Proceedings of the 2024 Winter Simulation Conference
H. Lam, E. Azar, D. Batur, S. Gao, W. Xie, S. R. Hunter, and M. D. Rossetti, eds.

AI-DRIVEN PHYSICS-BASED SIMULATION DESIGN WITH OPTICAL FLOW BASED MARKOV DECISION PROCESSES FOR SMART SURVEILLANCE

Seunghan Lee[1], Yinwei Zhang[2], Aaron LeGrand[1], and Seth Gibson-Todd[1]

[1]Dept. of Industrial and Systems Engineering, Mississippi State University, Mississippi State, MS, USA
[2]Dept. of Systems and Industrial Engineering, University of Arizona, Tucson, AZ, USA

ABSTRACT

This paper presents an integrated approach to enhancing situational awareness and decision-making in dynamic environments by combining optical-flow-based Markov Decision Processes (MDP) with physics-based simulations for proactive surveillance system design. By utilizing optical flow for real-time motion detection and analysis, our framework achieves immediate comprehension of environmental changes, which is essential for autonomous navigation and surveillance applications. Additionally, the framework employs MDP to model and resolve decision-making problems where outcomes are partly random and partly controlled by the decision-maker, optimizing actions based on predicted future states. The model was validated through Hardware-in-the-loop (HITL) simulations, providing a detailed understanding of the physical phenomena influencing the system. This ensures that decisions are data-driven and customized to specific situations and missions. Our approach offers a comprehensive understanding of system integration with AI, along with real-time analysis and decision-making capabilities, thereby advancing the simulation modeling methodology for engaging with complex, dynamic environments.

1 INTRODUCTION

Developing and applying autonomous control technologies for unmanned vehicles (UVs) has become crucial in the ever-changing landscape of defense strategy and operation (Maxwell et al. 2013). The uncertainty forces federal and state agencies to explore innovative methods to leverage the various capabilities of artificial intelligence (AI) to transform UV autonomous operations (Di and Shi 2021). Thus, the integration of AI expertise into these operations is also crucial to the operational success. Given the complexity, unpredictability, and diverse nature of today's battlefields, there is a pressing need for autonomous systems capable of working together effectively and adjusting quickly to changes. It also requires the achievement of high levels of situational awareness in diverse environments. Autonomous control technologies (Zhe et al. 2020), including unmanned aerial and terrestrial vehicle networks, present a resilient solution to these demands (Yun et al. 2022). However, unlocking their full potential depends on advancements in AI algorithms, coordination techniques, and understanding the operational context, calling for a holistic approach.

Integrating AI with UV autonomous control strategies is a key factor behind homeland security's simulation research. In today's era of complex defense engagements, AI-enhanced UVs bring unprecedented efficiency and autonomy, minimizing human risk and significantly outmatching opponents in the mission environment (Jeelani and Gheisari 2021; Primatesta et al. 2020). This fusion signifies a fundamental shift in the defense strategy toward conducting similar operations such as reconnaissance, surveillance, and mission completion. Several considerations, including - 1) simulating dynamic and unpredictable environments, 2) managing diverse UV autonomy, 3) integrating various sensor technologies, and 4) countering threats and operational limitations, represent challenges to implementing UV autonomous control strategies (Abdelmaksoud et al. 2020). Addressing these challenges necessitates focused research efforts in the integration of UV autonomy and AIs. As illustrated in Figure 1, the Observe, Orient, Decide,

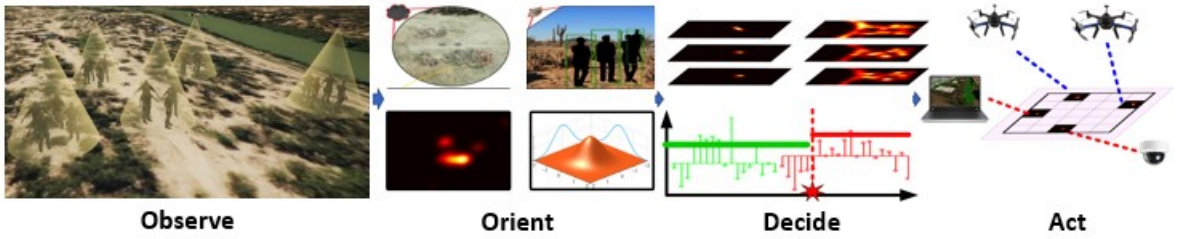


Figure 1: OODA loop supported by UV autonomy.

and Act (OODA) loop represents an effective framework for securing decision-making superiority and adjusting swiftly to enemy operations (Richards 2004). Within the defense strategy, the OODA loop highlights the essential steps of quickly observing current situations, understanding their implications through orientation, making informed decisions, and implementing them quickly and accurately (Miller et al. 2021). The effectiveness of defense strategies often depends on outpacing opponents in utilizing the OODA loop, disrupting their ability to make decisions, and securing a vital strategic advantage. Enhancing the efficiency of the OODA loop and leveraging UV autonomous technology is key for modern defense tactics (MeiJie et al. 2021). This method improves the speed and adaptability of decision-making and ensures a competitive edge in combat scenarios. It also reduces human risk by utilizing sophisticated UV technology, offering a cost-effective and flexible solution. The tactical benefits derived from optimizing OODA loop and integrating UV autonomous control strategies play a significant role in national defense, protecting a country's interests and preserving its superiority amid the changing dynamics of international security (Miller et al. 2021; MeiJie et al. 2021).

This paper introduces a comprehensive physics-based simulation design framework that utilizes optical flow analysis with Markov Decision Processes (MDP) to enhance the capabilities of smart surveillance systems. Leveraging optical flow results to define system states and employing MDP to encompass a wide array of enemy behaviors, our strategy aims to develop proactive control strategies that anticipate and mitigate potential security challenges ahead of their manifestation. This approach improves situational awareness and enables the development of dynamic, preemptive responses to imminent threats, elevating urban safety and security standards. Furthermore, the paper will investigate the critical role of sensor technologies and drones, utilizing physics-based simulation to assess their effectiveness in practical surveillance scenarios. Through this comprehensive methodology, our study provides a robust, adaptable framework that employs state-of-the-art technology to address the multifaceted security demands of modern urban settings.

