Utility-Based Camera Assignment in a Video Network: A Game Theoretic Framework

Yiming Li, *Student Member, IEEE*, and Bir Bhanu, *Fellow, IEEE*

***Abstract*—In this paper, an approach for camera assignment and handoff in a video network based on a set of user-supplied criteria is proposed. The approach is based on game theory, where bargaining mechanisms are considered for collaborations as well as for resolving conflicts among the available cameras. Camera utilities and person utilities are computed based on a set of user-supplied criteria, which are used in the process of developing the bargaining mechanisms. Different criteria and their combination are compared with each other to understand their effect on camera assignment. Experiments for multicamera multiperson cases are provided to corroborate the proposed approach. Intuitive evaluation measures are used to evaluate the performance of the system in real-world scenarios. The proposed approach is also compared with two recent approaches based on different principles. The experimental results show that the proposed approach is computationally more efficient, more robust and more flexible in dealing with the user-supplied criteria.**

***Index Terms*—Bargaining mechanism, camera handoff, dynamic camera selection.**

I. Introduction

DUE to the broad coverage of an environment and the possibility of coordination among different cameras, video sensor networks have attracted much interest in recent years. Although the field-of-view (FOV) of a single camera is limited and cameras may have overlapping/nonoverlapping FOVs, seamless tracking of moving objects is desired. This requires two capabilities: camera assignment and camera handoff, which are the subjects of this paper. We define camera assignment as a camera-object map, which tells us at each time instant which camera is being used to follow which object. Camera handoff is a dynamic process that the system transfers the right of tracking an object from one camera to another without losing the object in the network. The availability of camera handoff capability will provide the much needed situation assessment of the environment under surveillance. It is clear that the manual camera handoff will become unmanageable when the number of camera is large. In addition, it is unrealistic to display and manually monitor the surveillance videos captured from a large number of cameras simultaneously. Therefore, we need to develop surveillance systems that can automatically perform the camera assignment and handoff tasks, and then adapt to the appropriate video streams available from the currently used cameras.

In this paper, we provide a new perspective to the camera handoff problem based on game theory. The merit of our approach is that it is independent of the camera topology. When multiple cameras are used for tracking and where multiple cameras can "see" the same object, the algorithm can automatically provide an *optimal* as well as *stable* solution of the camera assignment. Since game theoretic approach allows dealing with multiple criteria optimization, we are able to choose the "best" camera based on multiple criteria that are selected *a priori*. The detailed camera calibration or 3D scene understanding is not needed in our approach.

In the rest of this paper, Section II describes the related work and contributions of this work. Section III formulates the camera assignment and handoff problem and then constructs the utilities and bargaining steps. Section IV discusses the implementation of this approach and shows the experimental results. Comparison results of the proposed approach with two recent approaches are also shown. Finally, Section V concludes this paper.

II. Related Work and Our Contribution

A. Related Work

There have been many papers discussing approaches for doing camera assignments in a video network. The traditional approaches generally fall into two categories: topology-based and statistics-based. The approaches belonging to the first category [1]–[4] rely on the geometrical relationships among cameras. These relationships tend to become quite complicated when the topology becomes complex and it is difficult to learn the topology based on the random traffic patterns [5]. The approaches belonging to the second category [6]–[10] usually depend on the objects' trajectories, while other factors such as orientation, shape, face, etc., which are also very important for visual surveillance, are not considered.

A comparison of related work with the proposed approach is summarized in Table I.

B. Contributions of This Paper

Our approach differs from the conventional approaches [1]–[4], [6]–[8], [12], [14], shown in Table I, in the following key aspects.

1) ***Game Theoretic Approach***: We propose a game theoretic approach for camera assignment and handoff

Manuscript received March 26, 2010; accepted May 03, 2010. Date of publication June 10, 2010; date of current version January 26, 2011. This work was supported in part by National Science Foundation (NSF) Grant 0551741, Grant 0622176, Grant 0905671, and ONR grants (DoD Instrumentation and the Aware Building). The contents of the information do not reflect the position or policy of the U.S. Government. The associate editor coordinating the review of this paper and approving it for publication was Dr. Subhas Mukhopadhyay.

The authors are with the Center for Research in Intelligent Systems, Department of Electrical Engineering, University of California at Riverside, Riverside, CA 92521 USA (e-mail: yimli@vislab.ee.ucr.edu; bhanu@cris.ucr.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSEN.2010.2051148

TABLE I
COMPARISON OF THIS PAPER WITH THE RELATED WORK

Author	Approach	Comments
Javed *et al.* [1]	Find the limits of overlapping FOVs of multiple cameras. Cameras are chosen based on the distance between the person and the edge of FOV.	Significant increase in computation with the increase in the number of cameras and persons.
Park *et al.* [2]	Create distributed look-up tables according to how well the cameras can image a specific location.	Based only on single criterion; depends on the information that may not be available for each network.
Jo & Han [3]	Construct a handoff function by computing the co-occurrence to occurrence ratio for selected pairs of points in the FOVs of two cameras.	Computation cost is high for increased handoff resolution. Handoff ambiguity and failure will arise when there are a large number of cameras or persons.
Qureshi & Terzopoulos [4]	Models conflicts that may arise during camera assignment as a constraints satisfaction problem; each camera assignment that passes the hard constraints is assigned a weight. The camera assignment with the highest weight is selected.	A leader node is required. One has to calculate all the solutions that can pass the constraints and select the best one. The ranking procedure can be time consuming as the problem becomes more complicated.
Kettnaker & Zabih [6]	Choose the object trajectory with the highest posterior probability.	Camera assignment is conducted based on the paths of the objects. Frontal view can be lost even when it is available.
Chang & Gong [7]	Use Bayesian modality fusion. A Bayesian network is used to combine multiple modalities for matching subjects among multiple cameras.	Geometry/recognition based, hard to combine the two together for a decision; mainly focuses on the occlusion problem, not other criteria for camera assignment.
Kim *et al.* [8]	Find a dominant camera for an object based on the ratio of the numbers of blocks on the object to that in the camera FOV.	The dominant camera has to be aided by views from other cameras. Only size and angle are taken into account.
Cai & Aggarwal [12]	Matches human objects by using multivariate Gaussian models.	Camera calibration is needed.
Tessens *et al.* [14]	Messages related to a distributed process are sent to a base station to determine the principal camera based on a score which is decided by experiments offline.	Use both distributed and centralized control. The criteria are determined offline so it is not suitable for active camera selection and control. The principal view is complemented by helper views.
This paper	Models the camera assignment problem as a potential game using the vehicle-target model. Uses bargaining mechanism to obtain a converged solution with a small number of iterations.	Can decide the camera assignment based on multiple criteria; no camera calibration is needed; independent of the camera network topology.

problem using the vehicle-target model [13]. We model the problem as a multiplayer *potential game* and allow for both coordination and conflicts among the players.

2) ***Multiple Criteria for Tracking***: Multiple criteria are used in the design of *utility* functions for the objects being tracked. The equilibrium of the game provides the solution of the camera assignment. The bargaining mechanism makes sure that we can get a stable solution, which is optimal or near optimal, after only a small number of iterations [13].
3) ***"Best" Camera Selection***: We do not use the traditional master-slave system [14]. Instead, by selecting the "best" camera, we can have a good enough view, based on the user-supplied criteria, for observation of some specific target and simultaneously free the other cameras in the network for other tasks. Thus, the system can perform the tracking task with a minimum number of cameras, or, can perform more tasks with the same number of cameras.
4) ***Experimental Results***: Unlike some of the previous work [4], we evaluate the proposed approach in the context of real network using real data and show promising results.

III. Technical Approach

A. Motivation and Problem Formulation

Game theory is well known for analyzing the interactions as well as conflicts among multiple agents [15], [16]. Analogously, in a video sensor network, collaborations as well as competitions among cameras exist simultaneously. The cooperation lies in the fact that all the available cameras, those which can "see" the target person, have to collaborate to track the person so that the person can be followed as long as possible. On the other hand, the available cameras also compete with each other for the rights of tracking this person, so that a camera can maximize its own *utility*, as a camera's utility is closely related to how well it can track a person. This enlightens us to view the camera assignment problem in a game theoretic manner. A game is the interactive process [17] among all the participants (*players*) of a game, who strive to maximize their *utilities*. The utility of a player refers to the welfare that the players can get in the game. In our problem, for each person to be tracked, there exists a multiplayer game, with the available cameras being the players. If there are multiple persons in the system, this becomes a multiple of multiplayer game being played simultaneously [18].

