Decision-Theoretic Approach for Controlling and Coordinating multiple Active Cameras in Surveillance (Extended Abstract)

Prabhu Natarajan
School of Computing, National University of Singapore, Singapore.
prabhu@comp.nus.edu.sg

ABSTRACT

This research presents a novel decision-theoretic approach to control and coordinate multiple active (pan-tilt-zoom) cameras to achieve high-quality surveillance. The decision-theoretic approaches provide robust mathematical frameworks that can model the interactions between the active camera network and the surveillance environment. It offers advantage of planning optimal control decisions for active cameras to achieve the desired surveillance task, in presence of uncertainties like targets' motion, location, etc. In this work, we provide an overview of proposed framework in which a surveillance tasks can be posed as stochastic optimization problems. Specifically we propose solutions for two novel problems/trade-offs in active camera surveillance: (i) Maximizing the number of targets observed in active cameras while maintaining the guaranteed resolution of these targets and (ii) Maximizing the number of targets observed in active cameras while minimizing the location uncertainty of the targets that are currently not observed in any of the active cameras. By exploiting the structure and properties that are inherently present in the surveillance problem, we have reduced the exponential policy computation time to polynomial time. This makes the decision-theoretic approach feasible for the surveillance applications.

Categories and Subject Descriptors

H.4 [**Information Systems Applications**]: Miscellaneous; I.4.8 [**Scene Analysis**]: tracking; I.2.9 [**Robotics**]: Commercial robots and applications, Sensors

Keywords

Surveillance and security, Robot coordination, Robot teams and multi-robot systems

1. INTRODUCTION

The use of active cameras are becoming popular in surveillance systems due to the improvements in smart camera technologies. These cameras are endowed with pan-tilt-zoom capabilities which can be exploited to: (i) provide high-resolution images of targets for bio-metric analysis like face recognition, gait analysis, etc., and (ii) dynamically configure the cameras to focus on sporadic region of interests. In order to achieve effective real-time surveillance, an efficient collaborative mechanism is required to control and coordinate these cameras' actions which is the primary focus of this research. Fig. 1 visualizes our grand vision of this thesis work.

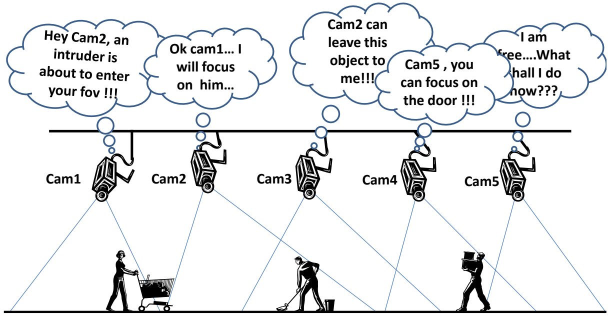

Figure 1: Multiple active camera coordination and control.

Controlling active cameras is a challenging and difficult task because: (i) there are inherent uncertainties in surveillance environment due to targets' motion, location, etc., (ii) the number of targets is much greater than the available number of cameras, and (iii) the camera coordination framework should be scalable to increasing number of targets. We address the above issues by proposing a novel decision-theoretic multi-agent planning framework for controlling multiple active cameras. To the best of our knowledge, we are the first to propose a decision-theoretic framework for controlling and coordinating network of active cameras. Decision-theoretic approach models the interaction between active camera network and the environment effectively. Some of the advantages it offers in surveillance are: (i) provides mathematical models for planning optimal control decisions in presence of uncertainties, (ii) multiple high-level application goals can be modeled as mathematical objective functions, and (iii) scalability can be achieved by exploiting the problem structure. Specifically, we account for uncertainty due to targets' motion via probabilistic motion model of targets and location uncertainty via probabilistic observation model of cameras. Scalability can be achieved by exploiting properties that are inherently present in the surveillance problems.

There are few notable works [1, 3, 4], etc. that have been carried out in the literature in controlling active cameras. These works are based on some heuristic approaches which are tailored specifically to their own objectives and cannot be modified to achieve other objectives. These previous works have not accounted for the uncertainty of targets' motion and location in controlling the cameras. In contrast, our approach is a generic mathematical framework in which different high-level surveillance goals can be modeled as low-level mathematical objective functions. All the previous works have serious drawbacks in scalability of number of targets, where as we exploit the properties of the surveillance problem and achieve scalability of upto 50 targets in real-time [2]. This paper provides an overview of the decision-theoretic framework for

Appears in: *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2012)*, Conitzer, Winikoff, Padgham, and van der Hoek (eds.), June, 4–8, 2012, Valencia, Spain.

controlling active cameras in surveillance applications (sec. 2) and potential surveillance problems/trade-offs that we intend to solve (sec. 3) as part of our thesis. Finally we discus how our thesis can contribute to multi-agent community with future directions of our research work (sec. 4).

