The 2010 IEEE/RSJ International Conference on
Intelligent Robots and Systems
October 18-22, 2010, Taipei, Taiwan

Target Acquisition in Resource Constrained Stationary Camera Systems

Aparna Veluchamy and Monica Anderson

Abstract— **Involving humans to monitor complex areas using a network of cameras is expensive and prone to errors and fatigue. Hence arises the need for unmanned, cooperative camera systems that can identify targets with intelligent camera motions. Acquiring the targets requires panning of the cameras in a manner that minimizes the time and the contiguous space that goes unmonitored.**

Average Linear Uncovered Length (ALUL) quantifies the effect of a specific unmonitored area by computing the average length of all lines drawn through the area. This paper considers ALUL as a metric that indicates the target acquisition ability of a particular system configuration. When used in conjunction with temporal constraints, we can intelligently automate camera coverage to improve target acquisition and tracking performance of the system. The results from the experiments confirm that the coverage in constrained environments when the existing camera configuration cannot view large portions of the region of our interest improves when ALUL is considered.

I. INTRODUCTION

Advancements in camera technology have led to expanded use of visual monitoring of publicly accessible areas. Such systems enable monitoring of complex, urban environments for security purposes to identify abnormal or potentially malicious activity. For an event to be identified, the region of interest (ROI) should be monitored such that the entire region is covered by at least one camera all the times. Given an unlimited number of resources to cover the area, the problem would be limited to positioning the cameras. However, we generally find there is a practical limit to the number of cameras deployed in an environment due to the cost and complexity of the deployment.

In manual surveillance systems, the human operators control the panning and tilting of the cameras to track targets of interest. But the disadvantage of such systems is that they do not guarantee uniform, unprioritized coverage of the area. We focus on achieving a uniform coverage of our ROI without any bias on the area being viewed. We assume that the knowledge about the environment such as areas of particular activities,etc. are not available to the system. Hence the entire area should be given equal priority in coverage to acquire any target activity inside the ROI.

This paper considers uniform coverage of an ROI in terms of a metric that quantifies the ability of a camera configuration to acquire new targets as well as an algorithm which maximizes acquisition by quantifying the effect of unmonitored areas. We focus on real environments with the assumptions that some area may be permanently occluded and the configuration does not allow for complete coverage and minimal redundancy. Coverage is accomplished so that it not only minimizes the time any of the region is left unmonitored but also the size of the space that is unseen by all of the cameras collectively.

The rest of the paper is organized as follows. Section II discusses related research in surveillance systems. Section III presents the algorithms for cooperative camera surveillance system. Section IV focuses on the experiments conducted whereas Section V analyzes the performance of the proposed algorithm. Section VI, finally concludes the paper.

II. RELATED WORK

Several researchers have presented their approaches on related issues. Some of them are discussed here.

The work by Tilak et al. [1] categorizes surveillance methods to notify the observer of a suspicious activity based on the feeds from the sensors. They are continuous, event-driven and observer-initiated. In the Continuous model, information feeds are sent to the observer at regular intervals. Whereas in event-driven method, the observer will be notified whenever a suspicious event occurs. Observer-initiated method allows for the user to get the feeds from the sensor whenever required.

A similar problem of target acquisition and tracking is addressed in [2]. But it assumes that enough resources are available to be dispatched for either acquiring or tracking at a given instant, but does not balance both. Thereby better surveillance is ensured through resource allocation. Moreover, human interruption is also allowed.

Whereas, [3] and [4] propose a method to balance these tasks given a constrained scenario. A prioritized surveillance approach with a presumed knowledge of the ROI has been focused in Davis et al.[3]. Areas of suspicious activity are identified and learned by the system with a probabilistic map which prioritizes the areas that needs to be more focused for any activity. A different approach to a prioritized surveillance by choosing each ROI within the entire area based on their activity levels is proposed in [4]. A ranking algorithm prioritizes the specific regions to be surveyed by the available cameras. The activity levels are determined by the feeds from a single static low resolution camera which views the entire region.

Qureshi and Terzopoulos [5] present a scheduling approach to calibrate one of the available Pan-Tilt-Zoom cameras to view a single pedestrian identified by a static camera with a wider field of view. It also assumes that the ROI can be viewed from at least one camera. The static camera, upon detection of a pedestrian movement, dispatches the viewing

This work was supported in part by the following NSF grants: IIS-0846976 and CCF-0829827.