Vehicle-target assignment [13] is a classical multiplayer game that aims to allocate a set of vehicles to a group of targets and achieves an optimal assignment. Viewing the persons being tracked as "vehicles" while the cameras as "targets," we can adopt the vehicle-target assignment model to choose the "best" camera for each person. In the following, we propose a game theory based approach that is well suited to the task at hand.

B. Game Theoretic Framework

Game theory involves *utility*, the amount of "welfare" an agent derives in a game. We are concerned with three utilities: 1) *Global utility*: the overall degree of satisfaction for tracking performance. 2) *Camera utility*: how well a camera is tracking persons assigned to it. 3) *Person utility*: how well a person is satisfied while being tracked by some camera. Our objective is to maximize the global utility while making sure that each person is tracked by the "best" camera. When competing with

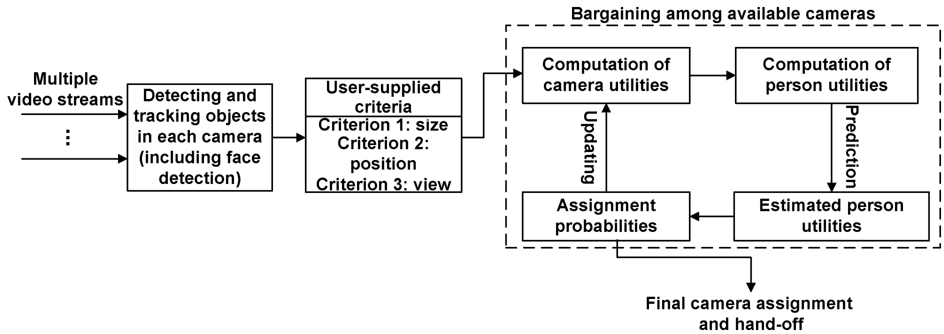


Fig. 1. Game theoretic framework for camera assignment and handoff.

TABLE II
NOTATIONS OF SYMBOLS USED IN THIS PAPER

Symbols	Notations
P_i	Person i
C_j	Camera j
N_P	Total number of persons in the entire network at a given time
N_C	Total number of cameras in the network at a given time
A_i	The set of cameras that can see P_i, $A_i = \{A_i^1, A_i^2, \ldots, A_i^{n_C}\}$
n_C	Number of cameras that can *see* person i, number of elements in A_i
n_P	Number of persons currently assigned to camera C_j
a_i	The currently assigned "best" camera for person i
a_{-i}	The assignment of cameras for the persons excluding P_i
a	The assignment of cameras for all persons, $a = (a_i, a_{-i})$
$U_{C_j}(a)$	Camera utility for camera j
$U_{P_i}(a)$	Person utility for person i
$U_g(a)$	Global utility
$\overline{U}_{P_i}(k)$	Predicted person utility for person i at step k, $\overline{U}_{P_i}(k) = \left[\overline{U}_{P_i}^1(k), \ldots, \overline{U}_{P_i}^l(k), \ldots, \overline{U}_{P_i}^{n_C}(k)\right]^T$, where $\overline{U}_{P_i}^l(k)$ is the predicted person utility for P_i if camera a_l is used
$p_i(k)$	Probability of person i's assignment at step k, $p_i(k) = \left[p_i^1(k), \ldots, p_i^l(k), \ldots, p_i^{n_C}(k)\right]^T$, where $p_i^l(k)$ is the probability for camera a_l to track person P_i

other available cameras, the cameras *bargain* with each other. Finally, a decision is made for the camera assignment based on a set of probabilities.

An overview of the approach is illustrated in Fig. 1. Moving objects are detected in multiple video streams. Their properties, such as the size of the minimum bounding rectangle and other region properties (color, shape, location within FOV, etc.) are computed. Various utilities are calculated based on the user-supplied criteria and bargaining processes among available cameras are executed based on the prediction of person utilities from the previous iteration step. The results obtained from the strategy execution are, in turn, used for updating the camera utilities and the person utilities until the strategies converge. Finally, those cameras with the highest converged probabilities are used for tracking. This assignment of persons to the "best" cameras leads to the solution of the handoff problem in multiple video streams. A set of key symbols and their notations used in the following discussion are given in Table II.

1) Computation of Utilities: We define the following properties of our system.

1) A person P_i can be in the FOV of multiple cameras. The available cameras for P_i belong to the set A_i. C_0 is a virtual camera that does not actually exist. We assume a virtual camera C_0 is assigned to P_i when there is no real camera in the network available to track P_i.
2) A person can only be assigned to one camera. The assigned camera for P_i is named as a_i.
3) Each camera can be used for tracking multiple persons.

We use a to denote the camera assignment for all the persons, and a_i denotes the assigned camera for P_i. For P_i, when we change the camera assignment from a_i' to a_i'' while assignments for other persons remain the same, if we have

$$U_{P_i} = (a_i', a_{-i}) < U_{P_i}\,(a_i'', a_{-i}) \Leftrightarrow U_g\,(a_i', a_{-i}) < U_g\,(a_i'', a_{-i}) \tag{1}$$

the person utility U_{P_i} is said to be aligned with the global utility U_g, where a_{-i} stands for the assignments for persons other than P_i, i.e., $a_{-i} = (a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_{N_P})$. So, the camera assignment result a can also be expressed as $a = (a_i, a_{-i})$. We define the global utility as

$$U_g(a) = \sum_{C_j \in C} U_{C_j}(a) \tag{2}$$

where $\mathrm{U_{C_j}(a)}$ is the camera utility and defined to be the utility generated by all the engagements of persons with a particular camera $\mathrm{C_j}$.

Now, we define the person utility as

$$\begin{aligned} U_{P_i}(a) &= U_g(a_i, a_{-i}) - U_g(C_0, a_{-i}) \\ &= U_{C_j}(a_i, a_{-i}) - U_{C_j}(C_0, a_{-i}) \end{aligned} \tag{3}$$

where C_0 is a virtual camera. The person utility $U_{P_i}(a)$ can be viewed as a marginal contribution of P_i to the global utility. To calculate (3), we have to construct a scheme to calculate the camera utility $U_{C_j}(a)$. We assume that there are N_{Crt} criteria to evaluate the quality of a camera used for tracking an object. Thus, the camera utility can be built as

$$U_{C_j}(a_i, a_{-i}) = \sum_{s=1}^{n_P} \sum_{l=1}^{N_{Crt}} Crt_{sl} \tag{4}$$

where n_p is the number of persons that are currently assigned to camera C_j for tracking and Crt are the criteria that are supplied by the user. Plugging (4) into (3), we can obtain

$$U_{P_i}(a_i, a_{-i}) = \sum_{l=1}^{N_{Crt}} \left(\sum_{s=1}^{n_P} Crt_{sl} - \sum_{\substack{s=1 \\ s \neq P_i}}^{n_P} Crt_{sl} \right) \quad (5)$$

where $s \neq P_i$ means that we exclude person P_i from the those who are being tracked by Camera C_j. One thing to be noticed here is that when designing the criteria, we have to normalize them. Besides this requirement, it does not matter what kind of criteria is used to be fed into the bargaining mechanism which is discussed below.

2) Criteria for Camera Assignment and Handoff: The choice of a criterion to be used for camera assignment and handoff depends on the users' requirements. There might be different criteria for different applications, such as criteria for power consumption, time delay, image resolution, etc. The camera assignment results may change due to applying different criteria. Our goal is to find the proper camera assignment solution quickly based on whatever criteria are supplied by the user. In the following, we provide four criteria, which include human biometrics, which can be used for camera assignment and handoff.

