2024 IEEE International Conference on Robotics and Automation (ICRA 2024)
May 13-17, 2024. Yokohama, Japan

Active Automotive Augmented Reality Displays using Reinforcement Learning

Ju-Hyeok Ryu, Chan Kim, and Seong-Woo Kim

***Abstract*— In order to enhance driving convenience and safety, automotive Augmented Reality displays, e.g., head-up displays, have garnered attention and are gradually being deployed. However, when vehicles encounter uneven roads, vertical vibrations lead to mismatches between external physical objects and augmented reality overlay images, adversely affecting the AR display's visibility. Resolving the problem is quite challenging because the optical system operates on a nanometer scale and is highly sensitive due to its multifunctional nature involving reflection and refraction through an intermediate medium. This paper aims to address the newly emerging problem of vertical mismatches in automotive AR displays. To tackle this issue, we begin by defining the problem and then examine the effectiveness of traditional control methods, on-policy and off-policy reinforcement learning as potential solutions. Finally, we validate our approach through experiments, demonstrating a significant reduction in vertical mismatches and an improvement in the overall visibility of automotive AR displays. Our findings provide valuable insights for enhancing driving convenience and safety in real-world conditions.**

I. Introduction

The adaptive optics system of interest in this paper is the Head-Up Display (HUD) used in automotive applications. Initially designed to deliver essential information directly into the driver's line of sight, HUDs have evolved to meet the growing demand for augmented reality displays. This evolution has expanded drivers' field of view and introduced active augmented reality information transmission, incorporating interaction with the surrounding environment. Despite these advancements, significant challenges remain, with one prominent example being the vertical vibration issue encountered by AR displays as vehicles navigate diverse road conditions.

The automotive augmented reality display system comprises three primary components: the object, the vehicle (Display), and the driver (Eyebox) [1]. The dynamic aspects related to the vehicle encompass factors such as the irregularities of the road surface, variations in slope, the suspension system's response, passenger count, and tire pressure. These dynamic variables, combined with the complex freeform design of the vehicle's optical system, pose significant challenges in aligning the augmented reality display with target objects. The virtual image generated by the Picture Generating Unit (PGU) is projected forward through the optical system, appearing as a virtual image within the driver's field of view. This field of view represents the range

J. Ryu is with LG Electronics, Seoul 07796, South Korea. (e-mail: juhyeok.ryu@lge.com)

C. Kim and S. Kim are with Seoul National University, 08826, South Korea. (email: chan_kim@snu.ac.kr, snwoo@snu.ac.kr)

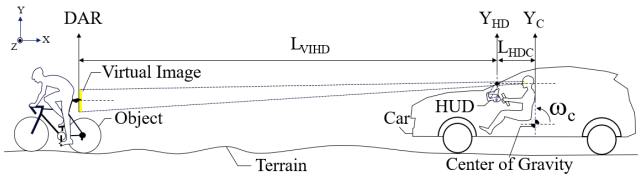

Fig. 1: Vertical oscillations of head-up display on the terrain.

within which the driver's eyes can comfortably move while driving. Given this unique context, real-time tracking of the virtual image's location becomes exceptionally challenging.

In practical terms, addressing this challenge often necessitates the use of simulation-based discrete correction information based on the vehicle's attitude data. However, this approach cannot account for all degrees of freedom and, as a result, cannot guarantee seamless continuity. Table I summarizes potential solutions for mitigating the vertical mismatches in automotive AR displays.

Among the methods, the most traditional numerical analysis approach is the Newton-Raphson method [2]. It demonstrated satisfactory performance in terms of computational efficiency when searching for optimal solutions within the fitted approximate function. However, this method exhibited limitations related to the selection of initial values and curve fitting, preventing it from surpassing the accuracy of the obtained approximation function.

The Extended Kalman Filter (EKF) [3] can be applied to optical systems with off-axis designs, which are highly sensitive to continuity disruptions. However, even a small error in linear optimization within a narrow section of the EKF method can lead to significant amplification along the optical path, making it less suitable for tasks that demand uniformity across the entire area. Conversely, tracking using the Reinforcement Learning (RL) method has demonstrated applicability in achieving superior objective function levels, particularly in scenarios characterized by increasing complexity due to randomness.

