

Reinforcement Learning of Dolly-In Filming Using a Ground-Based Robot

Philip Lorimer¹, Jack Saunders¹, Alan Hunter², and Wenbin Li¹

Abstract—Free-roaming dollies enhance filmmaking with dynamic movement, but challenges in automated camera control remain unresolved. Our study advances this field by applying Reinforcement Learning (RL) to automate dolly-in shots using free-roaming ground-based filming robots, overcoming traditional control hurdles. We demonstrate the effectiveness of combined control for precise film tasks by comparing it to independent control strategies. Our robust RL pipeline surpasses traditional Proportional-Derivative controller performance in simulation and proves its efficacy in real-world tests on a modified ROSBot 2.0 platform equipped with a camera turret. This validates our approach’s practicality and sets the stage for further research in complex filming scenarios, contributing significantly to the fusion of technology with cinematic creativity. This work presents a leap forward in the field and opens new avenues for research and development, effectively bridging the gap between technological advancement and creative filmmaking.

I. INTRODUCTION

Robotic technology has dramatically transformed the filmmaking landscape, offering unprecedented flexibility in camera movement and expanding creative possibilities [1]. Free-roaming dolly systems have evolved from traditional, track-constrained designs, enabling more flexible, three-dimensional camera movement. This technological advancement, however, introduces significant challenges in control adaptability and operational consistency, particularly in executing traditional film techniques like dolly-in shots. The reliance on operators with varying skills exacerbates these challenges, introducing variability in film quality that can detract from the system’s potential. While some classical control methods have been applied in similar domains, our research uses Reinforcement Learning (RL) to develop an advanced control mechanism that automates these techniques, providing a foundation for more precise and adaptable robotic cinematography.

Within this context, RL [2] emerges as a promising avenue to enhance the adaptability and precision of ground-based robotic filming systems. Prior studies have extensively explored the capabilities of autonomous systems in filmmaking, with particular emphasis on aerial cinematography [3] and static camera work. These investigations highlight the potential of computer vision and classical control methods for task planning and execution, demonstrating notable successes in their application [4], [5], [6]. Concurrently, the shift towards

data-driven approaches, particularly RL, offers significant adaptability by learning from data and reducing dependence on explicit models [4], [7]. Specifically, RL has shown considerable promise in mobile aerial filming and other applications, excelling at complex objectives such as framing and executing shots on diverse platforms [8], [9], [10], [11]. Despite these advancements, a notable gap remains in applying such innovations to ground-based robotic filming, particularly in managing complex coordination and control challenges.

Incorporating RL for controlling ground-based robots offers a pathway to execute dolly-in shots with greater versatility. In this work, ‘Independent’ control refers to the separate handling of actions such as throttle and steering. In contrast, ‘Combined’ control refers to the unified management of these actions within a single system. The distinction is critical, as the choice of these control architectures significantly impacts platform movement and camera framing. Aerial cinematography research often prefers independent control for its flexibility in meeting specific component needs [10], [12]. However, managing these controls separately can be challenging, particularly when integrating control disturbances, which can impair coordination and performance without sophisticated disturbance modelling [13]. In contrast, combined RL control treats the system as a unified entity, streamlining management and enhancing action coordination [14], offering a more practical approach for complex filming tasks.

Given these themes, our investigation addresses this gap by leveraging RL to autonomously control a ground-based robot in executing the traditional dolly-in shot, a staple in cinematic production. Central to our methodology is the abstraction of the robot and task dynamics. By distilling these into essential elements, we refine our RL approach to emphasise core cinematic movements, creating a scalable framework that accommodates the complexities of robotic cinematography. This abstraction showcases the effectiveness of RL within known contexts and supports extending our approach to tackle more intricate filming tasks and diverse environmental conditions.

The main contributions of this work are as follows:

- We evaluate combined and independent control strategies for dolly-in shots through throttle and steering actions on a ground-based filming platform, highlighting performance differences and their applicability in filmmaking.
- The direct application of a combined control RL agent to a ground-based filming robot achieves dolly-in shot performance matching that of a finely tuned PD con-

This work was supported by EPSRC Centre for Digital Entertainment with the grant number EP/L016540/1.

¹Department of Computer Science, University of Bath, UK, {pall20, js3442, w.li}@bath.ac.uk

²Department of Mechanical Engineering, University of Bath, UK, A.J.Hunter@bath.ac.uk

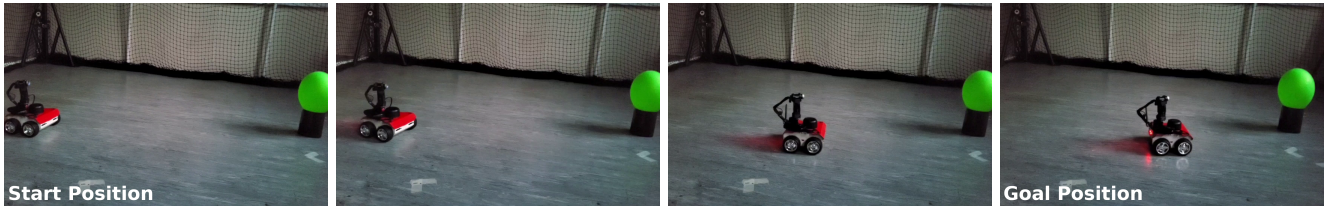


Fig. 1: Sequence of the dolly-in shot in the real world, showing the RL agent guiding our ground-based filming robot during the dolly-in shot from start to finish, using precise throttle, steering, pan, and tilt adjustments.

troller, as demonstrated across simulated and real-world tests.