A Veluchamy and M Anderson are with the Department of Computer Science, University of Alabama, Tuscaloosa, Alabama 35487 USA (email:aveluchamy@crimson.ua.edu; anderson@cs.ua.edu).

job to one of the available cameras. Although this approach also assumes controlled camera placement and unconstrained coverage, it does focus on uniform coverage since it does not use domain specific knowledge to prioritize area coverage.

A target tracking algorithm using multiple sensors assuming the domain-awareness of surveillance system is presented by David [6]. Regions are identified within the area being covered which are prone to target activity. The coverage by each of the configurations of the sensors are also acquired. Then, a heuristic search is used to find a sensor configuration that optimizes coverage function.

The trade-offs faced by using limited resources are discussed in [7] with respect to multi-sensor networks. A quality of service metric is proposed that analyzes the coverage performance by the time the target goes undetected by any of the resources when within the ROI. The sensor configurations are adjusted to minimize the unseen time of a target until an optimum.

A taxonomy of metrics for coverage analysis is presented in [8] namely contiguous undetected time, number of targets lost by the cameras, and time before targets are detected which can be utilized to control camera movement. A continuous (unintelligent) cycle of camera movement in all four directions proved to minimize the time taken to detect a target as compared to intelligent camera movements to track the targets.

A space-based coverage of the ROI coupled with a time-based approach in a constrained environment has not been proposed as a method for effective surveillance.

III. APPROACH

The aim of the paper being to maximize the target acquisition, we propose an algorithm for intelligently panning and tilting the cameras in such a way that (a) it minimizes the time and size of the contiguous space that goes unmonitored and (b) identifies as many targets as possible. A uniform unprioritized surveillance of the ROI ensures a fair coverage of the ROI.

The chosen ROI is discretized into uniform sized cells and the terrain elevation map of the ROI is acquired which can be used to determine the height of each cells in the ROI and the cells that are occluded with each of the cameras. The position and the initial angles of the cameras are also input to the algorithm.

$$Utility(ROI, \alpha, \bar{X}) = \sum_{z}^{ROI} visible(\alpha, \bar{X}, z) * 2^{u_z} \quad (1)$$

$$ROI = \{z_1, z_2, z_3, z_4, z_5 ....z_n\} \quad (2)$$

where
$\bar{X}$ - represents the position of the cameras
α - represents the angles of the cameras
u_z - unmonitored/occluded cells
z - each cell in the ROI

In surveilling urban environments, the cameras are to be positioned to acquire new targets whenever they are available in ROI. A static approach for such surveillance tasks will solve the problem if there are enough cameras to view the entire region without any movement of the cameras. But the issue in hand is when the region being surveilled is not collectively visible to all the cameras instantaneously. In order to efficiently detect targets, the holes in the ROI (those which can be seen if the cameras are panned accordingly) are to be quantified. Two metrics quantify the visible areas and the holes. First, the utility function gives the contiguous time for which a space goes unmonitored defined in Eqn 1, where the function visible calculates whether or not the particular cell is visible at this instant, given the position and the angles of the cameras. ROI represents the set of all the cells(z) in the selected region. The second metric, the Average Linear Uncovered Length (ALUL), quantifies the unmonitored and occluded cells in the ROI. Originally defined in [9], ALUL is calculated as shown in the Eqn 3 assuming each of those uncovered spaces to form a polygon. Max-ALUL, thus is the size of the largest hole in the ROI.

$$ALUL = \pi \frac{Area(Polygon)}{Perimeter(Polygon)} \quad (3)$$

Our real-time camera movement planning algorithm (Algorithm 1), relies on these metrics to decide the optimal movement of the cameras for the next time step. The camera movements are chosen to minimize the utility measure and the ALUL. For each time step, the ALUL is calculated as the camera moves and the maximum value is identified as the system ALUL(S-ALUL). The target acquisition algorithm is designed based on a hybrid Hill-climbing algorithm incorporating the approach to minimize S-ALUL. The Hill-climbing approach ensures minimizing the unmonitored time of the areas and S-ALUL ensures minimizing the space metric. ALUL metric aids us to position the cameras so that they minimize the unmonitored cells that can be covered by at least one of the cameras.

A. Applying ALUL to Target Acquisition within Regions of Interest

The ROI is defined by the domain and security needs. It may be contiguous or distributed. There are no assumptions about the structure of the ROI. In turn, we consider a non-optimized camera placement. The cameras are positioned within the environments based on placement, wiring and access needs. In addition, there is no assumption that any area is visible from some angle. Intervening obstacles such as buildings or trees may make specific sections occluded from any angle based on the existing camera placement. This approach analyzes how to maximize target acquisition based on existing resources (ROI selection, configuration, camera placement and orientation).

