# A Reinforcement Learning Framework for Medical Image Segmentation

Farhang Sahba, Member, IEEE, and Hamid R. Tizhoosh, and Magdy M.A. Salama, Fellow, IEEE

Abstract—This paper introduces a new method to medical image segmentation using a reinforcement learning scheme. We use this novel idea as an effective way to optimally find the appropriate local thresholding and structuring element values and segment the prostate in ultrasound images. Reinforcement learning agent uses an ultrasound image and its manually segmented version and takes some actions (i.e., different thresholding and structuring element values) to change the environment (the quality of segmented image). The agent is provided with a scalar reinforcement signal determined objectively. The agent uses these objective reward/punishment to explore/exploit the solution space. The values obtained using this way can be used as valuable knowledge to fill a Q-matrix. The reinforcement learning agent can use this knowledge for similar ultrasound images as well. The results demonstrate high potential for applying reinforcement learning in the field of medical image segmentation.

#### I. INTRODUCTION

Many applications in medical imaging need to segment an object in the image [1]. Ultrasound imaging is an important image modality for clinical applications. The accurate detection of the prostate boundary in ultrasound images is crucial for diagnostic tasks [2]. However, in these images the contrast is usually low and the boundaries between the prostate and background are fuzzy. Also speckle and weak edges make the ultrasound images inherently difficult to segment. The prostate boundaries are generally extracted from transrectal ultrasound (TRUS) images [2]. Prostate segmentation methods generally have limitations when there are shadows with similar gray level and texture attached to the prostate, and/or missing boundary segments. In these cases the segmentation error may increase considerably. Another obstacle may be the lack of a sufficient number of training (gold) samples if a learning technique is employed and the samples are being prepared by an expert as done in the supervised methods. Algorithms based on active contours have been quite successfully implemented with the major drawback that they depend on user interaction to determine the initial snake. Therefore, a more universal approach should require a minimum level of user interaction and training data

Farhang Sahba is with the Pattern Analysis and Machine Intelligence Laboratory, Department of System Design Engineering, University of Waterloo, Waterloo, Ontario, Canada (email: fsahba@uwaterloo.ca).

Hamid R. Tizhoosh is with the Pattern Analysis and Machine Intelligence Laboratory, Department of System Design Engineering, University of Waterloo, Waterloo, Ontario, Canada (email: tizhoosh@uwaterloo.ca).

Magdy M.A. Salama is with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, Ontario, Canada (email: msalama@hivolt.uwaterloo.ca).

Considering the above factors our new algorithm based on reinforcement learning (RL) is introduced to locally segment the prostate in ultrasound images. The most important concept of RL is learning by trial and error based on interaction with the environment [3], [4]. It makes the RL agent suitable for dynamic environments. Its goal is to find out an *action policy* that controls the behavior of the dynamic process, guided by signals (reinforcements) that indicate how well it has been performing the required task.

In the case of applying this method to medical image segmentation, the agent takes some actions (i.e., different values for thresholding and structuring element for a morphological operator) to change its environment (the quality of the segmented object). Also, states are defined based on the quality of this segmented object. First, the agent takes the image and applies some values. Then it receives an objective reward or punishment obtained based on comparison of its result with the goal image. The agent tries to learn which actions can gain the highest reward. After this stage, based on the accumulated rewards, the agent has appropriate knowledge for similar images as well.

In our algorithm we use this reinforced local parameter adjustment to segment the prostate. The proposed method will control the local threshold and the post-processing parameter by using a reinforcement learning agent. The main purpose of this work is to demonstrate this ability that as an intelligent technique, reinforcement learning can be trained using a very limited number of samples and also can gain extra knowledge during online training. This is a major advantage in contrast to other approaches (like supervised methods) which either need a large training set or significant amount of expert or a-priori knowledge.

This paper is organized as follows: Section II is a short introduction to reinforcement learning. Section III describes the proposed method. Section IV presents results and the last part, section V, concludes the work.

