Continuously Tracking and See-through Occlusion Based on A New Hybrid Synthetic Aperture Imaging Model

Tao Yang Yanning Zhang Xiaomin Tong Xiaoqiang Zhang Rui Yu
Shaanxi Key Laboratory of Speech & Image Information Processing
School of Computer Science, Northwestern Polytechnical University, Xi'an, 710129, China
yangtaonwpu@163.com ynzhang@nwpu.edu.cn xmtongnwpu@gmail.com

Abstract

Robust detection and tracking of multiple people in cluttered and crowded scenes with severe occlusion is a significant challenging task for many computer vision applications. In this paper, we present a novel hybrid synthetic aperture imaging model to solve this problem. The main characteristics of this approach include: (1) To the best of our knowledge, this algorithm is the first time to solve the occluded people imaging and tracking problem in a joint multiple camera synthetic aperture imaging domain. (2) A multiple model framework is designed to achieve seamless interaction among the detection, imaging and tracking modules. (3)In the object detection module, a multiple constraints based approach is presented for people localizing and ghost objects removal in a 3D foreground silhouette synthetic aperture imaging volume. (4) In the synthetic imaging module, a novel occluder removal based synthetic imaging approach is proposed to continuously obtain object clear image even under severe occlusion. (5) In the object tracking module, a camera array is used for robust people tracking in color synthetic aperture images. A network camera based hybrid synthetic aperture imaging system has been set up, and experimental results with qualitative and quantitative analysis demonstrate that the method can reliably locate and see people in challenge scene.

1. Introduction

The increasing use of multiple video sensors in surveillance applications has greatly increased researchers' attention on extracting and processing information from multiview video streams. Since the high level understanding of a surveillance scene relies on accurate low level detection and tracking of moving objects, developing a fully automatic and robust people tracking system has become a subject of great scientific and commercial interest [1-18].

Although many tracking algorithms and systems have been developed in the literature, in the case of cluttered and crowded scenes, tracking multiple people accurately is still a challenging task. The main problem is the frequent occlusion and complex intersection among objects, as a result people may be partially or even completely occluded in most of the camera views.

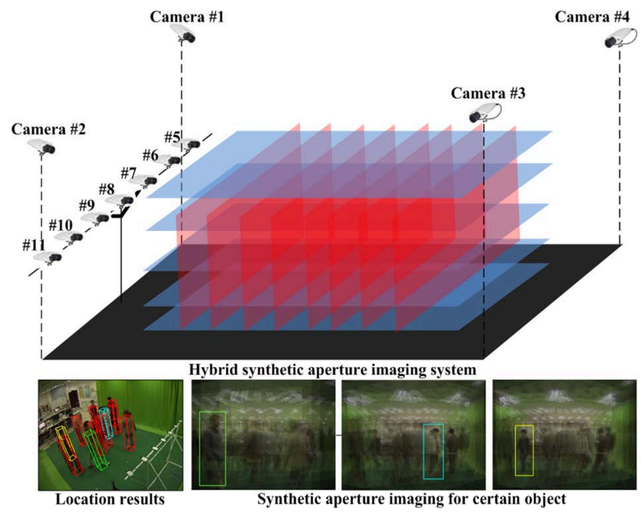


Figure 1. Four top down view AXIS network cameras are set to locate objects in the scene by foreground silhouette synthetic aperture images (shown as the blue horizontal planes). A linear camera array composed of seven AXIS network cameras is installed for object imaging and tracking in color synthetic aperture imaging space (shown as the red vertical planes). Results of object location and imaging through occlusion are shown at the bottom row.

1.1. Our Approach

To address this problem, in this paper we present a novel hybrid synthetic aperture imaging model.Our first contribution is that instead of tracking in the primitive color images, we solve this problem in a novel joint multiple camera synthetic aperture imaging domain (as shown in Fig.1), which contains complete multi-view color and foreground information of moving objects. Our second contribution is to design a multiple model framework and three interactive modules, which can continuously detect and track multiple objects. Our third contribution is a novel occluder removal based imaging approach which can significantly improve the imaging quality of objects even under severe occlusion.

1.2. Related Work

Broadly speaking, the previous works on tracking multiple occluding objects can be roughly classified into two categories: single-view approach and multi-view approach.

Single-view approach often adopts appearance features such as color, texture, shape [5,6] and dynamic models [7,8] to establish correspondence.Some approaches use incremental learning in subspace [10,11] to handle changes of object appearance or lighting condition. Single-view approach works well with isolated object, however in the case of occlusion the tracking process is severely hampered and often fails. Although the tracker may be recovered from transient occlusions through predicting the object's position, it is of little help when the object has non-linear motion and changes moving direction during occlusion.