- We demonstrate a pipeline to train RL agents on ground-based filming robots, enabling them to simulate and execute successful dolly-in shots in real-world tests, achieving a strong positive correlation in zero-shot Sim2Real transfers.

This work addresses the challenges posed by advanced robotic technologies in filmmaking, highlighting the untapped potential of reinforcement learning (RL) in this field. By framing our contributions within the broader discourse on autonomous camera systems, we illuminate a path toward achieving greater creative flexibility and technical precision in cinematography. Our findings lay a robust foundation for future research to leverage RL for more complex and creatively demanding tasks on mobile ground-based filming robots. This research not only pushes the boundaries of robotic cinematography but also establishes an abstraction approach as the backbone of our methodology, ensuring that our solutions are effective, adaptable, and scalable, thus setting the stage for continued advancements in this domain.

The paper proceeds as follows: Section II formalises the dolly-in shot. Section III covers reinforcement learning techniques and agent configurations, while IV details our experiments and results. Discussions and conclusions are in sections V and VI.

II. FORMALISATION OF DOLLY-IN SHOT

The dolly-in shot, a prevalent technique in filmmaking, emphasises the subject within a scene. Consider the subject as a simple green ball. The camera moves smoothly toward this subject [15], [16], with it centred in the frame, and concludes the shot once the subject occupies the desired portion of the image. To analyse and model this technique, we must define it precisely and quantifiably. Such a definition allows us to apply RL to refine the reward function systematically, focusing on measurable factors:

- The subject’s proportion within the frame.
- The subject’s position within the frame.
- The subject’s orientation relative to the camera’s movement.

Quantifying these elements enables the systematic optimisation of the dolly-in shot for cinematic impact.

A. Quantitative Definition of Dolly-In Shot

Image Processing: We begin by segmenting the image frame into two distinct regions: the subject and the back-

ground. For this, we use colour thresholding for our green ball subject, using OpenCV [17]. This process generates a binary mask $g(x,y)$, where pixels belonging to the subject are assigned a value of 1, while the background is set to 0. Any subsequent noise removal is carried out using OpenCV.

After segmentation, we calculate the subject’s geometric moments, which are computationally efficient descriptors for the subject’s shape, size, and position [18], defined by:

$$M_{ij} = \sum_x \sum_y x^i y^j g(x,y) \quad (1)$$

These geometric moments are used to extract essential features of the subject, facilitating subsequent analysis.

Subject’s Image Area: The proportion of the frame occupied by the subject is quantified using the zeroth-order geometric moment, M_{00} , which represents the sum of all non-zero pixels within the binary mask of the subject. It is calculated as: $M_{00} = \sum_x \sum_y g(x,y)$. To express this as a percentage of the frame, we calculate the area percentage A by dividing M_{00} by the total number of pixels s in the image:

$$A = \frac{M_{00}}{s} \quad (2)$$

Position of the subject in Frame: The subject’s position within the frame is determined by its centroid, representing the average position of all constituent points, calculated using the first-order moments μ_{pq} :

$$\mu_{pq} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q g(x,y) \quad (3)$$

Where $\bar{x} = \frac{M_{10}}{M_{00}}$ and $\bar{y} = \frac{M_{01}}{M_{00}}$ are coordinates of the centroid. Here, M_{10} and M_{01} represent the sums of x and y coordinates of all non-zero pixels in the binary mask $g(x,y)$.

Subject Offset Angle: Determining the subject’s offset angle relative to the robot’s heading is crucial for executing smooth dolly-in shots and minimising counter-steering. This involves a two-step process:

First, we calculate the camera offset angle α , which indicates the subject’s angular deviation from the centre of the camera’s Field of View (FoV). This calculation uses the camera’s fixed parameters: the pixel width w , the FoV f , the subjects’ x -axis position O_p and the image’s midpoint $M = \frac{w}{2}$. The pixel offset P is $O_p - M$. The camera offset angle α is calculated using the formula:

$$\alpha = P \times \left(\frac{f}{2/M} \right) \quad (4)$$

This formula adjusts the angular deviation based on the camera’s FoV and the pixel width.

Next, we calculate the subject’s offset angle orientation θ by incorporating the pan-angle θ_p from the robot’s joint sensors. The overall offset angle θ , which guides the robot’s alignment towards the subject, is then defined as:

$$\theta = \alpha + \theta_p \quad (5)$$

This combined offset angle ensures the robot maintains a direct line towards the subject, facilitating a centred and steady approach in dolly-in shots.

Quantifying Outcomes: The success of the performed dolly-in shot is quantified by assessing progress towards predefined goals concerning the subject’s area and position within the frame. This assessment employs a distance metric $\delta(p_A, p_E, p_{\max})$, defined as follows:

$$\delta(p_A, p_E, p_{\max}) = \begin{cases} \frac{\Delta p}{p_E}, & \text{if } p_A < p_E \\ \frac{\Delta p}{p_E - p_{\max}}, & \text{if } p_A \geq p_E \end{cases} \quad (6)$$

Where $\Delta p = |p_A - p_E|$ denotes the absolute difference between the actual value p_A and the desired parameter value p_E . This metric, acting as a normalised relative error, measures how close the current shot is to the desired cinematic goal. A value closer to 0 indicates proximity to the optimal state, as outlined in Reference [19].