## II. REINFORCEMENT LEARNING

Reinforcement learning (RL) is based on the idea that an artificial agent learns by interacting with its environment [3], [4]. It allows agents to automatically determine the ideal behavior within a specific context that maximizes performance with respect to predefined measures. Several components constitute the general idea behind reinforcement learning. The RL agent is the decision-maker of the process and attempts to take an action recognized by the environment. It receives a *reward* or *punishment* from its environment depending on the action taken. The RL agents discover which

actions bring more reward using *exploration* and *exploitation*. The agent also receives information concerning the state of the environment. At the beginning of the learning process the RL agent does not have any knowledge about how promising taking different actions are [3]. It takes the various actions, and observes the results. After a while, the agent has explored many actions which bring the highest reward and gradually begins to exploit them. In fact, the agent acquires knowledge of the actions and eventually learns to perform the actions that are the most rewarding. During this process it tries to meet a certain goal relating to the state of the environment. The reward and punishment could be defined objectively when they are defined using a function; or gained subjectively when they are given to the agent by an experienced operator.

Action policy  $\pi$  is the strategy used by the agent to select an action to change the current state. The agent must make a trade-off between immediate and long-term return. It must explore unseen states as well as the states which maximize its return by choosing what it already knows. Therefore, there needs to be a balance between exploration of unseen states and exploitation of familiar (rewarding) states.

Reinforcement learning learns online, and can continuously learn and adapt while performing the required task. This behavior is useful for many cases like medical imaging where precise learning samples are difficult or impossible to obtain [3], [7].

The design of RL agents is based on the definition of the problem at hand. Figures 1(a) and 1(b) show the general components of reinforcement learning and the model used in our proposed approach, respectively. The agent, which is the decision maker of the process, observes the state of the environment. Then it takes an action based on the former experience associated with the current observation and accumulated reinforcement (reward/punishment). Finally, The agent receives a reward or punishment from its environment depending on the action taken.

Q-Learning, a popular technique proposed by Watkins in 1989, is an iterative method for action policy learning [5], [6]. This off-policy method is one of the most commonly used RL methods used in temporal difference learning [4].

Boltzman policy is frequently used to estimates the probability of taking each action a given a state s. The probability used in this policy is calculated as follows [3]:

$$p(a) = \frac{e^{\frac{Q(s,a)}{\theta}}}{\sum e^{\frac{Q(s,:)}{\theta}}}.$$
 (1)

In this equation  $\theta$  is the temperature. It is initialized with high value and decreases when the numbers of iterations increases. Also there are other policies for Q-Learning such as  $\varepsilon$ -greedy and greedy. The  $\varepsilon$ -greedy performs for some applications better than greedy because in the greedy policy all actions are not explored, while  $\varepsilon$ -greedy selects the action with the highest Q-value in a given state, with probability of  $1-\varepsilon$  and selects other ones with probability of  $\varepsilon$ . Considering action  $a_t$  when visiting state  $s_t$  and

following an action policy such as Boltzman exploration, Q-learning algorithm can be defined as given in Table I.

# TABLE I Q-LEARNING ALGORITHM [3].

Initialize Q(s,a) randomly Repeat (for each episode) Initialize s Repeat (for each iteration of episode) Choose a from s using policy  $\pi$  derived from Q Take action a, observe reward r, and next state s'  $Q(s,a) \leftarrow Q(s,a) + \alpha[r + \gamma \max_{a'} Q(s',a') - Q(s,a)]$   $s \leftarrow s'$  until s is terminal

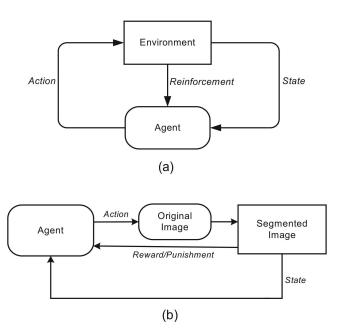


Fig. 1. (a) A general model for Reinforcement learning agent, (b) Model used in proposed approach.

#### III. PROPOSED APPROACH

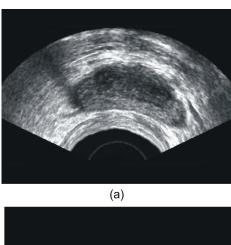
Reinforcement learning has already been used for some other image processing applications [7], [8], [9], [10]. In this paper, we show that it enables us to implement the task of prostate segmentation in a new way.