- *Criterion 1: The **size** of the tracked person*. It is measured by the ratio of the number of pixels inside the bounding box of the person to the size of the image. That is

$$r = \frac{\#\ of\ pixels\ inside\ the\ bounding\ box}{\#\ of\ pixels\ in\ the\ image\ plane}.$$

  Here, we assume that neither a too large nor a too small object is convenient for observation. Assume that λ is the threshold for best observation, i.e., when $r = \lambda$ this criterion reaches its optimal value.

$$Crt_{s1} = \begin{cases} \frac{1}{\lambda} r, & \text{when } r < \lambda \\ \frac{1-r}{1-\lambda}, & \text{when } r \geq \lambda \end{cases} \quad (6)$$

  where $\lambda \in (0, 1)$ is defined as the optimal ratio of the size of the minimum bounding box for the human body to the size of the image. These two sizes can be obtained by reading the coordinates of the bounding box and the size of the image. λ is dependent on the orientation of the camera and the location of a region-of-interest (ROI) in the image plane. Only in extreme rare situations a ROI will have a minimum width of one pixel. The value of λ remains valid at all times. Because of these reasons we do not do any camera calibration to find its extrinsic or intrinsic parameters. An example for the function Crt_{s1} when $\lambda = 1/15$ is shown in Fig. 2 as an illustration.
- *Criterion 2: The **position** of a person in the FOV of a camera*. It is measured by the Euclidean distance that a person is away from the center of the image plane

$$Crt_{s2} = \frac{\sqrt{(x - x_c)^2 + (y - y_c)^2}}{\frac{1}{2}\sqrt{x_c^2 + y_c^2}} \quad (7)$$

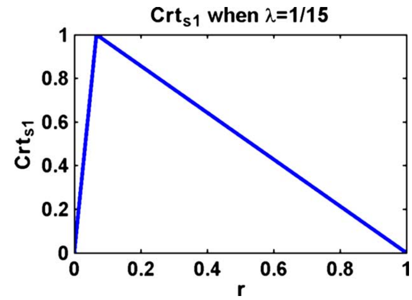


Fig. 2. Function of Crt_{s1} when $\lambda = 1/15$.

  where (x, y) is the current position (body centroid) of the person and (x_c, y_c) is the center of the image.
- *Criterion 3: The **view** of a person*. It is measured by the ratio of the number of pixels on the detected face to that of the whole bounding box. That is

$$R = \frac{\#\ of\ pixels\ on\ the\ face}{\#\ of\ pixels\ on\ the\ entire\ body}$$

  We assume that the threshold for the best frontal view is ξ, i.e., when $R = \xi (\xi \in (0, 1))$ the view of the person is the best

$$Crt_{s3} = \begin{cases} \frac{1}{\xi} r, & when\ R < \xi \\ \frac{1-R}{1-\xi}, & when\ R \geq \xi \end{cases}. \quad (8)$$

- *Criteirion 4: **Combination** of criterion (1), (2) and (3)*. It is given by the following equation,

$$Crt_{s4} = \sum_{l=1}^{3} w_l Crt_{sl} \quad (9)$$

  where w_l is the weight for different criterion.

It is to be noticed that all these criteria are appropriately normalized for calculating the corresponding camera utilities.

3) Bargaining Among Cameras: As stated previously, our goal is to optimize each person's utility as well as the global utility. Competition among cameras finally leads to the Nash equilibrium [21], as the solution of the camera assignment and handoff. Unfortunately, this Nash equilibrium may not be unique. Some of the solutions may not stable, which are not desired. To solve this problem, a bargaining mechanism among cameras is introduced, to make these cameras finally come to a compromise and generate a stable solution.

When bargaining, the assignment in the k^{th} step is made according to a set of probabilities

$$p_i(k) = \left[p_i^1(k), \ldots, p_i^l(k), \ldots, p_i^{n_C}(k)\right]^T$$

where n_C is the number of cameras that can "see" the person P_i and $\sum_{l=1}^{n_C} p_i^l(k) = 1$, with each $0 \leq p_i^l(k) \leq 1, 1, \ldots, n_C$. We can generalize $p_i(k)$ to be

$$p_i(k) = \left[p_i^1(k), \ldots, p_i^l(k), \ldots, p_i^{N_C}(k)\right]^T$$

by assigning a zero probability for those cameras which cannot "see" the person P_i, meaning that those cameras will not be assigned according to their probability. Thus, we can construct an $N_C \times N_p$ probability matrix

$$\begin{bmatrix} p_1^1(k) & \dots & p_{N_P}^1(k) \\ \vdots & \ddots & \vdots \\ p_1^{N_C}(k) & \dots & p_{N_P}^{N_C}(k) \end{bmatrix}.$$

At each bargaining step, we will assign a person to the camera which has the highest probability. We assume that one camera has no information of other cameras' utilities at the current step, which makes it hard to calculate all the possible current person utilities. So, we introduce the concept of predicted person utility $\overline{U}_{P_i}(k)$: Before we decide the final assignment profile, we predict the person utility using the previous person's utility information in the bargaining steps. As shown in (5), person utility depends on the camera utility, so, we predict the person utility for every possible camera that may be assigned to track it. Each element in $\overline{U}_{P_i}(k)$ is calculated by (10)

$$\overline{U}^l_{P_i}(k+1) = \begin{cases} \overline{U}^l_{P_i}(k) + \frac{1}{p_i^l(k)} \\ \quad \times \left(U_{P_i}(a(k)) - \overline{U}^l_{P_i}(k) \right), & a_i(k) = A_i^l \\ \overline{U}^l_{P_i}(k), & \text{otherwise} \end{cases} \tag{10}$$

with the initial state $\overline{U}^l_{P_i}(1)$ to be assigned arbitrarily as long as it is within the reasonable range for $\overline{U}_{P_i}(k)$, for $l = 1, \dots, n_C$. For the symbols used in (10), note that A_i^l is the l^{th} camera that is in the set of available cameras for person P_i, which is different from C_l, the l^{th} camera in the system. C_l can be in more than one available camera sets for different persons, while A_i^l is the l^{th} component in A_i, the set of available cameras for person P_i. It means that A_i^l is unique in the set for person P_i. Once these predicted person utilities are calculated, it can be proved that the equilibrium for the strategies lies in the probability distribution that maximizes its perturbed predicted utility [10]

$$p_i(k)'\overline{U}_{P_i}(k) + \tau H\left(p_i(k)\right) \tag{11}$$

where

$$H\left(p_i(k)\right) = -p_i(k)' log\left(p_i(k)\right) \tag{12}$$

is the entropy function and τ is a positive parameter belonging to [0,1] that controls the extent of randomization, where log means taking the log of every element of the column vector $p_i(k)$ and resulting in a column vector. The larger τ is, the faster the bargaining process converges; the smaller the τ is, the more accurate result we can get. So, there is a tradeoff when selecting the value of τ. We select τ, empirically, as 0.5 in our experiments. The solution of (11) is proved [10] to be

$$p_i^l(k) = \frac{e^{\frac{1}{\tau}\overline{U}^l_{P_i}(k)}}{e^{\frac{1}{\tau}\overline{U}^1_{P_i}(k)} + \cdots + e^{\frac{1}{\tau}\overline{U}^{n_C}_{P_i}(k)}}. \tag{13}$$

After several steps of calculation, the result of $p_i(k)$ tends to converge. Thus, we finally get the stable solution, which is proved to be at least suboptimal [13].

4) Game Theoretic Algorithm: This overall algorithm is summarized in Algorithm 1.

Algorithm 1: ***Game theoretic camera assignment and handoff***

Input: Multiple video streams.

Output: A probability matrix for camera assignments are made.

Algorithm Description:

- At a given time, perform motion detection and get the selected properties for each person that is to be tracked.
- For each person and each camera, decide which cameras can "see" a given person P_i.
- For those which can "see" the person P_i, initialized the predicted person utility vector $\overline{U}_{P_i}(1)$.

Repeat

1. Compute the Crt_{sl} for each available camera.
2. Compute the camera utilities $U_{C_j}(a)$ by (4).
3. Compute the person utilities $U_{P_i}(a)$ by (5).
4. Compute the predicted person utilities $\overline{U}_{P_i}(k)$ by (10).
5. Derive the strategy by $p_i(k)$ using (13).

Until The strategies for camera assignments converge.