In this paper, we conducted a comprehensive study aimed at exploring effective methods for nonlinear model approximation and identifying the optimal combination of methodologies to minimize the impact of environmental factors on the alignment of fixation states between the object and the display. This alignment is particularly susceptible to randomness induced by the vehicle dynamics mentioned earlier. To validate our findings, we conducted physical experiments alongside simulations. The key contributions of this paper can be summarized as follows:

- We formally define the vertical mismatch problem of

TABLE I: Comparison of candidate methodologies to cope with vertical mismatches of automotive AR displays

Item	Newton Raphson	KF	EKF	RL (PPO)	RL (SAC)
Primary update	Estimate of the root	Mean vector/covariance	Mean vector/covariance	Policy parameters	Policy parameters
Update strategy	Iterative refinement using derivatives	Two-step process of prediction and correction	Linearized prediction and correction	Policy gradient optimize-action /w a trust region	Off-policy actor-critic /w entropy maximization
Model Dependency	Very High	High	High	Low	Low
Robustness to randomness	Moderate to Low	Moderate	Moderate	High	High
Computational demand	Moderate to Low	Low to Moderate	Moderate to High	High	High

TABLE II: Comparison of related studies and our study.

Researches / Type of Pendulum	Uncertainty	Model	Method
Charles W. [4] Single Inverted	N/A	Model Free	Neural Network
Bowen Xum, et al. [5] Double Inverted	N/A	Model Based	Kalman Filter
Zabihifar, et al. [6] Rotary Inverted	Oscillation	Model Free	Neural Network
Omid Mofid, et al. [7] Rotary Inverted	External Disturbance	Model Based	ASMC
Wei Chu, et al. [8] Flexible Inverted	Vibration	Model Based	CBCM
Ours	Terrain, Object, Eyebox	Model Free	RL

automotive AR overlay images negatively affecting driver's fatigue and safety.

- We compare various potential solutions, ranging from conventional methods to on-policy and off-policy reinforcement learning, aimed at stabilizing this mismatch. through this, we found the optimal combination of methodologies for the most major case.
- Our simulation results are further validated through rigorous physical experiments.

II. Related Works

Augmented reality displays made their debut in 1901, with research in this domain steadily advancing, particularly since the 2000s. In the automotive industry, BMW introduced a HUD (Head-Up Display) to its 2011 Bimmerfile model, delivering essential information to the driver by projecting it onto the vehicle's windshield [9]. Subsequently, the adoption of AR HUDs by manufacturers has been on the rise [10], emphasizing the importance of a *world-fixed* HUD, which, unlike the screen-fixed type, requires seamless visual integration between AR graphics and the real world, accompanied by an expanded focal distance [1]. Research is also underway to assess the extent to which AR HUDs assist drivers in perceiving the road ahead [11].

One of the key challenges in automotive AR HUDs is maintaining consistency between a real-world target object projected onto the driver's retina and the display pixels projected onto the HUD. This challenge can be reframed as an inverted pendulum or tank stabilizer problem. In the inverted pendulum system, traditionally studied as an optimal control problem under unstable conditions, an appropriate force is applied to the cart, which supports the weight with rotational freedom, to prevent the pendulum from falling onto a fixed plate. Unlike a prior study [4] that employed a Neural Network under the assumption that system modeling might be unattainable, our study faces limitations in expressing the actual system model. The primary causes are compared in Table II. This paper employs a three-dimensional coordinate system with both a local coordinate system for control and a global coordinate system for the uncertain ground, making the search for optimal solutions challenging. Additionally, the presence of a freeform mirror in the light path further complicates the problem.

Prior research has minimized the mismatches using linear methods like the Kalman filter [12]–[14]. In a vehicle windshield projection HUD, consisting of a windshield with a free-form surface and an off-axis concave mirror, maintaining the aspect ratio of the image or the central light position value is crucial when forming a light path within each effective area. If the optical path deviates due to the surrounding environment, it must pass through a correction area. However, such areas frequently exhibit abnormal image scattering, leading to issues in maintaining correction continuity.

Methods such as the Extended Kalman Filter (EKF) demonstrate limitations in exploring detour routes to avoid such problematic areas. Therefore, this paper employs reinforcement learning to determine the optimal policy by assessing intermediate processes and their future impacts when searching for the optimal location. There have been various studies applying reinforcement learning to automotive vehicles [15]–[18]. Specifically, we apply PPO (Proximal Policy Optimization) [19] and SAC (Soft Actor-Critic) [20], corresponding to On and Off Policy approaches, and compare them with existing methods. Table I summarizes the comparison of candidate methodologies in this study and Table II summarizes the comparisons of approaches between related works and ours.