Each space in the ROI can be distinguished into visible and occluded areas. Visible areas can be seen from at least one camera from a specific set of angles. Occluded areas cannot be seen from any angle given the camera placement.

Algorithm 1 Planning camera movement

Require: t=t+1 {Increment time step}
Require: ROI={cells} {ROI is discretized into a set of cells}
Require: $p_{i,t} = \{angle_{i,t}, pos_{i,t}\}$ {each camera has a current position and angle}
u(cells)=u(cells)+1 {unseen cells have count incremented by 1}
for i in cameras **do**
{Calculate utility of incremental left pan (See Equation 1}
{Calculate utility of incremental right pan (See Equation 1)}
$p_{t+1,i} = max(Utility(leftpan), Utility(rightpan))$
ALUL(p) {Calculate overall ALUL for proposed change}
if $ALUL(p) < ALUL_{max}$ {Proposed pan violates system threshold ALUL} **then**
$p_{t+1,i}$=$p_{t+1,i}$+ -1 {Calculate overall ALUL for opposite direction}
if $ALUL(p) < ALUL_{max}$ {Opposite pan violates system threshold ALUL} **then**
$p_{t+1,i} = p_{t,i}$ {Leave pan angle unchanged}
else {Opposite angle is valid}
$p_{t+1,i} = p_{t+1,i}$ {Change pan angle to opposite angle}
end if
else {Proposed angle is valid}
$p_{t+1,i} = p_{t+1,i}$ {Change pan angle to proposed angle}
end if
end for

ALUL measures the contiguous areas within the ROI given the existing camera placement and pan angles that are not visible at this instant. These "holes" include both areas that are either unmonitored currently or occluded. The calculation produces a vector of values that characterize the coverage of the area both in terms of camera placement and pan angle. The Max-ALUL is the maximum value in this vector which measures the largest hole within the ROI.

We hypothesize that the Max-ALUL better estimates how long a target can operate within the ROI without detection. If true, then driving this value lower decreases the time targets go unmonitored. Max-ALUL is unrelated to methods that maximize viewing area to optimize detection. These approaches attempt to view the maximum area without considering the shape of the occluded area.

IV. EXPERIMENTAL RESULTS

To validate the hypothesis that Max-ALUL is a better measure of acquisition performance than the total viewing area (TVA), we perform two sets of experiments. The first set of experiments compares target acquisition and tracking from cameras with pan angles statically set to either minimize Max-ALUL or maximize TVA.

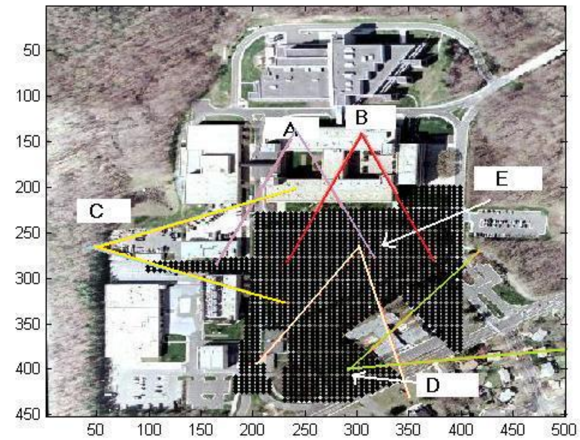

Fig. 1. Chosen ROI in the campus along with the positions of the cameras.

In addition to static pan angle surveillance, we also experiment with a panning surveillance algorithm that maximizes TVA while it minimizes time since last seen with Algorithm 1. The second set of experiments compares this base panning algorithm with an augmented algorithm that constrains movements that cause the Max-ALUL to exceed a threshold.

An urban corporate campus was chosen as our ROI as shown in Figure 1. The map is discretized into equi-sized cells of 5x5 meters each. The chosen ROI consisted of 1708 cells of the 7957 cells in the entire map. A height map of the ROI is used to determine visibility in terms of intervening obstacles from any camera position and pan angle. The ROI was chosen to be particularly challenging to cover by requiring a wide-field of view.

A variety of camera placements were used in the experiments. The camera positions were chosen manually to be realistic and to exercise the algorithms under both highly occluded, under and over constrained conditions. The number of cameras in each experiment varied from one to five.

