Angular Heuristics for Coverage Maximization in Multi-Camera Surveillance

Ahmed Abdelkader Moamen Mokhtar Hazem El-Alfy
Faculty of Engineering, Alexandria University
Alexandria, EGYPT 21544
Email: abdelkader@alexu.edu.eg, moamen.mokhtar@alex-eng.edu.eg, helalfy@alexu.edu.eg

***Abstract*—Multiple cameras are used to track targets moving amongst obstacles. Surveillance video streamed from a top-view camera is processed to control the orientation of multiple pan-tilt-zoom cameras to cover as many targets as possible at high resolutions. The problem of maximizing the number of covered targets with a set of cameras has been shown to be computationally expensive and hence, several approximations have been suggested in the literature. We develop our own ones, compare them to some existing approaches by extensive simulation and show their superiority. Our new heuristics make an attempt at continuous panning that is needed when moving to real world experimentation to achieve seamless target tracking.**

I. Introduction

Video surveillance systems are becoming more widespread every day. There are many uses to these systems in national safety, private surveillance and for providing special care for the disabled. Advances in network cameras, digital processing and communications have multiplied the number of cameras used. This situation has made the task of observing monitors by human operators not only outdated but also practically infeasible. For that reason, several methods have been developed to automate the detection and reporting of scenes, events or subjects of interest. Traditionally, hundreds of hours of surveillance video from static cameras are stored to backup systems and examined later if needed. Covering a large environment obviously requires the use and coordination of several cameras. With the increased numbers of cameras in use, interest has recently shifted to control individual pan-tilt-zoom (PTZ) cameras to focus on specific events or subjects, in addition to offline archival purposes.

Monitoring several targets moving in an environment with multiple PTZ cameras has many conflicting objectives which makes it a challenging task. Among those objectives are: (i) maximizing the number of covered targets, (ii) obtaining a high resolution view of each target, (iii) obtaining an unobstructed view of targets (occlusion by obstacles), (iv) efficient recomputing of camera to target assignments, (v) considering that targets constantly move, (vi) obtaining a smooth video per camera without abrupt jumps, (vii) considering camera hardware restrictions such as pan and zoom speeds and, (viii) considering processing configurations, whether centralized or distributed ([1]).

Many attempts to the problem of monitoring multiple targets with multiple cameras have been suggested to date, with each one looking at the problem from a different angle. We can fairly claim that this problem is still far from being satisfactorily solved.

A. Problem Definition

We consider the problem of tracking multiple targets moving among obstacles using multiple cameras. We are mainly concerned with small scale networks. Hierarchical methods can be used to handle large camera networks. Our long-term objective is to develop a method that can be applied in real time to an actual surveillance system. We assume a top-view wide-angle camera (*master camera*) returns 2D ground coordinates of moving targets. Top view tracking is more robust than its side-view counterpart and suffers less from challenges such as occlusion. We are not concerned with the details of tracking in the present work as much as we are concerned with strategies to pan cameras in order to maximize the number of covered targets. The returned target locations are then used to automatically control the orientation and, eventually, the zoom of a set of high resolution PTZ cameras (*slave cameras*) in order to follow (*track*) those targets.

We have studied many approaches presented in the literature to the problem of maximizing the number of covered targets with a set of cameras (or sensors). Those approaches are expensive and, for that reason, heuristics have been suggested to add efficiency at the expense of optimality. In addition, the dynamic nature of the problem requires an approach that can update itself as targets move, versus completely resolving the problem.

B. Contribution

We compare several heuristics found in the literature, suggest two new ones and verify, using our simulator, their superior performance. In addition to the common objective of maximizing the number of targets covered by cameras, our new heuristics present, to the best of the authors' knowledge, a first attempt at continuous panning as opposed to selecting from a discrete set of orientations as is common in the literature.

The rest of the paper proceeds as follows: we survey current literature in section II, then, in sections III and IV, we describe our heuristics to solve the problem. In section V, we evaluate our method and compare it to earlier approaches. Finally we conclude and suggest future work in section VI.