2 LITERATURE REVIEW

The intersection of physics-based simulation (PBS), smart surveillance, and AI-based technology has attracted significant attention in recent years (Wang et al. 2022; Young et al. 2022; Lee et al. 2019), particularly with the integration of MDP and reinforcement learning (RL) methodologies to improve intelligent decision-making (Chen et al. 2024; Terashima et al. 2024). Additionally, the utilization of drones for large-scale environment surveillance has emerged as a promising avenue for enhancing the fidelity of visual information, as shown in (Liu et al. 2021) where the authors employed a visual-geometric fusion learning technique to simultaneously map and navigate an urban environment. MDP methods offer numerous applications to improve the abilities and performance of drones, such as reducing interference and increasing energy efficiency in large unmanned aerial vehicle (UAV) swarms (Li et al. 2020), building trust among multi-vehicle systems (Delamer and Givigi 2020), and addressing issues related to learned policy carryover from simulation training to the execution phase (Wang et al. 2022). While thorough consideration for drone behavior and performance should be managed, recent literature has narrowed this focus to enhance the collection and use of optical flow data for UAV navigation (Chen et al. 2020) and obstacle detection (Li et al. 2019) via error analysis modeling and improved sparse optical flow, respectively. One particular

and noteworthy advancement in the field of object detection and recognition is the YOLOv7 model (Wang et al. 2023) which has seen successful application to vehicle detection in UAV image and video processing (Panthakkan et al. 2023), but use of the YOLOv7 model extends to areas such as defect detection in steel production (Zhou and Piao 2023). While a large portion of research using YOLO technology with UAVs is aimed at vehicle detection, (Ajmera and Singh 2020) combined YOLO with optical flow and applied RL to identify and track a target in urban-environment search and rescue missions.

Not only can hardware be used for conducting simulations, it can also be applied to the validation and verification stages of simulation and experimental design. Authors in (Lugo-Cardenas et al. 2016) used Hardware-in-the-loop (HITL) to validate different UV controllers with a range of UV platforms using their MAV3DSim simulator. HITL simulation is an effective strategy in evaluating the performance of control systems, such as flight control systems (Huang et al. 2020). Although hardware can play an important role in simulation development and testing, software or firmware can be equally vital in assessing, improving, and validating a simulation experiment. Similarly, software has also been implemented into simulation loops, defined by SITL, or software-in-the-loop. In (Silano et al. 2019), SITL simulation and methodology was applied to a quad rotor flight control system to uncover undesirable system behaviors that were previously hidden in the MATLAB/Simulink platforms. Another example of performance evaluation on the navigation and control system of a UAV through SITL simulation can be found in (Aminzadeh et al. 2018).

MDPs can be useful in a variety of applications such as analyzing energy efficiency in wireless sensor networks (Kalnoor and Subrahmanyam 2020), optimizing long-term operations of hydroelectric systems (Lamond and Boukhtouta 1996), solving treatment decision problems relating to depression (Li et al. 2022), and many more. However, the application at the forefront of this paper is in surveillance. UVs are often used to monitor areas due to their ease of operation, low cost, and ability to act freely. An efficient control logic must be applied to take full advantage of these abilities. MDPs have displayed enormous success in modeling and analyzing efficient path planning algorithms (dos Santos and Vivaldini 2022). More specifically, (RL) is a useful tool for UAV surveillance with single and multiple drone scenarios (Frattolillo et al. 2023).

3 MODELING METHODOLOGIES

3.1 Optical Flow-based Situational Awareness

Optical flow is defined as the perceived movement patterns of pixels across successive images or video frames. It operates through algorithms that calculate the motion vectors for either individual pixels or specific areas of an image across different frames, thereby identifying the motion trends within the scene. Horn and Schunck (Horn and Schunck 1981) proposed a method to accurately estimate the optical flow at each pixel by relying on two key assumptions: (i) the brightness of pixels within a small region remains fairly constant over time, and (ii) motion in the scene is uniformly smooth across the frame. The objective function can be defined below by considering these two assumptions:

$$\arg\min_{\mu_{i,j},\omega_{i,j}} \sum_{i,j} \left[\left(\frac{\partial f_{i,j}}{\partial x}\mu_{i,j} + \frac{\partial f_{i,j}}{\partial y}\omega_{i,j} + \frac{\partial f_{i,j}}{\partial t} \right)^2 + \eta^2 \left(\left(\frac{\partial \mu_{i,j}}{\partial x} \right)^2 + \left(\frac{\partial \omega_{i,j}}{\partial y} \right)^2 \right) \right]. \tag{1}$$

In this context, f represents the raw video frame, an element of the space of real-valued matrices of dimensions $I \times J$, where $f_{i,j}$ corresponds to the intensity of the pixel at position (i, j). $(\mu_{i,j}, \omega_{i,j})$ defines the horizontal and vertical components of the optical flow at the given pixel position. The terms $\frac{\partial f_{i,j}}{\partial x} = f_{i,j} - f_{i-1,j}$, $\frac{\partial f_{i,j}}{\partial y} = f_{i,j} - f_{i,j-1}$, and $\frac{\partial f_{i,j}}{\partial t} = f_{i,j} - f_{i,j}^{n-1}$ refer to the derivatives of the pixel intensity (i.e., finite difference between consecutive pixels) in the horizontal, vertical, and temporal dimensions, respectively. The coefficient η serves to adjust the smoothness level of the resulting optical flow.

Although the optical flow vectors $(\mu_{i,j}, \omega_{i,j})$ for each pixel can be estimated by solving equation (1), it may not detect the moving foreground effectively as the background and foreground are not explicitly

modeled. To address this limitation, (Zhang et al. 2023) proposed a method to explicitly decompose the optical flow magnitude into the foreground optical flow and the background optical flow. Specifically, the magnitude of the optical flow at each pixel can be computed as $\mathbf{Y}_{i,j} = \sqrt{\mu_{i,j}^2 + \omega_{i,j}^2}$. As a result, the value of $\mathbf{Y}_{i,j}$ increases with the speed of pixel movement at (i, j), i.e., a larger optical flow vector $(u_{i,j}, v_{i,j})$ corresponds to a higher pixel velocity and vice versa. Furthermore, such magnitude of the optical flow can be formulated by three components, as $\mathbf{Y} = \mathbf{B} + \mathbf{M} + \mathbf{E}$, where $\mathbf{B} \in \mathbb{R}^{I \times J}$ denotes the smoothness background due to the stationary motion of the UV, $\mathbf{M} \in \mathbb{R}^{I \times J}$ represents the optical flow pertinent to the mobile foreground subjects, and $\mathbf{E} \sim \mathcal{N}(0, \sigma^2)$ is the Gaussian distributed noises with mean zero and standard deviation σ. Given the typical scenario of a UV surveillance flight, the matrix $\mathbf{M}$ is characterized by sparsity, indicating that most of its elements are zero-valued (aligning with stationary background), with the exception of those few elements corresponding to the foreground movement, which are non-zero and tend to be spatially localized.

