A novel image segmentation algorithm based on multi-motive reinforcement learning and OTSU

Qiao Sun[1*], Feixiang Chen[1] , Hui Han[1], Fu Xu[1], Yanan Shi[2]

[1] *School of Information Science and Technology, Beijing Forestry University, Beijing 100083, China*

[2] *College of Computer Science and Technology, Jilin University, Changchun 130012, China*

Received 1 March 2014, www.cmnt.lv

Abstract

Image segmentation is one of the key technologies of computer vision. Among the image segmentation algorithms, the threshold-based approach is a simple and effective one. OTSU algorithm is considered to be one of the best approaches for threshold selection, but its drawbacks are the high time complexity and poor real time capabilities. In order to solve this issue, an efficient image segmentation algorithm based on multi-motive reinforcement learning is proposed in this paper, in the framework of OTSU, multi-motive reinforcement learning algorithm is adopted to get the optimal threshold for image segmentation. The learning motivation and action for threshold learning are defined in this article, and the original State-Action dual-layer structure is extended to State-Motive-Action triple-layer structure. Compared to traditional approach, the proposed approach has more flexibility, and is easier to integrate priori knowledge. The experimental result validated the effectiveness of the proposed approach.

Keywords: Image segmentation, Machine Learning, Computer Vision

1 Introduction

Image segmentation refers to the technique that dividing the image into a number of target areas, separating certain regions from other regions, and obtaining the goal from the images. Image segmentation is the fundamental technology and crucial research area in image processing and computer vision. Image segmentation has been widely used in communications, military, remote sensing image analysis, medical diagnosis, intelligent transportation, agriculture and industrial automation, etc [1].

In recent years, researchers have proposed a variety of image segmentation approaches. In 2007, Zhang et al [2] combined the characteristic that human visual system has different sensitivity to image information content area, improved the usual PCNN using the number of information of spatial neighbourhood with the consistency of neurons near the ignition degree, and proposed an adaptive image segmentation algorithm. In 2008, Berg H. et al [3] designed pulse coupled neurons for image segmentation, they take neurons as recursive function of ADATE system, then use ADATE evolving and producing better neurons automatically for image segmentation. In 2007, Wei et al [4] proposed a fingerprint image segmentation approach using support vector machine. In 2008, Liu et al [5] proposed a multi-scale remote sensing image segmentation by support vector machine. Moreover, in 2008, Liu et al [6] proposed a new remote sensing image segmentation approach based on graph theory. In 2007, Vctor et al [7] proposed an improved watershed algorithm, which by simulating the process of rain, replaces raindrops with pixels to calculate the watershed transform of the grey-scale digital images. Besides, in 2008, Parvati et al [8] proposed a watershed transformation algorithm, which utilizes grey-scale morphology and controls tags.

OSTU is the most commonly used threshold-based approach for image segmentation. However, its drawback is high time complexity, especially when calculates two-dimensional threshold, it has poor real time capabilities. As a result of that the process of choosing threshold is essentially seeking the optimal solution. Hence, in recent years, some researchers utilize evolutionary computation approaches such as genetic algorithm for image segmentation in order to achieve higher efficiency. But the traditional evolutionary algorithm is difficult to converge to global threshold because of the premature convergence, which could affect the accuracy of these kinds of algorithms.

To enhance the efficiency of traditional image segmentation algorithm, in this paper, we present an efficient image segmentation algorithm based on multi-motive reinforcement learning, in the framework of OTSU method, we adopted multi-motive reinforcement learning algorithm to obtain the optimal threshold for image segmentation.

This paper is structured as follows. Section 2 presents preliminaries about reinforcement learning. The detailed algorithms are described in Section 3. And experimental results are presented in Section 4. Section 5 summarizes this paper and draws conclusions.

* ***Corresponding author's*** e-mail: sunqiao19800608@163.com

2 Preliminaries

Reinforcement learning stems from a kind of "trial and error" approach. For instance, in the process of game training, when Agent wins, the trainer gives Agent a positive reward, when Agent fails, the trainer gives negative incentive, and otherwise the result is zero. Thus, Agent could learn to select a series of actions in order to get the highest reward from the indirect tardy rewards. Reinforcement learning process model could be represented by a six-tuple <A, S, R, P, γ, D>. In the six-tuple, A represents action space, S is state space, R is reward function R(s, *a*), indicates the immediate reward the Agent is expected to get when executes action *a* in the state *s*. P represents the probability function P (s, a, s'), indicates the possibility that after Agent executed action *a* then reaches state s' in the state s. γ is discount factor, $0 \le \gamma < 1$. D represents the initial state distribution.