In our proposed approach we divide the ultrasound image to several sub-images and use two main stages to locally segment the objects of interest. We first threshold the sub-images using local values. Due to some disturbing factors such as speckle and low contrast we usually have many artifacts after thresholding. Therefore, we use morphological opening in a second stage to locally post-process each thresholded sub-image. The reinforcement learning agent determines the local thresholding value and the size of structuring element for each individual sub-image.

To construct the RL agent, three components; states, actions and reward should be defined. The Q-matrix can be

constructed according to definition for states and actions. The RL agent starts its work using an ultrasound image and its manually segmented version. The agent works on each sub-image and using the gold standard (obtained from the manually segmented version) explores the solution space for that sub-image. During this time the RL agent changes the local thresholding values and the size of structuring element for each sub-image individually. By taking each action the agent receives corresponding reward/punishment for that state-action pair and updates the corresponding value in Q-matrix. After this process the agent has explored many actions and tries to exploit the most rewarding ones. This method is specifically useful for prostate ultrasound images where there are several images from a patient having inherently the same characteristics. In such a case, instead of parameter adjustment for each individual input image or using a large training data set to cover all possible cases, we can use some of them and acquire their knowledge to segment the other ones. It is also useful to gain extra knowledge during online training when the agent tries to segment new images.

Figure 2 (a) and (b) illustrate a prostate ultrasound image and its manually segmented version. They can be employed as a sample reference images to gain the knowledge for the RL agent.



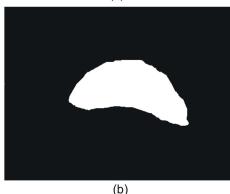


Fig. 2. (a) Original ultrasound image, (b) Its manually segmented version

#### A. Defining the States

To define the states, following features have been considered:

- 1) The location of the sub-images: To segment the image locally, we divide it into  $M_S$  rows and  $N_S$  columns (totally  $M_S \times N_S$  sub-images) and the RL agent works on each of them separately. The location of each sub-image is used as a state parameter.
- 2) Existence of attached parts to the prostate and/or missing boundary segments: Generally, prostate segmentation methods have limitations when the image contains irrelevant parts with similar gray level (usually caused by shadow) attached to the prostate, and/or missing boundary segments. When we threshold the sub-images these attached and missing parts may be revealed as well. The presence and intensity of these parts on the prostate boundary can be evaluated as a state parameter.

In our proposed algorithm we use a method to represents that how much these parts exist on the prostate. To recognize the irregularity in the boundary of the segmented object (in our case prostate) we use signature of a contour combined with an estimator based on Kalman filter [11], [12].

A signature is a functional representation of a contour, generated in various techniques [12], [13]. In our approach we use a signature based on the distance versus angle. In this method, we suppose that the geometric center of the prostate in original image is given by the user and the distance from the points on the boundary to the geometric center of the object, is represented as a  $2\pi$  periodic function. Generally, in a signature one angle  $\theta$  may have several distances r and we may represent it as a 2D function  $f(\theta,r)$  containing the values 0 and 1. But because we want to find the points where an irregularity is starting (due to the attached parts and missing boundary segments), for each angle we use the nearest corresponding contour point as measured data. Using this method the signature can always be described as a 1D function.

Because we use the geometric center of the object shape this representation is invariant to translation. Also we normalize r to make this transformation scale invariant. Because we just need to detect abrupt changes in the signature path as irregular points, our method is not sensitive to orientation as well.

To find the points corresponding to irregular parts we can use an estimator based on Kalman filter [11]. We can use some properties of this filter to evaluate the data on the object signature and detect the existence of the attached and/or missing parts. To implement such a technique, we simulate the problem of signature tracing as a dynamic tracking system. In this system the data located on the signature of the segmented object are used as the input (measurement data) for the tracking filter. Using such an estimator the Kalman filter can track the trajectory of the signature for a whole period. Each data on the signature brings updated information for the current and future data. We simulate it as a 1D dynamic movement. For this movement we can consider

the position and velocity as the variables which describe the state of the system. Using this method we can estimate the position and eventually the abrupt changes on the border of the segmented object.