II. Related Work

In the area of discrete combinatorial optimization, several attempts have been made at the problem of maximizing the number of covered targets using the minimum number of directional sensors. Relevant literature is found in [2], [3], [4]. The optimal solution is computationally expensive and hence, several sub-optimal heuristics are suggested. In addition, the solved problem is static.

When moving to the practical area of camera management, many challenges appear. The first step is to set-up and calibrate the cameras [5]. Then, comes the task of controlling cameras. It is worth noting that one of the first attempts that uses one wide-view camera and one high-resolution camera is a decade old [6]. Different combinatorial approaches have also been suggested. A self updating bipartite matching approach with clustering of neighboring targets is presented in [7]. An overview of the more generic assignment problem is presented in [8]. Probabilistic approaches can be found in [9], [10]. A centralized approach is suggested in [11], a 3D real system is presented in [12] and a large scale system can be found in [13]. Good surveys with comparisons of several methods are found in [10] and, for game theoretic approaches, in [14].

III. A Relaxed Coverage Function

The combinatorial nature of the discrete orientation assignment problem makes it difficult to define a continuous objective function and seek optimal or approximate solutions, as far as target coverage is concerned. This situation leaves us with either the computationally prohibitive integer programming formulation or trying different suboptimal heuristics that allow for computationally efficient and evidently practical solutions. The problem, however, could be attacked by relaxation but we could not find such a formulation in the literature. Taking visibility constraints into consideration to account for the possible presence of walls and obstacles makes it even harder to model.

It is generally believed that the assignment problem at hand is NP-hard. The authors in [3], [4] present several arguments to this point which all start by first mapping the problem to a structure similar to the Maximum Coverage Problem. This essentially rids the problem of its geometric nature and no formal reduction was presented. We anticipate a more rigorous proof similar in spirit to the complexity result of the vertex guard art gallery problem in [15].

The problem does bear great similarity to the well known Set Cover Problem, with the additional difficulty of multiple orientations per camera. A linear programming relaxation of $SETCOVER$ was first considered by Lovász [16] which could yield a logarithmic factor approximate solution through randomized rounding [17]. It was later shown that the greedy strategy of repeatedly selecting the set that covers the largest number of uncovered elements performs equally well [18]. This should justify why the greedy algorithms suggested in [2], [4] are very effective.

Taking hint from Lovász's relaxation, we seek a relaxed coverage function where covered and non-covered targets are not strictly mapped to 1 and 0, respectively, as one would expect. Such a function might be easier to handle while still preserving the coverage maximization objective. The following technical lemma provides useful hints towards the development of such flexible objective functions to our problem.

Lemma 1. *Given N cameras and T targets to cover, fix a positive real $w_{c,t}$ for each camera c and target t, such that $\frac{w_{min}}{w_{max}} > \frac{T-1}{T}$. Any weight function of the form*

$$f(t) = \begin{cases} 0, & t \text{ is not covered,} \\ w_{c,t}, & t \text{ is covered by } c, \end{cases} \tag{1}$$

describes at least one optimal solution maximizing $\sum_{t \in T} f(t)$.

Proof: Let n and n' be the sizes of two sets of targets the system could possibly cover, with $n, n' \in \{0, 1, 2, ..., T\}$. For $n > n'$, any coverage maximization objective function should always assign a larger score to the larger set. Now, assume all targets in the larger set were assigned the smallest possible weight i.e. w_{min} while all targets in the smaller set were assigned the largest possible weight i.e. w_{max}. The system must guarantee $n\, w_{min} > n'\, w_{max}$ or $\frac{w_{min}}{w_{max}} > \frac{n'}{n}$. This is trivially satisfied by the usual binary coverage function with $w_{min} = w_{max} = 1$. However, we desire to find the lower bound for this ratio which is clearly determined by the largest value of the ratio $\frac{n'}{n}$. Maximizing the numerator simplifies this to $\frac{n-1}{n}$, which is strictly increasing in n. Setting $n = T$ yields the largest ratio and forces $\frac{w_{min}}{w_{max}} > \frac{T-1}{T}$. When $n = n'$, the two sets do not necessarily get assigned the same weight. Hence, the set maximizing the objective function above also achieves maximum coverage. ■

Such a characterization of weight functions reveals the undesirable discontinuity over covered and non-covered targets. We suggest a workaround exploiting the geometric structure of the problem through what we call *Angular Relaxation*.

