Multi-tasking Smart Cameras for Intelligent Video Surveillance Systems

Wiktor Starzyk
Faculty of Science
University of Ontario Institute of Technology
Oshawa, ON, L1H 7K4, Canada
wiktor.starzyk@mycampus.uoit.ca

Faisal Z. Qureshi
http://faculty.uoit.ca/qureshi

Abstract

We demonstrate a video surveillance system—comprising passive and active pan/tilt/zoom (PTZ) cameras—that intelligently responds to scene complexity, automatically capturing higher resolution video when there are fewer people in the scene and capturing lower resolution video as the number of pedestrians present in the scene increases. To this end, we have developed behavior based-controllers for passive and active cameras, enabling these cameras to carry out multiple observation tasks simultaneously. The research presented herein is a step towards video surveillance systems—consisting of a heterogeneous set of sensors—that provide persistent coverage of large spaces, while optimizing surveillance data collection by tuning the sensing parameters of individual sensors (in a distributed manner) in response to scene activity.

1. Introduction

Automatic video surveillance systems available today typically comprise passive wide field-of-view (FOV) CCTV cameras. While these systems excel at pedestrian detection and tracking, these systems are unable to capture individual pedestrians at resolutions required to support subsequent biometric analysis. Consequently video surveillance systems comprising both active pan/tilt/zoom (PTZ) and passive wide-FOV cameras have been proposed to provide higher resolution video coverage of large spaces. Here, passive cameras are responsible for frame-to-frame pedestrian tracking (and in some cases, localization via stereo analysis); where as, active cameras are tasked with capturing closeup videos of objects for further biometric analysis (*e.g.*, facial recognition, gait analysis, etc.). Often there will be more pedestrians in the scene than the number of available PTZ cameras, and the automatic management of PTZ cameras becomes non-trivial. Operator monitoring and control is clearly infeasible as the number of cameras or scene complexity increases. It is, therefore, desirable to develop automated video surveillance systems capable of providing higher-resolution, persistent video coverage of large spaces. By necessity such systems need to be self-managed and will contain both passive wide-FOV and active PTZ cameras.

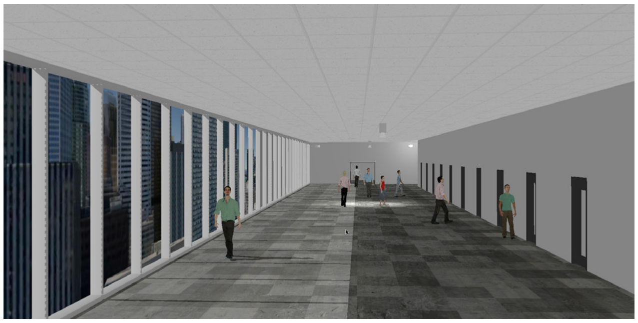


Figure 1. Our virtual vision simulator consists of a 60 meters by 10 meters office hallway populated with pedestrians that can be easily scripted to move along predefined paths. A state-of-the-art rendering pipeline [1] supports real-time shadows, lighting effects, and a myriad of lens artifacts.

In this paper, we present a novel video surveillance system comprising passive and active cameras. The proposed system automatically tune sensing parameters of its PTZ cameras in response to the scene activity, choosing to capture close-up video when the number of pedestrians present in the scene is low and electing to capture lower-resolution video as the number of pedestrians increase, thus always keeping every pedestrian in view. A noteworthy feature of the proposed system is its smart camera nodes—both passive and active, modelled as behavior-based agents—which can intelligently carry out multiple observation tasks simultaneously. These cameras enable the video surveillance system described herein to intelligently respond to scene complexity, automatically capturing close-up imagery of the pedestrians present in the scene when possible and behaving as wide-FOV cameras as the number of pedestrians increase.

