Quality-of-Information Aware Sensing Node Characterisation for Optimised Energy Consumption in Visual Sensor Networks

Anas Amjad

A thesis submitted in partial fulfilment of the requirement of Staffordshire University for the degree of Doctor of Philosophy

September 2017

Abstract

Energy consumption is one of the primary concerns in a resource constrained visual sensor network (VSN) with wireless transceiving capability. The existing VSN design solutions under particular resource constrained scenarios are application-specific, whereas the degree of sensitivity of the resource constraints varies from one application to another. This limits the implementation of the existing energy efficient solutions within a VSN node, which may be considered to be a part of a heterogeneous network. This thesis aims to resolve the energy consumption issues faced within VSNs because of their resource constrained nature by proposing energy efficient solutions for sensing nodes characterisation.

The heterogeneity of image capture and processing within a VSN can be adaptively reflected with a dynamic field-of-view (FoV) realisation. This is expected to allow the implementation of a generalised energy efficient solution that will adapt with the heterogeneity of the network. In this thesis, a FoV characterisation framework is proposed, which can assist design engineers during the pre-deployment phase in developing energy efficient VSNs. The proposed FoV characterisation framework provides efficient solutions for: 1) selecting suitable sensing range; 2) maximising spatial coverage; 3) minimising the number of required nodes; and 4) adaptive task classification. The task classification scheme proposed in this thesis exploits heterogeneity of the network and leads to an optimal distribution of tasks between visual sensing nodes. Soft decision criteria is exploited, and it is observed that for a given detection reliability, the proposed FoV characterisation framework provides energy efficient solutions which can be implemented within heterogeneous networks.

In the post-deployment phase, the energy efficiency of a VSN for a given level of reliability can be enhanced by reconfiguring its nodes dynamically to achieve optimal configurations. Considering the dynamic realisation of quality-of-information (QoI), a strategy is devised for selecting suitable configurations of visual sensing

nodes to reduce redundant visual content prior to transmission without sacrificing the expected information retrieval reliability. By incorporating QoI awareness using peak signal-to-noise ratio-based representative metric, the distributed nature of the proposed self-reconfiguration scheme accelerates the decision making process.

This thesis also proposes a unified framework for node classification and dynamic self-reconfiguration in VSNs. For a given application, the unified framework provides a feasible solution to classify and reconfigure visual sensing nodes based on their FoV by exploiting the heterogeneity of targeted QoI within the sensing region. From the results, it is observed that for the second degree of heterogeneity in targeted QoI, the unified framework outperforms its existing counterparts and results in up to 72% energy savings with as low as 94% reliability. Within the context of resource constrained VSNs, the substantial energy savings achieved by the proposed unified framework can lead to network lifetime enhancement. Moreover, the reliability analysis demonstrates suitability of the unified framework for applications that need a desired level of QoI.

Acknowledgements

Foremost, I would like to thank Almighty Allah for all the blessings throughout my life and for providing me wisdom.

I am most sincerely and heartily grateful to my principal supervisor, Dr. Alison Griffiths, for her support and guidance during the course of this research. She was always there to offer advice despite her extremely busy schedules. I also thank her for her confidence in my abilities. I owe a debt of gratitude for her valuable input and suggestions to enhance the quality of this thesis.

I would like to express my deepest appreciation and admiration to my second supervisor, and my mentor, Professor Mohammad Patwary, for sharing his knowledge, wisdom, experience, and enthusiasm. I am grateful to him for his immense contribution towards my academic and personal development. I am truly thankful to him for letting me attend his lectures, which improved my knowledge and skills significantly. I also wish to thank him for enhancing my critical thinking ability.

I owe my gratitude to my third supervisor, Dr. Abdel-Hamid Soliman, for his support during this research. His feedback on an earlier version of this thesis is greatly appreciated.

I would like to acknowledge Staffordshire University for the provision of PhD studentship and sponsoring this research. I especially thank the Head of Engineering Department at Staffordshire University, Mr. Cedric Belloc, for providing me opportunities to teach modules that fall within my area of research.

I would like to thank Mr. Muhammad Kamran Naeem, Mr. Siva Karteek Bolisetti and Mr. Raouf Abozariba, my fellow researchers in the Sensing, Processing and Communication Research Group at Staffordshire University for creating a pleasant working environment. I especially thank Mr. Muhammad Kamran Naeem for the useful and interesting discussions on this research. I would also like to thank my fellow researcher, Mr. Masum Billah.

I also wish to express my appreciation to my best friend, Mr. Zafar Khan, and his mother for making me feel like a part of their family. They have always been very supportive during my stay in the UK.

I would also like to express deepest gratitude to my parents and siblings who were a constant source of encouragement for me throughout this research. I am truly grateful to my parents for their unwavering support, love and prayers throughout my life. I especially thank my father, Mr. Muhammad Amjad Ali Siddiqui, for developing my interest in acquiring education from a very young age and for always motivating me during my educational career. I believe, without his support and faith in my abilities, my accomplishments in educational journey would not have been possible.

Anas Amjad
March 2017

List of Publications

J01 A. Amjad, A. Griffiths and M. Patwary, "QoI-Aware Unified Framework for Node Classification and Self-Reconfiguration Within Heterogeneous Visual Sensor Networks," in IEEE Access, vol. 4, pp. 9027-9042, Dec 2016.

J02 A. Amjad, M. Patwary, A. Griffiths and A. H. Soliman,"Characterization of Field-of-View for Energy Efficient Application-Aware Visual Sensor Networks," in IEEE Sensors Journal, vol. 16, no. 9, pp. 3109-3122, May 2016.

J03 A. Amjad, A. Griffiths and M. N. Patwary, "Multiple Face Detection Algorithm Using Colour Skin Modelling," in IET Image Processing, vol. 6, no. 8, pp. 1093-1101, Nov 2012.

C01 A. Amjad, M. Patwary and A. Griffiths "Energy Efficient Self-Reconfiguration Scheme for Visual Information based M2M Communication," in Proceedings of the 2017 IEEE 85th Vehicular Technology Conference (VTC Spring), Sydney, 2017.

Contents

List of Figures

List of Tables

Abbreviations

ACCM	**A**ctive **C**ompressive **C**alibration **M**atrix
BRISK	**B**inary **R**obust **I**nvariant **S**calable **K**eypoints
BRISKOLA	**BRISK O**ptimized for **L**ow-power **A**RM architectures
CBIR	**C**ontent-**B**ased **I**mage **R**etrieval
CCD	**C**harge-**C**oupled **D**evice
CCM	**C**ompressive **C**alibration **M**atrix
CMOS	**C**omplementary **M**etal **O**xide **S**emiconductor
DCT	**D**iscrete **C**osine **T**ransform
DWT	**D**iscrete **W**avelet **T**ransform
FoV	**F**ield-**o**f-**V**iew
FoVCC	**F**ield-**o**f-**V**iew **C**haracterisation **C**riteria
HM	**H**o**M**ogeneous
HT	**H**e**T**erogeneous
IoT	**I**nternet **o**f **T**hings
KB	**K**ilo **B**ytes
LDHF-dataset	**L**ong **D**istance **H**eterogeneous **F**ace dataset
LOTS	**L**ehigh **O**mnidirectional **T**racking **S**ystem
MGM	**M**ultiple **G**aussian **M**odel
MSE	**M**ean **S**quared **E**rror
PSNR	**P**eak **S**ignal-to-**N**oise **R**atio
PTZ	**P**an-**T**ilt-**Z**oom
QoI	**Q**uality-**o**f-**I**nformation

SGM	**S**ingle **G**aussian **M**odel
SIFT	**S**cale **I**nvariant **F**eature **T**ransform
SPIHT	**S**et **P**artitioning **I**n **H**ierarchical **T**rees
TI	**T**est **I**mages
VoI	**V**alue-**o**f-**I**nformation
VSN	**V**isual **S**ensor **N**etwork
WBAN	**W**ireless **B**ody **A**rea **N**etwork
WLAN	**W**ireless **L**ocal **A**rea **N**etwork
WPAN	**W**ireless **P**ersonal **A**rea **N**etwork

Notations

$\tilde{l}$	visual sensing node's index
$\mathrm{VS}_{\tilde{l}}$	$\tilde{l}th$ visual sensing node
$V_{\tilde{l}}$	3D coverage volume within the FoV of $\mathrm{VS}_{\tilde{l}}$
$\mathcal{N}$	number of visual sensing nodes in a network
$\mathbf{I}_{\tilde{l}}$	image captured by $\mathrm{VS}_{\tilde{l}}$
$\mathbf{I}$	a simplified notation for $\mathbf{I}_{\tilde{l}}$
f	focal length
(X, Y, Z)	a 3D scene point
(x, y)	a 2D point on the image plane
$P \times Q$	dimension of image $\mathbf{I}$
a	a pixel in $\mathbf{I}$
$f(\cdot)$	a function
VS	a simplified notation for VS_u
$\mathbb{R}^2$	2-dimensional Euclidean space
$\mathbb{R}^3$	3-dimensional Euclidean space
$P(X_c, Y_c, Z_c)$	a 3D scene point
$P'(x, z)$	a 2D point on the image plane
θ_h	horizontal Field-of-View
θ_v	vertical Field-of-View
ϕ_a	azimuth angle
θ_e	elevation angle
$ABCD$	a 2D plane

O_2	origin of the $ABCD$-plane
R	Sensing range of a visual sensing node
w_2	width of the $ABCD$-plane
h_2	height of the $ABCD$-plane
R_{min}	minimum sensing range
R_{max}	maximum sensing range
w_1	width of the image plane
h_1	height of the image plane
V	volume of the scene within a sensing node's FoV
R_{ref}	a reference distance
k	number of sensor classes in heterogeneous network
l	index for sensor class
n_l	number of nodes in sensing class l
j	index for representing number of sensing nodes
$\mathrm{VS}_{l,j}$	a sensing node in heterogeneous network
$\mathbf{VS}$	sensing nodes within the network
t	total sensing and processing tasks
t'_l	the tasks assigned to sensor class l
$\mathbf{T}$	task classification matrix
$\mathbf{T}'$	adaptive task classification matrix
$\mathbf{V}_c$	required 3D coverage of k sensor classes for t tasks
$\mathbf{v}$	chosen 3D coverage of k sensor classes for t tasks
b	a colour bin from a histogram
$h_c(b)$	a global colour histogram
$P(E)$	probability of a pixel belonging to object of interest
Cb	blue-difference chroma component in YCbCr colourspace
Cr	red-difference chroma component in YCbCr colourspace
γ^l_{Cb}	Cb lower bound for object detection
γ^u_{Cb}	Cb upper bound for object detection
γ^l_{Cr}	Cr lower bound for object detection

γ_{Cr}^{u}	Cr upper bound for object detection
$\mathbf{S}_m$	object segmentation matrix
$\mathbf{S}_g$	segmented image
d_h	horizontal density
d_v	vertical density
O_{po}	object pixel occupancy
$\lvert\varepsilon_d\rvert$	estimation error
PSNR_{dB}	peak signal-to-noise ratio in dB
A	area of an object of interest
ξ_o	object pixel occupancy lower bound
R_1	chosen range from object pixel occupancy
p_d	object's diameter in pixels
d_e	object's diameter in metres
d_a	object's actual measured diameter in metres
ξ_d	estimation error upper bound
R_2	chosen range from estimation error
ξ_p	PSNR_{dB} lower bound
R_3	chosen range from PSNR_{dB}
R_c	chosen sensing range from FoVCC
$\mathbf{R}_c$	estimated sensing range of k sensor classes for t tasks
$\mathbf{r}$	chosen sensing range of k sensor classes for t tasks
N_a	number of required active sensing nodes in the network
E_{Acq}	energy consumption for image acquisition
E_{Tx}	energy consumption for image transmission
E_{Rx}	energy consumption for receiving an image
q	index to denote image acquisition, transmission or receiving cost
E_q	image acquisition, transmission or receiving cost
$\tilde{E}_q$	overall image acquisition, transmission or receiving cost
E	total energy consumption within the VSN
LT	network lifetime

$\hat{t}$	number of tasks allocated to a sensor class
ζ^l	dynamic PSNR range's lower bound in dB
ζ^u	dynamic PSNR range's upper bound in dB
A_T	surface area of the total area of interest
V_T	volume of the total area of interest
R_r	reference distance within close proximity of a sensing node
R_d	sensor-to-object distance
λ	QoI index
λ_t	target QoI threshold in dB to be achieved for a given application
$\hat{\mathbf{r}}$	set of possible sensor-to-object distances
$\mathbf{\Lambda}_{\tilde{l}}$	compressive calibration matrix
$\theta_{y,e}$	vertical anti-clockwise angle from y-axis to sensor's LoS
$\theta_{y,a}$	horizontal anti-clockwise angle from y-axis to sensor's LoS
$\mathbf{\Omega}_t$	set of heterogeneous target QoI thresholds in dB
$\mathbf{C}_r^{\tilde{i}}$	3D coordinates of the $\tilde{i}th$ region of interest
$\mathbf{C}_s^{\tilde{l}}$	3D coordinates of the region within $\tilde{l}th$ visual sensing node's FoV
$\mathcal{O}_{\tilde{i}\tilde{l}}$	overlap between $\tilde{l}th$ sensing node's FoV and $\tilde{i}th$ region of interest
$\mathcal{H}$	degree of heterogeneity in target QoI thresholds
β	QoI delivered by a visual sensing node
$[e^-, e^+]$	confidence bound for the fidelity of CCM
$[\mathcal{P}_r^-, \mathcal{P}_r^+]$	performance reliability bound
$\hat{E}_{tx}$	average energy cost per node for transmitting an image frame
$\hat{E}_c$	average energy cost per node for transmitting $\mathcal{N}_t$ image frames

Dedicated to my beloved parents and siblings

Chapter 1

Introduction

This chapter presents an overview of the characteristics and constraints of visual sensor networks as well as some possible solutions to overcome such constraints; subsequently, it provides the motivation to work within the targeted research area. It also discusses various applications of wireless visual sensing as well as the challenges faced during the design and implementation phases. Afterwards, the chapter proceeds towards the aim and objectives of the research study. The methodology adopted to conduct this research study is presented. A brief summary of the contributions to knowledge by this research study is also given. Finally, a summary of the thesis organisation is presented at the end of this chapter.

1.1 Background and Motivation

A Wireless Sensor Network (WSN) consists of a group of sensor nodes with sensing, processing and communication capabilities. In traditional WSNs, sensors generally provide coverage in all directions to collect scalar measurements as 1D data, for example: temperature, pressure, humidity etc., which limits suitability of such sensors for many applications [1]. In order to enhance WSN's suitability

for a wider range of applications, traditional sensors are replaced by visual sensors resulting in a network suitable for a new scope of applications known as a Visual Sensor Network (VSN). In a VSN, each node captures image data that can be processed locally to extract relevant information (such as visual features) [2]. Moreover, sensing nodes in the network can collaborate to share such information. VSNs are used in surveillance [3–6], environmental monitoring [7, 8], object detection and tracking [9, 10], health care monitoring [11–14] and many other applications. Visual sensors within a VSN employ directional sensing to provide pixel based measurements as a 2D dataset and therefore, require a large bandwidth to transmit image data. The visual sensor's 3D viewing volume (i.e. the extent of the observable scene) is known as its Field-of-View (FoV) [2]. The FoV depends on a visual sensing node's location and orientation parameters. A modification in any of these parameters will lead to a change in the FoV. Therefore, precise knowledge of visual sensing nodes' location and orientation information is required for VSN management. Moreover, during the VSN design phase, some image processing algorithms require precise knowledge of the FoV. Cooperation can be exploited among visual sensing nodes for intelligent sensing and processing of the data acquired from the targeted sensing environment, independent of the given application. However, in order to facilitate such intelligent sensing within a visual sensing node, a dynamic coverage modelling approach is required to obtain the 3D FoV information. Owing to these fundamental differences between a traditional WSN and a modern VSN, the deployment of the latter is more challenging as compared to the former. Furthermore, due to the directional sensing nature of visual sensors, the existing WSN design solutions are not suitable for VSNs.

Some applications consider a large 3D plane as the targeted sensing environment, consequently requiring a large number of sensing nodes to provide full coverage. In case of battery failure within a visual sensing node, battery replacement may not be feasible and thus, a new sensing node must be added in the network to support the application. Therefore, VSN design solutions should be scalable and

the network performance must be independent of its size [15]. The algorithms for the management of VSNs can be categorised into centralised and distributed approaches [16, 17]. In centralised approaches, a central node is responsible for configuring visual sensing nodes and making appropriate decisions. On the other hand, in distributed approaches, visual sensing nodes are able to find suitable configuration parameters without any support from the central node. Distributed approaches are usually less complex and support scalability and hence, they are preferred over the centralised approaches [17, 18]. The intelligence of a VSN can be enhanced by utilising nodes with different sensing and processing capabilities. This results in a heterogeneous network which provides better functionality as compared to its homogeneous counterpart. Furthermore, heterogeneous VSNs play a prominent role in emerging Internet of Things (IoT) applications. However, designing algorithms for efficient distribution of sensing and processing tasks between visual sensing nodes within heterogeneous networks is a challenging task [1].

Energy is a scarce resource in a VSN due to the resource constrained nature of its nodes and the possibility of deployment in inadequately resourced areas to support complex algorithms [19, 20]. Therefore, the main challenge in designing VSNs is to utilise resources optimally while maintaining a certain degree of reliability, as per the given application. Efficient utilisation of network resources and optimisation of processing algorithms lead to the conservation of energy resulting in increased lifetime. In order to explore the challenges in more detail, consider a VSN deployed at a remote location for a surveillance application such as face detection, object detection and tracking etc. Since a power source may be unavailable, all nodes are assumed to be battery powered therefore, the network lifetime is limited. This imposes tight constraints on energy consumption and data storage capacity within a VSN. Furthermore, the aforementioned surveillance tasks vary in terms of complexity and desired reliability. As an example, distributed video coding is utilised in [21] and the energy consumption of a VSN is observed. Each

sensing node is battery powered with 20250 Joules of energy. Considering the transmission of 30 frames, the analysis revealed that the energy consumption is 16 Joules and 12 Joules for uncompressed and compressed frames respectively. The network lifetime with a transmission rate of 30 frames per second is found to be approximately 1687 seconds and 1265 seconds for uncompressed and compressed scenarios respectively. Therefore, the characterisation of FoV and task classification to provide an energy efficient application-aware design solution that can prolong the network lifetime is an important and challenging problem in VSNs.

The performance and lifetime of a VSN is characterised by its configuration. The configuration space of a network is defined by the set of parameters that actively control the quality and amount of acquired data. These parameters include resolution, frame rate, aperture, exposure time and level of visual data compression for networks comprising static visual sensing nodes. In addition to these, the orientation and FoV parameters are also considered for the configuration of networks comprising Pan-Tilt-Zoom (PTZ) capable visual sensing nodes [22]. Reconfiguration can be defined as the process of updating one or more parameters that form the configuration space of the network to achieve a specific goal, for example, maximum detection reliability with minimum possible energy consumption. Implementing self-reconfiguration schemes dynamically within a visual sensing node to achieve a specific goal for a given application can result in improved reliability and optimised energy consumption configuration.

Quality-of-Information (QoI) is defined in literature as the degree to which the data is suitable for a given application or a decision making process [23–25]. Within the context of surveillance applications, QoI reflects the degree to which the data produced at the output of a visual sensing node accurately quantifies the actual event being monitored [26]. In order to enhance the sensing and processing intelligence within a VSN, heterogeneity can be introduced in the target QoI based on the characteristics of the targets expected to be monitored within the sensing nodes' FoV. However, due to the strict orientation requirement and directional nature of

visual sensing nodes within a 3D plane, handling a wide range of heterogeneity within target QoI is a challenging task.

Although the existing work in literature addresses several key issues relating to VSNs such as FoV identification, coverage estimation, feature extraction, camera scheduling and visual data transmission etc.; it is found that the existing solutions are application-specific under particular resource constrained scenarios. Furthermore, many existing solutions have limited application capabilities as they assume a homogeneous network during the design phase. Moreover, despite the fact that various existing state-of-the-art techniques provide solutions to optimise the energy consumption in resource constrained scenarios, many schemes assume a simplified 2D sensing model and sensing environment; whereas, 3D modelling is more realistic for VSN applications. Although, techniques are found in the literature that utilise 3D modelling or provide solutions for heterogeneous VSNs; however, to the best of the author's knowledge, the existing schemes do not exploit the distributed nature of visual sensing nodes to adopt heterogeneity in the targeted QoI along with dynamic targeted threshold based optimisation.

In addition to these, the existing schemes do not consider the heterogeneous realisation of targeted QoI within the sensing region and dynamic targeted threshold based optimisation of visual data prior to transmission in VSNs. Hence, generalised and more realistic solutions are required to overcome the VSNs' resource constraints for a diverse range of applications.

Based on the aforementioned challenges faced in resource constrained VSNs, this thesis addresses the following research questions:

- Can FoV characterisation provide energy savings in resource constrained scenarios?
- Can the design and unification of node classification and self-reconfiguration schemes enhance the energy efficiency of resource constrained heterogeneous VSNs?

This thesis tackles the problems faced in resource constrained VSNs by proposing FoV characterisation, task classification, QoI-aware node classification and self-reconfiguration schemes to select optimised sensing node configurations leading to an improved network energy efficiency. The proposed schemes provide reliable solutions to the VSN design problems for the pre-deployment phase and VSN dynamic optimisation problems for the post-deployment phase. The proposed energy efficient FoV characterisation framework exploits heterogeneity of the network to obtain an application-aware design solution. Moreover, the unified framework proposed in this thesis incorporates 3D FoV modelling with dynamic realisation of visual data to achieve heterogeneous target QoI thresholds. The proposed unified framework leads to reduced energy consumption within the network whilst maintaining an acceptable degree of reliability.

In this thesis, QoI is characterised by the quality of the visual data provided by a sensing node and quantified by a Peak Signal-to-Noise (PSNR) based metric, as in [26]. The configuration of a VSN considered for FoV characterisation and task classification in this thesis is given by: a) the sensing range of the nodes and b) the allocation of sensing and processing tasks to the nodes which are part of a heterogeneous network. Whereas, the configuration of a visual sensing node for optimisation within the reconfiguration phases is considered to be the amount and quality of the visual data for transmission. In VSNs, the energy consumed during the communication phase (i.e. transmitting and receiving visual data) is significantly higher than the processing phase [27–29]. For example, the energy cost for transmitting 1 Kilo Byte data over a distance of 100 metres is equivalent to the energy required to process three million instructions [30–32]. Therefore, optimising the amount of visual data based on the targeted QoI thresholds prior to transmission can result in energy savings. In this context, the parameter that forms the network's configuration space is the level of compression employed by a visual sensing node.

1.2 Wireless Visual Sensing: Applications and Requirements

Wireless visual sensing serves an important purpose for a broad spectrum of applications such as surveillance, target classification, traffic monitoring, health care, and environment monitoring etc. The applications are discussed in more detail below.

- An application of wireless visual sensing is video surveillance which aims to monitor specific locations for intrusion detection and public safety. In conventional surveillance systems, visual data is sent from sensing nodes to the control centre (i.e. sink node) which is monitored by an operator. However, autonomous management of visual data is desirable due to the difficulties associated with manual monitoring of visual data. Many existing approaches provide design solutions for autonomous management of visual data at the sink node. However, the centralised nature of such solutions limits the scalability of VSNs [4]. A smart video surveillance system for distributed and collaborative processing of visual data within visual sensing nodes is proposed in [4] for autonomous scene analysis. The ground points of human subjects are estimated through an interference framework and a semantic topology is employed to construct a work-tree. Visual sensing nodes' existence probabilities are calculated and combined with the work-tree. Due to its distributed nature, the approach is highly efficient for smart video surveillance applications.

- Another surveillance application of wireless visual sensing is autonomous object detection and tracking. Due to the movement of objects within the sensing nodes' FoV, object detection is a challenging task. Once the objects have been detected, visual sensing nodes are required to keep a track of the

objects' movements to support the given application. However, this is a challenging task due to the high reliability requirements of tracking applications and resource constrained nature of sensing nodes. A cost-aware approach for collaborative target tracking in resource constrained VSNs is proposed in [10]. A decentralised resource allocation strategy is devised through the formation of coalitions by the sensing nodes. Negotiation between sensing nodes to form a coalition is dynamically supervised by a coalition manager for cooperative target tracking. In cost-free scenarios, the approach exhibits improvement in tracking reliability. Furthermore, under resource constrained scenarios, the approach provides a trade-off between reliability and energy efficiency.

- Target classification is an important application of wireless visual sensing which aims to classify targets within a sensing node's FoV into predefined categories. However, such classification is a challenging task due to the resource constrained nature of VSNs. The algorithms for the classification of targets can be divided into binary and multiclass categories. In binary classification algorithms, a decision is made to determine whether a target belongs to a particular category or not such as human/non-human classification as in [5]. Multiclass algorithms, on the other hand, provide solutions for the classification of multiple targets into suitable categories. A target classification framework for VSNs is proposed in [33] for both binary and multiclass animal classification. Firstly, wavelet transform is utilised for feature extraction from the captured images. Secondly, targets are classified using if-then rule based decision trees. A suitable dataset is utilised for the training of the classifier in the pre-deployment phase. After the training process, the sensing nodes are equipped with the knowledge of if-then rules. Hence, the approach leads to reduced learning phase time for energy efficient target classification with high accuracy.

- Visual sensing has become significantly important in traffic monitoring and control applications due to the growing traffic demands. Smart autonomous traffic monitoring and management is required in smart cities for congestion control and safety provision. Smart monitoring systems should support intelligent incident detection and robust classification of targets. However, traffic monitoring and control are challenging tasks due to the background changes, illumination variations and occlusion problems within the sensing environment [34]. An autonomous traffic monitoring system is proposed in [34] to support a diverse range of traffic monitoring applications for airports, highways and tunnels etc. An image processing unit is incorporated in the system which provides key information relating to the moving objects within the sensing nodes' FoV. The information is used to obtain statistical traffic information and trigger real-time alerts. The system architecture supports scalability without significantly affecting the computational complexity.

- Nowadays, due to the financial challenges faced by the health care sector, wireless visual sensing is a promising solution for smart health care monitoring applications. Sleep is a significantly important factor for the wellbeing of the elderly. Sleep disorders in elderly can cause fatigue and deteriorate physical functionality, consequently affecting the quality of life [14]. Therefore, the sleep patterns of elderly must be regularly monitored to reduce such risks. However, manually monitoring the elderly sleep patterns is not a cost effective solution. In [14], authors have proposed a low-cost solution for long-term sleep analysis for elderly care using low-resolution visual sensing nodes. The approach detects motion patterns and employs a rule-based approach to identify the wake-up and sleep times. The performance of the approach evaluated using real-life video data shows reasonably accurate estimations of sleep duration.

- Another application of wireless visual sensing within the context of pervasive health care delivery is the monitoring of elderly wellbeing through

gait impairment quantification. Such quantification aims to analyse post-operative recovery of patients by detecting the progression of diseases such as Parkinson's. Human motion is detected for the understanding of patients' activities to identify gait abnormalities. However, this is a challenging task as the supervised models of classifier requires both normal and impaired gait samples from a patient. Furthermore, a large training data set to model the behaviour of a patient's motion is required which is not feasible due to the resource constrained nature of visual sensing nodes [35]. In order to overcome these challenges, an unsupervised change analysis method is used in [35] to quantify a patient's gait impairment. Target blob extraction is utilised to support feature extraction and change analysis. A target blob skeletonisation method is employed for the quantification of target's internal motion. The method computes motion metrics and utilises change analysis method to observe pre-operative and post-operative gait changes.

- An important application of wireless visual sensing to ensure public safely is environment monitoring. Within the context of industrial environment monitoring applications, keeping a human operator within the proximity of the surveyed area poses potential concerns of compromising on the operator's health. Therefore, autonomous monitoring of environment for critical events detection is required [7]. The data (such as temperature, pressure etc) obtained from scalar sensing nodes may not always be sufficient for efficient environment monitoring. Thus, incorporating both scalar and visual sensing nodes is expected to improve the detection reliability of an environment monitoring system. An autonomous environment monitoring system is proposed in [7] by incorporating both scalar and visual sensing nodes for accidents detection such as fire and gas leaks. Scalar sensing nodes provide the temperature, poisonous gas concentration and humidity measurements. Visual sensing nodes analyse the colour information and flame behaviour for

fire detection. The approach improves the accident detection reliability and facilitates environment monitoring from a remote location.

1.3 Aim and Objectives

The aim of this research is to investigate the challenges faced in VSNs due to their resource constrained nature and to develop schemes for optimised sensing node characterisation by utilising visual features. As mentioned previously, the multi-dimensional nature of visual data and its transmission impose strict constraints on energy and bandwidth in VSNs. Therefore, in order to obtain the optimal configuration of visual sensing nodes, FoV characterisation, coverage modelling, node classification and self-reconfiguration in VSNs are required. Moreover, a unified framework can be developed for node classification and self-reconfiguration in VSNs to achieve targeted threshold based optimisation. Such unified framework is expected to enhance the energy efficiency of resource constrained VSNs while maintaining an acceptable degree of reliability.

The following objectives have been identified for this research:

- To investigate VSNs along with their characteristics, constraints and implementation challenges.
- To explore the existing solutions for VSN coverage, design, characterisation and optimisation problems with their strengths and weaknesses.
- To develop a FoV characterisation framework with feature utilisation for resource constrained homogeneous and heterogeneous VSNs, which minimises the energy consumption as well as maximises the network lifetime.
- To analyse the energy efficiency and failure probability of the proposed FoV characterisation framework for homogeneous and heterogeneous VSNs.

- To develop an energy efficient visual sensing node dynamic self-reconfiguration scheme for targeted threshold based redundant feature removal prior to transmission in VSNs.
- To develop 3D coverage modelling and QoI-centric node classification schemes by introducing heterogeneity in the network resulting in enhanced sensing and processing intelligence of VSNs.
- To propose a unified framework for node classification and self-reconfiguration within heterogeneous VSNs, and to formulate an analytical model for the quantification of the its performance reliability.
- To evaluate the performance of the proposed unified framework in terms of resource utilisation by comparing with the existing state-of-the-art techniques.

1.4 Research Methodology

In this section, the methodologies that support researchers during the research process are presented and the research process is discussed. The research methodology chosen for conducting this research project is also presented. Research can be defined as a process for the collection, analysis and interpretation of information that results in answering particular questions. The research process is required to possess particular characteristics which are given below [36].

- The process should be controlled and rigorous.
- The procedures utilised to conduct the investigation should be systematic.
- The findings of the research should be based on hard evidence so that they can be validated and verified.
- The research process should be able to withstand critical scrutiny.

