

Enhanced Model-Free Deep-Q Network based PTZ Camera Control Method

Dongchil Kim

Information&Media Research Center
Korea Electronics Technology Institute
Seoul, Korea
dckim@keti.re.kr

Sungjoo Park

Information&Media Research Center
Korea Electronics Technology Institute
Seoul, Korea
bpark@keti.re.kr

Abstract—Recently, the algorithm based on reinforcement learning for controlling PTZ(Pan/Tilt/Zoom) of a camera has been significantly increasing. To improve the accuracy of video analysis in a video surveillance environment, it is essential to get good video sources from a camera such as an appropriate object size or an object not covered by the other object. In this paper, we propose a model-free reinforcement learning based PTZ control method to support improving the accuracy of video analysis. The proposed method fine-tunes the DQN (Deep-Q Network) for smoothly controlling PTZ camera. The simulation results show that the proposed method can automatically and seamlessly manage the PTZ camera according to moving objects.

Keywords—*Deep-Q network; reinforcement learning; PTZ camera; video surveillance system*

I. INTRODUCTION

Reinforcement learning (RL) studies the way that natural and artificial systems can learn to predict the consequences of and optimize their behavior in environments in which actions lead them from one state or situation to the next, and can also lead to rewards and punishments [1]. Such environments arise in a wide range of fields, including video surveillance. The video surveillance system was designed to provide a means for preserving personal security from various security threats. It also allows to record and plays live videos from multiple surveillance cameras such as PTZ(Pan/Tilt/Zoom) camera. PTZ camera can adjust its view by varying pan, tilt, zoom parameters according to the target location. The PTZ camera control can support the enhancement of the video analysis, but currently, the manager is controlling the camera manually. To solve this problem and help to improve the accuracy of video analysis, we propose an enhanced model-free Deep-Q Network based PTZ camera control method. The proposed method places the object or event position in the camera at the center and appropriately scales it. The proposed method helps improve the accuracy of video analysis by solving the problem that the size of the object or event is too large or too small to perform the image analysis well. Also, the proposed method can track an object in real time. To control PTZ cameras more automatically and smoothly than the existing algorithm [2], the proposed method fine-tunes the DQN (Deep-Q Network). The remainder of this paper is organized as follows: Section 2 describes related works for deep reinforcement learning. In Section 3, we describe the concepts and algorithms introduced in the model-free Deep-Q network

based PTZ camera control method. Simulation results are described in Section 4. Finally, Section 5 presents our conclusion.

II. RELATED WORKS

Deep Reinforcement Learning (DPL) has already been applied to a wide range of problems such as robotics, where control policies for robots can now be learned directly from camera inputs in the real world [3]. Reinforcement learning methods can be divided into model-based and model-free. Model-based approaches assume an explicit model of the environment and the agent. The model describes the consequences of actions and the associated returns. From this, optimal policies can be inferred. The model-based descriptions apply to goal-directed decisions, in which choices reflect current preferences over outcomes. Model-free approaches forgo any explicit knowledge of the dynamics of the environment or the consequences of actions and evaluate how good actions are through trial-and-error learning. Model-free values underlie habitual and Pavlovian conditioned responses that are emitted reflexively when faced with certain stimuli. While model-based techniques have substantial computational demands, model-free methods require extensive experience [4]. DQN [5] successfully uses deep neural networks as Q-function to approximate the action values $Q(s, a; \theta)$, where the term θ is parameters of the network and (s, a) represent a state-action pair. Two essential ingredients used in DQN are target network and experience replay. The parameters of the neural network are optimized by using stochastic gradient descent to minimize the loss, where the term $\bar{\theta}$ represents parameters of a target network as (1) [6]. To automatically and smoothly control PTZ cameras, we fine-tune DQN architecture.

$$(r_t + \gamma \max_a Q_{\bar{\theta}}(s_{t+1}, a') - Q_{\theta}(s_t, a_t))^2 \quad (1)$$

III. PTZ CAMERA CONTROL METHOD

To automatically and smoothly control PTZ cameras, our system has to receive the position and size of an event or an object. In this paper, we assume that the location and size of the object or event information are received in real time from the video analysis system. The goal of PTZ camera controls is to place the object or event position in the camera at the center and appropriately scales it.

