# LINEAR PRODUCTION GAME SOLUTION TO A PTZ CAMERA NETWORK

*Yu-Chun Lai*[1], *Yu-Ming Liang*[2], *Sheng-Wen Shih*[3], *Hong-Yuan Mark Liao*[14], *and Cheng-Chung Lin*[1]

[1]Department of Computer Science, National Chiao Tung University, Taiwan
[2]Department of Computer Science and Information Engineering, Aletheia University, Taiwan
[3]Department of CSIE, National Chi Nan University, Taiwan
[4]Institute of Information Science, Academia Sinica, Nankang,Taipei, Taiwan

## ABSTRACT

Reconfiguring the PTZ parameters of a camera network is an combinatorial optimization problem and computing the optimal solution is very time consuming. Therefore, existing methods can only provide sub-optimal solutions. In this paper, a nonlinear objective function for better utilizing the cameras to track multiple targets is proposed. Furthermore, it is shown that by expanding the unknown parameters and imposing new constraints, the nonlinear objective function can be converted into a linear production game (LPG) problem. Since an LPG possesses an optimal solution which can be evaluated with polynomial time, the proposed method is efficient and accurate. Computer simulations have been conducted and the results show that the proposed method is very promising.

***Index Terms***— Visual Surveillance, Camera Network, Pan-Tilt-Zoom Camera, Linear Production Game.

## 1. INTRODUCTION

Intelligent video surveillance systems have been around for decades and have been densely deployed at important places all over the world. While it has been shown that a single camera can provide useful information for event detection and target tracking [1, 2], a surveillance system usually consists of a camera network to reduce blind spots and to improve its reliability [3, 4, 5, 6, 7, 8, 9]. A camera network is usually composed of heterogeneous cameras including panoramic cameras, fixed cameras, infrared cameras, and pan-tilt-zoom (PTZ) cameras. Among the different types of imaging devices, PTZ cameras are the most important ones for an intelligent surveillance system, because they can change their field of views (FOVs) actively in response to different task requirements. However, introducing PTZ cameras into a surveillance system also brings in a challenging issue about how to control and to coordinate the cameras to accomplish a given task. Most of the surveillance tasks related to PTZ cameras are related to three functions namely tracking multiple targets, improving evidential quality, and maximizing the surveillance coverage as explained in the following.

Target tracking involves target detection and temporal and/or inter-view target correspondence matching. Lim et al. [3] track the targets observed in FOVs and build a dynamic scene model containing the position, the velocity, and the view-dependent visibility of each target. Their system comprises three modules solving three main tasks. Cameras are assigned to tasks by solving a bipartite matching problem attempting to ensure that tasks of higher priority will be accomplished first. In the PTZ camera network by Ukita and Matsuyama [4], each time when the system detects a new target, the closest idle camera will be assigned to track the target. This system is simple an effective provided that the number of cameras are greater than the number of targets. Qureshi and Terzopoulos [5] demonstrated a multi-camera tracking system. In their system, calibrated wide-FOV cameras are used to locate targets, and PTZ cameras are used to fixate on the located targets. Their PTZ network operates with heuristic rules developed to track targets cooperatively. In summary, a cooperative target tracking approach can not only efficiently utilize the camera resources but also allow cameras to support each other to recover an otherwise failed task.

While a solution to the target tracking problem can provide the trajectory of each target, a surveillance system usually demands more information. For example, it is frequently required to record the face of a human target or the license plate of a vehicle target with a sufficient resolution. These applications are related to the improvement of evidential quality [7, 8]. Additionally, PTZ cameras are frequently used to extend the coverage of the surveillance area by pan-tilt scanning. Piciarelli et al. [9] proposed an approach to reconfigure the pan-tilt-zoom parameters of all PTZ cameras according to a given probability map of observing an event at a specific location. Song et al. [6] applied the competitive game theory to maximize the surveillance coverage. They adopted a sequential optimization strategy to achieve the Nash equilibrium [10]. With their method, a PTZ camera is randomly selected at a time for tuning its parameters while keeping other cameras' parameters unchanged. When a Nash equilibrium is achieved, the cameras may cover the entire area

978-1-4244-7994-8/10/$26.00 ©2010 IEEE

at acceptable resolution. When a human operator chose to track a specific target at a higher resolution, the target will be assigned to a proper PTZ camera which will be excluded from the game. Therefore, the other cameras will adjust their parameters trying to maintain the maximum surveillance coverage. The main advantage of their method is that it can be easily implemented and the amount of information required to be exchanged is very low. However, it should be noted that the Nash equilibrium is not necessarily an optimal solution. Also, tracking of a specific target with higher resolution is treated as an exception task which cannot be optimized using the same game theory framework.