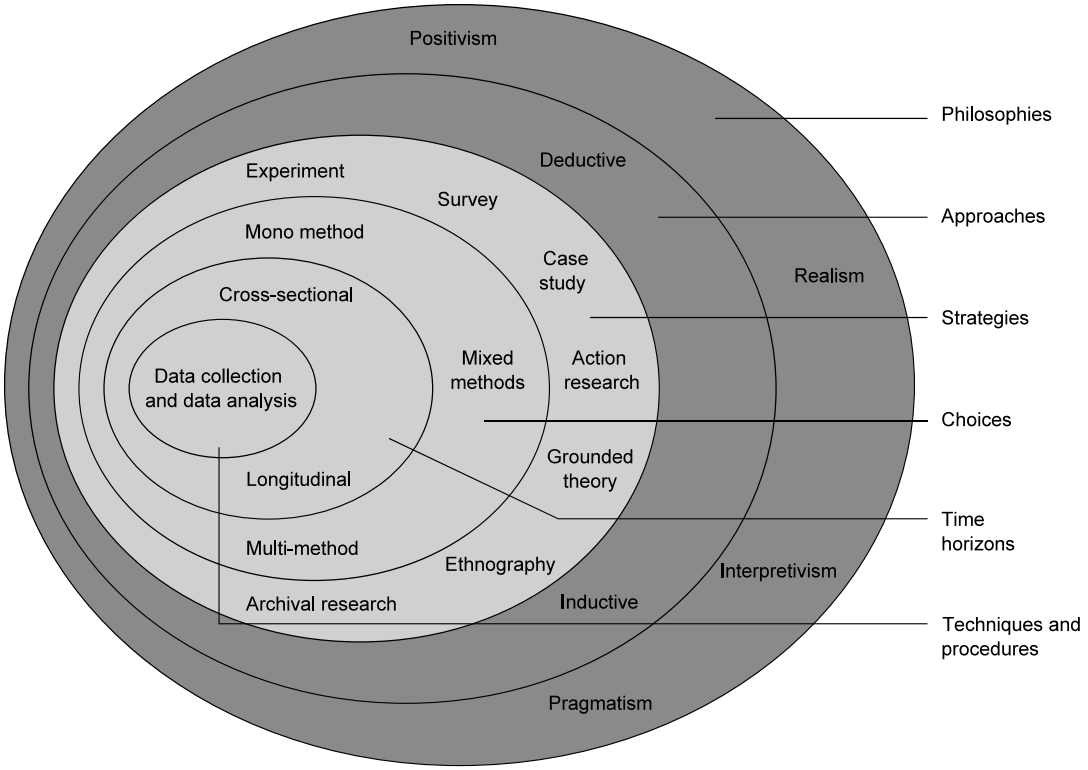

FIGURE 1.1: The research 'onion' [37]

In order to carry out a research project, the research study can be classified into qualitative and quantitative categories. Qualitative research is unstructured and flexible, whereas, quantitative research is structured and rigid. Moreover, qualitative research aims to investigate the variation in a particular phenomenon, whereas, quantitative research studies the degree of variation in a particular phenomenon. Qualitative research focuses on the description of variables under consideration, whereas, quantitative research concentrates on the measurement of variables [36]. Saunders *et al.* [37] related different stages involved in the research process with different layers of 'onion', as shown in Figure 1.1. There are six layers of the research onion model as given below [37].

1. research philosophy
2. research approach
3. research strategy

4. research choice
5. time horizon
6. techniques and procedures

The first layer of the research 'onion' defines the research philosophies which relate to knowledge development and include positivism, realism, interpretation and pragmatism. The second layer refers to the approaches that can be utilised to carry out research and include deductive and inductive approaches. In the deductive approach, theory and hypotheses are developed first and then the research strategy is designed to test the hypotheses. Whereas, inductive approach consists of theory development based on data collection and analysis. The third layer defines the strategies that can be employed for descriptive, exploratory and explanatory research [38]. The selection of a particular research strategy depends on many aspects such as the research questions, research objectives, research time duration and available resources etc. Action research, ground theory and ethnography are qualitative research approaches, whereas, experiment, survey and case study are classified as quantitative research approaches.

The fourth layer refers to research choice and depends on the selection of qualitative and quantitative approaches. This layer includes mono method, mixed methods and multi-method. The fifth layer relates to the time duration of the research project. This layer categorises research studies into cross-sectional and longitudinal studies. In cross-sectional research, a particular phenomenon is investigated at a particular time. Survey based strategy is usually employed in cross-sectional research studies [39, 40]. On the other hand, longitudinal studies are conducted over time and they have the ability to study particular behaviours that are not possible to be studied with a cross-sectional study. The last layer of the research onion model relates to the techniques and procedures for data collection and analysis. Data can be categorised into primary and secondary data, whereas, the data

collection and analysis techniques that can be adopted to conduct research depend on the type of the data.

In order to accomplish the objectives of this research (presented in Section 1.3), the quantitative research methodology is adopted. An experiment-based strategy is utilised to validate the robustness of the proposed FoV characterisation framework, dynamic self-reconfiguration scheme, 3D coverage modelling scheme, QoI-centric node classification scheme and the unified framework. In the experiments, the values of different parameters are changed and the behaviour of the developed frameworks is studied. The different stages of this research project are given below.

- Investigation of the existing research work and identification of the challenges within the context of VSN design and optimisation.
- Formulation of the research questions and identification of the aim and objectives of the research project.
- Design and development of a FoV characterisation framework to minimise the energy consumption of homogeneous and heterogeneous VSNs.
- Validation of the proposed FoV characterisation framework's robustness by utilising MATLAB simulation platform.
- Design and development of a dynamic self-reconfiguration scheme for visual sensing nodes.
- Analysis of the proposed dynamic self-reconfiguration scheme's energy efficiency in resource constrained scenarios using MATLAB simulation platform.
- Design and development of 3D coverage modelling and QoI-centric node classification schemes.
- Development of a unified framework for node classification and dynamic self-reconfiguration in heterogeneous VSNs.

- Evaluation of the proposed unified framework's energy efficiency and reliability using MATLAB simulation platform.

1.5 Research Contributions

The contribution of this thesis is the design of energy efficient and application-aware solutions for sensing nodes characterisation to resolve the issues faced within VSNs due to their resource constrained nature. A generalised framework for the characterisation of visual sensing nodes' FoV within homogeneous and heterogeneous VSNs is developed. Furthermore, prior to transmission within VSNs, a QoI-aware self-reconfiguration scheme to dynamically obtain optimal visual sensing node configuration for feature redundancy removal is also proposed. 3D coverage modelling and QoI-centric node classification schemes are developed to enhance the sensing and processing intelligence of heterogeneous VSNs. A unified framework for node classification and self-reconfiguration within heterogeneous VSNs is also proposed. The contributions by this research study are briefly described as follows:

- A generalised FoV characterisation framework for homogeneous and heterogeneous VSNs is proposed as a function of the required minimum object pixel occupancy, maximum allowable error tolerance and desired image quality. The proposed FoV characterisation framework provides the system design engineers with a resource trade-off model while obtaining an optimised sensing range of a visual sensing node for any given application. (Chapter 3)

- Considering the heterogeneity of the modern VSNs, an adaptive task classification scheme is proposed for the distribution of tasks between the nodes providing a trade-off model for reliability and energy efficiency. The proposed scheme provides solutions to the task classification problem feasible

for implementation in resource constrained scenarios. Moreover, a comparison of hard decision and proposed soft decision based techniques is presented. The proposed framework, when employed with the proposed task classification and soft decision based sensing range selection schemes results in an optimised VSN configuration by maximising the spatial coverage, reducing the energy consumption and increasing the network lifetime without compromising on the desired reliability. Analysis of the energy efficiency of the proposed framework validates its suitability for a diverse range of applications. (Chapter 3)

- A learning strategy is devised which can be employed by design engineers during the network pre-deployment phase. The proposed learning approach formulates a relationship between objects' variations within a sensing node's FoV, the level of detail in the acquired visual data and the effect of feature redundancy removal on the data. (Chapter 4)

- A dynamic self-reconfiguration scheme for resource constrained VSNs is proposed as a function of the targeted QoI threshold to be ensured based on the application design criteria. Utilising the proposed learning strategy and the in-node processing model, the scheme yields optimised configurations for visual sensing nodes resulting in substantial energy savings. Subsequently, providing system design engineers with a trade-off model between reliability and energy efficiency. (Chapter 4)

- A 3D coverage modelling scheme is proposed for visual sensing nodes to dynamically obtain their FoV information. The proposed scheme can be utilised in the network initialisation phase to support intelligent sensing by making the sensing nodes aware of the targeted sensing environment. (Chapter 5)

- By the heterogeneous realisation of the targeted QoI within the sensing regions, a QoI-centric scheme is proposed for the classification of visual sensing

nodes. The proposed coverage modelling and node classification schemes, coupled with the in-node processing model are incorporated in the proposed unified framework. The unified framework provides feasible solutions to guarantee targeted QoI satisfaction with optimised energy utilisation in resource constrained VSNs. (Chapter 5)

- An analytical model is formulated to quantify the performance reliability as a function of the targeted and delivered QoI thresholds. For a given application, the proposed analytical model provides system design engineers with the confidence bounds for fine-tuning to the required QoI thresholds while attaining the desired reliability. (Chapter 5)

The contributions by this thesis within the context of the limitations observed in existing literature are summarised in Figure 1.2.

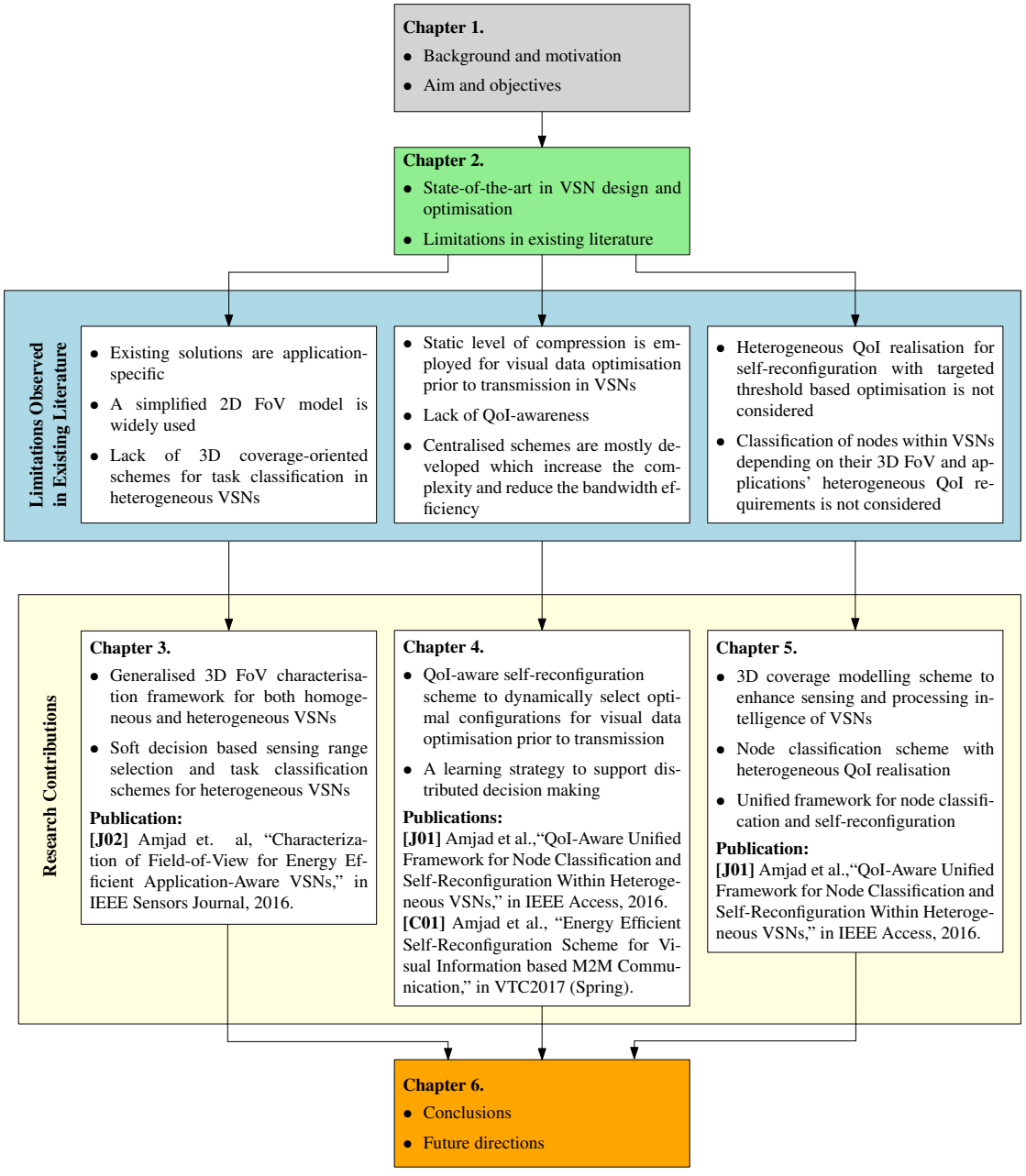

FIGURE 1.2: An overview of the contributions by this thesis within the context of limitations observed in existing literature

1.6 Thesis Organisation

The rest of the thesis is organised as follows:

Chapter 2 provides a comprehensive survey of some notable existing state-of-the-art techniques for VSN design and optimisation. Visual sensing node architecture is presented along with a discussion on some notable coverage and deployment methods. Existing techniques for visual data processing, task classification and dynamic self-reconfiguration are provided. The chapter also highlights limitations of the existing techniques within the context of resource constrained scenarios.

Chapter 3 presents the proposed FoV characterisation framework which assists the design engineers during the network pre-deployment phase in obtaining energy efficient design solutions for both homogeneous and heterogeneous VSNs. Soft decision criteria is exploited to maximise the spatial coverage of heterogeneous visual sensing nodes. Moreover, a task classification scheme is proposed for heterogeneous VSNs which supports the FoV characterisation framework during the post-deployment phase. Finally, the energy efficiency and reliability of the proposed FoV characterisation framework are evaluated.

Chapter 4 adopts the notion of QoI-awareness and describes the proposed self-reconfiguration scheme for visual sensing nodes to enhance their energy efficiency while providing acceptable level of reliability during the network post deployment phase. A generalised visual sensing node self-reconfiguration model is provided and various aspects of the reconfiguration process are discussed. In order to enable distributed decision making, the proposed self-reconfiguration scheme is supported by a training and calibration process which takes place in the network pre-deployment phase. The robustness of the proposed self-reconfiguration scheme is evaluated by observing the energy savings for various target QoI thresholds.

Chapter 5 presents a unified framework of node classification and self-reconfiguration with heterogeneous QoI realisation. A 3D coverage modelling scheme is proposed

to enhance visual sensing node's awareness of the sensing region within their FoV. A node classification scheme is developed which is driven by the application's QoI requirements and supported by the 3D coverage modelling scheme. An efficient framework is obtained by unifying the 3D coverage modelling, node classification and self-reconfiguration schemes. The chapter also provides an analytical model to gauge the proposed framework's performance reliability. The performance of the proposed framework in terms of energy consumption is analysed by comparing with the existing state-of-the-art techniques.

Finally, Chapter 6 summarises the thesis, discusses future scope of the work based on the schemes proposed in the thesis and provides the concluding remarks.

Chapter 2

State-of-the-art in VSN Design and Optimisation

2.1 Introduction

This chapter explores the existing VSN design and optimisation techniques along with their limitations within the context of resource constrained scenarios. In order to conduct the literature review, a systematic approach [41, 42] is adopted. Initially, the existing work which is relevant to the research study is identified through an exhaustive, rigorous and comprehensive search. The inclusion and exclusion criteria is employed which can depend on several factors such as relevance, validity of experimental design, analysis and results etc. The research studies that satisfy the inclusion criteria are selected from the existing work identified in the first stage of the review. Next, the quality of the selected existing research studies is assessed. Afterwards, the key findings from the existing work are extracted and limitations within the context of the problem under consideration are identified. In this chapter, the architecture of a visual sensing node is presented along with the hardware components and several examples. Existing state-of-the art techniques for visual sensing node coverage and deployment are presented. Various schemes

found in the literature for visual data processing are discussed. An overview of some notable task classification schemes from the literature is provided. Finally, state-of-the art resource-aware, coverage-oriented and target-based existing techniques for dynamic reconfiguration are discussed.

2.2 Visual Sensor Networks

In recent years, a significant number of research studies have been conducted on VSNs. The interest in this context lie in the optimisation of: sensing nodes' spatial coverage, task allocation, resource utilisation and visual data for transmission. In resource constrained scenarios, such optimisations aim to provide energy efficient solutions for the lifetime maximisation of VSNs. The hardware components that constitute a visual sensing node are discussed in Section 2.2.1. Furthermore, some notable schemes from the existing literature for optimisation are discussed in Section 2.2.2 to Section 2.2.5.

2.2.1 Visual Sensing Node Architecture

The hardware components of a visual sensing node consist of an image capturing device, a processing unit, a wireless transceiver and a power unit [43], as shown in Figure 2.1 [44]. Visual sensing nodes can also be equipped with additional

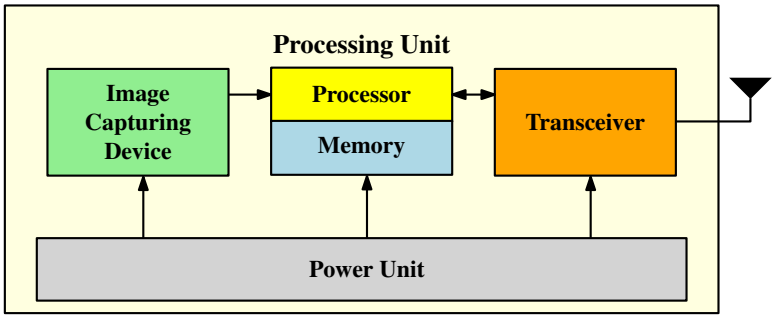

FIGURE 2.1: Hardware components of a visual sensing node [44]

application-specific components such as a location finding system and a mobiliser [44]. The imaging devices can be classified into Charge-Coupled Device (CCD) and Complementary Metal Oxide Semiconductor (CMOS) categories. Some examples of CCD-based imaging are Meerkats [45] and Panopes [46], whereas, CMOS-based imaging are Vision Motes [47] and Cyclops [48]. Although, CMOS-based imaging devices are compact and less costly as compared to their CCD counterparts, the latter provide superior image quality as compared to the former [43]. For example, Cyclops [48] utilises CMOS-based imaging providing visual data of resolution 352×288 and Meerkats [45] employs CCD-based imaging providing visual data of resolution 640×480. Image capturing devices that produce a high resolution image provide high quality visual content but result in high processing and transmission energy costs. Hence, there is a trade-off between the quality of visual data and resource utilisation. Due to the resource constrained nature of VSNs, an image capturing device with suitable resolution needs to be selected to support a given application.

The processing unit within a visual sensing node comprises of an embedded processor and a storage unit. The on-board processor performs various tasks assigned to a visual sensing node such as feature detection, extraction, visual data optimisation etc. Due to the multi-dimensional nature of data acquired by a visual sensing node, VSNs require much higher storage compared to WSNs. The memory requirements within visual sensing nodes is at least ten times higher than the traditional sensing nodes [43]. For example, Cyclops [48] provides 8-bit monochrome, 16-bit YCbCr colour, and 24-bit RGB colour images of resolution 352×288 and require 101.4 Kilo Bytes (KB), 202.8 KB and 304.1 KB memory respectively to store a single image frame. On the other hand, for a pixel representation with 24 colour bits, Meerkats [45] provides an image of resolution 640×480 and requires 921.6 KB storage.

The wireless transceiver modules for sensor networks can be categorised into Wireless Personal Area Networks (WPANs) and Wireless Local Area Networks

(WLANs). WPAN transceiver modules typically consist of IEEE 802.15.4 and Bluetooth (IEEE 802.15.1) which support a maximum data rate of 250 Kbps and 1 Mbps respectively resulting in low power consumption. Furthermore, IEEE 802.15.4 and Bluetooth have a limited transmission range of 10m to 20m. Therefore, such modules are used for applications that can tolerate a low data rate and transmission range but require low transmission energy cost. In contrast to WPAN modules, IEEE 802.11b is typically used as a WLAN transceiver module in VSNs. It offers a maximum data rate of 11 Mbps with a transmission range of upto 100m and is suitable for applications that demand real-time streaming. However, as it provides a high transmission range, this leads to a higher transmission energy cost. Hence, there is a trade-off between transmission range and energy consumption.

2.2.2 Visual Sensing Node Coverage and Deployment

The deployment of visual sensing nodes within the targeted sensing environment depends on the application's requirements and the environmental conditions. The deployment methods can be classified into deterministic and random categories. Deterministic node deployment is feasible for sensing environments with known characteristics, such as indoor environments. It results in coverage maximisation with minimum possible sensing nodes leading to a reduced network cost. Furthermore, it does not create FoV overlapping and occlusion problems. However, this type of deployment requires a strategy to be developed during the pre-deployment phase. On the other hand, random sensing node deployment is suitable for large scale VSNs with restricted access to the sensing environment. This technique is simple as it does not require design engineers' attention in the pre-deployment phase. However, it can result in FoV redundancy due to the deployment of excessive sensing nodes to compensate for the lack of deterministic positioning. Moreover, random deployment leads to occluded regions within the sensing environment [49]. Therefore, for a given application, a suitable deployment strategy must be

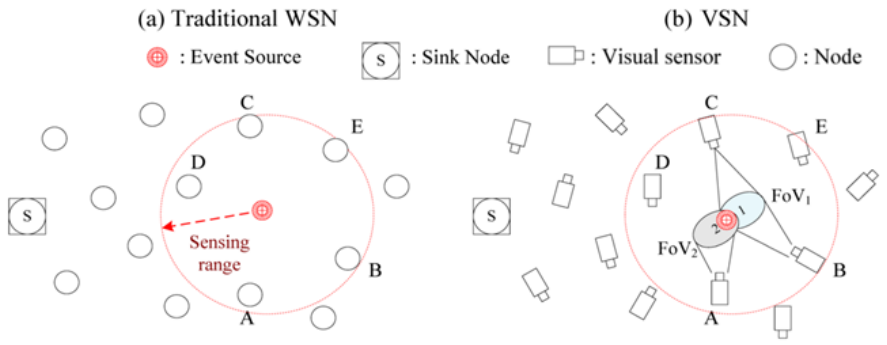

FIGURE 2.2: Comparison of a traditional WSN with a VSN in context of the sensing nature of the nodes [43]

employed. The post-deployment challenges faced in VSNs due to random deployment make coverage problem popular within the research community.

In traditional WSNs, an omni-directional sensor provides coverage to an event if it is within the sensing range. Whereas, apart from the sensing range, the event must be within the FoV of the visual sensing node in VSNs. Figure 2.2 [43] shows comparison of a traditional WSN with a VSN. Due to the directional nature of a visual sensing node's FoV shown in Figure 2.2, existing WSN solutions are not suitable for VSNs. Early scenarios related to the optimal sensing node's coverage and placement can be found in the art gallery problem [50] which deals with the estimation of optimal number of guards and their placement to secure an art gallery. However, as compared to the art gallery problem, the directional FoV and limited range of a visual sensing node makes the coverage and placement problems much more challenging.

A mathematical model is proposed in [51] which utilises optimisation-based heuristics to solve the VSN coverage problem by deploying each node sequentially and removing the overlapping nodes. The approach shows that in small grids, reduced number of sensing nodes will be required with fixed span angle and sensing range. Authors in [52] presented a deployment algorithm for video panorama creation in heterogeneous VSNs comprising sensing nodes with different costs, resolutions and FoVs. The algorithm takes into account the minimum desired resolution and

area boundaries to minimise the total deployment cost without compromising on the multi-perspective coverage of the area. A deployment methodology is provided in [53] for the secure monitoring of ground activities. The authors consider multi-objective global optimisation by incorporating the maximum covering location and backup coverage location problems to solve the deployment problem. In [54], the optimal placement and configuration problems for directional sensing nodes are investigated. An integer linear programming model is developed and the optimal range, FoV and orientation parameters of sensing nodes are determined. The model finds optimal locations for sensing nodes and base stations leading to a reduced network cost while guaranteeing desired coverage and connectivity requirements.

Authors in [55] considered energy balancing among different visual sensing nodes for network lifetime optimisation. Considering routing models with a Rayleigh fading channel, the authors observed that the amount and locations of visual sensing and relay nodes affect network optimisation. Therefore, assuming a 2D environment, a scheme has been proposed which employs a location-wise strategy for pre-determined deployment of sensing nodes attaining energy balancing and leading to optimised network lifetime. The two-tier deployment problem in large scale VSNs is considered in [56]. The VSN architecture comprises of tier-1 with visual sensing nodes and tier-2 with relay nodes. Visual sensing nodes are deployed with uniform random distribution to reduce the VSN cost, whereas, relay nodes are deployed with two dimensional Gaussian distribution to tackle energy-holes. The deployment strategy minimises the network deployment cost subsequently resulting in an improved network lifetime.

An optimisation problem is formulated in [57] to find optimal number of cameras required to provide sufficient coverage of the region of interest. Heterogeneity is considered by incorporating both directional and omnidirectional cameras to reduce the overall information processing cost. The authors concluded that omnidirectional cameras are suitable for providing large coverage, however, they are

inefficient for object recognition due to their limited resolution. On the other hand, directional cameras have limited sensing range but they can be utilised for moving object recognition and tracking. In order to observe a given floor plan, Erdem *et al.* [58] proposed a scheme to automatically compute the camera layout while satisfying the task-specific requirements. Considering polynomial regions of interest, a radial sweep algorithm is developed to estimate each camera's visible region. A grid-based methodology is applied for the representation of the feasible regions and the regions that require coverage. The solution is obtained for a discrete problem space and it results in total cost minimisation. In [59], an algorithm is presented to maximise the angle coverage in VSNs. The algorithm aims to satisfy the desired resolution requirement by identifying a set of sensors that capture the object from all view angles. The energy cost is minimised while preserving a wide view angle by transmitting only those images that satisfy the given resolution requirement. An unsupervised neural network based method to obtain optimal settings for PTZ capable cameras is proposed in [60]. The method utilises a growing neural gas model to maximise scene coverage in the presence of moving objects. Karakaya *et al.* [61] proposed a closed-form solution for visual sensing node's coverage estimation. A certainty-based target detection model is developed to support the formulation of a closed-form solution. Considering a crowded sensing environment, the proposed coverage estimation model provides an estimate of the minimum number of sensors required to guarantee visual K-coverage. The approach also considered the visual occlusions and boundary effect. In [62], the deployment problem of PTZ capable visual sensing nodes is considered. The coverage zone of a PTZ capable visual sensing node is classified into direct coverage and PTZ coverage regions. Time-awareness is incorporated in the deployment strategy to support real-time applications. The proposed strategy formulates an optimisation problem to minimise the visual sensing node deployment cost while satisfying the coverage objectives in time constrained scenarios.

Observations: Many existing solutions to visual sensing node coverage and deployment problems mostly consider the span angle coverage and do not take into account the sensing range coverage. Furthermore the existing schemes mostly consider a 2D FoV to simplify the coverage and placement problems. In contrast, the energy efficient solutions for coverage estimation presented in this thesis consider 3D FoV and sensing range coverage within a 3D sensing environment, thus making the proposed solutions more realistic for VSN applications. Although schemes are found in literature that consider 3D modelling to define a visual sensing node's FoV; however, to the best of the author's knowledge, they do not adopt heterogeneity in the targeted QoI for optimisation.

2.2.3 Visual Data Processing

The perception of information by other sensing nodes differs from that of a visual sensing node. Measurements can be obtained from most sensing nodes in the form of 1D data; however, a visual sensing node provides the information in the form of 2D data sets i.e. images. The rich information content in VSNs results in computationally intense signal processing [2], consequently imposing constraints on energy consumption within visual sensing nodes. Once the image data is captured by a visual sensing node in a VSN, it is processed on board to reduce the amount of data which is required to be transmitted through the network. Local processing can be simple image processing techniques such as edge detection, motion detection, thresholding, background subtraction etc. Furthermore, it can also involve complex image processing techniques such as feature extraction, object classification etc. Therefore, the visual sensing nodes may have different levels of intelligence depending on the complexity of the algorithms being used. Since the camera nodes are battery-powered, the lifetime of a visual sensing node is dependent on its energy consumption. In the case of large amount of visual data, transmission consumes significant amount of energy and bandwidth [63]. Hence,

efficient processing techniques should be employed in visual sensing nodes to enhance the energy and bandwidth efficiency of VSNs.

i) Feature extraction

Feature extraction is one of the most popular methods used for visual data processing in VSNs [43]. A feature is a function of one or more measurements, computed so that it quantifies some significant characteristic of the object [64]. Feature extraction is an area of image processing that deals with algorithms to extract various portions and descriptors from a still image or a video sequence. Since the useful information is limited, feature extraction methods transform the input data into reduced representation using a small set of features [65–71]. They reduce the bandwidth and redundancy of the data with minimum loss of information to provide relevant set of features which makes the decision process easier [66]. For a still image, the description of its content can be obtained from its visual features which play an important role in many image processing applications. These features are required to solve computational tasks which vary from one application to another [72]. There are a number of visual features which can be found in images; however, the selection of features required in the decision making depends on the given application. Features can be classified into low-level and high-level categories.

Low-level features are basic features that can be extracted directly from an image without requiring any information about its content [73]. They can be further subdivided into local features and global features. Local features represent the characteristics of a particular region of the image/object under consideration. They are robust to occlusion, clutter and image variations [63, 74]. Local features include colour, corner, edge, blob, texture, region/patch and spatial location. Within the context of object class recognition, a comparative analysis of various local features is presented in [75]. Skin tone is also a type of local feature which is used for face

detection in [76]. Local features are also used in Content-based Image Retrieval (CBIR) [77]. However, there are several challenges faced during the utilisation of local features. Features such as corners and edges can be significantly affected by noisy environment. Pixel intensity or colour features do not show good level of variance to illumination. Texture-based features are sometimes variant to scale and rotation. Region-based features can be affected by image blur. In contrast to the local features, the global features describe the image as a whole and they are represented by a single vector. Mean, histogram, variance and moment are global features [78]. These features are inefficient for occlusion and clutter [74] and are also sensitive to image variations [63].

Although there are many high level features, a brief description of shape features is provided. Shape features help to find shapes in images and they are represented by the shape descriptor. Shape descriptors should be as complete as possible to represent the content of information, should be stored compactly and should not be computationally expensive [72]. Shape description techniques can be classified into contour-based methods and region-based methods. In contour-based methods, the shape features are extracted from the contour only whereas; region-based methods extract shape features from the whole shape region. They are further subdivided into structural and global approaches. In structural approach, the shape is represented by segments; whereas, the shape is represented as a whole in global approaches. Simple global shape descriptors are area, eccentricity, circularity ratio, bending energy and major axis orientation. However, these features can only be used to distinguish between shapes having large differences. Therefore, they are usually combined with other shape descriptors or used as a filter to reduce false detections [72]. An example of shape-based features is human facial features extraction. The major facial features are the eyes, ears, nose, eyebrows, lips and mouth [72, 79]. These features can be found using their shapes that have some well-known properties, for example the white part of eyes is ellipsoidal, the appearance of mouth and the eyebrows resembles two lines. Moreover, the arrangement of

facial features is also helpful, for example the eyebrows are located above the eyes, the nose can be found below the eyes etc [73].

Within the context of VSNs, there are several studies in the literature that utilise features during the visual data processing phase to support different applications and achieve various objectives. In [80], the design of a multi-model video sensing node is presented to support surveillance in sensor networks with low power and cost. For human detection application, a PIR sensor activates the video sensing node to capture an image frame. The region-of-interest is isolated using a background subtraction scheme and a feature vector is extracted. It is observed from the analysis that feature extraction reduces the size of the data significantly. Multi-model event detection is investigated in [81] for environmental monitoring applications. Colour, histogram and edge features are extracted in the processing phase and a system is developed for event detection. Xie *et al.* [82] provided solutions for the design and implementation problems of a wireless camera network while considering object detection and recognition. The proposed solution provides a trade-off between computational complexity and recognition accuracy by utilising colour histogram-based features with Scale Invariant Feature Transform (SIFT) descriptors. In [83], the difficulties in resource constrained VSNs for a human detection application are investigated. A method is developed for robust human detection which comprises of two modules, a foreground segmentation module and a detection module. Gradient features are extracted to represent the characteristics of the human shape and performance analysis demonstrates that the solution is suitable for energy constrained visual sensing nodes. A visual sensing based abnormal event detection algorithm is presented in [84] for health care applications. The designed algorithm uses object extraction and classification techniques to distinguish normal and abnormal activities. Shape feature variations are observed within visual sensing nodes to detect abnormal events. A shadow removal algorithm is also developed to improve the detection accuracy. An energy efficient face detection method is proposed in [85] for VSNs which reduces the

processing and transition costs. An energy-aware strategy is employed for face detection and only the detected facial information is sent to the base station to enhance the lifetime of VSNs.