Multi-view approach has the advantage over single-view methods because of its ability to decrease the hidden regions and recover 3D space information from the scene. Recently, the use of plane homography for people detection and tracking has gained more attention [12-18]. Fleuret et al. [15] adopt an occupancy map to depict the probability of target existing in the scene. Eshel et al. [17] present a multi-view detection method with intensity correlation. Khan et al. [12] propose a homography constraint and fuse information from multiple views into a synergy map, and localize people on this image. The homography based approach provides us a way for occlusion handling. However, it also brings about artifacts or so called ghost objects, and the performance of this approach declines with increasing number of objects in the scene.

The foundation of our approach is based on the synthetic aperture imaging [19-21]. In this technique, a dense camera array is usually set up to observe a scene, and all camera images are aligned to a plane and then averaged together to approximate a camera with a very large aperture. These synthetically constructed images have very shallow depth of field. More importantly, objects off the plane of focus "disappear" due to significant blur. We believe this unique characteristic makes camera array synthetic aperture imaging to be a powerful way for object detection and tracking especially in complex surveillance scene.Joshi et al. [22] take the lead for applying synthetic aperture theory to visual tracking, and present a robust dense camera array based single object tracking approach.

There are two major differences between our work and previous research. First, we seek to continuously seeing and tracking multiple objects simultaneously even under severe occlusion, this is a more general and harder problem than those solved by standard tracking algorithm. To achieve this, we rely on the unique imaging ability of the hybrid synthetic aperture imaging model. Second, this is the first time to combine linear camera array and multiple surrounded cameras to monitor a given scene together, and create color and foreground synthetic aperture images in each of them respectively.

The rest of the paper is structured as follows: Section 2 introduces the framework of our hybrid synthetic aperture imaging model. Section 3 and Section 4 details the multiple object detection, tracking and imaging algorithm in synthetic aperture imaging domain. Section 5 shows the experimental results, and provides insight into the utility and robustness of our approach. We conclude this paper in Section 6.

2. Algorithm Overview

In this section, we give an overview of the hybrid synthetic aperture tracking approach by describing the flow of information through the system components shown in Fig.2. Our method mainly includes three parts: (1) foreground synthetic aperture (FSA) detection, (2) synthetic aperture imaging, and (3) synthetic aperture tracking.

A sensing cycle begins when multiple camera capture multi-view video streams. The *Foreground Synthetic Aperture Detection* module takes the multiple top down view images as input, through foreground segmentation and synthetic aperture imaging, this module firstly generate a new 3D foreground synthetic aperture image volume, and then detect multiple moving objects with occlusion in this domain via multiple constraints. The output of this module is the position of detected objects in the world coordinate. For each object, the *Synthetic Aperture Imaging* module will continuously estimate object's position and generate a clear object image in the color synthetic aperture image domain. The output of this module is the optimal depth and clear image of each individual. The depth estimation results are then fed as an input to the *Synthetic Aperture Tracking* module. Then, for each new object, this module initializes a new tracker based on incremental learning based tracking method. Besides that, this module also generates an assignment of the object observations to tracks for tracker initialization, updating or terminating. Finally, the processing results are stored in a database to support possible higher-level applications and tasks.

3. Foreground Synthetic Aperture Detection

The input for our foreground synthetic aperture detection module are synchronized multiple top down view video streams, and the goal is to estimate the probability map that

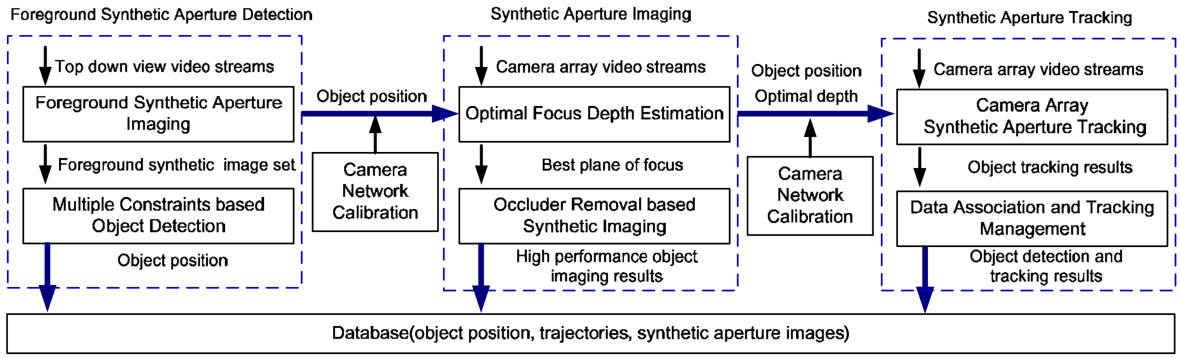


Figure 2. Framework of the hybrid synthetic aperture imaging and tracking algorithm.

the moving object appears at a certain position. The problem of object localization is formulated as the Bayesian estimation. Let $P(obj)$ as the probability map that objects appear at a certain location, in the condition that lack of any priori information, the likelihood of object position $P(obj)$ can be modeled as evenly distributed in the input image. In this section we refine the probability $P(obj)$ through successively add spatial and temporal constraints include foreground, shadow, occupancy, geometry and continuity in FSA imaging space. Fig.3 visualizes the process of the proposed multiple constraints.