Most other schemes for PTZ camera control assume that a PTZ camera can only carry out a single task at any given time, which is restrictive and may lead to catastrophic observation failures. Consider, for example, the simple case of two PTZ cameras swapping two pedestrians. Camera selection and handoff schemes that assume that each camera can only carry out a single observation task simply cannot switch the roles of the two cameras without interrupting one of the observation tasks. A PTZ camera that can observe multiple pedestrians simultaneously on the other hand can gracefully deal with situations such as these.

1.1. Contributions and Overview

The contributions of the research presented herein are twofold: 1) we demonstrate a video surveillance system comprising a heterogeneous set of cameras capable of tuning its performance (in a distributed fashion) in response to scene activity while at the same time providing persistent video coverage of the scene and 2) we develop behavior-based camera controllers that enable our (passive and active) cameras to carry out multiple observation tasks simultaneously.

Adhering to the Virtual Vision paradigm advocated by Qureshi and Terzopoulos [12], we demonstrate the proposed video surveillance system by deploying virtual cameras in a realistic 3D environment representing hallway of an office building, complete with windows, outdoor scenery, shadows, lights, and pedestrians (Fig. 1). The video surveillance system demonstrated in this paper builds upon the camera network model presented in [10]. Unlike there, however, in this paper we focus on the video surveillance system setup and behavior-based PTZ cameras.

The rest of the paper is organized as follows. We briefly discuss the related work in the next section. Sec. 3 develop behavior-based camera controllers, presenting a unified control framework for passive and active cameras. We introduce the negotiation protocol that sets up camera collaborations and solves the problem of camera assignment in Sec. 4. We demonstrate the proposed video surveillance system in Sec. 5. Sec. 6 concludes the paper with a discussion and possible future directions.

2. Related Work

Several authors (e.g., [4, 5, 7]) have studied multi-camera issues related to low-level sensing, distributed inference, and tracking. Recently, however, the research community has been paying increasing attention to the problem of controlling or scheduling active cameras in order to capture high-resolution imagery of interesting events. High-resolution imagery not only allows for subsequent biometric analysis, it also helps increase the situational awareness of the surveillance system operators. In a typical setup, information gathered by stationary wide-FOV cameras is used to control one or more active cameras [6, 11, 8]. Generally speaking, the cameras are assumed to be calibrated and the total coverage of the cameras is restricted to the FOV of the stationary camera. Nearly all PTZ scheduling schemes rely on site-wide multi-target, multi-camera tracking.

The problems of camera assignment and handoff have mainly been studied in the context of smart camera networks [15, 12, 9]. Qureshi and Terzopoulos propose a strategy for proactive PTZ camera control strategy. Here, long-term consequences of camera assignments are taken into account when determining how best to carry out a camera handoff task [13]. Singh *et al.* develop a coopetitive framework for multicamera assignment [14]. There framework relies upon competition among and cooperation between the available cameras when performing camera assignments. It assumes that at any given instance there are more cameras than the active tasks. A good place to start investigating behavior-based control is [2].

3. Behavior-based Camera Controllers

We treat each camera as a behavior-based autonomous agent. The overall behavior of the camera is determined by the vision routines and the current task. Since visual processing is not the primary theme of this work, this section focuses on behavior-based modeling of multi-tasking passive and active cameras.[1] Each camera has a repertoire of behaviors, from simple (*e.g.* switching on/off the sensing subsystem) to exceedingly complex (*e.g.* perform a visual search to locate a pedestrian). Fig. 2 shows behavior routines available to active and passive cameras. We take the layered approach to behavior design, which was first popularized by the *Subsumption* architecture [3]. Here, lower layer behaviors are used to build successively more complex and competent behaviors.

Camera controller, which is responsible for behavior selection and arbitration, is modeled as an *infinite state machine*. The state of a camera represents its current activities. For example, a camera may be tracking one individual and searching for another person at the same time. The set of activities that a camera can engage in are:

- $idle(c_i)$: The camera c_i is currently not performing any observation task.
- $observing(c_i, h_j)$: The camera c_i is currently observing pedestrian h_j.
- $evaluating(c_i, h_j)$: The camera c_i is currently evaluating its suitability to observe pedestrian h_j. A camera cannot take part in any negotiations involving h_j without first knowing its suitability to that task; suitability

[1] We assume vision processing performance similar to that found in a typical surveillance system, and that we have implemented appearance based pedestrian trackers that work for both passive and active cameras.