III. TWIN DELAYED DEEP DETERMINISTIC POLICY GRADIENT

The problem can be formulated as a Markov Decision Process (MDP)[2], defined as a tuple $(\mathcal{S}, \mathcal{A}, P, R, \gamma)$. At each timestep t , an agent in state $s_t \in \mathcal{S}$ chooses an action $a_t \in \mathcal{A}$. The agent then receives a reward r_{t+1} and transitions to the next state s_{t+1} , according to the probability distribution $P(s_{t+1}|s_t, a_t)$. The reward r_{t+1} is drawn from the reward distribution $R(s_t, a_t, s_{t+1})$. The agent’s objective is to maximise the expected future cumulative reward $\mathbb{E}[G_t|s_t]$, where G_t is the discounted return defined as $G_t = \sum_{k=t+1}^T \gamma^{k-t-1} R_k$ and $\gamma \in [0, 1]$ is the discount factor. We can therefore define the state-action value function as $Q_\pi(s, a) = \mathbb{E}[G_t|s_t, a_t]$, where the policy $\pi(a, s)$ maps a state s to a probability distribution over actions a .

To navigate the continuous action space inherent in our study, we implemented the Twin Delayed Deep Deterministic Policy Gradient (TD3) algorithm [20]. TD3 addresses the overestimation bias found in similar algorithms by the innovative use of two Q-functions Q_{ϕ_1} and Q_{ϕ_2} , along with a mechanism for delayed policy updates. These features significantly enhance the stability and accuracy of value estimation, which is crucial for achieving precise control in complex cinematographic tasks.

TD3 introduces target policy smoothing and action noise clipping to mitigate the exploitation of erroneous estimations, a critical advancement over previous methods. By adding clipped noise to target actions, $a'(s') = \text{clip}(\mu_{\theta_{\text{target}}}(s') + \text{clip}(\varepsilon, -c, c), a_{\text{low}}, a_{\text{high}})$, where $\varepsilon \sim \mathcal{N}(0, \sigma)$, the algorithm ensures actions remain within feasible bounds. Additionally,

TD3 employs the minimum of both Q-functions to update the target state-action value, fostering a more disciplined learning process.

The learning process trains the Q-functions by regressing to a target that reflects the discounted future return, while the policy μ_θ is refined to maximize Q_{ϕ_1} . To maintain stability and ensure effective convergence, target networks are periodically updated using Polyak averaging, which blends the old and new parameters to achieve a balanced update rate.

By implementing the TD3 algorithm, our research addresses the intricate control demands of robotic cinematography, demonstrating the potential of reinforcement learning in facilitating sophisticated, real-world tasks. By elucidating the algorithm’s structure and our application methodology, we underscore the versatility and robustness of TD3 in handling the nuanced requirements of our study’s objectives. This adaptation demonstrates the practicality of our approach and paves the way for future exploration in the realm of autonomous robotic filming, where precise control and adaptability are paramount.

A. Agent Configuration: Actions, State Variables, and Reward Shaping

This section outlines the RL agent configurations deployed in our experiments, focusing on their action and state spaces. It then defines the reward functions.

We detail independent and combined control strategies for operating throttle and steering actions, noting key distinctions: independent employs two separate agents for platform management. In contrast, combined control consolidates this within a single agent. Furthermore, we introduce a complex agent designed to utilise all degrees of freedom —throttle, steering, pan, and tilt— to execute dolly-in shots on a ground-based filming platform efficiently.

State Variable	Description	State Variable	Description
s_1	Object area	s_6	Centroid y-error
s_2	Object area error	s_7	Pan angle
s_3	Centroid x-position	s_8	Tilt angle
s_4	Centroid x-error	s_9	Object offset angle
s_5	Centroid y-position		

TABLE I: State variables and descriptions for the RL agent.

1) *Action spaces:* As detailed in Table II, the action spaces include throttle control (a_1), steering control (a_2), camera pan (a_3), and tilt control (a_4). These actions span from [-1, 1], allowing scaling and application within the simulation environment. Table III shows the distinct functional roles across different agents, highlighting their specialisation.

2) *State spaces:* Table I specifies the state variables used for all agents, tailored for their specific control tasks:

- *Independent Agents:* Function with unique state spaces. The Throttle agent’s state space, $\mathcal{S}_T = \{s_1, s_2\}$, contrasts with the Steering agent’s, $\mathcal{S}_S = \{s_3, s_4\}$, reflecting their specialised control focus.

Action Variable	Description
a_1	Throttle control (move forward/backward)
a_2	Steering control (turn left/right)
a_3	Camera pan control (turn left/right)
a_4	Camera tilt control (look up/down)

TABLE II: Action variables and descriptions for agents.

Agent	Throttle a_1	Steer a_2	Pan a_3	Tilt a_4
Throttle agent	✓			
Steering agent		✓		
Combined agent	✓	✓		
Complex agent	✓	✓	✓	✓

TABLE III: Comparison of actions used by different agents.

- *Combined Agent*: Utilises a comprehensive state space, $\mathcal{S}_C = \{s_1, s_2, s_3, s_4\}$, amalgamating throttle and steering controls for holistic task management.
- *Complex Agent*: Incorporates all available state and action variables for comprehensive platform control.

These configurations, with agents tailored to specific filming objectives, are key to controlling and understanding cinematographic tasks in this study.

3) *Reward Shaping*: To guide the agents towards an optimal policy, using the criteria detailed in Section II, we utilise a reward function formulated as a weighted sum of n sub-objectives: $R(s, a) = \sum_{i=1}^n w_i \cdot r_i(s, a)$. This structure tailors the reward strategy to the specific objectives of each agent, directly influencing their actions to achieve the desired outcomes. Each weight w_i signifies the importance of the corresponding sub-objective [21]. Each objective’s reward is based on Equation 6, customised for the control objective unless noted differently.