III. System Modeling and Problem Statement

A. Virtual Image

To aid in understanding the dynamic modeling of this problem, we restricted the model's dimensionality to two dimensions, focusing solely on displacement in the vertical

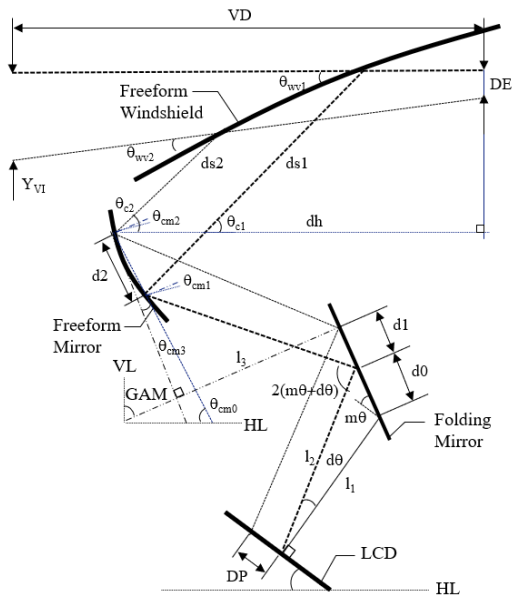

Fig. 2: Schematic of optical ray from pixel on LCD.

(up and down) directions. The schematic diagram is depicted in Figure 1.

When the vehicle experiences vertical displacement denoted as Y_C and undergoes a pitching angle change represented as ω_c, the vertical displacement of the virtual image DAR can be explained as follows:

$$DAR = \omega_c L_{VIHD} + \omega_c L_{HDC} + Y_C. \tag{1}$$

Note that any displacement resulting from external forces can be disregarded, as it is treated as part of a rigid body.

B. Control

The displacement of the virtual image, as determined by Eqn. (1), corresponds to the displacement of the pixels on the LCD, ensuring their alignment with the object assumed to be stationary in front. The typical optical system composing the HUD includes an LCD, a flat mirror, a concave mirror, and a windshield, with displacements d_1, d_2, and d_3 defined as follows:

$$d_1 = (d_0 \cdot DP)/(l_1 \cdot \tan d_\theta), \tag{2}$$

$$d_2 = \frac{(l_3 \tan m_\theta) - l_3(m_\theta + d_\theta) - d_1}{\sin\theta_{cm3} \cdot (\tan(90° + \theta_{cm3}) + \tan(m_\theta))}, \tag{3}$$

$$DAR = d_0 \cdot (\tan(2\theta_{wv2} - \theta_{c2}) - \tan(2\theta_{wv1} - \theta_{c1})) + DE, \tag{4}$$

where

$$\begin{aligned} DE = {} & ds_1 \cdot \sin(\theta_{c1}) - ds_2 \cdot \sin(\theta_{c2}) + \\ & \tan(2\theta_{wv1} - \theta_{c1}) \cdot (dh - (dc + ds_1 \cdot \cos(\theta_{c1})) \\ & - \tan(2\theta_{wc2} - \theta_{c2}) \cdot (dh - ds_2 \cdot \cos(\theta_{c2})), \end{aligned} \tag{5}$$

where

$$dc = d_2(\cos\theta_{cm0} + \sin\theta_{cm0}) \cdot \tan(GAM - 2\theta_{cm1} - m_\theta), \tag{6}$$

$$DAR = f(DP). \tag{7}$$

TABLE III: Differences between actual and estimated DAR values in DP

DP	Mode-ling	4th order Polynomial	Exponential (Half-Life)	Exponential (Proportio.)	Rectangu. hyperbola
-45	28.48	1.17	1.50	1.50	8.55
-30	24.37	5.05	0.10	0.10	0.78
-15	25.10	7.78	8.17	8.16	11.27
0	0.07	3.89	7.32	7.32	2.97
15	7.61	1.93	3.48	3.48	0.42
30	11.89	2.72	5.74	5.74	5.01
45	34.58	0.78	1.50	1.51	8.51

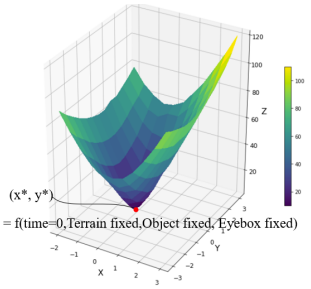

Fig. 3: Schematic of domain, range

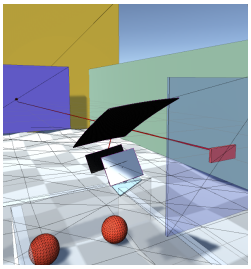

Fig. 4: Unity Simulator.

