IMPROVING TARGET LOCALIZATION ACCURACY OF WIRELESS VISUAL SENSOR NETWORKS

Wei Li[1,2], Jorge Portilla[2], Félix Moreno[2], Guixuan Liang[2], and Teresa Riesgo[2]

1 Research Institute of Electronic Science and Technology, University of Electronic Science and Technology of China, No. 4, Section 2, North Jianshe Road, 610054, Chengdu, China
2 Centro de Electrónica Industrial, Universidad Politécnica de Madrid, Jose Gutierrez Abascal 2, 28006, Madrid, Spain

Abstract: **This paper discusses the target localization problem of wireless visual sensor networks. Specifically, each node with a low-resolution camera extracts multiple feature points to represent the target at the sensor node level. A statistical method of merging the position information of different sensor nodes to select the most correlated feature point pair at the base station is presented. This method releases the influence of the accuracy of target extraction on the accuracy of target localization in universal coordinate system. Simulations show that, compared with other relative approach, our proposed method can generate more desirable target localization's accuracy, and it has a better trade-off between camera node usage and localization accuracy.**

I. INTRODUCTION

Wireless visual sensor network is a particular type of wireless sensor networks which includes some nodes that are equipped with visual sensors. These visual sensors nodes are responsible for capturing images of targets. They have a unique feature that the targets covered by the camera can be as far away from nodes as they can capture images of targets that are not necessarily in the camera's vicinity. Thus, wireless visual sensor networks are widely used for popular consumer application such as public security, facilities surveillance and monitoring. Most applications of surveillance networks, including event detecting and reporting, rely on the knowledge of target position [1].

Localization is an important part in visual sensor networks, since it provides with coordinates for both sensors and the targets in sensor networks [2]. In this paper, we focus on the problem which the sensor locations are already known. Approaches for sensor node localization, such as received-signal-strength (RSS) [3], can not solve the problem of target localization, because targets are usually passive and un-cooperative in localization. Our objective is to localize a target in the sensing field. Target localization is a technique that is used to estimate a target's position and merge the information regarding location and orientation of other cameras for effective handoff [4]. Generally, researches [1][4][5] on target localization in camera sensor networks are based on accurate image processing, and the position of the target can be extracted perfectly. However, vision-based target localization in camera sensor networks will face a great challenge. First of all, in wireless visual sensor networks, visual sensor node is usually equipped with a low-resolution camera due to the cost limitation [6]. Secondly, the image processing on the sensor node is a work of great challenge. Camera sensors generate a huge amount of data compared to scalar sensors. Such data processing is in general computationally expensive, requires floating point arithmetic, and is costly to implement locally [7]. Thirdly, although the central nodes or the base stations in the visual sensor network have more powerful compute capabilities, transmitting the image data, by generally low-power sensor nodes, is very difficult due to their bandwidth requirements [8]. Therefore, the accuracy of filtering and extraction of target's position relevant information can not be guaranteed in the sensor level.

Meanwhile, visual surveillance using multiple cameras has attracted much attention in the computer vision community. This is because by using multiple cameras, the area of surveillance is enlarged and information from multiple views is extremely helpful to address many issues [9]. For example, the accuracy of the target localization can be gradually improved by selecting the most informative cameras until the required accuracy level of target state is achieved. On the other hand, the limited energy and wireless channel capacity are also the major constraints of camera sensor networks. We have to make a trade-off between the accuracy of localization and network efficiency. Firstly, simple local image processing algorithm that produce only the essential position information is needed. Secondly, fusion algorithms that estimate the real coordinate of the target from the unperfected target extraction of different cameras are also needed.

Our motivation is to use the network's available visual data to determine the target's position that is in the coverage of camera's view. Using statistical method, we want to find the most correlated image point pairs from different cameras to reduce influence of noises, in order to reach the goal of improving the accuracy of target localization. We focus on 2-D target localization on the ground plane. Note that ceiling cameras could have non-occluded view towards the target, which is better than horizontal cameras. However, they are often impractical to deploy and they can only observe a small area limited by the field of view, compared with horizontal cameras. Thus, we assume that the cameras are placed horizontally around a room, which is the most relevant case for many real world applications. Besides, this paper makes the following assumptions about the wireless visual sensor network. First, the location and orientation of each camera node is known within a universal coordinate system. Once a node enters into the networks, its geographical position remains constant. Next, all of the cameras are well calibrated. Finally, all of the nodes are time synchronized.

