

A Taxonomy of Visual Surveillance Systems

Naeem Ahmad

Professor Mattias O'Nils

Dr. Najeem Lawal

Department of Electronics Design, in the
Faculty of Science, Technology and Media
Mid Sweden University, SE-851 70 Sundsvall, Sweden

Research Report in Electronics

ISBN 978-91-87103-89-6

stc@miun
Sensible Things that Communicate



Mittuniversitetet
MID SWEDEN UNIVERSITY

This report is a part of the research work which is conducted to develop a generalized model to design surveillance systems. A generalized taxonomy is developed which has organized the core features of surveillance systems at one place and can be used as an important tool for designing surveillance systems. The designers can use this taxonomy to design surveillance systems with reduced effort, time and cost.

A Taxonomy of Visual Surveillance Systems

© Naeem Ahmad, 2013

Research report in electronics

ISBN 978-91-87103-89-6

Electronics Design Department, in the
Faculty of Science, Technology and Media
Mid Sweden University, SE-851 70 Sundsvall
Sweden

Telephone: +46 (0)60 148561

Mittuniversitetet, Sundsvall, Sweden, 2012

ABSTRACT

The increased security risk in society and the availability of low cost sensors and processors has expedited the research in surveillance systems. Visual surveillance systems provide real time monitoring of the environment. Designing an optimized surveillance system for a given application is a challenging task. Moreover, the choice of components for a given surveillance application out of a wide spectrum of available products is not an easy job.

In this report, we formulate a taxonomy to ease the design and classification of surveillance systems by combining their main features. The taxonomy is based on three main models: behavioral model, implementation model, and actuation model. The behavioral model helps to understand the behavior of a surveillance problem. The model is a set of functions such as detection, positioning, identification, tracking, and content handling. The behavioral model can be used to pinpoint the functions which are necessary for a particular situation. The implementation model structures the decisions which are necessary to implement the surveillance functions, recognized by the behavioral model. It is a set of constructs such as sensor type, node connectivity and node fixture. The actuation model is responsible for taking precautionary measures when a surveillance system detects some abnormal situation.

A number of surveillance systems are investigated and analyzed on the basis of developed taxonomy. The taxonomy is general enough to handle a vast range of surveillance systems. It has organized the core features of surveillance systems at one place. It may be considered an important tool when designing surveillance systems. The designers can use this tool to design surveillance systems with reduced effort, cost, and time.

TABLE OF CONTENTS

ABSTRACT	III
TABLE OF CONTENTS	V
ACRONYMS	VII
1 INTRODUCTION	1
1.1 REPORT OUTLINE	3
2 SURVEILLANCE TAXONOMY	5
2.1 BEHAVIORAL MODEL.....	5
2.1.1 <i>Detection</i>	5
2.1.2 <i>Positioning</i>	5
2.1.3 <i>Identification</i>	5
2.1.4 <i>Tracking</i>	6
2.1.5 <i>Content Handling</i>	6
2.2 IMPLEMENTATION MODEL.....	8
2.2.1 <i>Sensor Type</i>	8
2.2.2 <i>Node Connectivity</i>	9
2.2.3 <i>Node Fixture</i>	9
2.3 ACTUATION MODEL.....	9
3 OPTIMIZATION	11
4 ANALYSIS OF EXISTING SYSTEMS	13
4.1 DIRTSS	13
4.2 AAFTFP.....	13
4.3 MSSOIE	14
4.4 ACMSPL.....	15
4.5 AISSOS	15
4.6 WT-BIRD	16
4.7 ID STAT	16
4.8 TADS.....	17
4.9 DTBIRD	17
4.10 VARS.....	18
4.11 ATCVLS	18
5 RESULTS AND DISCUSSIONS	21
6 CONCLUSIONS	25
REFERENCES	27

ACRONYMS

CCTV	Closed circuit television
IR	Infrared
UV	Ultraviolet
PTZ	Pan Tilt Zoom
k HZ	Kilo Hertz
IP	Internet Protocol
RFID	Radio frequency Identification
TADS	Thermal Animal Detection System
VARs	Visual Automated Recording System

1 INTRODUCTION

Growing need of security and availability of cheap surveillance components, such as sensors and processors, has increased the demand of intelligent video surveillance systems. Modern surveillance systems consist of thousands of cameras which are deployed in a surveillance area to collect data. Useful information is extracted from this data to detect, track and recognize objects of interest to understand and analyze their activities. The surveillance systems have a number of useful applications such as home security, traffic control, monitoring patients and elders. These applications have made it possible to monitor homes, hospitals, train stations, parking lots, markets and highways. [1]. Many examples of multi sensor surveillance systems are presented in [14], [15].

The designing of video surveillance systems depends on a number of multidisciplinary fields. The common examples include sensor development, image processing, signal processing, networking, communication, computer vision. The most important disciplines which are vital for any surveillance system are image processing and signal processing [6].

A wide range of products, designed for surveillance systems, are available in market today. Majority of these products do not fulfill the criteria which are required to design a surveillance system for a given application. The selection of components for designing a surveillance system is a complex task [8], [9].

Earlier surveillance systems were simple as compared to the systems available today. Depending on the advancement, the surveillance systems can be divided into three generations [2], [3], [4], [5]. First generation surveillance systems consist of analogue closed circuit television (CCTV) systems. These systems include a large number of cameras deployed in a monitoring area which send videos to the monitoring consoles where human operators watch these videos. Second generation surveillance systems use digital components. They are able to analyze the video data automatically and in real-time. The limitation of these systems is that they are unable to provide support of the algorithms which are required for robust detection and tracking. They are important for behavioral analysis. Third generation surveillance systems cover vast areas and contain a large number of distributed monitoring nodes. In these systems, the video data collected from camera sensors is converted into digital form at monitoring nodes and then transmitted to the destination via communication network.

In intelligent distributed surveillance systems [16], [17], [18], [19], a number of homogeneous and heterogeneous sensors are deployed in a large area, which collaborate with each other and function in a distributed fashion. The processing of collected data is distributed across the network among a number of nodes. This technique is helpful in designing more robust and scalable surveillance systems. Multi cameras systems using active cameras can be used to design more complex

surveillance systems. These systems will ensure the coverage of large area and will also handle the object occlusion problems. However, distributed surveillance systems also introduce new challenges such as designing handoff schemes to track an object in multiple sensors, sensor fusion algorithms and selection of the best view scene collected from multiple sensors [3].