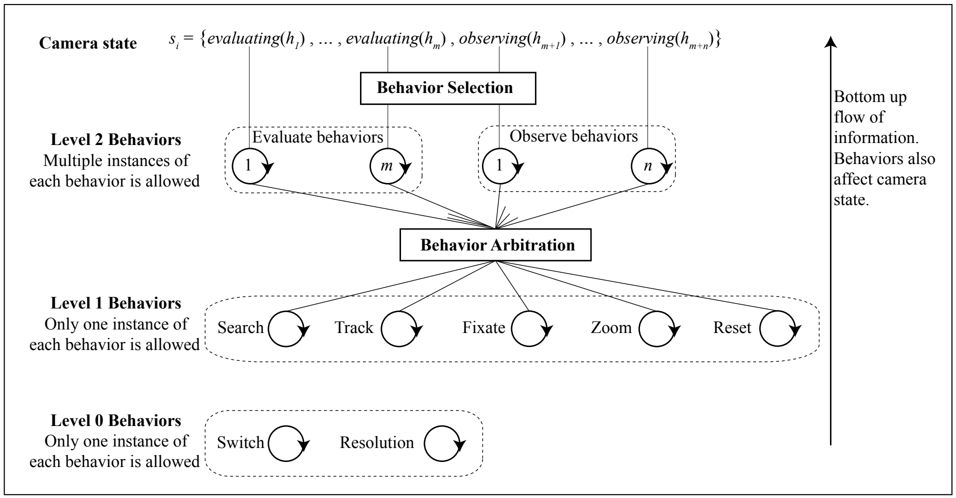

Camera state
$s_i = \{evaluating(h_1), \ldots, evaluating(h_m), observing(h_{m+1}), \ldots, observing(h_{m+n})\}$
Behavior Selection
Level 2 Behaviors
Multiple instances of each behavior is allowed
Evaluate behaviors
1
m
Observe behaviors
1
n
Behavior Arbitration
Level 1 Behaviors
Only one instance of each behavior is allowed
Search
Track
Fixate
Zoom
Reset
Level 0 Behaviors
Only one instance of each behavior is allowed
Switch
Resolution
Bottom up flow of information. Behaviors also affect camera state.

Figure 2. Behavior-based camera controller. The camera controller is responsible for both behavior selection and arbitration. Fixate, Zoom, and Reset behaviors are only available for active PTZ cameras. Level 0 and 1 behaviors implement a priority based arbitration when priority information is available. Behaviors in the higher layers are composed of those found in the lower layers.

encodes the success probability of a camera with respect to an observation task. Typically, a camera evaluates its suitability to an observation task when it receives the task request from a neighbouring camera.

Given this list of activities, we can define the activity set $\mathcal{A}_i$ for a camera c_i as

$$\mathcal{A}_i = \{idle, evaluating(h_j), observing(h_j) | h_j \in \mathcal{H}\},$$

where $\mathcal{H}$ is the set of pedestrians present in the scene. Let s_i represent the state of camera c_i then $s_i \in \mathcal{S}_i$, where

$$\mathcal{S}_i = \mathcal{P}(\mathcal{A}_i) - \Phi - \mathcal{P}^-(\mathcal{A}_i).$$

$\mathcal{P}(\mathcal{A}_i)$ is the powerset of $\mathcal{A}_i$, Φ represents the empty set, and $\mathcal{P}^-(\mathcal{A}_i)$ consists of logically invalid states. For example, a camera can not be both idle and observing a pedestrian at the same time. Specifically, Let s_i^- denote a logically invalid state for camera c_i then $s_i^- \in \mathcal{P}(\mathcal{A}_i) \cap \mathcal{P}^-(\mathcal{A}_i)$ and either of the following two conditions hold:

Condition 1: $idle(c_i) \in s_i^-$ and $|s_i^-| \geq 2$

Condition 2: $\exists u(h'), v(h'') \in s_i^-$ such that $h' = h''$, where $u, v \in \{evaluating(h_j), observing(h_j)\}$ and $h', h'' \in \mathcal{H}$.

