

Target Tracking in Wireless Visual Sensor Networks: Challenges, Steps, and Metrics of Evaluation

Afaf Mosaif^a, Said Rakrak^b

^aLaboratory of Computer and Systems Engineering (L2IS), Faculty of Science and Technology, Cadi Ayyad University, Marrakesh, Morocco. E-mail: afaf.mosaif@edu.uca.ac.ma

^bLaboratory of Computer and Systems Engineering (L2IS), Faculty of Science and Technology, Cadi Ayyad University, Marrakesh, Morocco. E-mail: s.rakrak@uca.ma

Article History: Received: 11 January 2021; Accepted: 27 February 2021; Published online: 5 April 2021

Abstract: In recent years, wireless sensor networks have been used in a wide range of applications such as smart cities, military, and environmental monitoring. Target tracking is one of the most interesting applications in this area of research, which mainly consists of detecting the targets that move in the area of interest and monitoring their motions. However, tracking a target using visual sensors is different and more difficult than that of scalar sensors due to the special characteristics of visual sensors, such as their directional limited field of view, and the nature and amount of the sensed data. In this paper, we first present the challenges of detection and target tracking in wireless visual sensor networks, then we propose a scheme that describes the basic steps of target tracking in these networks, we focus then on the tracking across camera nodes by presenting some metrics that can be considered when designing and evaluating this type of tracking approaches.

Keywords: Object Detection, Target Tracking, Visual Sensor Networks, Wireless Sensor Networks.

1. Introduction

The continuous evolution of wireless sensor networks (WSNs), imaging technologies, and the availability of low-cost CMOS cameras make wireless visual sensor networks (WVSNs) an important and active research area. They consist of a large number of low-power battery-operated camera nodes that are capable of collecting a large amount of image/video data from a monitored site and then processing it collaboratively and transmit just the useful information to each other and to the Base Station (BS) (also called the sink) for further analysis. Different from the 2D sensing range of scalar WSN, the camera nodes in WVSNs are characterized by a limited and directional 3D viewing volume called the Field of view (FoV) (Soro and Heinzelman, 2009). In 2D space, the FoV depends on the camera opening angle α , its direction \vec{d} , and its depth of view R_v as shown in Figure 1.

WVSNs support a large number of new vision-based applications, such as smart homes, smart meeting rooms, and telepresence.

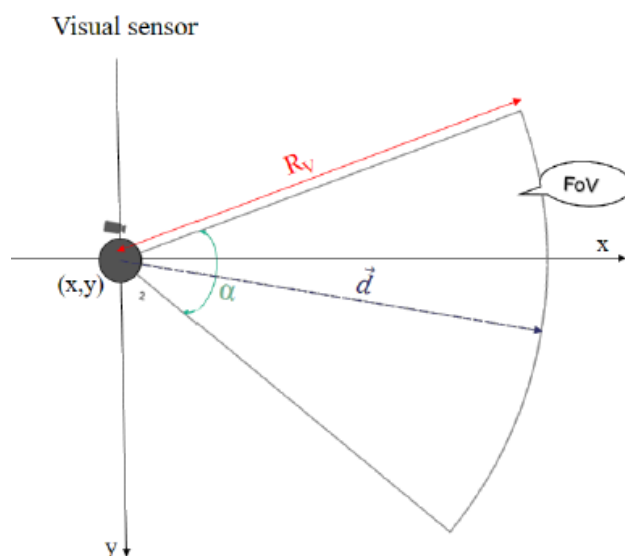


Figure 1. Visual Sensor Field of View

However, object(s) tracking is one of the most interesting applications in this type of networks, which mainly consists of detecting mobile targets in an area of interest and monitoring their motions. Figure 2 presents a scenario of target tracking in wireless visual sensor networks.