On the basis of control, surveillance systems are classified into three main categories: centralized, semi-distributed and distributed. A centralized surveillance system is controlled from a central server. One example of a centralized surveillance system is PRISMATICA [11]. In PRISMATICA, a centralized computer is dedicated which controls and supervises all functions of the system. In semi-distributed surveillance systems, some of the functions are controlled from a central server while other functions are controlled at the nodes. One example of such a system is ADVISOR [12]. The nodes of ADVISOR contain a central computer system which controls all the node functions. In distributed surveillance systems, there is no central server. All functions are controlled by collaborative sensor nodes. One such surveillance system is presented in [10], which is not using any server. The subsystems of this system are independent nodes which can perform their tasks independently. One node can communicate with any other nodes without the involvement of some central control. The advantage of this approach is more robust and autonomous system. If one or more nodes are failing, then the remaining nodes are able to perform their jobs. It is possible to distribute a task among many nodes, thus, providing flexibility. Moreover, a large number of sensors can be deployed in a wide area [2], [3].

A number of challenges are imposed on video surveillance systems. They should be ubiquitous and autonomous. Modern monitoring systems should be able to operate at remote sites, under varying and strict environmental conditions, such as varying lighting conditions, harsh temperatures, rain, fog, snow, dust, vibration. In case of using a large number of camera nodes, the design of a surveillance system should be scalable [13]. The development, deployment and running cost of these systems should be as low as possible. They must use low power hardware to ensure battery powered operation. The surveillance systems should have automatic fault detection facility. The cameras should be able to calibrate themselves automatically [3].

Optimal deployment of sensors plays fundamental role in designing surveillance systems. Extensive research is being carried out in this direction. The placement of cameras for maximum coverage are presented in Art Gallery Problem (AGP) [38], refined AGP for resolution [39], AGP with minimum guard [40] and Floodlight Illumination Problem (FIP) [47]. Sensor planning is important area in object recognition, tracking and surveillance applications [41], [42], [43], [44]. Multi-camera systems with fixed cameras are usually used for tracking of objects across cameras. The deployment optimization of sensors for such systems is helpful in reducing their implementation cost [37], [45], [46]. Camera placement for occlusion removal is considered in [48]. A design method for selecting the sensors and their optimized placement is discussed in [49]. Placement of cameras for

automated tracking for both coverage and overlap is presented in [50].

1.1 REPORT OUTLINE

The remainder of this report is organized as follows. Chapter 2 discusses in details about surveillance taxonomy, the models which formulate taxonomy and the functions and constructs which form the individual taxonomy models. Chapter 3 briefly discusses the optimization to develop solution for a surveillance problem. Chapter 4 presents a number of systems, their brief description and their analysis on the basis of taxonomy. Chapter 5 describes the results and discussion about the analysis of the systems discussed in chapter 4 and finally chapter 6 presents conclusions.

2 SURVEILLANCE TAXONOMY

This chapter discussed a general taxonomy which is formulated to ease the designing of modern surveillance systems. The taxonomy is shown in Figure 2.1 and can be used to formulate the solution of a surveillance problem. The taxonomy consists of three models: behavioral model, implementation model and actuation model. These models represent different aspects of a given surveillance problem. The behavioral model identifies the key functions which are required to solve a given surveillance problem. The implementation model suggests the physical solution to implement the functions which are identified in the behavioral model. The actuation model implements the actions and activities which are necessary to avoid unfair conditions or to fulfill the objectives of surveillance. The details of these models are discussed below.

2.1 BEHAVIORAL MODEL

The behavioral model is helpful to understand the nature of a surveillance problem. It identifies the key functions which are necessary for the solution of a given surveillance problem. The main functions which form the behavioral model include detection, positioning, identification and tracking. The detail of these functions is described below.

2.1.1 Detection

The intrusion of an object into a surveillance area defines an event. For effective security, it is necessary to detect this event. The detection function senses this intrusion and notifies to the system that some object has entered into the surveillance area. The detection can be of two types: full coverage based detection and partial coverage based detection. In case of partial coverage, the positioning of the nodes should ensure the detection of the objects according to a specific movement model of the objects.

2.1.2 Positioning

This function calculates the position of an intruding object in the surveillance area. The position of an object is described in terms of three parameters - latitude, longitude and altitude. Positioning can be of two types: mapped and triangulation based. The mapped positioning uses a reference with respect to some point or plane surface. The triangulation based positioning uses triangulation technique to find the position of an object. Triangulation uses the overlap of at least two cameras. The overlap can be of two types: fixed overlap or dynamic overlap. Fixed overlap is obtained with static cameras whereas dynamic overlap is obtained with PTZ cameras.

2.1.3 Identification

The identification function recognizes the intruding object. It identifies the shape, size and distance of the object relative the monitoring node. The identification can

either be done by indentifying a class of objects or by identifying a specific individual in a class. That is, if there is a need to detect any human then it corresponds to class based identification. On the other hand, if there is a need to identify a specific human among a group of humans, then it is individual based identification. A statistical model may be needed to identify an object. Temporal information, e.g., wing frequency, can be an important input parameter to the statistical model, which may increase the probability of object identification.

2.1.4 Tracking

The tracking function monitors the movement of an object in the surveillance area. This function depends on the coverage provided by cameras. There can be two types of tracking: full coverage based tracking or partial coverage based tracking. Full coverage based tracking covers each and every point in the surveillance area with the required minimum resolution. Partial coverage based tracking does not cover each and every point or may not provide the minimum resolution required for accurate object recognition. A number of problems can accompany with the coverage. For example, a given area is not fully covered due to improper placement of cameras, minimum required resolution is not assured at every point in the given area, or poor image quality due to camera optics focus. When camera optics is focused at infinity, then the objects near the camera will be imaged with poor quality. Similarly, when camera optics is focused at closer objects, then the objects far from the camera will be imaged with poor quality. Moreover, limited coverage of an area is provided due to the installation of smaller number of cameras than are actually needed. In all these cases, the coverage provided is partial.

Partial coverage can be used to reduce the implementation cost of a network. There are many ways to obtain partial coverage. Two examples of partial coverage are ring based design and islands based coverage, as shown in Figure 2.2. In case of ring based design, a ring of full coverage is formed around the surveillance area, which is used for detection, positioning and identification of objects. To reduce the implementation cost, the area inside the ring is partially covered with the coverage network. In the inner area the coverage network is unable to identify an object and it is only possible to track an object without identification. Thus, object monitoring is divided into two parts. The ring performs the identification task while the inner area performs tracking task. This strategy greatly reduces the implementation cost. In the islands of coverage example, the area contains small islands which are scattered into given area. Full coverage is provided on these islands while limited or zero coverage is provided in the areas other than islands. For such a system, objects are detected by using some kind of statistical model.

2.1.5 Content Handling

This function handles the video data, collected from surveillance area. A common example of this is scenario is CCTV application. Some applications collect video

data and display it on video terminal where people can monitor this video data. Some applications collect video data and also store it somewhere for later use.