Eriksson *et al.* [86] considered the completion time minimisation problem for feature detection and extraction in VSNs by distributing the processing tasks among sensing nodes. A multi-objective optimisation problem is formulated and a regression scheme is employed to support prediction of interest point distribution. Results demonstrate that the proposed approach provides good performance while minimising the computational complexity. Another distributed visual processing strategy is considered in [87]. Visual features are extracted by the sensing nodes and a distributed source coding scheme is developed to optimise the visual content prior to transmission based on the overlapping views. A visual feature extractor, BRISKOLA (Binary Robust Invariant Scalable Keypoints Optimized for Low-power ARM architectures), is proposed in [88] by optimising BRISK [89] for ARM architectures. The BRISKOLA feature extractor provides a solution to minimise the energy consumption without compromising on the performance. Resource allocation problem of VSNs is considered in [90] for visual analysis applications. In the designed solution, each node extracts local features in pixel-domain and apply lossy coding. An optimisation problem is formulated which aims to maximise VSN lifetime while considering the desired target accuracy along with the energy and bandwidth constraints.

Observations: As discussed previously, many techniques are found in the existing literature for feature detection and extraction [80, 83, 84, 86–88, 90]. However, due to strict constraints on available energy and limited hardware capabilities of visual sensing nodes, schemes that are less computationally complex and support faster implementation are desirable. Therefore, colour features are considered in this research study for utilisation in visual sensing nodes due to the simplicity of implementation, speed of feature detection/extraction and requirement of low storage space [91, 92], as in [81, 82]. Moreover, colour features show robustness in

case of rotation and scaling. In event-driven surveillance applications, the process of colour feature detection is followed by the extraction phase, where the object of interest is segmented from the captured image.

Segmentation is a key step while performing image analysis as it facilitates image understanding by providing object feature measurements. Segmentation techniques found in the literature can be classified into unsupervised, semi-supervised and supervised categories. In unsupervised techniques, an image is autonomously segmented into meaningful homogeneous regions, thus no human intervention is required [93]. Local region merging and splitting along with clustering were considered in the early unsupervised segmentation techniques. In contrast, the recent unsupervised segmentation techniques aim to optimise a particular global criteria [94–96]. In semi-supervised segmentation techniques, the un-annotated visual data is segmented by utilising the coarse annotations [93]. Such techniques can be further sub-divided into interactive and cosegmentation techniques [93]. Interactive segmentation techniques [97, 98] are driven by the user input and they are applied in commercial products (i.e. Adobe Photoshop and Microsoft Office). Recurring objects can be extracted from a set of images using the cosegmentation techniques [99, 100]. In supervised segmentation techniques [101, 102], an application-specific model for the object-of-interest is developed, therefore, these techniques achieve superior performance for particular given tasks [93]. In this research study, the segmentation approach presented in [103] is employed due to its computational efficiency. Moreover, the chosen segmentation approach is suitable for utilisation with the colour feature extraction technique considered to develop the proposed solutions.

ii) Quality-of-Information (QoI)

Many research studies have considered QoI but in different contexts. One of the earlier investigations on QoI for sensor networks is presented in [104]. Event detection in sensor networks is considered and application efficiency is modelled with QoI. Timeliness of the data, confidence on the data, characteristics of the sensor networks and events are considered as QoI attributes. Relationships are established between these attributes and their theoretical performance limits are obtained. In [105], a dual layer model based on QoI and Value-of-Information (VoI) is developed. QoI attributes of timeliness, accuracy and latency are considered; whereas, for a given QoI, VoI is characterised by the importance for a specific use. A VoI attributes-based framework for information scoring and ranking is presented by utilising a multi-criteria decision making process. A methodology for sensor network data models processing is proposed in [106]. The data models aim to combine data from sensing nodes while satisfying the QoI requirements for energy conservation and WSN lifetime maximisation. Several research studies are found in the literature that examine QoI for task management. Liu *et al.* [107] address the problem of dynamic multi-task management in WSNs and incorporate QoI-awareness in their solution. The solution provides QoI index of a task to gauge the satisfaction of requested QoI levels within a WSN. Network capacity in terms of QoI is modelled to evaluate the WSN's new task admission ability without affecting the QoI satisfaction of other tasks. An admission control strategy is developed for optimal utilisation of network resources resulting in optimised accommodation of tasks' QoI requirements. The performance of this approach is evaluated under the intruder detection scenario and the results demonstrate that it leads to lifetime enhancement. The task management solution presented in [107] is extended in [108] by introducing an energy management scheme. A distributed QoI-aware approach is employed to develop the energy management scheme which determines the participating state of sensing nodes. The approach exploits the trade-off between sensing nodes' energy management and QoI support resulting in energy savings.

In [25], an Internet-of-Things sensory environment is considered and an energy efficient multi-task oriented framework is proposed to manage and control the duty cycles of the sensors under particular QoI requirements. The performance of the framework is observed for an environmental monitoring application. Few schemes exploiting QoI for visual data are also present in the literature. In [26], the accuracy of representing an event at the output of a sensing node is observed as QoI. Practical examples are given and experiments are conducted to describe, estimate and evaluate the QoI. The versatility of a PSNR-based metric is demonstrated in terms of representing the QoI. Multi-sensor multimedia monitoring system is considered in [109] and a model is developed by characterising QoI attributes of certainty, accuracy and timeliness. A fusion approach is employed to determine the target information and QoI attributes are dynamically calculated depending on the participating sensors' observations. The suitability of this mechanism is demonstrated for detection and human identification.

Observations: Although many research studies have incorporated QoI to enhance an application's efficiency; however, most of the existing schemes are developed for WSNs, which are not suitable for VSNs due to the multi-dimensional nature of visual data. Therefore, a PSNR-based metric is developed in this thesis for the characterisation of QoI to reflect the quality of the data obtained from a visual sensing node. To the best of the author's knowledge, the existing schemes found in literature that define QoI for VSNs do not consider dynamic targeted threshold based optimisation along with the utilisation of visual sensing nodes' 3D FoV information. Moreover, a dynamic realisation of QoI is considered in this thesis to facilitate reconfiguration that can result in substantial energy savings.

iii) Pre-transmission visual data optimisation

Visual data aggregation is a processing technique which exploits the similarities between neighbouring visual sensing nodes' data to eliminate redundancies resulting in transmission energy conservation and efficient bandwidth utilisation. The energy consumption trade-off between image processing and aggregation within a visual sensing node is investigated in [110]. An energy consumption model for image aggregation is developed which incorporates image processing and transmission costs. The study also provides a solution to find the optimal number of nodes to be involved in aggregation resulting in an improved network lifetime. An entropy-based framework is developed in [111] for visual data aggregations in sensor networks. In the pre-transmission phase, clusters are formed by grouping sensing nodes with high joint entropy together. The framework tackles the problem of single image repeated transmissions by removing redundant data from correlated sensing nodes. In [112], a Wireless Multimedia Sensor Network is considered which consists of camera and scalar nodes. A distributed scheme is proposed for topology management and camera actuation. A packet aggregation strategy is employed within the network prior to reporting an event to a particular camera node. The approach leads to enhanced scalar nodes' lifetime, improved event coverage, increased packet delivery ratio and reduced latency. Multi-view video coding is considered in [113] to exploit the spatial correlation between partially overlapping FoVs of multiple cameras observing a particular scene of interest. A metric is used to represent the degree of correlation between different views of a scene of interest. The approach leads to a reduction in total bandwidth consumption. Authors in [114] employed a 3D sensing model for 3D visual correlation in resource constrained VNSs. The method enhances the energy efficiency by minimising the processing required for 3D visual correlation.

Image compression is one of the most widely utilised approaches in visual data optimisation prior to transmission in VSNs. Efficient image compression algorithms can reduce the amount of visual data for transmission to the sink node resulting

in network lifetime maximisation. Comprehensive surveys on image compression schemes for VSNs can be found in [115–117]. An energy efficient architecture for image processing and a protocol for communication in VSNs are proposed in [28]. Employing the proposed object detection architecture with Discrete Wavelet Transform (DWT) processing, the proposed approach leads to a reduction in image transmission cost. In [118], an energy efficient image compressive transmission scheme is proposed for resource constrained scenarios. The proposed scheme incorporates region of interest extraction with block-based compressive sensing to devise an energy-driven strategy for image quality control. Chow *et al.* [29] proposed an energy efficient framework for on-demand image transmission in visual sensor networks. Intra-neighbourhood processing and combining protocol is developed for overlapping regions. Subsequently, a particular level of compression is employed and images are transmitted upon request to the mobile sink resulting in reduced energy consumption. In [119], a computationally less expensive approach based on change detection is proposed for image compression in resource constrained sensor networks. Region of interest is determined by the change detection algorithm for visual data optimisation. Fast Discrete Cosine Transform (DCT) with optimised quantisation is used for compression. The scheme provides a trade-off between energy consumption and resulting image quality. Another approach based on region of interest is presented in [120]. An image compression algorithm is developed which employs a low compression level for the region of interest and a high compression level for the rest of the image. Thus, by reducing visual data, the algorithm minimises the transmission energy cost.

Wang *et al.* [121] developed an information theoretic generalised image compression framework which is independent of the image statistics and coding algorithms. The framework employs an entropy-based divergence measure scheme and a distributed multi-cluster coding protocol to maximise visual data compression in VSNs. An energy and bandwidth constrained wireless video sensor network is considered in [122] and a compressive sensing based image compression scheme is

proposed. The proposed scheme adaptively controls the image quality depending on the the sensing nodes' remaining energy and the network's link quality. In [123], an architecture for landslide monitoring utilising heterogeneous sensor networks is proposed. The sensor network considered in [123] consists of geographical and camera sensor nodes. Strategies for threshold self-learning, information fusion and sleep scheduling are utilised for anomaly detection. After the confirmation of a slope anomaly, the camera nodes are activated, object detection is performed and image compressive transmission takes place.

Several other approaches are presented in the literature for optimised transmission in VSNs. An energy efficient image prioritisation framework is proposed in [124] to select relevant information before transmission to the sink node. Based on salient motion detection, the proposed approach reduces the transmission cost of the visual information. A VSN lifetime maximisation strategy is proposed in [125] that optimises the source rates, encoding powers and routing schemes to prolong the network's lifetime. Ye *et al.* [126] proposed an energy-aware packet interleaving scheme for robust transmissions within VSNs to improve the end-to-end image transmission quality and prolong the network lifetime. In [127], Dai *et al.* proposed a routing algorithm by integrating correlation-aware inter-node differential coding and load balancing schemes. The proposed approach minimises the sensor network's energy consumption under certain constraints. Authors in [128] proposed a framework for image communication in bandwidth constrained WSNs. An image compression scheme utilising the set partitioning in hierarchical trees (SPIHT) technique is incorporated within the framework to enhance the energy efficiency of the network. Moreover, a protocol to reduce the packet drop rates is employed for reliable communication. In [129], a framework for video encoding and wireless transmission in VSNs is proposed which improves the energy efficiency of the network and results in lifetime enhancement. The transmission within the framework is supported by a reliable multi-priority routing protocol that considers the nodes' residual energy, the packet loss rates and the congestion

level.

Observations: Although, many research studies are present in the literature for visual data optimisation in the pre-transmission phase but most of them utilise a particular static level of compression for optimisation. In contrast, dynamic compression levels is considered for feature redundancy removal in this thesis to optimise the visual data, satisfy the criteria of desired reliability and enhance the energy efficiency of VSNs. To the best of the author's knowledge, the existing studies that adopt dynamic compression strategies do not consider heterogeneous levels of target QoI thresholds.

2.2.4 Task classification

Owing to the complexity of sensing and processing tasks within VSNs, tasks classification is an active area of research. A number of studies are found in the literature that investigate the task allocation problem. Dieber *et al.* [130] proposed a centralised resource-aware evolutionary algorithm that addresses coverage and task assignment problems in resource constrained VSNs. The proposed solution provides optimal configuration for cameras selection, frame rate, resolution and assignment of tasks. In [131], object tracking application for complex scenarios is considered and a dynamic task decomposition strategy is presented. A vision-based state estimation problem is formulated for maximised sharing of available resources. A task is divided into different subtasks which results in enhanced parallelisation, throughput and reliability. An optimisation problem is formulated in [132] and a distributed load allocation algorithm is developed to minimise the completion time of visual processing tasks. As compared to other schemes, the algorithm offers fast convergence at the cost of higher completion times. However, utilising a centralised coordinator, the algorithm can result in low completion times with reduced computational complexity.

Task allocation problem of VSNs is also discussed in [133]. A computational market-based approach is employed and a method is developed to maintain QoS levels of the tasks while considering the resource requirements and the desired level of service. Another market-based solution for the adaptive energy-aware assignment of tasks to cameras in resource constrained networks is proposed in [134]. An adaptive utility function is used by the cameras to bid for tasks which results in network lifetime maximisation. Authors in [135] investigated the task assignment problem in camera networks for surveillance applications. A model is formulated to address this problem and promising results are obtained in terms of priorities assignment, idle agents avoidance and waiting time reduction.

A multi-object tracking problem is discussed in [136] by employing a game theoretical approach. Energy levels, processing loads and nodes' accuracies are considered for reassigning object tracking tasks. Various scenarios are considered to evaluate the performance of the proposed solution which demonstrates that it leads to lifetime enhancement while reducing the communication between the nodes. In [137], the popular matching strategy is introduced to solve the multi-task assignment problem in camera networks. The approach allocates tasks to cameras based on the closest task preferences. A framework for in-network processing is proposed in [138] which utilises a speedy task reallocation strategy to recover from service disruptions that occur within the network. A self-learning algorithm is developed which dynamically maintains system parameters to achieve the desired criteria resulting in network lifetime improvement.

Observations: To the best of the author's knowledge, the existing solutions for task classification do not jointly consider heterogeneous visual sensing nodes, their 3D coverage and energy efficient FoV characterisation criteria for optimal task allocation in resource constrained scenarios. Hence, there is scope for a coverage-oriented generalised solution which can be implemented in heterogeneous visual sensing nodes to support a diverse range of applications. The task classification scheme developed in this thesis provides a generalised solution by exploiting the

sensing nodes' heterogeneity in an intelligent manner and it is suitable for a broad range of applications.

2.2.5 Dynamic reconfiguration

Many solutions for WSN design have been proposed in the literature that facilitate reconfiguration to optimise network performance [139–141]. As the nodes within WSNs employ omni-directional sensing to collect scalar data and due to the directional sensing nature of VSNs, the reconfiguration schemes for WSNs cannot be directly implemented in VSNs. Moreover, the acquisition, processing, transmission and reception of 2D image data impose stricter constraints on energy consumption and bandwidth within VSNs. Due to these fundamental differences, compared to WSNs, obtaining solutions to the reconfiguration problem for VSNs is much more challenging. Owing to these challenges, researchers have been actively engaged in developing dynamic reconfiguration schemes for VSNs to achieve various objectives. The existing solutions for dynamic reconfiguration can be classified into resource-aware, coverage-oriented and target-based methods [22], which are discussed in the following sections.

i) Resource-aware methods

Resource-aware methods support reconfiguration by considering the available and/or required resources on each node while satisfying the desired performance requirements [22]. In order to achieve a given objective, several heuristics for energy-aware resource allocation are discussed in [142]. The reconfiguration aims to guarantee energy minimisation as well as timely task completion. Dynamic power management incorporates strategies to dynamically regulate the state of system components depending on the workloads. An artificial neural network is utilised in [143, 144] and the reconfiguration workloads are obtained. Multi-camera traffic surveillance application is considered and the developed solutions result in optimal

selection of timeouts for each component to reduce the power consumption while ensuring the desired performance levels. A stochastic model is formulated in [145] for the network operational lifetime. The solution presented by the authors consider camera selection and energy distribution among selected sensing nodes for reconfiguration. The developed camera scheduling and energy allocation strategies lead to lifetime maximisation in VSNs.

Karuppiah *et al.* [146] developed a context-aware reconfiguration framework based on hierarchical smart resource coordination. The utilised approach comprises of a fault containment unit which offers dynamic fault tolerance in case of failures. The effectiveness of the approach is evaluated for a tracking application and it is observed that the developed system supports tracking by dynamically selecting camera pairs for optimal object localisation. In [130], a distributed approach to dynamically find optimal configuration of a network is developed which provides a trade-off between resource consumption and surveillance quality. The distributed approach is compared with its centralised counterpart to demonstrate effectiveness of the proposed approach. The distributed approach of [130] is coupled with a market-based strategy for tracking assignment in [147]. Fu *et al.* [148] proposed a tracking scheme which balances the tracking accuracy and energy consumption in wireless camera sensor networks. A decentralised approach is employed for tracking to enhance the energy efficiency of the target tracking scheme. Moreover, a cluster head selection strategy is proposed which formulates an optimisation problem for such selection depending on the residual energy and the sensor-to-target distance. Kim *et al.* [149] proposed an energy efficient management scheme to maximise the data quality and lifetime of solar-powered VSNs. The energy supply and demand is predicted to support the energy management scheme. Another energy management solution for VSNs is presented in [150]. A strategy is developed by formulating an optimisation problem for the minimisation of sensing nodes' power consumption. The process of reconfiguration takes place for optimal node scheduling and power allocation.

ii) Coverage-oriented methods

Global coverage optimisation for a particular scene of interest is the main aim of coverage-oriented methods which can be achieved by either increasing the number of sensing nodes providing coverage to specific regions or reducing the number of unobserved regions in the sensing environment [22]. Authors in [151] studied the dynamic coverage optimisation problem of directional sensors. The configuration space comprises of the position and orientation parameters of the sensing nodes, whereas the objective of reconfiguration is to maximise the targets' coverage with minimum possible sensing nodes. Adaptive scheduling is incorporated by introducing a sensing neighborhood cooperative sleeping protocol. Both centralised and distributed solutions are presented by the authors. Morsly *et al.* [152] considered the problem of providing reliable visual coverage of the sensing environment by minimum possible cameras. The sensors' position and orientation parameters are considered within the configuration space. The sensors' capabilities are modelled as constraints to formulate an optimisation problem which reduces the resource requirements. Another coverage maximisation algorithm is presented in [153] which takes into account the user defined priority regions. The configuration space considers pan, tilt and zoom level parameters. A centralised genetic algorithm is utilised in the reconfiguration process to find optimal coverage.

In [154], an occlusion-aware method is proposed to dynamically obtain the optimal configuration of cameras' pan, tilt and zoom parameters in the network for coverage maximisation. A multi-purpose approach to the reconfiguration of PTZ capable cameras within a network is proposed in [155]. The authors considered the sensors' orientation and zoom levels to define the configuration space with coverage and image quality maximisation objectives. An enhanced approach incorporating the solutions of [155] is presented in [156]. An optimisation technique based on particle swarm optimiser is proposed to obtain the most suitable camera position and configuration while achieving a given coverage objective. The proposed solution takes into account both the coverage and visual quality of the

acquired visual data, and it is capable of reconfiguration in case of cameras failures. Reconfiguration problem to guarantee barrier coverage is studied in [157]. Distributed methodologies are presented for self-configuration of mobile sensing nodes to form sensing node barriers over a rectangular region.

iii) Target-based methods

The aim of target-based methods is to reconfigure the sensing nodes focusing on a specific target. A typical application of these methods is target tracking which aims to follow the detected objects by reconfiguring PTZ capable sensing nodes. In case of static sensing nodes, reconfiguration can be employed to select the most suitable sensing nodes for acquiring good quality images of specific targets. In order to obtain high quality visual data representing a target, a master-slave approach is employed in [158] with PTZ cameras, where each camera within the network can function as master or slave. In the proposed strategy, the master camera observes a large portion of the sensing environment and estimates approximate location information for the slave cameras. An uncalibrated method is incorporated in the reconfiguration phase and time-variant homographes are estimated between camera views. The approach does not require any pre-defined 3D location information. The limitations of the master-slave approach are addressed in [159, 160] by developing coordination schemes for networks comprising PTZ capable multi-cameras. The schemes support multiple targets tracking by acquiring high quality visual data and works for occluded targets as well. Ding *et al.* [161] developed a distributed method to dynamically find optimal pan, tilt, zoom parameters for reconfiguration and system performance maximisation by acquiring high quality images of specific targets. The method provides a trade-off between image quality and tracking failure probability.

Solutions for target tracking problem with PTZ capable cameras are also presented in [162, 163]. The aim of dynamic reconfiguration is to minimise the number of

camera hand-offs as well as optimising the visual data quality. The strategy for camera tasking optimisation is supported by moving objects trajectories prediction. Authors in [164] adopted a negotiation-based approach for camera handover coordination. A status sharing strategy is developed which can be utilised in the cameras. Bidding takes place among the cameras where objects for sale are offered by the bidding camera to free allocated resources. In [165], a coalition-based framework for multi-target tracking in bandwidth constrained camera networks is proposed. Local interactions, tracking confidence and neighbourhood communication performance are considered to form coalition among cameras for performing distributed target tracking tasks. The approach leads to accurate tracking with reduced communication cost.

Observations: Despite the fact that a number of state-of-the-art techniques are found in the existing literature, most of them are developed for camera networks without considering a resource constrained scenario. Even though there are various resource-aware solutions but they consider the remaining or required energy resources and do not incorporate QoI-awareness in the process of reconfiguration. Furthermore, centralised approach for reconfiguration is employed in many schemes which is not suitable for bandwidth constrained VSNs. To the best of the author's knowledge, the existing techniques that utilise a distributed reconfiguration approach do not exploit heterogeneity of the target QoI thresholds. Hence, there is scope for a distributed self-reconfiguration scheme to be developed for visual sensing nodes which takes into account both resource and QoI requirements based on the application design criteria.

The notable existing schemes relevant to this research study that are found in the literature and discussed earlier in this chapter are listed in Table 2.1.

TABLE 2.1: Classification of some notable schemes found in literature within the context of the problem domain

Classification	**Methods**
Visual sensing node coverage and deployment	[50], [51], [52], [53], [54], [55], [56], [57], [58], [59], [60], [61], [62]
Feature extraction	[80], [81], [82], [83], [84], [85], [86], [87], [88], [90]
Quality-of-information	[25], [26], [104], [105], [106], [107], [108], [109]
Pre-transmission visual data optimisation	[28], [29], [110], [111], [112], [113], [114], [118], [119], [120], [121], [122], [123], [124], [125], [126], [127], [128], [129]
Task classification	[130], [131], [132], [133], [134], [135], [136], [137], [138]
Resource-aware reconfiguration	[142], [143], [144], [145], [146], [147], [148], [149], [150]
Coverage-oriented reconfiguration	[151], [152], [153], [154], [155], [156], [157]
Target-based reconfiguration	[158], [159], [160], [161], [162], [163], [164], [165]

2.3 Summary

The existing state-of-the-art techniques for VSN design and optimisation have been discussed in this chapter. The architecture of a visual sensing node is presented which comprises of an image capturing device, a processing unit, a transceiver and a power unit. Each component plays a prominent role in VSN applications. The coverage and deployment of visual sensing nodes depend on the characteristics of the image capturing device. The deployment algorithms for visual sensing nodes are classified into two categories - deterministic deployment and random deployment. Coverage estimation is a challenging task in VSNs due to the directional nature of visual sensing nodes. Therefore, the existing solutions for a traditional sensing node's coverage estimation cannot be implemented in a visual sensing node. Various coverage and placement algorithms for visual sensing nodes have been discussed in this chapter. However, it is found that many existing solutions for coverage estimation assume a simplified 2D FoV model, whereas, a 3D model is more realistic for VSN applications.

Visual data acquired by a sensing node can be processed to reduce the amount of data prior to transmission in VSNs. Several techniques are discussed for visual data processing and optimisation such as feature extraction, QoI utilisation, data aggregation and compression. QoI-awareness has been incorporated in several algorithms to enhance the network's intelligence and improve the management. However, most of the existing QoI-aware techniques are developed for WSNs and they may not be suitable for VSNs due to the multi-dimensional nature of visual data. It is observed that there is scope for a QoI index to be developed for VSNs characterised by a PSNR-based metric. Moreover, dynamic QoI realisation is required for optimisation. The existing schemes for visual data optimisation in the pre-transmission phase employ various compression schemes to reduce the amount of redundancy. However, many schemes employ a static level of compression and

do not take into account the dynamic nature of the visual data. Therefore, a dynamic compression level is considered for visual data optimisation in this thesis.

Task classification is an active area of research due to the complexity of sensing and processing tasks within VSNs. Several task classification algorithms have been discussed in this chapter. However, most of the existing solutions for task classification do not jointly consider heterogeneous visual sensing nodes and their 3D coverage for optimal task allocation in resource constrained scenarios. It is observed that there is scope for a generalised coverage-oriented task classification algorithm for heterogeneous visual sensing nodes to support a broad spectrum of applications. VSNs can be reconfigured to select optimal parameters while satisfying particular objectives. Resource-aware, coverage-oriented and target-based methods for dynamic reconfiguration have been discussed. It is observed that most of the existing reconfiguration schemes are developed for camera networks which may not be suitable for resource constrained scenarios. Furthermore, the schemes that take into account the resource requirements did not incorporate QoI-awareness to support reconfiguration. In addition, the centralised nature and suitability only for homogeneous sensing nodes limit the applications of some existing solutions. Hence, it is observed that there is scope for a distributed approach for reconfiguration with the joint realisation of resource and QoI requirements. Considering resource constrained VSNs, this research is focused on the development of energy efficient solutions for sensing node characterisation while considering heterogeneity, incorporating QoI-awareness and satisfying particular desired levels of reliability.

The next chapter presents a FoV characterisation framework to be employed by the system design engineers in the network pre-deployment phase for sensing range optimisation. Moreover, an adaptive task classification scheme is proposed to support the FoV characterisation framework in the post-deployment phase leading to substantial energy savings in heterogeneous VSNs.

Chapter 3

FoV Characterisation Framework

3.1 Introduction

The multi-dimensional nature of visual data imposes severe constraints on energy consumption in VSNs. Consequently, minimising energy consumption within the network is considered to be the most important design objective in resource constrained scenarios. Energy efficient solutions for network lifetime maximisation are required to assist the system design engineers in the pre-deployment phase. Such solutions can be obtained by developing a trade-off model for reliability and energy efficiency. Based on the application design criteria, the characterisation of visual sensing nodes' FoV can result in spatial coverage optimisation and lifetime maximisation. Moreover, the heterogeneity of visual sensing nodes can be exploited to distribute sensing and processing tasks among the nodes in an energy efficient manner.

This chapter presents the proposed FoV characterisation framework that provides generalised energy efficient design solutions for both homogeneous and heterogeneous VSNs. The 3D projection model of a visual sensing node is presented and various parameters that affect the projection model are discussed. The proposed

task classification and soft decision based sensing range selection schemes to enhance the energy efficiency of heterogeneous VSNs are described. Experiments and simulations are performed to evaluate the robustness of the proposed FoV characterisation framework and the results are presented in this chapter. Since energy consumption is a critical issue, an energy model is described in order to compute the energy dissipation and to demonstrate the energy efficiency of the proposed framework. In addition, an analysis of system failure probability is presented.

3.2 Visual Sensing Node's 3D Projection Model

The 3D projection model of a visual sensing node VS within a spherical sector is shown in Figure 3.1. VS employs directional sensing and transforms a projection of the 3-dimensional scene from $\mathbb{R}^3$ to a 2-dimensional image plane in $\mathbb{R}^2$.

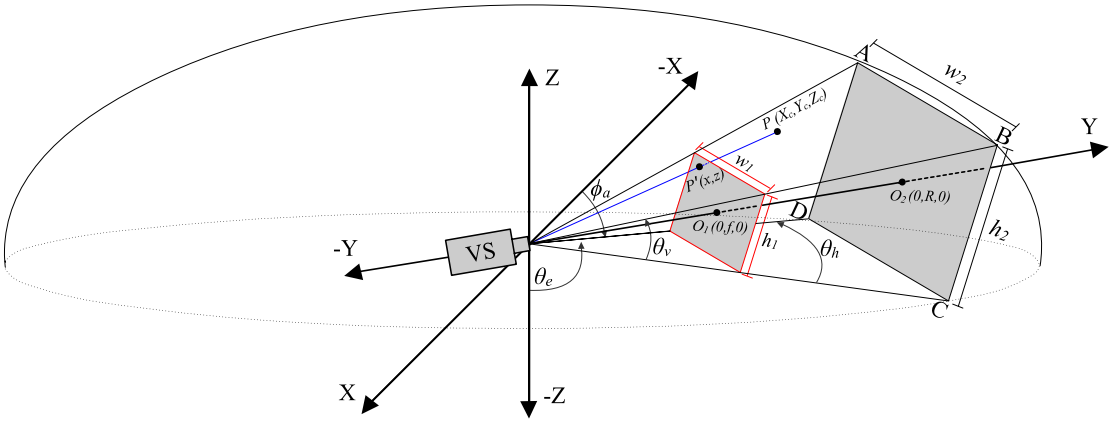

FIGURE 3.1: Visual sensing node's 3D projection model

In this model, the projection of the 3D scene points onto a physical 2D image plane is characterised by the pinhole camera model. In reality, the physical image plane lies inside the visual sensor behind its centre. The light rays hit the image plane through a pinhole and create an upside down image of the scene within the FoV. In order to simplify the mathematical model, it is assumed that the physical

image plane lies infront of the sensor's centre and provides the same image with respect to the scene within the FoV.

In order to describe the projection, Figure 3.2 [166, 167] shows a simplified form of Figure 3.1 where a scene point $P(X_c, Y_c, Z_c)$ projects onto the image plane at a point $P'(x, z)$. In addition, the relationship between the 3D scene point and 2D point on the image plane is illustrated in Figure 3.3 [167].

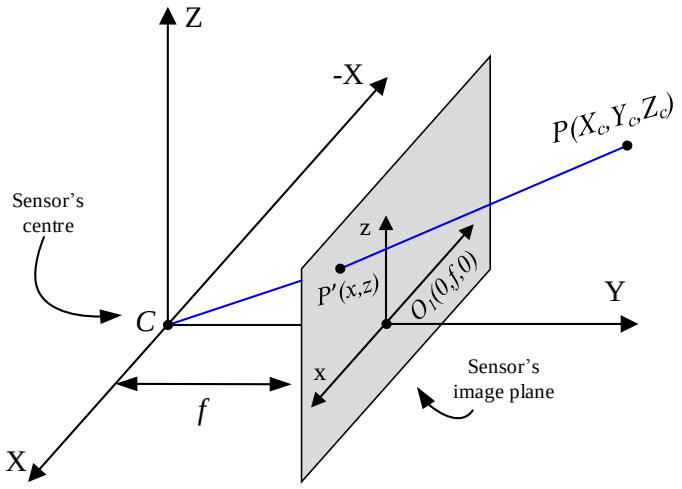

FIGURE 3.2: Sensing process characterised by a pinhole camera model [166,167]

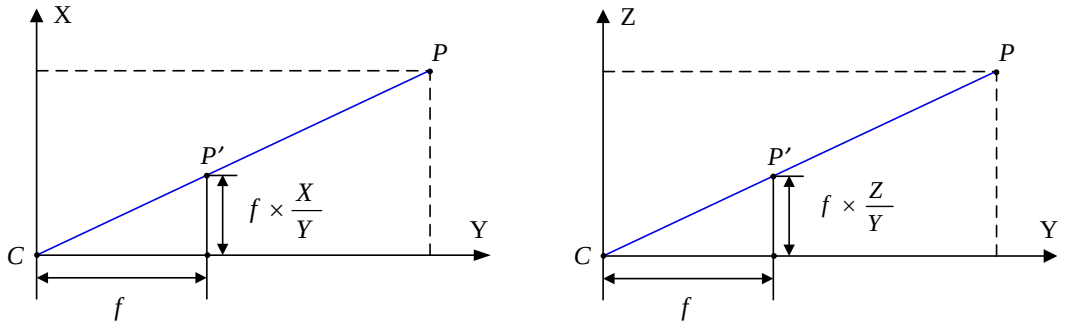

FIGURE 3.3: 2D view of the projection of point P [167]

The point $P'(x, z)$ on the image plane is calculated from Figure 3.3 using triangle similarity as shown in the following equation,

$$P'(x,z) = \begin{cases} x = f \times \dfrac{X_c}{Y_c} \\ \\ z = f \times \dfrac{Z_c}{Y_c} \end{cases} \tag{3.1}$$

The visual sensing node covers a certain part of the spherical area of interest. The region within the sensing node's 3D FoV is described by the horizontal FoV (θ_h) and the vertical FoV (θ_v) of the sensor.