This work was supported by the ICT R&D program of MSIP/IITP. [2017-0-00250, Intelligent Defense Boundary Surveillance Technology Using Collaborative Reinforced Learning of Embedded Edge Camera and Image Analysis.

A. Network

The proposed method fine-tunes DQN architecture. DQN can outperform humans in various Atari games. Using the visual information of the screen as input and the game's intrinsic reward, the DQN agent learns the Q-values of the individual actions for the input image. To deal with conventional instabilities of reinforcement learning when dealing with high-dimensional inputs, a technique called experience replay is utilized to decorrelate the observations and to smooth out the distribution of the training data. Instead of using the observation in sequence, the network's observations are saved in memory and randomly sampled when updating its weights. Unlike the Atari playing DQN, the proposed method does not use the image as input data but uses x, y, width, height information of the object or event received in the video analysis system. To train PTZ camera control actions, we modified the DQN structure as shown in Fig. 1. We update the parameters of a network(θ) that estimates the value function, directly from on-policy samples of experience, $s_t, a_t, r_t, s_{t+1}, a_{t+1}$, drawn from the algorithms interactions with the environment. We update the network parameters to minimize the error difference between the current Q value and the predicted Q value as (2). We train PTZ camera control value using modified DQN.

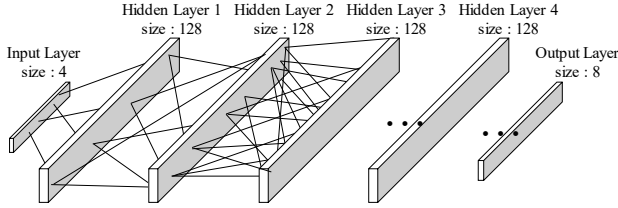


Fig. 1. Network architecture for controlling PTZ camera.

$$\min_{\theta} \sum_{t=0}^T \left[\hat{Q}(s_t, a_t | \theta) - (r_t + \gamma \max_a \hat{Q}(s_{t+1}, a | \theta)) \right]^2 \quad (2)$$

B. Reward function

Before defining reward policies, we defined action values for PTZ camera control. The action value of the PTZ camera is defined as (0.02, 0.06, 0.1, 0.14, 0.18, 0.22, 0.26, 0.3), and the defined action value is selected as (3). The action value a_t is randomly selected only if the reduced ε value is greater than the threshold value $rand$, otherwise, the best action value is selected at the current state s_t . ε value is updated as (4).

$$a_t = \begin{cases} \text{random action, } rand \leq \varepsilon \\ \arg \max_a Q(\phi(s_t)), \text{ otherwise} \end{cases} \quad (3)$$

$$\varepsilon = \varepsilon \times \varepsilon_{decay} \quad \text{subject to} \quad \varepsilon > \varepsilon_{min} \quad (4)$$

After moving the PTZ camera according to the action value selected from the defined action value, Reward is adaptively applied to the positional change of the moving object. First, we calculate Reward corresponding to the movement direction and

action value by using (5), (6), (7) when PTZ camera movement about Pan and Tilt occurs. Where $errC$ is the difference between the position of the object and the center of the frame in the state before the camera movement, $errN$ is the distance between the location of the object and the center of the frame changed after the camera movement.

$$r_{PT} = \begin{cases} +3, & errN \leq \tau_1 \\ f_1(errC, errN), & \text{otherwise} \end{cases} \quad (5)$$

$$f_1(errC, errN) = \begin{cases} -1, & errC \leq errN \\ +1, & \text{otherwise} \end{cases} \quad (6)$$

$$r = r_{PT} - \left(\frac{errN}{\alpha_1} \right) \quad (7)$$

The Reward for Zoom is calculated using (7), (8). Where $errCZ$ and $errNZ$ are the object size information before and after the camera movement respectively, α_1 and α_2 are the normalization constants 100 and 10 respectively, β is a constant 1.2 reflecting $errNZ$ and τ_2 is the target size 70 of the Zoom.

