Bandwidth Enhancement of Pixelated Patch Antennas Based on Localized Reinforcement Learning

Qipeng Wang
AVIC Research Institute for Special Structures of Aeronautical Composites, Jinan 250023, China
wqpseu637@163.com

Zhongxuan Pang
AVIC Research Institute for Special Structures of Aeronautical Composites, Jinan 250023, China.
Pangzx637@163.com

Di Gao
AVIC Research Institute for Special Structures of Aeronautical Composites, Jinan 250023, China.
seugaodi@163.com

Peng Liu
AVIC Research Institute for Special Structures of Aeronautical Composites, Jinan 250023, China.
pliu_sunny@163.com

Xuewei Zhang
AVIC Research Institute for Special Structures of Aeronautical Composites, Jinan 250023, China.
Zhangxuewei2016@163.com

Xiaoyu Pang
AVIC Research Institute for Special Structures of Aeronautical Composites, Jinan 250023, China.
pangxiaoyu2004@126.com

Xiaoxing Yin
State Key Laboratory of Millimeter Waves, Southeast University, Nanjing 210096, China
xxyin@seu.edu.cn

2024 14th International Symposium on Antennas, Propagation and EM Theory (ISAPE) | 979-8-3503-5312-9/24/$31.00 ©2024 IEEE | DOI: 10.1109/ISAPE62431.2024.10840429

***Abstract*—This paper employs a reinforcement learning algorithm to improve bandwidth for a pixelated patch antenna. Differing from traditional methods that depend on manual tuning and extensive computations, this approach autonomously discerns the relationships between numerically characterized pixelated patch antennas and their S-parameters. Specifically, the implementation of a double deep Q-network (DDQN) and prioritized experience replay allows the model to dynamically adjust learning processes based on a structured exploration of state and action spaces. In this case, pixelated patch antenna configurations are efficiently optimized for maximal bandwidth performance. The learning model systematically tests and refines antenna designs by iteratively adjusting the presence or absence of metal at each pixel, thereby achieving a significant bandwidth increase to 34%. This method not only optimizes the bandwidth efficiently but also suggests promising applications in complex antenna designs.**

Keywords—Bandwidth optimization, deep Q-network (DDQN), pixelated patch antenna, reinforcement learning

I. Introduction

As communication technologies continuously advance, the demand for broadband antenna designs has been steadily increasing [1], [2]. Rationally dividing the patch and optimizing the metal distribution within each unit have demonstrated the potential to enable broadband design. This approach, known as pixelated design, designates each unit as a pixel [3]-[5]. Conventional antenna design typically requires a profound knowledge reserve and a large number of EM simulation parameter scans to identify the relationship between structural parameters [6]. Inevitably, too much EM simulation is consumed during the optimization process. Even worse, it is commonly found that the design results fail to meet the set objectives. In addition, the design of pixelated patch antennas involves too many variables, rendering traditional design methods impractical. Consequently, there is a high demand for an effective method to optimize the bandwidth of pixelated patch antennas.

Recent efforts in optimizing the bandwidth of the pixelated patch antenna have mainly focused on two approaches: surrogate model-based optimization and evolutionary algorithm optimization [3]-[7]. For instance, by employing Gaussian process regression to construct a surrogate model, a relative bandwidth of 26.2% is achieved for the pixelated antenna by particle swarm optimization (PSO) [8]. In [3], [4], broadband pixelated patch antenna design is also realized by utilizing the brainstorming algorithm and genetic algorithm. Although these methods have identified broadband design solutions through hundreds of iterative optimizations, the bandwidth potential of pixelated patch antennas is far from fully realized. More critically, these algorithms require manual hyperparameter tuning and fixed design objectives, limiting their ability to learn autonomously [9]. Variations in parameter settings, initial sample selection, and iteration counts can all significantly influence the final optimization results.

This paper proposes a localized reinforcement learning (LRL) algorithm aimed at enhancing the bandwidth of pixelated patch antennas. This method strategically adjusts pixel configurations to optimize performance uniquely tailored to the specific properties of the antenna design. The 0-1 matrix representing the pixelated antenna serves as the action space, and the reflection coefficient within the desired frequency band functions as the state space. Starting from a traditional E-shaped antenna, the matrix configuration is modified. Based on the rewards obtained from different matrix configurations, a deep Q-network assigns specific Q-values to various pixel positions in each episode, ultimately producing the optimal bandwidth. To avoid being trapped in local optima, a backtracking strategy is employed. As a result, a pixelated patch antenna design with a relative bandwidth of 34% is achieved.