The rest of this paper is organized as follows. Section II highlights the related works. Section III describes the architectures of wireless visual sensor networks. Section IV presents the standard geometrical pinhole camera model to compute the possible position of target. Section V proposes the technique to find correlated point pair by statistical method. Section VI conducts experiments to verify the effectiveness of our proposed method. The paper concludes with section VII.

II. RELATED WORK

Due to the significant difference in information acquiring and processing methods from conventional sensor networks, the existing localization algorithms [10] cannot solve the target localization issue in visual sensor networks. Recently, research on image sensor networks has received large interest; however, only limited studies of the localization problem have been reported for these networks.

Ryan et al. [11] used a set of two cameras to localize the sensor nodes of a wireless sensor networks. However, they estimated the location of a target using their non-imaging sensors. Liu et al. [1] addressed the problem of localization-oriented optimal cameras selection based upon a tradeoff between the accuracy of target localization and the energy consumption in camera sensor networks. However, they manually extract the target from the background in the image to match the correlated point pair. It is prone to error and impractical in large deployment. Ercan [7] considered the measurement error in target localization by multiple sensors, but they assumed the covariance of measurement noise is known which usually can not be accepted. Oztarak [12] used distance from the camera to the extracted target to get the relative accurate target localization. In most cases, however, it is impossible to acquire this kind of distance information. Spors [13] proposed a solution by integrating different types of audiovisual sensor. The authors tried to reduce the error rate by using multiple sensors. Huiyu [14] presented an approach where target is located based on color-based change detection in the video modality and on the time difference of arrival estimation between the two microphones in the audio modality. Note that both of these methods require processing at a central node, not on the sensors node.

Although geometrical method of computing the coordinate of target is well studied, it is very hard to find correlated image point pairs. Some researchers used Scale-Invariant Feature Transform (SIFT) to find feature point correlations [15]. SIFT is an opportunistic feature point detection and correlation algorithm, but it has high processing and communication costs. Medeiros et al. [16] and Kurillo et al. [17] waved a bar with an LED light at each end to provide correlated feature points by the solutions that use the known length of the bar to fix the units of relative camera positions. Unlike rod-based solution, our solution of target localization does not rely on the additional rod, but uses statistical method to find the most correlated coordinates pair of the target.

III. ARCHITECTURE OF WIRELESS VISUAL SENSOR NETWORKS

Typical example architecture of wireless sensor networks [7] is illustrated in Figure 1. Each visual node includes a camera to capture the video and a processing component to encode the video in a desired way. These visual sensors compose a network, and they communicated with each other within a limited bandwidth. The video captured and encoded at each visual node is transmitted to a base station with computing capability enough for further analysis and decision making.

It is largely agreed that streaming all the data is impractical due to the severe energy and bandwidth constraints of wireless
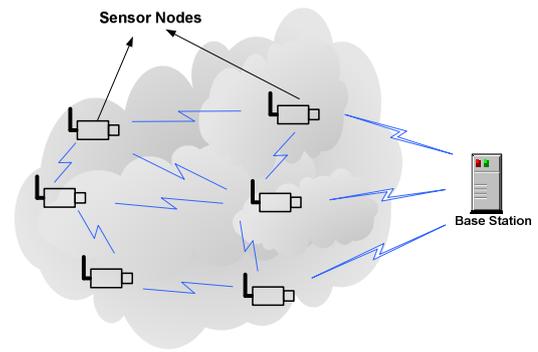

