Risk-Aware Collision Avoidance for Multi-drone Cinematography

Rebecca Martin, Micah Corah, Sebastian Scherer
Robotics Institute, Carnegie Mellon University
{rebecca2, micahc, basti}@andrew.cmu.edu

I. Introduction and Related Work

Movies have been a key source of entertainment for many decades now. While an interesting plot is certainly necessary, audiences have come to expect good filming quality that brings the story to life. Action movies especially continuously add increasingly more elaborate stunts. However, it is not always possible to achieve all of this when filming from the ground only. Sometimes, filming needs to be done from places unreachable by humans. Drone cinematography does just that. However, these situations are inherently more dangerous and pose higher risks than regular cinematography.

Single-drone cinematography has been studied extensively both in academia and in industry. [1] - [3] compute smooth aerial camera plans during motion given user-defined artistic guidelines. In the commercial sector, Skydio [4] and DJI [5] drones are able to track and film actors in complex cluttered environments. Recently, multi-drone cinematography has gained popularity due to its viewpoint diversity, with work that focuses on coordinating multi-drone, human-guided shot execution under various constraints, such as smoothness, battery life, and mutual camera visibility [6] - [10]. [11] increases practicality for multi-UAV tracking of unscripted targets by removing the need for predefined shots while maximizing 3D shot diversity online. However, these works on aerial cinematography doe not generally address risks and collision avoidance in detail.

As such, our work will seek to draw on the broad literature on collision-avoidance in multi-robot systems. [12] proposes online path planning with inter-drone collision avoidance for indoor settings. [13] uses learning to predict the trajectories of dynamic obstacles for collision avoidance in congested environments. [14] uses nonlinear model predictive control (NMPC) for distributed collision avoidance of dynamic obstacles, sending low level control commands at a high frequency. [15] uses a MINVO basis to generate UAV trajectories that avoid static and dynamic obstacles in a decentralized and asynchronous manner. While these methods achieve collision avoidance, they do not account for more general tasks such as for cinematography. We propose a system that is able to autonomously coordinate a team of drones that are filming an actor or group of actors for purposes such as cinematography or offline 3D reconstruction while successfully avoiding static and dynamic obstacles.

II. Problem: Risk-aware cinematography

The risks associated with drone cinematography can generally be separated into two categories: safety risks and task failure. Safety risks include collisions with the actor(s), other drones, or static/dynamic obstacles in the environment. Strong winds or other inclement weather also make it difficult to control the drone, creating a hazardous situation. Task failure includes situations where people or obstacles are incorrectly classified, failure of the tracking system, inter-drone communication failures, where drones' information about each other is not up to date, and failure of the planning system, where the planner is unable to find a feasible trajectory that maintains the goals of the cinematography task.

speculating about the how. Not really describing the problem anymore While it is important to make sure that each system in our pipeline is robust, fail safe protocols should be used to properly mitigate these risks. Depending on the threat, it might make sense to have one drone or the entire team to fall back to their fail safe protocol. For example, while landing the drone autonomously or manually seems like the obvious choice, that may not be trivial—such as when flying near people or over harsh terrain. Different actions may also make sense depending on how reasonable we assume the people in the environment are; they could do everything in their power to preserve their safety, they could do nothing, or they could be adversarial. [16] discusses a variety of maneuvers that can be used to navigate unknown environments in fail safe protocols. However, fail safe protocols are primarily emergency precautions and do not necessarily address the cinematography task. We propose a multi-agent planning system that generates trajectories for filming for 3D reconstruction while navigating complicated environments and avoiding static and dynamic obstacles.

III. Method

In section 3A, we briefly describe the cinematography objective and cost functions, that are discussed in more detail in our previous work [17]. Section 3B discusses our new collision avoidance strategy for multi-drone filming teams. Section 3C gives a mathematical formulation of the strategy described in Section 3B.

A. Objective

Similar to [17], we use an optimization framework to calculate our set of trajectories. Our mathematical objective is to minimize the reconstruction error E_{recon} calculated with respect to the true target joints over time, where $\theta(t) \in \mathbb{R}^{P\times 3}$ is a vector containing the target's 3D coordinates for P joints at time t:

$$E_{recon} = \sum_{t=1}^{T} \|\hat{\theta}(t) - \theta(t)\|^2 \quad (1)$$

In our weighted set of cost functions, we use the same cost for trajectory smoothness, camera occlusion, and team formation as [17]. We propose a new cost function for obstacle avoidance. This new cost function not only considers static and dynamic obstacles, but also introduces the notion of risk.