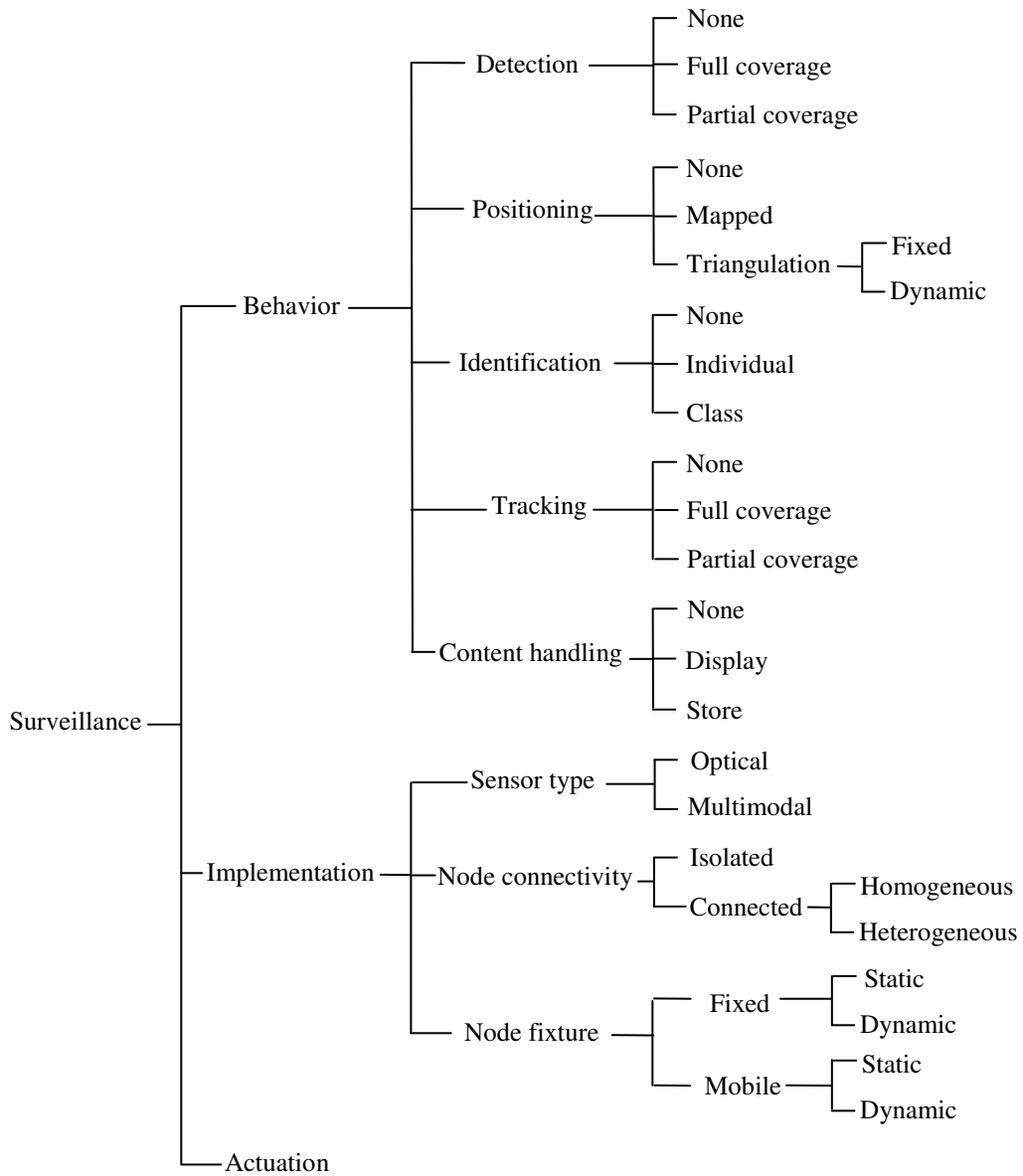


Figure 2.1 Taxonomy model for surveillance systems

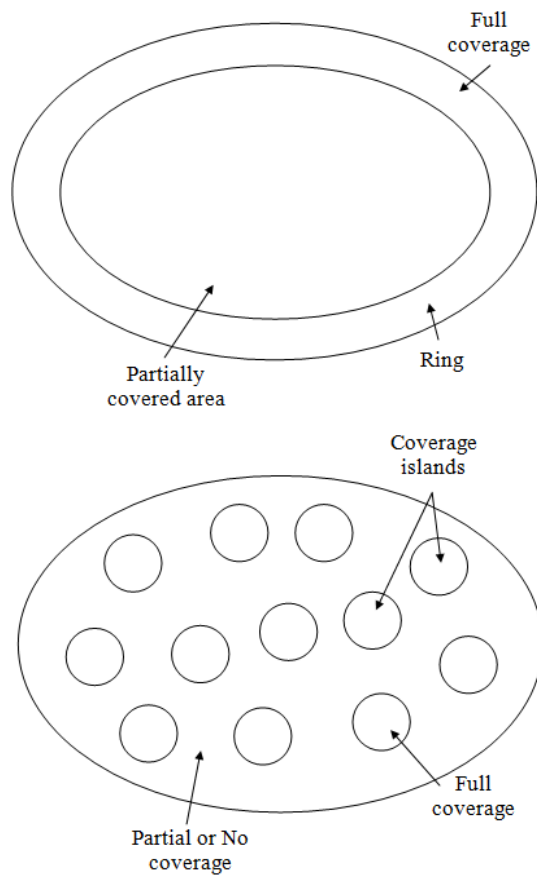


Figure 2.2 Ring design and coverage island models

2.2 IMPLEMENTATION MODEL

The actual surveillance of a given area is implemented with components like camera sensors, accompanying optics and the placement of these components in the surveillance area. A wide range of components is available in the market today and their suitable combinations are used to implement surveillance of a given area. The implementation model is used to structure the decisions which are necessary to implement the surveillance functions, recognized during behavioral analysis of the problem. The implementation model is a set of constructs which include sensor type, node fixture, and node connectivity. These constructs are described below.

2.2.1 Sensor Type

A sensor is a vital part of a monitoring node. It can fall in two main categories: optical and multimodal sensors. An optical sensor senses light which can fall in

three categories like visible light, infrared (IR) light or (UV) ultraviolet light. An optical sensor can be color, monochromatic, infrared, time of flight, laser scanner etc. A camera sensor type selected for an application depends on the nature of an application. Multimodal sensors include sensors which sense physical characteristics other than light such as sound, motion, pressure etc. Some examples of multimodal sensors include acoustic, motion, vibration, seismic, temperature and humidity sensors.