Since we are concerned with surveillance in the absence of domain specific information, the experiments utilized targets with largely random trajectories. Targets move from cell to cell with a bias for moving straight. For a small percentage of the time, targets will choose to stand still or turn. Using the height map, targets cannot move to adjacent cells with a difference in elevation of over five meters so they don't move up the sides of buildings. Generated target trajectories are reused in each algorithm for comparison purposes because performance is largely dependent upon whether the targets are in unmonitored, occluded or visible areas.

V. ANALYSIS

In this research, we propose that Max-ALUL better measures acquisition performance by estimating how long a target can operate within the ROI without detection.

TABLE I

COVERAGE ACHIEVED BY USING THE SPECIFIED CAMERAS IN TERMS OF THE VIEWED AND OCCLUDED/UNMONITORED CELLS

Cameras	Camera pan angles	Percentage of cell coverage over visible cells	Max ALUL
D	100	76.5	9.0783
	90	73.79	8.4866
A	-100	95.14	10.2482
	-80	59.24	8.3825
A,D	-100;90	75.46	7.9813
	-90;80	66.77	6.0805
A,C,D	-100; -10;60	80.58	4.2343
	-100; 0;60	79.85	3.9137
A,B,C,D	-100; -90;0;60	83.62	3.8397
	-80; -100; -10;60	82.41	3.7525
A,B,C,D,E	-100; -80;0; 80;-70	81.7	3.9477
	-90;-90; -10;70; -60	80.81	3.7525

TABLE II

TARGET TRACKING AND ACQUISITION PERFORMANCE USING CAMERA PAN ANGLES IDENTIFIED TO HAVE THE MINIMUM MAX-ALUL OR MAXIMUM VIEW-ABILITY

Cameras	Mean time steps to acquire targets		Mean time steps targets are continuously tracked	
	Max Area	Min Max-ALUL	Max Area	Min Max-ALUL
D	9.2803	10.127	5.8556	3.2511
A	7.6569	1.8087	8.12	9.7095
A,D	8.5104	1.6822	6.1125	14.1461
A,C,D	0.6425	0.5899	13.4	16.6511
A,B,C,D	0.4837	0.5470	15.6076	19.4975
A,B,C,D,E	0.6866	0.4226	16.4650	19.9925

TABLE III

AVERAGE TIME STEPS TAKEN TO DETECT TARGETS IN THE ROI USING THE SPECIFIED CAMERAS

Cameras	Hill Climbing Algorithm	S-ALUL		
		5	10	15
A	9.0308	9.0308	9.5343	9.0308
A,D	1.5776	1.5776	1.5866	1.5776
A,C,D	0.8864	0.7187	0.8864	0.8864
A,B,C,D	0.7782	0.5054	0.7882	0.7882
A,B,C,D,E	0.8816	0.5372	0.8816	0.8816

A. Hypothesis 1: Max-ALUL as a measure of uncovered area size, better approximates target acquisition performance than TVA

In looking at the resulting performance, we divide the configurations between those that provide more complete coverage (75% or greater) and more constrained coverage (less than 75%). To quantify Max-ALUL as a acquisition metric, static single and multi-camera configurations (detailed in Table I) were used for target acquisition and tracking. In Table II, both the results for target acquisition and target tracking time are shown.

TABLE IV

AVERAGE TIME STEPS ACQUIRED TARGETS ARE CONTINUOUSLY TRACKED INSIDE THE ROI USING THE SPECIFIED CAMERAS

Cameras	Hill Climbing	S-ALUL		
		5	10	15
A	4.1530	4.1530	1.7083	4.1530
A,D	10.5314	10.5314	10.4236	10.5314
A,C,D	13.1584	12.5170	13.1584	13.1584
A,B,C,D	13.6149	15.2586	13.6149	13.6149
A,B,C,D,E	13.6657	14.2393	13.6657	13.6657

Correlation between Max-ALUL and acquisition performance (measured by steps that targets are undetected) for configurations where 75% or more of the environment was visible was 0.68 compared to -0.27 for TVA. For relatively constrained configurations where less than 75% of the visible cells were visible, both metrics performed similarly with a correlation of 0.61 for TVA and 0.57 for Max-ALUL.

For camera configurations that are somewhat constrained (less than 75% of the area is visible), using Max-ALUL improved target acquisition and tracking performance. For all but one configuration, where more than 75% of the area is visible, using Max-ALUL improved acquisition and tracking but not significantly. In each of these cases, using Max-ALUL and TVA resulted in coverage percentages that were comparable (not more than 2% different).