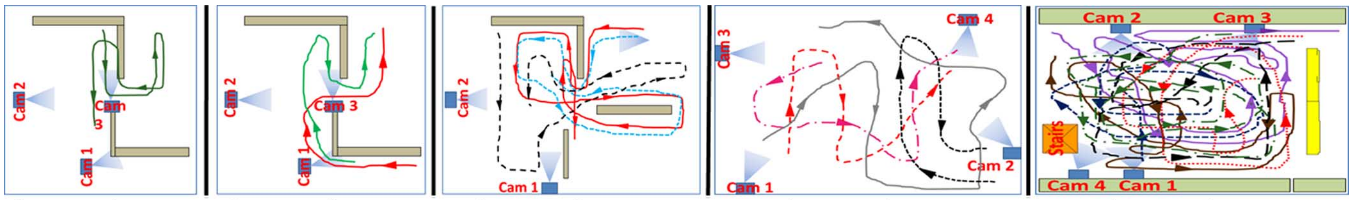
- Do camera assignment and handoff based on the converged strategies.

The bargaining mechanism and the criteria are tightly integrated in the proposed game theoretic approach. The bargaining process is based on a set of criteria, since the utilities used to update in each bargaining step are calculated using these criteria. Note that different criteria imply different emphasis and the definition of error (see Section IV-C) depends on them.

IV. Experimental Results

A. Data and Parameters

1) Data: In our experiments, we tested the proposed approach on five cases: (1) 3 cameras, 1 person; (2) 3 cameras, 2 persons; (3) 2 cameras, 3 persons; (4) 4 cameras, 4 persons; and (5) 4 cameras, 6 persons. These experiments include from the simple case, 3 cameras, 1 person, to a complicated case, 4 cameras, 6 persons. There are both cases with more people than cameras (see Fig. 14) and more cameras than people (see Figs. 6 and 10), which show that the performance of the proposed approach will not be influenced by relative numbers of cameras and persons. Both indoor and outdoor experiments are provided. The lengths of the video sequences vary from 450 frames to 700 frames. The frame rate for all indoor videos is 30 fps while that



3 cameras, 1 person | 3 cameras, 2 persons | 2 cameras, 3 persons | 4 cameras, 4 persons | 4 cameras, 6 persons

Fig. 3. Camera configuration and the persons' trajectories in the experimented cases.

for outdoor videos is 15 fps. The cameras used in our experiments are all Axis 215 PTZ cameras, which are placed arbitrarily. To fully test whether the proposed approach can help to select the "best" camera based on the user supplied criteria, some of the FOVs of these cameras are allowed to interact while some of them are nonoverlapping. The experiments are carried out in three different places with no camera calibration done before hand. The trajectories are randomly chosen by the persons for walking. We visualize the camera configuration and the persons' trajectories for the five cases in Fig. 3.

2) Parameters: In our experiments, we empirically give values to the parameters required by the criteria introduced in Section III-B-1. $\lambda = 1/15$, $\xi = 1/6$, $w_1 = 0.2$, $w_2 = 0.1$, $w_3 = 0.7$.

These parameters are held constant for all the experiments reported in this paper.

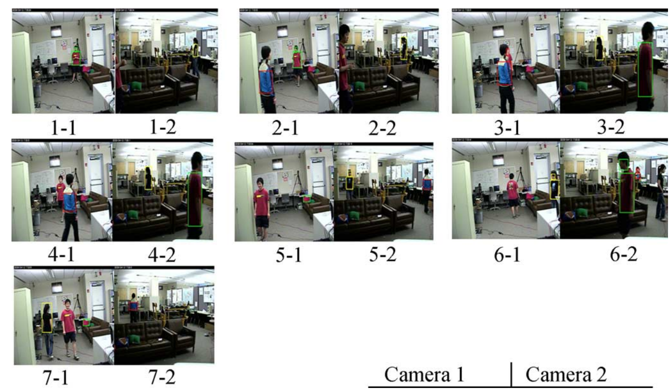


Fig. 4. An example for the failure of the Continuous Adaptive Meanshift (Camshift) tracker. We only draw ROIs in the cameras that are selected to track the persons. We generate bounding boxes only for those cameras that are selected to track a person.

B. Tracking and Face Detection

1) Tracking: All the experiments are conducted using the Continuous Adaptive Meanshift (Camshift) tracker [19] to evaluate the camera selection and handoff mechanisms. Theoretically, which tracker is used is not important as long as it can provide the tracking information that consists of size (size of the bounding box of a person) and location (position of the centroid of the bounding box) of a person. It is to be noticed that the same tracker is used for all the experiments and all the camera assignment approaches that are compared to filter out the influence of a tracker to the camera selection results.

The walking persons are initially selected by an observer manually when a person enters the FOV of a camera as detected by the background subtraction method. The persons who participated in the experiments wear clothes in distinct colors, so different persons can be identified by calculating the correlation of the hue histograms of the pixels inside their bounding boxes (ROIs) using the function CompareHist [20].

a) Errors caused by the tracker ($N_c = 2$, $N_p = 3$, indoor): There are some errors that are caused by the failure of the tracker. In Fig. 4, we show some error frames in a 2 cameras, 3 persons case, which are due to the failure of the Camshift tracker. The Camshift tracker is not robust when severe occlusion happens and it can be distracted by the object with similar colors as the target. However, the camera assignment results are correct if we ignore the errors that are caused by the tracker, i.e., if we assume that the tracker provides a correct ROI for the target, then the camera assignments, performed based on the user-supplied criteria, are correct. For instance, in Fig. 4 (4-1 and 4-2), the system should select camera 1 to track the person in red, where the person has a frontal view which is preferred according to the user-supplied criteria. However, the tracker for the person in red is distracted by the red pillow, which causes error for the camera assignment. But if we assume that the ROIs returned by the tracker are correct, then based on the size and position of the person in red (frontal face is not available because of the wrong tracking result), the system selects the correct camera.

2) Face Detection: Face detection is done in a particular region (top 1/3 height of the ROI), provided by the tracker. We use the cascade of haar feature based classifiers for face detection [20]. It can detect faces correctly in 90%+ cases when the tracker returns a correct ROI.

C. Performance Measures

In our experiments, the bottom line is to track walking persons seamlessly, i.e., the system will follow a person as long as the person appears in the FOV of at least one camera. In the case where more than one camera can "see" the persons, we assume that the camera that can "see" the person's face is preferable. This is because in surveillance systems, the frontal view of a person can provide us more interesting information than other views. So, based on this criterion, we define the camera assignment error in our experiments as: 1) failing to track a person, i.e., a person can be seen in some cameras in the system but there is no camera assigned to track the person or 2) failing to get the frontal-view of a person whenever it is available. We define these error terms in the following:

TABLE III
EXPERIMENT #1. OVERVIEW OF VIDEOS FOR EACH CAMERA AND THE NUMBER OF HANDOFFS THAT ARE TAKEN PLACE (NOF: NUMBER OF FRAMES)

	NOF(0 person)	**NOF**(1 person)	**NOF**(2 persons)	**NOF**(with occlusion)	No. of handoffs (Crt_{s4})
Cam1	56	22	12	0	2
Cam2	14	46	18	11	9
Cam3	44	23	17	6	6

N_{lost} the number of times that a target person is lost. It is determined if the bounding box returned by the tracker covers less than 30% of the person's actual size or is larger than 150% of the person's actual size during tracking. The term region-of-interest (ROI) and bounding box are used interchangeably in this paper.

N_{fvl} the number of times a frontal view is detected but not selected by any camera. Note that in our experiments, there is no case where a frontal view is detected but the person is lost during tracking. So, the intersection of the above two cases should be empty, i.e. $N_{\text{lost}} \cap N_{fvl} = \varnothing$.

$NP_i^{C_j}$ the number of persons appearing in Camera C_j in frame i.

N_f the total number of frames in an experimented video.

NPC_i the number of cameras with no persons in frame i.

N_C the total number of cameras in an experiment.

The total error of a video is defined as

$$Err = N_{\text{lost}} + N_{fvl}. \tag{14}$$

The error rate is defined as the error normalized by total the numbers of cameras and persons in all frames. Frames in which there are no persons are counted as correct frames, since there are no errors caused by losing a person or lose the frontal view of a person. Frames with more than one person in the FOV of a camera are multiply counted to normalize by the multiple persons. Error rate ER is defined as

$$ER = \frac{Err}{\sum_{i=1}^{N_f}\sum_{j=1}^{N_C} NP_i^{C_j} + \sum_{\substack{i=1 \\ NP_i^{C_j}=0}}^{N_f} NPC_i}. \tag{15}$$

D. Evaluation of Game Theoretic Framework

1) Experiment #1: Criterion Selection ($\mathrm{N_C} = 3$, $\mathrm{N_P} = 2$, *Indoor):* Since there are multiple criteria to be used in the experiments, we first test the performance for different criterion in a 3 cameras, 2 persons case. A general description of the videos is shown in Table III.