For independent agents, reward functions focus on individual control tasks:

- 1) *Throttle Control Agent*: Focuses on maintaining the subject’s desired size within the frame. The reward, $R_{\text{area, original}}(s, a)$, depends on the actual area a_A , expected area a_E , and the maximum permissible area a_{max} , defined as:

$$R_{\text{area, original}}(s, a) = \delta(a_A, a_E, a_{\text{max}}) \quad (7)$$

- 2) *Steering Control Agent*: Aims to keep the subject centred within the frame. Its reward, $R_{\text{position}}(s, a)$, is calculated based on the actual position p_A , expected position p_E , and the maximum allowed position p_{max} :

$$R_{\text{position}}(s, a) = \delta(p_A, p_E, p_{\text{max}}) \quad (8)$$

The Combined agent’s reward function combines the area and position rewards, balancing them through weights w_1 and w_2 :

Combined Agent: The reward function is a weighted sum of the rewards using Equations 7 and 8:

$$R_{\text{combined}}(s, a) = w_1 \cdot R_{\text{area, scaled}} + w_2 \cdot R_{\text{position}} \quad (9)$$

Here, $R_{\text{area, scaled}}$ is an adaptation of $R_{\text{area, original}}$ to promote optimisation of the subject area, especially during smaller adjustments:

$$R_{\text{area, scaled}} = \begin{cases} -0.5 + \left(\frac{|p-k|}{k}\right) \times (-0.5) & \text{if } p \leq k \\ \frac{p-a_E}{a_E} \times 0.5 & \text{if } p > k \end{cases}$$

In this context, k represents the threshold value of $R_{\text{area, original}}$, and a_E is the target area.

Complex Agent: The reward function for the complex task, $R_{\text{complex}}(s, a)$, combines previously defined reward components with additional penalties for object offset and excessive action variations between consecutive time steps:

$$\begin{aligned} R_{\text{complex}}(s, a) = & w_1 \cdot R_{\text{area, scaled}} \\ & + w_2 \cdot R_{\text{position}} \\ & + w_3 \cdot R_{\text{object offset}} \\ & + p \end{aligned} \quad (10)$$

This reward function includes components, each tailored to a specific aspect of the filming task:

$w_1 \cdot R_{\text{area, scaled}}$: Rewarding desired subject size in the frame

$w_2 \cdot R_{\text{position}}$: Rewarding desired subject position (Eq. 8).

$w_3 \cdot R_{\text{object offset}}$: Penalising robot heading deviations from the subject using θ from Eq. 5.

p : Penalty for actions exceeding the threshold between consecutive timesteps.

B. Implementation Details

We outline the architectures and training parameters for our TD3 algorithm implementation. Our framework employs distinct networks for the actor and critic, undergoing training across episodes - each with 1500 timesteps. Key hyperparameters include a batch size of 128, a learning rate of 0.0005, a discount factor of 0.99, and a tau value for soft updates of 0.005. An experience replay buffer with a capacity of 10 million ensures diverse transition sampling.

Both the critic and actor networks consist of three fully connected layers. The initial layer, with 400 nodes, processes state and action inputs. This is followed by a 300-node layer, leading to an output layer that estimates action values for the critic and outputs actions for the actor, scaled by the maximum action value. ReLU activation functions facilitate non-linear learning across all layers except the output. The Adam is the optimiser for both networks and exploration noise is added during training to improve the exploration of the action space.

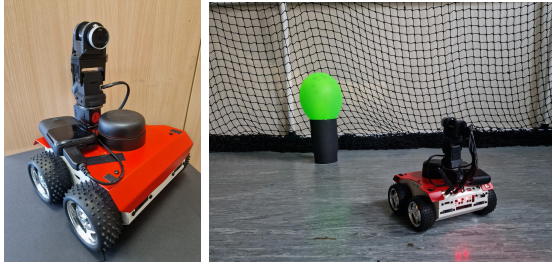
IV. EXPERIMENTAL RESULTS

This section details the experimental setup and outcomes, validating the effectiveness of RL agents in performing autonomous dolly-in shot tasks.

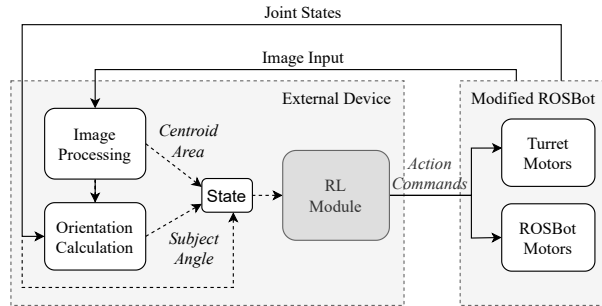
A. Simulation Setup

We chose Pybullet [22] for its accurate physics and environmental modelling, Pytorch [23] for its robust machine-learning capabilities, and the OpenAI Gym framework [24] for its standardised interface for reinforcement learning tasks.

B. Robot System Setup



(a) Modified ROSBot 2.0 Featuring Camera Turret for Filming Task.



(b) Data Flow and Software Component Diagram for the Modified ROSBot 2.0 in Filming Applications.

Fig. 2: Modified ROSBot 2.0 for Filming: Physical Modifications Overview (Top) and Software Components with Data Flow Diagram (Bottom).