It is usually preferred to have subjects at the center of the FOV rather than to the sides. This leads to a natural weight function we derive from the angular offset θ computed in the reference frame defined by the camera and its current direction d. We start with the following expression

$$f(t) = \left(\frac{T-1}{T}\right)^{\theta_t/\theta_{max}} \tag{2}$$

where θ_{max} is half the FOV. For covered targets, $\theta \in [0, \theta_{max})$ and we get $f(t) \in (\frac{T-1}{T}, 1]$. This means subjects at the center are assigned weights of 1 which decays exponentially to the sides of the field of view, and attains a value of $\frac{T-1}{T}$ right outside the FOV. Subjects further outside the FOV get smaller weights approaching 0. However, to achieve near zero weights for targets right outside the FOV we relax the bounding ratio for targets inside the FOV and use $e^{-\alpha}$ as the base instead of $\frac{T-1}{T}$, dropping other constants.

As the camera pans, θ becomes a function of d. For a target at $d = 0$, θ is the triangular wave $2\pi|\frac{d}{2\pi} - \lfloor \frac{d}{2\pi} + 0.5 \rfloor|$. We approximate θ by its Fourier series using just two terms to be $\frac{\pi}{2}(1 - \cos d)$, as shown in Figure 1.

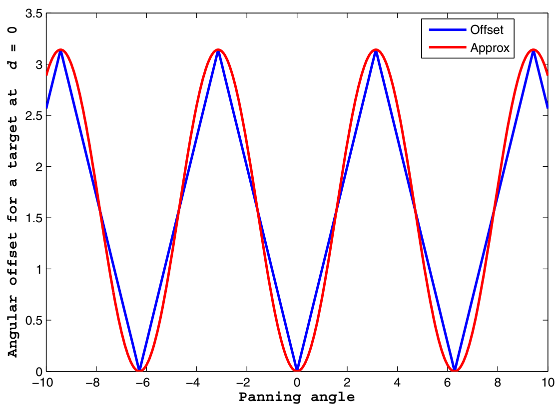

Figure 1. Angular offset θ as a function of the panning angle d.

With that, we reach the following form:

$$f(t) = e^{-\alpha(1-\cos\theta_t)} \tag{3}$$

Furthermore, we weight the whole set of targets covered by having any one camera assigned to a given direction by:

$$f(c, d) = \frac{1}{|T_c|} \sum_{t \in T_c} f_{c,d}(t) \tag{4}$$

where the subscript c, d requires angular offsets to be computed relative to camera c and direction d. T_c is the set of targets that can be covered by c as in [2]. Figure 2 shows an example with four targets at angular offsets $[1, 1.3, 1.8, 4]$ and $\alpha = 2$. Notice how each individual target contributes a peak of height 1 at the offset where it is located. The distribution of targets defines the shape of the objective function which peaks at the *hottest* region with a value approximating the anticipated coverage, before getting normalized by $\frac{1}{|T_c|}$.

Our heuristic is a direct application of the greedy algorithm for weighted set cover [19] using Equation 4. While the algorithm was not developed for our case it still serves as a good start and we consider more tailored formulations as future work e.g. [20]. The algorithm repeatedly selects the (c^*, d^*) pair maximizing $f(c, d)$ till all cameras are assigned. Gradient ascent initialized at the target yielding the highest value of the objective function can be used to find the direction d^* for each camera. As the number of iterations required to reach the optimal value can be kept small, the algorithm is $O(TN^2)$ with an $O(TN)$ cost per iteration to compute the best score for each camera in $O(T)$.