B. Hypothesis 2: Minimizing the Max-ALUL statically improves target acquisition in resource constrained environments

The efficacy of Max-ALUL in intelligent coverage algorithms(Tables III and IV) was evaluated by comparing a coverage algorithm that uses only temporal utility to manage panning with the addition of Max-ALUL to disallow camera changes to create uncovered areas larger than a set threshold (5, 10, 15). Large S-ALUL values such as 15 created results close to those obtained without considering ALUL (larger areas minimally constrained movement). Midrange values, such as 10, did not show improvement in either target acquisition or tracking and in some cases caused slight performance degradation. Lower values, that constrained movement to a fairly small region which did not view all cells but kept uncovered areas small improved target acquisition but not significantly so. Interestingly enough, target tracking improved in the larger camera sets but this improvement could be specific to the targets random movements. ALUL for each time step of the experiments are depicted in Figures 2 - 5.

Previous results in [10] studied the use of additional ALUL values for targets (T-ALUL) to manage competing tracking and acquisition priorities. The inclusion of T-ALUL allowed for better acquisition and tracking performance when

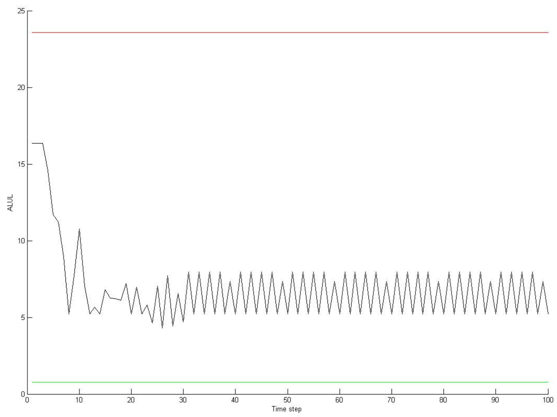

Fig. 2. ALUL by time step using Hill Climbing Algorithm with cameras A,C and D.

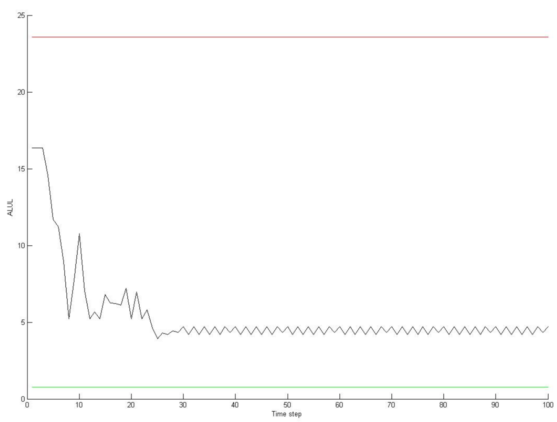

Fig. 3. ALUL by time step using S-ALUL=5 with cameras A,C and D.

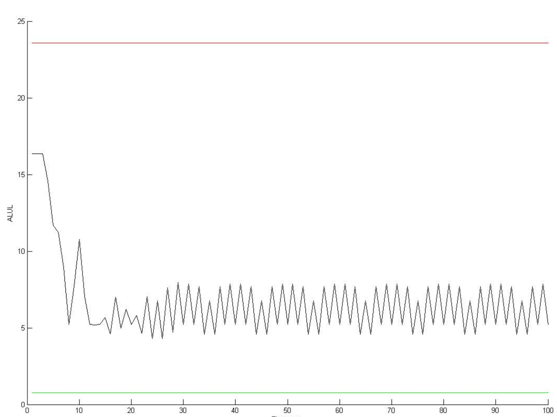

Fig. 4. ALUL by time step using Hill Climbing Algorithm using cameras A,B,C and D.

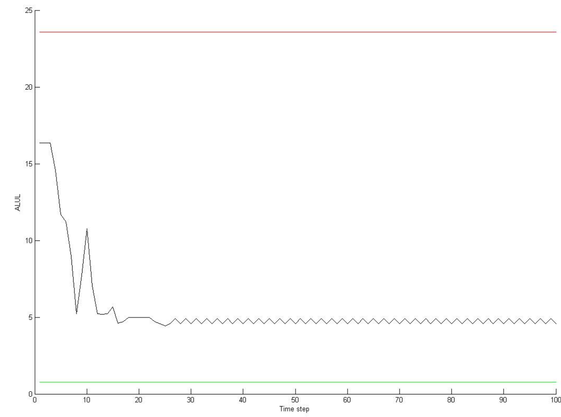

Fig. 5. ALUL by time step using S-ALUL=5 using cameras A,B,C and D.

resources were available and made managed trade-off when resources were constrained. In these experiments, tracking performance is reported. The addition of Max-ALUL does not adversely affect tracking performance when increasing acquisition. For a few isolated cases, tracking performance improved slightly.