The first condition excludes the states that show a camera as both idle and busy; where as, the second condition excludes the states that show a camera simultaneously observing and evaluating the same individual.

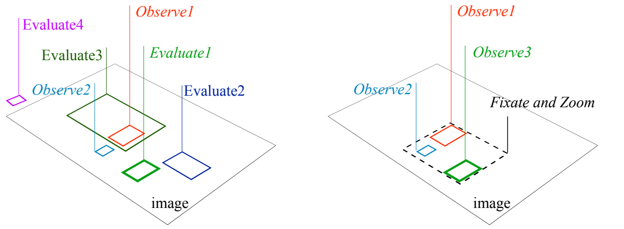
Each camera activity is mapped to a corresponding level 2 behavior. These behaviors sit between discrete activities and continuous operations and presents a well-behaved abstracted view of the underlying continuous reality to the camera controller. Camera controller uses state information, which contains camera activities, to activate/deactivate appropriate level 2 behaviors as necessary. Notice that the camera controller does not distinguish between passive wide-FOV and active PTZ cameras, since both expose the same level 2 behaviors: Observe and Evaluate.

3.1. Behavior Arbitration

As is shown in Fig. 2 multiple instances of level 2 behaviors—Observe and Evaluate—may be active simultaneously reflecting the multitasking nature of our cameras. Every instance of these behaviors, however, relies upon level 1 behaviors: Search, Track, Fixate, Zoom, and Reset (the last three behaviors are only available for active PTZ cameras). Since level 1 behaviors are singletons, instances of behaviors Observe and Evaluate must share level 1 behaviors, which suggests the need for behavior arbitration. Behavior arbitration aims at avoiding negative consequences of behavior interactions—*e.g.*, any one behavior taking over camera resources and locking out all other behaviors, thus leading to observation failures. Behavior arbitration also addresses the problem of *behavior dithering*, where a camera will constantly switch between two (or more) competing tasks.

Passive Cameras: Behavior arbitration is relatively straightforward in the case of passive cameras. Here, a contention arises when different instances of level 2 behaviors cannot agree on the resolution for video capture. We solve this issue by capturing the video at the highest requested resolution.

Active PTZ Cameras: The ability of an active PTZ camera to carry out multiple observation tasks presupposes that competing behaviors are able to work together in an intelligent, goal-driven manner. On the one hand, a PTZ camera



(a) 2 active, 4 potential tasks

(b) Behavior arbitrator activates another task and returns a new ROI to fixate and zoom behaviors

Figure 3. ROIs defined over the image coordinate system serve as the common currency used to arbitrate between multiple competing behaviors: (a) currently active level 2 behaviors express their ROI preference to the behavior arbitrator and (b) final ROI shown as a dotted black rectangle is sent to level 1 zoom/fixate behavior. Notice that ROI for Evaluate4 behavior sits outside the image boundaries indicating a request to perform a pan search in order to acquire pedestrian 4 (a).

can acquire closeup video of a single pedestrian, while on the other hand, it can track multiple pedestrians simultaneously at a much lower resolution by using a wide angle setting. In our case, a PTZ camera might be carrying out multiple observation tasks simultaneously, so it must be able to intelligently choose between the two extremes: observe a single pedestrian at a higher resolution or observe multiple individuals at lower resolutions.