Different experiments are carried out using the single and the combined criterion described in Section III-B-1. Some typical results are shown in Fig. 5. To make it convenient for a comparison, we show the tracking results for other cameras as well, no matter whether they are selected for tracking or not. The cameras, for which the bounding boxes are drawn in blue, are selected for tracking, while the ones in red or green are not as good as the blue ones.

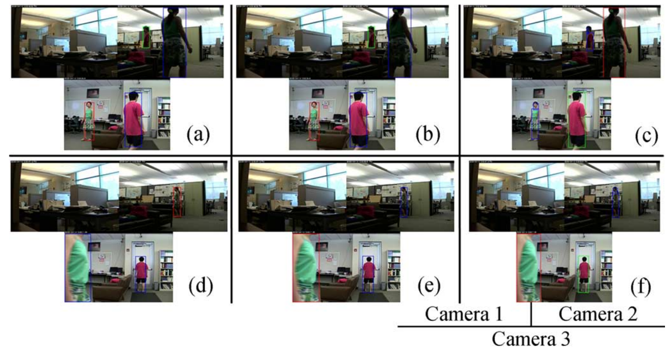


Fig. 5. Experiment #1. A comparison for using different criteria. The first row and the second row are for two time instants respectively. The first column through the third column are using criterion 1 to criterion 3, respectively.

Fig. 5(a)–(c) use criterion 1–3 at time instant 1, while Fig. 5(d)–(f) use criterion 1–3 at time instant 2. It can be observed from Fig. 5(d) that the problem for using criterion 1 only is that when the persons are getting close to the cameras, the size of the bounding box increases, and while the resolution is not very high, persons are not clear enough. Meanwhile, there are cases such that when a person is entering the FOV of a camera, the size of the person is not small but only part of the body is visible. This should not be preferred if other cameras can give a better view of the body. Thus, we introduced criterion 2, considering the relative position of persons in the FOVs of the cameras. The closer the centroid of a person is to the center of the FOV of a camera, the higher the camera utility is generated. We can observe that when applying criterion 2 in Fig. 5(e), the camera with the person near the center is chosen and we can obtain a higher resolution of the person compared to the results based on criterion 1 in Fig. 5(d). However, the problem for using criterion 1 or criterion 2 only is that we reject the camera(s) which can see a person's face, which is of general interest. This case is shown in Fig. 5(a), (b), and (d). To solve this problem, we developed criterion 3 (the view of the person). So, when applying criterion 3, we obtain a more desirable camera with a frontal view of the person in Fig. 5(c) and (f). Whereas criterion 3 can successfully select a camera with a frontal-view person, it may fail to track a person when no face can be detected. As shown in Fig. 5(f), although the person is in the FOV of some camera, the person is lost based on criterion 3.

So, finally, we come up with a weighted combination of these three criteria. As stated previously, we use 0.2, 0.1, and 0.7 as the weights for these three criteria respectively so that, in most cases, the system will choose the camera which can "see" a person's face. For those frames where there is person without the detected face, the combination criterion can also provide the "best" camera based on criteria 1 and 2 and, thus, realizing continuous tracking. All the camera handoffs, when applying the combined criterion, are shown in Fig. 6. The error rate in

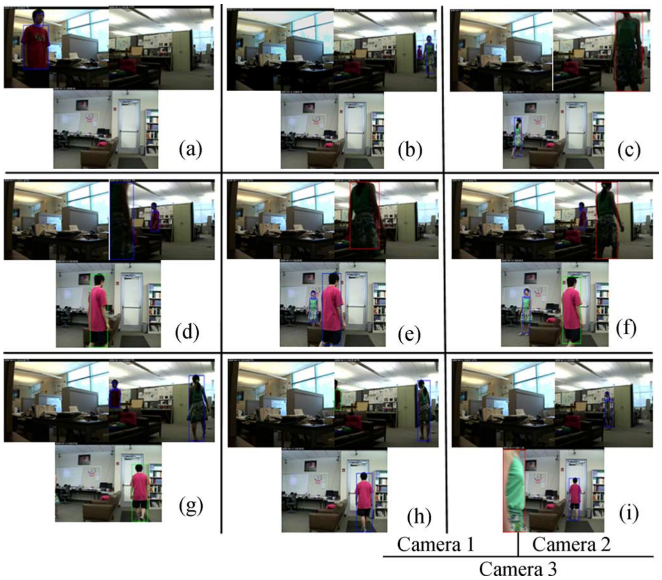


Fig. 6. Experiment #1. All camera handoffs when applying the combined criterion for 3 cameras, 2persons case. The cameras that are selected for tracking a person provides a blue bounding box for that person, otherwise it provides green bounding box for the person in red and red bounding box for the person in green.

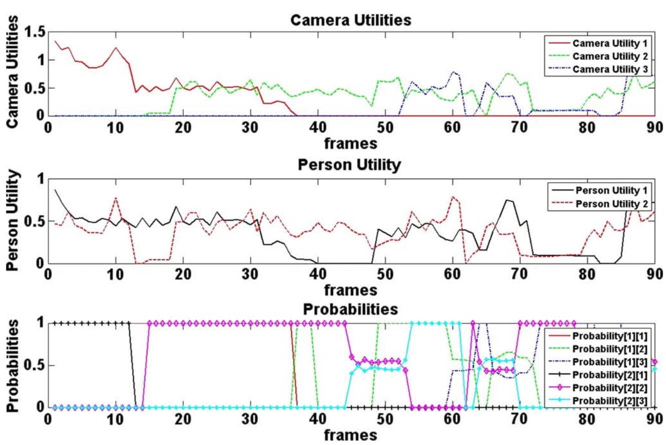


Fig. 7. Experiment #1. Utilities and assignment probabilities for each processed frame when using the combined criterion.

this case is 5.56%, while that for using criterion 1 to 3 only are 25.56%, 10.00%, and 30.00%, respectively.

The number of handoffs in this 3 cameras, 2 persons case is give in Table III. Camera utilities, person utilities and the corresponding assignment probabilities for the using the combined criterion is shown in Fig. 7, where Probability[i][j] stands for the probability that C_j is assigned to track P_i.

We use the combined criterion for all the other experiments in the rest of this paper.

2) Convergence of Results for Bargaining: For the above experiments, in most cases, the probabilities for making the assignment profile converges (with $\varepsilon < 0.05$, where ε is the difference between the two successive results) within five iterations. So, we use 5 as the number of iterations threshold when bargaining. Thus, for those cases that will not converge within five iterations, there may be an assignment error based on the

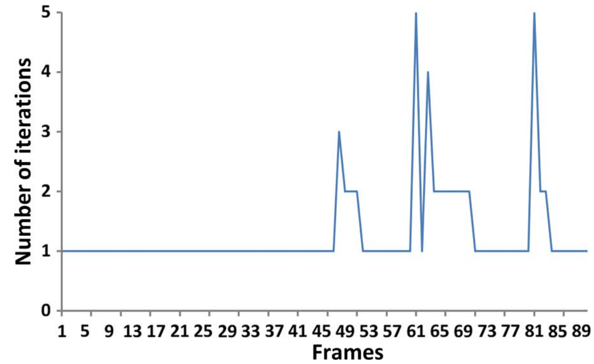


Fig. 8. Experiment #1. Number of iteration for the bargaining mechanism in each frame.

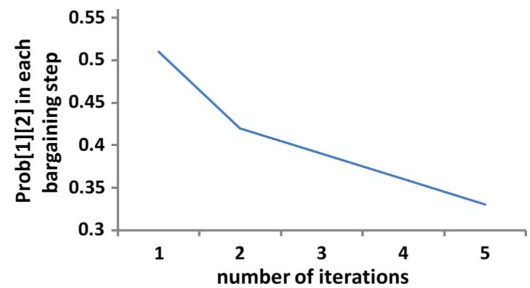


Fig. 9. Experiment #1. A typical convergence in the bargaining process (frame 56, camera 2, for the person in green).

unconverged probabilities. In Fig. 8, we plot the number of iteration with respect to every processed frame for Experiment #1. It turns out that the average number of iterations is 1.37. As the numbers of persons and cameras increase, this bargaining system will save a lot of computational cost to get the optimal camera assignments. A typical convergence for one of the assignment probabilities in a bargaining among cameras is given in Fig. 9. We also show an example of error caused by the failure of the bargaining mechanism in a more complicated (4 cameras, 6 persons) Experiment #6 discussed later in the comparison part.