The ultimate goal of this paper is to determine the value of DP that equals DAR to minimize the residual displacement between display and object without vibration. Table III presents the errors between actual data and predicted values obtained through modeling and regression techniques. We use a fourth-degree polynomial for our simulation in some approaches, except for reinforcement learning, due to the minimum errors observed. This analysis illuminates the distinctive characteristics of the problem, emphasizing the difficulty of formalizing the optical path. This challenge is particularly pronounced due to the inclusion of a freeform mirror along the path, which, in combination with its high sensitivity, complicates the problem further. Figure 3 illustrates this sensitivity by depicting the relationship between the domain and range of the normalized problem.

IV. Solving Approaches

In this section, we provide detailed equations for the comparison group of solution methodologies to address the problem described earlier.

A. Extended Kalman Filter

Calculating the optimal solution, first consider the traditional approach represented by EKF. Let a state variable be

$$x^\mathsf{T} = [DP, DAR, Y_C, \omega_c]. \tag{8}$$

The system model with the state variable is formulated as:

$$\begin{aligned} x_1' &= x_2'/\text{Mag}, \\ x_2' &= x_4'(L_{VIHD} + L_{HDC}) + x_3', \\ x_3' &= x_2' - x_4'(L_{VIHD} + L_{HDC}), \\ x_4' &= (x_2' - x_3')/(L_{VIHD} + L_{HDC}), \end{aligned} \tag{9}$$

TABLE IV: The description of parameters.

Symbol	Description
ω_c	Pitch angle of vehicle
Y_C	Position of Vehicle
θ_m	Degree of Mirror
r	Reward of RL
VD	Virtual Image Distance
DE	Displacement from Eyebox center
DP	Displacement of pixel
DAR	Displacement from AR Center
L_{VIHD}	Distance from AR to wind shield
L_{HDC}	Distance from wind shield to eyebox
HL	Horizontal Line
VL	Vertical Line
GAM	Global angle of mirror
Mag	Magnification

where Mag abbreviates Magnification, and the measurement model is formulated as follows:

$$z_k = \omega_c L_{VIHD} + \omega_c L_{HDC} + Y_C + v_k. \tag{10}$$

Jacobian matrix is formulated as follows:

$$\begin{aligned} F_1 &= [0, (1/\text{Mag})dt, 0, 0], \\ F_2 &= [0, 0, dt, (L_{VIHD} + L_{HDC})dt], \\ F_3 &= [0, dt, 0, -(L_{VIHD} + L_{HDC})dt], \\ F_4 &= [0, -dt/(L_{VIHD} + L_{HDC}), dt/(L_{VIHD} + L_{HDC}), 0]. \end{aligned} \tag{11}$$

B. Reinforcement Learning

In this paper, PPO and SAC are employed as the RL algorithms, representing on-policy and off-policy RL, respectively. In general, the state is defined as:

$$s_t^{\mathsf{T}} = [\omega_c, Y_C, \text{DAR}], \tag{12}$$

The action space is defined as:

$$a_t = DP \text{ or } \theta_m. \tag{13}$$

In situations of low complexity, only action DP is taken, while as complexity increases, an additional action, θ_m, is introduced. Please refer to Table IV for a detailed description of the parameters.

Reward Function: The reward function is determined by the negative L_1 norm or L_2 norm between the static object and the imaging point along the Center Ray, i.e., Displacement from AR Center, which is determined by the focal length. The reward function can be represented as follows:

$$r(s_t, a_t) = -||DAR||_1. \tag{14}$$

C. PPO Algorithm

The augmented reality optical system used in automobiles features a magnification that represents the ratio of the size of the virtual image generated by the PGU to its original size. Consequently, the targeted image's apparent movement is amplified by the magnification, compared to the displacement of the control variable. This results in the agent's displacement area, aimed at minimizing the objective function, occupying a relatively small portion of the overall search area.

TABLE V: The classification of simulation by degree of freedom

	Vehicle	Object	Eyebox	Control (local)	Practical Relevance
Class I, Class Ex	Dy Rx,Ry,Rz	All fixed	All fixed	(x)	Pertinent
Class II	Dy Rx,Ry,Rz	Dx,Dy Rx	All fixed	(x, y)	Moderate
Class III	Dy Rx,Ry,Rz	Dx, Dy Rx	Dy	(x, y)	Moderate

In the process of taking action and continuously updating within the broad search area, the most reliable area is selected. For policy determination, we have opted for the PPO algorithm [19]. The objective function is defined as:

$$L_t^{CLIP+VF+S}(\theta) = E_t[L_t^{CLIP}(\theta) - c_1 L_t^{VF} + c_2 S[\pi](s_t)], \tag{15}$$

where VF represents the error loss of the value function for the left-hand variable, and S denotes the entropy weight. This weight is determined as the difference between the sum of the terms c_1 and c_2, which control the extent of the proxy loss in the Clip term, limiting its correlation with the previous policy for the right-hand variable.