Reconfiguring PTZ cameras to achieve any of the above-mentioned three goals is intrinsically a combinatorial optimization problem and computing the optimal solution is very time consuming. Therefore, existing methods can only provide suboptimal solutions. In this paper, we propose an optimal and flexible solution to a PTZ network coordination problem. We show that the PTZ network problem can be formulated as a *linear production game* (LPG) problem [11]. The optimal solution of the PTZ network problem can be computed in polynomial time. The remainder of the paper is organized as follows. In Section 2, the PTZ network problem is formulated. Section 3 describes the proposed LPG method. Section 4 details the experiment results. Conclusions are given in Section 5.

## 2. PROBLEM FORMULATION

Suppose that there are $n$ calibrated PTZ cameras deployed in a region, each of which is controlled by a network-connected processor. Let $m$ denote the number of detected targets in the surveillance region. Each detected target is represented by a status vector denoted by $\mathbf{g}_t^k = \left[\mathbf{b}_t^k, \mathbf{v}_t^k\right]$, where $\mathbf{b}_t^k$ and $\mathbf{v}_t^k$ are the 3-D bounding box and the velocity of target $k$ estimated at time $t$. The target statuses $T_t = \left\{ \mathbf{g}_t^k \middle| k = 1, 2, ..., m\right\}$ and the static background scene model constitute a dynamic scene model which can be used to predict the statuses of all the $m$ targets expressed as $\hat{T}_{t+1} = \left\{ \hat{\mathbf{g}}_{t+1}^k \middle| k = 1, 2, ..., m\right\}$, at time $t+1$. The scene model is maintained by a central information processing node which gathers the information regarding to the detected targets and the camera parameters from each camera node. In order to integrate the detected target information, it is assumed that the camera network has been calibrated and the homography between any two of the cameras is known. The central information processing node is also in charge of determining the optimal camera parameters.

### 2.1. Parameters to be determined

Let $\phi^i$ denote the FOV of the $i$-th camera, which is controlled by the pan-tilt-zoom parameters of the camera. We assume that the relationship between the FOV and the pan-tilt-zoom parameters of a camera is known. Therefore, the problem of determining the optimal camera parameters is transformed into the problem of selecting the optimal FOV for each camera. Due to the limitation of the lens motor speed, a camera can only change its parameters locally in a short period of time. Therefore, given the current parameters of each camera, a set of feasible FOVs can be constructed, which is expressed as follows.

$$\Phi^i = \left\{ \phi_j^i \middle| j = 1, 2, ..., w_i \right\}, \tag{1}$$

for $i$ = 1, 2, ..., $n$, where $w_i$ is the number of feasible FOVs of the $i$-th camera. The PTZ camera coordination problem is formulated as the following combinatorial optimization problem.

$$\left(\phi^1, ..., \phi^n\right) = \arg \max_{\phi^i \in \Phi^i, i=1,...,n} Q\left(\phi^1, ..., \phi^n\right), \tag{2}$$

where $Q\left(.\right) : \Phi^1 \times \cdots \times \Phi^n \longmapsto \mathbb{R}$ is a function mapping $\left(\phi^1, ..., \phi^n\right)$ to a real quality value. In the next subsection, we will describe how to assess the quality of a set of FOVs under different goals.

### 2.2. Quality function of the camera fields of view

Given the dynamic scene model, a virtual image of the predicted bounding boxes of the targets can be computed for each camera FOV using a graphics card. The image region corresponding to an un-occluded bounding box is defined as a region of interest (ROI). For a visual surveillance system, assessing the quality of an FOV usually comprises the following two steps.

1. Examining whether the FOV includes some ROIs: an FOV contains no ROI should be evaluated to have the lowest quality.

2. Evaluating the dimension (width and height) of each ROI: the ROI should possess a sufficient resolution to accomplish the given task. Aldrige and Gilbert have suggested different resolution requirements for different tasks [12]. When the resolution is lower than the suggested value, a low quality value should be evaluated. Conversely, when the resolution is higher than the suggested value, the quality value should be upper bounded or be reduced to induce a camera zoom out for monitoring a larger area.