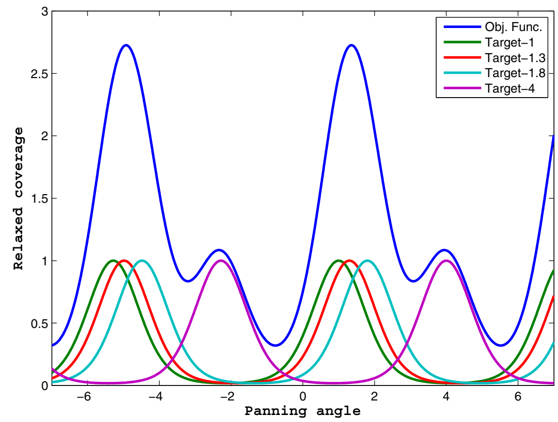

Figure 2. Composition of the objective function for four targets.

IV. Coverage by Angular Sweep

Taking hint from the algorithms suggested in [2] and [4], where panning was restricted to just 8 orientations, we make the following observation. Covering any group of subjects by a given camera can always be achieved by panning till one or more subjects lie on either side of the FOV. To greedily maximize its coverage, a camera can iterate over the targets it could possibly cover and compute the coverage achieved by panning to have the target at hand at the right or left sides of the FOV as in Figure 3. Using only one side of the FOV to sweep is enough. As the FOV rotates, all possible groupings of targets will be encountered. After the panning angle yielding the highest coverage is found, it may be arbitrarily shifted to center the FOV onto the targets to be covered. Greedily choosing the camera achieving the highest normalized coverage and proceeding till all cameras are assigned amounts to an $O(T^2N^2)$ algorithm with $O(T)$ panning angles and an $O(T)$ cost to compute the coverage. This method easily presents an upper bound for similar greedy heuristics with discrete orientations, regardless of how many they are e.g. 8 [2] or higher, while at the same time being more efficient.

V. Experimental Evaluation

We study the performance of our new heuristics, Angular Relaxation (AR) and Angular Sweep (AS), compared to the Centralized Greedy Algorithm (CGA) [4] and Centralized Force Algorithm (CFA) [2], all against optimal solutions

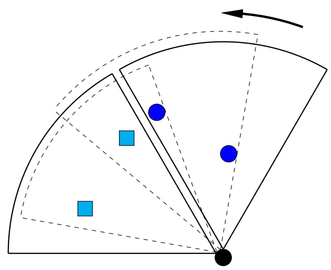

Figure 3. Sweeping finds *all* maximal groupings unlike limited orientations.

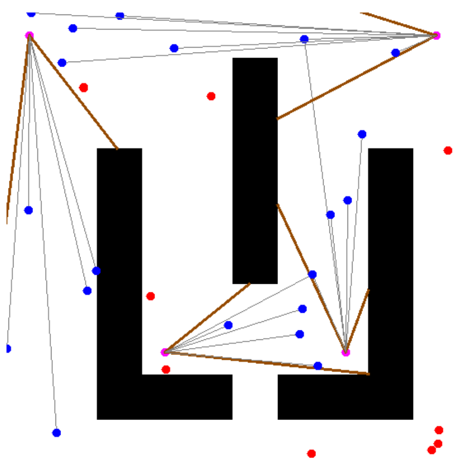

Figure 4. A sample test map with assignment and coverage.

obtained by brute-force search over the orientation space (OPT). Since all earlier methods we compare against are using a discrete number of panning angles, that number is indicated by the parameter used for each graph. We present results for both static and dynamic scenarios with and without walls to simulate different indoors and outdoors environments. We use the coverage metric as the number of covered subjects divided by the total number of subjects in any given frame/iteration. For all runs, we use $\theta_{max} = \frac{\pi}{8}$ and $\alpha = 4$, chosen empirically.