The following objective function is proposed by (Zhang et al. 2023) for the accurate estimation of the components, i.e., $\mathbf{B}$ and $\mathbf{M}$, as

$$\arg\min_{\mathbf{B},\mathbf{M}} \|\mathbf{Y} - \mathbf{B} - \mathbf{M}\|_F^2 + \beta_1 \|\mathbf{DB}\|_F + \beta_2 \|\text{vec}(\mathbf{M})\|_1 + \beta_3 \|\text{vec}(\mathbf{DM})\|_1, \tag{2}$$

where $\|\cdot\|_F$ denotes the Frobenius norm, $\|\cdot\|_1$ is the L_1 norm, and $\text{vec}(\cdot)$ is the operation that transforms a matrix into a vector. The $\beta_1, \beta_2, \beta_3$ are the tuning parameters. The matrix $\mathbf{D}$, representing first-order differences, is defined as:

$$\mathbf{D} = \begin{bmatrix} 1 & -1 & \cdots & \\ \ddots & \ddots & \ddots & \ddots \\ & & 1 & -1 \end{bmatrix}. \tag{3}$$

The first term, $\|\mathbf{Y} - \mathbf{B} - \mathbf{M}\|_F^2$, is for the ordinary least squares fitting that minimizes the residual sum of squares. The second term, $\|\mathbf{DB}\|_F$ ensures the smoothness of the background optical flow, accounting for the fact that background elements usually present consistent movement patterns in UAV footage. An L_1 norm was chosen to quantify the first term $\|\text{vec}(\mathbf{M})\|_1$, which guarantees that only the pixel with large optical flow magnitude (corresponds to significant movement) will be selected. Finally, $\|\text{vec}(\mathbf{DM})\|_1$ facilitates the model's tendency to cluster contiguous non-zero values in the matrix $\mathbf{M}$, thus reflecting the anticipated spatial configuration of moving elements. The objective function can be solved by the algorithm based on the ADMM efficiently (Zhang et al. 2023; Zhang et al. 2024).

3.2 Optical Flow-based Markov Decision Process

Integrating optical flow analysis with Markov Decision Processes (MDP) presents a sophisticated approach to enhancing surveillance systems. Optical flow, denoted by $\lambda(t)$ at time t, provides vector fields indicating motion between consecutive frames, capturing essential dynamics such as velocity (v) and direction of movement within the monitored area. These motion characteristics can be quantitatively defined to represent the states S in an MDP framework. The aim is for a UV's path to be adapted to amplify information based on the optical flow information, which, in turn, contributes to the operation's effectiveness. The state space $\mathbf{S}$ captures a comprehensive multi-dimensional model of each UV, including its position (x, y, z), velocity (v_x, v_y, v_z), and other pertinent state variables $(\boldsymbol{\theta})$ which are parameters derived from $\lambda(t)$, enabling a detailed representation of the environment's dynamic conditions. The action space $\boldsymbol{A}^t = \{\boldsymbol{a}_1^t, \boldsymbol{a}_2^t, \cdots, \boldsymbol{a}_N^t\}$ provides all necessary control information for k-th UV where $k = 1, 2, \cdots, N$, at every time step t. It presents a variety of actions, e.g., velocity and accelerations, for the collective UV swarm that includes camera adjustments or alert dispatch based on state assessments. R_t is the reward at time (t), incentivizing our surveillance aircraft to get closer to the intruder as the steps progress. For example, the ownship can receive a reward of +1 for finding the intruder, -1 if the intruder finds the stationary target before it is caught, and

$\frac{1}{d}$ for all other cases, where d is the relative distance between the ownship and the intruder. Moreover, one can add all the non-negative optical flow amounts (as a realization of Entropy I_t) detected in each grid to incorporate optical flow into the reward structure, which will be explained and further specified in Section 4. This allows us to compare the performance with and without incorporating optical flow information. The optimal policy π_j^* can be estimated by maximizing the following objective function as below:

$$\pi_j^* = argmax_\pi \mathbb{E}\left[\sum_{t=1}^{T}\sum_{k=1}^{N} R_t(\theta, \boldsymbol{a}_k^t) + \delta I_t(\theta, \lambda(t))\right], \tag{4}$$

where δ is a scaling factor that combines the information gained I_t (using Entropy) from the optical flow into the decision-making process. In specific, δ adjusts the influence of the optical flow information on the reward structure, enhancing the reward for detecting movement patterns and potential intruder locations. Consequently, a small δ may cause the UV to prioritize actions that maximize the immediate reward, such as directly pursuing the intruder based on its current location. In contrast, a large δ encourages the UV to rely more on the information obtained from the optical flow, leading to a more strategic approach that considers the broader movement patterns and predicts the future positions of the intruder. We have implemented a Q-learning algorithm and the following optimization model is considered to solve the MDP.

$$Q(\boldsymbol{\theta}, \boldsymbol{A}^t) \Leftarrow (1-\alpha)Q(\boldsymbol{\theta}, \boldsymbol{A}^t) + \alpha[R_t + \gamma \max Q(\boldsymbol{\theta}, \boldsymbol{A}^{t+1})], \tag{5}$$

where α and γ are the learning rate and discount factor, respectively. By using the proposed MDP method combined with Q-learning techniques, Equation 5 can be applied to the independent learning of each UV agent in a swarm via the unique Q-table assigned to each agent. In this work, Q-learning techniques are applied only to a single own ship and a single intruder to improve operational performance in our testing environment. Through this process, each agent learns effective tactics for decision-making under dynamic conditions. By updating the Q-values with the immediate reward and the discounted future rewards, the model becomes adept at discerning long-term action outcomes. The framework enables us to contemplate various mission types, including Dive Bomber, Pursuer-Evader, Simple Herder, and Greedy Cluster, each presenting distinct challenges from the perspective of adversarial agents.