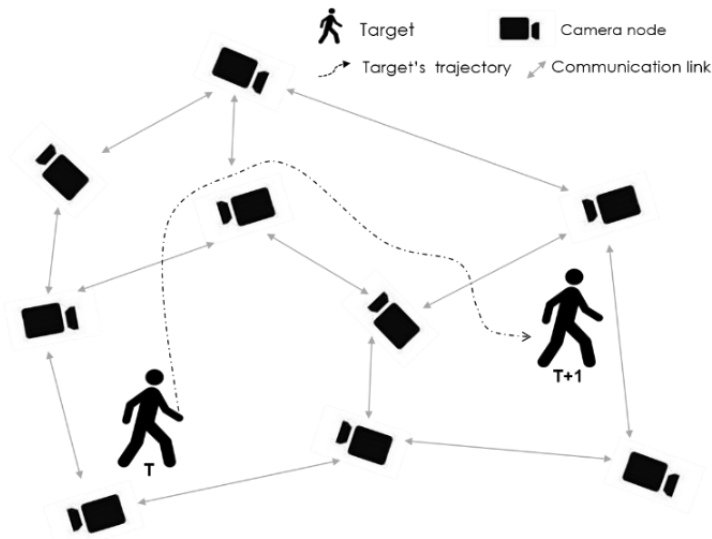


Figure 2. Scenario of Target Tracking in Wireless Visual Sensor Networks

Usually sensing tasks in traditional target tracking methods is performed by one camera node at a time which may result in power loss and induces less accuracy and heavy computation burden on that node (Ramya, Kumar and Rao, 2012). While in WVSNs, the use of multiple cameras provides different views of the scene, which enhances the reliability of the captured events (Soro and Heinzelman, 2009) and gives better results in terms of energy-saving and accuracy, due to the collaboration between nodes. And unlike the traditional camera-based systems that require several kilometers of cables connecting each camera to a central BS, the deployment of cameras in WVSN is easier which makes the extension or the reparation of the system possible by just adding new cameras without the need of new cables installation (Peixoto and Costa, 2017).

Several applications use object tracking such as video surveillance, robot vision, traffic monitoring, etc., and most of these applications are in security and military contexts.

This paper presents an overview of target tracking in WVSN, and especially focuses on the tracking across camera nodes by presenting some metrics that can be taken into account when designing and evaluating this type of tracking approaches.

The remainder of this paper is organized as follows. Section 2 presents the challenges of detecting and tracking targets in WVSNs. Section 3 shows our proposed scheme of the basic steps of target tracking in WVSN and describes the metrics that can be considered when designing a solution for tracking across camera nodes, followed by the presentation of some metrics that can be used to evaluate the performance of this type of tracking approaches in Section 4. Finally, the conclusion is presented in Section 5.

2. Detection and Target Tracking Challenges

Target tracking involves the detection and localization of objects that occupy the monitored space. Therefore, object detection presents the first step of visual data processing and is mostly based on lightweight background subtraction algorithms (Soro and Heinzelman, 2009). These algorithms consist of measuring deviations from a built model of the background scene for each image frame in the video stream. In realistic environments, there are several problems in accurately detecting regions containing moving objects explained in (Javed and Shah, 2008, chap. 2) such as:

- **Gradual or Sudden Illumination Changes** that cause deviation from the background model by altering the appearance of the scene which results in false detection of foreground pixels or in the worst case, the whole image appears as foreground.
- **Uninteresting Movement** such as reflection of moving objects from wet or shiny surfaces or motion of uninteresting objects like flags waving.

- **Shadows** casted by objects might be classified as foreground due to the illumination change in the shadow region.
- **Initialization with Moving Objects** makes accurate modeling of background in no longer possible because a part of the background is occluded by moving objects present during initialization.
- **Camouflage** occurs when an object is very similar to the background, therefore, no regions of interest will be marked because there will not be any significant deviation from the background model.
- **Relocation of the Background Object** induces change in two different regions in the image, its newly acquired position that should be identified as a foreground region and its previous position. While both are detected foreground by any background subtraction system based on color variation.

The detected objects can be then classified in various categories such as vehicles, humans, birds, etc. In a realistic surveillance scenario, the **large variation in object appearance** such as variety of shapes and sizes, etc., and the **wide variety of viewing conditions** such as the illumination in the scene, make classification difficult and more challenging (Javed and Shah, 2008, chap. 3).

Once the objects have been detected and classified, it is useful to track them to know where they are in the image at each instant in time (Javed and Shah, 2008, chap. 1). Tracking algorithms aim at estimating the trajectory of a moving object in an interested area. One of the major issues that a tracking algorithm needs to solve is occlusion (Javed and Shah, 2008, chap. 4). Which occurs when the camera node cannot capture the desired object even if this object is within its field of view, because an obstacle (object/structure) is blocking its view. Therefore, the position and velocity of the occluded object cannot be determined which leads to discontinuity in the observation of the object.

The use of multiple camera nodes in WVSNs induces problems such as **camera handoff**, also known as “**Camera Handover**”, which refers to the ability to identify an object from a camera node to another one that shares or not a considerable common field of view.