For dynamic scenarios, we do 1000 iterations with camera assignments being recomputed every 10 frames. We do not pan cameras until the next assignment and assume that panning is performed instantaneously. We do not bound the FOV by maximum/minimum ranges as we are interested in small scale setups and study centralized algorithms. When studying the speed of targets we assume all targets move with the same speed for every step they take. We use manually designed maps with different camera setups and random initial target locations with controlled separation, as in Figure 4. Targets then repeatedly select random destination points and move towards them in straight lines.

Figure 5 shows the superior performance of AS. AR

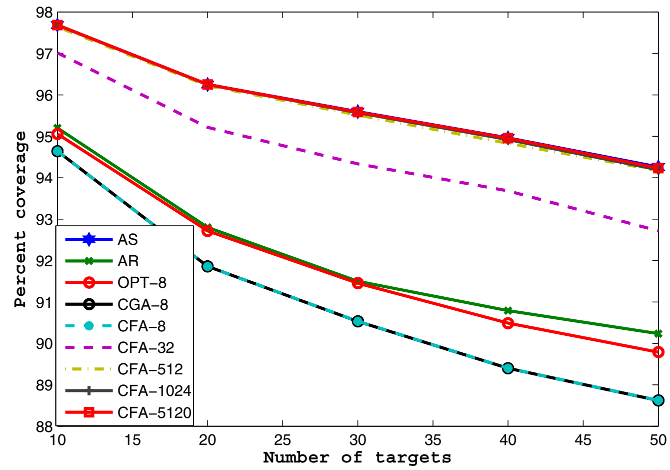

Figure 5. Varying the number of static targets without walls.

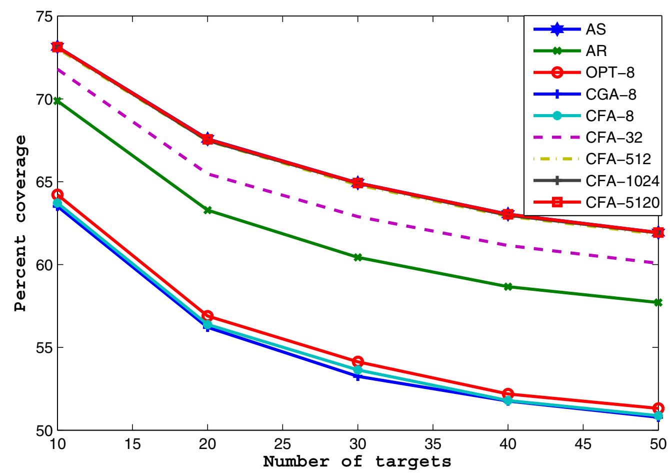

Figure 6. Varying the number of static targets with walls.

performed much better in the presence of walls as in Figure 6. We continue the comparison with the CFA-32.

Figures 7 and 8 confirm the advantage of AS when the number of moving targets varies. It becomes clear that AR does not perform as well as the CFA-32, but better than the CFA-8.

All three algorithms came close with varying speeds as shown in Figure 9. The indoors case was not much different as in Figure 10. This also shows that AR performs worse with higher target speeds. Figure 11 and 12 show runtimes for different numbers of static targets, performed on an Intel i7 Quad Core with 4 GB of RAM. Our methods are more efficient than the CFA-32 and above, with the AR performing faster in the presence of walls.