The main objective is to optimize decision-making through maximal information gain using optical flow, aiming to balance operational costs in a mission-based context. Our experiments compared two strategies: 1) General reward-based policy and 2) Optical flow-based policy. Indeed, the UV's decisions are either guided by immediate rewards based on its current state and position relative to the intruder or allowed by intruders' dynamic movement patterns. Moreover, by incorporating hyperparameter analysis, we demonstrated how adjusting parameters like gamma, epsilon, alpha, and decay can enhance UV performance (Section 4). The optical flow-based policy led to more efficient UV routing and decision-making, highlighting the benefits of integrating additional environmental information into the UV's strategy. This difference can be seen through the reduction of the number of steps required for the UV to find the intruder. Further refining surveillance operations, the determination of optimal UV hovering altitudes post-waypoint sequence generation becomes crucial. A higher altitude offers a broader area coverage at the expense of detail due to reduced resolution, while a lower altitude enhances detail but limits the scope of the surveillance area, constrained by the onboard sensors' field of view (FOV). Both operational limitations and the strategic objectives of the surveillance mission influence this decision. Incorporating physics-based simulations can aid in assessing these trade-offs by providing a virtual environment to experiment with different altitudes and their impact on data quality and area coverage. By simulating various operational scenarios, the CSU can make informed decisions that balance the need for detailed data against the broader surveillance objectives, optimizing the surveillance system's effectiveness within the operational constraints of UV autonomy.

4 VERIFICATION & VALIDATION

In order to validate the framework, the author's implementation efforts are threefold: 1) Design and development of Human & Hardware-in-the-loop simulator 2) Implementation of our extended optical flow for Detection, 3) Incorporation of MDP. Our focus on the hardware design is to incorporate physical controls that aid in simulations, such as the Honeycomb Aeronautics yoke and throttle set, as seen in Figure 2. The authors have been working closely with experts in the aviation field to gather feedback to enhance the validity of the proposed algorithms and framework. Consequently, Section 4.1 focuses on the design and implementation of the simulator, offering detailed information about the hardware used and its integration into the simulations for this paper.

4.1 Human and Hardware in the Loop Simulator (HHITLS)

Simulation development was primarily carried out in Unreal Engine (UE), which can simulate real-world physics interactions among assets and agents. This game engine provides a significant amount of functionality, customization, and a plethora of pre-built Application Programming Interface (API) plugins. As a physics-based simulation (PBS), UE is capable of simulating accurate flight dynamics through the JSBSim plugin. JSBSim is an object-oriented (C++), non-linear 6-DOF, Flight Dynamics Model (FDM) with a fully configurable flight control system. Cesium ion serves as an online platform tailored for the management and visualization of 3D geospatial data, which can be seen as it was applied to the UE model in Figure 2(a). Continuously integrated as a pre-built plugin for UE, this software can stream global 3D content from cloud sources. Cesium's data originates from Bing Aerial maps as a full-scale WGS84 globe and users have the option to effortlessly upload their own data, enabling the creation of custom 3D environmental tiles. This dataset encompasses a spectrum of geospatial elements, including terrain imagery, atmospheric effects, and detailed 3D OpenStreetMap (OSM) structures. Serving as the foundational framework, this data enables the swift development of location-specific simulation environments. An aeronautical yoke and aeronautical throttle quadrant (Figure 2(b), (c), (d)) were used along with JSBSim FDM and Cesium ion 3D geospatial data. For the flight controls, the yoke features a solid steel yoke shaft with dual linear ball bearings and 180° yoke rotation. The throttle quadrant features six aviation levers, seven two-way programmable switches, and 14 warning lights. This hardware, along with the human operator, was crucial in the validation of our HHITLS, which acted as a test-bed for the MDP environment.

To validate all the algorithms in Section 3, soft crowd simulations have been developed using UE. Using this powerful tool, we can modify our agent control dynamics to allow for physical crowd interactions to effect agent movements. This phenomenon is described as a social force model (SFM). When using an SFM, the actions of pedestrians are under the influence of a set of forces they produce and receive. These forces include a repulsive force and attractive force between pedestrians as well as a repulsive force between pedestrians and physical boundaries (walls, light poles, tables, etc). Additionally, there is a special attraction force for pedestrians when in a calm state. When pedestrians are in a panicking state, such as an evacuation, these formulas change to emphasize repulsive forces and diminish the effects of the attractive

(a) (b) (c) (d)

Figure 2: Hardware-and-human-in-the-loop simulator.

Figure 3**:** Normal vs abnormal scenarios under different environments and sensor locations.

forces that are not applicable in states of panic. Leveraging SFMs and the PhysX physics engine provided by UE, we can create visually and physically realistic soft crowd models.

4.2 Validation of Extended Optical Flow Models

We have progressed to the validation phase within our meticulously crafted simulation environments to validate the proposed optical flow algorithm. Our focus was on emulating two distinct types of events: firstly, a Book Fair hosted at the University of Arizona (UA) Figure 3, and secondly, an Engineering Day celebration at Mississippi State University (MSU) Figure 4. These events were chosen for their typical crowd dynamics and the variety of movements expected at such gatherings. To incorporate realistic and challenging variables, we introduced environmental conditions by simulating both rainy and sunny conditions. Weather plays a significant role in surveillance as it affects visual acuity and consequently the performance of optical flow algorithms. Additionally, we diversified the camera perspectives to encompass two different angles: a side view to observe lateral movements and interactions and an inside view to closely monitor the dynamics within enclosed spaces.

The collection of eight figures in Figure 3 presents a simulation environment to test an optical flow algorithm under various conditions that could occur during social events. These conditions include changes in weather, camera perspectives, and the presence or absence of an emergency situation, such as a bombing attack. The simulation aims to evaluate the robustness and accuracy of the algorithm in different scenarios that a surveillance system might encounter. Figure 3(a) and (e) show a side view of an event in sunny and rainy weather conditions, respectively, under normal circumstances. The optical flow algorithm's performance here provides a baseline for movement patterns when there is no immediate threat or panic among the attendees. Conversely, Figure 3(b) and (f) introduce an element of chaos to the simulation, depicting an abnormal situation from the same side perspective. Here, the sunny and rainy conditions, combined with the hypothesized aftermath of a bombing attack tests the algorithm's ability to distinguish between normal crowd behavior and the erratic movement patterns that would signify an emergency. Ground sensor scenarios represented in Figure 3(c) and (g) simulate events taking place inside a venue.