•Foreground Constraint

The first constraint is the moving foreground extracted by long-term background modeling and subtraction in the monitoring scene. Here, we use an extension of the Gaussian mixture model (GMM) for pixel level background modeling [23]. The original GMM background model can deal with illumination changes, periodic motions from a cluttered background and long term scene changes. However, its processing speed is influenced by the number of Gaussian components. To handle this problem, a modified GMM model [23] with additional selection of the number of the Gaussian components for each pixel is implemented in our system, which achieves very fast foreground segmentation with slightly worse performance. According to the GMM background model, the posteriori probability of the object belonging to the foreground is given by

$$P(obj|foreground) \propto L(f) \tag{1}$$

$$L(f) = 1 - \sum_{i=1}^{B} \omega_{i,t} \cdot \eta(X_t, \mu_{i,t}, \textstyle\sum_{i,t}) \tag{2}$$

where X_t denotes the pixel intensity value, $\eta(X_t, \mu_{i,t}, \sum_{i,t})$ represents the gaussian probability density function with mathematical expectation $\mu_{i,t}$ and covariance matrices $\sum_{i,t} = \sigma_{i,t}^2 \mathbf{I}$ of the ith Gaussian components at time t. For computational reasons the covariance matrices are kept isotropic. The first B distributions are estimated by:

$$B = \arg\min_b \left(\sum_{i=1}^{b} \omega_{i,t} > T\right) \tag{3}$$

where T is the minimum portion of the background model.

•Shadow Constraint

Since shadow points are often misclassified as object in the background subtraction step, and will cause server errors in foreground segmentation, thus it's critical to detect and remove shadow for accurate object detection in video streams. The output of existing shadow removal method is a binary decision mask, once we select a certain approach, the Equation(1) can be modified as (4) with shadow constraint:

$$\begin{aligned} &P(obj|foreground, shadow) \\ &\propto L(S) = \begin{cases} L(f) & foreground \\ 0 & shadows \end{cases} \end{aligned} \tag{4}$$

Without the loss of generality, any shadow removal method can be used here according to diffident application fields, and a comprehensive survey of moving shadow detection approaches can be found in [24]. Because our system is installed in an indoor laboratory environment with relative simple background color and lack of texture information, we select a color based shadow removal approach [25] for its efficiency and low false alarm rate.

•Occupancy Constraint in FSA Image

The above two constraints are based on information extracted from single camera. In order to take advantage of multi-view information, we propose two novel occupancy and geometry constraints in FSA image domain.

In this paper, we extend the traditional camera array into multiple stationary cameras circling around the observing scene, and build a novel foreground silhouette synthetic aperture imaging space. Suppose a cylindrical object standing on the ground plane of a scene. To focus on a given plane Ω_j , we can project the foreground probability $L_i(S)$ of camera $i, i = 1, ..., n$ to this plane using a homography and then multiply as:

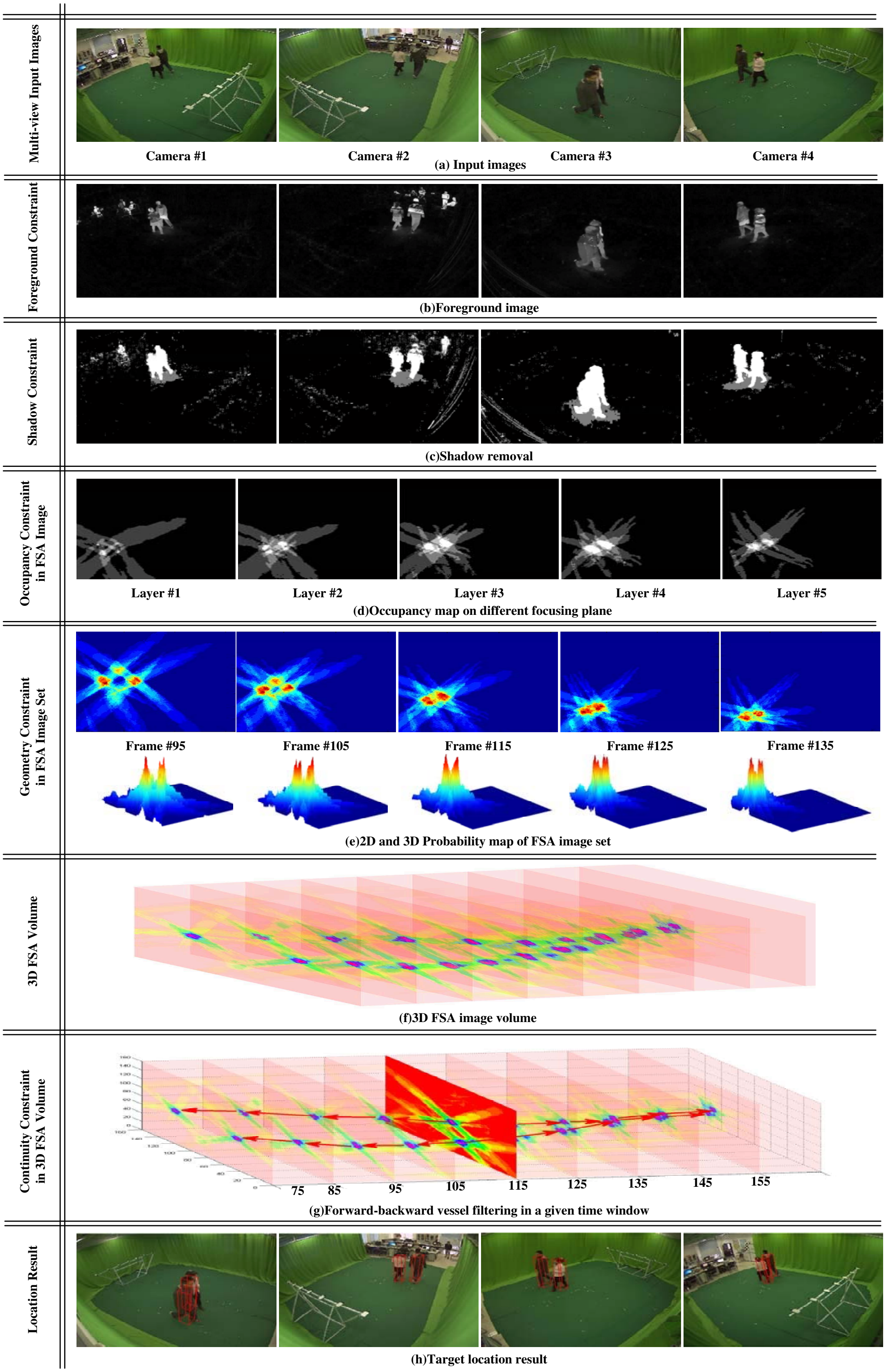


Figure 3. Multiple constraints based object detection in FSA image space.

$$L(\Omega_j) = \prod_{i=1}^{n} f(H_i^j, L_i(S)) \tag{5}$$

where $f(H_i^j, L_i(S))$ is a function to warp $L_i(S)$ to plane Ω_j according to matrix H_i^j.According to the synthetic aperture imaging theory, only the part of the cylindrical object on the plane Ω_j becomes sharp, whereas body parts away from the plane appear significantly blur due to parallax inside the plane. Thus we can state the following occupancy constraint in FSA image as:

$$P(obj|foreground, shadow, occupancy) \propto L(\Omega) \tag{6}$$

•Geometry Constraint in FSA Image Set

Using the occupancy constraint, we can achieve good detection result in simple scene. However, ghost objects would appear frequently when the scene is crowded or occlusion occurs. The ghost objects are caused by the intersections of different moving objects on the focus plane. To handle this problem, we adopt the parallax among multiple focus planes to enhance the object occupancy constraint.

The reason of this design is that when we focus on different parallel planes in different height(see Fig.3d), the appearances of object projections change along with different planes of focus, thus the positions of ghost objects change simultaneously. In contrast, the projection images of real object in each layer will be warped together, which denotes the true position of the object in the scene. Based on the difference of the ghost object's dynamic characteristic and the real object's position invariance on different planes of focus, we can eliminate the ghost objects through changing the height of focus plane among different views.

Theoretically we can get $(C_m^1)^n$ synthetic probabilistic images through various configurations with different camera views and planes of focus, where m and n denotes the total number of plane and camera respectively. However, it is difficult to decide which configuration should be used for object detection due to lack of priori information. Here, using Bayesian theory, we have the following derivation.