Other challenges that can affect the target tracking quality in WVSNs are presented in (Ez-Zaidi and Rakrak, 2016) such as:

- **Node Failure** due to hardware failures, battery exhaustion, physical disasters, etc.
- **Target Missing and Recovery** caused by prediction errors, sudden speed or changes in target trajectory, etc.
- **Tracking Latency** is another challenge to deal with, because the moving target may change its location when the execution of target tracking algorithms takes too long.
- **Coverage and Connectivity** have to be considered because high coverage results in high tracking accuracy.

In particular, compared with the traditional WSNs that can only transmit scalar information such as temperature, WVSNs are more resource-constrained in terms of energy, bandwidth, memory and processing power. Therefore, it is important to eliminate data redundancy by using **data aggregation** (fusion) mechanism, to minimize the data load that has to be communicated among the cameras and to use lightweight processing algorithms (Soro and Heinzelman, 2009) while preserving **energy**.

3. Target Tracking Steps

In the literature, there is no standardized classification of target tracking algorithms in WVSNs. In Figure 3 we present our proposed scheme that describes the basic steps of target tracking in WVSNs with the classifications of some methods presented in previous researches (Himani S. Parekh¹, Darshak G. Thakore², 2014; Ez-Zaidi and Rakrak, 2016). To the best of our knowledge, there are no other authors in the literature who have presented such a scheme, which regroups the metrics of both the tracking across and within the camera nodes.

As shown in Figure 3, there are two types for object tracking in WVSNs. The first one is the tracking within the camera node, which is applied when it detects an object to track in its field of view. In this type, different vision and image processing algorithms can be used depending on the tasks assigned to the camera node (e.g. vehicle detection, face recognition, etc.) and depending on the number of objects to be tracked (single or multiple). The second type is the tracking across camera nodes, which is applied to keep tracking the object when is out of the field of view of a camera node and moves in the area of interest across the cameras of the network. This type of tracking is more challenging than within camera tracking because, in addition to the significant differences that may exhibit in the appearance of the tracked object (due to illumination variation

from a camera to another), there are other metrics that must be taken into account during the object tracking process. Therefore, in this paper, we will focus on the tracking across-camera nodes, and hence we will present the metrics to consider in this type of tracking.