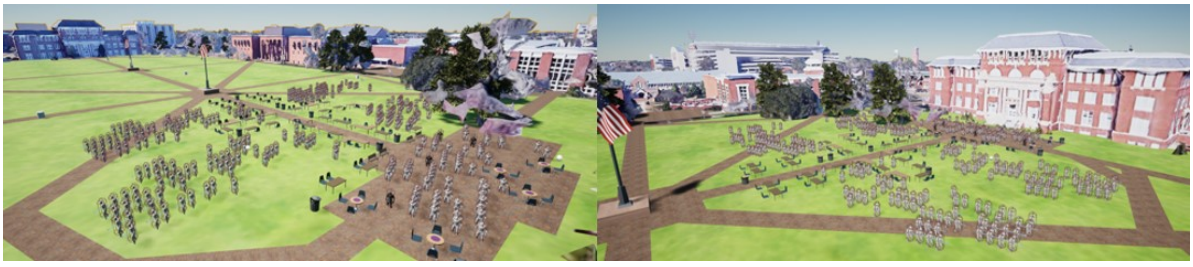


Figure 4**:** Engineering day event at Mississippi State University.

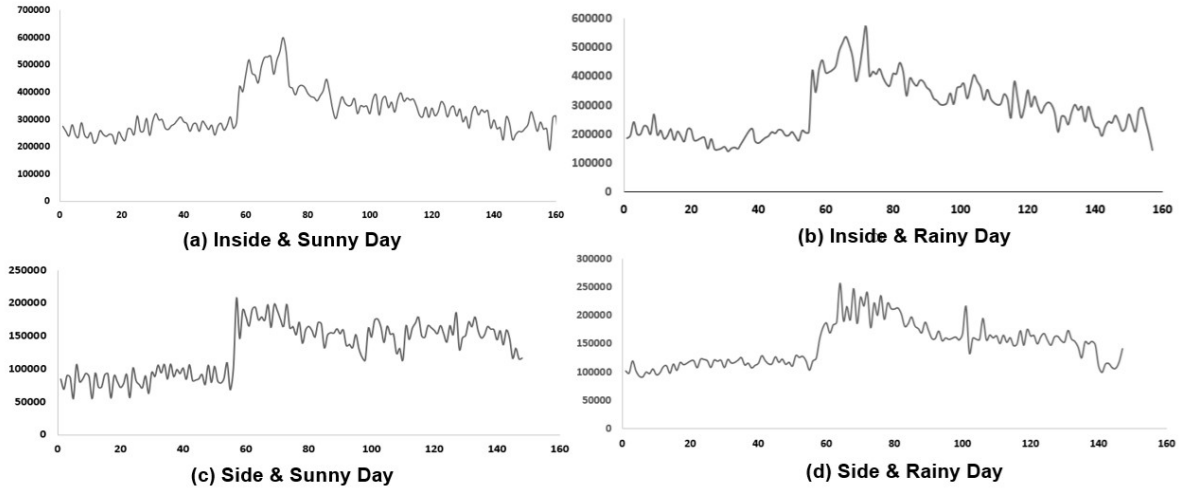


Figure 5: Optical flow magnitudes over time under different configurations.

Figure 3(c) assumes a sunny day outside, while Figure 3(g) assumes a rainy day, affecting the indoor lighting and potentially the camera's ability to capture clear images. Figure 3(d) and (h) challenge the algorithm further with abnormal situations indoors. Whether in sunny or rainy weather, the simulation assesses if the algorithm can effectively detect disturbances and track movement changes within an enclosed space where panic and rapid motion might be confined by walls and obstacles.

In the validation phase of our research, we hypothesize that the extended optical flow algorithm could accurately identify abnormalities, such as a simulated bombing incident within various surveillance settings. The complexity of these settings is dictated by differing camera placements—either inside or at the side of the event venue—and by varying meteorological conditions, specifically sunny versus rainy days. Figure 5 presents a concise summary of the algorithm's performance. The x-axis delineates the time expressed in video frames, while the y-axis quantifies the magnitude of the optical flow, which indicates movement within the frame. A visual examination of the four graphs reveals a spike in optical flow values at video frame 58 across all scenarios. This spike corresponds to the simulated moment of a bombing event, highlighting the algorithm's capability to discern an anomaly amidst diverse conditions. In each instance, regardless of whether the environment is bathed in sunlight or shrouded by rain, the algorithm consistently identifies the abnormal activity, as shown by the surge in optical flow magnitude. This adaptability to varied FOVs and weather conditions, while maintaining elevated sensitivity to disruptions validates the algorithm as a tool for security surveillance systems.

4.3 Experiments in Markov Decision Environments

The authors have set up a comprehensive experimental scenario to validate the proposed MDP framework. The environment is designed as a voluminous three-dimensional state space, encompassing a 10 *km* by 10 *km* by 10 *km* area (cubic). Within this space, every cubic kilometer is determined as a distinct unit, considering the operational limits of each UV. Central to this environment are two autonomous agents:

Table 1: Model parameters and configurations.

Model Parameters	Configuration
Discount factor: Gamma (γ)	$\{0.01, 0.1, 0.25, 0.75, 0.99\}$
Exploration rate: Epsilon (ε)	$\{0.01, 0.1, 0.25, 0.75, 0.99\}$
Learning rate: Alpha (α)	$\{0.1, 0.25, 0.5, 0.75, 0.9\}$
Decay factor: Lambda (λ)	$\{0.01, 0.025, 0.05, 0.075, 0.1\}$