2.2.2 Node Connectivity

The nodes deployed for a surveillance application can function as isolated nodes or connected nodes. The isolated nodes have no collaboration with neighboring nodes and are unable to form a network. The connected nodes collaborate and coordinate with neighboring nodes by forming a network. A network can be of two types, homogeneous or heterogeneous. In homogeneous network, all monitoring nodes use same type of sensors and accompanying optics. In heterogeneous network, the monitoring nodes use different type of sensors or accompanying optics. Some nodes can use optical sensors while others can use non optical sensors. Even in case of using only optical sensors, the node can form heterogeneous networks due to using different type of image sensors. If the nodes are using same type of image sensors, the network can be heterogeneous due to different type of optics used with these image sensors.

2.2.3 Node Fixture

Node fixture relates to the possibility of change of position of a monitoring node. Two major node fixtures which can be considered include fixed and mobile fixtures. A fixed node is unable to change its position once deployed. In contrast, a mobile node is moveable and is able to change its position. It is fitted on some kind of movable platform such as a robot. A mobile node is most suitable for surveillance of indoors of large and complex buildings. For each fixture type, there are further two categories are available which include static and dynamic nodes. These categories define the activity of a node in terms of change of field of view after installation of nodes. A static node has fixed field of view. A dynamic node can change its field of view horizontally as well as vertically with respect to space after its deployment. It can also change its zoom property. Such a node is implemented with PTZ cameras.

2.3 ACTUATION MODEL

The actuation model relates to implementing precautionary measures when a surveillance system detects some abnormal situation. For example in case of DT Bird system [33], the dissuasion module emits warning and dissuasion signals to the birds which are flying in the high collision risk area. The other example is the

generation of stop signal to a wind turbine from the stop control module of DT Bird system in case of collision risk.

3 OPTIMIZATION

The behavioral model is important to understand a surveillance problem and is helpful to identify the functions which are necessary to solve a given problem. The implementation model proposes the physical solution to implement the functions identified by the behavioral model. The selection of suitable camera sensors, accompanying optics and the placement of nodes in a given area are not simple undertakings. The components like higher resolution image sensors, optical lenses with longer focal lengths, improve the resolution aspect but at the same time increase the implementation cost of the surveillance. Optimization techniques are necessary to select balanced combinations of components and their deployment in the surveillance area. Before applying the optimization techniques, it is necessary to formulate the optimization problem of the required surveillance assignment. Camera sensor types and the related optics define the variables of the optimization problem while the behavioral model constructs define the constraints for the optimization problem. The objective function forms the cost function with camera sensor types and the accompanying optics variables. The objective of the optimization problem is to minimize the cost required to implement a surveillance solution. The optimization solution helps to select the optimized combination of camera sensors and related optics. A general cost function is given below.

$$\text{cost} = f(\text{camera sensor, optics})$$

Additional variables can be added to the optimization model for more sophisticated solutions. For example, a variable which can be input to the cost function is the number of camera configurations. One camera configuration has one implementation cost while for two camera configurations there is new cost. Then the above cost function will change to the following form.

$$\text{cost} = f(\text{camera sensor, optics, number of camera configurations})$$

Another variable can be considered which captures the surface topography of the surveillance area. In the beginning phase, flat surface topography can be assumed. However, real topography of the surface can be added to the cost function for more complex scenarios and for better solutions.

4 ANALYSIS OF EXISTING SYSTEMS

After description of surveillance taxonomy, a number of surveillance systems are studied and are analyzed on the basis of taxonomy. In each case, a brief review of the system is presented first and then its analysis is performed on the basis of developed taxonomy. The studied systems and their analysis are given below.

4.1 DIRTSS

The DIRTSS [20] implements a real time video surveillance system. The non optical sensors used by the system include microphone, tone detector while the optical sensors include two cameras. A group of cooperating sensors detect and track mobile objects and report their positions to the sink node. The sink node uses two IP cameras DCS-5300G to record these events. Each camera has a pan and tilt function that can cover 270° angle side-to-side and a 90° angle up and down. The system is using 16 sensor nodes in addition to two IP cameras. A 4 kHz sound signal is used to trigger an event which is generated by a hand-held device. A microphone is used to detect the sound. The sensor board of each mote has a microphone and a tone detector that can filter out the 4 kHz sound signal and can generate 1-bit digital output.

The analysis of the DIRTSS on the basis of taxonomy shows that the system is implementing distributed detection function. The positioning of objects is provided by non optical sensors and the positioning is mapped based. The system can identify more than one object and the identification is of class based type. Tracking of objects is implemented with two active cameras which are providing partial coverage. The application is not providing content handling function. The system is using multimodal sensors. The optical sensors include visual light pan tilt cameras. The non optical sensors include microphone and tone detectors. Node connectivity is of connected type. The nodes are connected with heterogeneous network which is connecting optical and non optical sensors. Node fixture is of fixed type implemented with dynamic nodes which are pan tilt cameras. The system is providing actuation function in the form of 1-bit digital output.

4.2 AAFTFP

The AAFTFP application [21] implements a system to track multiple football players with multiple cameras. The system provides positions of players and ball during a football match. The system is composed of eight cameras, which are statically positioned around a stadium. The cameras have overlapping fields-of-view. The system is using two processing stages. The first stage processes the data received from each individual camera while the second stage processes the data received from multiple cameras. Single-view processing includes change detection

while multi-view processing uses Kalman trackers to model player position and velocity. Data from each camera is input to a central tracking process, to update the state estimates of the players. The system can identify the colors of five different types of uniforms of the players (two types of uniforms for two teams, two types for goal-keepers, and one type for three referees). The central tracking process outputs the positions of 25 players per time step. The tracker is able to identify a player of a team but is unable to identify an individual player in a team. The ball tracking methods are not considered in the paper.

The analysis of AAFTFP application on the basis of taxonomy shows that the system is providing detection of distributed type. To find the position of players, the system is using triangulation which is implemented with fixed overlap of coverages. The identification of players is class based. Tracking of players is provided on partial coverage basis. The application is not providing content handling function. The system is using optical sensors in the form of eight visible light cameras. Node fixture is fixed type implemented with fixed nodes. The node connectivity of eight camera nodes is of connected type. The camera nodes are similar which are connected with homogeneous network. The system does not provide any actuation function.