Fixate and Zoom are image-driven behaviors that strive to maintain an imaginary Region of Interest (ROI) within the field of view. The ROI is defined within the image coordinate system. Specifically, the fixation routine aims to keep ROI in the center of image, while the goal of the zoom routine is to maintain the desired pixel coverage for ROI. When a PTZ camera is acquiring a closeup video of a single pedestrian, ROI is simply the bounding box for that pedestrian as estimated by the pedestrian tracker. For more complex scenarios—*e.g.*, when a PTZ camera is observing/evaluating multiple pedestrians (Fig. 3)—each level 2 behavior sends its desired ROI to the behavior arbitration module. The behavior arbitration routine picks the largest subset of ROIs that meets task requirements (*e.g.* minimum resolution) and has the maximum return (utility). The ROI sent to level 1 behaviors is the smallest region enclosing the selected ROIs (indicated as the dashed rectangle in Fig. 3(b)).

Fixate and Zoom behaviors are modeled as proportional–integral–derivative controllers [12]. These behaviors operate independently of each other, each trying to achieve its respective goals. Consequently, an active PTZ camera can simultaneously zoom and fixate on the ROI computed by the behavior arbitration routine. When a task failure is detected, the camera controller goes into a recovery mode where the fixation routine is deactivated and the camera begins to increase its FOV setting with the aim to keep the pedestrian(s) within the field of view. Visual search is performed to reacquire the pedestrian using the stored pedestrian signature. If unsuccessful, the camera reports a failure and returns to its default state.

A Model for Behavior Arbitration: We now develop a model for behavior arbitration and task selection. Let $\mathcal{R}_a$ be the set of ROIs corresponding to different instances of level 2 observe behavior and $\mathcal{R}_e$ be the set of ROIs corresponding to different instances of level 2 evaluate behavior. $\mathcal{R}_a$ then represents active tasks; whereas, $\mathcal{R}_e$ lists potential tasks. The behavior arbitration routine is responsible for selecting tasks from $\mathcal{R}_e$ to activate and for choosing tasks from $\mathcal{R}_a$ to deactivate. Specifically, if $\mathcal{R}^*$ represents the decision of this arbitration process then

$$\mathcal{R}^* = \underset{\mathcal{R}\subseteq\{\mathcal{R}_a\cap\mathcal{R}_e\}}{\arg\max} \; \mathcal{F}(\mathcal{R}) \text{ subject to } L(\mathcal{R}),$$

where functional $\mathcal{F}(.)$ evaluates the overall "utility" of carrying out a set of tasks simultaneously and $L(\mathcal{R})$ represents the set of constraints. We can, for example, define $\mathcal{F}(\mathcal{R})$ to be the success probability of carrying out the set of tasks $\mathcal{R}$ simultaneously. Then, under the conditional independence assumption,

$$\mathcal{F}(\mathcal{R}) = \prod_{r\in\mathcal{R}} P(r),$$

where $P(r)$ is the success probability of the task corresponding to ROI r. We refer the reader to [12, 8, 15, 9] for discussions about how best to compute $P(r)$. $L(\mathcal{R})$ can be defined in terms of the actions that constitute $\mathcal{R}$, *e.g.*

$$L(\mathcal{R}) \equiv \arg\min_{r\in\mathcal{R}} Resolution(r) > 1000$$

sets minimum acceptable pixel resolution of any task at 1000 pixels. Here, $Resolution(r)$ returns the size of ROI r in pixels.

Task deactivation may occur for the following three reasons: 1) prolonged observation failure, 2) completion, or 3) interference with other active tasks.

4. Persistent Surveillance

The video surveillance system described herein relies upon a simplified version of the negotiation model presented in [10]. Camera assignments and handoffs are carried out via strictly local negotiations between two cameras. Negotiations are task specific and take place between a client camera node, which wishes to establish a collaboration, and one or more neighbouring cameras, called server camera nodes.