VI. CONCLUSION

In this paper, the prediction of target acquisition performance in terms of the amount of time a target can move undetected within the ROI is explored. ALUL measures the average length of the rays through an area giving an indication of its effect on acquisition performance. This approach specifically targets existing camera deployments where the camera system is not optimized to view all areas or provide redundancy. For camera deployments and configurations where 75% or greater of the visible ROI is monitored, Max-ALUL was a better measure of acquisition performance. In cases where the system was more constrained, both TVA and Max-ALUL performed similarly. The ability to approximate the ability to acquire a target can be used to provide operator insight into system quality when manually intervening through redirecting system resources.

Experiments compared the use of ALUL in lieu of and in conjunction with more traditional temporal parameters that maximize TVA and minimize the time an area goes unseen. Experiments that utilize S-ALUL to constrain camera configurations to those that limit unmonitored areas can provide some improved tracking performance but not at a statistically significant level. Previous results have shown that using an ALUL specific to targets (T-ALUL) that allows the creation of larger unmonitored areas to track targets improves tracking performance without degrading acquisition when resources are available or prioritizes tracking over acquisition if the T-ALUL is large enough.

Future work includes extension to a distributed robotic system utilizing Unmanned Aerial Vehicles(UAVs) and Unmanned Ground Vehicles(UGVs) through formulation as a dynamic Constraint Satisfaction Problem. The current algorithm is centralized and assumes a connected networked camera system. Inclusion of UAVs and UGVs requires consideration of additional configuration parameters while eliminating the invisibility constraint.

REFERENCES

[1] S. Tilak, N. Abu-Ghazaleh, and W. Heinzelman, “A taxonomy of wireless micro-sensor network models,” *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 6, no. 2, pp. 28–36, 2002.

[2] R. Collins, A. Lipton, T. Kanade, H. Fujiyoshi, D. Duggins, Y. Tsin, D. Tolliver, N. Enomoto, O. Hasegawa, and T. R. Cmu Ri, “A system for video surveillance and monitoring,” Robotics Institute, Carnegie Mellon University, Report CMU-RI-TR-00-12, May 2000.

[3] J. Davis, A. Morison, and D. Woods, “An adaptive focus-of-attention model for video surveillance and monitoring,” 2006.

[4] S. Iacob and A. Sadden, “Attention and anticipation in complex scene analysis- an application to video surveillance,” in *Proc. IEEE Int’l Conf. on Systems, Man, and Cybernetics*, 2004.

[5] F. Z. Qureshi and D. Terzopoulos, “Surveillance camera scheduling: a virtual vision approach,” *VSSN ’05: Proceedings of the third ACM international workshop on Video surveillance & sensor networks*, vol. Hilton, Singapore, pp. 131–140, 2005. [Online]. Available: http://doi.acm.org/10.1145/1099396.1099420

[6] P. David, “Multiple sensor cueing using a best-first search,” in *Proceedings of SPIE, Applications of Artificial Intelligence IX*, 1991.

[7] C. Gui and P. Mohapatra, “Power conservation and quality of surveillance in target tracking sensor networks,” in *Proceedings of the 10th annual international conference on Mobile computing and networking*, U. ACM Press New York, NY, Ed., 2004, pp. 129–143.

[8] J. Everist, T. N. Mundhenk, C. Landauer, and K. Bellman, “Visual surveillance coverage: Strategies and metrics,” in *Intelligent Robots and Computer Vision XXIII: Algorithms, Techniques, and Active Vision*, D. Casasent, Ed., 2005.

[9] A. Mazzolo, “Properties of uniform random walks in bounded convex bodies,” *Journal of Physics A:Mathematical and General*, vol. 37, pp. 7095–7111, 2004.

[10] M. Anderson and P. David, “Role of quality of service metrics in visual target acquisition and tracking in resource constrained environments,”,” in *Proceedings of SPIE Defense and Security Symposium*, vol. 6562, 2007.