$$r = \begin{cases} -1, & errCZ \leq errNZ \\ 2 + f_2(errCZ, errNZ), & \text{otherwise} \end{cases} \quad (8)$$

$$f_2(errCZ, errNZ) = \begin{cases} 3 + \left(\frac{errNZ}{\alpha_1} \right), & errNZ \leq \tau_2 \\ - \left(\frac{errNZ \times \beta}{\alpha_2} \right), & \text{otherwise} \end{cases} \quad (9)$$

C. Training

To train PTZ camera control values, we collected the number of PTZ camera control actions about 4000 as shown in Fig. 2. We generate the random number of events based on a predefined camera preset so that the position of events can be randomly generated. It randomly assigned 335 locations for each preset to collect about 4000 training data from a total of 12 presets.

| Event status information before camera moves (position x, position y, width, height) | | | | Moving prediction value and Reward value (moving value, reward) | | Event status information after camera movement (position x, position y, width, height) | | | | Episode Termination Conditions (episode done flag) | |
|---|------|-----|----|--|---|---|------|-----|----|---|-------|
| 1 | 159 | 899 | 62 | 62 | 1 | -5.04 | 356 | 899 | 63 | 63 | FALSE |
| 2 | 355 | 899 | 63 | 63 | 2 | -1.78 | 682 | 900 | 64 | 64 | FALSE |
| 3 | 682 | 899 | 63 | 63 | 6 | -0.18 | 682 | 422 | 63 | 63 | FALSE |
| 4 | 682 | 422 | 63 | 63 | 3 | -0.78 | 1138 | 423 | 63 | 63 | FALSE |
| 5 | 1138 | 423 | 64 | 64 | 0 | 2.87 | 1073 | 422 | 63 | 63 | FALSE |
| ⋮ | | | | | | | | | | | |
| 2996 | 912 | 402 | 63 | 63 | 1 | 3.71 | 912 | 511 | 63 | 63 | FALSE |
| 2997 | 913 | 511 | 62 | 62 | 5 | 5.11 | 892 | 494 | 89 | 89 | TRUE |
| 2998 | 446 | 415 | 62 | 62 | 7 | -3.65 | 1425 | 414 | 63 | 63 | FALSE |
| 2999 | 1425 | 415 | 63 | 63 | 1 | -1.69 | 1229 | 415 | 62 | 62 | FALSE |
| 3000 | 1229 | 415 | 63 | 63 | 7 | -8.09 | 251 | 415 | 62 | 62 | FALSE |
| ⋮ | | | | | | | | | | | |
| 3996 | 945 | 422 | 62 | 62 | 2 | 3.34 | 945 | 606 | 63 | 63 | FALSE |
| 3997 | 945 | 606 | 63 | 63 | 3 | -0.28 | 945 | 627 | 81 | 81 | TRUE |
| 3998 | 181 | 62 | 63 | 63 | 6 | 3.3 | 1030 | 62 | 63 | 63 | FALSE |
| 3999 | 1031 | 62 | 63 | 63 | 6 | 3.99 | 1030 | 539 | 63 | 63 | FALSE |
| 4000 | 1030 | 539 | 63 | 63 | 6 | 5.07 | 1091 | 535 | 93 | 93 | TRUE |

Fig. 2. Collected training DB sets.

IV. SIMULATION RESULTS

We evaluate the proposed method that can automatically and smoothly control the PTZ camera according to the movement of objects or events. We use the programming language of Python 3.6 to implement the enhanced model-free DQN based PTZ camera control method. The proposed method received the position and size of the object and event from the video analysis system. To test the proposed method, we generated a random number of events from eight pre-defined presets. It collected 240 test data from a total of 8 presets by specifying ten random positions of objects per preset about three cameras. All tested network configurations were trained for 35 epochs with fixed learning rates, as shown in Table 1, with no early stops applied.