D. SAC Algorithm

Unlike the on-policy PPO algorithm, the off-policy SAC (Soft Actor-Critic) [20] algorithm introduces an entropy measurement term, enabling exploration of various potential areas while avoiding low-probability paths. This addition is represented in the objective function $J(\theta)$ as:

$$J(\theta) = \sum_{t=1}^{T} E_{(s_t,a_t)\sim p_{\pi_\theta}}[r(s_t, a_t) + \alpha H(\pi_\theta(\cdot|s_t))], \tag{16}$$

where entropy is quantified by the term H on the right-hand side, and the importance of entropy in exploration is controlled by the weight parameter α. Notably, when selecting the optimal policy, the SAC algorithm takes past experiences into account, resulting in high sample efficiency and a reduced risk of falling into a local minimum.

E. Evaluation Criteria

The problem tackled in this paper pertains to the challenge of maintaining overlap between the displayed information, which conveys information to the user, and the object. To ensure a comprehensive approach that accounts for the driver's perception, we incorporate the Dynamic Visual Acuity, abbreviated as 'DVA' variable [21], in addition to displacement considerations aimed at aligning the object expressed in DAR with the display.

V. Simulation

A. Simulation Environment

The experiment was conducted using Unity Simulator [22] on Intel(R) Core(TM) i7-10750H CPU@ 2.60GHz, GTX 1660 TI hardware to determine the degree of achievement of the objective function after learning.

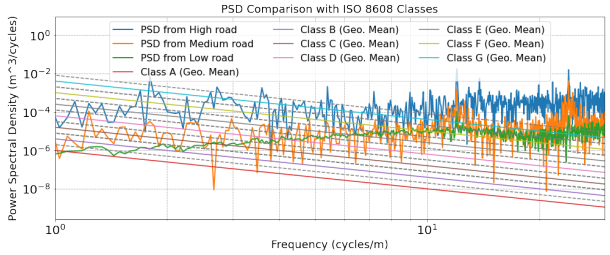

Fig. 5: Power Spectral Density of each terrain.

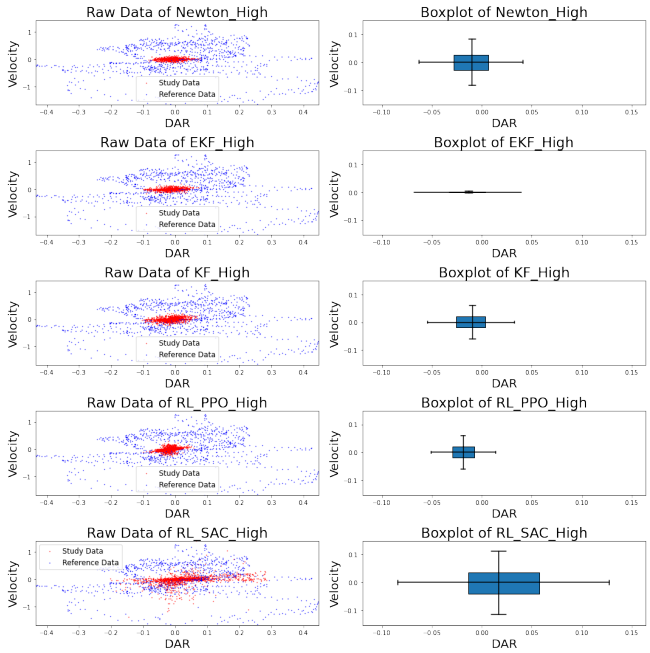

Fig. 6: Plot of Class I.

B. Simulation Classification

The main conditions affecting the results of the simulation are summarized in Table V. Terrain, the uncertainty factor mentioned in Table II, was classified based on the acceleration occurring in the vertical direction to the ground. In the road pavement condition classified by ISO8608 [23], three ground conditions (High, Medium, Low) were set to exceed class B in sections exceeding 10 cycle/m, and the power spectral density(PSD) is shown in Figure 5.

C. Simulation Results

1) Class I: We initiated the simulation by considering the degree of freedom constraints based on complexity. The measurement results, using the case where no adjustments were made to the display projected onto the three surfaces as a reference. Specifically, in Class I, the outcomes for methods under High road conditions are presented in Figure 6 and summarized in Table VI.

TABLE VI: Results of Class I simulation as high road.