B. Collision Avoidance

Filming with multiple drones requires the robots and actors to be within a certain distance of each other. This is especially important, because getting too close to the actors may compromise their safety. Likewise, robots may need to avoid other people in the environment and inanimate obstacles, whether static or dynamic. While the filming task may require the drones to operate near people, the minimum allowable distance should increase with the speed of the drone and the speed of the person or dynamic obstacle. This is important because the stopping distance of the robot increases and the uncertainty of the person's trajectory is greater with higher speed. Giving a wider berth at high speeds also allows the drone more space and time in which to react if a person changes direction or speed suddenly.

Within our planning horizon, we also want to consider if the trajectories of obstacles come near those of the UAVs. If the trajectory of the obstacle does not pass or cross the trajectories of the actor/UAV team, i.e. the obstacle is behind the UAVs and traveling in the opposite direction, then it poses a lower risk than if it were in front of the actor/UAV team and traveling toward them, regardless of speed. In this way, based on the relative risk to the actor/UAV team, we can determine how large of a distance the drones need to keep away in order to maintain safety.

In our system, the drones plan their trajectories and execute them autonomously. Safety pilots also form a component of our fail safe system. In the future, we will also explore autonomous fail safe mechanisms, because manual takeovers in a multi-robot team could require significant coordination or to enable continued operation of the team when one drone cannot proceed.

C. Formulation

Let $\xi_{qi} : [0, t_f] \rightarrow \mathbb{R}^3 \times SO(2)$ be the trajectory of the i-th UAV, i.e., $\xi_{qi}(t) = \{x(t), y(t), z(t), \psi_q(t)\}$, and $\Xi = \{\xi_{q1}, \ldots, \xi_{qn}\}$ be the set of trajectories from n UAVs. Let $\xi_a : [0, t_f] \rightarrow R^3$ be the trajectory of the actor, i.e., $\xi_a(t) = \{x(t), y(t), z(t)\}$, which is inferred using onboard cameras. Let grid $\mathcal{G} : \mathbb{R}^3 \rightarrow \mathcal{R}$ be a voxel occupancy grid that maps every point in space to a probability of occupancy. Let $\mathcal{M}(\mathcal{G}) : \mathbb{R}^3 \rightarrow \mathbb{R}$ be the signed distance values of a point to the nearest obstacle.

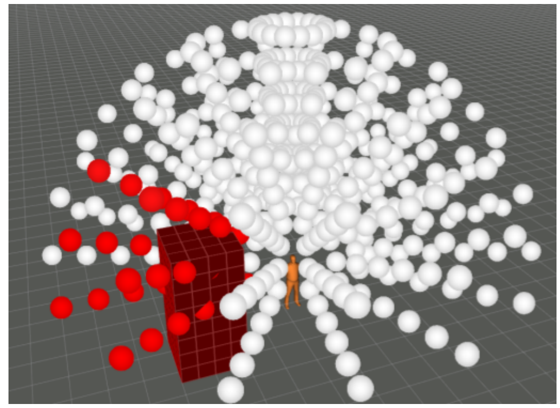

Fig. 1. Visualization of occupancy and occlusion avoidance costs in spherical grid $\mathcal{G}_t^s$, from [11]

We transform the environment's occupancy grid into a time-dependent spherical domain centered around the target $\mathcal{G} \rightarrow \mathcal{G}_t^s \in [0, 1]$, as shown in Fig. 1. For each obstacle, we weight the risk based on its position and speed relative to the UAV team. by applying a multiplier to its position in the occupancy grid, based on the proximity of the obstacle. This gives us the resulting cost function:

$$J_{obstacle}(\Xi) = \sum_{t=1}^{T} \sum_{i=1}^{n} \int_0^{r_{\max}} \mathcal{W}_t^o \mathcal{G}_t^s(\xi_{qi}(t)) d(\text{volume}) \quad (2)$$

We define $\mathcal{W}_t^o : \mathbb{R}^2 \rightarrow \mathbb{R}$, $\mathcal{W}_t^o = |v_o - v_a| + m(\mathbb{1}_{pos})$ where v_a is the current velocity of the actor(s), v_o is the current velocity of the obstacle relative to the speed of the actor, $\mathbb{1}_{pos}$ is an indicator variable for the relative position of the obstacle to the actor, and m is our proximity risk multiplier. Based on this objective, each drone will choose the appropriate position in the occupancy grid projected out for each time step that optimizes for both safety and filming quality.