In our case, we have one variable for position and one for velocity. We represent the state variables based on the data located on a signature. Whenever we want to extract the state parameter we can use the geometric center O of the whole segmented area to define such state variables and consider the following state vector:

$$\mathbf{x} = \left[ \begin{array}{c} r \\ \dot{r} \end{array} \right],\tag{2}$$

where r is the distance between the geometric center  $O(x_c, y_c)$  and pixels  $(x_p, y_p)$  located on the border of the prostate (signature value). For each value of r there is a corresponding angel  $\theta$  between the vertical axis and  $\vec{r}$ . Hence, the following equations can be considered for  $r, \dot{r}$  and  $\theta$ :

$$r = \left[ (x_p - x_c)^2 + (y_p - y_c)^2 \right]^{\frac{1}{2}},$$
 (3)

$$\dot{r} = \dot{y}\cos\theta + \dot{x}\sin\theta,\tag{4}$$

$$r = \left[ (x_p - x_c)^2 + (y_p - y_c)^2 \right]^{\frac{1}{2}},$$

$$\dot{r} = \dot{y}\cos\theta + \dot{x}\sin\theta,$$

$$\theta = \tan^{-1}\left(\frac{y_p - y_c}{x_p - x_c}\right),$$
(5)

where  $\dot{r}$  is the radial velocity. Using the above state vector we represent the sequential data on the signature of the detected object in terms of  $\theta$ . This estimator considers a discrete dynamic model contains state and measurement equations:

$$\mathbf{x}_{k} = \mathbf{A}_{k-1} \mathbf{x}_{k-1} + \mathbf{B}_{k-1} \mathbf{W}_{k-1}, \tag{6}$$

$$\mathbf{Z}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{V}_k,\tag{7}$$

where  $\mathbf{x}$  is the state vector in equation 2. Other components are defined as follows:

$$\mathbf{A}_k = \left[ \begin{array}{cc} 1 & T \\ 0 & 1 \end{array} \right],\tag{8}$$

$$\mathbf{B}_k = \left[ \begin{array}{c} \frac{T^2}{2} \\ T \end{array} \right],\tag{9}$$

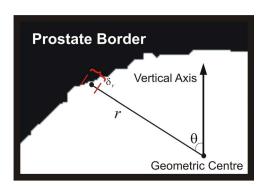
$$\mathbf{H}_k = \left[ \begin{array}{cc} 1 & 0 \end{array} \right], \tag{10}$$

where  $\mathbf{W}_k \sim N(0,Q_k)$  and  $\mathbf{V}_k \sim N(0,R_k)$  are the process and measurement noise, respectively. T is the interval which represents the changes in the state and measurement equations. The value of T does not affect the final result and therefore we can choose it as T = 1 for simplicity. The values of R and Q are the square of measurement and process noise covariance, respectively. The accelerations in the radius is modelled as a zero-mean, white, Gaussian noise **W**. Also the measurement data  $\mathbf{Z}_k$  which is calculated based on the location of boundary pixels is assumed to be a noisy version of the actual position.

Kalman filter starts the estimation using the signature of the segmented object in the thresholded image. In each sequential iteration, the points on the signature (corresponding to the points along the prostate border) are used as measured data and the Kalman filter estimates the next r . These predicted values determine a point as the next one on the signature. Also it predicts  $\dot{r}$  for the next iteration. When we go to the next iteration the new data on the signature is the new measured data for the filter. This data is compared to the predicted position from the previous iteration. If there is sufficient correlation between them the measured data is incorporated to update the filter state, otherwise based on the shape of prostate the prediction point is considered as measured data and filter starts the next iteration. To measure the correlation we implement an association process between the predicted and measured data. For this association process we use an interval  $\delta_r$ , the so-called "association interval" around the predicted point. Only the data located on the signature and inside of this interval are considered as valid measurements for updating the filter. For good performance, the association interval must be adaptive. This means that its size must be varied. It can be changed based on the covariance of the Kalman filter so that it maximizes the presence of valid data and minimize invalid data. In one dimensional problems, it can be represented as the following

$$\alpha.\delta_r \le L,$$
 (11)

The value of L is a constant and  $\alpha$  is the element of the Kalman filter covariance. In the case when there is no data inside the association interval we need to capture it. Therefore the value of L should be larger gradually. Figure 3 shows the r,  $\theta$  and association interval for a sample point on the prostate border.



r ,  $\theta$  and association interval for a sample point on the prostate Fig. 3.