	AVG DAR		AVG Velocity		SD DAR		SD Velocity	
	m	%	m/s	%	m	%	m/s	%
Nwtn	0.023	17.1	0.034	7.5	0.028	16.5	0.044	7.8
KF	0.022	15.8	0.037	8.3	0.027	15.8	0.057	10.2
EKF	0.025	18.3	0.018	4.0	0.021	12.3	0.035	6.2
PPO	0.022	15.8	0.04	8.8	0.019	10.9	0.062	11.0
SAC	0.055	40.5	0.091	20.3	0.075	44.2	0.173	30.8
Ref.	0.136	100.	0.449	100.	0.170	100.	0.561	100.

TABLE VII: Results of Class II, III SIMULATION.

	Class II				Class III			
	AVG		SD		AVG		SD	
	DAR e-3m	Vel m/s	DAR e-3m	Vel m/s	DAR e-3m	Vel m/s	DAR e-3m	Vel m/s
L	26	0.240	38	0.103	72	0.184	70	0.186
M	36	0.260	42	0.086	95	0.248	96	0.285
H	53	0.267	61	0.143	117	0.286	135	0.280

TABLE VIII: Results of Class Ex SIMULATION.

	Newton-RL				EKF-RL			
	AVG		SD		AVG		SD	
	DAR e-3m	Vel m/s	DAR e-3m	Vel m/s	DAR e-3m	Vel m/s	DAR e-3m	Vel m/s
L	6	6e-4	8	0.002	31	0.059	32	0.942
M	8	0.005	11	0.008	32	0.665	51	1.257
H	9	0.008	11	0.011	55	0.773	730	1.433

The simulation results generally showed that in terms of DVA, it does not exceed the target level of 2 m/s, and in terms of DAR, the PPO algorithm of reinforcement learning showed the best results, with the average DAR compared to the reference being 15.8% and the standard deviation of DAR being 10.9%, as shown in Table VI The traditional methods, KF and EKF, collected data that showed convergence, but as time went by and the Gaussian distribution condition deteriorated, the initial stability decreased and the results diverged, showing that they were not stable.

The Newton method showed acceptable performance in both DVR and DAR but showed limitations in application due to increased complexity due to increased degrees of freedom and randomness beyond CLASS II. Performance at Low, where ground conditions are most favorable, has consistently improved. In terms of DAR, 0.05 m was set as the target displacement based on the size of a car license plate, and in terms of speed, 1m/s was set as the target speed considering the deterioration of DVA function in driving situations [24].

2) Class II: In Class II, we introduced additional degrees of freedom for the object located in front of the vehicle, as well as adjustment displacement of the pixels on the LCD denoted as (x, y). Consequently, achieving precise system modeling, which was essential for Newton and EKF, became unfeasible. As a result, we conducted experiments exclusively with the PPO reinforcement learning algorithm. The results for Class II are presented in Figure. 7 and summarized in Table VII - Class II. While the complexity increased compared to Class I, performance remained acceptable, particularly under medium and low road flatness conditions.

3) Class III: In Class III, we introduced the displacement of the eyebox in addition to the factors considered in Class II, making it the most complex case in this study. As complexity

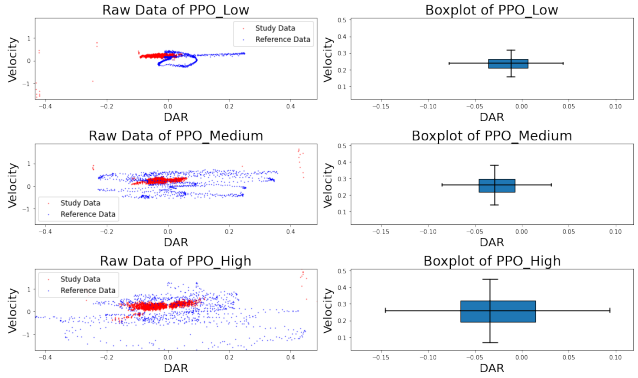

Fig. 7: Plot of Class II.

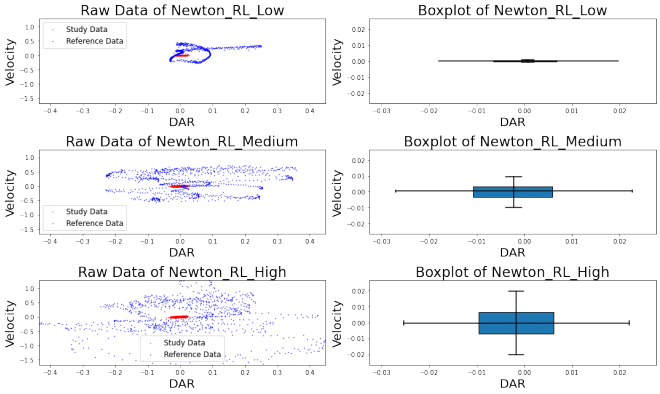

Fig. 8: Plot of Class Ex.

increased, we observed occasional conflicts between actions among agents, resulting in less stable learning compared to previous scenarios. Nevertheless, both eyebox and object tracking performed well, even under medium road conditions. While performance deteriorated under high complexity conditions, there was potential for further optimization. The test results are presented in Table VII-Class III.