4.3 MSSOIE

MSSOIE [22] is an autonomous mobile robotic system which operates in an indoor environment with key points marked by RFID tags. These tags serve two purposes. They provide information about the surrounding region as well as instruct the robot to perform certain tasks. The robot is equipped with a camera, a laser scanner, encoders and an RFID device. The camera is used to image the current scene, which is compared with a stored image. The robustness of the visual detection is improved by considering both geometrical and color information. The laser sensor matches the local reference and current range data to look for scene variations. The robot detects points which are extracted both from current and the stored images and are matched to locate the differences in the scenes. The robot builds a map of the environment to identify areas of interest which are marked by RFID tags, to monitor the target zones to detect unexpected changes such as object addition or removal. If no variation is detected, the system generates actuation command for robot motor to navigate to the next key point. If some type of variation is detected, the system generates actuation command to trigger an alarm.

The analysis of MSSOIE on the basis of taxonomy shows that the system is providing area based detection. The positioning function implemented is mapped with RFID tags. The system can identify many key points and identification is class based. The system is not providing tracking and content handling functions. The system is using multimodal sensors. The optical sensors are visual light cameras

and laser scanners. The non optical devices include encoders and RFID tags. Node connectivity is connected type. Node fixture is both fixed and mobile type. The system is providing actuation in the form of command to the robot motor or sound an alarm.

4.4 ACMSPL

ACMSPL [23] is a cooperative system of static and active cameras for video surveillance of parking lots. The system is able to track multiple targets in real-time. Object positions are computed by using static cameras which are used to monitor a wide parking area and to track multiple moving objects. Data from different cameras are fused to enhance system performance. Out of static cameras, some are using b/w image sensor and others are using color sensor. The information about object positions is provided by static cameras, which is later used by active cameras to track the objects. The active cameras record close-up and high resolution video of the suspicious events. Once a target is selected by static cameras, the active camera is activated by computing the rotation angles α and β for pan and tilt, respectively, such that the target is located at the center of the field of view of the active camera. In this scenario, the active camera is able to identify and track a target.

The analysis of ACMSPL on the basis of taxonomy shows that the system is implementing distributed type detection. The object position measurement is triangulation based with both fixed and dynamic overlap. The static cameras make a fixed overlap while static-active and active-active cameras make dynamic overlap. The identification is class based. The tracking function is of partial coverage type, which is performed with active cameras. The system is not implementing content handling function. The system is using optical sensors which are visible light, low resolution image sensors. The node fixture is fixed type which is implemented with both static and dynamic nodes. The node connectivity is connected type. Different types of camera nodes are connected together in the form of heterogeneous network. The actuation is provided by the system to activate pan and tilt functions of active cameras.

4.5 AISSOS

AISSOS [24] is an integrated system for detection, tracking and identification of people in wide outdoor environments such as parking lot. The system uses two cameras, static and dynamic, connected to two processing nodes. The static camera has a wide view and thus covers a wide area. The active PTZ camera is used to image a closer view of people to get higher resolution images of their faces. Static camera monitors the environment to detect suspicious events. The detection of an anomalous event raises an alarm which is communicated to the active node along

with the information about the position of the area of interest. The PTZ camera focus the people involved, with the objective to recognize their identity by using face detection and recognition techniques. The face recognition module determines the identity of the detected person's face by comparing it with face images of known identity stored in a database.

The analysis of AISSOS system on the basis of developed taxonomy reveals that the system is using area based detection. It is not implementing positioning function. Identification function is individual based. Tracking is partial coverage based. The system is not implementing content handling function. The sensors used are of optical type, which include visible light, lower and higher resolution image sensors. Node fixture is fixed type which is implemented with both static and dynamic nodes. Node connectivity is connected type. Two different types of cameras are connected in the form of heterogeneous network. The system is implementing actuation functions in the form of pan, tilt and zoom commands to active camera.

4.6 WT-BIRD

WT-Bird is a system for detecting and registering bird collisions with wind turbines. This system uses a combination of accelerometers and microphones to detect collisions, and two active infrared video cameras to record video [25], [26], [27], [28], [29]. The cameras, along with illumination, are mounted on the lower part of the turbine tower to capture images of the area swept by the rotors. The sensors, located within the rotors and turbine towers, detect collisions. In case of occurrence of an event, both sound and images are captured and saved. A notification about the event is sent to the user.

After analyzing the WT-Bird system on the basis of taxonomy it is found that it is implementing collision detection function which is of distributed type. The system is not implementing positioning function. The identification function implemented is class based. Tracking and content handling functions are not implemented in this system. The system is using multimodal sensors. The optical sensors are infrared image sensors. Non optical sensors are microphone and accelerometers. Node fixture is fixed type, implemented with dynamic nodes. Node connectivity is connected type and nodes are forming heterogeneous network. Actuation is provided in the form of message transmission to the user in case of collision detection and actuation commands to active camera.

4.7 ID STAT

ID Stat system is designed to detect birds across wind turbines [25], [26], [30]. To detect collisions, directional microphones are placed within the hub of the turbines

at the base of each rotor. The microphones detect potential collisions and the accompanying software filter out the background noise as well as noise from rain. Once a collision is detected, the relevant information such as date, time, turbine ID, sensor ID are stored using data loggers and a message can be sent to the user via GSM network. This system can only detect collisions and is unable to record visual information.

After analyzing the ID-Stat system on the basis of taxonomy it is found that it is implementing collision detection function which is of distributed type. The functions like positioning, identification, tracking and content handling are not implemented in this system. The system is using non optical sensors which are directional microphones. Node fixture mobile type implemented with static nodes. Node connectivity is isolated type. Since the nodes are working in isolated fashion so they are not forming a network. Actuation is provided in the form of message transmission to the user in case of collision detection.

4.8 TADS

Thermal Animal Detection System (TADS) [31] is developed to identify bird collisions. The system uses a combination of infrared video cameras which are mounted on the base of a turbine tower. Coverage of the entire rotor area is obtained with three cameras while the coverage from all directions is implemented with six cameras [32]. TADS includes the possibility of using multiple cameras to give effectively a larger field of view. The sensor nodes are able to identify smaller species.

After analyzing the TADS system on the basis of taxonomy it is found that it is implementing distributed type collision detection function. The functions like positioning, tracking and content handling are not implemented in this system. The identification function is implemented which is class based. The system is using optical sensors which are infrared cameras. Node fixture is fixed type implemented with static nodes. Node connectivity is connected type. The nodes are using similar type of cameras which are connected together forming homogeneous network. The system is not implementing actuation function.

4.9 DTBIRD

DTBird [33] system is able to detect flying birds across a wind turbine and is using two visual light cameras, each covering 180° around the turbine. DTBird detects flying birds in real-time and can respond by carrying out pre-programmed actions if birds are detected within a pre-defined risk-zone. The dissuasion module scares birds in close proximity of the turbines. The stop control module can stop wind turbine when a bird flies in a pre-defined risk area. The collision control module

detects collisions and records data from events when a flying bird is detected close to a pre-defined area, such as the area around the rotors of a wind turbine. Depending on the position of the camera, more than one wind turbine may be monitored simultaneously.