Figure 1. The architecture of wireless visual sensor networks

wireless visual sensor networks. Since processing costs are significantly lower than communication costs, it makes sense to reduce the size of data before sending them to the base station [18]. In this perspective, it is desirable to provide with a simple local processing algorithm that produces only the essential information needed for target localization in wireless visual sensor networks [8]. Then, only the position information of the target is needed to transmit to the base station for target localization in the world coordination. Therefore, it dramatically decreases the burden of data communication. The following is the architecture of wireless visual sensor network where only position information of the target flows. There are two key issues needed to be addressed in this kind of architecture: (1) target's position extraction algorithm in local sensor node, (2) fusion algorithm of target localization at base station. The more accurately the target's position is extracted from the image in the sensor node level, the more the estimated result approaches to the real position. However, it is difficult to realize. It is desirable to design a fusion algorithm to release the influence of the target extraction algorithm at the local sensor level. This process is divided two levels: (1) to merge target's position information originated from multiple cameras together to produce target's coordinate in the world frame; (2) to find an optimal method to improve the accuracy of the target's coordinate.
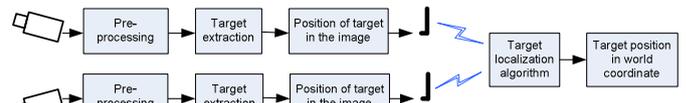

Figure 2. The architecture of wireless visual sensor networks where only position information flows

IV. CAMERA GEOMETRICAL MODELS

It is assumed that the known conditions include positions and projection directions of cameras and only one target is captured by cameras. Since there is only one target to localize, we do not need to consider occlusion from other targets. The principle of computing the coordinates of the target is to estimate the orientation of the target in each image. The orientations in the two images of different cameras are intercrossed in a unique point that is the possible target position. Its coordination can be calculated by geometrical method that will be described in the following.

As shown in Figure 3, each camera is represented as a feature point source within the global *xy*-coordinate system. A camera has a field of view (FOV) that represents the area on the *xy*-plane where a target captured by the camera was been
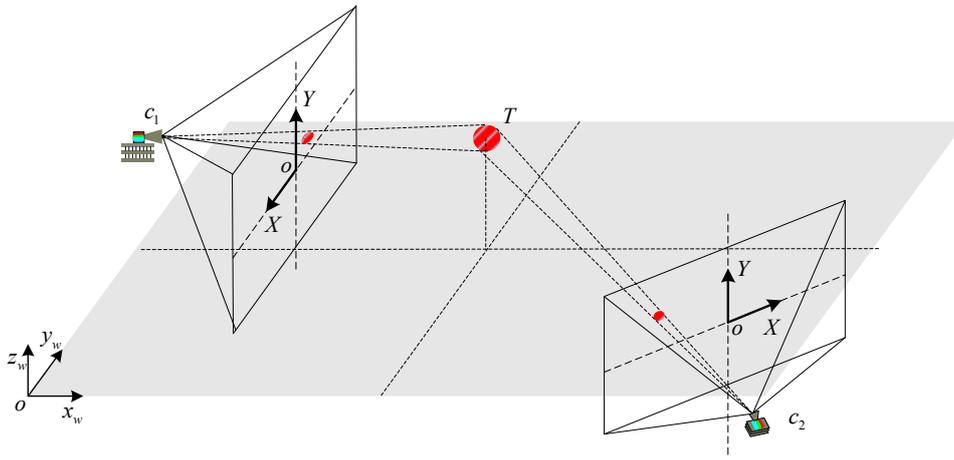
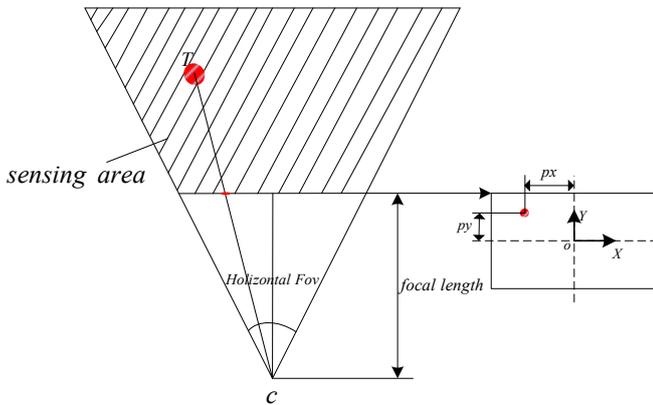figure 3. Model of target localization using two cameras