Given a set of focus planes Ω , the observation probability of the object is computed as:

$$L(G) = P(obj|\Omega_1, \Omega_2, ..., \Omega_M), M = (C_m^1)^n \tag{7}$$

By conditional independence, we can write this term as

$$L(G) = P(obj|\Omega_1){\cdot}P(obj|\Omega_1)...P(obj|\Omega_M) = \prod_{j=1}^{M} L(\Omega_j) \tag{8}$$

Plugging Equation(5) to Equation(8):

$$L(G) \propto \prod_{j=1}^{M}\prod_{i=1}^{n} f(H_i^j, L_i(S)) \tag{9}$$

Merging the same projection recurring in different layers:

$$L(G) \propto \prod_{j=1}^{M}\prod_{i=1}^{n} f(H_i^j, L_i(S)) = \prod_{j=1}^{m}\prod_{i=1}^{n} (f(H_i^j, L_i(S)))^{C/m} \tag{10}$$

Applying log likelihood, we get

$$L(G) \propto \sum_{j=1}^{m}\sum_{i=1}^{n} f(H_i^j, L_i(S)) \tag{11}$$

Thus we have

$$P(obj| \begin{matrix} foreground & shadow \\ occupancy & geometry \end{matrix}) \propto L(G) \tag{12}$$

Equation (12) provides the geometry constraint in foreground synthetic aperture image set, which has the unique ability for ghost objects removal.Fig.3e displays the probability map with geometry constraint.

•Continuity Constraint in 3D FSA Volume

Given the evidence of Equation (12) at a certain frame, although we can carry out hierarchical clustering in the 2D probability map to detect moving objects, this detection result is only optimal in frame level.

Considering the fact that object trajectories should be consistent over time, we propose to analyze the continuity in the 3D foreground synthetic aperture volume spanned by the temporal aggregation of the probability map. Fig.3f displays the 3D FSA volume, as it can be seen that the linear structures shows the trajectories. Thus the object detection problem has been transformed into a linear structure extraction problem. Linear structures such as vessels in medical images have been well studied, and inspired by those medical image processing techniques, we propose a forward-backward vessel filter to locate object in the defined 3D FSA volume.

The proposed method is based on iterative searching. At a given frame, firstly candidate objects are detected using hierarchical clustering in the likelihood map of Equation (12) with a given window width, then for each candidate, the probability for an object at its neighboring frames and locations will be checked, and the location of the search point will be updated by the new discovered candidates. This process is iterated until a given time window is researched. Fig.3g visualizes the above searching processing. The continuity score is the sum of confidence along the searching tracks in the 3D FSA volume.

$$L(\Theta) = \sum_{t} P_t(obj| \begin{matrix} foreground & shadow \\ occupancy & geometry \end{matrix}) \tag{13}$$

This function provides the output of the observation probability:

$$P(obj| \begin{matrix} foreground & shadow & occupancy \\ geometry & continuity \end{matrix}) \propto L(\Theta) \tag{14}$$

4. Synthetic Aperture Imaging and Tracking

Once we obtain the object location, we can use this information for further object imaging and tracking.

•Optimal focusing depth estimation

To create a clear synthetic image, we need to estimate the object's depth to the camera array precisely. In our system, we set up a linear array to simulate a large virtual lens of three meters. Since the camera array has a very limited depth of field, thus slight changes of focus depths will result severe blur in synthetic aperture image, we take advantage of this sensibility to estimate object's position precisely.

Fig.4 shows the process of the focus depth estimation . The first row shows five example input images of the camera array. Based on the Plane + Parallax calibration [19], we can compute synthetic aperture image on arbitrary focus planes that parallel to the camera array plane. Similar as the Section 3, we also perform foreground segmentation in each camera. Through varying the depth of focus plane, a set of FSA images are created. The second and third row of Fig.4 shows the FSA images on the depth from 2.0m to 3.4m. With the camera calibration information,we project the positions of detected objects to the FSA image of each depth (see blue and red box of Fig.4), and then estimate the optimal focus depth for each object through analyzing the blur properties of the sub-image inside the box.

An intuitive way is to calculate the average intensity variance inside the box. Considering the computational cost, we use a simple property to estimate the best depth. Since significant blur appears when the focus plane is away from the real depth of the object, the blur can be measured by the average gray level inside the box. The last row of Fig.4 gives the average gray value of each object in the scene, with depth ranging from 2.0m to 4.0m. Note that the maximum gray value corresponds the best focus depth.

•Occluder removal based hidden object imaging

As we have mentioned earlier, one unique characteristic of our approach is to acquire a high quality image of each individual people even under occlusion. Although the standard synthetic aperture imaging [21] has the ability to reconstruct an image of a partially occluded object surface, an obvious limitation is that the shadow cast by an occluder on the synthetic aperture image plane will cause artifacts on the final imaging result. To solve this problem, instead of accumulating all camera views indiscriminately, we analyze the visibility of each object respectively through the following occluder segmentation and removal method.

Fig.5 shows the process of our imaging method. In the first row of Fig.5 three people stand in a line, and the people in the last row (as shown in red box, Fig.5) is occluded by the other two people (as shown in green and yellow box, Fig.5). Firstly, after optimal focusing depth estimation, we get the complete object positions relative to the camera array, and the optimal focus FSA image of the front two occluders (as shown in Fig.5a). Secondly, through thresholding and morphological filtering, the binary mask of each occluder under a certain depth is acquired (as shown in Fig.5b). These binary images are then projected to each camera view through off-line calibrated homography to create multiple masks. Finally, during the synthetic aperture imaging step, the intensity value of each individual pixel is computed as the fraction of cameras which avoid the occluders. Fig.5d shows our imaging result. Please note that after occlude removal, our approach result in much better contrast and clarity compared against the standard synthetic aperture imaging (as shown in Fig.5c).