For the video surveillance system demonstrated here, client camera initiates task assignment by sending out a **Msg_New_Task(hid)** message to the neighbouring cameras. **hid** encodes information—such as, pedestrian signature, location, etc.—needed to perform a successful visual search to acquire the initial track. The server camera instantiates an Evaluate(**hid**) behavior and attempts to determine whether or not it can successfully observe the pedestrian in question. It returns **Msg_Server(Success hid)** to the client if it is able to carry out the requested observation task; otherwise, the server camera returns **Msg_Server(Failed hid)**. Every camera periodically broadcasts its status, including the observation tasks that it is currently performing, to other cameras in the vicinity through the **Msg_Status(Observe: hid list, Evaluate: hid list)** message.

5. Results

We evaluate our multi-tasking PTZ controller by deploying simulated cameras within a 3D virtual environment, consisting of a 60 meters by 10 meters office hallway, complete with left and right entrances, office doors, floor-to-ceiling panoramic windows showing a photorealistic cityscape, sunlight filtering into the hallway through these windows casting shadows on the floor (see Fig. 1). The 3D environment is populated with self-animating pedestrians that follow scripted paths.

Fig. 4 illustrates the multi-tasking aspect of our PTZ camera nodes. Initially, camera 1 is tracking object 0 and camera 2 is tracking object 1. Camera 2, however, decides to observe object 2 upon object 2's entry. At this point, the PTZ camera is fixating and zooming in on two objects. *Observe* behavior corresponding to each object attempts to pull the camera towards itself; however, the behavior arbitration routine intervenes and arrives at a compromise where both objects are observed, albeit at lower resolution. Later, the two cameras perform a successful switch when object 2 leaves the observational range of camera 2. It is fair to say that without multi-tasking the two PTZ cameras cannot perform the *switching* behavior seen here without interrupting one of the observation tasks. We repeated this experiment 10 times with different initial conditions and camera parameters (pan/tilt and zoom limits), and the cameras successfully carried out the switching task every time. Under extreme situations when the two objects observed by these PTZ cameras are far from each other, the cameras act as wide-FOV cameras.

We simulated a video surveillance system comprising 3 passive and 4 active PTZ cameras in an L configuration (see Fig. 5(a)). First, we scripted two pedestrians to move along the route (shown in the figure as a dotted line) in opposite directions. As expected, both pedestrians were successfully observed throughout their stay in the scene. More importantly, PTZ cameras were able to capture high resolution video for these pedestrians as seen from the graph that shows the average FOV values over time (Fig. 5(b)). FOV is inversely related to camera zoom, so a lower FOV corresponds to higher zoom setting resulting in higher resolution imagery. Next, we scripted 10 pedestrians to move along the route. These pedestrians were divided into three groups. Again, all pedestrians were successfully observed, albeit at a lower resolution, as suggested by higher average FOV values shown in Fig. 5(c). It is important to realize that the proposed video surveillance system is flexible and gracefully adapts to scene complexity. It is able to capture high resolution video when there are fewer people. However, as the number of pedestrians grow, PTZ cameras behave like wide-FOV cameras to ensure that everyone is observed, although at lower resolutions.

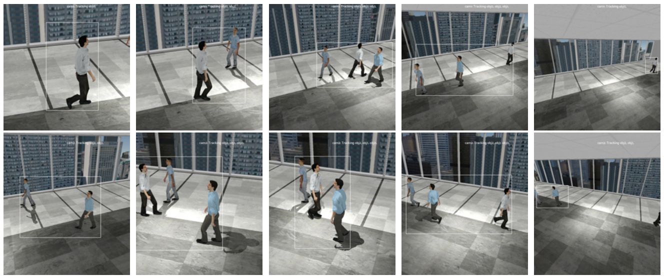


Figure 6. Row1: PTZ camera 1 observing pedestrians present in the hallway. Row2: PTZ camera 2 observing pedestrians present in the hallway. White rectangles depict the ROI computed by behavior arbitration routine and passed down to fixate and zoom behaviors.

