

Title	運動者に対するランニング経路推薦のための方策勾配法に基づくランニング経路生成方法の研究
Author(s)	小倉, 裕平
Citation	
Issue Date	2018-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/15138
Rights	
Description	Supervisor:Ho Bao Tu, 先端科学技術研究科, 修士 (知識科学)

Generating Running Route based on Policy Gradient for Running Route Recommendation

Yuhei Ogura

Graduate School of Advanced Science and Technology,
Japan Advanced Institute of Science and Technology
February 2018

Keywords: running route recommendation, reinforcement learning, policy gradient, deep learning

Recently, the number of runners is increasing because of rising health consciousness. With the development of smartphones and wearable devices, it has become possible to measure and record running distance, heart rates and so on. There are many apps for runners available on smartphones and wearable devices. Some apps recommend running routes to runners, but they take into consideration only time and distance. “TabiRun”-running while traveling- is prevalent and thus the demand for recommending running routes including interesting places is increasing. As stated above, a system for recommending running routes using not only time and distance but also other data is significantly beneficial.

This research objective is to develop effective methods for recommending running routes using not only distance and time but also more data such as speed change and running direction in the running route at the time of runner's past running.

In this research, in order to recommend the running route to runners, we proposed a method to generate the running route using policy gradient method which is one kind of reinforcement learning method. Since the state which is received by the reinforcement learning agent and the output value of the policy are taken as continuous values, the policy gradient method which is trained in handling continuous state and action space was adopted. In addition, inverse reinforcement learning method was adopted for reward function estimation. Because a probability density function is widely used as the policy function when using

Actor-Critic method to handle continuous values and a deep neural network has high expression power, we designed four kinds of policy function: functions with two features and fourteen features using a probability density function, and functions with two features and fourteen features using deep neural networks.

The evaluation of the running route generated by the reinforcement learning agent was evaluated in the form of measuring trajectory similarity with the running route of the expert using the trajectory similarity evaluation index called Trajectory Similarity Measures. For the purpose of an experimental comparative study, we employed three trajectory similarity measures: DTW (Dynamic Time Warping), Fréchet distance and edit distance. We designed the random policy function that randomly takes action for the comparative evaluation.

As the result of the comparative evaluation, the running route generated using the policy function with the deep neural network has higher similarity to the expert's running route than the reference value which is the trajectory similarity between the running route of the expert and the running route by the random policy function regardless of the number of features. Therefore, by using the proposed method, it can be said that learning of expert data and parameters of each parameterized function were successfully learned.

The running route generated using the policy function with the probability density function has a lower degree of similarity with the running route of the expert than the reference value. This is considered to be caused by the fact that the probability density function does not have higher expressive power than the deep neural network. However, the minimum value of the similarity of the running route generated using the policy function using the probability density function is lower than the minimum value of the reference, and when evaluated with the minimum value, the similarity with the expert's running route is high. In addition to this, learning and using policy function with the probability density function is faster and simpler than policy function with the deep neural network. This suggests that although the policy function using the probability density function is inferior to the policy function using the deep neural network, the minimum learning can be done. Our study shows that the policy function with less features has higher similarity than that with more features. It allows us to develop effective methods for recommending running routes using not only distance and time but also more data such as speed change and running direction in the running route at the time of runner's past running was achieved.

In this research, comparison experiment with the method of related work is not done. This is because it is difficult to implement recommendation system of related work, and it is difficult to compare recommendation systems with different characteristics. Therefore, it may be difficult to say proposed method is better, but it is potentially better. The point that

we used advanced machine learning techniques such as reinforcement learning and deep learning which are not used in related work. To our best knowledge, no other work employs advanced techniques of reinforcement learning and deep learning in recommendation of running routes. This research proposed the method of generating the running route by reinforcement learning, but by improving the proposed method it can be used for a wider range of purposes: generating suitable practice menus that can give the best results in important sports meets for runners and recommending a course that can be enjoyed to drivers. Thus, it can be said that this research contributes as one step in long-term research.