We customised a ROSBot 2.0 [25] for this study (Fig. 2a), integrating it with an Interbotix PhantomX Vision Tracking Robot Turret [26] and a Microsoft LifeCam Webcam for advanced tracking and filming capabilities. The ROSBot handles basic control and sensory functions. At the same time, an external computer powered by an Intel(R) Core(TM) i7-10875H CPU and Nvidia RTX 2080 Super GPU supports image processing and RL computations. We employ the Robot Operating System (ROS) [27] for efficient control and messaging. Figure 2b showcases the software architecture, highlighting the RL module and the system’s information flow. After training in a simulated environment identical to our real world, the agent is ready for real-world validation to ensure its effectiveness outside the simulated conditions.

By bridging advanced hardware and software, this tailored robot system setup directly contributes to our research objectives of showcasing the practical applications of RL in robotic cinematography.

C. Simulation Experiments

Through simulation experiments, this section explores the efficacy of independent and combined control strategies for ground-based free-roaming robotic cinematography. Additionally, we assess an agent’s performance with complete degrees of freedom against a baseline PD controller, focusing on the dolly-in shot task.

1) *Independent versus Combined*: This experiment assesses RL agents’ utility, performance, and stability in developing optimal ground-based robotic cinematography systems, focusing specifically on the effectiveness of combined versus Independent control strategies. Our primary objective is to determine which strategy more effectively facilitates the execution of cinematography tasks, like the dolly-in shot, using throttle and steering actions with precision and stability.

To achieve this, we trained agents across 5,000 episodes, each lasting 1,500 timesteps. The agent setup, detailed in Section III-A, contrasts two independent agents - a steering agent for subject alignment and a throttle agent for distance adjustment - with a combined agent tasked with collective goal achievement. This approach facilitates a nuanced comparison of their performance on the dolly-in task and their potential applicability to more complex challenges, such as incorporating pan and tilt mechanisms.

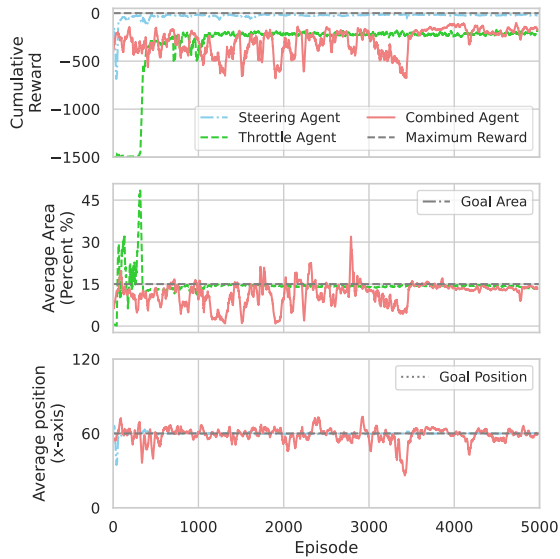
Our findings reveal distinct performance and stability patterns between the control strategies. Independent agents demonstrated quicker convergence (Fig. 3a) and higher stability than their combined counterparts, as shown by a 100-trial stability test (Fig. 3b). However, when evaluating task performance over 100 trials using normalised reward metrics for area and position accuracy (Section II), the combined strategy outperformed the independent approach, with a mean reward of -120 compared to -160. Scores closer to 0 indicate higher performance and efficacy, showing the centralised strategies performance in this context (Fig. 4).

The performance variance, particularly in centroid accuracy, highlights coordination challenges inherent in decentralised setups. While the independent throttle agent often met area-size goals, steering accuracy was compromised, indicating a pressing need for cohesive control strategies. These results confirm the potential of our RL-based pipeline for simulating robotic cinematography tasks, setting a solid foundation for exploring basic dolly-in tasks and beyond.

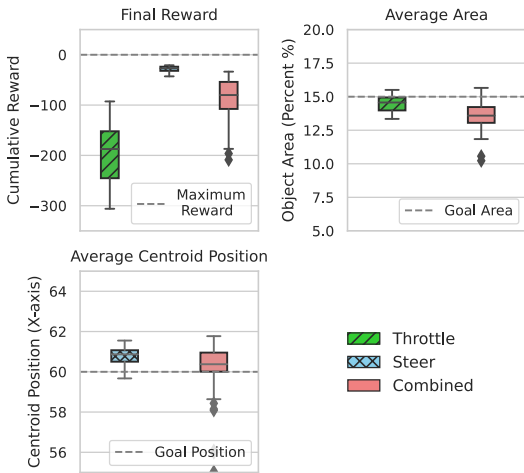
2) *Full Dolly-in Task comparison*: Our analysis of the dolly-in task demonstrates that an RL-trained agent, utilising its full range of motion capabilities (throttle, steering, pan, and tilt), achieves performance comparable to a finely hand-tuned PD controller. This comparison underscores the RL agent’s potential to match or surpass traditional control methods in robotic cinematography, marking a significant advancement in the field.

The complex agent was trained on the dolly-in task to conduct this comparison, incorporating strategies such as early stopping to prevent overfitting and ensure effective convergence. Progress in training the agent is detailed in Figure 5a. Concurrently, a PD controller was configured using the ‘simple-pid’ library ([28]), hand-tuned to achieve a performance goal aligned with that of the complex agent in a simulated environment.

The stability and performance of the complex agent and the classical PD controller were evaluated over 100 trials, as presented in Figure 5b. The results of this evaluation highlight the competitive performance of our trained RL



(a) Training results for each agent during training. Cumulative Reward, Average Centroid Position (X-axis), and Average Area per episode.