2. PROPOSED FRAMEWORK

The proposed multi-agent framework consists of n active cameras and m targets such that $n \ll m$. We have $n \ll m$ because, the problem becomes more challenging and realistic when there are more targets that are to be observed by few active cameras. The state of the system consists of the state of active cameras and the targets: $\mathcal{S} = \mathcal{C} \times \mathcal{T}$. The state of active cameras is given by set $\mathcal{C}$ of discretized pan/tilt/zoom values. The state of target is represented by set of locations, directions and velocities: $\mathcal{T} = \mathcal{T}_l \times \mathcal{T}_d \times \mathcal{T}_v$. The physical state of the surveillance system has been discretized into grid cells, such that the center of the cell represents the target's location. The set of actions $\mathcal{A}$ of active cameras are the respective pan/tilt/zoom commands to move the camera to the specified state. Since the targets are non-deterministic, the motion of the targets are not affected by the actions. The transition model of the system has been factorized into transition model of cameras and targets based on the conditional independence assumption: transition model of cameras is conditionally independent of transition model of targets. The transition model of cameras is deterministic as these cameras can move to their respective states accurately. The transition model of a target is given by velocity-direction based motion model with velocity and direction distributed normally. The observation model of the cameras incorporates the noisy measurement of targets' location determined by the underlying view-geometry of respective cameras. The objective functions are real-valued functions that represent the high-level surveillance goals. For example, observing m targets in active cameras with predefined image resolutions can be defined as follows:

$$R((T,C)) = \sum_{i=1}^{m} \widetilde{R}((t_i,C))$$

$$\widetilde{R}((t_i,C)) = \begin{cases} 1 & \text{if location of target } t_i \text{ lies in } fov(C), \\ 0 & \text{otherwise;} \end{cases}$$

where T and C represents the joint state of targets and cameras, t_i represents the state of i^{th} targets and $fov(C)$ represents the set of target's locations that are observed by the active cameras in a predefined resolution. We compute optimal policy $\pi^*(S) \rightarrow A$, such that the action $A \in \mathcal{A}$ associated with the policy maximizes the expected reward for the given state $S \in \mathcal{S}$.

3. PROBLEM DESCRIPTION

This thesis focuses on solving interesting and challenging non-trivial trade-offs in active camera surveillance:

- Maximizing the number of targets observed in active cameras while maintaining the guaranteed resolution of these targets.
- Maximizing the number of targets observed in active cameras while minimizing the location uncertainty of these targets.

The major challenges in addressing the above trade-offs are the real-time policy computation and scalability to increasing number of targets. The first trade-off is due to the fact that increasing the resolution of observing some targets through panning, tilting, or zooming may result in the loss of other targets being tracked in high-resolution. We had addressed this trade-off in our current work in [2] by developing a Markov Decision Process (MDP) framework in which the targets' state is fully observable by having static cameras. These static cameras obtains approximate 3D-locations of targets (that are rounded-off to nearest discretized grid cell), which is then used by transition function to predict next locations of targets. But, static cameras cannot observe targets in high-resolution, which is our primary goal. So the active cameras are controlled by MDP to maximize the expected number of observed targets in high-resolution, based on predicted locations of targets (i.e. distribution of locations of targets in next-time step). We have exploited the conditional independence property in transition model and linearity property in reward function to reduce the exponential computation time to polynomial time for greedy policy. We were able to scale up to 50 targets in simulation with on-line policy computation time of $\mathcal{O}(|\mathcal{A}|m)$. We encourage the reader to look into our real-camera demo video[1].

The second trade-off is due to the fact that observing some of the targets in active cameras will reduce the location uncertainty of these observed targets, while increasing the location uncertainty of the targets that are currently not observed by these cameras. This is our on-going work where we are developing a Partially Observable Markov Decision Process (POMDP) based planning framework where the targets' states are partially observable, i.e. we do not have static cameras that can observe the targets' location and hence targets' state are tracked in the belief space based on the observations from the active cameras. This model provides a realistic camera controlling framework that can account for both targets' motion uncertainty and location uncertainty.

4. FUTURE WORKS AND CONCLUSION

We have developed a novel decision-theoretic planning framework for controlling active cameras in surveillance. We have addressed first trade-off in our current work [2] using MDP model and currently developing a POMDP model to address the second trade-off. Currently we are working on centralized architecture which need to be expanded for very large active camera networks. In future we intend to: (i) explore on decentralized multi-agent coordination approaches for controlling large network of active cameras and (ii) develop an automatic way to choose respective objective function to control the cameras based on the external surveillance situation. Through this work in multi-camera surveillance, we hope to (i) contribute to the theory of multi-agent system in terms of key concepts like decentralized coordination, efficient planing, factored MDP/POMDP model, etc. and also (ii) improve the practical significance and public awareness of multi-agent system research.

5. REFERENCES

[1] N. Krahnstoever, T. Yu, S.-N. Lim, K. Patwardhan, and P. Tu. Collaborative Real-Time Control of Active Cameras in Large Scale Surveillance Systems. In *Proc. M2SFA2*, 2008.

[2] P. Natarajan, T. N. Hoang, K. H. Low, and M. Kankanhalli. Decision-theoretic approach to maximizing observation of multiple targets in multi-camera surveillance. In *Proc. AAMAS*, 2012.

[3] F. Qureshi and D. Terzopoulos. Planning ahead for PTZ camera assignment and handoff. In *Proc. ICDSC*, 2009.

[4] E. Sommerlade and I. Reid. Probabilistic surveillance with multiple active cameras. In *Proc. ICRA*, 2010.

---

[1] `http://www.youtube.com/watch?v=M_g4q40KF-0`