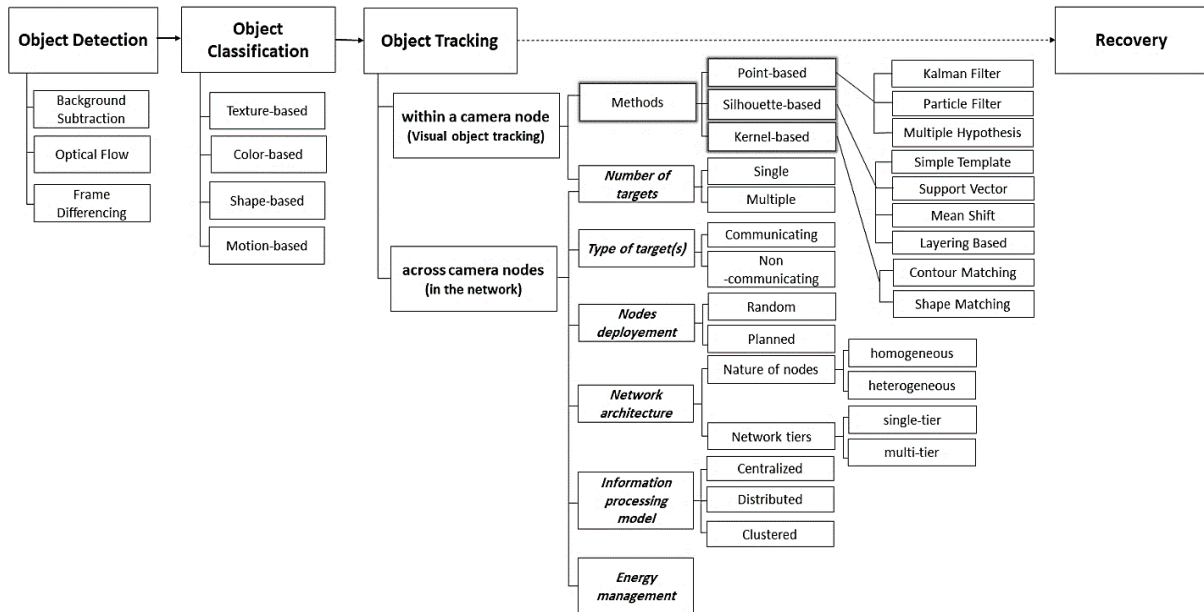


Figure 3. Classification and Basic Steps for Target Tracking in WWSN

3.1. Nodes Deployment

Nodes deployment is a preliminary step in which the sensor nodes are deployed in the area of interest in order to ensure a given coverage quality. This step is so important since it performs the coverage and the connectivity of the network which will affect the tracking accuracy. In the literature, there are two main deployment strategies: Random and planned (Boulanouar *et al.*, 2015). In the random strategy, the camera nodes are deployed in a random manner and it is used when nothing is previously known about the region of interest such as in a region that is impossible for a human being to access. This strategy is suitable for some outdoor applications such as military applications. The planned strategy is used when the camera nodes performances are affected by their location details. Therefore, the camera nodes are placed following a predefined method (e.g. Voronoi-based (Sung and Yang, 2014) and Virtual-Forces-based (Deng *et al.*, 2019) methods). This strategy is often used in indoor applications.

Maximizing the network coverage while minimizing coverage holes and overlap covered areas and thus minimizing the number of nodes deployed in the network are the main objectives of the Planned deployment strategies. Therefore, there are some parameters that have to be considered in such strategies including the type (e.g. PTZ (Pan, Tilt, and Zoom), rotatable or mobile camera), the communication range, the Field of view, the orientation, and the occlusion of each deployed camera.

3.2. Network Architecture

In (Araújo *et al.*, 2015), the authors claimed that the choice of a wireless multimedia sensor network (WMSN) architecture depends on the network application characteristics and they classified the WMSN architecture into three main classes: Single-tier flat architecture, single-tier clustered deployment architecture, and multi-tiered network with heterogeneous multimedia sensors. Therefore, in our scheme, we choose *the nature of the network nodes (homogeneous or heterogeneous nodes)* and the number of the network tiers (single-tier or multi-tier) as the main parameters to choose a suitable WWSN architecture for its tracking system.

In WWSN *homogeneous architecture*, the nodes have the same physical capabilities such as energy, processing, storage, etc. Whereas, in WWSN *heterogeneous architecture*, the nodes have different capabilities and therefore, some nodes may be equipped with more processing power, memory, or energy, compared to the others. Also, the nodes in the network may generate different types of data such as images, videos, or scalar data (Chew *et al.*, 2013), (Aghajan and Cavallaro, 2009).

Depending on the number of the WWSN tiers, these architectures can be classified into two categories (Chew *et al.*, 2013): *single-tier* and *multi-tier* architectures.

In homogeneous architecture, the nodes are often organized in a single-tier flat architecture and perform the same sensing function and transmit the collected data to the BS. However, flat architectures are not suited for computer vision applications, such as target tracking due to the huge amount of data captured by the camera nodes that need to be processed and transmitted to the BS, which may be overloaded and affect the performance of the network by causing latency in both communication and tracking process. Therefore, hierarchical networks (Cluster-based) are used in order to reduce the data flow transmitted to the BS depending on the number of tiers (single-tier or multi-tier) in the clustering hierarchy.

The nodes in Cluster-based networks are organized into groups, called clusters, that can be formed statically at the network deployment stage or dynamically as the target moves. Each cluster contains a cluster head and member nodes. The cluster heads have more processing and communication capabilities and serve as relays to extract useful information from the data collected by the cluster members (the notion of data aggregation) and transmit them the BS.

3.3. Information Processing Model

Target tracking can also be studied from the processing model aspect. According to (Aghajan and Cavallaro, 2009), we can distinguish three types of processing models: centralized processing, distributed processing, and clustered processing.

In *the centralized processing model*, all the camera nodes that detect the target in their FoV send the information to the BS, which takes the appropriate decisions after analyzing the received data. This processing model resembles the one in the traditional WSN architecture, the difference is that the nodes in WWSN can perform some simple local processing to reduce raw image/video data transmitted to the BS, by extracting and transmitting just the important features from the sensed data. Therefore, the amount of processing at the BS is greatly reduced, and the latency performance is improved.

In *the distributed processing model*, each camera node is autonomous in decision making and collaborates with the other camera nodes in the network to track the target. This model is more robust than the centralized one, however, it is important to minimize the data load that has to be communicated among the camera nodes and to use lightweight processing algorithms at the camera nodes level as they are resource-constrained.