(b) Stability evaluation box plot for each agent.

Fig. 3: Training and stability evaluation for each agent.

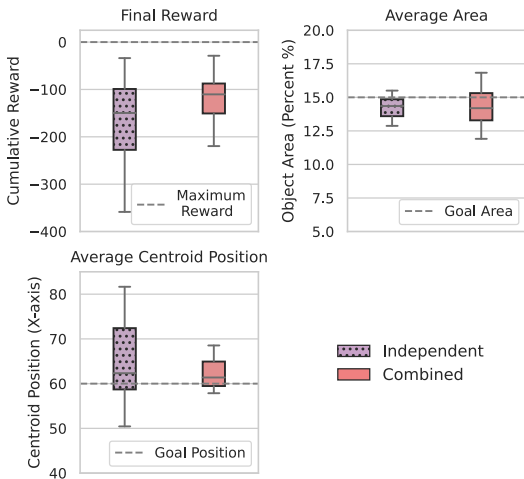


Fig. 4: Comparison of Combined and Independent control methods for dolly-in shot, over 100 trial runs.

agent compared to the PD controller. By closely matching the performance of traditional methods, our agent exhibits its capability to execute complex cinematographic tasks and signals a promising direction for the future of robotic cinematography.

D. Real-world Experiments

Bridging the 'reality gap' between simulation and real-world applications, our Sim2Real experiments aimed to validate the practical applicability of our RL model. Conducted in both simulated environments and real-world settings using the modified ROSBot 2.0 (Fig. 2a) and its digital twin, these experiments were designed to assess the generalisability and performance translation of our simulated agent into real-world scenarios.

Using the Simulation vs. Real-world Correlation Coefficient (SRCC, [29]) as a quantifier of performance fidelity, where values closer to 1 indicate high simulation-to-real-world translation accuracy, our findings demonstrate the robustness and adaptability of the RL model. This validation marks a pivotal step in autonomous robotic filming, showing that our model can seamlessly transition from simulated training to effective real-world execution without intermediary fine-tuning.

In detail, the Sim2Real transfer experiment evaluated the agent's performance from three distinct starting positions (left, right, and centre) across both environments, with ten runs per position totalling 30 runs. The results, documented in Table IV, reveal a consistently positive correlation between simulation and real-world performance, illustrating the agent's ability to achieve high rewards in both settings. This consistent performance underlines the model's capability for direct application in real-world cinematographic tasks, such as the dolly-in shot, further emphasising the potential of RL in bridging theoretical models with practical film-making applications (Fig. 1).

Our comprehensive approach—from simulation training to zero-shot Sim2Real transfer—highlights a scalable pathway for integrating RL into robotic cinematography, advancing the field towards more autonomous and adaptable robotic filming solutions.

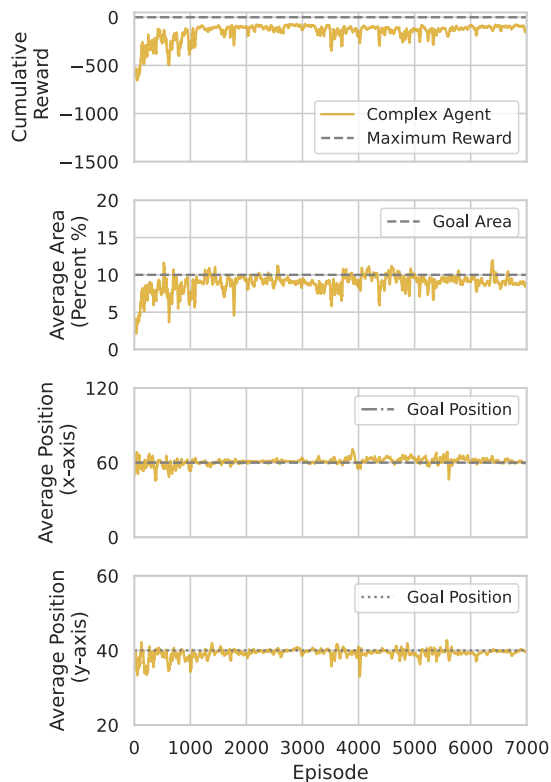
V. DISCUSSION AND LIMITATIONS

Our research demonstrates the potential to use RL to automate traditional dolly-in shots with a ground-based, free-roaming filming robot. It highlights the synergy between robotics and data-driven filmmaking. This approach demonstrates RL's feasibility in complex cinematographic tasks and sets a precedent for future interdisciplinary innovations.

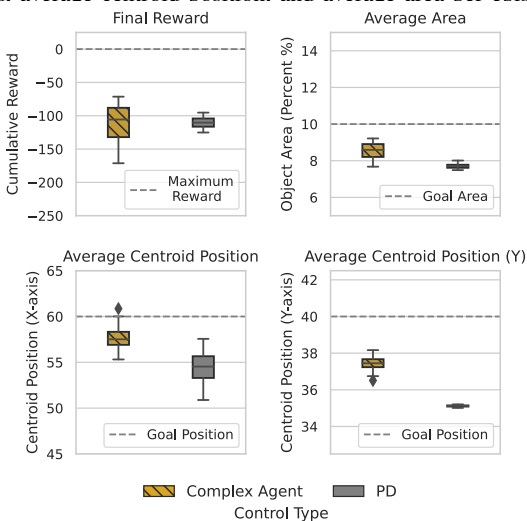
Our exploration of combined and independent control strategies reveals critical insights into their effectiveness and challenges. Combined control approaches offered enhanced action-objective synchronisation between actions and objectives but required longer training periods due to increased variables and larger search spaces. Despite these inherent challenges, combined control may be a viable alternative to independent systems, especially in complex, multi-action