After analyzing DTBird system on the basis of taxonomy it is found that it is implementing distributed type collision detection function. The functions like positioning, tracking and content handling are not implemented in this system. The identification function is implemented which is class based. The system is using optical sensors which are visual light cameras. Node fixture fixed type implemented with static nodes. Node connectivity is isolated type. Since nodes are working in isolated mode, so no network is formed. The system provides actuation function in the form of dissuasion signals and wind turbine stop signal.

4.10 VARS

Visual Automated Recording System (VARS) uses two active infrared video cameras along with infrared lamps for detecting flying birds. The cameras are mounted on the nacelle of the turbines and have a relatively narrow field of view. The VARS system has been used to assess a number of flying birds close to the turbines [34], [35]. The sensor nodes are attached to the turbine. One camera is attached to the nacelle and covers an area just behind the rotors (30° FoV parallel to the rotor-swept area). Second camera is positioned at the base of the turbine and faces upward towards the rotors. The motion-controlled cameras are always on and a sequence of images is recorded once the trigger threshold is reached.

After analyzing the VARS system on the basis of taxonomy it is found that it is implementing distributed type collision detection function. The functions like positioning, tracking and content handling are not implemented in this system. The identification function implemented is class based. The system is using optical sensors which are infrared cameras. Node fixture fixed type implemented with dynamic cameras. Node connectivity is isolated type. Since nodes are working in isolated mode, so no network is formed. No actuation function is implemented in this system.

4.11 ATCVLS

A three camera video lobby surveillance system is presented in [36]. The system includes three cameras which are installed in a lobby to monitor people which are leaving and entering the lobby and surveillance of the reception area. One camera has large field of view than the other two cameras and is able to cover more area. The cameras are connected to video terminals and video recording equipment,

placed in the security room where security people can monitor the video output from cameras and they are also able to record video when needed.

After analyzing the ATCVLS system on the basis of taxonomy it is found that the functions like detection, positioning, identification and tracking are not implemented in this system. The system is implementing content handling function in the form of display and storage facility. The system is using optical sensors which are visible light camera sensors. Node fixture is fixed type implemented with static nodes. Node connectivity is connected type as cameras are working in collaboration. The cameras are forming a heterogeneous network as the camera types are different. The system is not implementing any actuation function.

5 RESULTS AND DISCUSSIONS

A number of surveillance systems are investigated and analyzed on the basis of developed taxonomy. The analysis results are summarized in Table 5.1. Starting from behavioral model, the analysis of each function in this model is presented one at a time. Considering detection function, the table shows that almost all the surveillance systems provide detection function which is of distributed type while only a few systems provide area based detection. Considering positioning function, most of the surveillance systems do not provide positioning function. Out of the systems which provide positioning function, some systems provided mapped positioning while others provide triangulation based positioning. This type of positioning uses overlap of camera coverage to position an object. The coverage overlap can be of two types: fixed overlap and dynamic overlap. Fixed overlap is implemented with static nodes while dynamic overlap is implemented with active nodes. Most systems use either fixed overlap or dynamic overlap. Only limited number of systems uses both fixed and dynamic overlap to implement the positioning function. Considering identification function, the table shows that most of the systems are providing identification function. Out of the systems which provide identification function, almost all the systems provide class based identification while only a few systems provide individual based identification. Considering tracking function, the table shows that most systems are not providing tracking function. Out of the systems which are providing tracking function, all systems are providing partial coverage based tracking. Considering content handling function, most of the surveillance systems are not providing content handling function.

After presenting the analysis of the behavioral model, the analysis of the implementation model is presented here. Considering sensor type, TABLE 5.1 shows that most of the systems are using optical sensors while fewer systems are using multimodal sensors. Also, there are some systems which are using both optical as well as multimodal systems. Considering node fixture, most of the systems are using fixed node fixture with static nodes while some systems are also using dynamic nodes. Very few systems are using mobile node fixture with static nodes. No system is found to have mobile node fixture with dynamic nodes. A single system is found to have both fixed as well as mobile node fixtures with static nodes. Considering node connectivity, very few systems are using isolated nodes which are not collaborating with their neighboring nodes. The majority of the systems are using collaborative nodes which are collaborating with their neighboring nodes by forming a network with them. Most of the system are using heterogeneous network to provide communication among the nodes while fewer systems are using homogeneous network.

Considering actuation model, most of the systems are providing actuation functions to handle unfair situation or to implement the desired action in response to an event in the surveillance area.

In summary, most surveillance systems are providing distributed detection instead of area based detection. Very few systems are providing positioning function. Out of the systems, which are providing positioning function, most systems are using triangulation based positioning function with fixed overlap. There is a potential to implement triangulation technique with dynamic overlap by using dynamic nodes. Moreover, there is space available to consider combination of both static and active nodes. Most of the systems are providing class based identification function. Some systems are implementing partial coverage based tracking function which is good trend in terms of reducing the implementation cost, because it uses lesser nodes for tracking. Most surveillance systems are not using content handling function. Many surveillance systems are depending more on optical sensors than multimodal sensors. There is a potential to design surveillance systems which exploit the power of multimodal sensors in addition to optical sensors. Most systems are using fixed node fixture with static nodes. There is space available to implement fixed node fixture with dynamic nodes as well. There is a potential to explore mobile node fixture with static as well as dynamic nodes. Combination of both fixed and mobile node fixtures with both fixed and dynamic nodes should also be explored. Most systems are using collaborative nodes, which is a good trend for very effective surveillance of an area. More surveillance systems are using heterogeneous networks and are providing actuation function.

TABLE 5.1. Summary of analysis results

No.	Applications	Behavioral Model										Implementation Model						Actuation Model											
		Detection			Positioning			Identification			Tracking			Content handling			Sensor Type		Node Fixture		Node connectivity								
		None	Area	Distributed	None	Mapped	Fixed overlap	Dynamic overlap	None	Individual	Class	None	Full coverage	Partial coverage	None	Display	Store	Optical	Non Optical	Static	Dynamic	Fixed	Static	Dynamic	Isolated	Heterogeneous	Connected	Heterogeneous	
1	DIRTSS			✓		✓				✓		✓						✓	✓	✓		✓				✓	✓	✓	
2	AAFTFP			✓			✓			✓		✓						✓	✓	✓						✓			
3	MSSOIE		✓			✓				✓		✓						✓	✓	✓						✓	✓	✓	
4	ACMSPL			✓			✓			✓		✓						✓	✓	✓							✓	✓	✓
5	AISSOS		✓															✓	✓	✓							✓	✓	✓
6	WT-Bird			✓						✓		✓						✓	✓	✓							✓	✓	✓
7	ID-Stat			✓						✓		✓						✓	✓	✓					✓				✓
8	TADS			✓						✓		✓						✓	✓	✓							✓		✓
9	DTBird			✓						✓		✓						✓	✓	✓							✓		✓
10	VARS			✓						✓		✓						✓	✓	✓							✓		✓
11	ATCVLS	✓					✓					✓						✓	✓	✓							✓		✓