VI. Conclusion and Future Work

We have addressed the problem of covering as many targets as possible, moving through obstacles, using a set

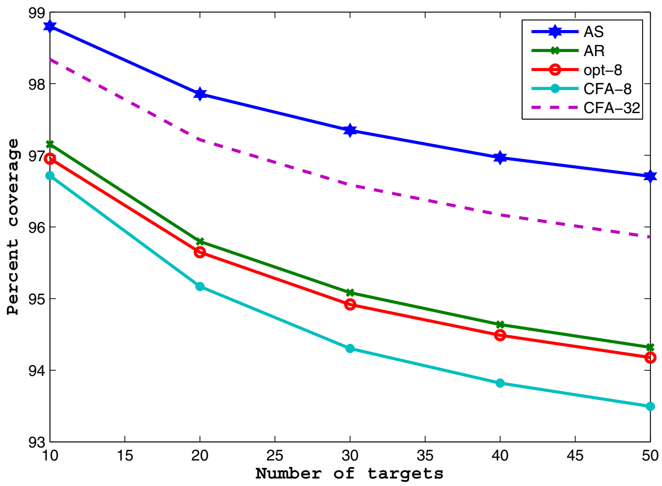

Figure 7. Varying the number of mobile targets without walls.

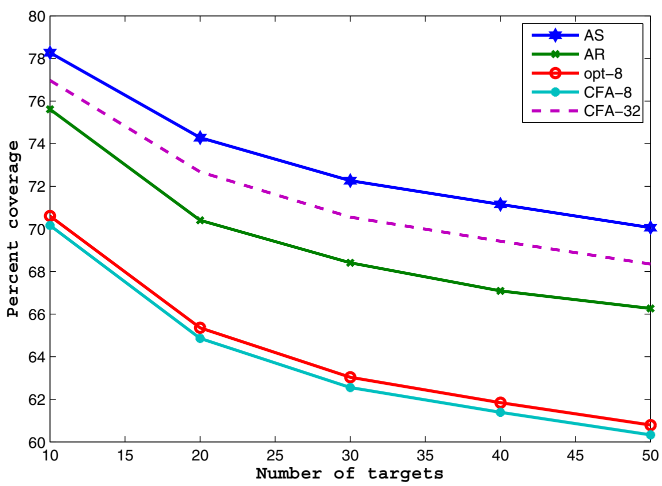

Figure 8. Varying the number of mobile targets with walls.

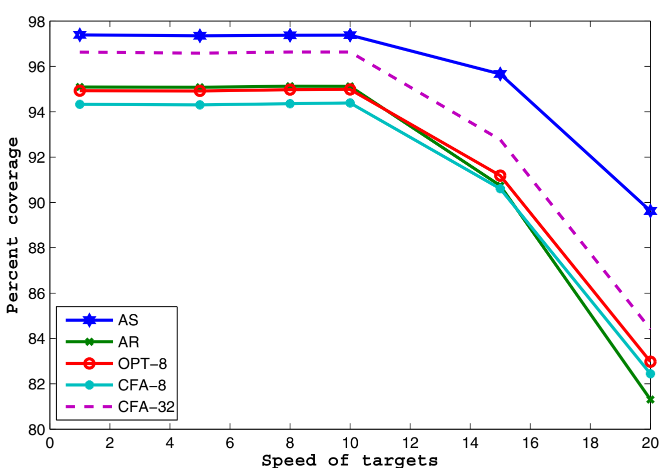

Figure 9. Varying speed of targets without walls.

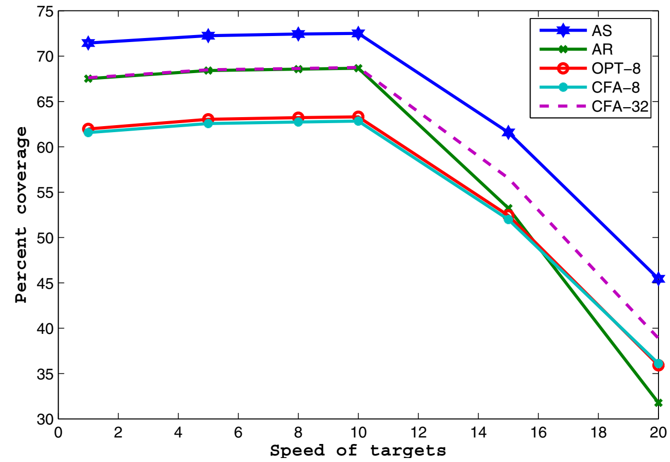

Figure 10. Varying speed of targets with walls.