Where θ_h and θ_v are the angular extents of the scene measured horizontally and vertically by the sensor respectively. The FoV of a sensing node is characterised by various parameters which are described in the following sections.

3.2.1 Sensing Node's Location

In Figure 3.1, sensing node VS is located at the origin of the cartesian coordinate system i.e. $(0, 0, 0)$ and the sensor's optical axis overlaps onto the y-axis with $X = 0$ and $Z = 0$. Within the context of a VSN, where $\mathcal{N}$ sensing nodes are present, each sensing node $\text{VS}_{\tilde{l}}$ ($\tilde{l} = \{1, 2, 3, \ldots, \mathcal{N}\}$) is identified by its location which is described by the cartesian coordinates $(X_{\tilde{l}}, Y_{\tilde{l}}, Z_{\tilde{l}})$, azimuth angle ϕ_a and elevation angle θ_e. These parameters define sensing nodes' distribution within the network and are tuned to fit the respective regions of interest within each sensing node's FoV.

3.2.2 Sensing Range and ABCD-plane Dimension

In Figure 3.1, the origin O_2 of the ABCD-plane intersects the y-axis at $(0, R, 0)$; where R is the distance between the visual sensor and the ABCD-plane and is

known as the *sensing range*, w_2 is the width of the ABCD-plane and h_2 is its height. For a target object, the sensing range spans from R_{min} to R_{max} for a certain acceptable level of sharpness.

Varying R affects the sensing node's coverage due to the change in ABCD-plane dimensions. Increasing this distance R results in a larger coverage volume within a sensing node's FoV, whereas, decreasing R reduces the coverage volume. Therefore, R is a key parameter for FoV characterisation.

3.2.3 Image plane Dimension and Sensor's Resolution

As mentioned earlier, a visual sensing node captures a 3D scene onto a physical image plane. The quality of the captured image depends on the dimension of the physical image plane and the sensor's resolution in pixels. In Figure 3.1, the width and height of the physical image plane are represented by w_1 and h_1 respectively; and are typically measured in milli-metres (mm). The physical distance $f \in \mathbb{R}^+$ between the sensor's optical centre and the image plane is known as the *focal length*, typically measured in milli-metres (mm). Each sensor maps $P \times Q$ pixels onto the image plane, where $P \times Q$ is known as the *resolution* of the sensor. Each sensor has a particular pixel size typically measured in micro-metres (μm). A sensor with higher pixel size has larger image plane dimension and results in a better quality image as compared to the sensor with smaller pixel size. Similarly, a sensor with higher resolution provides a better quality image as compared to the one with lower resolution. Sensors that provide a high quality image are capable of observing large regions within their FoVs and result in reduction of the number of sensing nodes required for full coverage. However, such sensors increase the overall network design cost. Therefore, the selection of sensors for VSN design requires careful consideration of all the aforementioned parameters.

3.3 Proposed FoV Characterisation Framework

The proposed FoV characterisation framework to design and calibrate energy efficient VSNs is presented in Figure 3.4 and Figure 3.5. Based on the type of VSN, there are two approaches for its design and calibration: (i) Approach I for homogeneous sensor networks shown in Figure 3.4 (ii) Approach II for heterogeneous sensor networks shown in Figure 3.5. The proposed framework consists of image capture, projection modelling, ABCD-plane modelling, adaptive task classification (for heterogeneous networks), feature detection, extraction and FoV characterisation which are described in the following sections.

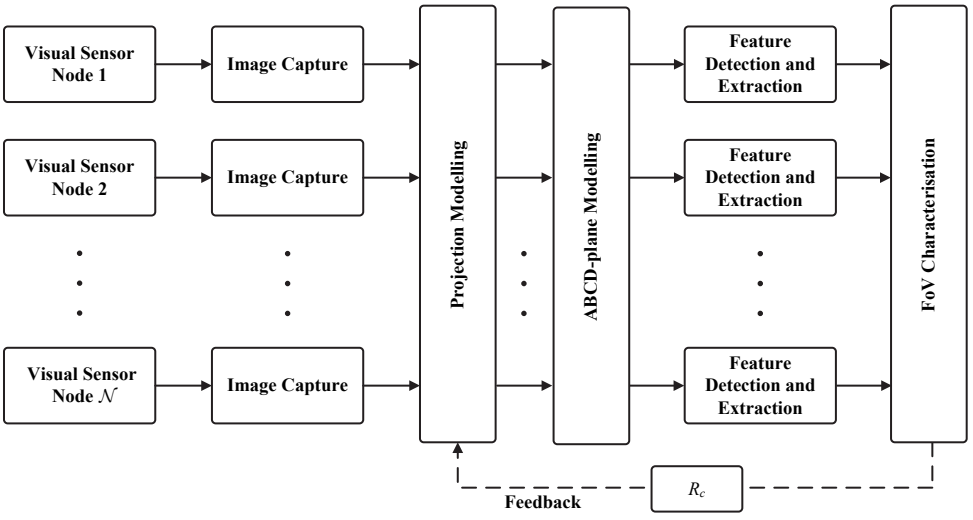

FIGURE 3.4: Proposed FoV Characterisation Framework for Energy Efficient Homogeneous VSNs

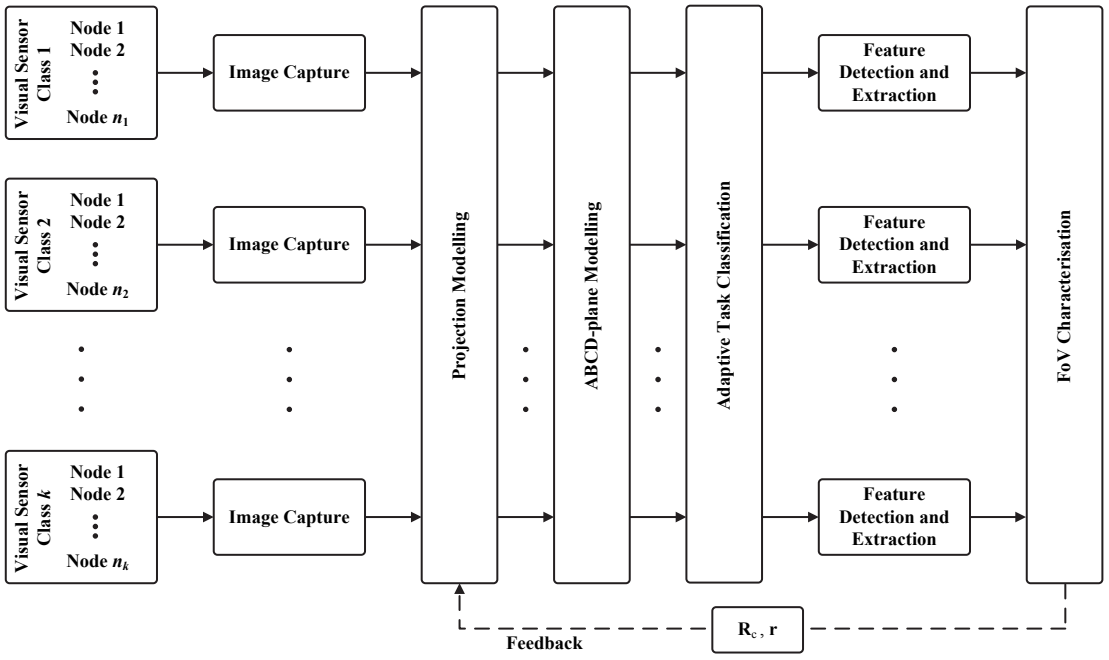

FIGURE 3.5: Proposed FoV Characterisation Framework for Energy Efficient Heterogeneous VSNs

3.3.1 Image Capture

Each sensing node $\text{VS}_{\tilde{l}}$ captures an image $\mathbf{I}_{\tilde{l}}$ of dimension $P \times Q$ which is a function of the following parameters: the distance R, the horizontal FoV θ_h, the vertical FoV θ_v.

$$\mathbf{I}_{\tilde{l}} = f\left(R, \theta_h, \theta_v\right) \tag{3.2}$$

3.3.2 Projection Modelling

As mentioned earlier, a visual sensor projects 3D scene points onto its image plane. Volume V of the scene within the sensing node's FoV projected onto its image plane is given by,

$$V = \frac{4R^3 \sin\theta_h \sin\theta_v}{3(1+\cos\theta_h)(1+\cos\theta_v)} \tag{3.3}$$

In order to characterise sensing node's coverage, FoV (θ_h, θ_v) is required to be known. The projection modelling approach is different for homogeneous and heterogeneous networks, as discussed in the following sections.

i) Homogeneous Networks (Approach I)

A homogeneous sensor network has identical nodes in terms of their sensing parameters and hardware capabilities. Using the following equations [168], FoV (θ_h, θ_v) is calculated with prior knowledge of the following sensor specifications: image plane dimensions (w_1, h_1) and focal length (f).

$$\theta_h = 2\arctan\left(\frac{w_1}{2f}\right) \tag{3.4}$$

$$\theta_v = 2\arctan\left(\frac{h_1}{2f}\right) \tag{3.5}$$

ii) Heterogeneous Networks (Approach II)

A heterogeneous network has two or more types of nodes in terms of their sensing parameters and hardware capabilities. The nodes with lower specifications are less costly and consume less energy. On the other hand, the nodes with higher specifications can perform feature detection and extraction with higher reliability but consume more energy and cost more. Keeping a certain reliability level in a heterogeneous network, few higher specification nodes can be used in each cluster along with the lower specification nodes to reduce the overall network cost.

In the case of using a variety of sensing nodes within the network, any of the following sensor specifications: (w_1, h_1) and (f) may be unknown and FoV (θ_h, θ_v) cannot be calculated through Approach I. For such case, an alternative approach is presented for calculating the FoV of each type of sensing node using the following equations,

$$\theta_h = 2 \arctan\left(\frac{w_2}{2R}\right) \tag{3.6}$$

$$\theta_v = 2 \arctan\left(\frac{h_2}{2R}\right) \tag{3.7}$$

This method requires an experimental setup (described later in Section 3.4) which utilises a known reference distance $R = R_{\text{ref}}$ for FoV calculation.

If θ_h, θ_v and f are known, the dimension of the physical image plane can be calculated using the following equations,

$$w_1 = 2f \tan\left(\frac{\theta_h}{2}\right) \tag{3.8}$$

$$h_1 = 2f \tan\left(\frac{\theta_v}{2}\right) \tag{3.9}$$

If θ_h and w_1 or θ_v and h_1 are known, the focal length f can be calculated using one of the following equations,

$$f = \left(\frac{2}{\theta_h}\right) \arctan\left(\frac{w_1}{2}\right) \tag{3.10}$$

or,

$$f = \left(\frac{2}{\theta_v}\right) \arctan\left(\frac{h_1}{2}\right) \tag{3.11}$$

3.3.3 ABCD-plane Modelling

Once θ_h and θ_v are known, the dimension of the ABCD-plane is calculated for a range of values of R i.e. R_{min} to R_{max} using the following equations,

$$w_2 = 2R \tan\left(\frac{\theta_h}{2}\right) \tag{3.12}$$

$$h_2 = 2R \tan\left(\frac{\theta_v}{2}\right) \tag{3.13}$$

The aspect ratio of the captured image can be calculated using the following equation,

$$\text{Aspect ratio} = \frac{w_1}{h_1} = \frac{w_2}{h_2} = \frac{\tan\left(\frac{\theta_h}{2}\right)}{\tan\left(\frac{\theta_v}{2}\right)} \tag{3.14}$$

3.3.4 Adaptive Task Classification

As discussed earlier, heterogeneous networks comprise of sensing nodes with different capabilities and their performance is better as compared to the homogeneous networks due to the classification of sensing and processing tasks assigned to each visual sensor class based on its sensing capabilities. Adaptive task classification is employed by the FoV characterisation framework to enhance the intelligence of heterogeneous VSNs.

Consider a heterogeneous sensor network with k sensor classes; each sensing node is denoted by $\mathrm{VS}_{l,j}$ such that $l = \{1, 2, 3, \ldots, k\}$ represent the sensor class and $j = \{1, 2, 3, \ldots, n\}$ represent n_l sensing nodes belonging to a sensing class l. Let n denote the maximum number of sensing nodes belonging to a particular sensing class given by $n = \max\{n_l \mid l = 1, 2, 3, \ldots, k\}$. The sensing nodes within the VSN can be represented by,

$$\mathbf{VS} = \begin{bmatrix} \mathrm{VS}_{1,1} & \mathrm{VS}_{1,2} & \ldots & \mathrm{VS}_{1,n} \\ \mathrm{VS}_{2,1} & \mathrm{VS}_{2,2} & \ldots & \mathrm{VS}_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{VS}_{k,1} & \mathrm{VS}_{k,2} & \ldots & \mathrm{VS}_{k,n} \end{bmatrix}$$

Assume the sensor network is divided into clusters and each cluster head receives control signals from the cluster nodes to determine whether they are active or inactive. $\mathrm{VS}_{l,j}$ is assigned a value based on the following condition,

$$\mathrm{VS}_{l,j} = \begin{cases} -1, & j > n_l \\ 1, & \text{the sensing node is active} \\ 0, & \text{the sensing node is inactive} \end{cases} \tag{3.15}$$

Suppose t represent the total sensing and processing tasks within the VSN. Let an i-dimensional task classification matrix $\mathbf{T}$ such that $i = \{1, 2, 3, \ldots, t\}$ is given by,

$$\mathbf{T} = \begin{bmatrix} T_{1,1} & T_{1,2} & \ldots & T_{1,n} \\ T_{2,1} & T_{2,2} & \ldots & T_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ T_{k,1} & T_{k,2} & \ldots & T_{k,n} \end{bmatrix}$$

where each $T_{l,j}$ is given by,

$$T_{l,j} = \begin{cases} 1, & ith \text{ task is assigned to sensing node } \mathrm{VS}_{l,j} \\ 0, & \text{otherwise} \end{cases} \tag{3.16}$$

In the proposed approach, upto $\left\lceil \sqrt{k} \right\rceil$ sensor classes are assigned an ith task; where ($\lceil\ \rceil$) refers to the ceiling function. Let $\mathbf{T}'_i$ denote the adaptive ith task classification matrix which optimises $\mathbf{T}_i$ for active sensing nodes within the VSN and is given by,

$$\mathbf{T}'_i = \left\lfloor \frac{1}{2} \Big[\mathbf{T}_i \cdot \mathbf{VS} + \mathbf{J} \Big] \right\rfloor \tag{3.17}$$

where $\mathbf{J}$ is a $k \times n$ all-ones matrix and ($\lfloor\ \rfloor$) refers to the floor function.

Feedback $\mathbf{R}_c$ and $\mathbf{r}$ are substituted in equation Equation 3.3 to calculate the required 3D scene coverage $\mathbf{V}_c$ of k sensor classes to perform t tasks and the chosen 3D scene coverage $\mathbf{v}$ of k sensor classes respectively. Algorithm 3.1 presents the proposed adaptive task classification scheme that calculates $\mathbf{T}$ and then $\mathbf{T}'$ in an optimised way.

Algorithm 3.1 Adaptive task classification scheme for heterogeneous networks

Require:

The number of: sensor classes k, sensing nodes of each type n_l, tasks t required to be performed by the VSN; the required 3D scene coverage $\mathbf{V}_c$ of k sensor classes to perform t tasks and the chosen 3D scene coverage $\mathbf{v}$ of k sensor classes.

Ensure:

For $\forall\ j \in \{1, 2, 3, \ldots, n\}$ and $\forall\ i \in \{1, 2, 3, \ldots, t\}$

$$0 < \sum_{l=1}^{k} T'(l, j, i) \leq \lceil \sqrt{k} \rceil$$

1: $n \leftarrow \max\{n_1, n_2, n_3, \ldots, n_k\}$

Algorithm 3.1 (continued) Adaptive task classification scheme for heterogeneous networks

VS $\leftarrow \varnothing$
s $\leftarrow [-1 \;\; 1 \;\; 0 \;\; 0 \;\; 0]$
for $l \leftarrow 1$ **to** k **do**
    for $j \leftarrow 1$ **to** n **do**
        if $j > n_l$ **then**
            $\mathrm{VS}(l, j) \leftarrow s_1$
        else if sensing node is active **then**
            $\mathrm{VS}(l, j) \leftarrow s_2$
        else if sensing node is inactive **then**
            $\mathrm{VS}(l, j) \leftarrow s_3$
        end if
    end for
end for
T $\leftarrow \varnothing$
T$' \leftarrow \varnothing$
for $i \leftarrow 1$ **to** t **do**
    for $l \leftarrow 1$ **to** k **do**
        $s_5 \leftarrow k - l + 1$
        if $V_c(s_5, i) \geq v(s_5)$ & $s_4 < \lceil \sqrt{k} \rceil$ **then**
            $T(s_5, 1{:}n, i) \leftarrow s_2$
            $s_4 \leftarrow s_4 + 1$
        else
            $T(s_5, 1{:}n, i) \leftarrow s_3$
        end if
    end for
    $\mathbf{T}'_i = \left\lfloor \frac{1}{2} \left[\mathbf{T}_i \cdot \mathbf{VS} + \mathbf{J} \right] \right\rfloor$
    $s_4 \leftarrow 0$
end for
return T$'$

In the proposed adaptive task classification scheme, initially, the number of sensing nodes belonging to different sensor classes are compared and the maximum number of sensing nodes belonging to a particular sensor class are found and represented by n. A matrix $\mathbf{VS}$ of dimension $k \times n$ is defined. Moreover, a vector $\mathbf{s}$ is defined to keep sensing nodes' status information. Afterwards, each sensing node is assigned a particular status, where s_2 denote that the sensing node is active and s_3 represent an inactive sensing node. Once the status of each sensing node within the network is determined, matrices $\mathbf{T}$ and $\mathbf{T}'$ with dimensions equal to the dimension of $\mathbf{VS}$ (i.e. $k \times n$) are defined. Subsequently, for an ith task such that $i = \{1, 2, 3, \ldots, t\}$, the criteria $\left\{V_c(s_5, i) \geq v(s_5)\ \&\ s_4 < \lceil\sqrt{k}\rceil\right\}$ is used for ith task allocation to an lth sensor class such that $l = \{1, 2, 3, \ldots, k\}$. However, as discussed earlier, the active and inactive sensing nodes within the network have already been identified in $\mathbf{VS}$. Therefore, for an ith task, utilising $\mathbf{T}_i$, $\mathbf{VS}$, $\mathbf{J}$ and the floor function ($\lfloor\ \rfloor$), $\mathbf{T}'_i$ providing information regarding the classification of ith task between active sensing nodes of the network is obtained. The algorithm is terminated once $\mathbf{T}'_i$ is calculated for all the tasks required to be performed within the network.

3.3.5 Feature Detection and Extraction

Global colour histogram is used for object detection and feature extraction. A global colour histogram represents the distribution of colours within each captured image $\mathbf{I}_i$ of size $P \times Q$ and is given by [169],

$$h_c(b) = \sum_{x=1}^{P} \sum_{z=1}^{Q} \begin{cases} 1, & I_i(x, z) \text{ is in bin } b \\ 0, & \text{otherwise} \end{cases} \tag{3.18}$$

where a colour bin defines a region of particular colour.

In this framework, histogram-based features have been extracted in YCbCr colour space as it distinguishes the luminance and chrominance. The extracted features

have been analysed and a range of values of Cb and Cr has been defined to detect a particular object of interest through image segmentation.

The probability $P(E)$ of a pixel at location (x, z) belonging to an object of interest is given by,

$$P(E) = \begin{cases} 1, & \gamma_{Cb}^{l} \leq \text{Cb} \leq \gamma_{Cb}^{u} \cap \gamma_{Cr}^{l} \leq \text{Cr} \leq \gamma_{Cr}^{u} \\ 0, & \text{otherwise} \end{cases} \tag{3.19}$$

The pixels probabilities are indexed at their respective locations in the object segmentation matrix $\mathbf{S}_m$. The object of interest is extracted from $\mathbf{I}_i$ by image segmentation using the following equation [103],

$$\mathbf{S}_g = \mathbf{I}_i.\mathbf{S}_m \tag{3.20}$$

where $\mathbf{S}_g$ is the segmented image and (.) refers to the dot product.

The segmented image can be morphologically processed to remove noise and small regions. Blob detection and circularity test (for homogeneous objects) can be used to detect the object of interest.

3.3.6 FoV Characterisation

As mentioned earlier, a visual sensing node captures scene points within its FoV onto the physical image plane. Although the dimensions of ABCD-plane can be regulated by R, the dimensions of physical image plane are always constant. The relationship between sensor's resolution and ABCD-plane dimensions for a distance R is given by,

$$d_h = \frac{P}{w_2}$$
$$d_v = \frac{Q}{h_2} \tag{3.21}$$

where d_h is the horizontal and d_v is the vertical density measured in $pixels/mm$.

As P and Q are constant, it is found that $d_h \propto 1/w_2$ and $d_v \propto 1/h_2$. Increasing the distance R increases w_2 and h_2 which results in the reduction of horizontal and vertical density. If R goes outside a certain range, the captured image may not provide sufficient feature descriptors.

Hence, the need arises to propose a criteria for optimised range defined as the *Field-of-View Characterisation Criteria* (FoVCC). FoVCC must ensure the presence of sufficient feature descriptors within the captured image as well as guarantee optimised utilisation of resources while maintaining a certain quality.

The proposed FoV characterisation method for energy efficient VSNs utilises either one or a combination of the following parameters.

(i) Object Pixel Occupancy (O_{po})

(ii) Estimation Error ($|\varepsilon_d|$)

(iii) Peak Signal-to-Noise Ratio (PSNR_{dB})

The parameters and their respective criterion are discussed in the following sections.

i) Object Pixel Occupancy

The number of pixels (O_{po}) an object of interest occupies in the image captured from a particular distance R is given by,

$$O_{po} = A \times d_h \times d_v \quad (pixels) \tag{3.22}$$

where A is the area of the object in mm^2; the product of d_h and d_v is calculated from Equation 3.21 in $pixels/mm^2$.

Rearranging Equation 3.22,

$$\frac{O_{po}}{A} = d_h \times d_v \tag{3.23}$$

Substituting d_h and d_v from Equation 3.21,

$$\frac{O_{po}}{A} = \frac{P}{w_2} \times \frac{Q}{h_2} \tag{3.24}$$

Substituting w_2 and h_2,

$$O_{po} = A \times \frac{P}{2R\tan\left(\frac{\theta_h}{2}\right)} \times \frac{Q}{2R\tan\left(\frac{\theta_v}{2}\right)} \tag{3.25}$$

Let ξ_o define the required minimum pixel occupancy for a particular application, the chosen range R_1 of a visual sensing node must guarantee the criteria $O_{po} \geq \xi_o$ and it can be calculated using the following equation,

$$R_1 = \sqrt{\frac{P \times Q \times A}{4 \times O_{po} \times \tan\left(\frac{\theta_h}{2}\right) \times \tan\left(\frac{\theta_v}{2}\right)}} \tag{3.26}$$

Table 3.1 provides the minimum object pixel occupancy required for various detection algorithms. ξ_o for face detection depends on the image size used to train the classifier. The detection accuracy calculated on PETS 2005 data set in [170] with $\xi_o = 25$ for LOTS, SGM and MGM is 91.2%, 86.8% and 85.0% respectively.

TABLE 3.1: Required minimum object pixel occupancy for various detection algorithms

Detection Method	ξ_o
Viola-Jones face detector [171]	315
Lehigh Omnidirectional Tracking System (LOTS) [172]	25
Single Gaussian Model (SGM) [173]	25
Multiple Gaussian Model (MGM) [174]	25

ii) Estimation Error

Increasing range R_1 reduces the object pixel occupancy O_{po} which may lead to detection/estimation error. Hence, the need arises to provide a method for the estimation of maximum sensor-to-object distance based on a certain acceptable error tolerance level. In order to propose such method, an application that estimates the detected object's diameter from the acquired visual data is considered.

After feature detection, if O_{po} denote the number of pixels representing the detected object; the framework estimates pixels representing the diameter p_d by,

$$p_d = 2\sqrt{\frac{O_{po}}{\pi}} \tag{3.27}$$

The diameter d_e of the object is estimated by,

$$d_e = \frac{4R\tan\left(\frac{\theta_h}{2}\right)}{P}\sqrt{\frac{O_{po}}{\pi}} \quad (metres) \tag{3.28}$$

If d_a is the actual measured diameter of the object of interest, the absolute percentage estimation error $|\varepsilon_d|$ is given by,

$$|\varepsilon_d| = \frac{|d_a - d_e|}{d_a} \times 100 \quad (\%) \tag{3.29}$$

It is expected that as the range increases, the estimation error will increase. Let ξ_d define the maximum acceptable estimation error in percentage for a particular application, the chosen range R_2 of a visual sensing node must guarantee the criteria $|\varepsilon_d| \leq \xi_d$.

Substituting Equation 3.28 in Equation 3.29, the range R_2 can be calculated using the following equation,

$$R_2 = \left[\sqrt{\frac{\pi}{O_{po}}}\right]\left[\frac{Pd_a}{4\tan\left(\frac{\theta_h}{2}\right)}\right]\left[1 - \frac{|\varepsilon_d|}{100}\right] \tag{3.30}$$

The above equation is valid for $|\varepsilon_d| \neq 100\%$.

iii) PSNR

Suppose image $\mathbf{I}_1$ of dimension $P \times Q$ is captured at a distance R_p which contains a particular object of interest. The aforementioned histogram-based feature extraction scheme is employed to extract the region of interest containing only the object under consideration in the form of image $\mathbf{I}'_1$ of dimension $P' \times Q'$. As $\mathbf{I}'_1$ contains the object captured at distance R_p, the dimension $P_s \times Q_s$ of image $\mathbf{I}_s$ containing the extracted object at distance R_s (such that $R_s > R_p$) is estimated by,

$$P_s = P'\left(\frac{w_2^{R_p}}{w_2^{R_s}}\right) \tag{3.31}$$

$$Q_s = Q'\left(\frac{h_2^{R_p}}{h_2^{R_s}}\right) \tag{3.32}$$

where $w_2^{R_p}$, $w_2^{R_s}$, $h_2^{R_p}$ and $h_2^{R_s}$ are calculated using Equation 3.12 and Equation 3.13. As $P_s < P'$ and $Q_s < Q'$, the object captured and extracted at distance R_s appears

smaller in size. In order to measure the quality, $\mathbf{I}'_1$ and $\mathbf{I}_s$ are compared to find the PSNR value. As PSNR requires both images to have the same size, $\mathbf{I}_s$ is resized to $P' \times Q'$. First, Mean Squared Error (MSE) is calculated and then the PSNR.

$$\text{MSE} = \frac{1}{3P'Q'} \sum_{x=1}^{P'} \sum_{y=1}^{Q'} \Big[I'_1(x,y) - I_s(x,y) \Big]^2 \tag{3.33}$$

$$\text{PSNR}_{\text{dB}} = 10 \times \log_{10} \left(\frac{\text{MAX}^2_{\mathbf{I}'_1}}{\text{MSE}} \right) \tag{3.34}$$

where $\text{MAX}^2_{\mathbf{I}'_1}$ is the maximum possible pixel value in $\mathbf{I}'_1$.

Let ξ_p define the required minimum PSNR_{dB} for a particular application, the chosen range R_3 of a visual sensing node must guarantee the criteria $\text{PSNR}_{\text{dB}} \geq \xi_p$. As this method is based on image quality assessment, an experiment needs to be conducted to find the range R_3 from graph analysis which is discussed later in Section 3.4.

Apart from PSNR_{dB}, there are many other image quality assessment methods such as [175–181] that can be used with the proposed FoV characterisation framework based on their respective confidence bounds for sensing range estimation.

The FoV characterisation methods and their respective equations for range estimation are summarised in Table 3.2. The selection of one or more characterisation methods depends on the application design criteria as stated below.

- The application where design considers the detection method's minimum pixel requirement, object pixel occupancy based method is used.
- If the design criteria depends on a particular tolerance level, then estimation error based method is used.
- The design considering image quality utilises the PSNR based method.

These methods are interdependent i.e. selecting R for a lower estimation error is expected to provide a larger pixel occupancy and better image quality (in terms of PSNR). Similarly, selecting R for a poor image quality (in terms of PSNR) is expected to provide lower pixel occupancy and higher error.

TABLE 3.2: A summary of FoV characterisation methods

Characterisation Method	**Selection Method Based on**	**Characterisation Criteria**	**Range Estimation**
Object pixel occupancy based	Algorithm's minimum pixel requirement	$O_{po} \geq \xi_o$	$R_1 = \sqrt{\frac{P \times Q \times A}{4 \times O_{po} \times \tan\left(\frac{\theta_h}{2}\right) \times \tan\left(\frac{\theta_v}{2}\right)}}$
Estimation error based	Maximum acceptable error tolerance	$\lvert\varepsilon_d\rvert \leq \xi_d$	$R_2 = \left[\sqrt{\frac{\pi}{O_{po}}}\right]\left[\frac{Pd_a}{4\tan\left(\frac{\theta_h}{2}\right)}\right]\left[1 - \frac{\lvert\varepsilon_d\rvert}{100}\right]$
PSNR based	Image quality assessment	$\text{PSNR}_{\text{dB}} \geq \xi_p$	R_3 - from graph analysis

iv) FoV Charaterisation Criteria

Let R_c define the chosen value of a sensing node's range, the FoV Characterisation Criteria (FoVCC) is proposed as,

$$R_c = \begin{cases} R_1, & O_{po} \geq \xi_o \\ R_2, & \lvert\varepsilon_d\rvert \leq \xi_d \\ R_3, & \text{PSNR}_{\text{dB}} \geq \xi_p \end{cases} \tag{3.35}$$

The sensing range for applications where the design engineer utilises more than one characterisation method is selected by,

$$R_c = \min\{R_1, R_2, R_3\} \tag{3.36}$$

The designed VSN's FoV is said to be optimised based on the following criteria,

$$\begin{cases} R_c = R_1 \cup R_c = R_2 \cup R_c = R_3 & \text{optimised} \\ R_c \neq R_1 \cap R_c \neq R_2 \cap R_c \neq R_3 & \text{unoptimised} \end{cases} \tag{3.37}$$

v) Adaptive Range Selection

a) Hard decision based sensing range selection: In homogeneous network design that considers t tasks to be performed within the VSN, sensing range $\{R_{c(i)} \mid i = 1, 2, 3, \ldots, t\}$ is required to be calculated for each task. The sensing range R_c can be obtained by hard decision as shown below,

$$R_c = \min \{R_{c(1)}, R_{c(2)}, R_{c(3)}, \ldots, R_{c(t)}\} \tag{3.38}$$

The chosen sensing range R_c is the feedback to projection modelling.