Figure 4. Geometric algorithm of target localization in the sensor image

located. The FOV is represented as an isosceles triangle where both the equal sides join at the point representing the camera location. The angle between these equal sides is known as the FOV angle and is a factory specification defined for every camera. Our localization scheme works and assumes that cameras are calibrated at system setup time and their orientations are known relative to a global reference frame.

The basic goal of the localization algorithm is to use multiple cameras data to ascertain the location of a target on a global plane, as shown in figure 3. T is the target we want to localize. c_1 and c_2 are the cameras that capture the target. $ox_wy_wz_w$ represent the world coordination and OXY is the image plane. Pyramids in the figure represent the coverage of cameras.

From Figure 3, it can be inferred that the image taken by a camera contains information on two of the three dimensions but the third or depth dimension of each target is lost. If the target is located in the frame, the target can reside anywhere in the FOV. Thus, it is necessary to include at least two cameras to decide the third dimension information which is used for target localization.

The target in the world coordinate system is represented as (p_x, p_y) on the image plane by perspective projection model, as shown in figure 4. p_x and p_y are the observation measurements of the target on the image coordinate system. The process of

computing the orientation of target in the image of camera can be computed according to the following formula:

$$k = \tan\left(\varphi - \arctan\left(\left(2px / p\right) \cdot \tan(\theta / 2)\right)\right) \quad (1)$$

Where k is the orientation of the target, φ is the angle of camera rotating around z_w axis, when the direction of the camera is along with the x_w, $\varphi = 0$, and rotating in counter-clockwise is positive; θ is the horizontal field of view, and p is the number of pixels in horizontal. px is the horizontal pixel coordinate in the image. In our localization scheme, only px is communicated to the base station.

If two cameras capture the same target at the same time, the target orientations generated from two cameras could be intercrossed. We can infer the coordinate of target from the known positions of two cameras and the intersected point of orientations. The computation process is described by the follow equations:

$$AX = b, A = [-k,1], X = [x,y]^T, b = [y_{ci} - kx_{ci}] \quad (2)$$

Where x_{ci}, y_{ci} are the coordinates of the ith cameras in the world coordinate frame. x and y are the pending coordinates of the target in the world coordinate frame. If there is only one camera that could detect the target, the values x and y cannot be uniquely determined because there are two unknowns. Thus, at least two cameras that detect the camera are needed to determine the location of the target. The target position's computation matrices are shown as follows:

$$\begin{bmatrix} -k_1 & 1 \\ -k_2 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} y_{c1} - k_1 x_{c1} \\ y_{c2} - k_2 x_{c2} \end{bmatrix} \quad (3)$$

Then,

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -k_1 & 1 \\ -k_2 & 1 \end{bmatrix}^{-1} \begin{bmatrix} y_{c1} - k_1 x_{c1} \\ y_{c2} - k_2 x_{c2} \end{bmatrix} \quad (4)$$

V. STATISTICAL METHOD OF TARGET LOCALIZATION

The purpose of our method is to reduce the influence of the reverberations and background noise of the image on the accuracy of target localization. In traditional methods, after

target extractions, one feature point would be selected to match the correlated feature point pair which will be involved in the calculating the possible position of the target. However, the location of target generated by the image processing is always corrupted by some additive noise which dramatically affects the accuracy of the target location in the image, specially, when the cameras resolutions are not good enough. If we use this kind of location information to determine the target position in the world coordination, it is very likely that the target will not be where it is really located. Although the accuracy of the image processing may not good enough, the quality of localization also can be improved with multiple measurements from different cameras. As we know, since there is always error between the real position and result generated by the image processing, it is very difficult to improve the accuracy of the target extraction from the background. Thus, it is desirable to find some optimal algorithms to merge candidate coordinates of the target which come from different cameras, in order to get a more accurate target coordinate.