	Cum. Reward		Object Area			Object Position (X-axis)			Object Position (Y-axis)		
	Sim.	Real.	Sim.	Real.	SRCC	Sim.	Real	SRCC	Sim.	Real	SRCC
Left	-170.77	-176.27	10.12	9.63	0.69	59.42	61.22	0.52	44.0	44.54	0.65
Right	-166.03	-166.35	10.09	10.38	0.80	59.32	59.08	0.72	44.1	44.23	0.83
Centre	-129.93	-134.96	10.09	11.09	0.46	59.79	55.67	0.56	44.0	46.51	0.69

TABLE IV: Comparison of 30 runs (10 per starting position: left, right, centre) in both simulation and real-world settings, demonstrating Sim2Real transfer effectiveness with SRCC values indicating strong simulation-to-reality correlation.



(a) Training outcomes for the complex task agent: cumulative reward, average centroid position, and average area per episode.



(b) Performance comparison between the RL agent and the baseline PD controller for dolly-in shots, using throttle, steering, pan, and tilt actions over 100 trials.

Fig. 5: Training and evaluation for complex agent.

scenarios such as those found in ground-based robotic filming, which require a level of coordination similar to that in aerial cinematography. However, independent control and Multi-Agent Reinforcement Learning (MARL) remain promising for tasks with clear objective divisions, potentially mitigating coordination challenges and warranting further study.

Our findings demonstrate that our RL agent, trained for the complex task, can equal or even outperform traditional PD controllers in simulation, as shown in Figure 5b. The Sim2Real experiments critically validate the RL model’s effectiveness for zero-shot transfer to real-world tests, showing a strong positive SRCC correlation. These findings underscore the utility of simulation-based training, demonstrating that performance in simulation can translate effectively to real-world applications, even on a limited, low-powered platform like the modified ROSBot 2.0. The TD3 algorithm, chosen for its strengths in continuous action spaces, proved practical for learning the dolly-in task. While this approach has proven effective, future work will explore comparisons with other RL methods to enhance performance further. However, it is important to note that there are currently no established benchmarks for this specific task, highlighting the novelty of our contribution.

Several challenges arose during real-world implementation, such as environmental variability and hardware limitations. The simplistic subject detection method required stable and consistent conditions, while the filming robot’s underpowered equipment further constrained the task’s complexity. These limitations, however, sharpened our focus on the core aspects of the problem. Through iterative testing and refinement of the RL model, we demonstrated its robustness and adaptability, which are crucial for real-world success.

The abstraction of the dolly-in shot task allowed us to isolate fundamental elements, setting the groundwork for effectively linking theoretical understanding with practical application, particularly in our real-world experiments. This approach led to the design of reward functions that emphasised key cinematic outcomes - subject size, position, and heading deviation - ensuring a platform-agnostic solution. As filming tasks and environments grow more complex, the reward function will need to be refined, introducing challenges in balancing reward design and component weightings.

With a more powerful robotic platform and a more robust and generalisable approach to subject detection, the need for high-fidelity digital twins becomes evident. While our approach is effective, it also provides ample opportunities for future development.

Imitation Learning (IL) offers a strategic approach to simplify reward design by utilising expert demonstrations to derive policies that capture desired filming behaviours. This approach could make the technology more accessible to non-experts. IL offers significant advantages over traditional RL by streamlining the learning process and more efficiently guiding the agent through the search space. This adaptability enhances the feasibility of RL-driven cinematography. It sets the stage for future research focused on expanding the applicability and robustness of our approach in dynamic and evolving filmmaking environments.

Our study illustrates the feasibility of RL in robotic cinematography for a ground-based filming platform and establishes a foundation for integrating theoretical understanding with practical filmmaking applications. The efficacy of combined strategies and the prospects of IL's in improving learning suggest new research and application opportunities.

VI. CONCLUSIONS AND FUTURE WORK

This study demonstrates the successful application of reinforcement learning (RL) to automate dolly-in shots with a ground-based filming robot, achieving performance comparable to traditional PD controllers. The effectiveness of the RL model, validated through simulation and real-world tests, represents a significant advancement in robotic cinematography. Combined control strategies showed considerable promise in synchronising actions and objectives for complex cinematic tasks, despite their complexity. This work lays a strong foundation for integrating RL into creative filmmaking, advancing the transition from theory to practice.

A. Future Work

To further refine the training process for cinematography agents, we plan to leverage imitation learning, utilising expert demonstrations to simplify and more effectively translate desired behaviours to RL agents. This approach aims to lower barriers for filmmakers, making advanced filmmaking technologies more accessible. Future research will focus on extending the applicability of RL in creative filmmaking, particularly in dynamic and evolving environments, thereby broadening the scope and impact of data-driven methods in the industry.