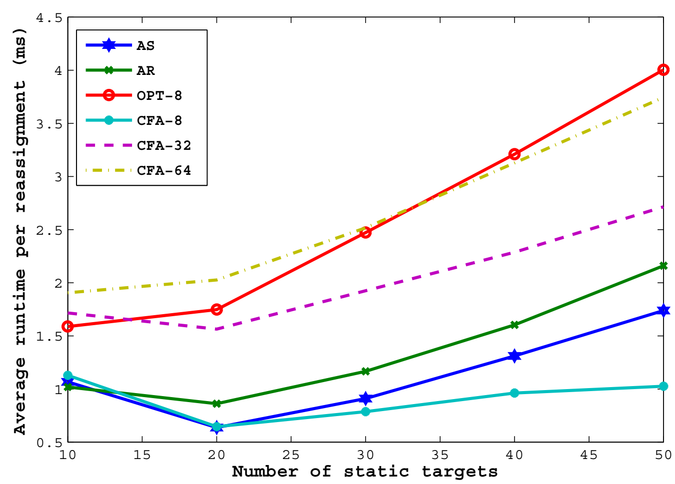

Figure 11. Runtime varying the number of static targets without walls.

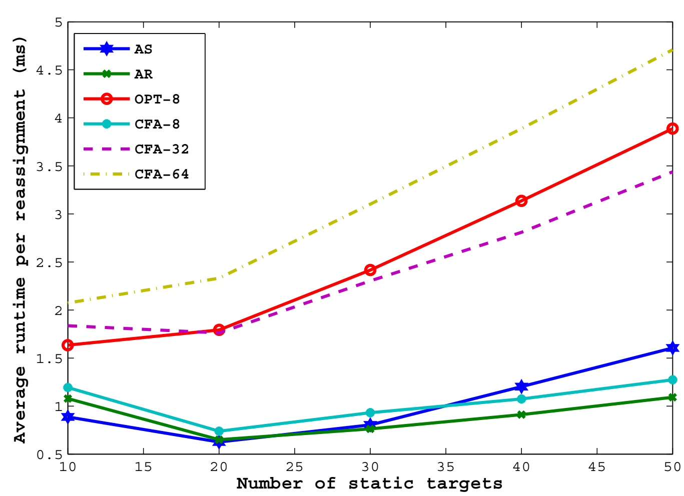

Figure 12. Runtime varying the number of static targets with walls.

of PTZ cameras (or directional sensors). We introduced two heuristics. One performs numerical optimization of a carefully designed relaxed coverage function using gradient ascent. The other utilizes the sweeping concept from computational geometry to yield all maximal groupings of targets. Experimental results showed that our heuristics supersede earlier approaches in terms of coverage rates and running times and have the advantage of providing continuous camera panning for smoother target tracking.

In the future, we plan to develop heuristics that would provide a more stable system by reducing the sum of the angles that the cameras have to pan to cover the targets. In real surveillance systems, we seek to obtain better views of targets (facing camera, illuminated) and to trigger camera panning by scene events (rather than by mere people motion). Plug-in applications (such as face/plate recognition) can benefit from the zoom-in view from the slave cameras, however these require a better iris focus at high zooms.

Acknowledgments

This work was supported in part by the research enhancement program offered by the Faculty of Engineering, Alexandria University, Egypt (AlexEng REP2). The authors would like to thank Ahmed Zakaria, Abdelrahman Alaa Eldin and Hamdy Elgammal for creating the test maps and an early prototype of the code.

References

[1] A. C. Sankaranarayanan, A. Veeraraghavan, and R. Chellappa, "Object detection, tracking and recognition for multiple smart cameras," *Proceedings of the IEEE*, vol. 96, no. 10, pp. 1606–1624, 2008.

[2] V. P. Munishwar and N. B. Abu-Ghazaleh, "Scalable target coverage in smart camera networks," in *Fourth ACM/IEEE International Conference on Distributed Smart Cameras (ICDSC'10)*. ACM, 2010, pp. 206–213.