6 CONCLUSIONS

In this report, we formulate a taxonomy to ease the design of surveillance systems. The taxonomy contains three models: behavioral model, implementation model and actuation model. The behavioral model consists of functions such as detection, positioning, identification, and tracking. This model can be used to exactly pinpoint which functions are necessary for a particular situation. The implementation model handles the practical details to implement the functions identified in the behavior model. It includes the constructs such as sensor type, node connectivity and node fixture. The actuation model implements specific actions in response to occurrence of a particular situation in the surveillance area. After formulation of taxonomy, a number of surveillance systems are investigated and analyzed on the basis of this taxonomy. The analysis shows that the taxonomy is general enough to handle a vast range of surveillance systems. The taxonomy has organized the core features of surveillance systems at one place and can serve as an important tool for designing surveillance systems. The designers can use it to design surveillance systems with reduced effort, cost and effort.

Analysis results of surveillance systems based on the developed taxonomy show that most surveillance systems provide detection and positioning functions but a limited number of systems provide identification and tracking functions. The detection function is of distributed type while the identification function is of class based. Most surveillance systems are not implementing content handling function. Also, most systems are using optical sensors for their implementation, collaborative nodes of static type, heterogeneous networks for nodes connectivity, and provide actuation functions to handle undesirable situations. The analysis shows that the surveillance systems are not using complete taxonomy map. It means that more design options can be explored by using the unused constructs of taxonomy which can result in more reliable and robust surveillance systems. For example, more versatile surveillance systems can be designed by introducing multimodal systems in addition to optical sensors. Another option is to explore the possibility of mobile node fixtures and dynamic nodes for designing surveillance systems. A combination of fixed and dynamic nodes can result in more versatile surveillance applications. Partial coverage can be exploited to implement tracking function which can result in substantial reduction in the implementation cost.

REFERENCES

- [1] X. Wang, "Intelligent Multi-Camera Video Surveillance: A Review", *Pattern Recognition Letters*, vol. 34, no. 1, pp. 3-19, 2013.
- [2] M. Valera, S.A. Velastin, "Real-time architecture for a large distributed surveillance system", *Proceedings of IEE Intelligent Distributed Surveillance Systems*, pp. 41-45, 2004.
- [3] T.D. Rätty, "Survey on Contemporary Remote Surveillance Systems for Public Safety", *IEEE Transactions on Systems, Man, and Cybernetics, Part C*, vol. 40, no. 5, pp. 493-515, 2010.
- [4] M. Valera and S. A. Velastin, "Intelligent distributed surveillance systems: A review," *IEE Proceedings on Vision, Image and Signal Processing*, vol. 152, no. 2, pp. 192-204, 2005.
- [5] M. Bramberger, A. Doblander, A. Maier, B. Rinner, and H. Schwabach, "Distributed embedded smart cameras for surveillance applications," *IEEE Computer*, vol. 39, no. 2, pp. 68-75, 2006.
- [6] K.N. Plataniotis and C.S. Regazzoni, "Visual-Centric Surveillance Networks and Services", *IEEE Signal Processing Magazine*, special issue on video and signal processing for surveillance networks and services, vol. 22, no. 2, 2005.
- [7] C. Micheloni, G.L. Foresti and L. Snidaro, "A Network of Co-operative Cameras for Visual Surveillance", *Proceedings of IEE Vision, Image and Signal Processing*, vol. 152, no. 2, pp. 205-212, 2005.
- [8] M.H. Sedky, M. Moniri and C.C. Chibelushi, "Classification of Smart Video Surveillance Systems for Commercial Applications", *Proceedings of IEEE International Conference on Advanced Video and Signal based Surveillance*, pp. 638-643, 2005.
- [9] C. H. Heartwell, A. J. Lipton, "Critical Asset Protection, Perimeter Monitoring, and Threat Detection Using Automated Video Surveillance - A Technology Overview with Case Studies," *International Carnahan Conference on Security Technology*, pp. 87, 2002.
- [10] M. Christensen and R. Alblas, "V2-design issues in distributed video surveillance systems", *Technical Report, Department of Computer Science, Aalborg University, Aalborg, Denmark*, 2000.
- [11] S. A. Velastin, B. A. Boghossian, B. P. I. Lo, J. Sun, and M. A. Vicencio-Silva, "PRISMATICA: Toward ambient intelligence in public transport environments", *IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and Humans*, vol. 35, no. 1, pp. 164-182, 2005.

- [12] Annotated Digital Video for Intelligent Surveillance and Optimised Retrieval (ADVISOR) <http://www-sop.inria.fr/orion/ADVISOR/index.html>.
- [13] P. Remagnino, S. S. Velastin, G.L. Foresti, M. Trivedi, "Novel concepts and challenges for the next generation of video surveillance systems", *Machine Vision and Applications*, vol. 18, no. 3-4, pp. 135-137, 2007.
- [14] C. Regazzoni, V. Ramesh, and G. Foresti, "Special issue on video communications, processing, and understanding for third generation surveillance systems", *Proceedings of the IEEE*, vol. 89, no. 10, 2001.
- [15] S. Haritaoglu, D. Harwood, and L. Davis "W4: Real-time surveillance of people and their activities", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 809-830, 2000.
- [16] I. Mikic, S. Santini, and R. Jain, "Tracking objects in 3d using multiple camera views", *Asian Conference on Computer Vision*, 2000.
- [17] R. Collins, A. Lipton, H. Fujiyoshi, and T. Kanade "Algorithms for cooperative multisensor surveillance", *Proceedings of the IEEE*, vol. 89, no. 10, pp. 1456-1477, 2001.
- [18] S. Dockstader, and A. Tekalp "Multiple camera tracking of interacting and occluded human motion", *Proceedings of the IEEE*, vol. 89, no. 10, pp. 1441-1455, 2001.
- [19] T. Matsuyama, and N. Ukita "Real-time multitarget tracking by a cooperative distributed vision system", *Proceedings of the IEEE*, vol. 90, no. 7, pp. 1136-1150, 2002.
- [20] W.T. Chen, P.Y. Chen, W.S. Lee, and C.F. Huang, "Design and implementation of a real time video surveillance system with wireless sensor networks," *Proceedings of IEEE, Vehicular Technology Conference*, pp. 218-222, 2008.
- [21] M. Xu, L. Lowey, and J. Orwell, "Architecture and algorithms for tracking football players with multiple cameras," *Proceedings of IEE Workshop on Intelligent Distributed Surveillance Systems*, London, pp. 51-56, 2004.
- [22] D. Di Paola, D. Naso, A. Milella, G. Cicirelli and A. Distante, "Multi-Sensor Surveillance of Indoor Environments by an Autonomous Mobile Robot," *International Conference on Mechatronics and Machine Vision in Practice*, pp. 23-28, 2008.
- [23] C. Micheloni, G.L. Foresti, and L. Snidaro, "A co-operative multicamera system for video-surveillance of parking lots," *IEE Symposium on Intelligent Distributed Surveillance Systems*, pp. 21-24, 2003.
- [24] C. Micheloni, E. Salvador, F. Bigaran, and G.L. Foresti, "An Integrated Surveillance System for Outdoor Security," *Proceedings of the IEEE*