In the case of heterogeneous network design, $\mathbf{R}_c$ of dimension $k \times t$ is the feedback to projection modelling which provides the estimated sensing range of k sensor classes for t tasks.

$$\mathbf{R}_c = \begin{bmatrix} R_{c(1,1)} & R_{c(1,2)} & \ldots & R_{c(1,t)} \\ R_{c(2,1)} & R_{c(2,2)} & \ldots & R_{c(2,t)} \\ \vdots & \vdots & \ddots & \vdots \\ R_{c(k,1)} & R_{c(k,2)} & \ldots & R_{c(k,t)} \end{bmatrix}$$

Sensing range $r(l)$ can be calculated for each sensor class by hard decision as shown below,

$$r(l) = \min \{R_{c(l,1)}, R_{c(l,2)}, R_{c(l,3)}, \ldots, R_{c(l,t)}\} \tag{3.39}$$

The individual sensing range values for different sensor classes obtained through hard decision can be represented collectively by $\mathbf{r}$ as,

$$\mathbf{r} = \begin{bmatrix} r(1) \\ r(2) \\ \vdots \\ r(k) \end{bmatrix}$$

The proposed hard decision based scheme is suitable for homogeneous networks as they have identical sensing nodes and hard decisions need to be made for sensing range selection. However, the hard decision based scheme for heterogeneous networks does not take advantage of the multiple sensor classes present within the network. The approach provides the minimum range for each sensor class and does not prolong the network lifetime by maximising the sensing range. To maximise the sensing range and prolong the lifetime of heterogeneous networks, a soft decision based scheme for sensing range selection is proposed in the following section.

b) Soft decision based sensing range selection: A soft decision based sensing range selection scheme is proposed in Algorithm 3.2 which calculates a suitable range $r(\cdot)$ for each sensor class based on the estimated $\mathbf{R}_c$. The algorithm provides range $\mathbf{r}$ for k sensor classes by maximising it for $\left(k - \left\lceil \sqrt{k} \right\rceil\right)$ sensor classes.

Algorithm 3.2 Proposed soft decision based sensing range selection scheme for heterogeneous network design

Require:

The number of sensor classes k, the number of sensing nodes of each type n_l, the number of tasks t required to be performed by the VSN, $\mathbf{R}_c$ providing the estimated sensing range of k sensor classes for t tasks.

Algorithm 3.2 (continued) Proposed soft decision based sensing range selection scheme for heterogeneous network design

Ensure:

For $\left\lceil \sqrt{k} \right\rceil$ values of $l \in \{1, 2, 3, \ldots, k\}$ and $\forall\ i \in \{1, 2, 3, \ldots, t\}$

$r(l) \leq R_{c(l,i)}$

$\mathbf{s}_1 = \varnothing$

for $l \leftarrow 1$ **to** k **do**

  $\mathbf{s}_1 \leftarrow \sum_{i=1}^{t} R_{c(l,i)}$

end for

$\mathbf{s}_2 \leftarrow$ Sort $\mathbf{s}_1$ in ascending order

$\mathbf{s}_3 \leftarrow$ First $\left\lceil \sqrt{k} \right\rceil$ values' indices from $\mathbf{s}_2$

for $l \leftarrow 1$ **to** k **do**

  if $l \in \mathbf{s}_3$ **then**

    $r(l) \leftarrow \min\left\{R_{c(l,1)}, R_{c(l,2)}, R_{c(l,3)}, \ldots, R_{c(l,t)}\right\}$

  else if $l \notin \mathbf{s}_3$ **then**

    $r(l) \leftarrow \left(\frac{1}{t}\right) \times \sum_{i=1}^{t} R_{c(l,i)}$

  end if

end for

return r

In the proposed soft decision based sensing range selection scheme presented in Algorithm 3.2, initially a vector $\mathbf{s}_1$ defined. The algorithm needs matrix $\mathbf{R}_c$ as an input which provides the estimated sensing range of k sensor classes for t tasks required to be performed within the network. Afterwards, for lth sensor class such that $l \in \{1, 2, 3, \ldots, k\}$, sum of the sensing range values for t number of tasks is calculated. This process is repeated for k sensor classes and the results are obtained in $\mathbf{s}_1$. Subsequently, the values in $\mathbf{s}_1$ are sorted in ascending order and the resulting values are stored in $\mathbf{s}_2$. As the proposed soft decision based sensing range selection scheme is to be designed in such a way that it maximises the sensing range for $\left(k - \left\lceil \sqrt{k} \right\rceil\right)$ sensor classes, the indices of the first $\left\lceil \sqrt{k} \right\rceil$

values from $\mathbf{s}_2$ are stored in $\mathbf{s}_3$. Thereafter, for lth sensor class, the criterion $l \in \mathbf{s}_3$ and the criterion $l \notin \mathbf{s}_3$ are utilised for soft decision based sensing range selection. In case if criterion $l \in \mathbf{s}_3$ is true, the sensing range for lth sensor class is obtained by employing $r(l) = \min\{R_{c(l,1)}, R_{c(l,2)}, R_{c(l,3)}, \ldots, R_{c(l,t)}\}$. Whereas, if criterion $l \notin \mathbf{s}_3$ is true, the sensing range for lth sensor class is calculated using $r(l) = \left(\frac{1}{t}\right) \times \sum_{i=1}^{t} R_{c(l,i)}$. The algorithm is terminated once the sensing range is calculated for all the sensor classes.

3.4 Experimental Setup, Results and Analysis

This section evaluates the robustness of the proposed FoV characterisation framework for homogeneous and heterogeneous VSNs. An analysis of system failure probability is also presented. In order to evaluate the performance of the proposed FoV characterisation framework for homogeneous and heterogeneous VSNs, the analysis is performed using the MATLAB simulation platform on an Intel Core i3-2100U CPU @ 3.10GHz with 4 GB RAM and 3 MB cache. Simulation modelling is chosen to analyse the performance of the proposed schemes and frameworks in this thesis as: it reduces the cost of experimentation; offers a safe way for testing and provides details of system behaviour with different parameter values. Moreover, networks can be analysed without deployment limitations using simulations [10]. In addition, several notable existing schemes [10, 29, 55, 58, 61, 124] found in literature have been evaluated using simulations.

3.4.1 Image Capture

Specification of the visual sensor used for experiments is presented in Table 3.3. The visual sensor is utilised to validate the Projection Modelling Approach II (as discussed in Section 3.4.3). Once the horizontal FoV (θ_h) and vertical FoV (θ_v) have been estimated, the parameter values of the visual sensor as well as

the captured images are used to analyse the feasibility and energy efficiency of proposed FoV characterisation framework in resource constrained scenarios.

TABLE 3.3: Visual sensor specification

Parameter	**Specification**
Imaging device	3.6mm (1/5 type) CMOS sensor
Focal length	3.2mm
Resolution	2304×1728 pixels
Pixel size	$1.25\mu m \times 1.25\mu m$
Sensor dimensions	$2.88mm \times 2.16mm$

3.4.2 Projection Modelling utilising Approach I

Using Projection Modelling Approach I for homogeneous VSNs, after substituting focal length (f) and sensor dimensions ($w_1 \times h_1$) in Equations 3.4 and 3.5, the calculated values of horizontal and vertical FoVs are: $\theta_h = 48.39°$, $\theta_v = 37.25°$ respectively.

3.4.3 Projection Modelling utilising Approach II

As mentioned in Section 3.3.2, in the case of a heterogeneous network with some unknown sensor specifications, Equations 3.4 and 3.5 cannot be used. Therefore, an experiment has been conducted utilising Projection Modelling Approach II outlined in the proposed estimation framework to calculate θ_h and θ_v. In order to measure the accuracy of the calculated FoV values from Projection Modelling Approach II, they will be compared with those calculated from Projection Modelling Approach I. The experimental procedure is described in Table 3.4.

TABLE 3.4: Experimental procedure for calculating horizontal FoV (θ_h) and vertical FoV (θ_v)

Procedure: Horizontal FoV (θ_h) and vertical FoV (θ_v) calculation
1: The visual sensing node is placed near a known vertically uniform object at a certain height h_t and a known reference distance R_{ref} without any tilt or pan.
2: A certain portion of the object is captured within the sensing node's FoV.
3: The width w_2 and height h_2 of the object's portion within the FoV are measured for R_{ref}.
4: Using Equations 3.6 and 3.7, θ_h and θ_v are calculated.
5: Steps 1 to 4 are repeated for a set of values of R_{ref} to guarantee the accuracy of the calculated θ_h and θ_v values.

The experimental results for six cases have been summarised in Table 3.5; where each case is distinguished by its reference distance R_{ref}.

TABLE 3.5: FoV calculation utilising Projection Modelling Approach II

Parameter	**Case 1**	**Case 2**	**Case 3**	**Case 4**	**Case 5**	**Case 6**	**Average**
R_{ref}	$0.27m$	$0.30m$	$0.45m$	$0.68m$	$0.92m$	$1.91m$	-
w_2	$0.24m$	$0.27m$	$0.40m$	$0.61m$	$0.82m$	$1.72m$	-
h_2	$0.18m$	$0.20m$	$0.30m$	$0.46m$	$0.62m$	$1.29m$	-
h_t	$1.04m$	$1.04m$	$1.04m$	$1.04m$	$1.04m$	$1.04m$	-
θ_h	$48.37°$	$48.38°$	$48.43°$	$48.41°$	$48.29°$	$48.48°$	$48.39°$
θ_v	$37.11°$	$37.23°$	$37.29°$	$37.30°$	$37.16°$	$37.40°$	$37.25°$
ϕ_a	$41.54°$	$41.54°$	$41.54°$	$41.54°$	$41.54°$	$41.54°$	-
θ_e	$52.75°$	$52.75°$	$52.75°$	$52.75°$	$52.75°$	$52.75°$	-

Based on the experimental results it is found that the error for each case is negligible and by averaging the estimated values, the projection modelling approach II leads to accurate FoV measurements.

3.4.4 ABCD-plane Modelling

ABCD-plane modelling plays a vital role for FoV characterisation. As θ_h and θ_v have been calculated, extensive numerical simulations have been performed for ABCD-plane modelling utilising Equations 3.12 and 3.13 for a range of values of R. The simulation results are presented in Figure 3.6.

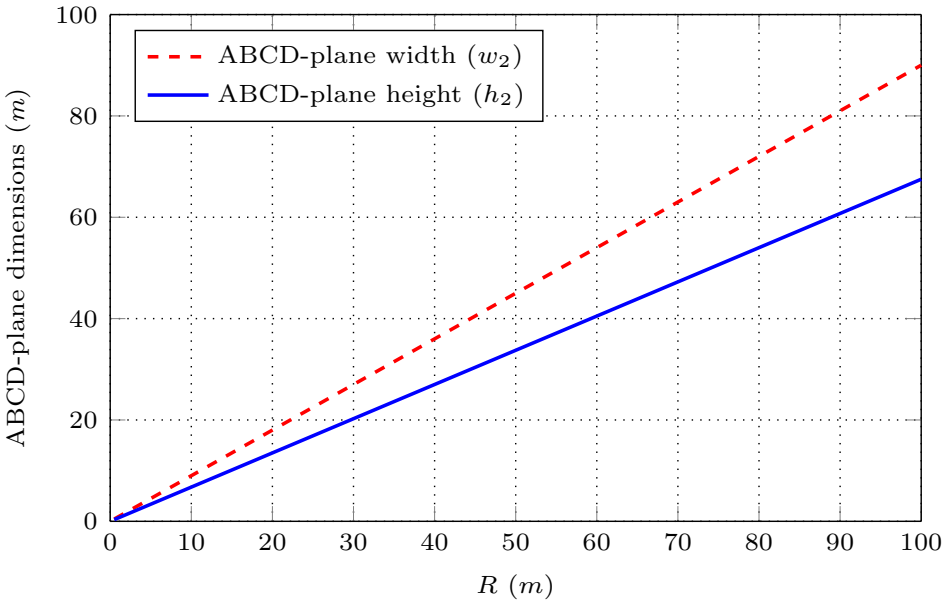

FIGURE 3.6: ABCD-plane modelling

From the results, it is found that increasing the sensing range increases the ABCD-plane's width (w_2) and height (h_2) as well. This behaviour is logical due to the relationship between the sensing range and the ABCD-plane's dimensions in Equations 3.12 and 3.13. As the visual sensing nodes are assumed to be static, i.e. θ_h and θ_v are constant, it can be noticed from Equation 3.12 that the sensing range

R is directly proportional to the ABCD-plane's width w_2. Similarly, it can be observed from Equation 3.13 that the sensing range R is directly proportional to the ABCD-plane's height h_2. Hence, the ABCD-plane's dimensions increase linearly by increasing the sensing range. This relationship can be utilised to estimate the dimensions of the ABCD-plane for a particular sensing range value, thus, supporting the FoV characterisation process. In this case, $w_2 > h_2$ for any value of R due to the fact that $\theta_h > \theta_v$.

3.4.5 Feature Detection and Extraction

Using the global colour histogram, the probability $P(E)$ of a pixel at location (x, z) belonging to the object of interest is found to be,

$$P(E) = \begin{cases} 1, & 43 \leq \text{Cb} \leq 90 \cap 138 \leq \text{Cr} \leq 159 \\ 0, & \text{otherwise} \end{cases} \tag{3.40}$$

After dataset creation (discussed in the following section) probability $P(E)$ can be used for feature detection and extraction.

3.4.6 FoV Characterisation

Although the FoV characterisation depends on several factors, sensing range (R) is the key parameter for the characterisation process. The sensing range estimation and optimisation requires practical measurements and simulations. The experimental procedure for these calculations is described in Table 3.6.

In order to estimate the sensing range for optimised FoV characterisation, a dataset is created by capturing object for a range of values of R i.e. $0.25m$ to $9.55m$. Sensing range can be estimated for optimised FoV characterisation using one or a combination of the following parameters: object pixel occupancy, estimation

TABLE 3.6: Experimental procedure for FoV characterisation

Procedure: FoV Characterisation
1: A particular type of feature needs to be considered based on the desired application. In this experiment, colour features are considered.
2: Images of objects under consideration are captured for a range of values of R.
3: The captured images are processed to classify those containing sufficient colour feature descriptors.
4: A particular value of R_c is chosen for VSN design based on the FoVCC.

error, PSNR. The experimental and simulation results for FoV characterisation are presented and analysed in the following sections.

i) Object Pixel Occupancy

Figure 3.7 shows a comparison of the theoretical and experimental object pixel occupancy for a range of values of R with $A = 34cm^2$ for a homogeneous test object. The theoretical results are obtained from Equation 3.22 whereas, the experimental results are achieved by extracting the object of interest from the captured images (using $P(E)$ for feature detection and extraction) and counting the pixels it occupies.

It has been noticed that initially, at lower sensing range values, the object pixel occupancy is higher, however, as the sensing range increases, the object pixel occupancy reduces sharply. Due to the fact that an object-of-interest appears smaller if captured by a visual sensing node with a higher sensor-to-object distance, the reduction in object pixel occupancy is justified. It is evident from the graph that the object pixel occupancy estimated by the proposed method matches with the experimental results. The proposed object pixel occupancy based characterisation method provides FoV based mapping between the sensing range and the object pixel occupancy and it can be used to estimate an optimised sensing range based on the application's object pixel occupancy requirement.

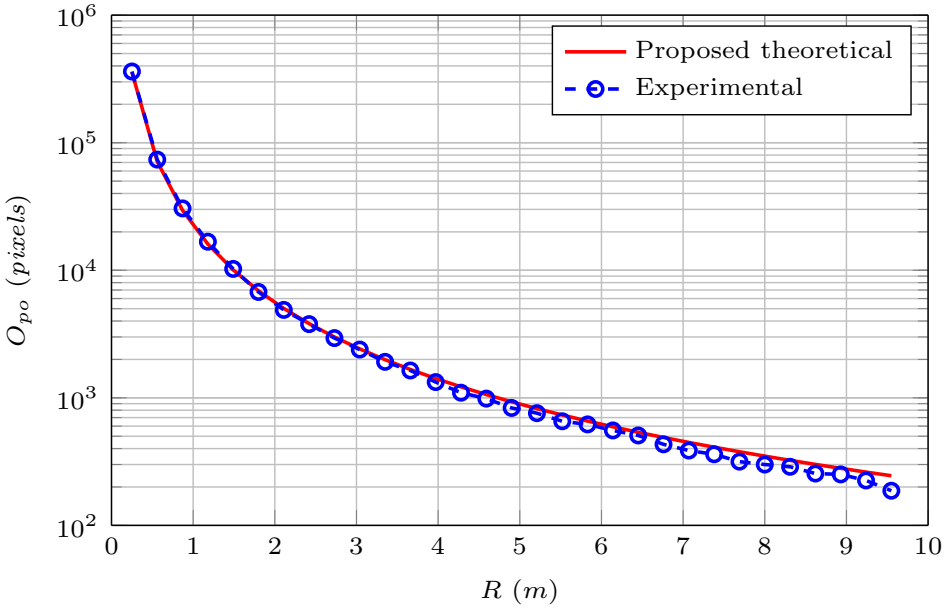

FIGURE 3.7: A comparison of theoretical and experimental values of object pixel occupancy

As an example, a face detection application [171] requires object pixel occupancy to be at least 315 *pixels* i.e. $O_{po} \geq 315$, the acceptable sensing range in that case will be $R_1 \leq 4.05m$.

ii) Estimation Error

As mentioned earlier, in order to observe the relationship between the sensing range and the estimation error, an application is considered where an object's diameter estimation is required after the detection phase. Using Equation 3.28, Figure 3.8 presents a comparison of actual and estimated diameter for images captured for a range of values of R. It is observed from Figure 3.8 that at lower sensing range values, the estimated diameter of the detected object is closer to the actual diameter value. However, as the sensing range increases, the estimated diameter deviates from the actual diameter, thus increasing the error.

The absolute percentage estimation error $|\varepsilon_d|$ is shown in Figure 3.9. It has been observed from Figure 3.9 that as the range increases, the estimation error increases.

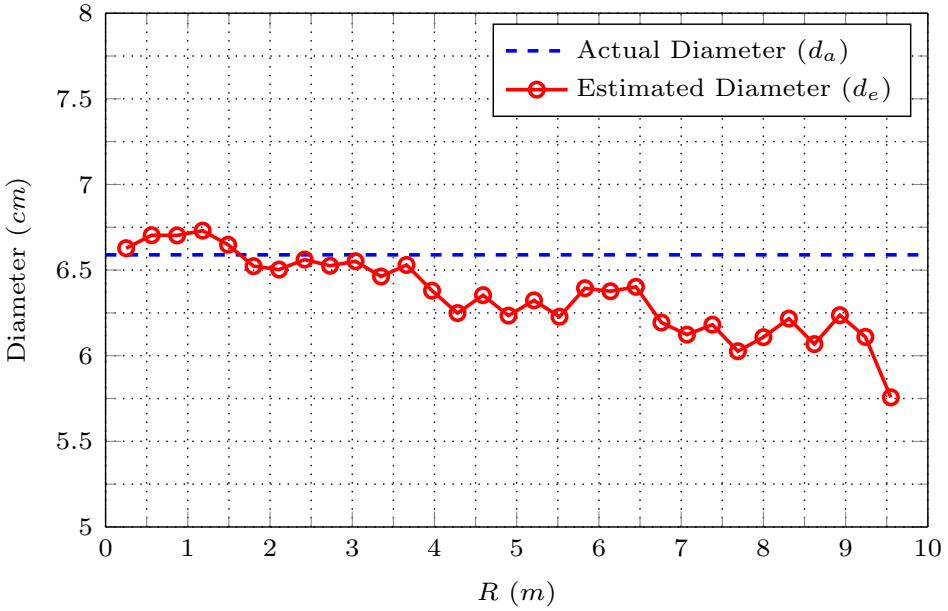

FIGURE 3.8: A comparison of actual and estimated diameter for different sensing range values

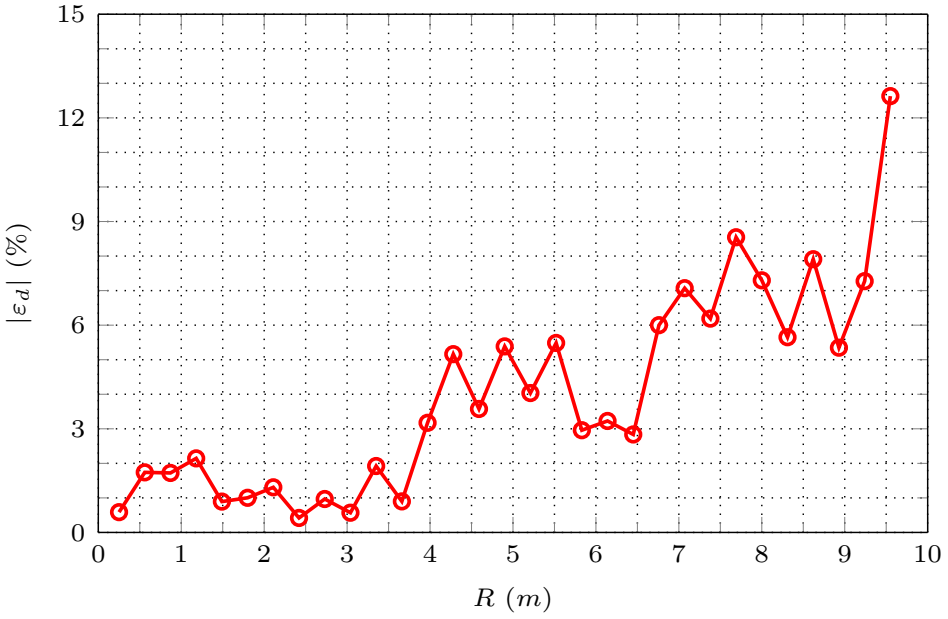

FIGURE 3.9: Estimation error for different sensing range values

This is due to the fact that the object of interest appears too small beyond a certain range which leads to inaccurate feature detection and extraction results. Therefore, depending on the maximum error tolerance for a given application, the estimation error based FoV characterisation method can be utilised by the design

engineers to obtain suitable range of visual sensing nodes.

As an example, suppose a particular application can tolerate maximum 6% error i.e. $|\varepsilon_d| \leq 6\%$, the acceptable sensing range will be $R_2 \leq 6.76m$.

iii) PSNR

This method utilises an image quality assessment technique for FoV characterisation. Figure 3.10 shows the estimated PSNR_{dB} for a range of values of R.

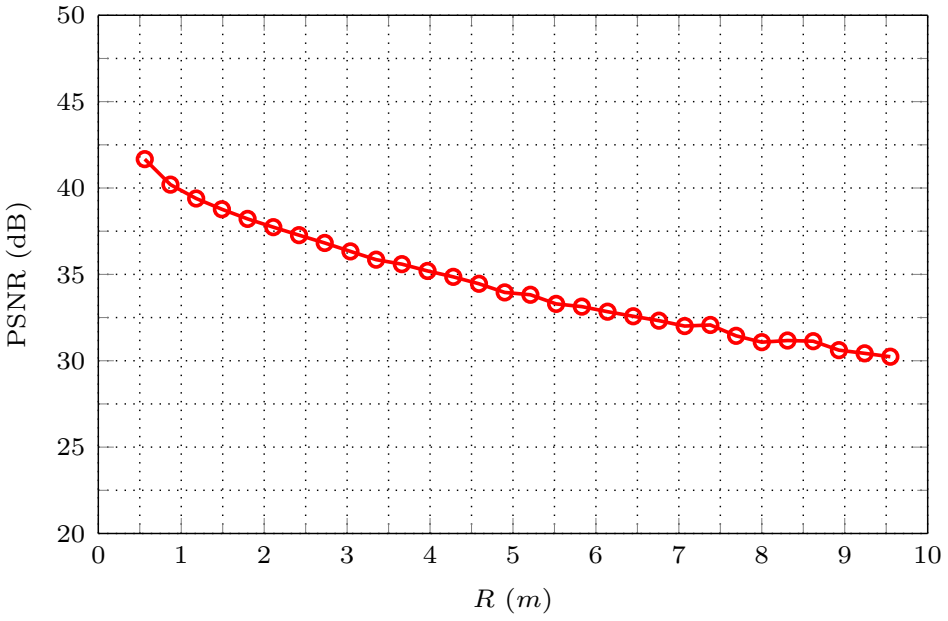

FIGURE 3.10: PSNR estimation for different sensing range values

It is noticed from Figure 3.10 that initially, at lower sensing range values, the peak signal-to-noise ratio is higher. However, as the sensing range increases, the peak signal-to-noise ratio decreases gradually. Assuming static visual sensing nodes (i.e. with fixed sensing characteristics), this behaviour is logical as an increase in sensing range is expected to affect the image quality. Therefore, the selection of suitable sensing range depending on the desired image quality is an important process during VSN design. This assures that the sensing node is able to provide visual information regarding the region within its FoV with an acceptable degree

of reliability. As PSNR_{dB} is an index for image quality assessment, this method can assist the design engineer to tune the network for a suitable image quality. As an example, an image transmission application [182] requires PSNR_{dB} to be at least 30 dB i.e. $\text{PSNR}_{\text{dB}} \geq 30$ dB, the acceptable sensing range in that case will be $R_3 \leq 9.55m$.

iv) Adaptive Range Selection

a) Homogeneous Networks: Consider a homogeneous network design for a surveillance application that requires $t = 2$ tasks to be performed within the VSN i.e. face detection [171] (Task I) and occluded target surveillance and tracking [172] (Task II). Suppose a medium resolution sensor with the following parameters: $P \times Q = 640 \times 480$, $\theta_h = 48.39°$ and $\theta_v = 37.25°$ is selected for the VSN design. Let the object pixel occupancy based characterisation method be used with required minimum $O_{po} = 315$ and $O_{po} = 25$ for Task I and Task II respectively. The area (A) to be considered for detection is found to be $406cm^2$ for Task I and $3922.6cm^2$ for Task II. By substituting these parameters in Equation 3.26, the sensing range estimated for Task I is $R_{c1} = 8.09m$ and for Task II is $R_{c2} = 89.21m$. According to the hard decision based sensing range selection method for homogeneous networks, the chosen sensing range R_c is $\min\{R_{c1}, R_{c2}\}$ i.e. $R_c = 8.09m$. The chosen range R_c is a feedback to projection modelling and it is also used to find the number of active sensing nodes (N_a) required within the VSN to perform the desired tasks.

b) Heterogeneous Networks: Now consider a heterogeneous network that has to perform the same $t = 2$ tasks described earlier for homogeneous network i.e. face detection and occluded target surveillance and tracking. Suppose the network consists of the following $k = 3$ sensor classes: a low resolution sensor with $P \times Q = 320 \times 240$, a medium resolution sensor with $P \times Q = 640 \times 480$ and a high resolution sensor with $P \times Q = 2304 \times 1728$. The horizontal FoV ($\theta_h$) and the vertical FoV ($\theta_v$) are assumed to be same for $k = 3$ sensor classes and

are given by $\theta_h = 48.39°$ and $\theta_v = 37.25°$. Again, object pixel occupancy based characterisation method is considered with the same O_{po} and A values described earlier for the homogeneous network scenario.

In this case, the matrix $\mathbf{R}_c$ which is the feedback to projection modelling and provides the estimated sensing range of $k = 3$ sensor classes for $t = 2$ tasks is found to be,

$$\mathbf{R}_c = \begin{bmatrix} R_{c11} & R_{c12} \\ R_{c21} & R_{c22} \\ R_{c31} & R_{c32} \end{bmatrix} = \begin{bmatrix} 4.05 & 44.60 \\ 8.09 & 89.21 \\ 29.12 & 321.16 \end{bmatrix}$$

In this design solution, $\mathbf{R}_c$ is calculated in *metres*.

Using the hard decision based sensing range selection scheme, if $t = 2$ tasks have to be performed by a sensor class, the chosen range should be,

$$\begin{aligned} r(1) &= \min\{R_{c11}, R_{c12}\} \\ r(2) &= \min\{R_{c21}, R_{c22}\} \\ r(3) &= \min\{R_{c31}, R_{c32}\} \end{aligned}$$

This can be represented in the form of vector $\mathbf{r}$ and the values are calculated to be,

$$\mathbf{r} = \begin{bmatrix} r(1) \\ r(2) \\ r(3) \end{bmatrix} = \begin{bmatrix} 4.05 \\ 8.09 \\ 29.12 \end{bmatrix}$$

The hard decision based approach does not take advantage of having multiple sensor classes within the network. The proposed soft decision based range selection

scheme for heterogeneous networks (presented in Algorithm 3.2) maximises the sensing range for $\left(k - \left\lceil\sqrt{k}\right\rceil\right)$ sensor classes. The range $\mathbf{r}$ computed using the soft decision based scheme is,

$$\mathbf{r} = \begin{bmatrix} r(1) \\ r(2) \\ r(3) \end{bmatrix} = \begin{bmatrix} 4.05 \\ 8.09 \\ 175.14 \end{bmatrix}$$

This vector $\mathbf{r}$ is also a feedback to projection modelling. It can be observed from the comparison of $\mathbf{r}$ computed using the hard and soft decision based approaches that the latter maximises the sensing range for $\left(k - \left\lceil\sqrt{k}\right\rceil\right)\Big|_{k=3}$ sensor classes. The hard decision based approach computed $r(3)$ to be $29.12m$ whereas, the soft decision based scheme calculated $r(3)$ to be $175.14m$. This shows that the soft decision based scheme maximised the range approximately 6 times compared to the hard decision based approach.

3.4.7 Sensing Node's 3D Coverage Volume Estimation

The sensing node's 3D coverage volume can be calculated from Equation 3.3 by substituting θ_h and θ_v for a suitable sensing range R. Assuming a fixed sensing range $R = 5$, Figure 3.11 shows the sensing node's 3D coverage volume in m^3 for different horizontal and vertical FoV angles. It is observed from Figure 3.11 that the utilisation of a visual sensing node with larger horizontal and vertical FoV angles results in a larger portion of the 3D sensing environment within the sensing node's FoV. As this study is focused on the selection of a suitable sensing range for FoV characterisation, the 3D coverage volume for a range of values of R is plotted in Figures 3.12 and 3.13 with fixed horizontal and vertical FoV angles. The results demonstrate that increasing the sensing range leads to a larger volume of the 3D environment within a visual sensing node's FoV. This justifies the requirement of a FoV characterisation framework that can provide suitable range of a visual

sensing node resulting in an improved 3D coverage, without compromising on the desired reliability. As observed in the previous section, the proposed soft decision based sensing range selection scheme maximises the sensing range, subsequently, enhancing the 3D coverage of a visual sensing node.