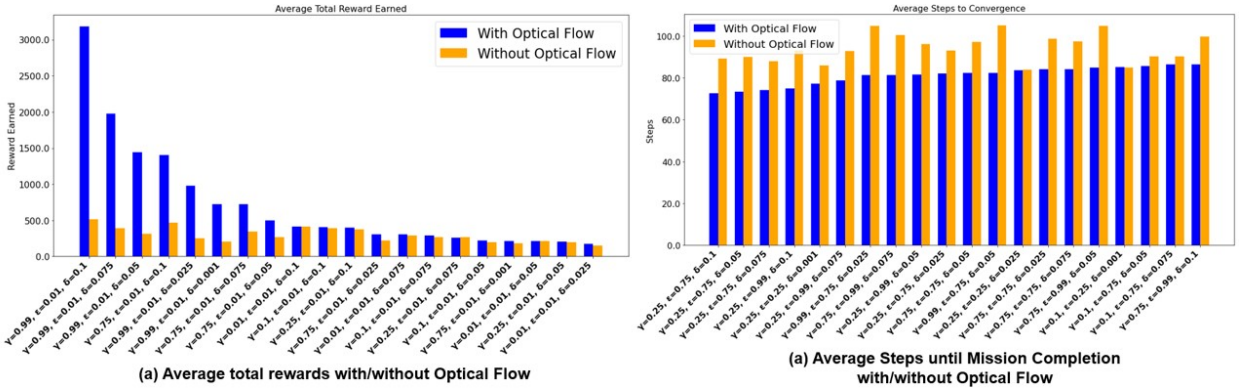


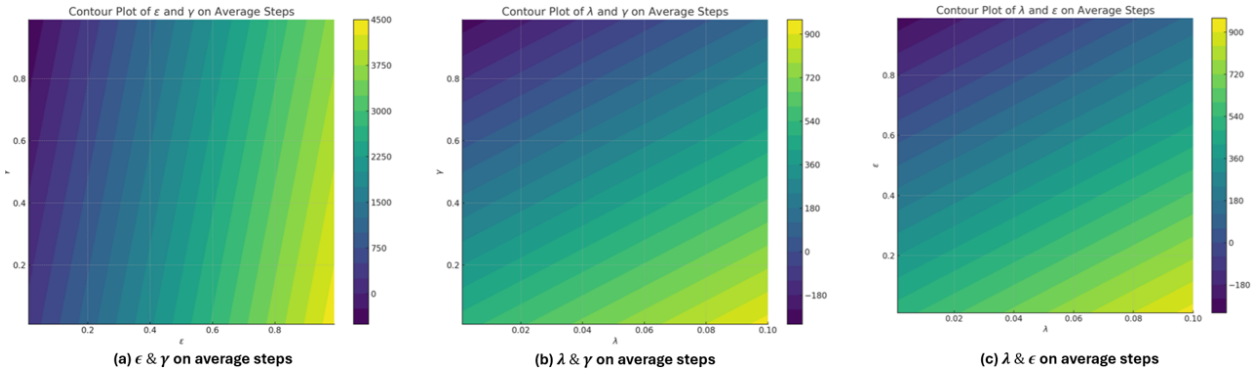
Figure 6**:** The average reward collected (a) and the average number of steps (b) until mission completion over the top 20 hyperparameter configurations with/without optical flow.

an own-ship UV, embodying the surveillance apparatus, and an intruder UV, which serves as the enemy of the surveillance operation. Each UV is allowed the capacity to traverse the grid by moving to an adjacent unit within the three-dimensional lattice at each discrete time step, reflecting realistic constraints on mobility. The authors have selected to adopt a similar action structure for this experiment, as described in Section 3. Here, each UV is allowed to choose from six potential action directions: upward, downward, leftward, rightward, forward, and backward, providing a comprehensive suite of navigational possibilities. The surveillance UV, referred to as the own-ship, is designed with a singular imperative: to locate the intruder within this vast expanse. On the other hand, the intruder drone is imbued with a twofold mission: to persistently evade the prying sensors of the own-ship and to navigate towards a predetermined 'target' situated at the origin point of the state space. Initially, the intruder's maneuvers are randomized, signifying an element of unpredictability and spontaneity in its behavior.

As shown in Table 1, we selected five different values to test for each of the four parameters (γ, ε, α, and λ), resulting in 625 different simulation configurations. Specifically, the parameters γ and ε both tested the same five values: $\{0.01, 0.1, 0.25, 0.75, 0.99\}$. Due to the long convergence rates observed with higher decay values ($\lambda > 0.1$), we tested five smaller values for decay: $\{0.01, 0.025, 0.05, 0.075, 0.1\}$. For α, we used the values: $\{0.1, 0.25, 0.5, 0.75, 0.9\}$. We collected 400 runs for each configuration to construct statistical inferences on the hyperparameters. This was achieved by performing 20 replications of 20 epochs for each parameter set. The 625 configurations are made up of each combination of possible values for gamma (γ), epsilon (ε), decay (λ), and alpha (α). Thus, we conducted 400 runs per configuration, resulting in a total of 250,000 simulation runs (400 runs $\times$ 625 configurations) for the hyperparameter analysis.

In Figure 6, we present the outcomes of our experiments comparing the standard MDP framework with the Optical Flow-based MDP framework. The authors monitored two performance measures: the average reward until mission completion and the average number of steps until mission completion. Figure 6(a) shows the top 20 configurations yielding the highest average rewards, while Figure 6(b) illustrates the bottom 20 configurations yielding the lowest average steps until mission completion. Each bar represents a distinct configuration, distinguished by varying levels of all four hyperparameters ($\gamma, \varepsilon, \alpha,$ and λ).

Figure 6(a) demonstrates that our proposed optical flow-based MDP framework enhances average reward outcomes across all top 20 configurations. This improvement indicates that utilizing visual information via optical flow contributes to tracking the intruder closer in the UV operations, enhancing surveillance effectiveness. Figure 6(b) focuses on operational efficiency, measured by the average number of steps required to complete the mission. The optical flow-based MDP consistently shows a significant reduction in the steps needed to intercept the intruder. This efficiency not only potentially reduces operational costs but also enables more agile responses to dynamic environmental cues, which are crucial for applications such as UV surveillance and interception missions.



(a) ϵ & γ on average steps (b) λ & γ on average steps (c) λ & ϵ on average steps

Figure 7: Contour plots based on hyperparameter analysis.

We also conducted Analysis of Variance (ANOVA) tests to determine the influence of different hyperparameters on two performance measures: average rewards and steps until mission completion. The results indicate that, except for α, all three hyperparameters (λ, γ, and ε) significantly affect both performance effectiveness (in terms of average reward at the end of the mission) and efficiency (in terms of average steps until mission completion), as all of them have p-values less than 0.01. To investigate further, we developed three contour plots (Figure 7) to demonstrate the effects of these hyperparameters on the average number of steps until mission completion for performance efficiency.