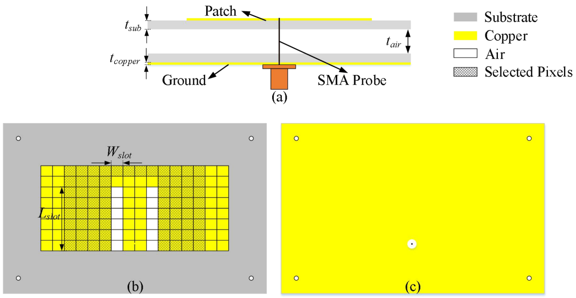

Fig. 1. The configuration of the E-shaped patch antenna. (a) Side view. (b) Top view. (c) Bottom view. Dimensions (mm): W_{slot} = 6.25, L_{slot} = 5.625, t_{sub} = 1.524, t_{copper} = 0.018, t_{air} = 8.

I. LRL Algorithm Framework

A. Knowledge-based Pixelated Patch Antenna.

The proposed pixelated patch antenna consists of upper and lower layers of Rogers 4003C dielectric substrates, with an intermediate air layer separated by nylon spacers to enhance operational bandwidth. The patch and ground plane are positioned on the surfaces of the upper and lower substrates, respectively [10]. The excitation signal is provided to the patch antenna by means of bottom-feeding with an SMA probe, as shown in Fig. 1(a). The overall dimensions of the patch are set to 140 mm × 90 mm, and it is evenly divided into 8 × 16 pixel units. To reduce design dimensionality and minimize cross-polarization, pixelated patch antennas are commonly designed symmetrically. That is, the matrix dimensions are set to 8 × 8. By optimizing the presence or absence of metal in each unit, a broadband antenna design can be achieved. However, the large number of variable pixels complicates optimization and hinders convergence. The E-shape antenna is a typical broadband antenna, in which an initial E-shape pattern can be easily constructed based on prior knowledge [11]-[13]. Subsequently, the E-shaped patch is pixelated and some of the pixel positions that have little effect on bandwidth or are highly susceptible to deteriorating performance are fixed, as depicted in Fig. 1(b). In this way, the dimensionality and complexity of the optimization process can be effectively reduced.

B. Localized Reinforcement Learning

An LRL algorithm is implemented to optimize the design of an 8 × 16 pixelated patch antenna. The framework of the LRL algorithm is depicted in Fig. 2. The primary objective is to iteratively adjust the pixel configuration to enhance antenna performance by minimizing the S_{11} parameter across the desired frequency range. Each episode in the LRL framework begins with an initialized 8 × 8 matrix representing the initial configuration of the pixelated antenna, where matrix elements are binary: '1' indicates the presence of metal, and '0' is absence. The action space of the LRL comprises all possible flip operations, allowing a '1' to be changed to a '0'. The CST electromagnetic (EM) simulation acts as the environment. The simulated reflection coefficients ($|S_{11}|$) of the pixelated patch antenna are the state. The reward function for each frequency between 1.85 GHz to 2.6 GHz is defined by

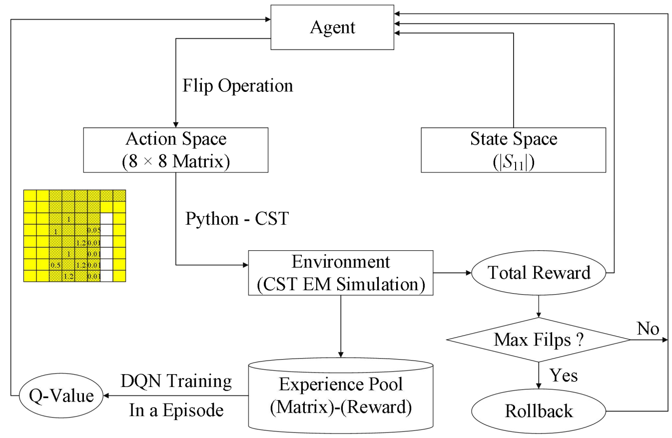

Fig. 2. The framework of the LRL algorithm.

$$reward = \begin{cases} 1 & |S_{11}| < 0.31 \\ -16.67 \cdot \left(|S_{11}| - 0.31\right) + 1 & 0.31 \le |S_{11}| < 0.55 \\ -3 & |S_{11}| \ge 0.55 \end{cases} \quad (1)$$