E. Comparison of Game Theoretic Approach With Other Related Approaches

In this section, we will compare our approach with two other approaches: the first approach [3] performs camera handoff by calculating the co-occurrence to occurrence ratio (COR). We will call this the COR approach. The second approach performs the camera assignment problem by solving the Constraint Satisfaction Problem (CSP) [4]. We will call this approach the CSP approach in the following. As concluded in Section IV-D, we will use the combined criterion (9) for the following comparisons.

1) Comparison With the COR Approach: In [3], the mean probability that a moving object is detected at a location x in the FOV of a camera is called an occurrence at x. The mean probability that moving objects are simultaneously detected at x in the FOV of one camera and x' in the FOV of another camera is called a co-occurrence of x and x'. The COR approach decides whether two points are in correspondence with each other by calculating the co-occurrence to occurrence ratio. If the COR

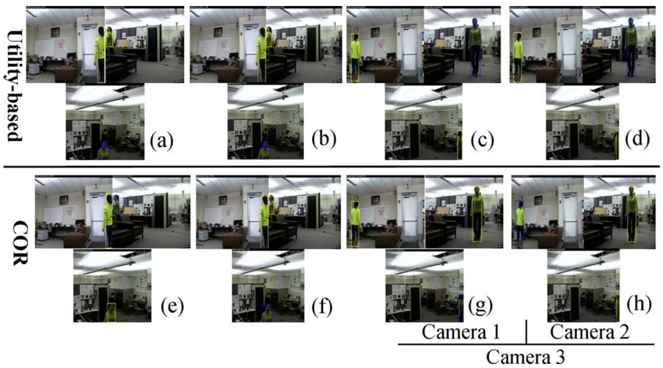


Fig. 10. Experiment #2. Two camera handoffs by using the co-occurrence to occurrence ratio (COR) approach and the comparison with our approach. The first row are the results by our approach and the second row are the results by the COR approach. The camera selected to track the person provides a blue bounding box, otherwise it provides a yellow bounding box.

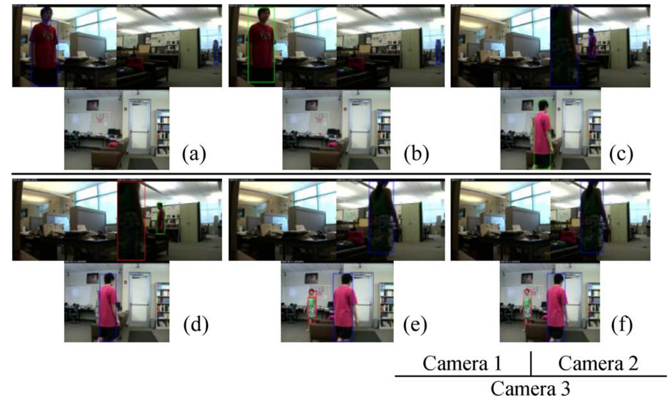


Fig. 11. Experiment #3. Some camera handoff errors by the co-occurrence to occurrence ratio (COR) approach in a 3cameras, 2 persons case. The cameras that are selected for tracking a person provides a blue bounding box for that person, otherwise it provides green bounding box for the person in red and red bounding box for the person in green.

is higher than some predefined threshold, then the two points are decided to be in correspondence with each other. When one point is getting close to the edge of the FOV of one camera, the system will handoff to another camera that has its corresponding point. However, the COR approach in [3] has been applied to two cameras only. We generalize this approach to the cases with more cameras by comparing the accumulated COR in the FOVs of multiple cameras. We randomly select 100 points on the detected person, train the system for ten frames to construct the correspondence for these 100 points, calculate the cumulative CORs in the FOVs of different cameras and select the one with the highest value for handoff.

Experiments have been done to compare the COR approach with our approach for the 3 cameras, 1 person case (Experiment #2) and the 3 cameras, 2 persons case (Experiment #3).

a) Experiment #2: Comparison with COR approach ($N_C = 3$, $N_P = 1$, indoor): The handoff process by using the COR approach and the corresponding frames by using our approach (may not be the handoff frames) are shown in Fig. 10. In Fig. 10(g) and (h), the COR approach switches to camera 1, while our proposed approach sticks to camera 2 [Fig. 10(c) and (d)] to get the frontal view of the person. The COR approach needs a time period to construct the correspondence between different views. We let this period to be ten frames. As a result, there is some time delay for the handoff. For instance, in Fig. 10(a) and (b), our approach has already selected camera 3 in (a), where a frontal view of the person is already available and the size of the person is acceptable, while the COR approach switched to camera 3 in (d) when the person is detected as leaving the FOV of camera 2 and entering the FOV of camera 3.

b) Experiment #3: Comparison with the COR approach ($N_C = 3$, $N_P = 2$, indoor): In Fig. 11, we show some error frames by using the COR approach. These results can be compared with Fig. 6 (Experiment #1) where we use the same video for the proposed approach. By the comparison, we can notice that the COR approach can only switch the camera to another one when the person is about to leave the FOV, but cannot select the "best" camera based on other criteria. So, the number of handoffs by our approach is larger than that of the COR approach (see Table IV). If we use the definition of error as stated in Section IV-C, the error rates for these two cases are compared in Table IV. Based on this error definition, the COR approach loses the frontal view of a person more easily. Examples are Fig. 11(b) (lose the person in red), (d) (lose the frontal view of the person in red), and (f) (lose the frontal view of the person in green).

TABLE IV
COMPARISON OF ERROR RATES FOR THE CO-OCCURRENCE TO OCCURRENCE RATIO (COR) APPROACH AND THE PROPOSED APPROACH

	Experiment #2		Experiment #3	
	# of handoffs	Error rate	# of handoffs	Error rate
COR	4	38.77%	5	45.62%
Proposed	6	2.87%	8	5.56%

2) Comparison With the CSP Approach: The approach in [4] solves the camera assignment problem by using the constraint satisfaction approach. According to the key assumptions made in Section III-B, we allow one camera to track multiple persons but one person can only be tracked by one camera. So, for each camera C_j, we let all those persons that can be seen by this camera form a group g_j. For instance, if, in our case, the camera C_j can see person P_1 and P_2, then the domain of g_j, noted as $\mathrm{Dom}[g_j]$, is $\{\{P_1\}, \{P_2\}, \{P_1, P_2\}\}$. The constraint is set to be $d_i \cap d_j = \{\varnothing\}$, for $i \neq j$, where $d_i \in b_i \cup \varnothing$ is the camera assigned to track person P_i, and b_i and b_j belong to $\mathrm{Dom}[g_j]$ and $i \neq j$. By doing so, we mean that the persons to be tracked are assigned to different cameras. We changed some of the notations from [4] so that the notations in this section are not in conflict with the notations used in the previous sections of this paper.