When the border reaches an attached shadow, or a missing boundary segment is encountered, then there is an abrupt change in the pixels' path. These sharp changes are considered as new paths for tracking process. After a few iterations Kalman filter detects that such cases do not belong to the true path because the data do not correlate with the followed path. The association interval is made larger until it again captures the true data on the object signature which have enough correlation with the prediction point. Using the association technique the data belonging to shadows and missing segments on the prostate border are detected.

In the area that border is changing smoothly the measurement data is placed inside the association interval and the value of measurement noise in matrix R should be small. In the case when there is no data inside the association gate we cannot be sure about the validity of measurement data. Therefore, in these cases the value of measurement noise in R should be large. Also the value of process noise in matrix Q simulates the small variation around the estimated point.

Figure 4 (a) illustrates the irrelevant parts that may be revealed after thresholding. In this figure the parts AB,CD and GH are attached parts and EF is a missing boundary segment. Figure 4 (b) shows the points used to make the signature and consequently for the Kalman filter. Also 4 (c) shows the result of kalman filter on the signature of the segmented object in part AB. The estimation of filter for the border is marked with 'x' sign.

The above process needs to be applied on the whole segmented object. Therefore, to find the points corresponding to the attached and missing parts we look at the whole segmented image. When we detect the potential points we note that in which sub-image they are located and follow the local operations.

Using this method, if there exists an attached or missing part on the prostate border we can estimate its thickness as:

 $\Delta_{thickness}$  = Thickness of attached or missing parts

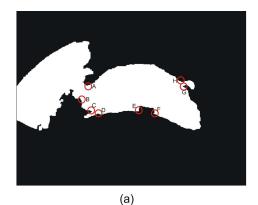
The discretized value of this thickness is used as a parameter to define the state for the RL agent.

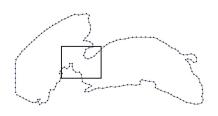
#### B. Defining the Actions

For each sub-image the agent must adjust the threshold value and the size of structuring element for morphological opening. This can be done by increasing and decreasing of the assigned local thresholding value for each sub-images. We can add/subtract a specific value  $(\pm \Delta Tr)$  to increase/decrease the threshold (Tr). Also we can use a simpler way by taking some predefined values  $(T_1, T_1, ..., T_n)$  between the maximum and minimum gray levels in each iteration. For morphological opening we increase/decrease the size of structuring element in a specific interval or choose among some predefined values  $(s_1, s_2, ..., s_n)$ .

#### C. Defining Reward/Punishment

In order to define an objective reward/punishment we need to have a criterion for how well the object has been segmented in each sub-image. We can use several criteria for this purpose. A straightforward method is to compare the results before and after action based on the quality of segmented objects. To measure this for each sub-image we note that how much the quality is changed after the action. For each sub-image, for high increase in the quality of segmented object the agent receives high rewards, for the





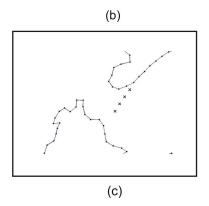


Fig. 4. (a) Attached parts and missing boundary segments, (b) The points used to make the data for object signature and consequently for the Kalman filter, (c) The result of kalman estimator on the signature of the segmented object in part AB. The estimation is marked by 'x' sign.

medium increase it will receive less, and for the decreasing quality it will be punished:

$$reward = \begin{cases} \epsilon_1.D_{\Delta} & D_{\Delta} \ge 0, \\ \epsilon_2.D_{\Delta} & D_{\Delta} \le 0, \end{cases}$$
 (12)

where  $D_{\Delta}$  is the normalized difference between the quality measure before and after taking the action which is automatically determined based on increasing or decreasing of the attached or missing parts. In this equation  $\epsilon_1$  and  $\epsilon_2$  are the constant values.

### D. Offline Procedure and Testing

Now the system is completely designed and can start using a reference image and its segmented version. The states and actions are based on what we designed in section III-A and III-B, respectively. The perfect output image is available using manually segmented version. For reward/punishment function, we use the same equation 12 but for the quality measure of each sub-image we calculate in how far the

similarity with the perfect output image is changed after the action was taken. To measure this similarity we can calculate the percentage of the pixels that are the same in the perfect output image and the image segmented by the RL agent.