[3] Y. Cai, W. Lou, M. Li, and X.-Y. Li, "Target-oriented scheduling in directional sensor networks," in *26th IEEE International Conference on Computer Communications (INFOCOM'07)*, 2007, pp. 1550–1558.

[4] J. Ai and A. A. Abouzeid, "Coverage by directional sensors in randomly deployed wireless sensor networks," *Journal of Combinatorial Optimization*, vol. 11, no. 1, pp. 21–41, 2006.

[5] F. Porikli and A. Divakaran, "Multi-camera calibration, object tracking and query generation," in *Proc. International Conference on Multimedia and Expo (ICME'03)*, Baltimore, MD, USA, Jul. 2003, pp. 653–656.

[6] R. T. Collins, A. J. Lipton, H. Fujiyoshi, and T. Kanade, "Algorithms for cooperative multisensor surveillance," *Proceedings of the IEEE*, vol. 89, no. 10, pp. 1456–1477, Oct. 2001.

[7] H. El-Alfy, D. Jacobs, and L. Davis, "Assigning cameras to subjects in video surveillance systems," in *Proc. IEEE International Conference on Robotics and Automation (ICRA'09)*, Kobe, Japan, May 2009, pp. 837–843.

[8] D. W. Pentico, "Assignment problems: A golden anniversary survey," *European Journal of Operational Research*, vol. 176, no. 2, pp. 774–793, 2007.

[9] E. Sommerlade and I. Reid, "Probabilistic surveillance with multiple active cameras," in *Proc. IEEE International Conference on Robotics and Automation (ICRA'10)*, Anchorage, AK, USA, May 2010, pp. 440–445.

[10] 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 *Eleventh International Conference on Autonomous Agents and Multiagent Systems (AAMAS'12)*, Valencia, Spain, June 2012.

[11] N. Takemura and J. Miura, "View planning of multiple active cameras for wide area surveillance," in *Proc. IEEE International Conference on Robotics and Automation (ICRA'07)*, Roma, Italy, Apr. 2007, pp. 3173–3179.

[12] L. Fiore, D. Fehr, R. Bodor, A. Drenner, G. Somasundaram, and N. Papanikolopoulos, "Multi-camera human activity monitoring," *Journal of Intelligent & Robotic Systems*, vol. 52, no. 1, pp. 5–43, 2008.

[13] F. Z. Qureshi and D. Terzopoulos, "Planning ahead for ptz camera assignment and handoff," in *Third ACM/IEEE International Conference on Distributed Smart Cameras (ICDSC'09)*, Como, Italy, September 2009, pp. 1–8.

[14] Y. Li and B. Bhanu, "Utility-based camera assignment in a video network: A game theoretic framework," *Sensors Journal, IEEE*, no. 99, pp. 1–1, 2010.

[15] D. Lee and A. Lin, "Computational complexity of art gallery problems," *Information Theory, IEEE Transactions on*, vol. 32, no. 2, pp. 276–282, 1986.

[16] L. Lovász, "On the ratio of optimal integral and fractional covers," *Discrete mathematics*, vol. 13, no. 4, pp. 383–390, 1975.

[17] P. Raghavan and C. Tompson, "Randomized rounding: a technique for provably good algorithms and algorithmic proofs," *Combinatorica*, vol. 7, no. 4, pp. 365–374, 1987.

[18] N. Young, "Randomized rounding without solving the linear program," in *Proceedings of the sixth annual ACM-SIAM symposium on Discrete algorithms*. Society for Industrial and Applied Mathematics, 1995, pp. 170–178.

[19] V. Chvatal, "A greedy heuristic for the set-covering problem," *Mathematics of operations research*, pp. 233–235, 1979.

[20] M. Gairing, "Covering games: Approximation through non-cooperation," *Internet and Network Economics*, pp. 184–195, 2009.