As mentioned before, the first step of target localization is to extract the target from the image of camera nodes. There are some algorithms to extract the target from the video frames, such as the brightest pixel identification algorithm, the circular filter algorithm, and the two-dimensional convolution algorithm. By comparisons, the accuracy of the target extraction of the circular filter algorithm is better than others' according to the works of Massey [19]. Therefore, we use the Canny detector, one kind of circular filter algorithms, to determining the target edge. Then we want to find correlated point pair to compute the real position of target. In traditional method, one feature point has to be selected to represent the target in the correlated point pair matching in wireless visual sensor networks. If the feature point is not selected rationally, or it is corrupted by some noise, the accuracy of the target position can not be guaranteed. To avoid the drawbacks of the above method, our solution selects multiple feature points to represents the target in the image from distinct cameras, then it uses statistical analysis method to find the most correlated feature points pair, in order to improve accuracy of the target localization. On the one hand, multiple feature points ensure the raw data from image processing contains target location information; on the other hand, statistical approach determines which feature point will truly reflect the target position.

The statistical method of merging target's position information originated from multiple cameras is described as follows.

- **Step 1**:

  In our solution, we extract the edge of target from the background by Canny edge detector. The edge of target is a finite set of m points. We select two feature points to represent the target: (1) the center of the target edge; (2) the intersection of middle line of the leftmost and rightmost edge point in horizontal and the middle line of the highest and the lowest edge point in vertical is defined as another pending position of the target. Only the horizontal pixel positions of these two feature points are communicated to the base station.

  The center of the target can be formulized as follows:

$$
\begin{cases}
px = \sum_{i=1}^{m} px_i / m \\
py = \sum_{i=1}^{m} py_i / m
\end{cases}
\tag{5}
$$

px and py mean the pending coordinate in horizontal and vertical respectively, px_i and py_i are the horizontal and vertical pixel positions of the edge points of the target.

The other pending position of the target can be written as follows:

$$
\begin{cases}
px = \dfrac{px_l + px_r}{2} \\
py = \dfrac{py_u + py_b}{2}
\end{cases}
\tag{6}
$$

px_l and px_r are the leftmost and rightmost edge point in horizontal, py_u and py_b are the highest and the lowest edge point in vertical. Only the px is needed to communicate to the base station to make the further computation.

- **Step 2**

  It chooses one feature point of the target from each candidate camera that captures the target, and it makes the target coordinate pair. Then it uses the target localization algorithm to generate possible coordinates of the target. We assume that n cameras capture the target simultaneously. There will be $C_n^2 C_2^1 C_2^1$ possible target position.

- **Step 3**

  Since every two different cameras could generate a possible coordinate of the target, after step 2, we would have 2^n possible coordinate pairs of the target. Each feature point pair has P_n^2 possible target coordinates. Then, it computes the expectation and the mean square error of these coordinates of the target. The expectation of the pending coordinates of target is formularized as follows:

$$
\begin{cases}
X = \dfrac{x_1 + x_2 + \cdots + x_n}{n} \\
Y = \dfrac{y_1 + y_2 + \cdots + y_n}{n}
\end{cases}
\tag{7}
$$

X and Y are the expectation of the pending position of the target, x_i and y_i are the horizontal and the vertical coordinates of the pending position of target respectively. The mean square error of a set of pending targets positions can be expressed as follows:

$$
MSE = \dfrac{\sqrt{\dfrac{\sum_{i=1}^{n} (x_i - X)^2 + \sum_{i=1}^{n} (y_i - Y)^2}{2}}}{n}
\tag{8}
$$

- **Step 4**

  Go back to step 2, it selects different feature points of the target from candidate cameras. Then it continues the process until all the feature points are utilized.

- **Step 5**

After step 4, we would have 2^n expectations and mean square error. Then we utilize the Mean-Variance Model to get the minimum variance. In our model, we use mean square error to substitute for the variance, and use expectation of pending positions to substitute for the mean.