Experiments for 3 cameras, 2 persons (Experiment #4) and 4 cameras, 4 persons (Experiment #5) cases are carried out under the above constraint to maximize the criterion 4 [eq. (9)], using the *BestSlov* algorithm in [4].

a) Experiment #4: Comparison with the CSP approach ($N_C = 3$, $N_P = 2$, indoor): Since our approach requires five iterations for the 3 cameras, 2 persons case (Experiment #3) to

TABLE V
COMPARISON OF ERROR RATES FOR THE CONSTRAINT SATISFACTION PROBLEM (CSP) APPROACH AND THE PROPOSED APPROACH

	Experiment #4		Experiment #5	
	# of handoffs	Error rate	# of handoffs	Error rate
CSP	9	8.38%	17	10.66%
Proposed	8	5.56%	19	7.32%

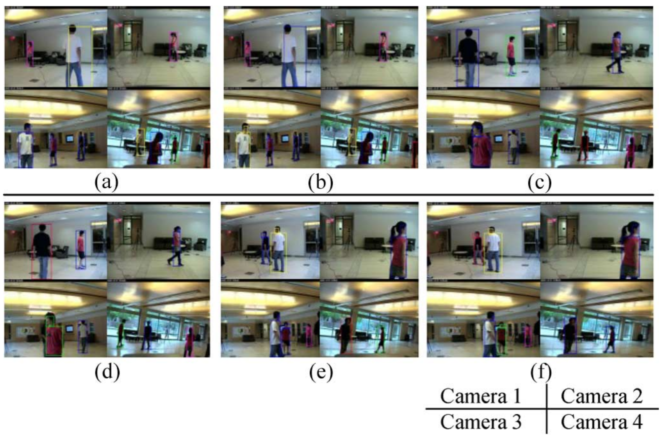


Fig. 12. Experiment #5. Some error frames by using the CSP approach for the 4 cameras, 4 persons case. The camera with blue bounding box for a person is selected to track the person. The cameras selected for tracking a person provides a blue bounding box for that person.

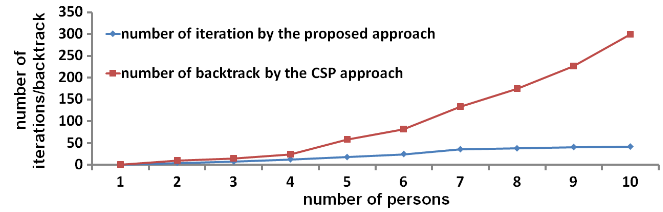


Fig. 13. Experiment #6. Comparison for the number of iteration or backtrack by the proposed utility-based approach and by the CSP approach.

get acceptable results, we also use five backtracking steps in the CSP approach.

Both the CSP approach and our proposed approach are able to accommodate different criteria. Most of the time, the CSP approach can select the "best" camera, based on our criterion and the error definition. So, we only compare the number of handoffs and the error rates for this case in Table V. The results show that the CSP approach has higher error rates than our approach.

b) Experiment #5: Comparison with the CSP approach ($N_C = 4$, $N_P = 4$, indoor): Since in this case, there are more persons and cameras involved, we increase the number of backtracking steps and the number of iterations to 10. Because the performance of the CSP approach heavily depends on the number of backtracks (the more backtracks it takes, the more accurate the results can be), as the number of cameras and persons goes up, the CSP approach will miss the "best" camera with a high probability. Some of the errors for this case are shown in Fig. 12. There are errors when a person's frontal view is available but it is not chosen such as in Fig. 12(b), (d), and (f), or when a person's frontal view is unavailable, the system chooses the camera with a smaller size person and farther from the center of the FOV, such as in Fig. 12(d) for the person in black. The high error rate for the CSP approach is due to its computational cost.

c) Experiment #6: Further comparison between the CSP and the proposed game theoretic approach—Number of iterations ($N_c = 3$, $N_P = 1$–10): Fig. 13 gives a comparison of number of iterations for our approach and the number of backtracks for the CSP approach for the case when the number of cameras is fixed to 3 and the number of persons goes up from 1 to 10. We can see that although the CSP approach can solve the camera assignment problem based on different user-defined tasks, it is computationally expensive as the complexity of the system increases.

3) Comparison Among Game Theoretic, COR and CSP Approaches:

a) Experiment #7: $N_C = 6$, $N_P = 4$, outdoor: In this section, we consider a more complicated case with 4 cameras and 6 persons. Because there are too many people in the system, it will be hard to observe if we mark the person in all the cameras that can see them. So, we only draw the bounding boxes for those cameras which are assigned to track the specific person. Different colors are used to distinguish different persons. We only display some typical results (Fig. 14) for each of the approaches that are compared. Because there are more cameras and persons involved in this experiment than the previous ones, we increase the number of iterations to 20 for all the CSP and the proposed approach.

For the proposed approach, we can notice that whenever there is a camera available to track a specific person, the camera assignment can be performed based on the predefined criteria. In Fig. 14 A6 (the utility-based approach group), we provide a case when the bargaining mechanism fails, i.e., the number of iterations is not large enough to converge to the optimal result. In this figure, the person in red bounding box should be tracked by Camera 1 based on an exhaustive calculation which can be regarded as the ground-truth.

The COR approach cannot decide which camera to select based on the user supplied criteria. So most of the handoffs take place when a person is leaving the FOV of one camera and entering the FOV of another camera. The CSP approach can deal with the supplied criteria to some extent, but since 20 backtracks are too few to reach the optimal answer, the CSP loses the "best" camera easily.

The overall performance of these approaches is presented in Table VI.

V. CONCLUSION

In this paper, we proposed a new principled approach based on game theory for the camera assignment and handoff problem. We developed a set of intuitive criteria in this paper and compared them with each other as well as the combination of them. Our experiments showed that the combined criterion is the best based on the error definition provided in Section IV. Since the utilities, input of the bargaining process, largely depend on the user-supplied criteria, our proposed approach can be task-oriented. Unlike the conventional approaches which perform camera handoffs only when an object is leaving or

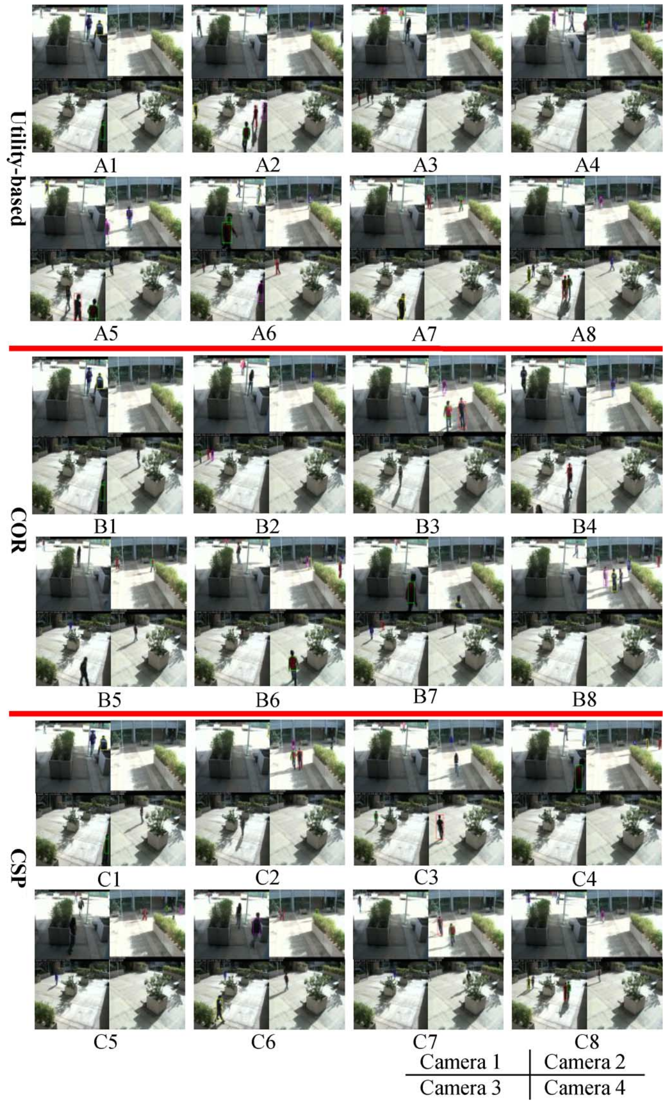


Fig. 14. Experiment #7. A comparison for the proposed utility-based game theoretic approach, the COR approach and the CSP approach. Only those cameras selected to track a person provide a bounding box for that person.

TABLE VI
COMPARISON OF ERROR RATES FOR THE COR, CSP, AND THE PROPOSED APPROACH

Experiment #6	COR	CSP	Proposed
Error rates	45.67%	12.96%	7.89%

entering the FOV, we can select the "best" camera based on the predefined criteria.

The key merit of the proposed approach is that we use a theoretically sound game theory framework with bargaining mechanism for camera assignment in a video network so that we can obtain a stable solution with a reasonably small number of iterations. The approach is independent of: a) the spatial and geometrical relationships among the cameras and b) the trajectories of the objects in the system. It is robust with respect to multiple user-supplied criteria. The approach is flexible since there is no requirement for a specific criterion that a user is obligated to use. A wide variety of experiments show that our approach is computationally more efficient and robust with respect to other existing approaches [3], [4].