Fig. 6 show a video surveillance system consisting of two PTZ and four passive wide-FOV virtual cameras. The PTZ cameras, which are uncalibrated, are tasked to record close-up imagery of the pedestrians who pass through the hallway. Here too, PTZ cameras automatically choose between the two alternatives: 1) observe a single pedestrian at higher resolution or 2) observe multiple pedestrians at lower resolution.

6. Conclusions

Future video surveillance systems will be capable of providing perceptive coverage of large environments over extended periods. In addition to robust visual analysis routines, we believe, such systems will be capable of self-management, in terms of power, sensing, bandwidth, processing, and storage. In other words future video surveillance systems need to be *autonomic*. While we are still far from realizing this vision, the research presented in this paper is inspired by such future video surveillance systems.

Specifically, we have developed behavior-based smart camera nodes capable of carrying out multiple observation tasks simultaneously. Our control methodology treats passive and active PTZ cameras within a unified framework, which allows one to develop camera network control strategy without worrying about the actual camera types and their respective sensing capabilities. We demonstrate a video surveillance system comprising passive and active PTZ cameras that tunes its sensing performance in response to scene complexity. Consequently, the video surveillance system avoids catastrophic sensing failures and its perfor-

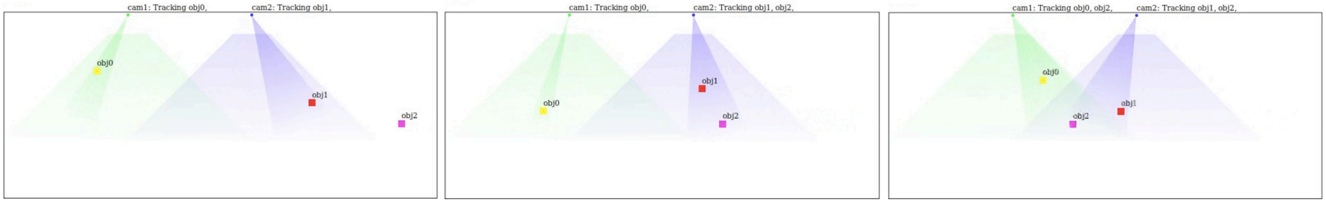
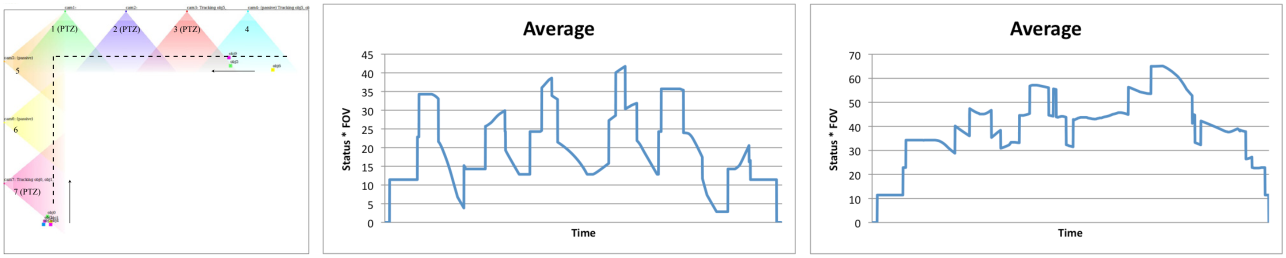


Figure 4. Camera switching. Left: Camera 1 (Green) is tracking object 0 (Yellow rectangle) and camera 2 (Blue) is tracking object 1 (Red rectangle). Middle: Object 2 (Purple rectangle) enters the scene and camera 2 automatically decides to observe it in addition to observing object 1. Right: Cameras 1 and 2 successfully handoff object 2 while still observing objects 0 and 1, respectively. This figure is best viewed in color.