Although reinforcement learning can be applied to image segmentation and has good performance, but it still has some limitations, including: when faced with dynamically changing environment, the target itself may change. In this case, if the target itself has changed, the learned strategies would fail. Therefore, it needs to re-learn the optimal policy. When the targets are not only one, i.e. it is not a problem with a single target but multiple targets; it is difficult for most of the reinforcement learning approaches to adapt the situation. If the environment is complicated, and the state space is huge, it would lead to a bad real time of reinforcement learning. Besides, there are some cases, which the state space is too huge to use reinforcement learning algorithms. When addressing the issue of image segmentation, there often exists some uncertain factors, which makes it difficult for algorithm to obtain accurate environment status and feedback information, it may cause that reinforcement learning algorithm cannot converge.

In 2013, Zhao and Tan[9] proposed a multi-motive reinforcement learning approach, adding motive space to former state space to two-layer map structure of action space, so that transforming it into the three-layer map structure which from the state space to action space, then from motives space to action space. It makes the algorithm could take advantage of priori knowledge. In addition, it enhances the flexibility of reinforcement learning algorithms and improves the efficiency of the algorithm.

3 The proposed methodology

3.1 PROBLEM MODELING

According to traditional threshold based image segmentation approach[10], in general, the grey levels of an image is assumed to be 1-m, and the number of the pixels whose grey level is i is represented as n_i, then the total number of the pixels of the image is as follows:

$$N = \sum_{i=1}^{m} n_i \,. \tag{1}$$

The probabilities of each grey level are as follows:

$$P_i = \frac{n_i}{N} \,. \tag{2}$$

Using variable k to divide them into two groups $A_0=[1,\ldots,k]$, $A_1=[k+1,\ldots,m]$, then the probability of A_0 is:

$$w_0 = \frac{\sum_{i=1}^{k} n_i}{N} = \sum_{i=1}^{k} p_i \,. \tag{3}$$

The probability of A_1 is:

$$w_1 = \sum_{i=k+1}^{m} p_i = 1 - w_0 \,. \tag{4}$$

The average value of grey level of A_0 is:

$$u_0 = \frac{\sum_{i=1}^{m} P_i \times i}{w_0} \,. \tag{5}$$

The average value of grey level of A_1 is:

$$u_1 = \frac{\sum_{i=k+1}^{m} P_i \times i}{w_1} \,. \tag{6}$$

So the grey level of the whole image is:

$$u = \sum_{i=1}^{m} P_i \times i \,. \tag{7}$$

The average value of grey scale whose threshold is k is as follows:

$$u(k) = \sum_{i=1}^{k} P_i \times i \,. \tag{8}$$

The sampled average value is $\mu = w_0 u_0 + w_1 u_1$, the variance is:

$$d(k) = w_0 (u_0 - u)^2 + w_1 (u - u_1)^2 \,. \tag{9}$$

Therefore, the fitness function can be as follows:

$$d(k) = w_0 w_1 (u_1 - u_2)^2 \,. \tag{10}$$

The value of variable k ranges from 1 to m, and we assumed that the k^* can maximize $d(k)$, so k^* is the optimal threshold for image segmentation.

This paper utilizes multi-motivate reinforcement learning algorithm to obtain the value of k *. Reinforcement learning for image segmentation trains each set of data. First randomly assigned to the Agent an initial segmentation threshold, dividing original image, Agent calculates the current state, then select an action to change the current segmentation threshold based on the action selective policy. Defining reward (return value) r as the fitting degree of the current segmented target area and the actual optimal image segmentation result; the better the fitting degree of the segmented target area and the actual optimal image segmented, which the obtained threshold makes, the better the obtained threshold, the bigger positive reward Agent gets. Besides, updates matrix of value of Q. Repeat this cycle until matrix of value of Q converges, then end the learning process.

3.2 STATE REPRESENTATION

Every step of the reinforcement learning is that Agent chooses and executes an optimal action for the current state, and so forth, until it reaches the ultimate aim, i.e. the image is segmented successfully. Thus, in the reinforcement learning for image segmentation, the status of Agent is representation of current image. This paper used the approach, which is presented by Zhu et al., using two-tuple {S1, S2} to represent the status of reinforcement learning. S1 is the overlapping ratio of the object contour edge of the segmentation result of current threshold to the edge that is obtained by edge detection. The ratio is higher means the surface segmentation is better. S2 represents the ratio of the target area of current segmentation result to the target area, which is segmented by OSTU. In summary, state S= {S1, S2}.