For most of the surveillance tasks, the quality of an FOV can be evaluated individually and the total quality function, $Q\left(\phi^1, ..., \phi^n\right)$, can be simplified as follows.

$$Q\left(\phi^1, ..., \phi^n\right) = f\left(q_1, q_2, ..., q_n\right), \tag{3}$$

where $q_i = Q\left(\phi^i\right)$ for $i$ = 1, ..., $n$, and $f(\cdot) : \mathbb{R}^n \longmapsto \mathbb{R}$ is a function that maps the $n$ individual quality values into a total quality value. Possible choices of $f(\cdot)$ will be discussed later. Furthermore, the quality function of each FOV, say $\phi^i$, can also be expressed as a function of the qualities of individual ROIs. Since the quality of each ROIs can be evaluated independently, it is reasonable to compute the quality of an FOV as follows.

$$Q\left(\phi^{i}\right)=\sum_{k=1}^{m} Q_{k}\left(\phi^{i}\right), \tag{4}$$

where $Q_{k}\left(\phi^{i}\right) \triangleq Q\left(\hat{\mathbf{b}}_{t+1}^{k} ; \phi^{i}, \hat{T}_{t+1}-\left\{\hat{\mathbf{g}}_{t+1}^{k}\right\}\right)$ is the non-negative quality of observing target $k$ with FOV $\phi^{i}$, which is zero when $\hat{\mathbf{b}}_{t+1}^{k}$ is not observable from the FOV $\phi^{i}$ or it is completely occluded by targets in $\hat{T}_{t+1}$ except for itself.

The simplest form of $f(.)$ in equation (3) is a linear summation function given by

$$Q\left(\phi^{1}, \ldots, \phi^{n}\right)=\sum_{i=1}^{n} Q\left(\phi^{i}\right). \tag{5}$$

However, the drawback of the above equation is that an eye-catching target may attract too many cameras causing unattended targets. To overcome this drawback, the following nonlinear quality function is adopted in this work.

$$Q\left(\phi^{1}, \ldots, \phi^{n}\right)=\sum_{k=1}^{m} \max_{i} Q_{k}\left(\phi^{i}\right), \tag{6}$$

where only the maximum ROI quality of each target counts toward the total quality. Thus, the quality of a solution biased on a specific target will be lower than that of a solution apportions the cameras to monitor different targets.

## 3. THE PROPOSED APPROACH

In this section, we will show that the nonlinear objective function (6) can be converted to a linear function by expanding the set of feasible solutions and imposing new constraints.

Let $T_{j}^{i}$ and $\left|T_{j}^{i}\right|$ denote the set of targets covered the FOV $\phi_{j}^{i}$ and the number of targets in $T_{j}^{i}$, respectively. The total number of the subsets of $T_{j}^{i}$ is $2^{\left|T_{j}^{i}\right|}$. For each subset $S_{j,h}^{i} \subset T_{j}^{i}$, $1 \leq h \leq 2^{\left|T_{j}^{i}\right|}$, a virtual FOV, $\phi_{j,h}^{i}$ can be constructed which ignores any target not in $S_{j,h}^{i}$, i.e., $Q_{k}\left(\phi_{j,h}^{i}\right)=0$, for all $k \notin S_{j,h}^{i}$.

By replacing $\phi_{j}^{i}$ with the virtual FOVs, $\phi_{j,h}^{i}$, $h=1, \ldots, 2^{\left|T_{j}^{i}\right|}$, the number of feasible FOVs to be assigned to the $i$-th camera become $\sum_{j=1}^{w_{i}} 2^{\left|T_{j}^{i}\right|}$. To select the optimal FOVs, we define binary variables $x_{j,h}^{i} \in \mathbb{B}$ ($\mathbb{B} \triangleq \{0,1\}$) to denote the status whether the $(j, h)$-th virtual FOV of the $i$-th camera, i.e., $\phi_{j,h}^{i}$, is selected as the optimal FOV. Therefore, the optimization problem of FOV selection can be rewritten as follows.