In *the clustered processing model*, the camera nodes are grouped into clusters where the camera nodes in each cluster can collaborate and send the important data the cluster head. This latter can perform more complicated tasks such as data fusion and target tracking in the area covered by its cluster and then transmit the processed data to the BS. Each cluster can also collaborate and share data with its neighboring clusters to track the target in all the area covered by the WWSN. The advantage of this model is that the processing is dispatched across the network levels which not only balances the traffic load but also reduces the energy consumption for communication and improves scalability when the network grows (Chew *et al.*, 2013).

3.4. Energy Management

Since target tracking applications consume a lot of power, and camera nodes in most cases run on batteries that are typically not rechargeable, power efficiency is a critical issue in WWSN. Therefore, it is important to take into account the energy consumed in the network during the tracking process not only to avoid target loss due to the energy hole problem (which occurs when the nodes die out early due to excessive load or complex local processing) but also to maximize the lifetime of the entire network.

Many techniques can be used to prolong the lifetime of the network, such as minimizing the number of active camera nodes in the network by activating just the nodes that will participate in the target tracking process (e.g. (Boulanouar *et al.*, 2013), (Li, 2013), (Sabokrou, 2013)), using lightweight processing algorithms at the camera nodes level, minimizing the amount of data transmitted in the network by choosing the suitable processing model according to the nature of the nodes used and the size of the network, and managing the network according to the energy of its nodes, for example, selecting low energy node for performing sensing tasks and short-range communication and nodes with higher energy for data processing and long-range communication.

3.5. Type of Targets

The type of target must be also considered. In the literature, there are two types of targets: the *communicating* and the *non-communicating* ones. The difference between these two types is that communicating targets are equipped with a communication module (e.g. GPS) that allows it to transmit signals and communicate with the network, which facilitates the tracking process (Boulanouar *et al.*, 2013).

3.6. Number of Targets

Based on the number of targets to track in the area of interest, target tracking approaches can be divided into two types: *single* and *multiple target* tracking. Single *target* tracking approaches are in general energy-efficient because only a low traffic load is generated in the network during the target tracking process. Whereas, *multiple targets* tracking approaches are more complex and challenging, because multiple objects have to be tracked in the area of interest at the same time, and these objects may have different directions, positions, classes (e.g. pedestrian, bicycle, car), and speed variations. In addition, the traffic generated during the tracking process is much more important than the single target tracking process.

4. Evaluation Metrics

In the literatures (Boulanouar *et al.*, 2015), (Boulanouar *et al.*, 2013) (Boulanouar *et al.*, 2014), (Lu *et al.*, 2019), some metrics have been used to evaluate the performance of tracking algorithms, such as:

- *The tracking accuracy*: is used to evaluate the accuracy and effectiveness of the proposed target tracking solution. It can be calculated by counting the number of positions (coordinates) retrieved for a given mobile target trajectory or by calculating the average deviation, which is obtained by calculating the Euclidean distance between the real position of the given mobile target at a given instant and that calculated by the proposed tracking approach at the same instant.
- *Tracking latency*: represents the time between the actual entry of the target into the area of interest and the time of its first detection by a camera node. This time depends on the local processing time of the camera node which also depends on the hardware characteristics of the node (e.g. computation speed, message sending/receiving). This metric is also related to the target movement speed. Therefore, the execution of target tracking algorithms must be performed rapidly while preserving positioning accuracy. Because, if the tracking operation takes too long, the moving node may change its location.
- *The number of active camera nodes*: represents the efficiency of the camera nodes utilization in the network during the tracking process. A small number of activated camera nodes means minimal consumption of network resources, which reflects the good performance of the tracking algorithm.
- *The number of exchanged messages*: represents the communication cost of the deployed tracking algorithm. It is obtained by calculating the number of collaborative messages exchanged during the tracking process according to the total number of nodes. This number has not only an impact on the network overload rate but also on the network energy consumption. Therefore, it must be kept to a minimum.
- *Energy consumption*: in WWSN, the lifetime of the network depends on the energy consumed by the deployed tracking algorithm. Therefore, this is an important evaluation metric that can be evaluated in terms of energy consumed, residual energy, or the average lifetime of the nodes in the network. Many factors directly influence the energy such as the number of messages exchanged during the tracking process, the type of data processed at the camera node level (e.g. image, video, etc.), the number of active camera nodes in the network, etc.
- *Target-loss rate*: represents the number of times a target is not detected while it is within the FoV of a camera node. This loss may be due to errors in the camera sensor wake-up mechanism and therefore errors in the target detection, or to an error in the prediction of the target location.