The Deep Q-Network (DQN) explores this action space within each episode, with actions selected based on the current state. An epsilon-greedy policy is used to balance exploration and exploitation by either selecting a random action (exploration) or choosing the action that maximizes the Q-value (exploitation). In this context, the Q-value estimates the expected cumulative reward from a given state-action pair. The reward function plays a critical role in guiding the learning process. The DQN updates its Q-values based on this reward, refining its policy over time to favor actions that yield higher rewards [14]. A rollback mechanism is employed in each episode to avoid getting trapped in local optima. For example, We set the initial maximum number of flips from '1' to '0' as 15. As the flip reaches the upper limit, it will rollback to the matrix with the maximum reward. Then, the exploration of new flip options continues based on it. This mechanism ensures that the learning process is robust against suboptimal decisions and helps in avoiding local minima. The Q-network is reinitialized at the start of each episode, preventing the transfer of learned Q-values between episodes. This approach ensures that each episode remains independent and concentrates on optimizing the antenna configuration within its specific context. After multiple episodes, the best-performing configuration is identified and further analyzed. The final output includes the optimized 8 × 8 matrix configuration and the corresponding $|S_{11}|$ curve, demonstrating the effectiveness of the LRL algorithm in achieving the desired antenna performance.

II. Algorithm Implementation and Simulation Verification

The implementation of the LRL algorithm begins with the initialization of the environment and agent parameters. The action space consists of 34 variable pixel flip actions, where the agent explores the configuration space by flipping pixels from '1' to '0'. The optimization goal is to minimize the $|S_{11}|$ parameter within the 1.85 GHz to 2.6 GHz frequency range. A DQN is utilized to learn and predict Q-values for each state-action pair.

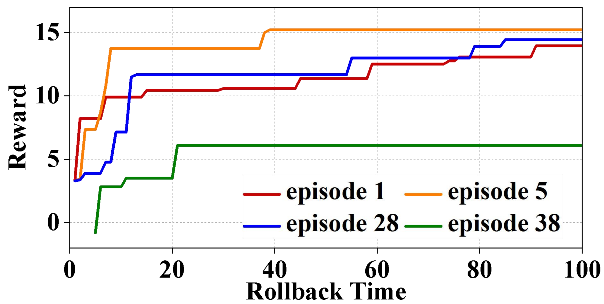

Fig. 3. The highest rewards from the rollback processes across multiple episodes

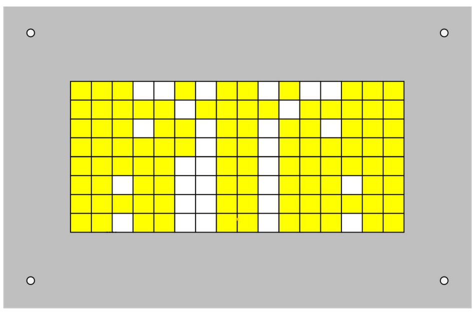

Fig. 4. The optimized pixelated patch antenna.

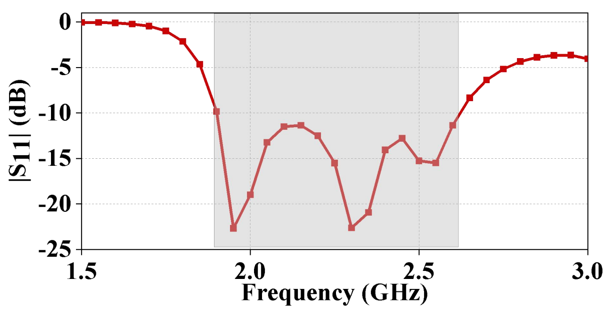

Fig. 5. The reflection coefficient of the optimized pixelated patch antenna.

Employing an epsilon-greedy strategy balances the exploration of new configurations with the exploitation of existing knowledge, enabling the agent to incrementally optimize the antenna configuration [14].

In each episode, the agent executes a series of pixel flips, starting with a maximum of 15 flips and later extending to 20. If no significant improvement in the reward is observed, a rollback mechanism reverts to the configuration with the highest current reward to avoid suboptimal states. The Q-network is re-initialized at the start of each episode, ensuring that optimizations are conducted independently. The highest rewards from the rollback processes across multiple episodes are selected to plot the reward curve, as depicted in Fig. 3. The initial selection of pixel flip positions significantly influences the upper limit of the reward. However, as can be observed in the reward curves of various episodes, the DQN effectively learns the optimization strategy.