(a) Setup

(b) 2 Pedestrian, lower FOV, higher zoom

(c) 10 Pedestrian, higher FOV, lower zoom

Figure 5. (a) Camera setup. (b) Cameras are able to capture (on average) higher resolution video when there are only two pedestrians in the scene. (c) Cameras capture (on average) lower resolution video when there are 10 pedestrians in the scene.

mance degrades gracefully when it is overwhelmed by a multitude of observation tasks. We evaluate the proposed video surveillance system by deploying simulated cameras in a realistic 3D virtual environment.

In the future, we plan to evaluate the proposed approach on larger simulated and physical camera networks.

References

[1] Panda3D, 2011. 1

[2] R. C. Arkin. *Behavior Based Robotics*. MIT Press, Cambridge MA, 1998. 2

[3] R. A. Brooks. A Robust Layered Control System for a Mobile Robot. In *IEEE Journal of Robotics and Automation*, volume RA-2 (1), Apr. 1986. 2

[4] R. Collins, O. Amidi, and T. Kanade. An Active Camera System for Acquiring Multi-View Video. In *Proc. International Conference on Image Processing*, pages 517–520, Rochester, NY, Sept. 2002. 2

[5] R. Farrell and L. S. Davis. Decentralized Discovery of Camera Network Topology. In *Proc. of the Second International Conference on Distributed Smart Cameras (ICDSC08)*, Menlo Park, CA, Sept. 2008. 2

[6] A. Hampapur, S. Pankanti, A. Senior, Y.-L. Tian, L. Brown, and R. Bolle. Face cataloger: {M}ulti-scale imaging for relating identity to location. In *Proc. IEEE Conference on Advanced Video and Signal Based Surveillance*, pages 13–21, Washington, DC, 2003. 2

[7] K. Heath and L. Guibas. Multi-person tracking from sparse 3D trajectories in a camera sensor network. In *Proc. of the Second International Conference on Distributed Smart Cameras (ICDSC08)*, Menlo Park, CA, Sept. 2008. 2

[8] N. O. Krahnstoever, T. Yu, S. N. Lim, K. Patwardhan, and P. H. Tu. Collaborative Real-Time Control of Active Cameras in Large-Scale Surveillance Systems. In *Proc. ECCV Workshop on Multi-camera and Multi-modal Sensor Fusion*, pages 1–12, Marseille, France, Oct. 2008. 2, 4

[9] Y. Li and B. Bhanu. Utility-based Dynamic Camera Assignment and Hand-off in a Video Network. In *Proc. of the Second International Conference on Distributed Smart Cameras (ICDSC08)*, pages 1–9, Menlo Park, CA, Sept. 2008. 2, 4

[10] F. Z. Qureshi. On the Role of Negotiations in Ad Hoc Networks of Smart Cameras. In *IEEE International Conference on Distributed Computing in Sensor Systems (DCOSS 10)*, pages 1–2, Santa Barbara, June 2010. 2, 4

[11] F. Z. Qureshi and D. Terzopoulos. Surveillance Camera Scheduling: {A} Virtual Vision Approach. *ACM Multimedia Systems Journal*, 12(3):269–283, Dec. 2006. 2

[12] F. Z. Qureshi and D. Terzopoulos. Smart Camera Networks in Virtual Reality. *Proceedings of the IEEE (Special Issue on Smart Cameras)*, 96(10):1640–1656, Oct. 2008. 2, 4

[13] F. Z. Qureshi and D. Terzopoulos. Planning Ahead for PTZ Camera Assignment and Control. In *Proc. Third ACM/IEEE International Conference on Distributed Smart Cameras (ICDSC 09)*, pages 1–8, Como, Italy, Aug. 2009. 2

[14] V. Singh, P. Atrey, and M. Kankanhalli. Coopetitive multimedia surveillance using model predictive control. *Machine Vision and Applications*, 19(5-6):375–393, 2008. 2

[15] B. Song, C. Soto, A. K. Roy-Chowdhury, and J. A. Farrell. Decentralized camera network control using game theory. In *Proc. of the Second IEEE/ACM International Conference on Distributed Smart Camers (ICDSC08)*, Menlo Park, CA, Sept. 2008. 2, 4