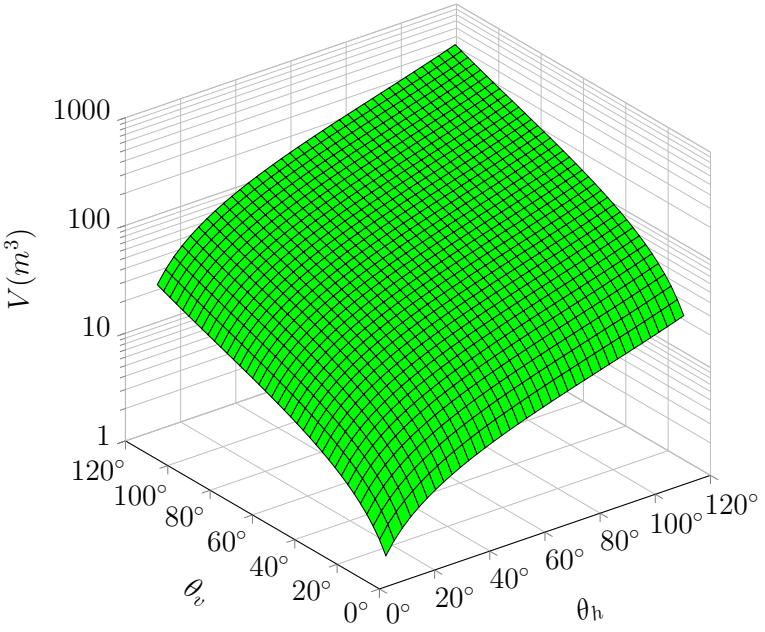

FIGURE 3.11: 3D volume within a visual sensing node's FoV for range $R = 5m$

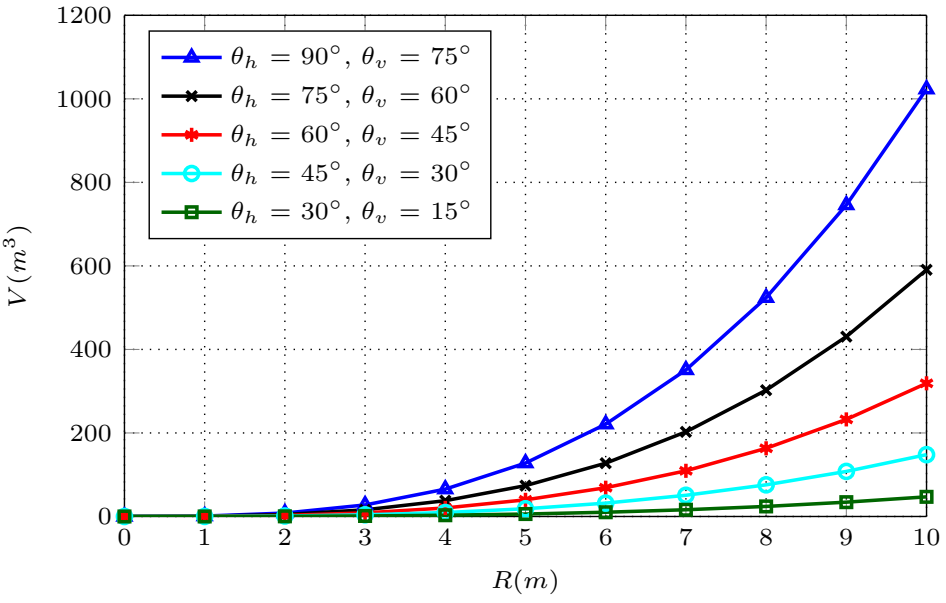

FIGURE 3.12: 3D volume within a visual sensing node's FoV, $\theta_h \neq \theta_v$

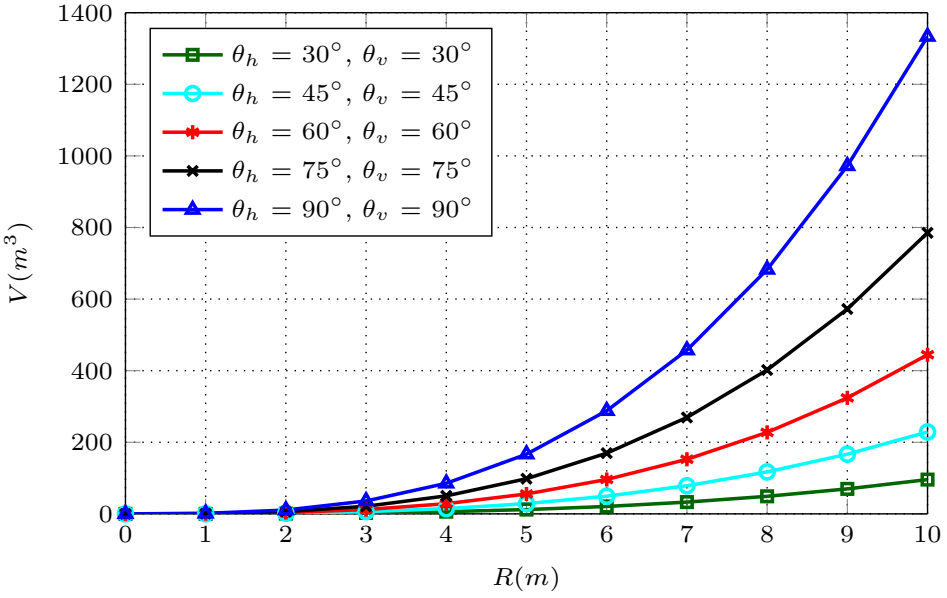

FIGURE 3.13: 3D volume within a visual sensing node's FoV, $\theta_h = \theta_v$

In the case of homogeneous sensor networks, feedback R_c is utilised for projection modelling. Considering the homogeneous network design solution presented in Section 3.3 and using Equation 3.3, the chosen sensing range $R_c = 8.09m$ leads to 3D coverage volume $106.90m^3$.

On the other hand, heterogeneous networks require feedback $\mathbf{R}_c$ and $\mathbf{r}$ for projection modelling. Considering the heterogeneous network design solution presented in Section 3.3 and using Equation 3.3, feedback $\mathbf{R}_c$ leads to the following required 3D scene coverage $\mathbf{V}_c$ (in m^3) of $k = 3$ sensor classes to perform $t = 2$ tasks.

$$\mathbf{V}_c = \begin{bmatrix} 13.41 & 1.79 \times 10^4 \\ 106.90 & 1.43 \times 10^5 \\ 4.99 \times 10^3 & 6.69 \times 10^6 \end{bmatrix}$$

After calculating $\mathbf{r}$ using the hard decision based scheme, the chosen 3D scene coverage $\mathbf{v}$ (in m^3) of $k = 3$ sensor classes is found to be,

$$\mathbf{v} = \begin{bmatrix} 13.41 \\ 106.90 \\ 4.99 \times 10^3 \end{bmatrix}$$

Similarly, after calculating $\mathbf{r}$ using the soft decision based scheme, the chosen 3D scene coverage $\mathbf{v}$ (in m^3) of $k = 3$ sensor classes is found to be,

$$\mathbf{v} = \begin{bmatrix} 13.41 \\ 106.90 \\ 1.08 \times 10^6 \end{bmatrix}$$

3.4.8 Adaptive Task Classification

In the proposed adaptive task classification scheme for heterogeneous networks, upto $\left\lceil \sqrt{k} \right\rceil$ sensor classes are assigned a certain task. Considering the heterogeneous design solution presented in Section 3.3, $\left\lceil \sqrt{k} \right\rceil \Big|_{k=3}$ evaluates to the allocation of 2 sensor classes for each task. The proposed scheme utilises $\mathbf{V}_c$ and $\mathbf{v}$ (calculated from the soft decision based sensing range selection scheme) for task classification.

In order to analyse the proposed task classification scheme, four different cases are compared. These are being hard decision based approach without $\left\lceil \sqrt{k} \right\rceil$ upperbound (case 1), hard decision based approach with $\left\lceil \sqrt{k} \right\rceil$ upperbound (case 2), soft decision based approach without $\left\lceil \sqrt{k} \right\rceil$ upperbound (case 3) and soft decision based approach with $\left\lceil \sqrt{k} \right\rceil$ upperbound (case 4). Case 1 utilises hard decision based $\mathbf{v}$ but does not impose a limit on the number of sensor classes performing a certain task i.e. the upper bound $\left\lceil \sqrt{k} \right\rceil$ is not considered. In case 2, hard decision based $\mathbf{v}$ as well as the upper bound $\left\lceil \sqrt{k} \right\rceil$ are utilised for task classification. Case 3 utilises soft decision based $\mathbf{v}$ but does not consider the upper bound $\left\lceil \sqrt{k} \right\rceil$. Finally,

Case 4 takes advantage of the soft decision based scheme for sensing range selection and also considers the upper bound $\left\lceil \sqrt{k} \right\rceil$ for task classification. Table 3.7 summarises the task classification results for these cases where 'Task I' refers to face detection, 'Task II' refers to occluded targets surveillance and tracking, '1' refers to an allocated task and '0' refers to an unallocated task.

TABLE 3.7: A comparison of task classification using four different cases for a network consisting of $k = 3$ sensor classes performing $t = 2$ tasks

Sensor class	**Sensor type**	**Hard decision based approach without $\left\lceil \sqrt{k} \right\rceil$ upper bound (case 1)**		**Hard decision based approach with $\left\lceil \sqrt{k} \right\rceil$ upper bound (case 2)**		**Soft decision based approach without $\left\lceil \sqrt{k} \right\rceil$ upper bound (case 3)**		**Soft decision based approach with $\left\lceil \sqrt{k} \right\rceil$ upper bound (case 4)**	
		Task I	**Task II**	**Task I**	**Task II**	**Task I**	**Task II**	**Task I**	**Task II**
1	low resolution	1	1	0	0	1	1	1	0
2	medium resolution	1	1	1	1	1	1	1	1
3	high resolution	1	1	1	1	0	1	0	1

Case 1 for task classification leads to a trivial solution where each sensor class has to perform every single desired task. Clearly, this is not a desired solution for VSN design. Although, case 2 provides a better solution as compared to case 1, it totally neglects the sensor class 1 by not allocating even a single task. It can be noticed that the task classification solution from case 2 will always neglect $\left(k - \left\lceil \sqrt{k} \right\rceil\right)$ sensor classes due to the hard decision. The solution obtained from case 3 is somewhere between the solutions of case 1 and case 2. Utilising the proposed soft decision based approach with $\left\lceil \sqrt{k} \right\rceil$ upper bound suggested in the proposed framework for task classification, case 4 leads to a promising solution. It is clear from the allocation results that this case provides optimised and the most suitable solution to the task classification problem by utilising all the sensor classes intelligently. An analysis of the energy efficiency of these cases is presented

in the following section which will further justify the superiority of case 4 over its counterparts.

3.4.9 Energy Efficiency of the Proposed Framework

Consider a visual sensor network that requires N_a active nodes to cover an area of size $100 \times 100m^2$. The number of nodes N_a required to be active depends on: the chosen sensing range R_c for homogeneous networks, or the chosen sensing range $\mathbf{r}$ of k sensor classes for heterogeneous networks. Figure 3.14 shows the number of nodes N_a required to be active within the VSN for different sensing range values. The results demonstrate that as the sensing range increases, the number of sensing nodes required to be active within the region under VSN coverage reduces. The results are logical and justify the significance of the proposed FoV characterisation framework to maximise the range of visual sensing nodes while satisfying the application design criteria, thus, reducing the number of nodes required to be active within the network.

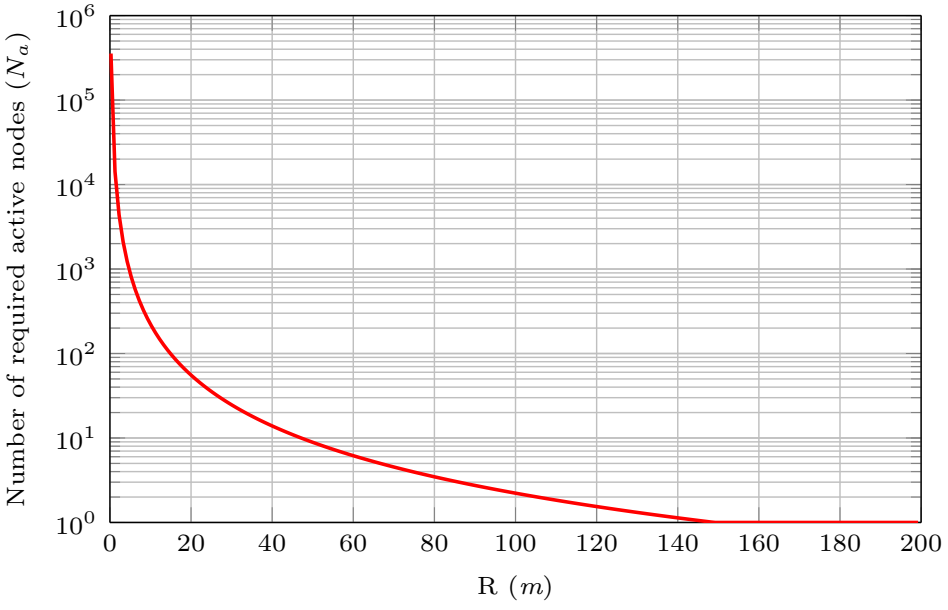

FIGURE 3.14: The number of required active nodes N_a to cover an area of size $100 \times 100m^2$

Furthermore, in order to validate the proposed framework which provides optimised energy consumption within certain desired confidence bounds, an energy-measurement testbed employed in [183] is considered. Each visual node within the testbed consists of a multimedia subsystem and a radio subsystem. The testbed's parameters are listed in Table 3.8.

TABLE 3.8: Energy-measurement testbed's parameters

Parameter	**Value**
Image acquisition cost	5.00 x 10^{-3} J
Initialising cost (JPEG)	1.40 x 10^{-2} J
Overall JPEG acquisition cost	1.90 x 10^{-2} J
Transmission cost	2.20 x 10^{-7} J/bit
Receiving cost	2.92 x 10^{-6} J/bit

The parameters given in [183] are used to measure the energy consumption of visual sensing nodes as the parameter values are obtained after substantial dedicated energy-measurement experimentation. Moreover, the parameter values are obtained such that a reliable data transmission platform can be facilitated. The consistency of the parameter values over different activation intervals has also been verified.

Suppose E_{Acq}, E_{Tx} and E_{Rx} denote the energy consumption of a single visual node to acquire, transmit and receive a single image frame respectively. Consider a scenario where each node within the VSN acquires, transmits and receives one image frame, the overall acquisition, transmission or receiving cost is given by,

$$\tilde{E}_q = N_a \times E_q \quad ; \quad q \in \{Acq, Tx, Rx\} \tag{3.41}$$

The total energy consumption within the VSN will be,

$$E_c = N_a \times (E_{Acq} + E_{Tx} + E_{Rx}) \tag{3.42}$$

The energy efficiency of the proposed framework for both homogeneous and heterogeneous networks is discussed in the following sections.

i) Homogeneous Networks

Consider a homogeneous network for a surveillance application that utilises sensors with the following parameters: $P \times Q = 320 \times 240$, $\theta_h = 48.39°$ and $\theta_v = 37.25°$. A comparison of image acquisition, transmission and receiving costs for different sensing range values is shown in Figure 3.15.

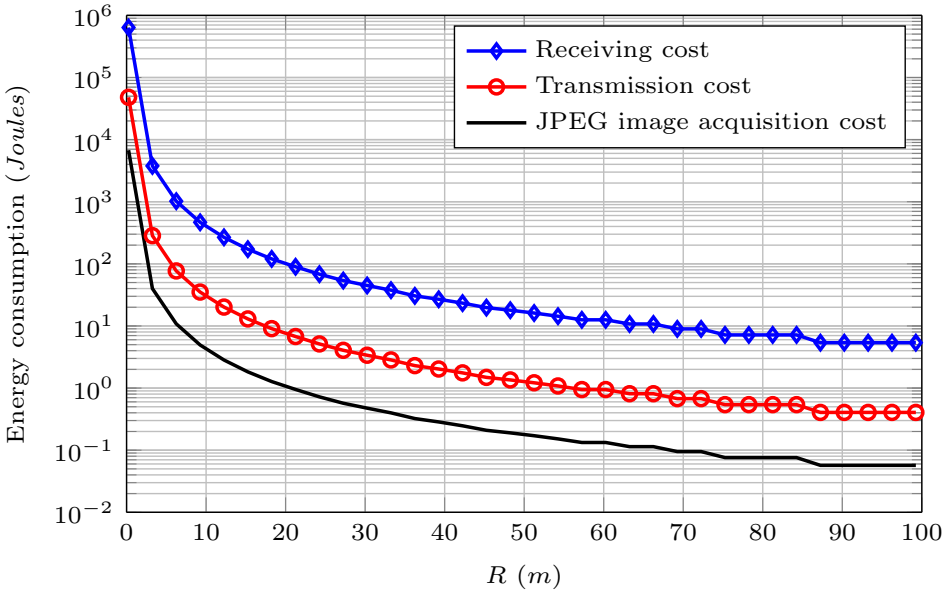

FIGURE 3.15: A comparison of JPEG image acquisition, transmission and receiving cost for different sensing range values

The results demonstrate that the reception and transmission of visual data requires much more energy compared to the energy consumed during the acquisition phase. This problem is addressed in the next chapter and a self-reconfiguration scheme is developed to reduce the redundant visual content prior to transmission in a VSN. It is also found from the results that increasing the sensing range results in less number of required active nodes leading to reduced energy consumption. However, if the sensing range goes beyond a certain threshold, the sensor may not provide accurate feature descriptors and may lead to miss detections. In that case, the

proposed framework can be used for application-aware sensing range estimation during the VSN design and calibration process. It maximises the spatial coverage leading to the reduced energy consumption configuration without compromising on the desired accuracy. Moreover, reducing the energy consumption will prolong the network's lifetime.

Table 3.9 lists the estimated sensing range for various applications based on certain criteria along with the number of required active nodes N_a and the total energy consumption E_c within the VSN.

TABLE 3.9: Application-aware sensing range estimation and energy consumption

Application	**Characterisation method**	**Criteria**	**Range and energy**	**Energy savings**	**Network lifetime**
Face detection [171]	object pixel occupancy based	$O_{po} \geq 315$	$R_1 = 4.05m$ $N_a = 1355$ $E_c = 2.64$ kJ	–	LT
Feature extraction & size estimation	estimation error based	$\lvert\varepsilon_d\rvert \leq 6\%$	$R_2 = 6.76m$ $N_a = 487$ $E_c = 948.78$ J	64.06%	LT × 2.78
Image transmission using IEEE 802.15.4a [182]	PSNR based	$\text{PSNR}_{\text{dB}} \geq 30$ dB	$R_3 = 9.55m$ $N_a = 244$ $E_c = 475.36$ J	81.99%	LT × 5.55
Occluded targets surveillance & tracking [172]	object pixel occupancy based	$O_{po} \geq 25$	$R_1 = 44.60m$ $N_a = 12$ $E_c = 23.38$ J	99.11%	LT × 112.92

The applications are listed in descending order of their energy consumption. The results show that the application-aware proposed FoV characterisation framework estimates the sensing range based on the desired criteria to maximise the spatial-coverage within the VSN and thus optimises the energy consumption. Suppose the lifetime of a VSN employing face detection algorithm is LT. It is evident from the results that the proposed approach leads to increased VSN lifetime for other applications in comparison with the first. The proposed framework has also optimised the number of required active nodes N_a leading to reduced energy consumption

configuration. The LOTS method proposed in [172] for occluded targets surveillance and tracking finds its applications in military where energy efficiency is highly desirable. As shown in the results, utilising the proposed approach with LOTS has resulted in optimised energy consumption. Hence, the application-aware sensing range estimation from the proposed approach makes it suitable for a wide range of applications and it can be utilised to design and calibrate an energy efficient VSN.

ii) Heterogeneous Networks

Heterogeneous networks provide much more flexibility to the design engineer compared to the homogeneous networks due to the presence of different types of sensing nodes within the network. The analysis of energy efficiency presented in the previous section for homogeneous networks considered the design solution for four different applications. For heterogeneous networks, suppose $\hat{t}$ represent the number of tasks allocated to a sensing class; the task classification solutions obtained from four different cases given in Table 3.7 are used to analyse the energy efficiency of the proposed framework and the results are presented in Table 3.10.

It is found from the results that the proposed soft decision based sensing range selection scheme with $\left\lceil \sqrt{k} \right\rceil$ upper bound for task classification maximised the spatial coverage and allocated tasks efficiently that lead to the minimum energy consumption configuration. The proposed task classification solution presented in case 4 minimises the energy consumption to 8 kJ and doubles the network's lifetime compared to case 1. The energy savings with the solution presented in case 4 compared to case 1, case 2 and case 3 are 49.8%, 25.0% and 24.8% respectively.

TABLE 3.10: Analysis of the energy efficiency of the proposed framework for heterogeneous network design

Sensor class	**Sensor type**	**Hard decision based approach without $\lceil\sqrt{k}\rceil$ upper bound (case 1)**	**Hard decision based approach with $\lceil\sqrt{k}\rceil$ upper bound (case 2)**	**Soft decision based approach without $\lceil\sqrt{k}\rceil$ upper bound (case 3)**	**Soft decision based approach with $\lceil\sqrt{k}\rceil$ upper bound (case 4)**
1	low resolution	r(1) = 4.05m $\hat{t} = 2$ $N_a = 1355$ $E_c = 5.28$ kJ	r(1) = 4.05m $\hat{t} = 0$ $N_a = 0$ $E_c = 0$ J	r(1) = 4.05m $\hat{t} = 2$ $N_a = 1355$ $E_c = 5.28$ kJ	r(1) = 4.05m $\hat{t} = 1$ $N_a = 1355$ $E_c = 2.64$ kJ
2	medium resolution	r(2) = 8.09m $\hat{t} = 2$ $N_a = 340$ $E_c = 5.26$ kJ	r(2) = 8.09m $\hat{t} = 2$ $N_a = 340$ $E_c = 5.26$ kJ	r(2) = 8.09m $\hat{t} = 2$ $N_a = 340$ $E_c = 5.26$ kJ	r(2) = 8.09m $\hat{t} = 2$ $N_a = 340$ $E_c = 5.26$ kJ
3	high resolution	r(3) = 29.12m $\hat{t} = 2$ $N_a = 27$ $E_c = 5.40$ kJ	r(3) = 29.12m $\hat{t} = 2$ $N_a = 27$ $E_c = 5.40$ kJ	r(3) = 175.14m $\hat{t} = 1$ $N_a = 27$ $E_c = 100$ J	r(3) = 175.14m $\hat{t} = 1$ $N_a = 27$ $E_c = 100$ J
Overall energy consumption		15.94 kJ	10.66 kJ	10.64 kJ	8 kJ
Network lifetime		LT	LT × 1.5	LT × 1.5	LT × 2

3.4.10 Analysis of System Failure

After the design process, the VSN is expected to perform tasks within a certain confidence bound. Let ζ^l to ζ^u be the dynamic PSNR range in dB for a particular application and $\delta_1 = \text{antilog}\left(-\frac{\zeta^u}{10}\right) - \text{antilog}\left(-\frac{\zeta^l}{10}\right)$ be the dynamic difference. Suppose λ_t denote the threshold for system quality assessment representing the desired PSNR in dB. The probability that a system with quality β (representing the achieved PSNR in dB) will fail to perform a certain task is derived as,

$$P(\lambda_t > \beta) = \begin{cases} 0, & \lambda_t < \beta \\ \dfrac{\delta_2}{\delta_1}, & \lambda_t \geq \beta \end{cases} \tag{3.43}$$

where $\delta_2 = \text{antilog}\left(-\frac{\lambda_t}{10}\right) - \text{antilog}\left(-\frac{\beta}{10}\right)$

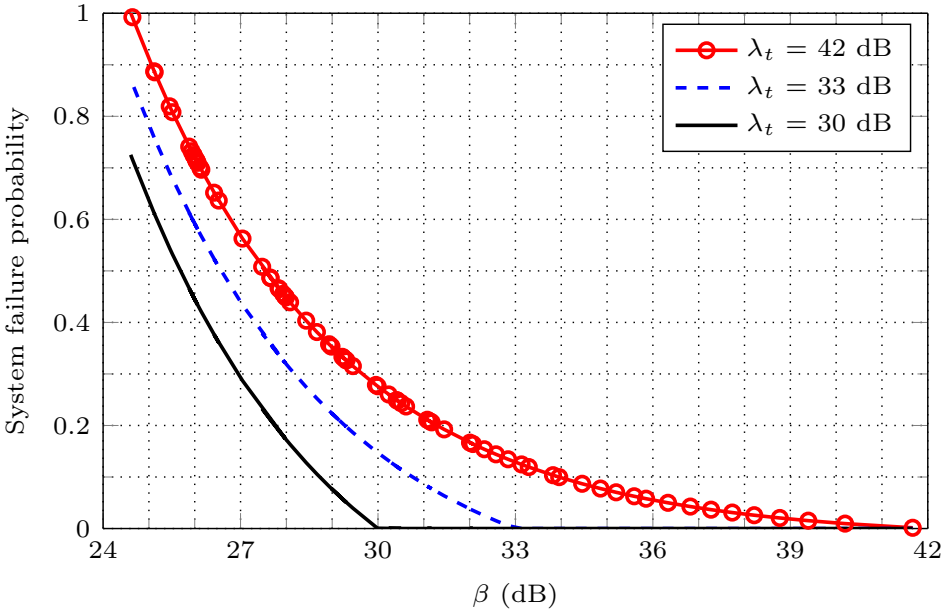

FIGURE 3.16: Analysis of failure for a system with target quality threshold λ_t and achieved quality β

Figure 3.16 shows an analysis of system failure for several values of target quality threshold λ_t and system's achieved quality β. It can be observed from the graph that the system failure probability is maximum when the target quality threshold λ_t and system's achieved quality β are at the opposite ends of the dynamic PSNR range. However, the system failure probability reduces when the target quality threshold λ_t and the system's achieved quality β lies between the dynamic range. Moreover, the results demonstrate that the failure probability minimises to zero when the system's achieved quality β is higher than the target quality threshold λ_t.

3.5 Summary

In this chapter, the issues of FoV characterisation and task classification for VSNs have been addressed. A novel framework for the FoV characterisation of both homogeneous and heterogeneous networks is proposed. The image captured by

a visual sensing node is found to be a function of the sensing range, the horizontal FoV and the vertical FoV. Therefore, accurate FoV parameters estimation and sensing range maximisation is paramount which can lead to a reduction in the total number of sensing nodes required to provide full coverage resulting in minimisation of overall network energy requirements. Initially, in the projection modelling phase, a suitable approach is selected to calculate the FoV parameters based on whether the network is homogeneous or heterogeneous. YCbCr colour space is chosen and features are extracted using global colour histogram. Next, for a given level of reliability, the proposed framework defines a criteria, referred to as the Field-of-View characterisation Criteria, which estimates the optimal sensing range of a visual sensing node. This chapter also proposes adaptive task classification and soft decision based sensing range selection schemes for heterogeneous networks. Without compromising on the desired reliability, the proposed adaptive task classification and soft decision based sensing range selection schemes provide an energy efficient solution for the maximisation of visual sensing node's spatial coverage resulting in an enhanced network lifetime.

The performance of the proposed FoV characterisation framework is evaluated through experiments and simulations. An energy model is presented which is utilised to observe the robustness of the framework. For any given application, it is noticed that the proposed solution for FoV characterisation enhances the spatial coverage, optimises the energy consumption and increases the lifetime in homogeneous networks. Furthermore, the configuration of heterogeneous network obtained by utilising the proposed FoV characterisation framework with the task classification and sensing range selection schemes for a surveillance application resulted in substantial energy savings compared to the trivial design solution. The energy efficiency of the proposed FoV characterisation framework demonstrates that it can be utilised during the network design and calibration phase to achieve an application-aware solution. An analysis of system failure probability is also presented which can assist the design engineers in predicting and minimising the

network failure probability.

The next chapter presents a QoI-aware self-reconfiguration scheme for dynamically selecting optimal configurations of visual sensing nodes in the network post-deployment phase. The distributed nature of the proposed reconfiguration scheme is obtained by introducing a learning strategy which involves the training and calibration of visual sensing nodes in the network pre-deployment phase.

Chapter 4

QoI-Aware Self-Reconfiguration Scheme

4.1 Introduction

In the network post-deployment phase, the main purpose of a visual sensing node is the transmission of visual data to the sink node; where the visual data provides information about a portion of the sensing environment within the sensing node's FoV. Due to energy and bandwidth constraints of visual sensing nodes, in-node energy conservation is one of the prime concerns in VSNs with wireless transceiving capability. The amount of visual data transmitted by a sensing node is characterised by local processing in the pre-transmission phase. Local processing can remove redundant features (by employing a suitable compression scheme) from the acquired visual data prior to transmission in VSNs leading to a reduced transmission energy cost and enhanced bandwidth efficiency. Nevertheless, such processing may deteriorate the quality of visual data; therefore, the reliability and energy efficiency trade-off must be analysed within the context of visual data optimisation. The analysis can be performed at a centralised unit (i.e. the sink node) or a distributed approach can be employed at each sensing node. The latter is

more suitable as compared to the former as it facilitates speedy decision making. Due to the dynamic nature of visual data, a static level of compression utilised within visual sensing nodes for feature redundancy removal may not be suitable. Therefore, each visual sensing node is required to be equipped with a distributed strategy to dynamically and independently select optimal configuration. The intelligence of a distributed strategy can be enhanced by developing a training and calibration process for sensing nodes to be employed by the design engineers in the network pre-deployment phase. Moreover, the suitability of visual data for a given application or a decision making process can be reflected by the QoI. Incorporating QoI-awareness in the self-reconfiguration strategy is expected to provide a reliable solution for visual data optimisation while enhancing the energy efficiency of visual sensing nodes.

This chapter provides an energy efficient solution to the visual sensing nodes' self-reconfiguration problem for resource constrained scenarios. A generalised reconfiguration model for visual sensing nodes is presented and various parameters that constitute the model are discussed. A learning strategy is devised to be employed by design engineers during the network pre-deployment phase. The proposed learning approach formulates a relationship between objects' variations within a sensing node's FoV, the level of detail in the acquired visual data and the effect of feature redundancy removal on that data. Moreover, incorporating QoI-awareness, an energy efficient scheme is proposed to reconfigure the nodes of a VSN for optimal performance in resource constrained scenarios. Depending on the application design criteria, the proposed self-reconfiguration scheme is capable of maintaining target QoI thresholds while achieving energy savings. The distributed nature of the proposed learning-aided reconfiguration scheme expedites the decision making process at visual sensing nodes. The robustness of the proposed reconfiguration scheme in terms of energy savings is evaluated by performing simulations and the results are presented in this chapter.

4.2 Visual Sensing Node Reconfiguration Model

Energy conservation is a primary issue within resource constrained VSNs which is expected to be achieved by dynamic self-reconfiguration, for example, by the realisation of the targeted QoI thresholds at each visual sensing node to fine tune its parameters. The self-reconfiguration model employed within each visual sensing node is illustrated in Figure 4.1.