After multiple episodes, the matrix with the best performance is selected to realize the corresponding pixelated patch antenna, as shown in Fig. 4. The $|S_{11}|$ curve in Fig. 5 indicates that the proposed pixelated patch antenna maintains a $|S_{11}|$ below -10 dB in the 1.9 GHz to 2.6 GHz, achieving a relative bandwidth of 34%.

III. Conclusion

This study introduces a knowledge-driven localized reinforcement learning approach that effectively optimizes the design of E-shaped pixelated patch antennas, showcasing significant improvements in bandwidth. By integrating reinforcement learning with prior design knowledge, this method efficiently explores the configuration space, resulting in significant improvements in bandwidth. The simulation results validate the effectiveness of this approach, highlighting its potential for practical applications in advanced antenna design.

References

[1] J. Logeshwaran, and S. Karthick, "A smart design of a multi-dimensional antenna to enhance the maximum signal clutch to the allowable standards in 5G communication networks," *ICTACT J. Microelectron.,* vol. 8, no. 1, pp. 1269-1274. 2022.

[2] S. Karapantazis, and F. Pavlidou, "Broadband communications via high-altitude platforms: A survey," *IEEE Commun. Surv. Tutor.,* vol. 7, no. 1, pp. 2-31. 2005.

[3] A. Aldhafeeri, and Y. Rahmat-Samii, "Brain Storm Optimization for Electromagnetic Applications: Continuous and Discrete," *IEEE Trans. Antennas Propag.,* vol. 67, no. 4, pp. 2710-2722. 2019.

[4] F. J. Villegas, T. Cwik, Y. Rahmat-Samii, and M. Manteghi, "A Parallel Electromagnetic Genetic-Algorithm Optimization (EGO) Application for Patch Antenna Design," *IEEE Trans. Antennas Propag.,* vol. 52, no. 9, pp. 2424-2435. 2004.

[5] Q. Wu, W. Chen, C. Yu, H. Wang, and W. Hong, "Machine-Learning-Assisted Optimization for Antenna Geometry Design," *IEEE Trans. Antennas Propag.,* vol. 72, no. 3, pp. 2083-2095. 2024.

[6] !!! INVALID CITATION !!! [6], [7].

[7] B. Liu, H. Aliakbarian, Z. Ma, G. A. E. Vandenbosch, G. Gielen, and P. Excell, "An Efficient Method for Antenna Design Optimization Based on Evolutionary Computation and Machine Learning Techniques," *IEEE Trans. Antennas Propag.,* vol. 62, no. 1, pp. 7-18. 2014.

[8] !!! INVALID CITATION !!! [8], [9].

[9] Y. Su, Y. Yin, S. Li, H. Zhao, and X. Yin, "Bandwidth Improvement for Patch Antenna via Knowledge-Based Deep Reinforcement Learning," *IEEE Antennas Wireless Propag. Lett.* 2024 (Early Access).

[10] Y. He, J. Huang, W. Li, L. Zhang, S.-W. Wong, and Z. N. Chen, "Hybrid Method of Artificial Neural Network and Simulated Annealing Algorithm for Optimizing Wideband Patch Antennas," *IEEE Trans. Antennas Propag.*, pp. 1-1. 2023.

[11]B. Babayigit, E. Senyigit, and G. Mumcu, “Optimum broadband E-patch antenna design with Taguchi method,” *J. Electromagn. Waves Appl.,* vol. 30, no. 7, pp. 915-927. 2016.

[12]A. G. Koutinos, D. E. Anagnostou, R. Joshi, S. K. Podilchak, G. A. Kyriacou, and M. T. Chryssomallis, “Modified easy to fabricate E - shaped compact patch antenna with wideband and multiband functionality,” *IET Microw. Antennas Propag.,* vol. 12, no. 3, pp. 326-331. 2018.

[13]A.-A. E. Fubara, and R. I. Salawu, “Design and Simulation of the Slotted E-Patch Microstrip Dual Band Antenna for Wifi Applications,” *Int. J. Inf. Syst. Eng.,* vol. 6, no. 2, pp. 22-36. 2018.

[14]A. G. Barto, "Reinforcement learning," *Neural systems for control*, pp. 7-30: Elsevier, 1997.