IV. Future Work

We will quantitatively evaluate our proposed system in a photo-realistic simulation environment and in real world experiments to ensure that it is robust. We will evaluate it in a number of scenarios, from cul de sacs to narrow passageways, with a variety of static and dynamic obstacles. We will also test this with a varying number of actors being filmed for 3D reconstruction.

Acknowledgements

This work is supported by the National Science Foundation under grant no. 2024173.

References

[1] R. Bonatti, W. Wang, C. Ho, A. Ahuja, M. Gschwindt, E. Camci, E. Kayacan, S. Choudhury, and S. Scherer, "Autonomous aerial cinematography in unstructured environments with learned artistic decision-making," *Journal of Field Robotics*, vol. 37, no. 4, pp. 606–641, 2020.

[2] R. Bonatti, C. Ho, W. Wang, S. Choudhury, and S. Scherer, "Towards a robust aerial cinematography platform: Localizing and tracking moving targets in unstructured environments," in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2019, pp. 229–236.

[3] R. Bonatti, Y. Zhang, S. Choudhury, W. Wang, and S. Scherer, “Autonomous drone cinematographer: Using artistic principles to create smooth, safe, occlusion-free trajectories for aerial filming,” in *International Symposium on Experimental Robotics*. Springer, 2018, pp. 119–129.

[4] “Skydio,” https://www.skydio.com/technology, May 2022.

[5] “Dji mavic,” https://www.dji.com/br/mavic, May 2022.

[6] A. Torres-González, A. Alcántara, V. Sampaio, J. Capitán, B. Guerreiro, R. Cunha, and A. Ollero, “Distributed mission execution for aerial cinematography with multiple drones,” 2019.

[7] A. Alcántara, J. Capitán, A. Torres-González, R. Cunha, and A. Ollero, “Autonomous execution of cinematographic shots with multiple drones,” *IEEE Access*, vol. 8, pp. 201 300–201 316, 2020.

[8] A. Alcántara, J. Capitán, R. Cunha, and A. Ollero, “Optimal trajectory planning for cinematography with multiple unmanned aerial vehicles,” *Robotics and Autonomous Systems*, vol. 140, p. 103778, 2021.

[9] L.-E. Caraballo, Á. Montes-Romero, J.-M. Díaz-Báñez, J. Capitán, A. Torres-González, and A. Ollero, “Autonomous planning for multiple aerial cinematographers,” in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2020, pp. 1509–1515.

[10] I. Karakostas, I. Mademlis, N. Nikolaidis, and I. Pitas, “Shot type constraints in UAV cinematography for autonomous target tracking,” *Information Sciences*, vol. 506, pp. 273–294, 2020.

[11] A. Bucker, R. Bonatti, and S. Scherer, “Do you see what i see? coordinating multiple aerial cameras for robot cinematography,” in *2021 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2021, pp. 7972–7979.

[12] T. Nägeli, L. Meier, A. Domahidi, J. Alonso-Mora, and O. Hilliges, “Real-time planning for automated multi-view drone cinematography,” *ACM Transactions on Graphics (TOG)*, vol. 36, no. 4, pp. 1–10, 2017.

[13] X. Xie, C. Zhang, Y. Zhu, Y. N. Wu, and S.-C. Zhu, “Congestion-aware multi-agent trajectory prediction for collision avoidance,” in *2021 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2021, pp. 13 693–13 700.

[14] B. Lindqvist, P. Sopasakis, and G. Nikolakopoulos, “A scalable distributed collision avoidance scheme for multi-agent UAV systems,” in *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2021, pp. 9212–9218.

[15] J. Tordesillas and J. P. How, “MADER: Trajectory planner in multiagent and dynamic environments,” *IEEE Transactions on Robotics*, 2021.

[16] S. Arora, S. Choudhury, D. Althoff, and S. Scherer, “Emergency maneuver library - ensuring safe navigation in partially known environments,” in *2015 IEEE International Conference on Robotics and Automation (ICRA)*, 2015, pp. 6431–6438.

[17] C. Ho, A. Jong, H. Freeman, R. Rao, R. Bonatti, and S. Scherer, “3D human reconstruction in the wild with collaborative aerial cameras,” in *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2021, pp. 5263–5269.