It makes a comparison of the mean square error of the possible coordinates, and it finds the minimum one. Then the expectation corresponding to the minimum mean square error is set as the coordinate of the target. By comparison, the greater mean square error of the pending target positions means that the raw data which is used to generate the target position contains much noise information. Since the noise of image in different cameras is independent with each other, the mean square error reflects the correlated extend of pending feature points. The little mean square error of the set of pending target positions means that most of feature points from different cameras represent the same feature point pair. Therefore, the mean of the pending target positions with little mean square error could be regarded as the coordinate of target in the world coordination.

VI. CASE STUDY AND SIMULATIONS

We perform simulation experiments by using C program language to implement the proposed approach. Due to the space limitation of experiment, we only deployed 4 cameras in an area of 100*100 centimetres, as shown in figure 5. A simulation scenario of deployment of cameras and target is shown in Figure 5. The positions and the direction of projection of cameras are summarized in table 2.

The images of the target taken from the four cameras are shown in figure 6. Then we use Canny edge detector to extract edge of target. According to Eq (5) and (6), we can get two feature points of the target. The calculated results are shown in Table 4. The second column shows the horizontal coordinate of the geometry center of the target. The third column shows the horizontal coordinates of middle line of the leftmost and rightmost edge points.

According to Eq(4), we can get the possible target position in the world frame. As each camera generates two feature points, we can get 24 possible positions of the target. Then using Eq(7) and Eq(8), positions of the target can be estimated. In order to verify the effectiveness of the proposed method, we also use relative approaches presented by Liu [9] to localize the target.

Figure 5. Topology of the experiment

(1) image from camera 1   *(2) image from camera 2*
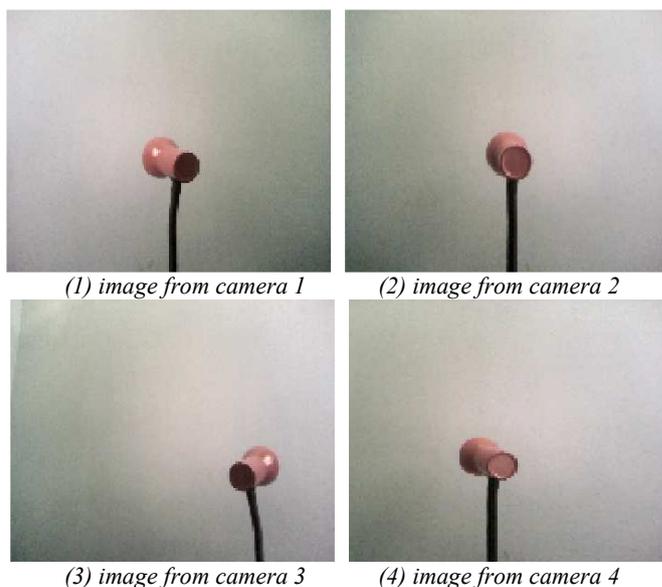
(3) image from camera 3   *(4) image from camera 4*
Figure 6. Images taken from four cameras

Table 2. Geometrical Parameters of Cameras

	Position (*cm*)	Projection
Camera 1	(20, 30)	45°
Camera 2	(60,0)	90°
Camera 3	(100,0)	135°
Camera 4	(40,10)	60°

Table 1. Parameters of Cameras

Parameter	value
FOV in horizontal	57.4°
Image format	CIF
Pixel in horizontal	352

Table 3 multiple feature points of target extracted by cameras

	Feature point 1	Feature point 2
Camera 1	26	30
Camera 2	22	29
Camera 3	102	96
Camera 4	-50	-46