During this procedure, the system must explore the parameter space. It can be achieved using the Boltzman policy with a high temperature or  $\varepsilon$ -greedy policy. After a sufficiently large number of iterations, the Q-matrix is filled with appropriate values. It means that the agent can estimate the best action for each given state. Then we can use the system on new samples. The agent must find the appropriate thresholding and post-processing parameter (size of structuring element) for each sub-image such that the prostate can be correctly segmented. The system takes its action based on the knowledge it has gained already. After a limited number of iterations the system can recognize the optimal values and segment the prostate.

#### IV. RESULTS AND DISCUSSIONS

In this section we present and discuss the results of the proposed approach. The ultrasound and the manually segmented version illustrated in Figure 2(a) and (b) can be used as sample. We implemented an  $\varepsilon$ -greedy policy to explore/exploit the solution space. The ultrasound image was divided to  $M_S=3$  rows and  $N_S=4$  columns. The number of discrete levels for thickness (as over-segmentation and under-segmentation) was set to 9. Because we have 12 subimages in our case, there are  $9\times 12=108$  states in total. The RL agent was trained using a total of 5000 iterations for all sub-images.

The threshold action is defined as increasing/decreasing of a specific value for the current local threshold. This value is equal to  $\frac{1}{10}$  of the difference between the maximum and minimum gray levels for each sub-image or 0 for no change. For the post-processing action (morphological opening operator) we chose the size of structuring element among values 0,5,10 or 20. For calculation of reward we choose  $\epsilon_1=\epsilon_2=10$  (see Eq. 12). After the performing of procedure the Q-matrix was filled with appropriate values. In fact, the agent gained enough knowledge to recognized the optimum values for each sub-image.

In the test stage we used 6 similar sample images from the same patient in order to verify the segmentation results. Figure 5 shows these test images (images I1-I6). In all cases, after a limited number of iterations (usually less than 20 for the conducted experiments) the agent could segment the prostate and terminate the process.

To quantitatively evaluate our results, we have used a similarity measure,  $\eta$ , based on the misclassification rate as a general criterion in image segmentation [14], [15]:

$$\eta = 100 \times \frac{|B_O \cap B_T| + |F_O \cap F_T|}{|B_O| + |F_O|},$$
(13)

where  $B_O$  and  $F_O$  denote the background and foreground of the perfect image (manually segmented),  $B_T$  and  $F_T$  denote the background and foreground area pixel in the result image, and |.| is the cardinality of the set. Table II

shows the summarized results for these images.

TABLE II
RESULTS FOR TEST IMAGES I1-I6 (SEE FIGURE 5)

Sample	I1	12	I3	I4	I5	I6
$\eta\%$	95.5	94.8	96.6	95.7	94.3	95.6

Table II shows that for simple cases the proposed approach has acceptable results to use as the input of a fine tune segmentation algorithm. For instance the result of proposed approach may be used as initial snake for the well-known method introduced in [16] or as a coarse estimation for the methods introduced by authors in [17]. Even in some parts that the original image has good quality, the results of proposed approach can be matched with final segmentation.

#### V. CONCLUSIONS

In this work, a reinforcement learning method as a novel idea for prostate segmentation was proposed and some results were illustrated. First, the image is divided to some subimages. Then in an offline stage, the agent takes some actions (i.e. changing the thresholding value and the size of structuring element) to change its environment (the quality of the segmented parts) in each sub-image. After this stage, the agent takes actions with maximum reward for each possible state for each sub-image. It can choose the appropriate values for the input image with similar characteristics based on its accumulated knowledge. The proposed method can be trained for object segmentation in medical images to achieve an acceptable level of performance. The idea in this method has the potential to be used as the main segmentation approach, or as an interim stage to serve other segmentation methods. This method was applied to some similar test ultrasound images containing prostate. Based on a simple similarity measure, we showed the effectiveness of proposed approach. Our future work will concentrate on extension of the algorithm. Adaptive selection of number of sub-images, and integration of more and robust features will be investigated. Adding other operations like noise filtering to be controlled by the RL agent will be tested. Also, more appropriate quality measures (usually used in medical imaging) must be apply to evaluate the performance more accurately.