Three different contour plots illustrate the relationships Figure 6 between two significant parameters where the other is given. Regions with darker colors represent lower average steps, indicating more efficient mission completions. The first plot shows that lower values of ε help maintain operational efficiency regardless of the γ value, which suggests that an optimal configuration for these parameters may involve taking fewer randomized actions at each step. The second contour plot shows the interaction between the decay rate (λ) and the gamma value (γ). Darker regions are observed with an increase in λ combined with a larger γ value, indicating that less exploration of the environment while still prioritizing future rewards is key to maintaining operational efficiency in this setting. The last contour plot shows the effect of ε and λ on the average steps until mission completion. As observed in the plot, a combination of mid to high values of ε with lower values of λ seems to be associated with a lower number of steps. This suggests that if one has some freedom to take random actions at each point, it is better to prioritize immediate gains to minimize the number of average steps until mission completion.

5 SUMMARY AND FUTURE WORK

Our research successfully integrated an extended optical flow-based surveillance algorithm with an enhanced MDP model to optimize surveillance systems' operational efficiency and effectiveness. We investigated the role of optical flow algorithms in detecting abnormalities, such as simulated bombing attacks, across various environmental conditions and camera configurations. We also analyzed the interplay between the MDP variables and their impact on the surveillance system's performance by seamlessly adjusting key hyperparameters in both optical flows and MDP within a simulation environment that mirrors real-world conditions. Our findings highlighted the significance of decay, as well as the interactions between reward size with epsilon and decay, in influencing the model's convergence speed. This comprehensive study not only demonstrates the potential of combining optical flow algorithms with MDP for enhanced surveillance capabilities but also sets a new benchmark in the field of autonomous UV surveillance.

One future direction is to incorporate environmental effects, such as wind, weather conditions, and moving objects, to determine transition probabilities and aerodynamics realistically. This integration can also relate to assessing the UV's battery consumption, which is crucial for determining the end conditions of mission operations. By doing so, researchers can further harness the benefits of the HITL framework to derive more efficient control strategies. Moreover, considering multiple agents in the scene (e.g., multiple

intruders with different missions) could be another direction to investigate, focusing on the coordination and integration among UVs. This would allow for studying the impact of collaborative and adversarial interactions in multi-agent environments, leading to more resilient UV strategies.

REFERENCES

Abdelmaksoud, S. I., M. Mailah, and A. M. Abdallah. 2020. "Control Strategies and Novel Techniques for Autonomous Rotorcraft Unmanned Aerial Vehicles: A Review". *IEEE Access* 8:195142–195169.

Ajmera, Y. and S. P. Singh. 2020. "Autonomous UAV-based Target Search, Tracking and Following using Reinforcement Learning and YOLOFlow". In *2020 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR)*, 15–20. IEEE.

Aminzadeh, A., M. A. Atashgah, and A. Roudbari. 2018. "Software in the Loop Framework for the Performance Assessment of a Navigation and Control System of an Unmanned Aerial Vehicle". *IEEE Aerospace and Electronic Systems Magazine* 33(1):50–57.

Chen, D., K. Qian, K. Liao, and Y. Zhu. 2020. "Research on Navigation Algorithm of UAV Optical Flow/Inertia Multi-Information Fusion". In *2020 IEEE 4th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, Volume 1, 1594–1597. IEEE.

Chen, Y., X. Yin, S. Li, and X. Yin. 2024. "Optimal Control Synthesis of Markov Decision Processes for Efficiency with Surveillance Tasks". *arXiv preprint arXiv:2403.18632*.

Delamer, J.-A. and S. Givigi. 2020. "Trust in Multi-Vehicle Systems Using MDP Control Strategies". In *2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2377–2382. IEEE.

Di, X. and R. Shi. 2021. "A Survey on Autonomous Vehicle Control in the Era of Mixed-Autonomy: From Physics-based to AI-Guided Driving Policy Learning". *Transportation Research Part C: Emerging Technologies* 125:103008.

dos Santos, M. A. A. and K. C. T. Vivaldini. 2022. "A Review of the Informative Path Planning, Autonomous Exploration and Route Planning Using UAV in Environment Monitoring". In *2022 International Conference on Computational Science and Computational Intelligence (CSCI)*, 445–450. IEEE.

Frattolillo, F., D. Brunori, and L. Iocchi. 2023. "Scalable and Cooperative Deep Reinforcement Learning Approaches for Multi-UAV Systems: A Systematic Review". *Drones* 7(4):236.

Horn, B. K. P. and B. G. Schunck. 1981, August. "Determining Optical Flow". *Artificial Intelligence* 17(1):185–203.

Huang, M., C.-w. Zhu, and F.-m. Zhao. 2020. "A Hardware-in-the-Loop Simulation Method for the Evaluation of Flight Control Systems". In *2020 5th International Conference on Automation, Control and Robotics Engineering (CACRE)*, 320–324. IEEE.

Jeelani, I. and M. Gheisari. 2021. "Safety Challenges of UAV Integration in Construction: Conceptual Analysis and Future Research Roadmap". *Safety science* 144:105473.

Kalnoor, G. and G. Subrahmanyam. 2020. "A Review on Applications of Markov Decision Process Model and Energy Efficiency in Wireless Sensor Networks". *Procedia Computer Science* 167:2308–2317.

Lamond, B. F. and A. Boukhtouta. 1996. "Optimizing Long-term Hydro-power Production using Markov Decision Processes". *International Transactions in Operational Research* 3(3-4):223–241.

Lee, S., S. Jain, Y. Yuan, Y. Zhang, H. Yang, J. Liu *et al*. 2019. "Design and Development of a DDDAMS-based Border Surveillance System via UVs and Hybrid Simulations". *Expert Systems with Applications* 128:109–123.

Li, F., F. Jörg, X. Li, and T. Feenstra. 2022. "A Promising Approach to Optimizing Sequential Treatment Decisions for Depression: Markov Decision Process". *PharmacoEconomics* 40(11):1015–1032.

Li, L., Q. Cheng, K. Xue, C. Yang and Z. Han. 2020. "Downlink Transmit Power Control in Ultra-dense UAV Network based on Mean Field Game and Deep Reinforcement Learning". *IEEE Transactions on Vehicular Technology* 69(12):15594–15605.