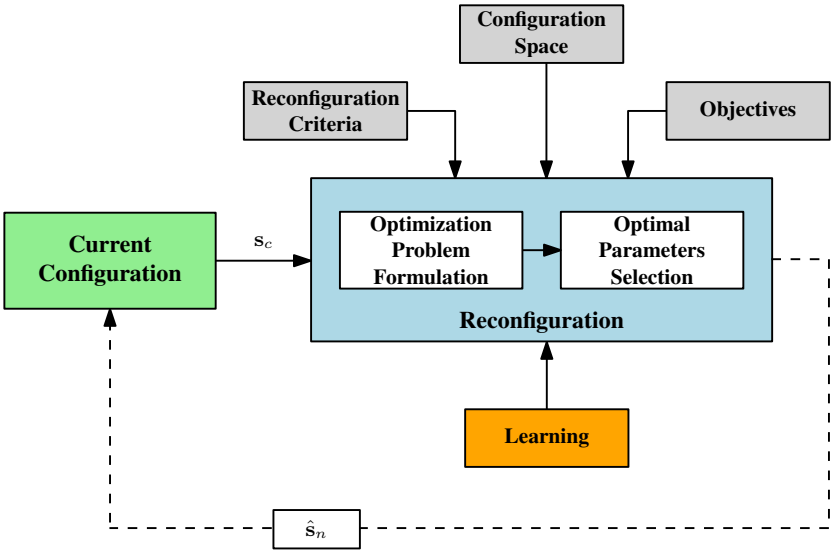

FIGURE 4.1: Visual sensing node self-reconfiguration model

Let $\mathcal{C}_c = \{C_1^c, C_2^c, C_3^c, \ldots, C_{\hat{n}}^c\}$ be the set of criteria provided by the design engineer for a given application and $\mathcal{C}_s = \{C_1^s, C_2^s, C_3^s, \ldots, C_{\tilde{n}}^s\}$ denote the set of parameters in the network configuration space. Suppose $\mathcal{C}_o = \{C_1^o, C_2^o, C_3^o, \ldots, C_{\breve{n}}^o\}$ represent the set of objectives to be achieved, for example, coverage maximisation, energy conservation, redundancy management and optimal task classification between visual sensing nodes. Based on particular reconfiguration criteria $\mathcal{C}_c$, the considered set of parameters $\mathcal{C}_s$ and the current configuration state $\mathbf{s}_c$; the process of reconfiguration takes place dynamically within a visual sensing node to obtain a new configuration state $\hat{\mathbf{s}}_n$ and achieve particular goals $\mathcal{C}_o$. The self-reconfiguration model depicted in Figure 4.1 incorporates a learning-assisted strategy for the selection of optimal parameters during the decision making process. The algorithms

for the management of nodes within VSNs can be classified into centralised and distributed categories. Distributed algorithms are preferred over the centralised ones due to the complexity and scalability issues of the latter [184]. Moreover, the distributed algorithms reduce bottlenecks and improve resilience against network failures as they do not rely on a single central node for decision making. Therefore, in this thesis, a distributed decision making strategy for reconfiguration is devised where each node takes the decision independently to accelerate the decision making process.

4.3 Proposed QoI-Aware Self-Reconfiguration Scheme

Suppose a region of interest with surface area A_T and volume V_T is to be monitored for a surveillance application. For this purpose, consider the deployment of a VSN which consists of $\mathcal{N}$ visual sensing nodes and one sink node. The placement of nodes within the region of interest can either be random or deterministic. Suppose each visual sensing node within the network is represented by $\{\mathrm{VS}_{\tilde{l}} | \tilde{l} = 1, 2, 3, \ldots, \mathcal{N}\}$. The sensing nodes $\mathrm{VS}_1, \mathrm{VS}_2, \mathrm{VS}_3, \ldots, \mathrm{VS}_{\mathcal{N}}$ capture images, represented by $\mathbf{I}_1, \mathbf{I}_2, \mathbf{I}_3, \ldots, \mathbf{I}_{\mathcal{N}}$ respectively, and process them for feature detection and object extraction. Moreover, the effective sensing range R of a visual sensing node for a given application can be estimated from the methodology presented in Chapter 3 and is assumed to be known.

The system model of the proposed QoI-aware self-reconfiguration scheme to enhance the energy efficiency of VSNs by achieving targeted threshold based optimisation is presented in Figure 4.2. In the proposed scheme, training and calibration take place in the pre-deployment phase which consists of training dataset selection, object appearance modelling, redundant feature removal, quality estimation and

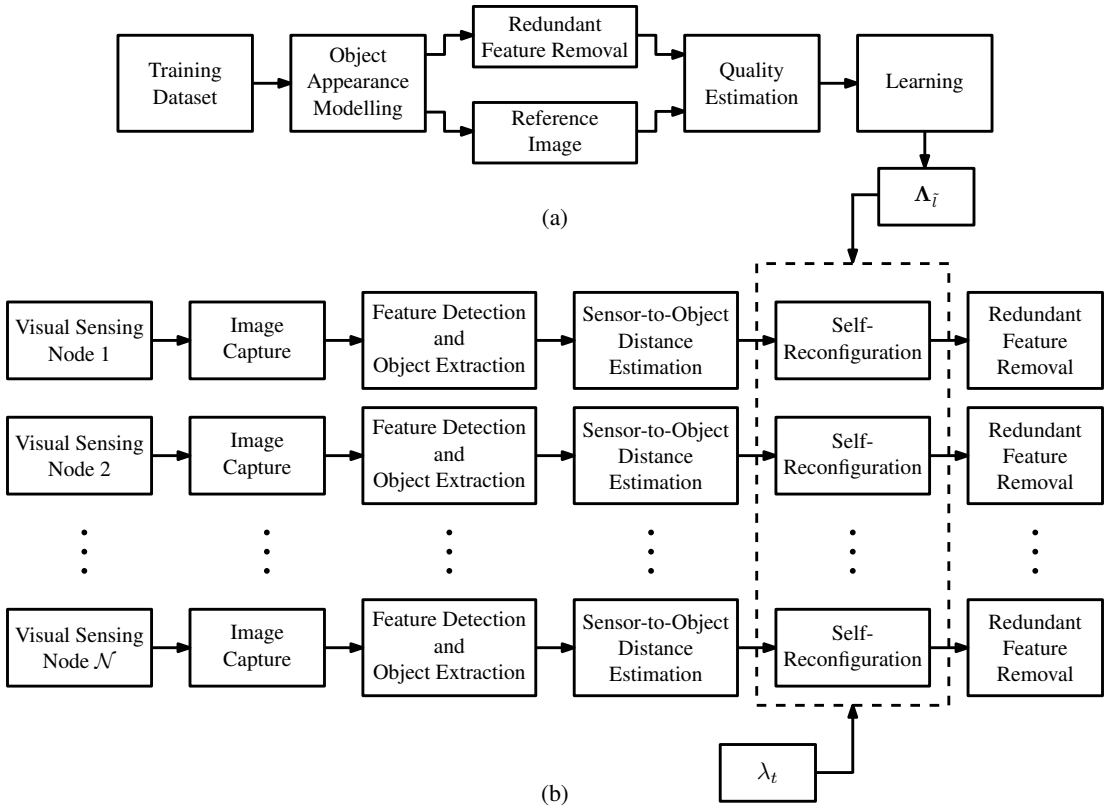

FIGURE 4.2: (a) Training and calibration process in the network pre-deployment phase (b) Proposed in-node processing model for QoI-aware self-reconfiguration within resource constrained VSNs to achieve targeted threshold based optimisation in the network post-deployment phase

learning, as shown in Figure 4.2a. In the post-deployment phase, the scheme incorporates image capture, feature detection and object extraction, sensor-to-object distance estimation, self-reconfiguration and redundant feature removal within the in-node processing model, as shown in Figure 4.2b.

4.3.1 Training and Calibration

In the proposed scheme, training and calibration take place only in the pre-deployment phase. Therefore, once this task is accomplished, the proposed framework does not require any further training in the post-deployment phase; consequently, facilitating the feasibility of the proposed scheme for resource constrained

scenarios. The training and calibration process is discussed in the following sections.

i) Dataset selection

In order to initiate the training and calibration process, a suitable dataset is selected. For a given application, the types of targets expected to be monitored within the FoV of visual sensing nodes have to be considered for such selection.

ii) Object appearance modelling

In the proposed scheme, reconfiguration is initiated based on the appearance of the object within a visual sensing node's FoV. As an object moves closer to a visual sensing node, its pixel occupancy within the captured image increases resulting in an increased number of redundant features. On the other hand, if the object moves away from a visual sensing node, the reduction in its pixel occupancy within the captured image leads to a reduced number of redundant features. Therefore, object appearance modelling plays a prominent role during training and calibration of visual sensing nodes.

Suppose $\mathbf{I}_r$ denote a training image of size $D^+ \times D^-$ captured from a reference distance $R_r \geq R_l$ containing only the object of interest; where R_l is the minimum possible distance to capture a suitably sharp image of the object of interest. Let R_d be the sensor-to-object distance such that $R_r < R_d \leq R_u$; where R_u is the maximum distance to capture an object of interest's suitably sharp image. For an object captured from sensor-to-object distance R_d, its appearance at the reference distance R_r can be modelled as,

$$\hat{\mathbf{I}}_r = \left(\mathbf{I}_r \downarrow D_d^{\pm}\right) \uparrow D_u^{\pm} \tag{4.1}$$

where $\downarrow$ denote the image down-sampling operator, $\uparrow$ represent the image up-sampling operator; $D_d^{\pm}$ and $D_u^{\pm}$ are the down-sampling and up-sampling factors given by,

$$\begin{aligned} D_d^{\pm} &= D^{\pm}\left(\frac{R_r}{R_d}\right) \\ D_u^{\pm} &= D^{\pm}\left(\frac{R_d}{R_r}\right) \end{aligned} \tag{4.2}$$

iii) Redundant feature removal

In resource constrained VSNs, energy conservation can be achieved through redundant feature removal, which leads to the minimisation of the transmission cost. Suppose α_c^l to α_c^u be the dynamic compression range; the possible compression ratios can be written as,

$$\mathbf{a}_c = \{\alpha_c^l, \alpha_c^l + \mathcal{S}_c, \alpha_c^l + 2\mathcal{S}_c, \ldots, \alpha_c^u\} \tag{4.3}$$

and the resulting bits per pixel values can be denoted by $\mathbf{a}_b$ as,

$$\mathbf{a}_b = \{\alpha_{b(l)} | l = 1, 2, 3, \ldots, l_m\} \tag{4.4}$$

where $\mathcal{S}_c$ is a positive scalar step-size, $l_m = \lfloor(\alpha_c^u - \alpha_c^l + \mathcal{S}_c)/\mathcal{S}_c\rfloor$ and $(\lfloor\ \rfloor)$ refers to the floor function.

iv) Quality estimation

As discussed earlier, redundant features can be removed to optimise energy consumption of the visual sensing nodes. However, such removal may affect the quality of the visual data. Therefore, the impact of redundant feature removal on the quality of the acquired visual data must be taken into consideration to achieve an

acceptable level of reliability for the given application. In the proposed scheme, PSNR is used as a quality metric, which is realised with the system's dynamic PSNR range in dB (ζ^l to ζ^u) to obtain the QoI index λ as,

$$\lambda = \frac{10 \log_{10} \left(\frac{\mathcal{A}^2}{\mathcal{M}} \right) - \zeta^l}{\zeta^u - \zeta^l} \tag{4.5}$$

where $\lambda \in [0, 1]$, $\lambda = 0$ refers to the QoI index of ζ^l, $\lambda = 1$ refers to the QoI index of ζ^u, $\mathcal{A}$ is the maximum possible pixel value in $\mathbf{I}_r$ and $\mathcal{M}$ represent the Mean Squared Error (MSE) given by,

$$\mathcal{M} = \frac{1}{3(D^+ D^-)} \sum_{x=1}^{D^+} \sum_{z=1}^{D^-} \left[I_r(x, z) - \hat{I}_r(x, z) \right]^2 \tag{4.6}$$

Suppose the set of possible sensor-to-object distances is given by $\hat{\mathbf{r}} = \{R_r, R_r + \mathcal{S}_r, R_r + 2\mathcal{S}_r, \ldots, R_u\}$ and let $\boldsymbol{\lambda} = \{\lambda_p \mid p = 1, 2, 3, \ldots, p_m\}$ be the respective quality metrics calculated using Equation 4.5; such that $\lambda_{(1)} \geq \lambda_{(2)} \geq \lambda_{(3)} \geq \ldots \geq \lambda_{(p_m)}$; where $\mathcal{S}_r$ is a positive scalar step-size and $p_m = \lfloor (R_u - R_r + \mathcal{S}_r)/\mathcal{S}_r \rfloor$. For a dynamic sensor-to-object range R_r to R_u and dynamic compression range α_c^l to α_c^u, the resulting quality metrics can be given by,

$$\boldsymbol{\Lambda} = \begin{bmatrix} \lambda_{1,1} & \lambda_{1,2} & \ldots & \lambda_{1,p_m} \\ \lambda_{2,1} & \lambda_{2,2} & \ldots & \lambda_{2,p_m} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{l_m,1} & \lambda_{l_m,2} & \ldots & \lambda_{l_m,p_m} \end{bmatrix} \tag{4.7}$$

where $\lambda_{(\cdot,1)} \geq \lambda_{(\cdot,2)} \geq \ldots \geq \lambda_{(\cdot,p_m)}$ and $\lambda_{(1,\cdot)} \geq \lambda_{(2,\cdot)} \geq \ldots \geq \lambda_{(l_m,\cdot)}$.

v) Learning

Let $\hat{t}_m$ be the set of $\tilde{t}$ training images used for learning and $\{\mathbf{\Lambda}_i | i = 1, 2, 3, \cdots, \tilde{t}\}$ be their respective quality metrics. Suppose $\mathbf{\Lambda}_{\tilde{l}}$ denote the Compressive Calibration Matrix (CCM) which serves as a reference to reflect the impact of feature redundancy removal on the quality of the acquired visual data for $\tilde{t}$ training images; and it is calculated by,

$$\mathbf{\Lambda}_{\tilde{l}} = \frac{1}{\tilde{t}} \sum_{i=1}^{\tilde{t}} \mathbf{\Lambda}_i \tag{4.8}$$

The training and calibration process terminates after learning of the CCM and it is assumed that CCM is known to each visual sensing node in the network for utilisation during its self-reconfiguration phases. The fidelity of CCM is quantified within a confidence bound $[e^-, e^+]$, measured in dB, given by,

$$e^{\pm} = \mathfrak{D}^{\pm}(\zeta^u - \zeta^l) \tag{4.9}$$

where $\mathfrak{D}^- = \min\left\{(\overline{\mathbf{\Lambda}_i - \mathbf{\Lambda}_{\tilde{l}}}) \middle| i = 1, 2, 3, \cdots, \tilde{t}\right\}$,
$\mathfrak{D}^+ = \max\left\{(\overline{\mathbf{\Lambda}_i - \mathbf{\Lambda}_{\tilde{l}}}) \middle| i = 1, 2, 3, \cdots, \tilde{t}\right\}$ and $(\bar{\cdot})$ represent the mean.

4.3.2 In-node Processing Model

The in-node processing model is employed by the proposed scheme during the network post-deployment phase to dynamically attain optimal configurations of visual sensing nodes that result in energy conservation while satisfying a desired reliability level. The in-node processing model in discussed in the following sections.

i) Image Capture

During the data acquisition phase, an image $\mathbf{I}_{\tilde{l}}$ of dimension $D^+ \times D^-$ is captured by each visual sensing node $\mathrm{VS}_{\tilde{l}}$.

ii) Feature detection and object extraction

Let O_b be an object, $\mathbf{S}_m$ be the object segmentation matrix and $\mathbf{S}_g$ be the segmented image. The probability of a pixel (x, z) in $\mathbf{I}_{\tilde{l}}$ belonging to the object of interest O_b can be given by,

$$\mathcal{D}(x,z) = \begin{cases} 1, & (x,z) \in O_b \\ 0, & \text{otherwise} \end{cases} \tag{4.10}$$

This can be accomplished by utilising an appropriate detection method depending on the given application; for example, [185], [186] can be considered for face detection, [172] can be used for human detection and vehicles can be detected using [187], [188]. The pixels probabilities from Equation 4.10 are indexed at their respective locations in the object segmentation matrix $\mathbf{S}_m$. Consider a target-driven approach where each node transmits the acquired data only if an object of area A is detected within its FoV. The decision is made based on the following criteria,

$$e_d = \begin{cases} 1, & \gamma_{\hat{p}} \geq \gamma_t \\ 0, & \text{otherwise} \end{cases} \tag{4.11}$$

where $\gamma_{\hat{p}}$ represent the number of pixels an object occupies and γ_t denote the detection threshold given by,

$$\gamma_t = \frac{A \times D^+ \times D^- \times \sin\theta_h \times \sin\theta_v}{4R^2\,(1-\cos\theta_h)\,(1-\cos\theta_v)} \tag{4.12}$$

The object of interest is extracted from $\mathbf{I}_{\tilde{l}}$ by image segmentation using the following equation,

$$\mathbf{S}_g = e_d \cdot \mathbf{I}_{\tilde{l}} \cdot \mathbf{S}_m \tag{4.13}$$

where $(\cdot)$ represent the dot product.

iii) Sensor-to-object distance estimation

The level of detail in the captured image $\mathbf{I}_{\tilde{l}}$ is a function of the sensor-to-object distance R_d. The object of interest is obtained after feature detection and extraction as image $\mathbf{S}_g$. Therefore, for an object of area A being monitored within the FoV of a visual sensing node, utilising the output of feature detection and extraction process, R_d is estimated prior to reconfiguration, as shown below,

$$R_d = \frac{1}{2}\sqrt{\frac{A \times D^{+} \times D^{-} \times \sin\theta_h \times \sin\theta_v}{\gamma_{\hat{p}}\left(1-\cos\theta_h\right)\left(1-\cos\theta_v\right)}} \tag{4.14}$$

iv) Self-reconfiguration

Suppose λ_t denote the target QoI in dB to be achieved for a given application and γ_b represent the level of compression employed. Based on the spatial location of an object within a sensing node's FoV, each visual sensing node is to be reconfigured dynamically by maximising the compression level for redundant feature removal while achieving the target QoI threshold. Let $\mathbf{M}$ be a matrix of dimension $l_m \times p_m$ and $M_{l,p}$ is assigned a value based on the following condition,

$$M_{l,p} = \begin{cases} 1, & \Lambda_{\tilde{l}(l,p)} \geq \hat{\lambda}_t \\ 0, & \text{otherwise} \end{cases} \tag{4.15}$$

where $\hat{\lambda}_t = (\lambda_t - \zeta^l)/(\zeta^u - \zeta^l)$.

Considering λ_t for a particular application, the Active Compressive Calibration Matrix (ACCM) $\mathbf{\Lambda}_a$ is calculated by,

$$\mathbf{\Lambda}_a = \mathbf{\Lambda}_{\tilde{l}} \cdot \mathbf{M} \tag{4.16}$$

and the respective transmission energy cost can be expressed by $\mathbf{E}$ as,

$$\mathbf{E} = \begin{bmatrix} E^t_{1,1} & E^t_{1,2} & \dots & E^t_{1,p_m} \\ E^t_{2,1} & E^t_{2,2} & \dots & E^t_{2,p_m} \\ \vdots & \vdots & \ddots & \vdots \\ E^t_{l_m,1} & E^t_{l_m,2} & \dots & E^t_{l_m,p_m} \end{bmatrix} \tag{4.17}$$

where $E^t_{l,p} = \left\{ \mathcal{L}(\alpha_{b(l)}) \mid l = 1, 2, \dots, l_m \right\}$ and $\mathcal{L}(\cdot)$ denote a function for energy cost calculation.

Within the context of the visual sensing node self-reconfiguration model depicted in Fig 4.1, the compression level γ_b forms the configuration space, target QoI threshold λ_t is the reconfiguration criterion and minimisation of the transmission energy cost $E^t_{l,p}$ is the ultimate objective. Hence, the following optimisation problem can be stated,

$$\begin{aligned} &\text{minimise} \quad E^t_{l,p} \\ &\text{subject to} \quad \beta = \lambda_t \end{aligned} \tag{4.18}$$

where β is the QoI delivered by a visual sensing node.

The proposed scheme for dynamic self-reconfiguration of a visual sensing node within a resource constrained network is described in Algorithm 4.1 to find optimal compression level $\gamma_b \in \mathbf{a}_b$ that solves the optimisation problem expressed in Equation 4.18.

Algorithm 4.1 Proposed dynamic self-reconfiguration scheme

Input:

The target QoI threshold λ_t, the set of possible sensor-to-object distances $\hat{\mathbf{r}}$, the compressive calibration matrix $\mathbf{\Lambda}_{\tilde{l}}$, the system's dynamic PSNR range (ζ^l, ζ^u) and the estimated sensor-to-object distance R_d.

Output:

The new configuration state $\hat{s}_n$ of a visual sensing node to achieve targeted QoI threshold with optimised energy consumption.

$\mathbf{M} \leftarrow \varnothing$

$\hat{\lambda}_t = (\lambda_t - \zeta^l)/(\zeta^u - \zeta^l)$.

for $p \leftarrow 1$ **to** p_m **do**

  if $\Lambda_{\tilde{l}(1,p)} \geq \hat{\lambda}_t$ **then**

    $\mathbf{M} \leftarrow [\mathbf{M} \;\; \mathbf{1}]$

    where $\mathbf{1}$ is a $l_m \times 1$ all-ones vector

  else

    $\mathbf{M} \leftarrow [\mathbf{M} \;\; \mathbf{0}]$

    where $\mathbf{0}$ is a $l_m \times 1$ all-zeros vector

  end if

end for

$\mathbf{\Lambda}_a \leftarrow \mathbf{\Lambda}_{\tilde{l}} \cdot \mathbf{M}$

$\gamma_1 \leftarrow \arg\min_p \left[|R_d - \hat{r}_{(p)}| \right]; p = \{1, 2, 3, \ldots, p_m\}$

$\mathbf{t} \leftarrow \mathbf{\Lambda}_{a(1:l_m, \gamma_1)}$

Algorithm 4.1 (continued) Proposed dynamic self-reconfiguration scheme

15: $\gamma_2 \leftarrow \arg\min_{\hat{p}} \left[|\hat{\lambda}_t - t_{(\hat{p})}| \right]; \hat{p} = \{1, 2, 3, \ldots, l_m\}$

16: $\beta \leftarrow t_{(\gamma_2)}$

17: $E^t_{min} \leftarrow \mathcal{L}\left(\alpha_{b(\gamma_2)}\right)$

18: $\gamma_b \leftarrow a_{b(\gamma_2)}$

19: $\hat{s}_n \leftarrow \gamma_b$

20: **return** $\hat{s}_n$

In the proposed dynamic self-reconfiguration scheme presented in Algorithm 4.1, initially a matrix $\mathbf{M}$ is defined. Afterwards, the target QoI threshold λ_t and the system's dynamic PSNR range (ζ^l, ζ^u) are utilised to calculate $\hat{\lambda}_t$. Subsequently, for each value of p, such that $p = \{1, 2, 3, \ldots, p_m\}$ and $p_m = \lfloor (R_u - R_r + \mathcal{S}_r)/\mathcal{S}_r \rfloor$, the criterion $\Lambda_{\tilde{l}(1,p)} \geq \hat{\lambda}_t$ is employed. In case if the criterion is true, an all-ones vector of dimension $l_m \times 1$, where $l_m = \left\lfloor (\alpha^u_c - \alpha^l_c + \mathcal{S}_c)/\mathcal{S}_c \right\rfloor$, is appended to the matrix $\mathbf{M}$. Whereas, if $\Lambda_{\tilde{l}(1,p)} < \hat{\lambda}_t$ is true, an all-zeros vector of dimension $l_m \times 1$ is appended to the matrix $\mathbf{M}$. Afterwards, the Active Compressive Calibration Matrix $\mathbf{\Lambda}_a$ is obtained using $\mathbf{\Lambda}_a = \mathbf{\Lambda}_{\tilde{l}} \cdot \mathbf{M}$. Subsequently, function $\arg\min_p \left[|R_d - \hat{r}_{(p)}| \right]$ is utilised and the resulting value of p is stored in γ_1. Moreover, the first l_m rows from column γ_1 of the Active Compressive Calibration Matrix $\mathbf{\Lambda}_a$ are extracted and stored in vector $\mathbf{t}$. Next, function $\arg\min_{\hat{p}} \left[|\hat{\lambda}_t - t_{(\hat{p})}| \right]$ is employed, where $\hat{p} = \{1, 2, 3, \ldots, l_m\}$, and the resulting value of $\hat{p}$ is stored in γ_2. The value at index γ_2 is extracted from $\mathbf{t}$ and stored in β. Afterwards, the energy cost is calculated and the compression level γ_b is obtained and utilised as the new configuration state $\hat{s}_n$ of a visual sensing node to achieve targeted QoI threshold with optimised energy consumption.

v) Redundant feature removal

After the dynamic reconfiguration of a visual sensing node to obtain a new configuration state $\hat{s}_n$, the removal of redundant features from the segmented image $\mathbf{S}_g$ (containing the detected object) can be expressed as,

$$\mathbf{S}_r = \mathcal{G}(\mathbf{S}_g, \hat{s}_n) \tag{4.19}$$

where $\mathbf{S}_r$ is a matrix of dimension $\hat{D}^+ \times \hat{D}^-$ representing the reduced set of features to be transmitted to the sink node, such that $(\hat{D}^+ \times \hat{D}^-) < (D^+ \times D^-)$; and $\mathcal{G}(\cdot)$ is a function representing the compression method employed such as Discrete Wavelet Transform (DWT), Discrete Cosine Transform (DCT) etc.

The reconstruction takes place at the sink node and the reconstructed image $\hat{\mathbf{S}}_g$ can be given by,

$$\hat{\mathbf{S}}_g = \mathcal{G}^{-1}(\mathbf{S}_r) \tag{4.20}$$

where $\mathcal{G}^{-1}(\cdot)$ denote the inverse of $\mathcal{G}(\cdot)$.

4.4 Results and Analysis

This section demonstrates the performance of the proposed QoI-aware self-reconfiguration scheme for resource constrained VSNs. In order to evaluate the performance of the proposed QoI-aware self-reconfiguration scheme, the analysis was conducted using the MATLAB simulation platform on an Intel Core i5-4200U CPU @ 1.60GHz 2.30GHz with 8 GB RAM and 3 MB cache. The Long Distance Heterogeneous Face (LDHF) dataset [189, 190] is used for training and calibration and its specification is presented in Table 4.1. This particular dataset is

TABLE 4.1: Specification of the dataset used for training and calibration in the pre-deployment phase

Parameter	**Specification**
Imaging environment	Indoor and Outdoor
Lighting	Day-light
Actual Resolution	5184×3456
Resolution after down-sampling	512×512
Number of subjects	100
Indoor sensor-to-object distance	1 m
Number of indoor images	100
Outdoor sensor-to-object distance	60 m, 100 m, 150 m
Number of outdoor images	300

selected as it provides a variety of facial images captured with various sensor-to-object distances, hence making it suitable to be employed for the evaluation of the proposed framework. Even though the dataset contains images captured at both daytime and nighttime, the specification of images captured at daytime is considered. Keeping the capabilities of visual sensing nodes into consideration, the original resolution of the test images is found to be large, therefore, the test images are down-sampled to a resolution of 512×512.

An image subset extracted from the dataset containing cropped faces is shown in Figure 4.3. In order to conduct experiments for object appearance modelling, reference distance $R_r = 1\ m$, maximum distance $R_u = 40\ m$ and step size $\mathcal{S}_r = 1\ m$ is considered in the simulation model. Suppose $\hat{t}_s$ denote a subset consisting of 15 test images selected randomly from the dataset, i.e. $\hat{t}_s \subset \hat{t}_m$. Utilising the proposed object appearance modelling approach, the appearance of faces contained in test images at $R_r = 1$ is modelled for a range set of sensor-to-object

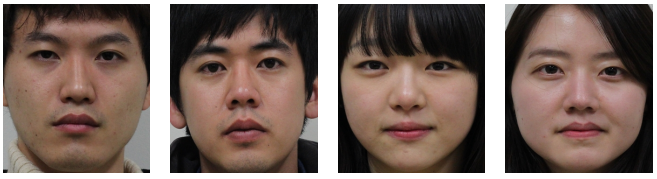

FIGURE 4.3: An image subset from the LDHF dataset containing cropped faces for $R_r = 1\ m$

distance $R_d \in [2, 40]$ and the resulting QoI index is shown in Figure 4.4. Due to the fact that moving an object away from a visual sensing node results in reduced object pixel occupancy, it can be observed from the results that increasing R_d results in QoI index reduction; where a higher value of QoI index refers to a higher QoI within the image. This justifies the requirement of incorporating QoI-awareness during the self-reconfiguration phase so that the proposed solution considers the estimated sensor-to-object distance R_d and is capable of achieving particular objectives. Moreover, it can be noticed that the decaying characteristics of the test images are identical and the QoI metrics are bounded within a range of $[-0.043, 0.048]$ from the mean.

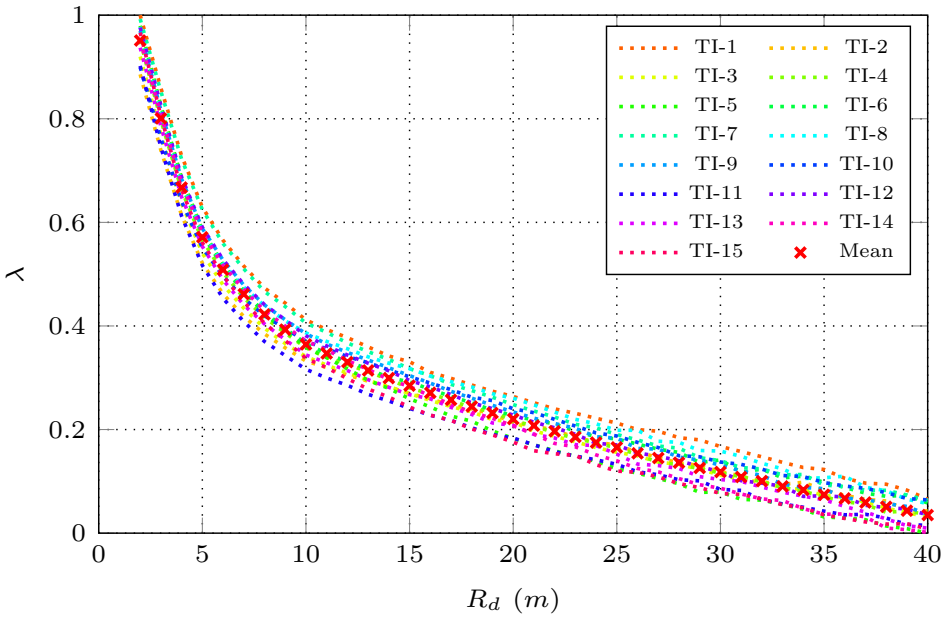

FIGURE 4.4: QoI index estimation for object appearance modelling utilising Test Images (TIs) from the LDHF dataset

In order to analyse the performance of the proposed framework, 2D-DWT and 2D-DCT are considered for the removal of redundant features from the images captured by visual sensing nodes. These methods are chosen for conducting experiments due to their suitability for visual sensing nodes [191] and utilisation in many existing schemes [28, 29, 55]. Nevertheless, the proposed framework is expected to support any compression scheme provided that its implementation

is feasible within visual sensing nodes. In the following discussion, even though 2D-DWT and 2D-DCT are treated together, they are required to be employed individually and only one redundancy removal method is required to be used for a given application. In the experiments, the dynamic compression range $\alpha_c^l = 2$ to $\alpha_c^u = 100$ and $\alpha_c^l = 2$ to $\alpha_c^u = 70$ is considered for 2D-DWT and 2D-DCT respectively, with step size $\mathcal{S}_c = 0.5$. This particular range is considered because the quality deteriorates beyond the upper limit and the resulting image may not be suitable for object detection. After redundancy removal and QoI index estimation, Equation 4.8 is used to obtain the compressive calibration matrix, which is shown in Figure 4.5 and Figure 4.6 for 2D-DWT and 2D-DCT respectively.