Conference on Advanced Video and Signal Based Surveillance, pp. 480-485, 2005.

- [25] M. P. Collier, S. Dirksen, K. L. Krijgsveld, "A review of methods to monitor collisions or micro-avoidance of birds with offshore wind turbines, Part 1: Strategic Ornithological Support Services Project SOSS-03A", Bureau Waardenburg bv, 2011.
- [26] M. P. Collier, S. Dirksen, K. L. Krijgsveld, "A review of methods to monitor collisions or micro-avoidance of birds with offshore wind turbines, Part 2: Feasibility study of systems to monitor collisions", Bureau Waardenburg bv, 2012.
- [27] E.J. Wiggelinkhuizen, L.W.M.M. Rademakers, S.A.M. Barhorst and H.J. den Boon, "Bird collision monitoring system for multi-megawatt wind turbines WTbird: Prototype development and testing", Report ECN-E-06-027. Energy research Center of the Netherlands, 2006.
- [28] E.J. Wiggelinkhuizen, L.W.M.M. Rademakers, S.A.M. Barhorst, H.J. den Boon, S. Dirksen and H. Schekkerman, "WT-Bird: Bird collision recording for offshore wind farms", Report ECN-RX-06-060. Energy research Center of the Netherlands, 2006.
- [29] E.J. Wiggelinkhuizen, L.W.M.M. Rademakers, S.A.M. Barhorst, H.J. den Boon, S. Dirksen and H. Schekkerman, "WT-Bird Bird Collision Recording for Offshore Wind Farms", Report ECN-RX--04-121. Energy research Center of the Netherlands, 2004.
- [30] B. Delprat, "ID Stat: innovative technology for assessing wildlife collisions with wind turbines", Oral presentation at Conference on Wind energy and Wildlife impacts, Trondheim, Norway, 2011.
- [31] M. Desholm, "Thermal Animal Detection System (TADS). Development of a method for estimating collision frequency of migrating birds at offshore wind turbines", National Environmental Research Institute (NERI) Denmark, Technical Report No 440, 2003.
- [32] M. Desholm, "Preliminary investigations of bird-turbine collisions at Nysted offshore wind farm and final quality control of Thermal Animal Detection System (TADS)", National Environmental Research Institute (NERI) Denmark and Ministry of Environment Denmark, 2005.
- [33] DTBird, bird detection and dissuasion, <http://www.dtbird.com>, 2011.
- [34] T. Coppack, C. Kulemeyer, A. Schulz, T. Steuri and F. Liechti, "Automated in situ monitoring of migratory birds at Germany's first offshore wind farm", Oral presentation at Conference on Wind energy and Wildlife impacts, Trondheim, Norway, 2011.

- [35] T. Coppack, C. Kulemeyer and A. Schulz, "Monitoring migratory birds through fixed pencil beam radar and infrared videography at offshore wind farm Alpha Ventus", 2011.
- [36] H. Kruegle, "Analog Three-Camera Video Lobby Surveillance System", CCTV Surveillance: Analog and Digital Video Practices and Technology, pp. 516-518, 2007.
- [37] Bodor, R., Drenner, A., Schrater, P. and Papanikolopoulos, N., "Optimal camera placement for automated surveillance tasks," J. Intell. Robotics Syst., 50(3), 57-295 (2007).
- [38] O'Rourke, J., "Art Gallery Theorems and Algorithms," Oxford University Press, New York, (1987).
- [39] Fleishman, S., Cohen-Or, D., Lischinski, D., "Automatic camera placement for image-based modeling," Proc. of Pacific Graphics, 12-20 (1999).
- [40] Isler, V., Kannan, S., Daniilidis, K., "Vc-dimension of exterior visibility," IEEE Transactions on pattern analysis and machine intelligence, 26(5), 667-671 (2004).
- [41] Tarabanis, K.A., Allen, P., Tsai, R.Y. "A survey of sensor planning in computer vision," IEEE Transactions on robotic automation, 11(1), 86-104(1995).
- [42] Roy, S., Chaudhury, S., Banerjee, S., "Active recognition through next view planning: a survey," Pattern Recognition., 429-446(2004).
- [43] Scott, W., Roth, G., Rivest, J.-F., "View planning for automated three-dimensional object reconstruction and inspection," Computer Surveillance, 35(1), 64-96 (2003).
- [44] Sharma, R., Hutchinson, S. "Motion perceptibility and its application to active vision-based servo control," IEEE Transaction on Robotic Automation 13(4), 607-617 (1997).
- [45] Stauffer, C., Tieu, K., "Automated multi-camera planar tracking correspondence modeling," Proc. of IEEE International Conference on Computer Vision and Pattern Recognition, vol. 1, (2003).
- [46] Ukita, N., Matsuyama, T, "Incremental observable-area modeling for cooperative tracking," Proc. of International Conference on Pattern Recognition, (2000).
- [47] Bose, P., Guibas, L., Lubiw, A., Overmars, M., D. Souvaine and Urrutia, J., "The floodlight problem," International Journal of Computational Geometry and Applications, 153-163 (1997).

- [48] Murray, A., Kim, K., Davis, J., Machiraju, R., Parent, R.. "Coverage optimization to support security monitoring," *Computers, environment and urban systems*, 31(2), 133-147 (2007).
- [49] Ram, S., Ramakrishnan, K., Atrey, P., Singh, V., Kankanhalli, M., "A design methodology for selection and placement of sensors in multimedia systems," *Proc. of International workshop on video surveillance and sensor networks*, (2006).
- [50] Yao, Y., Chen, C., Abidi, B., Page, D., Abidi, B., Abidi, M., "Sensor planning for automated and persistent object tracking with multiple cameras," *Proc. of IEEE International conference on computer vision and pattern recognition*, 1-8 (2008).