3.3 MOTIVE AND ACTION REPRESENTATION

The learner changes current segmentation threshold by perform an action. For instance, we defined 14 motives, of which the first 9 motives are m_1 = {- 50, -30, -10, -1, 0, 1, 10, 30, 50}, representing the increment value to current threshold and corresponding to 9 actions. If the current threshold value is k, then after selecting motive m, the corresponding action is adding m to k, i.e., the threshold value is k + m. So the actions according to motive m_1 are a_1 = {k-50, k-30, k-10, k-1, k, k + 1, k + 10, k + 30, k + 50}.

Since we hope that when the current threshold is poor and far from the optimal value, the threshold would be increased significantly. Besides, when the current threshold is close to optimal value, we are willing to change threshold in smaller range near the local optimal solution, and the change should be slightly. So we set the other five motives as m_2 = {double, increase by 50%, increase by 1%, decrease by 1%, decrease by 50%}, if current threshold is k, then these five motives correspond with action a_2 = {2k, 1.5 k, 1.01k, 0.99k, 0.5k}.

In addition, we want to introduce a priori knowledge, so that action selection could be linked directly with the function d(k) that measures the quality of threshold value. Then set two motives as m_3 = {increase d(k), decrease d(k)}, the corresponding actions are making value d(k) increase or decrease of a_1 and a_2.

In summary, motive m = {m_1, m_2, m_3}, action a = {a_1, a_2}. Of course, more appropriate motives and actions can be designed based on the actual application scenarios.

3.4 POLICY FOR MOTIVE AND ACTION SELECTION

According to the MMQ-voting [9] method that Zhao proposed, the main purpose of learning for function Q(s, m) is learning the Q table corresponding to state-action. The selection of action a is not just simply a choice based only motive m, but multi-motive, which combine many motives which might participate in the selection of action together in the state *s*, based on the value of function Q of these motives, weighted voting for each action, then selecting the appropriate action. Setting the voting weights W (Q) makes sure there is a positive relation between the weights of W(Q) and value of Q. During each step of the learning process, select n motives which have the maximum value of Q, then utilize these n motives and their corresponding actions to weighted vote:

$$vote(a) = \sum_{\substack{\text{All the motives } m \text{ that} \\ \text{corresponding to action } a}} e^{Q(s,m)} \quad . \tag{11}$$

We select the action, which has the maximum value of Q, there is a positive relation between each action and its value of function Q. Thus, the algorithm tends to make the value Q to be bigger in the learning process.
In the MMQ-voting method, the optimal policy to learn can be expressed as:

$$\pi^*(\mathrm{s})=\arg \max_a vote(a) \quad . \tag{12}$$

The right side of the equation represents the selected action after the voting. Function Q (s, m) is rewritten as the following equation:

$$\begin{aligned} Q(s,m) &= \mathrm{r}(\mathrm{s}, \arg \max_a vote(a)) + \\ &\gamma V^*(\delta(s, \arg \max_a vote(a))). \end{aligned} \tag{13}$$

With updating rules for function Q, when updating the evaluation function Q (s, m), you should update all the value Q of motives that all the actions which take part in the selection are corresponding to. The algorithm can be described as follows:

Algorithm Image Segmentation Based on Multi-Motive Reinforcement Learning and OTSU

Initialization the value of Q (s, m);
Set the initial state;
while not convergence do
begin
Calculate the state value *s*.
According to the current state value *s*, choose *n* motives which have maximum value of Q(s,m);
Calculate the vote value of each action by following equation:

$$vote(a) = \sum_{\substack{\text{All the motives } m \text{ that} \\ \text{corresponding to action } a}} e^{Q(s,m)}$$

Select the action *a* which has the highest value of vote(a).
Execute action *a*, get a new image segmentation threshold value k;
According to k, generate a new state value s';
Calculate the fitness of the current divided target area and the actual optimal image segmentation result using k and, then obtain reward value *r*;
According to *r*, update value Q(s,m) of every motive m which participates in the vote:

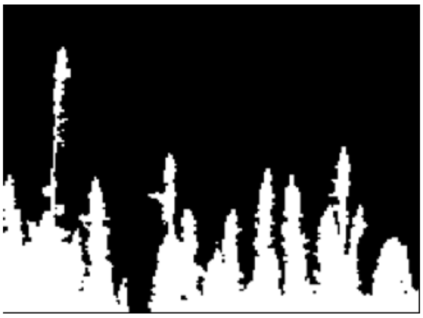
$$\begin{aligned} Q(s,m) &= \mathrm{r}(\mathrm{s}, \arg \max_a vote(a)) + \\ &\gamma V^*(\delta(s, \arg \max_a vote(a))) \end{aligned}$$

S←S';
end
The current value of *k* is the optimal image segmentation threshold for this kind of images. ▌

4 Experimental results and analysis

Experiments are conducted on Intel(R) Core i5 2.50 GHz CPU with a RAM of 4.00 GB. We select two images as experimental images, use the proposed algorithm compared with traditional reinforcement learning image segmentation algorithm. The images both before and after the segmentation are shown in Figure1 and Figure 2.