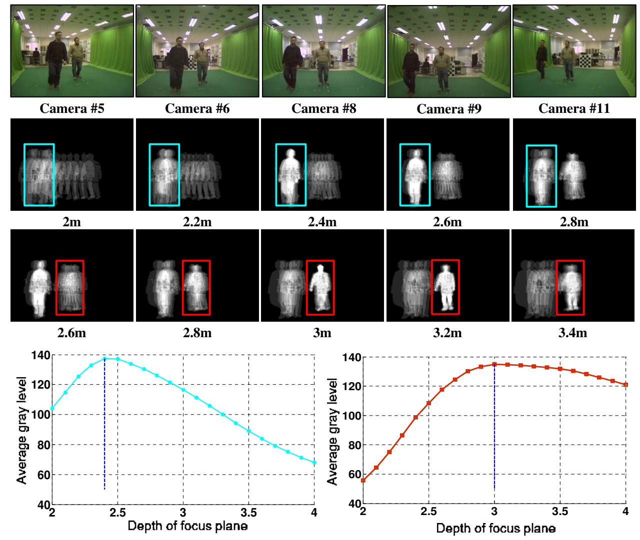


Figure 4. Optimal focusing depth estimation based on the foreground silhouette synthetic aperture imaging.

•Synthetic Aperture Tracking

In the synthetic aperture tracking module, we track objects through occlusion in the color synthetic aperture imaging space using the incremental learning based tracker [10] due to its robustness under challenging conditions like appearance or illumination changes.

The main difference of our work lies in the data association and tracking management. Through apply Hungarian association algorithm between the tracking and the detecting results, we classify the objects into three categories: new appearing objects, existing objects and disappearing objects. For each new appearing object, we estimate the optimal focusing depth and obtain the initial appearance model on the optimal focusing synthetic aperture imaging. With the initial appearance model and the optimal imaging depth, a new tracker is initialized. For each existing object, the corresponding tracker based on the incremental learning approach [10] is updated using the current color synthetic imaging. For each disappearing object, we will output the tracking results and record the lost times of the tracker. A tracker will be terminated when it doesn't get matches for several consecutive frames.

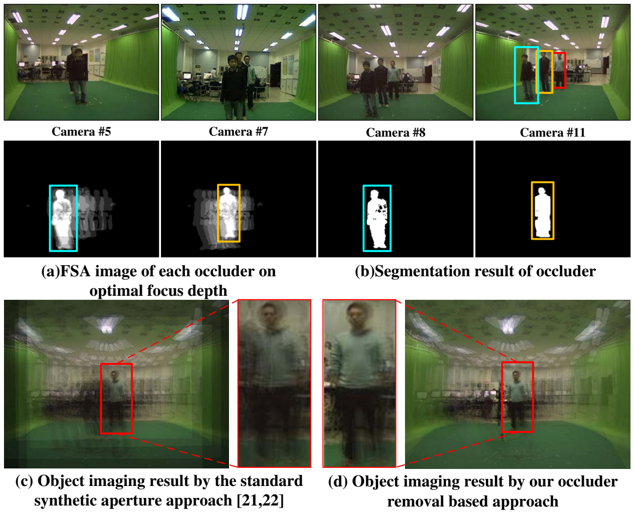


Figure 5. Occluder removal based hidden object imaging.

5. Experimental Results

Since the existing public multi-view databases do not satisfy the hybrid synthetic aperture imaging requirements of our approach, to evaluate the performance of the approach we have set up a camera network in our laboratory with eleven AXIS 211M digital network cameras (shown as in Fig.1). Four cameras are top down view and the other seven are formed as a linear camera array.

•Experiment A-Moving Object Detection

In this experiment, we adopt precision and recall, two widely used metrics to quantitatively evaluating the accuracy of our detection algorithm.The recall corresponds to the detection rate and the precision represents to the trustfulness of a detection. A detection is accepted if it has enough overlapping with the real object.

The dataset was captured in our system, which contains eleven volunteers in a small area of 4mx4m simultaneously.We have manually labeled the ground truth of objects, and evaluate the performance of our multiple constraints based detector on the entire dataset. Fig.6a visualizes the comparison result of recall, precision and f-measure under successive constraints. Please note although complex interactions among people such as merging,splitting and severe occlusion happens frequently in the test dataset, our multiple constraints achieves the f-measure 84.2% from occupancy constraint's 80.5% based on the significant improvement of precision from 68.3% to 92.4%. Fig.6b,c,d show an example the detection results under different constraints. The red and green 3D bounding boxes correspond to real object and false alarms respectively. Note that there are seven ghost objects in occupancy constraint result (see Fig.6b), three in the fusion geometry constraint(see Fig.6c), and no ghost objects in our multiple constraint result (see Fig.6d). Statistical results demonstrate the robustness of the proposed multiple constraints approach.