$$\max_{\mathbf{x}} \sum_{i=1}^{n} \sum_{j=1}^{w_{i}} x_{j,h}^{i} \sum_{k=1}^{m} Q_{k}\left(\phi_{j,h}^{i}\right), \tag{7}$$

subject to

$$\sum_{j=1}^{w_{i}} \sum_{h=1}^{2^{\left|T_{j}^{i}\right|}} x_{j,h}^{i} \leq 1, \tag{8}$$

for $i=1,2, \ldots, n$, and

$$\sum_{i=1}^{n} \sum_{j=1}^{w_{i}} \sum_{h=1}^{2^{\left|T_{j}^{i}\right|}} x_{j,h}^{i}\, o_{ijhk} \leq 1, \tag{9}$$

for $k=1,2, \ldots, m$, where $\mathbf{x}=\left[\cdots x_{j,h}^{i} \cdots\right]$ and the following binary coefficient, $o_{ijhk} \in \mathbb{B}$, indicating whether target $k$ is observable in the virtual FOV $\phi_{j,h}^{i}$. The constraint specified in (8) ensures that each camera can only be assigned one FOV at a time, whereas (9) guarantees that the quality value of a target can only be evaluated with a single FOV.

Since the coefficients in the objective function (7) and in the constraints (8) and (9) are all non-negative, the optimization problem is equivalent to an LPG [11]. In an LPG, all players (cameras) are coordinated to enforce a cooperative behavior in a cooperative game. An LPG has the following two very attracting properties.

1.The optimal solution always exists [11].

2.The optimal solution can be computed in polynomial time [13].

In this work, the branch-and-cut algorithm [14] is adopted to compute the optimal integer solution. The relation between the solutions to (7) and to (6) can be reveal by changing the summation order of (7) as follows.

$$\sum_{k=1}^{m}\left[\sum_{i=1}^{n} \sum_{j=1}^{w_{i}} x_{j,h}^{i}\, Q_{k}\left(\phi_{j,h}^{i}\right)\right]=\sum_{k=1}^{m} q_{k}^{*}, \tag{10}$$

where $q_{k}^{*}$ is the quality of target $k$ evaluated at one of the optimal FOVs. Since an optimal solution guaranteesc that (10) is an upper bound of (6), the LPG solution is equivalent to the optimal solution of the nonlinear objective function (6).

## 4. EXPERIMENTS

In order to compare the performances of different approaches, we implemented a simulation system which contains three virtual PTZ cameras and many manually generated pedestrians with different paths and moving speeds. Figure 1 shows the simulation environment where the whole coverage of each PTZ is depicted as a fan-shaped region. In addition to the method proposed in Section 3 (nonlinear objective function LPG, abbreviated as NOF-LPG), we have also implemented the linear sum method (linear objective funcion LPG, abbreviated as LOF-LPG), which maximizes (5), and the method proposed in [6] (abbreviated as SONG) for comparison. The simulation system is implemented at a PC with a CPU of Intel core 2 DUO E8400 3.00Ghz. For each time instance, it takes about 220 ms to process the image data and to update the scene model. If the computation load is distributed to different nodes, the preprocessing time will be reduced to approximately 75 ms. Additionally, computing the PTZ parameters using NOF-LPG, LOF-LPG and SONG takes 13 ms, 2.2 ms and 0.5 ms, respectively. Therefore, the frame rate in the computer simulation is set to 10. Notably, the number of targets in each view, i.e., $\left|T_j^i\right|$, determines the computation load of NOF-LPG. However, the number of targets in an FOV is usually small. In the simulation, the maximum number of targets in an FOV is only five. Therefore, although the computation time of the proposed NOF-LPG method is much longer than that of either LOF-LPG or SONG.

Figure 2 shows the performance of the three methods, where the ground truth data is computed by counting the number of targets inside any of the three fan-shaped regions. In this experiment, the LOF-LPG method outperformed the SONG method because LOF-LPG utilizes a central optimization method. Furthermore, the proposed NOF-LPG method outperforms the other two methods because NOF-LPG chooses a better objective function.

## 5. CONCLUSIONS

In this paper, the process of assessing the quality of a set of FOVs is described and a nonlinear objective function for reducing the number of unattended targets in a surveillance region is developed. Furthermore, we have shown that the nonlinear optimization problem can be converted into a linear production game problem which is guaranteed to have a optimal solution. The branch-and-cut method has been adopted to solve the PTZ parameters selection problem. Computer simulations have been conducted to test the proposed method against the other two methods and the results show that the proposed method has achieved that highest tracking rate.