(a) Scenery image before segmentation



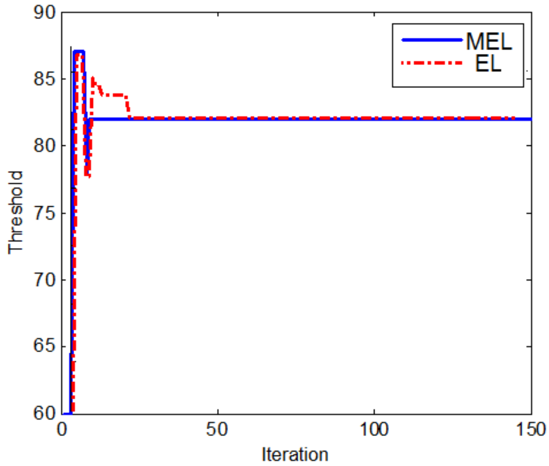
(b) Scenery image after segmentation

FIGURE 1 Segmentation of a scenery image of trees

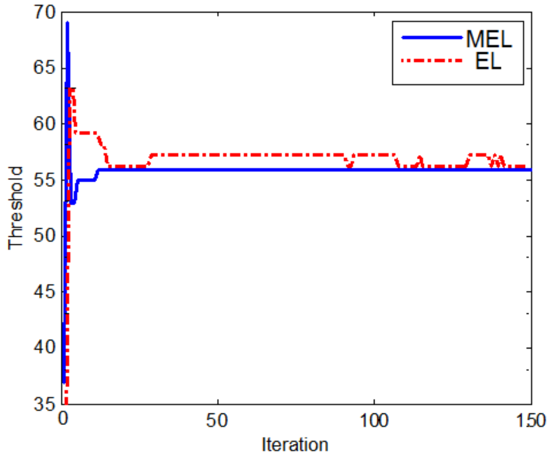
(a) Sky and birds image before segmentation

(b) Sky and birds image after segmentation

FIGURE 2 Segmentation of a sky and birds image



(a) Threshold at each learning iteration for the image1l



(b) Threshold at each learning iteration for the image12

FIGURE 3 Comparison of the learning speed for segmentation, MEL is presented as the proposed Multi-motive

Figure 3 shows comparison about speed of learning segmentation threshold between the proposed approach in this paper and traditional reinforcement learning image segmentation approach. It can be seen that, compared with traditional approaches, the proposed approach can be quicker to learn the optimal threshold. Since the images of the experiment is relatively simple gray-scale image, so the two approaches could both quickly obtain the optimal segmentation threshold, but the proposed algorithm, by introducing multi-motive, makes the learning process more flexible, indirectly related with function d(k), so that the learning process has better and real-time capacity.

5 Conclusions and future works

To improve the efficiency of traditional image segmentation algorithm, in this paper, we present an efficient image segmentation algorithm based on multi-motive reinforcement learning, in OTSU framework, multi-motive reinforcement learning algorithm is adopted to obtain the optimal segmentation threshold. The state of reinforcement learning is presented as tuple $\{S_1,S_2\}$. S_1 is the overlapping ratio of the object contour edge of the segmentation result of current threshold to the edge that is obtained by edge detection. S_2 represents the ratio of the target area of current segmentation result to the target area, which is segmented by OSTU. The reward is defined as the fitness of the current segmented target area and the actual optimal image segmentation result. The selection of action is based on multiple motives according to the value of Q function of these motives and weighted voting for each action. Compared to traditional approaches, the proposed approach has more flexibility and is facilitate to integrate priori knowledge. The experimental result illustrated the effectiveness of the approach. Future research will be focus on improve the efficiency of the algorithm combined with Bayesian networks based machine learning algorithms [11-15].

Acknowledgments

This research is supported by the Fundamental Research Funds for the Central Universities (TD2014-02). Thanks for Yu He's contribution to part of the experimental source codes.

