Proceedings of the 2016 Winter Simulation Conference
T. M. K. Roeder, P. I. Frazier, R. Szechtman, E. Zhou, T. Huschka, and S. E. Chick, eds.

EFFECTIVE VISUAL SURVEILLANCE OF HUMAN CROWDS USING COOPERATIVE UNMANNED VEHICLES

Sara Minaeian

Systems and Industrial Engineering, The University of Arizona
1127 E James E. Rogers Way
Tucson, AZ 85721, USA

ABSTRACT

The goal of this work is to propose an effective and efficient visual surveillance system for detection, geo-localization, and data association of moving human crowds using teams of unmanned aerial vehicles (UAVs) and unmanned ground vehicles (UGVs) in a border patrol application. Such complex system suffers from various emerging challenges such as: heterogeneous dynamic data, non-rigid target shapes, dynamic background due to moving sensors, and occlusion. Therefore, different fidelity levels have been considered in this work for UAVs and UGVs based on their different characteristics, and a number of related computer vision algorithms have been proposed based on the dynamic data driven application system (DDDAS) paradigm. Moreover, a testbed involving real hardware (UAVs, UGVs, and cameras) and an agent-based simulation model is developed to verify, validate, and demonstrate the system. The experimental results reveal the effectiveness of the proposed approaches for visual surveillance of human crowds by unmanned vehicles.

1 INTRODUCTION

In recent years, unmanned vehicles (UVs) are being increasingly used in autonomous surveillance. Detecting, geo-localizing, and multi-target tracking of targets are among the most critical objectives of UVs in a visual surveillance system. Although visual surveillance of human crowd in dynamic scenes have recently attracted increasing attention of researchers, there are a number of unsolved problems in such field. Since the visual sensors move with the UVs, and thus the targets and the environment are dynamic, it adds to the complexity and uncertainty of the video processing. Moreover, the limited onboard computation resources require more efficient and effective algorithms to be proposed. The goal of this work is to design and develop an effective and efficient visual surveillance system based on DDDAS paradigm to be used by the UVs for the application of real-time autonomous crowd control and border patrol. In particular, different vision algorithms are proposed and applied in this context, including: 1) a new moving object detection method to detect the crowd from the moving camera onboard UAV, 2) a human classification method to recognize the individuals using the camera onboard UGV, 3) a new moving-landmark-based method for geo-localization of crowd from the UAV, 4) a heuristic method based on triangulation for target geo-location from the UGV, and 5) a multi-target data association method for differentiating detected targets over successive frames. In order to implement and demonstrate this system, a cooperative team of one UAV and multiple UGVs with onboard visual sensors is considered, to take advantage of their complementary characteristics. The proposed methodology and experiments are discussed below.

2 CONTRIBUTIONS IN METHODOLOGY

As the first vision-based module for a human surveillance system, human detection techniques are categorized into two group of modules: 1) motion detection, to detect and segment targets in motion and 2) classification, to recognize the detected targets belonging to the human class based on their features. In this

work, UAV onboard sensor with lower resolution is used for moving crowd detection, while higher resolution visual sensors on the UGVs are utilized for higher fidelity human classification. Considering the challenge of moving background due to the moving camera onboard UAV, the proposed motion detection algorithm combines various computer vision techniques over a temporal sliding window, to subtract the moving background regions without using a prior model. Moreover, a classification module based on histogram of oriented gradients (HOG) is used by the UGVs as described in (Minaeian, Liu, and Son 2015). In this algorithm, parameters are adjusted to improve the human detection effectiveness and efficiency, considering the real-time requirements.

As the next milestone, UVs need the real-time geographic location of the detected target, in order to create the same reference coordinate system for higher-level decision-making. To this end, a moving-landmark-based algorithm is proposed for geo-localizing the detected crowd from the UAV (Minaeian, Liu, and Son 2016). In this work, the cooperative UGVs with known GPS locations and fewer sources of estimation error compared to the UAV are considered as independently moving landmarks. Based on these landmarks, an efficient perspective transformation of the target position is estimated in real-time. Moreover, a heuristic triangulation method is proposed to estimate the 2-D geo-locations of detected targets, based on UGVs camera pose. These estimated geo-locations are then used in generating human agents in our simulation testbed (Minaeian et al. 2015) for improving the tracking performance.

Another potential problem in this context is to differentiate between various detected targets, so that we can track them across the temporal frames. Multi-target tracking introduces a higher level complexity to the problem compared to the single-target tracking, as it includes the *data association*. In this work, a novel approach for measurement-to-track data association is proposed, so that newly detected targets (measurements) at each time stamp can be associated to previous detections (tracks). In our proposed approach, we aim at applying the more appropriate data association affinity score (e.g. appearance or motion information) dynamically based on DDDAS concept, and considering the detection results.

3 EXPERIMENTS USING SIMULATION-BASED TESTBED

In order to demonstrate the proposed system and algorithms, a testbed involving assembled hardware (UAVs, UGVs, and sensors) and agent-based simulation model using Repast Simphony® is developed. The reasons to use agent-based model for designing experimental scenarios, are that: 1) it is less expensive and tedious to run simulation-based experiments compared to the sole-hardware-based real testbed in a border area, and 2) it is more convenient to verify estimated locations (without the GPS sensors) and run experiments with changing parameters. The sensitivity analyses were conducted for a series of parameters, including: 1) flight altitude, 2) crowd's velocity, 3) number of landmarks, and 4) landmark assignment method. The preliminary experimental results demonstrate the effectiveness of the proposed approach. More specifically, we achieved better geo-localization results (2.2 m to 3.5 m average Euclidean error) compared to the relevant literature (up to 30 m).

As future work we will develop and use a physics-based simulation model, so that the proposed computer vision algorithms can be validated, using the realistic models of sensors, vehicles, and targets.

REFERENCES

Minaeian, S., J. Liu, and Y.-J. Son. 2015. "Crowd Detection and Localization Using a Team of Cooperative UAV/UGVs." In *Proceedings of the 2015 IIE Annual Conference*. edited by S. Cetinkaya and J. K. Ryan, 595-604. Nashville, Tennessee: Institute of Industrial Engineers.

Minaeian, S., J. Liu, and Y.-J. Son. 2016. "Vision-Based Target Detection and Localization via a Team of Cooperative UAV and UGVs." *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 46, No. 7: 1005-1016.

Minaeian, S., Y. Yuan, J. Liu, and Y.-J. Son. 2015. "Human-in-the-Loop Agent-Based Simulation for Improved Autonomous Surveillance using Unmanned Vehicles." In *Proceedings of the 2015 Winter Simulation Conference*, edited by L. Yilmaz, W. K. V. Chan, I. Moon, T. M. K. Roeder, C. Macal, and M. D. Rossetti, 3126-3127. Piscataway, New Jersey: IEEE, Inc.