Table 4 result of target localization

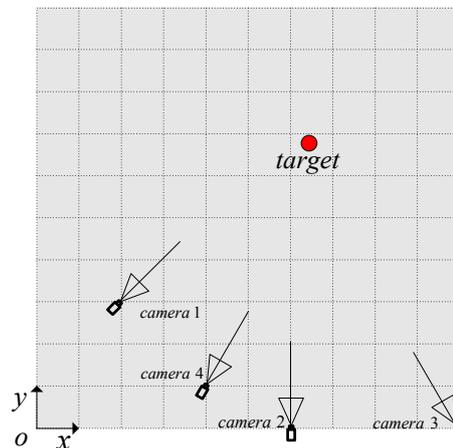|	target position(*cm*)	relative approach (*cm*)	Error (*cm*)	Our approach (*cm*)	Error (*cm*)
2 cameras	(64,68)	(63.05,65.75)	(-0.95,-2.25)		
3 cameras	(64,68)	(64.89,65.14)	(0.89,-2.86)	(64.30, 66.28)	(0.30,-1.72)
4 cameras	(64,68)	(64.41,67.22)	(0.41,-0.78)	(63.92, 67.68)	(-0.08,-0.32)

If only two cameras capture the target, for example, camera 1 and camera2, the target position computed by relative approach is (63.05,65.75) whose feature points' pair is 30 from camera 1 and 29 from camera 2. It is about 2.44 *cm* deviated from the real position. Note that it is impossible to compute the mean square error of pending target positions by the multiple feature points' method when there are only two cameras involved in the computation. When the 3 cameras capture the target, the estimated result by relative approach is (64.89, 65.14) whose feature points' pair is 30 from camera 1, 29 from camera 2 and 96 from camera 3. Error between the estimated result and the real position is about 2.99 *cm*. In contrast, by our proposed approach, the computed result is (64.30, 66.28) whose feature points' pair is 30 from camera 1, 22 from camera 2 and 102 from camera 3. The error is about 1.75 *cm*. If all the four cameras capture the target, the target position generated by the relative approach is (64.41, 67.22), whose feature points' pair is 30 from camera 1, 29 from camera 2, 96 from camera 3 and -46 from camera 4. The error is about 0.88 *cm*. Using our method, the result is (63.92, 67.68), whose feature points' pair is 26 from camera 1, 22 from camera 2, 102 from camera 3 and -46 from camera 4. The error is about only 0.33 *cm*. This case indicates that our proposed method could get more accurate localization of target than the other approach when the target is captured by the same amount of cameras.

The amount of cameras affects the accuracy of target localization. For example, the accuracy of target localization computed from 4 cameras is better than that produced from 3 cameras by our approach. That is because the more cameras capture the target, the less influence of error would play on the result. In the relative approach, although there is also improvement of localization accuracy by using more cameras, it needs more cameras than our proposed method to reach the same accurate level. Therefore, our method has better tradeoff between localization accuracy and camera sensor usage.

VII. CONCLUSIONS

This paper introduces the technique of multiple feature points to compute the target location in the wireless visual sensor networks. This technique combines the statistical method with image processing algorithm. Target extraction algorithm determines two feature points to present positions in the image at each camera node. Only the coordination information of the feature points is communicated to the base station. In the base station, pinhole camera model is applied to compute the target localization in the world frame. Due to error's influence on the localization accuracy, we provide statistical method to find the most correlated point pair to determine the target position. While in this paper a stationary target has simulated, four cameras nodes are involved in exterminating the target position. Simulations show that our proposed approach could get more desirable localization accuracy than other relative approach in the same amount of cameras. Besides, in terms of tradeoff between camera node usage and localization accuracy, our proposed method achieves better performance than other relative approach.

Future works are oriented to analyze target extraction algorithms in order to further relax the computing burden in node level. In addition, we also plan to develop multiple targets localization algorithm for tracking targets in a wireless visual sensor network, based on the Cookies platform [20].