References

[1] Wen J, Yan Z, Jiang J 2014 Novel lattice Boltzmann method based on integrated edge and region information for medical image segmentation *Bio-Medical Materials and Engineering* **24**(1) 1247-52
[2] Zhang J, Fan X, Dong J, Shi M 2007 Image segmentation based on modified pulse-coupled neural networks *Chinese Journal of Electronics* **16**(1) 119-22
[3] Berg H, Olsson R, Lindblad T 2008 Automatic design of pulse coupled neurons for image segmentation *Neurocomputing* **71**(6) 1980-93
[4] *Deleted by CMNT Editor*
[5] Liu Ting ,Wen Xian- bin, Quan Jin-juan 2008 Multiscale SAR image segmentation using support vector machines *Proceedings of the 2008 Congress on Image and Signal Processing USA: IEEE, 2008* 706-9
[6] Liu B, Tian Z, Li X, Zhou Q 2008 Multiscale SAR Image Segmentation Based on Gomory-Hu Algorithm *Journal of Astronautics* **29**(3) 1002-7
[7] *Deleted by CMNT Editor*
[8] Parvati K, Prakasa Rao B S, Mariya D M 2008 Image segmentation using gray- scale morphology and marker- controlled watershed transformation. *Discrete Dynamics in Nature and Society* 1- 8
[9] Zhao F, Tan Z 2013 Multi-Motive Reinforcement Learning Framework. *Computer Research and Development* **56**(2) 240-7
[10] Otsu N 1979 A Threshold Selection Method from Gray Level Histogram *IEEE Transactions on System Man and Cybernetics* **9**(1) 62-6
[11] Zhu Y, Liu D, Chen G, Jia H, Yu H 2013 Mathematical modeling for active and dynamic diagnosis of crop diseases based on Bayesian networks and incremental learning *Mathematical and Computer Modeling* **58**(3-4) 514-23
[12] Zhu Y, Liu D, Jia H, Trinugroho D 2012 Incremental Learning of Bayesian Networks based on Chaotic Dual-Population Evolution Strategies and its Application to Nanoelectronics *Journal of Nanoelectronics and Optoelectronics* **7**(2) 113-8
[13] *Deleted by CMNT Editor*
[14] Zhu Y, Liu D, H Jia 2011 A New Evolutionary Computation Based Approach for Learning Bayesian Network *Procedia Engineering,* **15**(8) 4026 - 30
[15] Gamez A, Mateo J, Puerta M. 2011 Learning Bayesian networks by hill climbing: efficient methods based on progressive restriction of the neighborhood *Data Mining and Knowledge Discovery* **22**(1-2) 106-48

Authors

Qiao Sun, 1977.06, Changchun City, Jilin Province, P.R.China

Current position, grades: Lecturer of School of Information Science & Technology, Beijing Forestry University, China.
University studies: B.Sc. and M.Sc.in Electrical Engineering from Chanchun University of Technology in China. She received her PHD from Beihang University in China.
Scientific interest: Her research interest fields include machine learning and data mining.
Publications: more than 6 papers published in various journals.
Experience: She has teaching experience of 12 years, has completed two scientific research projects.

Feixiang Chen, 1977.11, Jingmen City, Hubei Province, P.R. China

Current position, grades: Professor of School of Information Science and Technology, Beijing Forestry University, China.
University studies: B.Sc. and M.Sc. in Computer Science and Technology from China University of Geosciences. He received his PhD. from Chinese academy of sciences.
Scientific interest: His research interest fields include Moible GIS, 3D GIS.
Publications: more than 30 papers published in various journals.
Experience: He has teaching experience of 8 years, has completed 6 scientific research projects.

Fu Xu, 1979.07, Weihai City, Shandong Province, P.R. China

Current position, grades: the Associate Professor of School of Information and Technology, Beijing Forestry University, China
University studies: received her B.Sc. and M.Sc.in Electrical Engineering from Chanchun University of Technology in China. She received her PHD from Beihang University in China.
Scientific interest: Her research interest fields include machine learning and data mining.
Publications: more than 6 papers published in various journals.
Experience: She has teaching experience of 12 years, has completed two scientific research projects.

Hui Han, 1976.09, Taian City, Shandong Province, P.R. China

Current position, grades: the Lecturer of School of Information Science and Technology, Beijing Forestry University, China.
University studies: received her D.E. in department of automation from Tsinghua University in China.
Scientific interest: His research interest fields include data mining, machine learning.
Publications: more than 6 papers published in various journals.
Experience: She has teaching experience of 7 years, has completed two scientific research projects.

Yanan Shi, 1993.12, Linfen City, Shanxi Province, P.R. China

Current position, grades: undergraduate student, College of Computer Science and Technology, Jilin University, China.
University studies: will receive her B.Sc. in Computer Science and Technology from Jilin University in China.
Scientific interest: Her research interest fields include machine learning
Experience: she is doing one scientific research project.