## 6. ACKNOWLEDGMENTS

This work is supported in part by the National Science Council, Taiwan, under Grant No. NSC 98-2221-E-260-022-MY3 and by Ministry of Economic Affairs, Taiwan, under Grant No. 98-EC-17-A-02-S1-032.

## 7. REFERENCES

[1] I. Haritaoglu, D. Harwood, and L. S. Davis, "W4: real-time surveillance of people and their activities," *IEEE T-PAMI*, vol. 22, no. 8, pp. 809–830, 2000.

[2] Omar Javed and Mubrark Shah, "Tracking and object classification for automated surveillance," in *Proceedings of ECCV*, 2002, vol. 2423, pp. 343–357.

[3] Ser-Nam Lim, L. S. Davis, and A. Elgammal, "A scalable image-based multi-camera visual surveillance system," in *Proceedings of AVSS*, 2003, pp. 205–212.

[4] Norimichi Ukita and Takashi Matsuyama, "Real-time cooperative multi-target tracking by communicating active vision agents," *CVIU*, vol. 97, no. 2, pp. 137–179, 2005.

[5] Faisal Z. Qureshi and Demetri Terzopoulos, "Surveillance in virtual reality: System design and multi-camera control," in *Proceedings of CVPR*, 2007, pp. 1–8.

[6] Bi Song, Cristian Soto, Amit K. Roy-Chowdhury, and Jay A. Farrell, "Decentralized camera network control using game theory," in *Proceedings of ICDSC*, 2008, pp. 1–8.

[7] A. D. Bagdanov, A. D. Bimbo, and W. Nunziati, "Improving evidential quality of surveillance imagery through active face tracking," in *Proceedings of ICPR*, 2006, vol. 3, pp. 1200–1203.

[8] Nicola Bellotto, Eric Sommerlade, Ben Benfold, Charles Bibby, Ian Reid, Daniel Roth, Luc Van Gool, Carles Fernandez, and Jordi Gonzalez, "A distributed camera system for multi-resolution surveillance," in *Proceedings of ICDSC*, 2009, pp. 1–8.

[9] C. Piciarelli, C. Micheloni, and G.L. Foresti, "PTZ camera network reconfiguration," in *Procedings of ICDSC*, 2009, pp. 1–7.

[10] Eric Rasmusen, *Games and Information - An Introduction to Game Theory*, Wiley-blackwell, 1991.

[11] Guillermo Owen, "On the core of linear production games," in *Mathematical Programming*, 1975, vol. 9, pp. 358–370.

[12] J. Aldrige and C. Gilbert, *Performance Testing CCTV Perimeter Surveillance Systems*, Number 14–95. White Crescent Press, Home Office Police Scientific Development Branch, Woodcock Hill, Sandridge, St Albans, Herfordshire, UK, 1996.

[13] Jean Derks and Jeroen Kuipers, "On the core of routing games," *Int'l J. of Game Theory*, vol. 26, no. 2, pp. 193–205, 1997.

[14] J. Mitchell, *Branch-and-cut algorithms for combinatorial optimization problems in Handbook of Applied Optimization*, Oxford University Press, 2002.

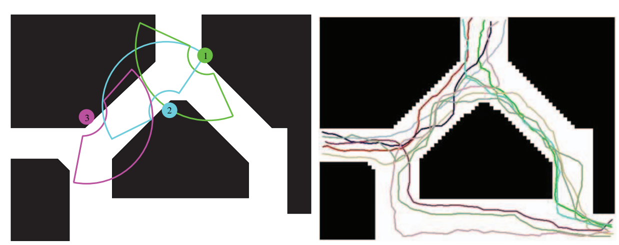

**Fig. 1**. Simulation surveillance environment with three PTZ cameras and computer generated random targets.

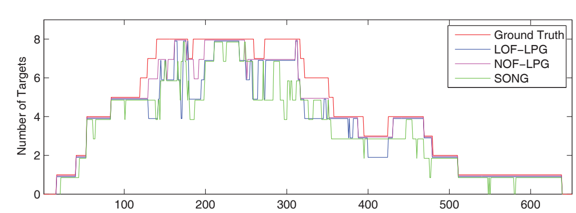


**Fig. 2**. Tracking results of different methods.