Fig. 3 shows the results of PTZ camera control based on moving objects or events. The average accuracy of the proposed method is 61.4%. The reason why the average accuracy is low is that the location and size of the object or the event received by the video analysis system are inaccurate or lack of data learning in various environments. Fig. 4 shows that the implementation results of the proposed method in real situations. The proposed method can position the object in the center of the camera regardless of the position of the object in the camera. The proposed method can automatically and smoothly track the object by control PTZ camera. Through this experiment, it is possible to improve the accuracy of video analysis by appropriately adjusting the position and size of the object using the proposed method.

TABLE I. HYPER-PARAMETERS VALUES EXPLORED

| Hyper-parameters | Vales explored |
|---------------------|--|
| Optimizer | Adam |
| Loss function | MSE(Mean Square Error) |
| Activation function | ReLU(Hidden layer) Linear(Output layer) |
| Learning rate | 0.001 |
| Batch size | 128 |
| Epochs | 35 |
| Episode | 1000 |

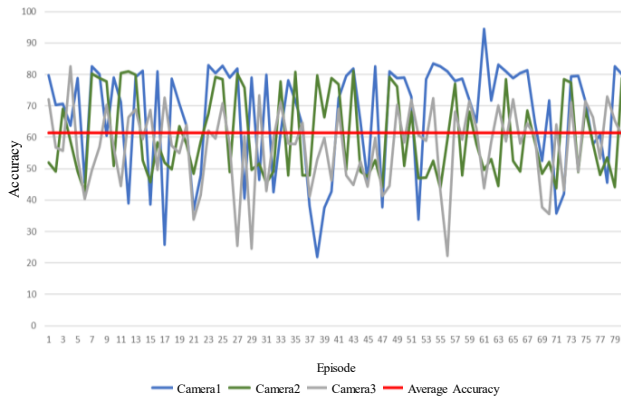


Fig. 3. Simulation results of the proposed method.



Fig. 4. Implementation results of the proposed method.

V. CONCLUSIONS

This paper proposes the model-free reinforcement learning based PTZ control method to support improving the accuracy of video analysis. The goal of the proposed method is to automatically and smoothly control PTZ camera to minimize the distance between the event position and camera center position and adjust the size of objects properly. Through simulation and experiment results, the proposed method automatically and smoothly regulates the location and size of the object by controlling the PTZ camera based on the modified DQN algorithm.

REFERENCES

- [1] P. Dayan and Y. Niv, "Reinforcement learning: The Good, The Bad and The Ugly," *Current Opinion in Neurobiology*, vol. 18, pp. 1-12, Sept. 2008.
- [2] D. Kim and S. Park, "Automatic PTZ Camera Control Based on Deep-Q Network in Video Surveillance System," *Proceedings of International Conference on Electronics, Information and Communication*, pp. 479-481, Jan. 2019.
- [3] K. Arulkumaran, M. P. Deisenroth, M. Brundage, A. A. Bharath, "A Brief Survey of Deep Reinforcement Learning," *IEEE Signal Processing Magazine*, Special Issue on Deep Learning for Image Understanding, pp.1-16, Sept. 2017.
- [4] Q. J. M. Huys, A. Cruickshank, and P. Seriès, "Reward-based Learning, Model-based and Model-free," In *Encyclopedia of Computational Neuroscience*, pp. 2634-2641. Feb. 2014.
- [5] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland et al, "Human-level Control Through Deep Reinforcement Learning," *Nature*, pp. 5829-5833, Feb. 2015.
- [6] Z. Lin, T. Zhao, G. Yang, and L. Zhang, "Episodic Memory Deep Q-Networks," *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence*, pp. 2433-2439, May 2018.