We analyzed the influence of a tracker on the proposed approach in Section IV-B and compared our work with two other recent approaches both qualitatively and quantitatively. All the experiments used a physical camera network with real data in real time. This included both indoor and outdoor environments with different numbers of cameras and persons. As compared to the other approaches, it is shown that the proposed approach has smaller error rates in all the experiments. The computational efficiency of the proposed approach is also verified quantitatively. This comparison shows that: a) COR approach cannot do any criterion-dependent camera assignment. b) As the number of cameras and persons in the system increases, the assignment ambiguity and failure also increase in the COR approach. c) The CSP approach is task-dependent and can select the "best" camera based on whatever criterion is provided by the user. d) The CSP approach is computationally much more expensive than our approach.

Our future work will allow communication among cameras, which will make the computational framework and computational resources decentralized and distributed.

REFERENCES

[1] O. Javed, S. Khan, Z. Rasheed, and M. Shah, "Camera hand-off: Tracking in multiple uncalibrated stationary cameras," in *Proc. IEEE Workshop on Human Motion*, 2000, pp. 113–118.

[2] J. Park, P. C. Bhat, and A. C. Kak, "A look-up table based approach for solving the camera selection problem in large camera networks," in *Proc. Int. Workshop on Distrib. Smart Cameras*, 2006.

[3] Y. Jo and J. Han, "A new approach to camera hand-off without camera calibration for the general scene with non-planar ground," in *Proc. ACM Int. Workshop on Video Surveillance and Sensor Networks*, 2006, pp. 195–202.

[4] F. Z. Qureshi and D. Terzopoulos, "Multicamera control through constraint satisfaction for pedestrian surveillance," in *Proc. AVSS*, 2008, pp. 211–218.

[5] X. Zou and B. Bhanu, "Anomalous activity detection and classification in a distributed camera network," in *Proc. ICIP*, 2008, pp. 781–784.

[6] V. Kettnaker and R. Zabih, "Bayesian multi-camera surveillance," in *Proc. CVPR*, 1999, pp. 253–259.

[7] T. Chang and S. Gong, "Bayesian modality fusion for tracking multiple people with a multi-camera system," in *European Workshop on Advanced Video-Based Surveillance Systems*, 2001.

[8] J. Kim and D. Kim, "Probabilistic camera hand-off for visual surveillance," in *Proc. ICDSC*, 2008.

[9] J. Kang, I. Cohen, and G. Medioni, "Continuous tracking within and across camera streams," in *Proc. CVPR*, 2003, pp. 267–272.

[10] O. Javed, K. Shafique, and M. Shah, "Appearance modeling for tracking in multiple non-overlapping cameras," in *Proc. CVPR*, 2005, pp. 26–33.

[11] B. Son and A. Roy-Chowdhury, "Stochastic adaptive tracking in a camera network," in *Proc. ICCV*, 2007.

[12] Q. Cai and J. K. Aggarwal, "Tracking human motion in structured environments using a distributed camera system," *Proc. PAMI*, pp. 1241–1247, 1999.

[13] G. Arslan, J. R. Marden, and J. S. Shamma, "Autonomous vehicle-target assignment: A game-theoretical formulation," *ASME Trans. Dynamic Syst., Measure., Control, Special Issue on Analysis and Control of Multi-Agent Dynamic Systems*, pp. 584–596, 2007.

[14] L. Tessens *et al.*, "Principal view determination for camera selection in distributed smart camera networks," in *Proc. ICDSC*, 2008.

[15] R. B. Myerson, *Game Theory-Analysis of Conflict*. Cambridge, MA: Harvard Univ. Press, 1991.

[16] M. J. Osborne, *An Introduction to Game Theory*. Oxford, U.K.: Oxford Univ. Press, 2003.

[17] [Online]. Available: http://plato.stanford.edu/entries/game-theory/#Uti

[18] Y. Li and B. Bhanu, "Utility-based dynamic camera assignment and hand-off in a video network," in *Proc. ICDSC*, 2008.
[19] G. R. Bradski, "Computer vision face tracking for use in a perceptual user interface," *Intel Technology J., Q2*, 1998.
[20] [Online]. Available: http://opencv.willowgarage.com/wiki/CvReference
[21] M. J. Osborne and A. Rubinstein, *A Course in Game Theory*. Cambridge, MA: The MIT Press, 1994.

Yiming Li (S'08) received the B.S. degree in electrical engineering from the South China University of Technology, Guangzhou, China, in 2006 and the M.S. degree in electrical engineering from the University of Florida, Gainesville, in 2007. She is currently working towards the Ph.D. degree at the Center for Research in Intelligent Systems, University of California, Riverside.

Her research interests are in computer vision, pattern recognition and image processing. Her recent research has been concerned with camera networks.

Bir Bhanu (S'72–M'82–SM'87–F'96) received the S.M. and E.E. degrees in electrical engineering and computer science from the Massachusetts Institute of Technology, Cambridge, the Ph.D. degree in electrical engineering from the Image Processing Institute, University of Southern California, Los Angeles, and the MBA degree from the University of California, Irvine.

He is the Distinguished Professor of Electrical Engineering and serves as the Founding Director of the Interdisciplinary Center for Research in Intelligent Systems, University of California, Riverside (UCR). He was the founding Professor of Electrical Engineering at UCR and served as its first chair (1991–1994). He has been the Cooperative Professor of Computer Science and Engineering (since 1991), Bioengineering (since 2006), Mechanical Engineering (since 2008), and the Director of Visualization and Intelligent Systems Laboratory (since 1991). Previously, he was a Senior Honeywell Fellow with Honeywell Inc., Minneapolis, MN. He has been with the faculty of the Department of Computer Science, University of Utah, Salt Lake City, and with Ford Aerospace and Communications Corporation, Newport Beach, CA; INRIA-France; and IBM San Jose Research Laboratory, San Jose, CA. He has been the principal investigator of various programs for the National Science Foundation, the Defense Advanced Research Projects Agency (DARPA), the National Aeronautics and Space Administration, the Air Force Office of Scientific Research, the Office of Naval Research, the Army Research Office, and other agencies and industries in the areas of video networks, video understanding, video bioinformatics, learning and vision, image understanding, pattern recognition, target recognition, biometrics, autonomous navigation, image databases, and machine-vision applications. He is coauthor of the books *Computational Learning for Adaptive Computer Vision* (Forthcoming), *Human Recognition at a Distance in Video* (Springer-Verlag, 2010), *Human Ear Recognition by Computer* (Springer-Verlag, 2008), *Evolutionary Synthesis of Pattern Recognition Systems* (Springer-Verlag, 2005), *Computational Algorithms for Fingerprint Recognition (Kluwer, 2004), Genetic Learning for Adaptive Image Segmentation* (Kluwer, 1994), and *Qualitative Motion Understanding* (Kluwer, 1992), and coeditor of the books on *Computer Vision Beyond the Visible Spectrum* (Springer-Verlag, 2004), *Distributed Video Sensor Networks* (Springer-Verlag, 2010), and *Multibiometrics for Human Identification* (Cambridge University Press, 2010). He is the holder of 11 U.S. and international patents. He has more than 350 reviewed technical publications.

Dr. Bhanu is a Fellow of American Association for the Advancement of Science, the International Association of Pattern Recognition, and the International Society for Optical Engineering. He has been on the editorial board of various journals and has edited special issues of several IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE; IMAGE PROCESSING; SYSTEMS, MAN AND CYBERNETICS—PART B; ROBOTICS AND AUTOMATION; INFORMATION FORENSICS AND SECURITY), and other journals. He was the General Chair for the IEEE Conference on Computer Vision and Pattern Recognition, the IEEE Conference on Advanced Video and Signal-Based Surveillance, the IEEE Workshops on Applications of Computer Vision, the IEEE Workshops on Learning in Computer Vision and Pattern Recognition; the Chair for the DARPA Image Understanding Workshop, the IEEE Workshops on Computer Vision Beyond the Visible Spectrum and the IEEE Workshops on Multi-Modal Biometrics. He was the recipient of best conference papers and outstanding journal paper awards and the industrial and university awards for research excellence, outstanding contributions, and team efforts.