4) Class Ex: Based on the lessons learned from previous classes, a combination experiment was conducted taking advantage of the strengths of reinforcement learning and traditional methods. The results are shown in Fig. 8 and TableVIII. The main idea of combining the two methodologies is to distribute the main update parameters separately between the two methodologies, and the combination of Newton and reinforcement learning shows successful results.

VI. Experimental Validation

To verify the accuracy of the raytracing performed in the simulation, we used a validated optical inspection system, The resolution of the system is 0.1 degrees, which is a level that does not affect the measurement results. Measurements of the range according to the domain, which is the action range of the agent, were compared. The left figure of Fig. 9 depicts simulation and measurement values, and the right one of Fig. 9 shows the difference between simulation and measurement values. As a result of the comparison, it was confirmed that the simulation measurements tended to be convex compared to the actual measurements. However, the converted absolute amount was confirmed to be less than 0.8

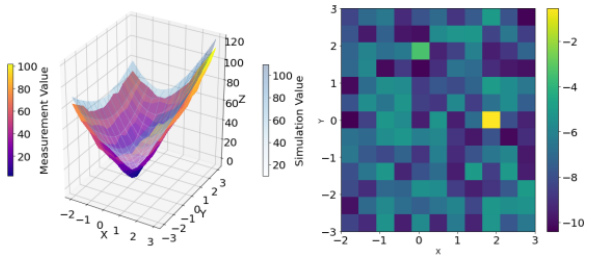

Fig. 9: The left figure depicts both measurement and simulation values at the same time. The right figure shows the difference between simulation and measurement values.

mm, which is negligible even considering the magnification which is mainly applied to our system.

VII. Conclusion

In the realm of automotive augmented reality optical systems, we integrated reinforcement learning as a means to maintain the alignment between virtual images and real-world objects, the fundamental objective of augmented reality. Throughout our investigation, we established the intricate relationship between crucial parameter values and the objective function, thereby streamlining the time required to approach the optimal objective function value under various conditions. This paper delved into techniques for achieving stable learning and offered a comprehensive performance comparison between objective functions derived from traditional methodologies and those rooted in reinforcement learning.

By systematically categorizing classes based on the escalating complexity arising from the introduction of random elements, we were able to pinpoint the respective limitations and advantages of each approach. This comparison led us to apply the most suitable methodologies in accordance with the defined classes. Building on this foundation, we subjected the combination of reinforcement learning and existing methods to rigorous testing, ultimately discerning specific conditions under which this amalgamation outperformed traditional methods. The insights gained through this exploration not only advance our understanding of the interplay between reinforcement learning and existing approaches but also pave the way for the practical application of augmented reality in automotive settings.

Acknowledgement

This work is supported by the National Research Foundation of Korea (NRF) through the Ministry of Science and ICT under Grant 2021R1A2C1093957, by Korean Ministry of Land, Infrastructure and Transport(MOLIT) as Innovative Talent Education Program for Smart City, and by Korea Institute for Advancement of Technology(KIAT) grant funded by the Korea Government(MOTIE) (P0020536, HRD Program for Industrial Innovation). The Institute of Engineering Research at Seoul National University provided research facilities for this work.

REFERENCES

[1] J. L. Gabbard, G. M. Fitch, and H. Kim, "Behind the glass: Driver challenges and opportunities for ar automotive applications," *Proceedings of the IEEE*, vol. 102, no. 2, pp. 124–136, 2014.

[2] T. J. Ypma, "Historical development of the newton–raphson method," *SIAM review*, vol. 37, no. 4, pp. 531–551, 1995.

[3] F. Daum, "Nonlinear filters: beyond the kalman filter," *IEEE Aerospace and Electronic Systems Magazine*, vol. 20, no. 8, pp. 57–69, 2005.

[4] C. W. Anderson, "Learning to control an inverted pendulum using neural networks," *IEEE Control Systems Magazine*, vol. 9, no. 3, pp. 31–37, 1989.

[5] B. Xu, Y. Lyu, and S. A. Gadsden, "Estimation and control of a double-inverted pendulum," 2018.

[6] S. H. Zabihifar, A. S. Yushchenko, and H. Navvabi, "Robust control based on adaptive neural network for rotary inverted pendulum with oscillation compensation," *Neural Computing and Applications*, vol. 32, pp. 14 667–14 679, 2020.

