These problems have less influence on our method using the prior



(a)

(b)

(c)

Fig.2 Segmentations of (a) rhombus, (b) ellipse and (c) circularity in synthetic image

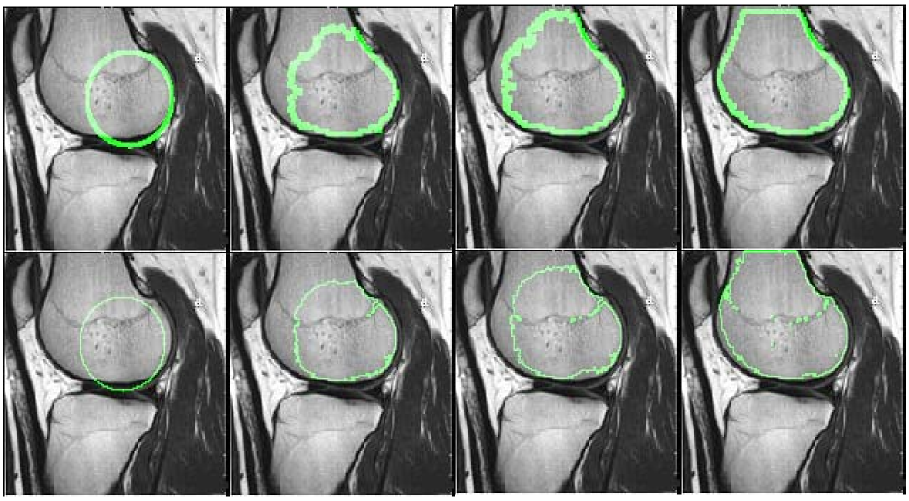


Fig.3 Segmentation of femur in MR image of the knee

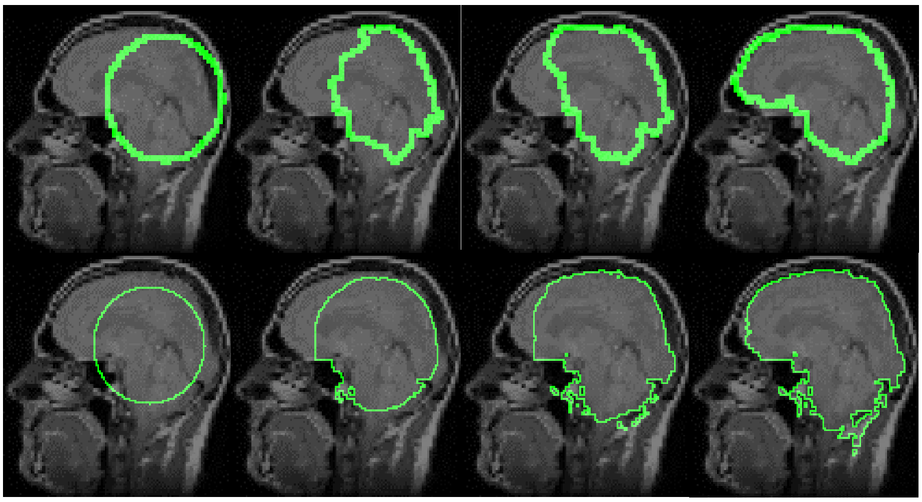


Fig.4 Segmentation of brain in sagittal MR image

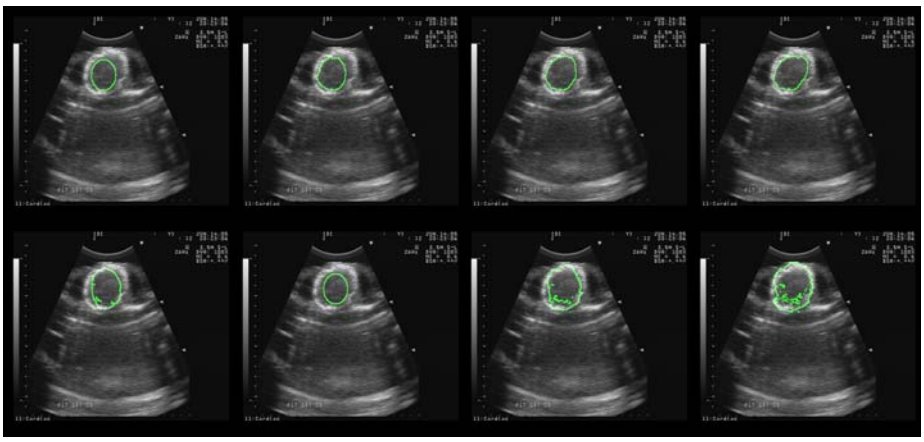


Fig.5 Segmentation of egg in ultrasonic image

model.

The results presented here were on 2D imagery. Now more and more 3D images need to be segmented in medical image processing applications. We will extend the algorithm to 3D in future work.

References

Adalsteinsson, D., Sethian, J.A., 1995. A fast level set method for propagating interface. *Journal of Computational Physics*, **118**(2):269-277. [doi:10.1006/jcph.1995.1098]

Chan, T.F., Vese, L.A., 2001. Active contours without edges. *IEEE Transactions on Image Processing*, **2**(10):141-151.

Hampton, C., Persons, T., Wyatt, C., Zhang, Y., 1998. Survey of Image Segmentation. Available: http://citeseer.ist.psu.edu/hampton98survey.html.

Leventon, M.E., Faugeras, O., Grimson, W.E., Wells, W.M., 2002. Level Set Based Segmentation with Intensity and Curvature Priors. Biomedical Imaging, IEEE International Summer School, Brittany.

Sethian, J.A., 1999. Level Set Methods and Fast Marching Methods: Evolving Interfaces in Computational Geometry, Fluid Mechanics, Computer Vision, and Materials Science. Cambridge University Press, New York.

Vapnik, V., Mukherjee, S., 2000. Support Vector Method for Multivariate Density Estimation. Advances in Neural Information Processing Systems, MIT Press, Massachusetts, p.659-665.