#### REFERENCES

- [1] C. Mettlin, American society national cancer detection project, Cancer, vol. 75, pp. 1790-1794, 1995.
- [2] M. F. Insana and Brown D. G., Acoustic scattering theory applied to soft biological tissues, Ultrasonic Scattering in biological tissues, CRC Press, Boca Raton,pp 76-124, 1993.
- [3] R.S. Sutton, and A.G. Barto , Reinforcement Learning, MIT Press, Cambridge, MA, 1998.
- [4] S. Singh, P. Norving, D.Cohn, Introduction to Reinforcement Learning, Harlequin Inc., 1996.
- [5] C.J.C.H. Watkins, P. Dayan,"Q-Learning" Machine Learning, 8, 279-292, 1992.
- [6] C. J. C. H. Watkins, Learning from Delayed Rewards Cambridge: Cambridge University, 1989.
- 7] Tizhoosh, H.R., Taylor, G.W.," Reinforced Contrast Adaptation", Proc. To be published in International Journal of Image and Graphic, 2006.

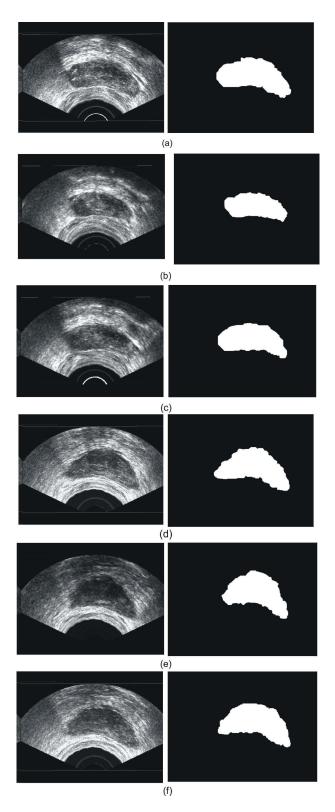


Fig. 5. The original image and its result for test images I1-I6: (a) Image 1, (b) Image 2, (c) Image 3, (d) Image 4, (e) Image 5, (f) Image 6

- [8] M. Shokri., H. R. Tizhoosh, "  $Q(\lambda)$ -based Image Thresholding", Canadian Conference on Computer and Robot Vision, 2004.
- [9] F. Sahba., H. R. Tizhoosh, "Filter Fusion for Image Enhancement Using
- Reinforcement Learning", *CCECE03*, Montreal, May 2003.

  [10] F. Sahba, H. R. Tizhoosh, M. M. A. Salama, "Using Reinforcement Learning for Filter Fusion in Image Enhancement", *The Fourth IASTED* International Conference on Computational Intelligence, July 2005, Calgary, Canada.
- [11] Gelb, A. Applied Optimal Estimation, MIT Press, 1974.
- [12] Rourke, J., "The Signature Of A Plane Curve", SIAM J. Comput, Vol. 15, No. 1, Feb 1986. pp. 34-51
- [13] Holt, R. J. and Netravali, A. N., "Using Line Correspondences In Invariant Signatures For Curve Recognition", Image and Vision Computing, Vol. 11, No. 7, Sept. 1993. Pp. 440-446.
  [14] B. Sankur, M. Sezgin, "Survey over image thresholding technique and
- quantitative performance evaluation", journal of Electronic Imaging, vol. 13, no. 1, pp 146-165, 2004.
- [15] W. A. Yasnoff, J. K. Mui and J. W. Bacus," Error measures for scene segmentation", Pattern Recognition, 9, 1977.
- [16] H. M. Ladak, F. Mao, Y. Wang, D. B. Downey, D. A. Steinman, A. Fenster, "Prostate boundary segmentation from 2D ultrasound images", Medical Physics 27, pp. 1777-1788, 2000.
- [17] F. Sahba, H. R. Tizhoosh, M. M. A. Salama, " A Coarse-to-fine approach to prostate boundary segmentation in ultrasound images", BioMedical Engineering OnLine, 4:58, October 2005.
- [18] M. Sezgin, B. Sankur, "Survey over image thresholding techniques and quantitative performance evaluation", Journal of Electronic Imaging, January 2004, Volume 13, Issue 1, pp. 146-168