Li, P., X. Hao, J. Wang, Y. Gu and G. Wang. 2019. "UAV Obstacle Detection Algorithm Based on Improved ORB Sparse Optical Flow". In *2019 IEEE 4th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*, Volume 1, 562–569. IEEE.

Liu, Y., K. Xie, and H. Huang. 2021. "VGF-Net: Visual-Geometric Fusion Learning for Simultaneous Drone Navigation and Height Mapping". *Graphical Models* 116:101108.

Lugo-Cardenas, I., S. Salazar, and R. Lozano. 2016. "The Mav3dsim Hardware in the Loop Simulation Platform for Research and Validation of UAV Controllers". In *2016 International Conference on Unmanned Aircraft Systems (ICUAS)*, 1335–1341. IEEE.

Maxwell, P., D. Larkin, and C. Lowrance. 2013. "Turning Remote-Controlled Military Systems into Autonomous Force Multipliers". *IEEE Potentials* 32(6):39–43.

MeiJie, Z., M. Ye, and T. Chang. 2021. "Multi-Agent Unmanned Swarm Combat Architecture Based on OODA Loop". *Advances in Computer, Signals and Systems* 5(1):81–87.

Miller, B., S. Hasbrouck, and B. Udrea. 2021. "Development of Human-Machine Collaborative Systems through Use of Observe-Orient-Decide-Act (OODA) Loop". In *ASCEND 2021*, 4092.

Panthakkan, A., W. Mansoor, and H. Al Ahmad. 2023. "Accurate UAV-Based Vehicle Detection: The Cutting-Edge YOLOv7 Approach". In *2023 International Symposium on Image and Signal Processing and Analysis (ISPA)*, 1–4. IEEE.

Primatesta, S., A. Rizzo, and A. la Cour-Harbo. 2020. "Ground Risk Map for Unmanned Aircraft in Urban Environments". *Journal of Intelligent & Robotic Systems* 97:489–509.

Richards, C. 2004. *Certain To Win: The Strategy of John Boyd, Applied to Business*. Xlibris Corporation.

Silano, G., P. Oppido, and L. Iannelli. 2019. "Software-in-the-Loop Simulation for Improving Flight Control System Design: a Quadrotor Case Study". In *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*, 466–471. IEEE.

Terashima, K., K. Kobayashi, and Y. Yamashita. 2024. "On Reward Distribution in Reinforcement Learning of Multi-agent Surveillance Systems with Temporal Logic Specifications". *Advanced Robotics* 38(6):386–397.

Wang, C.-Y., A. Bochkovskiy, and H.-Y. M. Liao. 2023. "YOLOv7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-time Object Detectors". In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 7464–7475.

Wang, J., Y. Li, R. X. Gao, and F. Zhang. 2022. "Hybrid Physics-based and Data-driven Models for Smart Manufacturing: Modelling, Simulation, and Explainability". *Journal of Manufacturing Systems* 63:381–391.

Wang, W., L. Wang, J. Wu, X. Tao, and H. Wu. 2022. "Oracle-guided Deep Reinforcement Learning for Large-scale Multi-UAVs Flocking and Navigation". *IEEE Transactions on Vehicular Technology* 71(10):10280–10292.

Young, A., J. Taves, A. Elmquist, S. Benatti, A. Tasora, R. Serban *et al*. 2022. "Enabling Artificial Intelligence Studies in Off-Road Mobility Through Physics-Based Simulation of Multiagent Scenarios". *Journal of Computational and Nonlinear Dynamics* 17(5):051001.

Yun, W. J., S. Park, J. Kim, M. Shin, S. Jung, D. A. Mohaisen *et al*. 2022. "Cooperative Multiagent Deep Reinforcement Learning for Reliable Surveillance via Autonomous Multi-UAV Control". *IEEE Transactions on Industrial Informatics* 18(10):7086–7096.

Zhang, Y., S. Xia, B. Zhang, and J. Liu. 2023. "Detecting Moving Objects from Moving Background by Optical Flow Decomposition". In *2023 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, 0990–0994.

Zhang, Y., T. Zhang, J. Liu, W. Kang, R. Liang, and B. G. Potter. 2024. "Profile Extraction and Defect Detection for Stereolithography Curing Process based on Multi-Regularized Tensor Decomposition". *Journal of Manufacturing Systems* 74:100–111.

Zhe, Z., N. Yifeng, and S. Lincheng. 2020. "Adaptive Level of Autonomy for Human-UAVs Collaborative Surveillance using Situated Fuzzy Cognitive Maps". *Chinese Journal of Aeronautics* 33(11):2835–2850.

Zhou, Y.-T. and J.-C. Piao. 2023. "A Lightweight YOLOv7 Algorithm for Steel Surface Defect Detection". In *2023 IEEE 6th International Conference on Pattern Recognition and Artificial Intelligence (PRAI)*, 285–289. IEEE.

AUTHOR BIOGRAPHIES

SEUNGHAN LEE is an assistant professor in the Industrial and Systems Engineering Department at Mississippi State University. His research encompasses modeling and analysis of data-driven systems under uncertainty using simulation and stochastic. Major application areas include disaster relief and mitigation efforts, smart city and homeland security, and healthcare operations. His email address is slee@ise.msstate.edu.

YINWEI ZHANG is a Ph.D. in the Department of Systems and Industrial Engineering, at the University of Arizona. His research interests include machine learning, statistical modeling and analysis, optimization, and image processing. His e-mail address is zhangyinwei@arizona.edu.

AARON LEGRAND is a graduate research assistant in the Department of Industrial Engineering at Mississippi State University. His research interests include reinforcement learning, physics-based simulations of crowd evacuations, and Unmanned Aerial Vehicle (UAV) surveillance simulations. He is currently working to obtain his Master's of Science in Industrial Engineering. His email address is awl139@msstate.edu.

SETH GIBSON-TODD is a graduate research assistant pursuing his Master of Science degree in the Department of Industrial and Systems Engineering at Mississippi State University. His research interests include Autonomous Mobile Robots (AMR) and their applications to material handling in manufacturing, MDP simulation, and development of training methods using Virtual Reality (VR) and Augmented Reality (AR). His email address is smg639@msstate.edu.