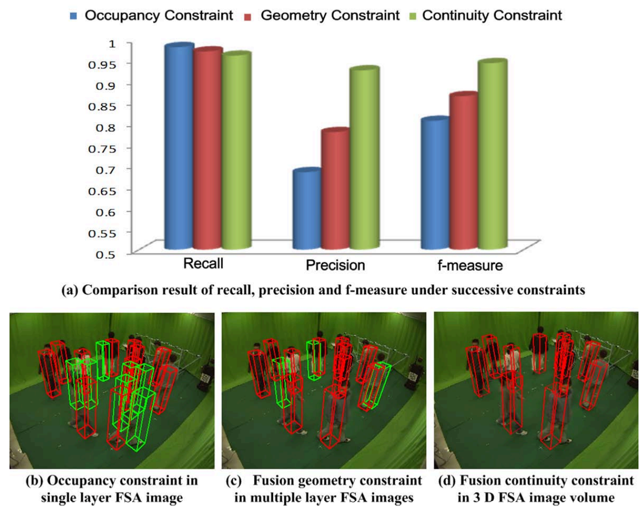


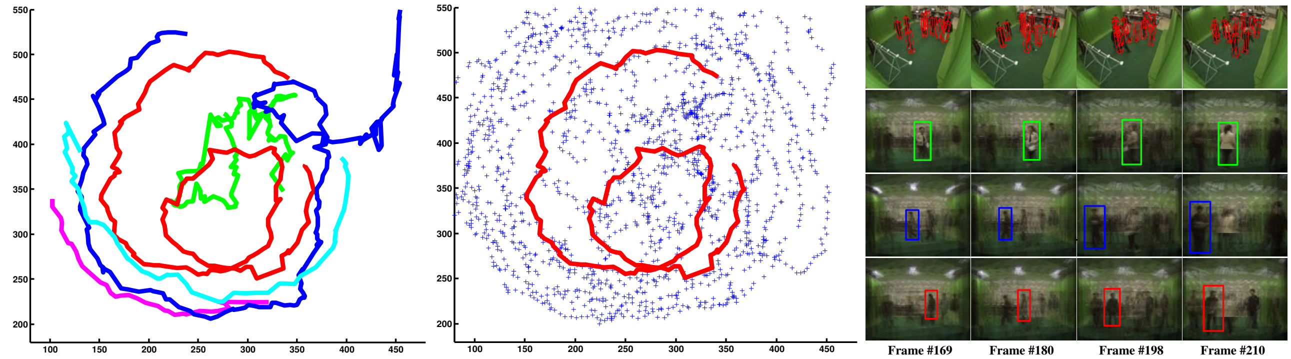
Figure 6. Comparison of object detection result under various constraints. The red and green 3D bounding boxes correspond to real object and false alarms respectively. Note that our multiple constraints successfully filter ghost objects.

•Experiment B-Object Tracking and Imaging

We provide the complete object detection, tracking and imaging results by our approach in Fig.7. In this test dataset, eleven people are moving in a small space, and they are often occluded simultaneously in several views.The color curves in Fig.7a show objects' trajectories on the ground plane by our approach. It can be seen that through seamless interaction of the detection and tracking module, our approach successfully handles the complex crowded condition and continuously track the multiple objects robustly. In particular, we display a single object tracking result in Fig.7b, in which the red curve shows object's trajectory and the blue cross shows the detection results of the other objects at the same period of time. Note that our approach continuously track the given object in the clustered scene. The corresponding locating results in the top-down views and object imaging results during tracking in the color synthetic imaging view are shown as Fig.7c. Note that our approach achieves continuously seeing object even through occlusion, this unique characteristic is especially usefully in surveillance application fields.

6. Conclusions and Future Work

We present a novel hybrid synthetic aperture imaging based algorithm for automatically detection, tracking and seeing multiple people through occlusion with multiple surrounding cameras and a camera array. This approach is designed for complex scene where significant occlusions occur frequently. Extensive experiments show that the proposed foreground synthetic aperture detection, the camera array synthetic aperture imaging and tracking module achieve low false alarm rate and good imaging and tracking result in challenge crowded scene.



(a)Tracking Trajectories (b)Trajectories of single object (c) Detection and imaging results

Figure 7. Multiple people detection, tracking and synthetic aperture imaging results with our algorithm.