[7] O. Mofid, K. A. Alattas, S. Mobayen, M. T. Vu, and Y. Bouteraa, "Adaptive finite-time command-filtered backstepping sliding mode control for stabilization of a disturbed rotary-inverted-pendulum with experimental validation," *Journal of Vibration and Control*, vol. 29, no. 5-6, pp. 1431–1446, 2023.

[8] W. Chu, C. Li, and Y. Zhou, "A rapid stabilization method of the flexible inverted pendulum based on constrained boundary circumferential motion," *Mechanical Systems and Signal Processing*, vol. 187, p. 109895, 2023.

[9] I. Sünger and S. Çankaya, "Augmented reality: historical development and area of usage," *Journal of Educational Technology and Online Learning*, vol. 2, no. 3, pp. 118–133, 2019.

[10] V. Charissis, J. Falah, R. Lagoo, S. F. Alfalah, S. Khan, S. Wang, S. Altarteer, K. B. Larbi, and D. Drikakis, "Employing emerging technologies to develop and evaluate in-vehicle intelligent systems for driver support: infotainment ar hud case study," *Applied Sciences*, vol. 11, no. 4, p. 1397, 2021.

[11] A. Feierle, D. Beller, and K. Bengler, "Head-up displays in urban partially automated driving: Effects of using augmented reality," in *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*. IEEE, 2019, pp. 1877–1882.

[12] T.-H. S. Li and C.-C. Chen, "Extended kalman filter based hand-shake detector for optical image stabilization using a low cost gyroscope," *IEEE Transactions on Consumer Electronics*, vol. 59, no. 1, pp. 113–121, 2013.

[13] C. Wang, J.-H. Kim, K.-Y. Byun, J. Ni, and S.-J. Ko, "Robust digital image stabilization using the kalman filter," *IEEE Transactions on Consumer Electronics*, vol. 55, no. 1, pp. 6–14, 2009.

[14] S. Ertürk, "Real-time digital image stabilization using kalman filters," *Real-Time Imaging*, vol. 8, no. 4, pp. 317–328, 2002.

[15] S.-W. Yoo, C. Kim, J. Choi, S.-W. Kim, and S.-W. Seo, "Gin: Graph-based interaction-aware constraint policy optimization for autonomous driving," *IEEE Robotics and Automation Letters*, vol. 8, no. 2, pp. 464–471, 2022.

[16] C. Kim, J.-K. Cho, H.-S. Yoon, S.-W. Seo, and S.-W. Kim, "Unicon: Uncertainty-conditioned policy for robust behavior in unfamiliar scenarios," *IEEE Robotics and Automation Letters*, vol. 7, no. 4, pp. 9099–9106, 2022.

[17] C. Kim, J. Cho, C. Bobda, S.-W. Seo, and S.-W. Kim, "Sero: self-supervised reinforcement learning for recovery from out-of-distribution situations," in *International Joint Conference on Artificial Intelligence*, 2023.

[18] S. R. Kim, C. Kim, S. Shin, and S.-W. Kim, "Deep reinforcement learning for semi-active suspension: A feasibility study," in *2023 International Conference on Electronics, Information, and Communication (ICEIC)*, 2023, pp. 1–5.

[19] I. G. Petrazzini and E. A. Antonelo, "Proximal policy optimization with continuous bounded action space via the beta distribution," in *2021 IEEE Symposium Series on Computational Intelligence (SSCI)*. IEEE, 2021, pp. 1–8.

[20] T. Haarnoja, A. Zhou, P. Abbeel, and S. Levine, "Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor," in *International conference on machine learning*. PMLR, 2018, pp. 1861–1870.

[21] M.-S. Jin and I.-C. Jeon, "A study on the measurement of dynamic visual acuity according to the change of accommodative stimulus," *The Korean Journal of Vision Science*, vol. 20, no. 4, pp. 523–530, 2018.

[22] A. Juliani, V.-P. Berges, E. Teng, A. Cohen, J. Harper, C. Elion, C. Goy, Y. Gao, H. Henry, M. Mattar *et al.*, "Unity: A general platform for intelligent agents," *arXiv preprint arXiv:1809.02627*, 2018.

[23] P. R. Pawar, A. T. Mathew, and M. Saraf, "Iri (international roughness index): an indicator of vehicle response," *Materials Today: Proceedings*, vol. 5, no. 5, pp. 11 738–11 750, 2018.

[24] J. L. Demer and F. Amjadi, "Dynamic visual acuity of normal subjects during vertical optotype and head motion." *Investigative ophthalmology & visual science*, vol. 34, no. 6, pp. 1894–1906, 1993.