VIII. REFERENCE

[1] Liang Liu; Huadong Ma; Xi Zhang, "Optimal Camera Selection for Target Localization in Camera Sensor Networks". 42nd Annual Conference on Information Sciences and Systems, 19-21 March 2008. Princeton, NJ, USA. pp.140-145

[2] Stanislava Soro; Wendi Heinzelman, "A Survey of Visual Sensor Networks", Advances in Multimedia, vol 2009, pp.1-21

[3] Gang Wang, Kehu Yang, "A New Approach to Sensor Node Localization Using RSS Measurements in Wireless Sensor Networks," IEEE Transactions on Wireless Communications, vol. 10, no. 5, pp. 1389-1395, 2011

[4] Purushottam Kulkarni, "Senseye: A Multi-tier Heterogeneous Camera Sensor Network," University of Massachusetts, Feb, 2007

[5] H. Lee and H. Aghajan, "Vision-Enabled Node Localization in Wireless Sensor Networks," Proc. of the Conference on Cognitive Systems with Interactive Sensors, Paris, France, Mar. 2006.

[6] Akyildiz. I.F, Melodia. T, Chowdury. K.R, "A Survey on Wireless Multimedia Sensor Networks," IEEE Wireless Communications, vol. 14, issue. 6, Dec, 2007, pp. 32-39

[7] A. Ercan, D. Yang, A.E. Gamal, L. Guibas, "Optimal Placement and Selection of Camera Network Nodes for Target Localization," in: DCOSS, 2006, pp. 389–404

[8] Youssef Charfi, Naoki Wakamiya, Masayuki Murata, "Challenging Issues in Visual Sensor Networks," IEEE Wireless Communications, vol. 16, issue 2, April 2009, pp. 44-49

[9] Liang Liu, Xi Zhang, Huadong Ma, "Optimal Node Selection for Target Localization in Wireless Camera Sensor Networks," IEEE Trans on Vehicular Technology, vol.59, no.7, 2010, pp.3562-3576

[10] S.Rajeev, A.Ananda, C.Mun, T.Wei, "Mobile, Wireless, and Sensor Networks: Technology, Applications, and Future Directions," John Wiley and Sons, 2005

[11] R. Farrell, R. Garcia, D. Lucarelli, A. Terzis, I-J. Wang, "Target localization in camera wireless networks," Pervasive and Mobile Computing, Special Issue, vol 5, issue 2, April 2009, pp. 165-181

[12] Oztarak. H, Akkaya. K, Yazici. A, "Lightweight Object Localization with a Single Camera in Wireless Multimedia Sensor Networks," IEEE Global Telecommunications Conference, Hawaii, USA, Dec. 2009

[13] S. Spors, R. Rabenstein, N. Strobel, "A Multi-sensor Object Localization System," Proceedings of the Vision Modeling and Visualization Conference, Stuttgart, Germany, Nov 2001, pp. 19-26,

[14] Z.Huiyu, M. Taj, A.Cavallaro, "Target Detection and Tracking with Heterogeneous Sensor," IEEE Journal of Selected Topics in Signal Processing, vol. 2, no. 4, 2008, pp. 503-513

[15] Dhanya Devarajan, Richard J. Radke, Haeyong Chung, "Distributed Metric Calibration of Ad Hoc Camera Networks," ACM Transactions on Sensor Networks, vol. 2, Aug 2006, pp.380-403

[16] H. Medeiros, H. Iwaki, J. Park, "Online Distributed Calibration of a Large Network of Wireless Camera Using Dynamic Clustering," ACM/IEEE International Conf' on Distributed Smart Cameras, 2008

[17] Kurillo. G, Zeyu Li, Bajcsy. R, "Wide-area External Multi-camera Calibration Using Vision Graphs and Virtual Calibration Object," Second ACM/IEEE International Conference on Distributed Smart Cameras, Stanford, USA, Dec. 2008. pp.1-9

[18] G.J. Pottie and W. J. Kaiser, "Wireless Integrated Network Sensors," Commnun. ACM, vol.43, May 2000, pp. 51-58.

[19] T.Massey, R.Kapur, F.Dabiri, L.N.Vu, M.Sarrafzadeh. "Localization Using Low-Resolution Optical Sensors". The 4th IEEE International Conference on Mobile Ad-Hoc and Sensor Systems, Italy. 2007

[20] Y. Krasteva, J.Portilla, E. de la Torre, T. Riesgo, "Embedded Run-time Reconfigurable Nodes for Wireless Sensor Networks Applications," IEEE Sensors Journal, issue:99, 2011