In future work we plan to improve our tracking results by using more sophisticated tracking methods in the joint synthetic aperture imaging volume, and explore our work to interesting outdoor surveillance scenarios.

Acknowledgements

This work is supported by the National Natural Science Foundation of China(No.60903126, No.60872145), China Postdoctoral Science Foundation Special Class (No.201003685), China Postdoctoral Science Foundation(No.20090451397), and National High Technology Research and Development Program(863) of China (No.2009AA01Z315).

References

[1] A. Yilmaz, X. Li, and M. Shah. Object tracking: A survey. *ACM J. Computing Surveys*, 2006.

[2] X.G.Wang, K.Tieu, and W.E.L.Grimson. Correspondence-free activity analysis and scene modeling in multiple camera views. *TPAMI*, 32(1):56-71, 2010.

[3] T.Zhao, R. Nevatia, B.Wu. Segmentation and tracking of multiple humans in crowded environments. *TPAMI*, 30(7):1198-1211, 2008.

[4] B. Leibe, K. Schindler, N. Cornelis, and L.V. Gool. Coupled detection and tracking from static cameras and moving vehicles. *TPAMI*,30(10):1683-1698, 2008.

[5] A. Senior, A. Hampapur, Y.L. Tian, L. Brown, S. Pankanti, and R. Bolle. Appearance models for occlusion handling. *Image and Vision Computing*,24(11):1233-1243, 2006.

[6] R.Cucchiara, C. Grana, and G. Tardini. Track-based and object-based occlusion for people tracking refinement in indoor surveillance. *VSSN*,81-87, 2004.

[7] A. Perera, C. Srinivas, A. Hoogs, G.Brooksby, and W. Hu. Multi-object tracking through simultaneous long occlusions and split and merge conditions. *CVPR*,2006.

[8] F. Fleuret, R. Lengagne and P. Fua. Fixed point probability field for complex occlusion handling. *ICCV*,694-700, 2005.

[9] T. Zhao and R. Nevatia. Tracking multiple humans in complex situations. *TPAMI*,26(9):1208-1221, 2004.

[10] D.Ross, J.Lim, R.S.Lin, M.H.Yang. Incremental learning for robust visual tracking. *IJCV*,2007.

[11] M.Li,W.Chen, K.Q.Huang,T.N.Tan. Visual tracking via incremental self-tuning particle filtering on the affine group. *CVPR*,1315-1322, 2010.

[12] S.M. Khan and M. Shah. Tracking multiple occluding people by localizing on multiple scene planes. *TPAMI*,31(3):505-519,2009.

[13] R.Eshel,Y.Moses. Tracking in a dense crowd using multiple cameras. *IJCV*,88(1):129-143,2010.

[14] D.Delannay, N.Danhier, C.D.Vleeschouwer. Detection and recognition of sports (wo)men from multiple views. *ICDSC*,2009.

[15] F. Fleuret, J. Berclaz, R. Lengagne, and P. Fua. Multi-camera people tracking with a probabilistic occupancy map. *TPAMI*,30(2):267-282,2008.

[16] M.Taj, A. Cavallaro. Multi-camera track-before-detect. *ICDSC*,2009.

[17] R.Eshel, Y. Moses. Homography based multiple camera detection and tracking of people in a dense crowd. *CVPR*,2008.

[18] J.C.Ren, J.Orwell, G.A. Jones, M.Xu. Tracking the soccer ball using multiple fixed cameras. *CVIU*,113(5):633-642,2009.

[19] V.Vaish, B.Wilburn, N.Joshi, M.Levoy. Using plane + parallax for calibrating dense camera arrays. *CVPR*,2004.

[20] V. Vaish, R. Szeliski, C.L. Zitnick, S.B.Kang, and M. Levoy. Reconstructing occluded surfaces using synthetic apertures: Stereo, focus and robust methods. *CVPR*,2331-2338, 2006.

[21] B. Wilburn, N. Joshi, V. Vaish, E.V. Talvala, E. Antunez, A. Barth,A. Adams, M. Horowitz, and M. Levoy. High performance imaging using large camera arrays. *ACM Transactions on Graphics*,24(3):765-776, 2005.

[22] N. Joshi, S. Avidan, W. Matusik, and D.J. Kriegman. Synthetic aperture tracking: Tracking through occlusions. *ICCV*,2007.

[23] Z. Zivkovic, F.van derHeijden. Efficient adaptive density estimation per image pixel for the task of background subtraction. *PR Letters*,773-780, 2006.

[24] A. Prati, I. Mikic, M. Trivedi, R. Cucchiara. Detecting moving shadows: formulation, algorithms and evaluation. *TPAMI*,25(7): 918-923, 2003.

[25] R. Cucchiara, C. Grana, M. Piccardi, and A. Prati. Improving shadow suppression in moving object detection with HSV color information. *IEEE ITS*,334-339,2001.