REFERENCES

- [1] R. Bogue, "The role of robots in entertainment," *Industrial robot an international journal.*, vol. 49, no. 4, pp. 667–671, 2022.
- [2] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press, 2018.
- [3] K. C. W. Goh, R. B. C. Ng, Y. K. Wong, N. J. H. Ho, and M. C. H. Chua, "Aerial filming with synchronized drones using reinforcement learning," *Multimedia Tools and Applications*, vol. 80, pp. 18125 – 18150, 2021.
- [4] J. Chen and P. Carr, "Autonomous camera systems: A survey," *AAAI Workshop - Technical Report*, vol. WS-14-06, pp. 18–22, 2014.
- [5] P. Carr, M. Mistry, and I. Matthews, "Hybrid robotic/virtual pan-tilt-zoom cameras for autonomous event recording," in *Proceedings of the 21st ACM International Conference on Multimedia*, MM '13, (New York, NY, USA), p. 193–202, Association for Computing Machinery, 2013.
- [6] A. Farag and A. Abdel-Hakim, "Virtual forces for camera planning in smart vision systems," pp. 269–274, 02 2005.
- [7] C. Chen, O. Wang, S. Heinzle, P. Carr, A. Smolic, and M. Gross, "Computational sports broadcasting: Automated director assistance for live sports," *Proceedings - IEEE International Conference on Multimedia and Expo*, 2013.
- [8] B. Singh, R. Kumar, and V. P. Singh, "Reinforcement learning in robotic applications: A comprehensive survey," *Artif. Intell. Rev.*, vol. 55, p. 945–990, feb 2022.
- [9] T. Zhang and H. Mo, "Reinforcement learning for robot research: A comprehensive review and open issues," *International Journal of Advanced Robotic Systems*, vol. 18, no. 3, p. 17298814211007305, 2021.
- [10] M. Gschwindt, E. Camci, R. Bonatti, W. Wang, E. Kayacan, and S. Scherer, "Can a Robot Become a Movie Director? Learning Artistic Principles for Aerial Cinematography," *arXiv e-prints*, p. arXiv:1904.02579, Apr. 2019.
- [11] N. Passalis and A. Tefas, "Deep reinforcement learning for controlling frontal person close-up shooting," vol. 335, pp. 37–47.
- [12] D. L. Leottau, A. Vatsyayan, J. R. del Solar, and R. Babuka, "Decentralized reinforcement learning applied to mobile robots," in *Robot Soccer World Cup*, 2016.
- [13] J. Zhao, T. Zhao, and N. Liu, "Fractional-order active disturbance rejection control with fuzzy self-tuning for precision stabilized platform," *Entropy*, vol. 24, no. 11, 2022.
- [14] C. F. Hayes, R. Rădulescu, E. Bargiacchi, J. Källström, M. Macfarlane, M. Reymond, T. Verstraeten, L. M. Zintgraf, R. Dazeley, F. Heintz, E. Howley, A. A. Irissappane, P. Mannion, A. Nowé, G. Ramos, M. Restelli, P. Vamplew, and D. M. Roijers, "A practical guide to multi-objective reinforcement learning and planning," *Autonomous Agents and Multi-Agent Systems*, vol. 36, apr 2022.
- [15] C. Manetta and R. A. Blade, "Glossary of Virtual Reality Terminology," *International Journal of Virtual Reality*, vol. 1, no. 2, pp. 35–39, 1995.
- [16] T. V. Nguyen, M. Xu, G. Gao, M. Kankanhalli, Q. Tian, and S. Yan, "Static saliency vs. Dynamic saliency: A comparative study," *MM 2013 - Proceedings of the 2013 ACM Multimedia Conference*, pp. 987–996, 2013.
- [17] G. Bradski and A. Kaehler, "Opencv 4.5.0," 2023.
- [18] L. G. Kotoulas and I. Andreadis, "Image analysis using moments," 2005.
- [19] L. Matignon, G. J. Laurent, and N. Le Fort-Piat, "Reward function and initial values: Better choices for accelerated goal-directed reinforcement learning," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 4131 LNCS, pp. 840–849, 2006.
- [20] S. Fujimoto, H. Hoof, and D. Meger, "Addressing function approximation error in actor-critic methods," in *Proceedings of the 35th International Conference on Machine Learning*, vol. 80, 2018.
- [21] A. Kusari and J. P. How, "Predicting optimal value functions by interpolating reward functions in scalarized multi-objective reinforcement learning," pp. 7484–7490. ISBN: 9781728173955.
- [22] E. Coumans and Y. Bai, "Pybullet, a python module for physics simulation for games, robotics and machine learning." <http://pybullet.org>, 2016–2023.
- [23] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, and S. Chintala, "Pytorch: An imperative style, high-performance deep learning library," in *Advances in Neural Information Processing Systems 32*, pp. 8024–8035, Curran Associates, Inc., 2019.
- [24] G. Brockman, V. Cheung, L. Pettersson, J. Schneider, J. Schulman, J. Tang, and W. Zaremba, "Openai gym," *arXiv preprint arXiv:1606.01540*, 2016.
- [25] Husarion, "Rosbot 2.0 official documentation." <https://husarion.com/manuals/rosbot-manual/>, 2020.
- [26] Interbotix, "Interbotix ros turrets." https://github.com/Interbotix/interbotix_ros_turrets/tree/main/interbotix_ros_xsturrets, 2022. GitHub repository.
- [27] Stanford Artificial Intelligence Laboratory et al., "Robotic operating system."
- [28] J. Frech, *simple-pid: A Simple PID Controller API for Python*, 2023.
- [29] A. Kadian, J. Truong, A. Gokaslan, A. Clegg, E. Wijmans, S. Lee, M. Savva, S. Chernova, and D. Batra, "Sim2real predictivity: Does evaluation in simulation predict real-world performance," *IEEE Robotics and Automation Letters*, vol. PP, pp. 1–1, 08 2020.