5. Conclusion

Target tracking in wireless visual sensor networks is an active and challenging research area. However, in the literature, there is no standardized classification of target tracking algorithms in this type of network. Therefore, this paper presented our proposed scheme, which describes the basic steps and metrics that can be used for the classification of WWSN target tracking approaches. In addition, the challenges of target detection/tracking in WWSN, and the metrics that can be considered when designing and evaluating tracking across camera nodes solutions are presented. By doing so, we created a mini-platform over which new research can be built.

References

1. Aghajan, H., & Cavallaro, A. (2009). *Multi-Camera Networks*, Multi-Camera Networks. Elsevier.
2. Araújo, T. (2015). *Wireless Sensor Multimedia Networks*, Wireless Sensor Multimedia Networks. Edited by M.M.A. Azim and X. Jiang. CRC Press.
3. Boulanouar, I., Lohier, S., Rachedi, A., & Roussel, G. (2013). PTA: A predictive tracking algorithm in wireless multimedia sensor networks. In *Global Information Infrastructure Symposium-GIIS*, 1-6.
4. Boulanouar, I., Lohier, S., Rachedi, A., & Roussel, G. (2014). A collaborative tracking algorithm for communicating target in *Wireless Multimedia Sensor Networks*. In *2014 7th IFIP Wireless and Mobile Networking Conference (WMNC)*, 1-7.
5. Boulanouar, I., Lohier, S., Rachedi, A., & Roussel, G. (2015). DTA: Deployment and Tracking Algorithm in *Wireless Multimedia Sensor Networks*. *Ad Hoc Sens. Wirel. Networks*, 28(1-2), 115-135.
6. Ang, L.M., Seng, K.P., Chew, L.W., Yeong, L.S., & Chia, W.C. (2013). *Wireless multimedia sensor networks on reconfigurable hardware*. Heidelberg: Springer.
7. Deng, X., Yu, Z., Tang, R., Qian, X., Yuan, K., & Liu, S. (2019). An optimized node deployment solution based on a virtual spring force algorithm for wireless sensor network applications. *Sensors*, 19(8), 1817.
8. Ez-Zaidi, A., & Rakrak, S. (2016). A comparative study of target tracking approaches in wireless sensor networks. *Journal of Sensors*, 1-11.
9. Parekh, H.S., Thakore, D.G., & Jaliya, U.K. (2014). A survey on object detection and tracking methods. *International Journal of Innovative Research in Computer and Communication Engineering*, 2(2), 2970-2979.
10. Javed, O., & Shah, M. (2008). *Automated multi-camera surveillance: algorithms and practice*. Springer Science & Business Media, 10.
11. Li, W. (2013). Camera Sensor Activation Scheme for Target Tracking in *Wireless Visual Sensor Networks*. *International Journal of Distributed Sensor Networks*, 9(4), 397537.
12. Lu, X., Zhang, Y., Liu, J., Yuan, F., & Cheng, L. (2019). Mobile target tracking algorithm for wireless camera sensor networks with adjustable monitoring direction of nodes. *International Journal of Communication Systems*, 32(10), 1-14.
13. Peixoto, J.P.J. and Costa, D.G. (2017) 'Wireless visual sensor networks for smart city applications: A relevance-based approach for multiple sinks mobility', *Future Generation Computer Systems*. Elsevier B.V., 76, 51–62
14. Ramya, K., Kumar, K.P., & Rao, V.S. (2012). A survey on target tracking techniques in wireless sensor networks. *International Journal of Computer Science and Engineering Survey*, 3(4), 93-108.
15. Sabokrou, M. (2013). Mobile Target Tracking in non-overlapping *Wireless Visual Sensor Networks* using Neural Networks, In *3rd International Conference on Computer and Knowledge Engineering (ICCKE)*, 1–6.
16. Soro, S., & Heinzelman, W. (2009). A Survey of *Visual Sensor Networks*', *Advances in Multimedia*, 21.
17. Sung, T.W., & Yang, C.S. (2014). A Voronoi-based sensor handover protocol for target tracking in distributed visual sensor networks. *International Journal of Distributed Sensor Networks*, 10(3), 586210.