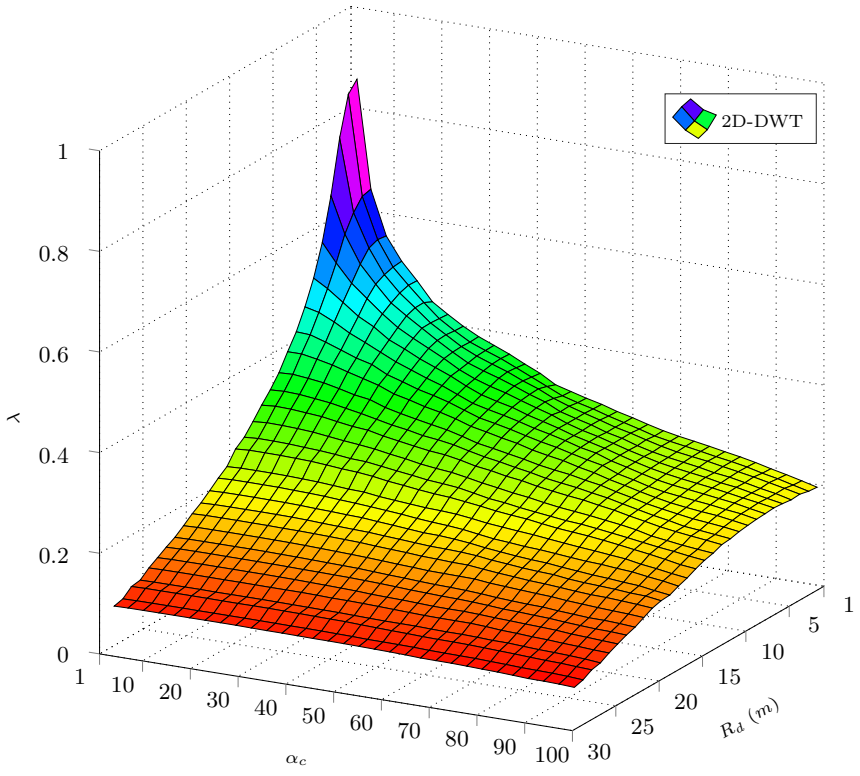

FIGURE 4.5: QoI index estimation after object appearance modelling and redundant feature removal with 2D-DWT to obtain compressive calibration matrix

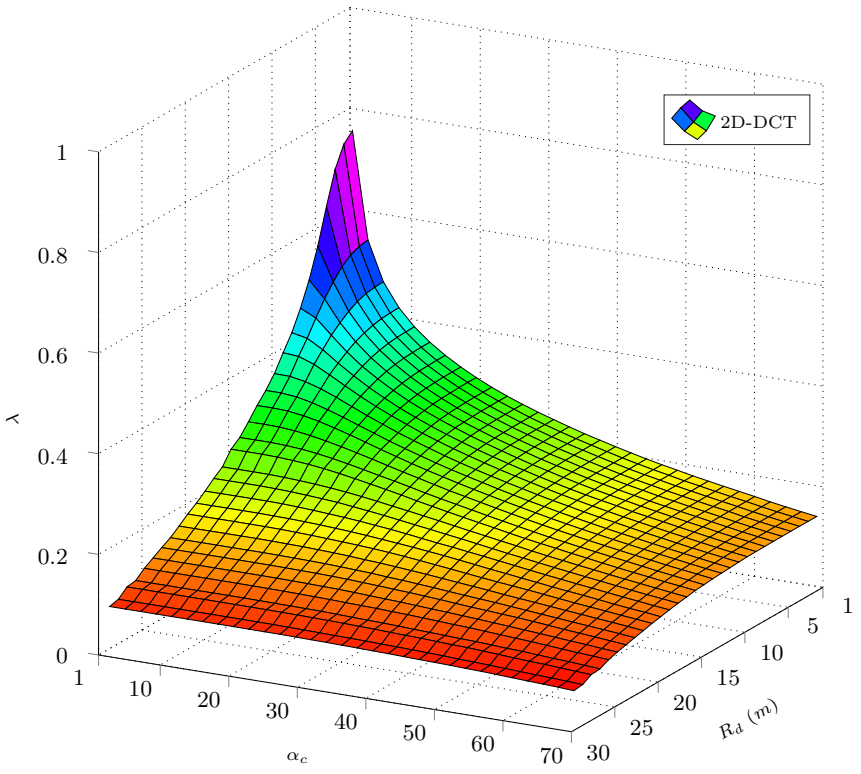

FIGURE 4.6: QoI index estimation after object appearance modelling and redundant feature removal with 2D-DCT to obtain compressive calibration matrix

It is observed from Figure 4.5 and Figure 4.6 that increasing the level of redundancy removal results in QoI index reduction for 2D-DWT and 2D-DCT respectively. Moreover, the results demonstrate that the decaying characteristics of the QoI index are similar for both 2D-DWT and 2D-DCT. This behaviour is logical as an increase in either the sensor-to-object-distance or the level of redundancy removal leads to a reduction in the visual data quality resulting in QoI index degradation. Therefore, the selection of inappropriate parameters for feature redundancy removal within visual sensing nodes is expected to affect the taget QoI achievement reliability. As energy conservation can be achieved by redundancy removal, CCM shown in Figure 4.5 and Figure 4.6 are used for the training and calibration of visual sensing nodes so that each node is dynamically self-reliant for the selection of an optimal configuration to minimise its energy consumption. Furthermore, it is found that compared to the DCT-aided scheme, the scheme supported by DWT is much more efficient in terms of the resulting QoI. This is due

to that fact that DWT, at high compression ratios, results in significantly higher compression efficiency compared to DCT [116]. Therefore, the overall transmission cost of the DCT-aided scheme is expected to be higher than the DWT-aided scheme.

The movements of objects within the FoV of visual sensing nodes are modelled randomly in the simulation. As the primary focus is the self-reconfiguration problem, it is assumed that after the capturing of an image by a visual sensing node in the post-deployment phase, an appropriate feature detection scheme can be used for object extraction from the image. It is also assumed that the sensor-to-object distance is estimated with a reasonable degree of reliability. Each time a target is detected, the configuration of a visual sensing node is obtained dynamically using Algorithm 4.1.

Suppose a VSN comprises of $\mathcal{N} = 100$ visual sensing nodes within a 3D sensing environment of size $50 \times 50 \times 10$ m^3. Uniform random deployment is considered and the visual sensing nodes are represented by $\mathrm{VS}_1, \mathrm{VS}_2, \mathrm{VS}_3, \ldots, \mathrm{VS}_{100}$. The horizontal FoV and vertical FoV values that have been obtained in Chapter 3, i.e. $\theta_h = 48.39^\circ$ and $\theta_v = 37.25^\circ$, are utilised for simulations in this chapter. Figure 4.7 illustrates the network model and shows the node distribution within the network. A homogeneous realisation of target QoI is considered i.e. target QoI thresholds are consistent throughout the network. In order to define connectivity between independent sensing nodes and a sink node, star topology is widely used in the existing literature [192–199]. Star topology provides a simplified solution to realise data communication in sensor networks. Therefore, in this study, a star-shaped topology in considered where the visual data collected by sensing nodes is transmitted to a sink node.

As in Chapter 3, the energy model given by Redondi *et al.* in [183] is used to calculate the energy consumption of visual sensing nodes. Redondi *et al.* considered star topology in their previous work on the energy consumption of VSNs [194]

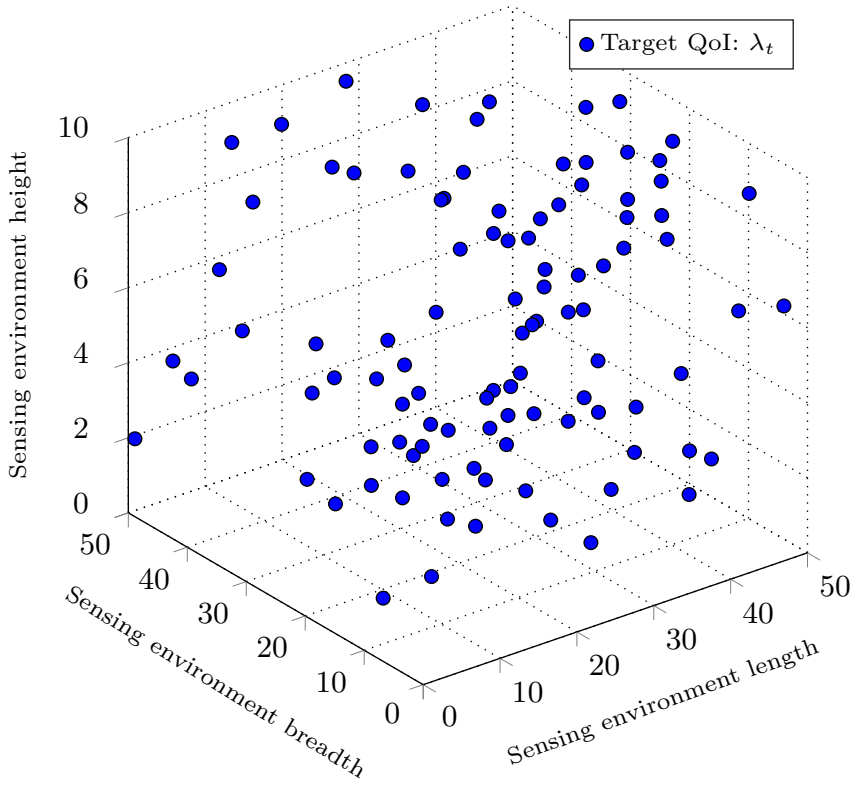

FIGURE 4.7: Visual sensing nodes distribution within the network

which justifies the selection of star topology for the chosen energy model. Since, the energy cost of communication is higher than the energy cost of data processing within VSNs [27–29], the optimisation of transmission energy cost is considered for reconfiguring visual sensing nodes and analysing the proposed scheme's energy efficiency. Let $\hat{E}_{tx}$ be the average energy cost incurred per node for the transmission of one image frame and $\hat{E}_c$ be the total transmission cost of $\mathcal{N}_t$ image frames, where an image frame is represented by $\mathcal{I}$. In order to demonstrate the energy efficiency of the proposed QoI-aware scheme compared to the conventional scheme, their transmission costs are observed for various target QoI thresholds. Unlike the proposed scheme, the conventional scheme does not dynamically tune the visual sensing nodes' parameters for redundant feature removal and utilises a constant value of γ_b to achieve target QoI threshold λ_t. The results obtained from the comparative analysis for $\mathcal{N}_t = 150$ image frames are presented in Figure 4.8 to Figure 4.11.

Let μ_p and μ_c denote the energy consumption averaged over the total number of image frame transmissions for the proposed and conventional schemes respectively. Figure 4.8 and Figure 4.9 illustrate the average transmission energy cost comparison for target QoI threshold $\lambda_t = 31$ dB with 2D-DWT and 2D-DCT respectively.

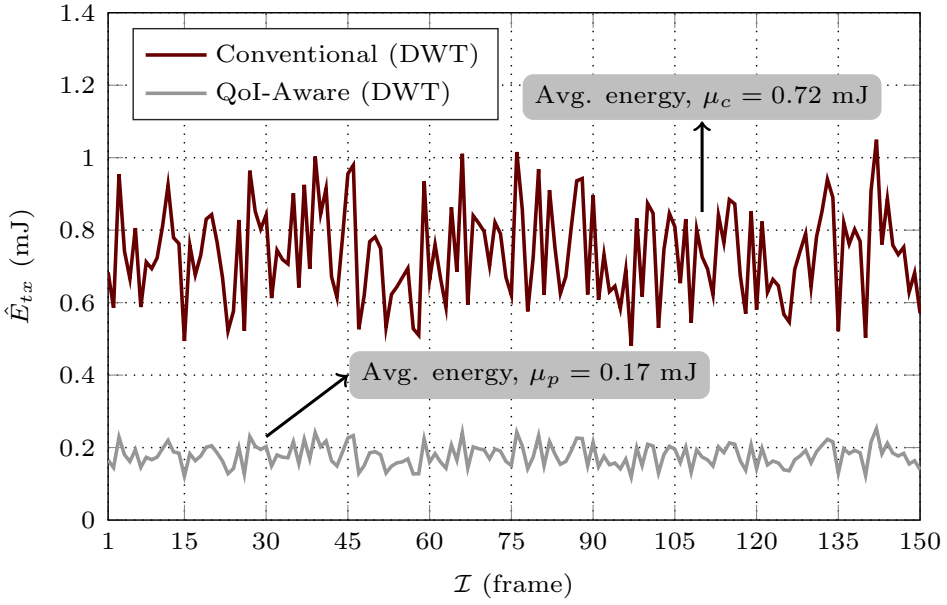

FIGURE 4.8: Comparison of the average energy cost incurred per node to transmit an image frame with the proposed QoI-aware scheme and conventional scheme for target QoI $\lambda_t = 31$dB with 2D-DWT

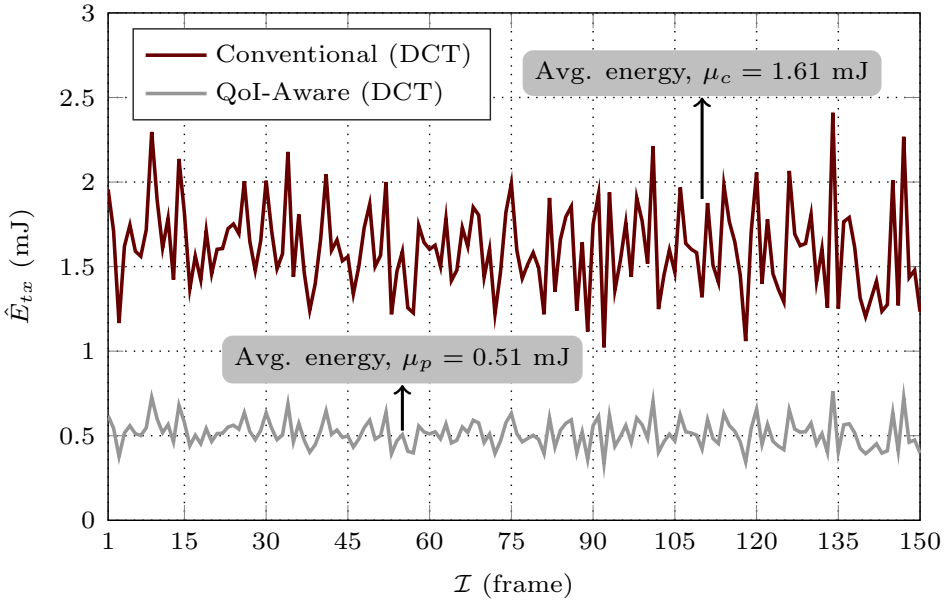

FIGURE 4.9: Comparison of the average energy cost incurred per node to transmit an image frame with the proposed QoI-aware scheme and conventional scheme for target QoI $\lambda_t = 31$dB with 2D-DCT

It is observed from Figure 4.8 and Figure 4.9 that average transmission energy consumption per frame of the proposed and conventional schemes to achieve homogeneous target QoI threshold $\lambda_t = 31$ dB is: 0.17 mJ and 0.72 mJ respectively with 2D-DWT; and 0.51 mJ and 1.61 mJ respectively with 2D-DCT. Thus the proposed scheme leads to energy conservation of 76.39% and 68.32% with 2D-DWT and 2D-DCT respectively. It is noticed from the comparison between the proposed and conventional schemes that the former leads to substantial energy savings compared to the latter as the redundancy removal process is more effective due to the lower target QoI threshold requirement.

Similarly, the average transmission energy cost comparison for target QoI threshold $\lambda_t = 37$ dB with 2D-DWT and 2D-DCT is presented in Figure 4.10 and Figure 4.11 respectively.

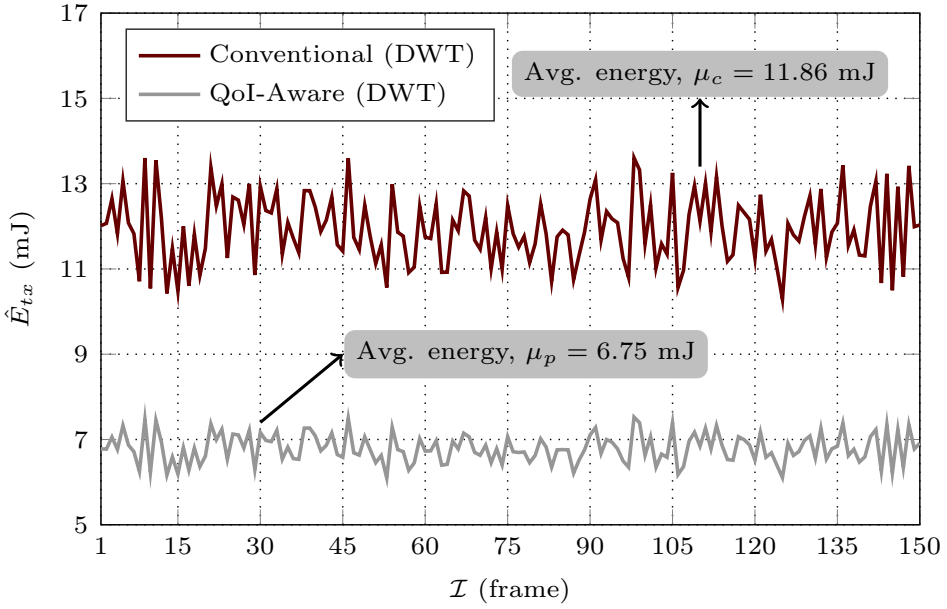

FIGURE 4.10: Comparison of the average energy cost incurred per node to transmit an image frame with the proposed QoI-aware scheme and conventional scheme for target QoI $\lambda_t = 37$dB with 2D-DWT

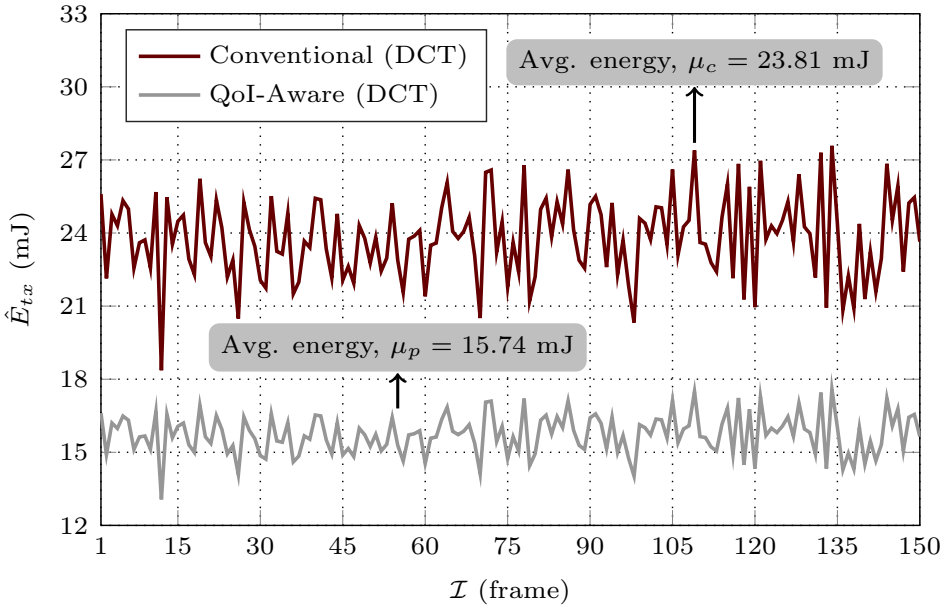

FIGURE 4.11: Comparison of the average energy cost incurred per node to transmit an image frame with the proposed QoI-aware scheme and conventional scheme for target QoI $\lambda_t = 37$dB with 2D-DCT

As shown in Figure 4.10 and Figure 4.11, to achieve target QoI threshold $\lambda_t = 37$ dB, the average cost of transmission energy per frame for the proposed and conventional schemes is: 6.75 mJ and 11.86 mJ respectively with 2D-DWT; and 15.74 mJ and 23.81 mJ respectively with 2D-DCT. Hence, energy savings of 43.09% and 33.89% are achieved with 2D-DWT and 2D-DCT respectively. In this comparison between the proposed and conventional schemes, the proposed scheme shows superior performance; however, in this scenario, the energy savings with the proposed scheme have reduced due to the higher QoI threshold requirement compared to the previous scenario. The results demonstrate that in scenarios where high energy savings are desirable, a lower target QoI threshold should be chosen to prolong the network lifetime.

Moreover, a comparison of the overall transmission cost of the proposed and conventional schemes per node to achieve a set of given homogeneous target QoI thresholds $\lambda_t \in [30, 40]$ for $\mathcal{N}_t = 150$ image frames with 2D-DWT and 2D-DCT is depicted in Figure 4.12 and Figure 4.13. In order to enhance readability, the performance of the proposed and conventional schemes for $\lambda_t = \{30, 31, 32\}$ is shown within Figure 4.12 and Figure 4.13 separately. The results demonstrate that for any given target QoI threshold, the proposed scheme minimises the transmission energy consumption with both 2D-DWT and 2D-DCT. Moreover, it is observed that the proposed DWT-aided scheme shows superior performance compared to the proposed scheme supported by the DCT. However, both DWT and DCT-based proposed schemes are capable of enhancing the energy efficiency of the visual sensing nodes. The proposed self-reconfiguration scheme outperforms the conventional scheme due to its dynamic realisation of the QoI which assists in selecting optimal configurations of visual sensing nodes leading to an energy efficient solution feasible for implementation in VSNs with strict constraints on available energy and bandwidth. Summarising the findings from Figure 4.8 to Figure 4.13, the results demonstrated that the energy savings with the proposed scheme are higher at lower QoI targets, thus making it suitable for resource constrained scenarios.

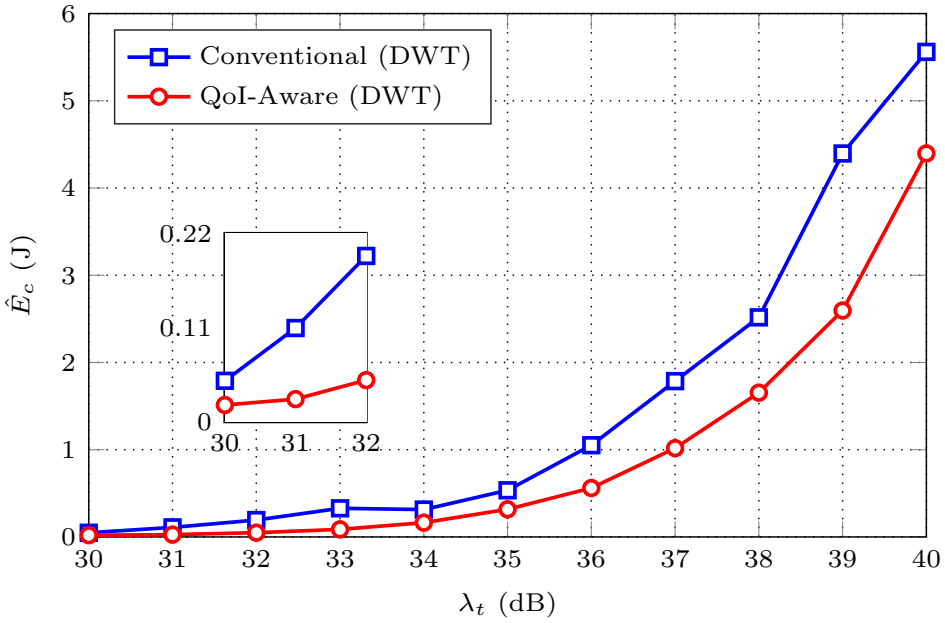

FIGURE 4.12: Comparison of the average energy cost incurred per node to transmit $\mathcal{N}_t = 150$ image frames with the proposed QoI-aware scheme and conventional scheme to achieve homogeneous target QoI thresholds with 2D-DWT

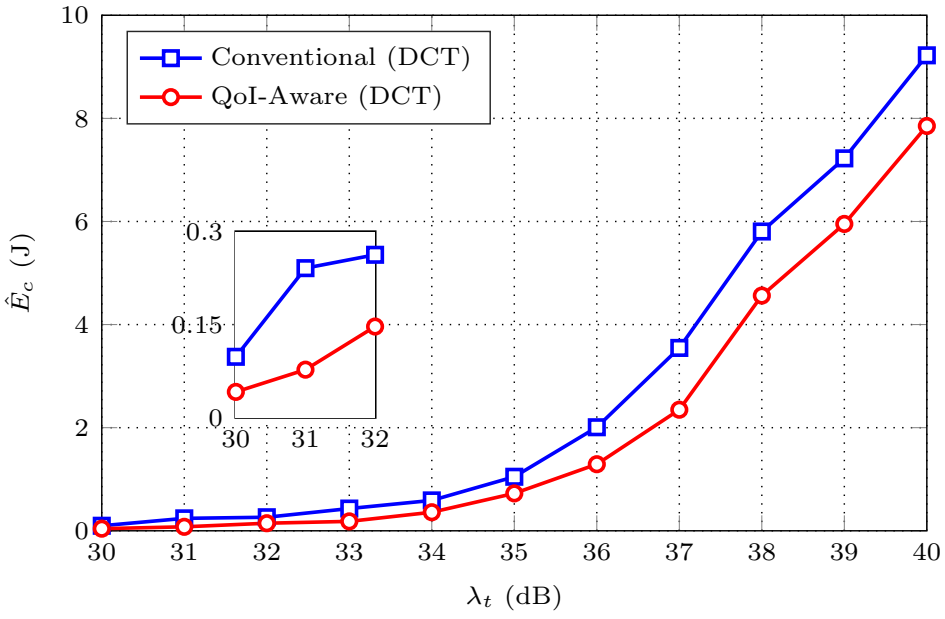

FIGURE 4.13: Comparison of the average energy cost incurred per node to transmit $\mathcal{N}_t = 150$ image frames with the proposed QoI-aware scheme and conventional scheme to achieve homogeneous target QoI thresholds with 2D-DCT

4.5 Summary

In this chapter, the issue of visual data optimisation prior to transmission in VSNs has been addressed. A QoI-aware dynamic self-reconfiguration scheme is proposed to obtain suitable configurations of visual sensing nodes for visual data optimisation in the pre-transmission phase. In the proposed scheme, QoI is characterised by the quality of the visual data and gauged by a PSNR-based metric. Since the communication energy cost is significantly higher than the processing cost in VSNs, the configuration space in the proposed scheme comprises of the amount and quality of visual data to be transmitted to the sink node. By making the visual sensing nodes self-reliant through the training and calibration process in the pre-deployment phase, the proposed reconfiguration scheme is fully decentralised, which accelerates the decision making process. During the training and calibration process, first, a suitable dataset is chosen depending on the given application. Subsequently, exploiting the variations in sensor-to-object distance and their impact on the feature redundancy within the acquired visual data, the appearance of objects within sensing nodes' FoV is modelled. Thereafter, redundant features are removed from the visual data by considering a dynamic compression range. Since feature redundancy removal may degrade the visual data quality, the resulting quality should be observed to satisfy an application's desired reliability level. Therefore, a QoI index is developed by taking into account the system's dynamic PSNR range. For a range set of dynamic sensor-to-object distances and compression ratios, compressive calibration matrix is obtained which is learnt by the visual sensing nodes to be employed in the post deployment phase.

In addition, the proposed in-node processing model provides an energy efficient solution for self-reconfiguration of visual sensing nodes in the post-deployment phase. Once the visual data is acquired by a sensing node, a suitable feature detection scheme is utilised and the object of interest is extracted. Thereafter, the sensor-to-object distance is estimated which is utilised by the proposed self-reconfiguration

scheme. The proposed scheme utilises the compressive calibration matrix learnt during the training and calibration process to select optimal configuration of visual sensing nodes for feature redundancy removal. Hence, a trade-off model between reliability and energy efficiency is provided which assists in dynamically optimising visual data while achieving the target QoI threshold.

Simulations were performed using the LDHF dataset to evaluate the performance of the proposed self-reconfiguration scheme. After employing the object appearance modelling approach, the resulting QoI index is analysed and it is observed that an increase in sensor-to-object distance results in QoI index reduction. Moreover, the decaying characteristics of QoI index are found to be identical for all the test images. Utilising 2D-DWT and 2D-DCT due to their suitability for visual sensing nodes, compressive calibration matrices are obtained. The QoI-aware self-reconfiguration scheme supported by 2D-DWT is found to be more efficient in terms of resulting QoI compared to the 2D-DCT based scheme. Considering a conventional scheme which utilises a static level of compression, the transmission energy cost compared to the proposed scheme is analysed for various homogeneous target QoI thresholds. It is observed from the comparison that for any given target QoI threshold, the proposed QoI-aware self-reconfiguration scheme results in transmission energy cost minimisation leading to energy savings in visual sensing nodes.

The next chapter builds on the self-reconfiguration scheme presented in this chapter to propose a unified framework of node classification and self-reconfiguration. A 3D coverage modelling scheme is introduced to enhance the sensing and processing intelligence of VSNs. Incorporating heterogeneity in target QoI thresholds, the proposed node classification scheme categorises visual sensing nodes into different levels of heterogeneity. The unified framework enables visual sensing nodes to optimise visual data prior to transmission based on heterogenous target QoI thresholds resulting in substantial energy savings.

Chapter 5

Unified Framework with Heterogeneous QoI Realisation

5.1 Introduction

In VSNs, depending on the given application, sensing nodes are required to achieve particular objectives whilst satisfying the desired reliability criteria. This can be accomplished by reconfiguring visual sensing nodes dynamically to obtain the optimal configuration, as discussed in Chapter 4. Visual sensing nodes are typically provided with a homogeneous reliability level which is to be maintained network-wide during the post-deployment phase for reliable network operation. Nevertheless, visual sensing nodes observe particular portions of the 3D target environment based on their location and orientation parameters and thus, provide different views of the sensing environment. Consequently, the achievement of particular objectives by satisfying homogeneous levels of desired reliability may not lead to efficient solutions. In these scenarios, exploiting the distributed nature of visual sensing nodes and adopting heterogeneity in the desired level of reliability is expected to enhance the global efficiency. Since, some regions within the monitoring space may dynamically facilitate relaxation in the desired reliability levels,

heterogeneous realisation is expected to achieve the objectives more convincingly as compared to its homogeneous counterpart. Although global efficiency can be improved by introducing heterogeneity, manually identifying the suitability of each visual sensing node for a particular heterogeneity level is impractical. Therefore, a strategy could be devised to dynamically classify visual sensing nodes into various categories depending on the application design criteria. Hence, by unifying the visual sensing node classification and self-reconfiguration schemes, a heterogeneous solution will emerge. The heterogeneous solution is expected to enhance the sensing and processing intelligence of VSNs resulting in an improved global efficiency.

In light of the above discussion, this chapter aims to incorporate heterogeneity and presents a unified framework for node classification and self-reconfiguration in VSNs. A scheme for 3D coverage modelling of visual sensing nodes is developed which is capable of dynamically obtaining their individual FoV information. The visual sensing nodes are made aware of their sensing environment by the proposed scheme. The context of heterogeneity is considered to be the allocation of suitable target QoI thresholds to visual sensing nodes in the network.

In order to introduce heterogeneity, a node classification scheme is presented which categorises visual sensing nodes depending on their 3D coverage. The proposed coverage modelling and node classification schemes are coupled with the in-node processing model (previously presented in Chapter 4) to formulate a unified framework. An energy model is developed to quantify the energy consumption of the unified framework. The robustness of the proposed unified framework is evaluated by performing simulations and comparing the results with existing state-of-the-art techniques. This chapter also presents an analytical model to evaluate the proposed framework's performance reliability. As compared to the homogeneous QoI realisation, the unified framework improves the global efficiency of achieving the objective (i.e. transmission energy cost minimisation) and leads to substantial energy savings.

5.2 Proposed Unified Framework

The proposed unified framework for node classification and self-reconfiguration in VSNs with heterogeneous QoI realisation for achieving targeted threshold based optimisation is presented in Figure 5.1.

In the proposed framework, training and calibration take place in the pre-deployment phase (as discussed in Chapter 4) which consists of training dataset selection, object appearance modelling, redundant feature removal, quality estimation and learning, as shown in Figure 5.1a. In the post-deployment phase, the unified framework incorporates 3D coverage modelling and QoI-centric node classification with the in-node processing model (comprising image capture, feature detection and object extraction, sensor-to-object distance estimation, self-reconfiguration and redundant feature removal), as shown in Figure 5.1b. Since, the in-node processing model for self-reconfiguration is presented earlier in Chapter 4, 3D coverage modelling and QoI-centric node classification are discussed in the following sections.

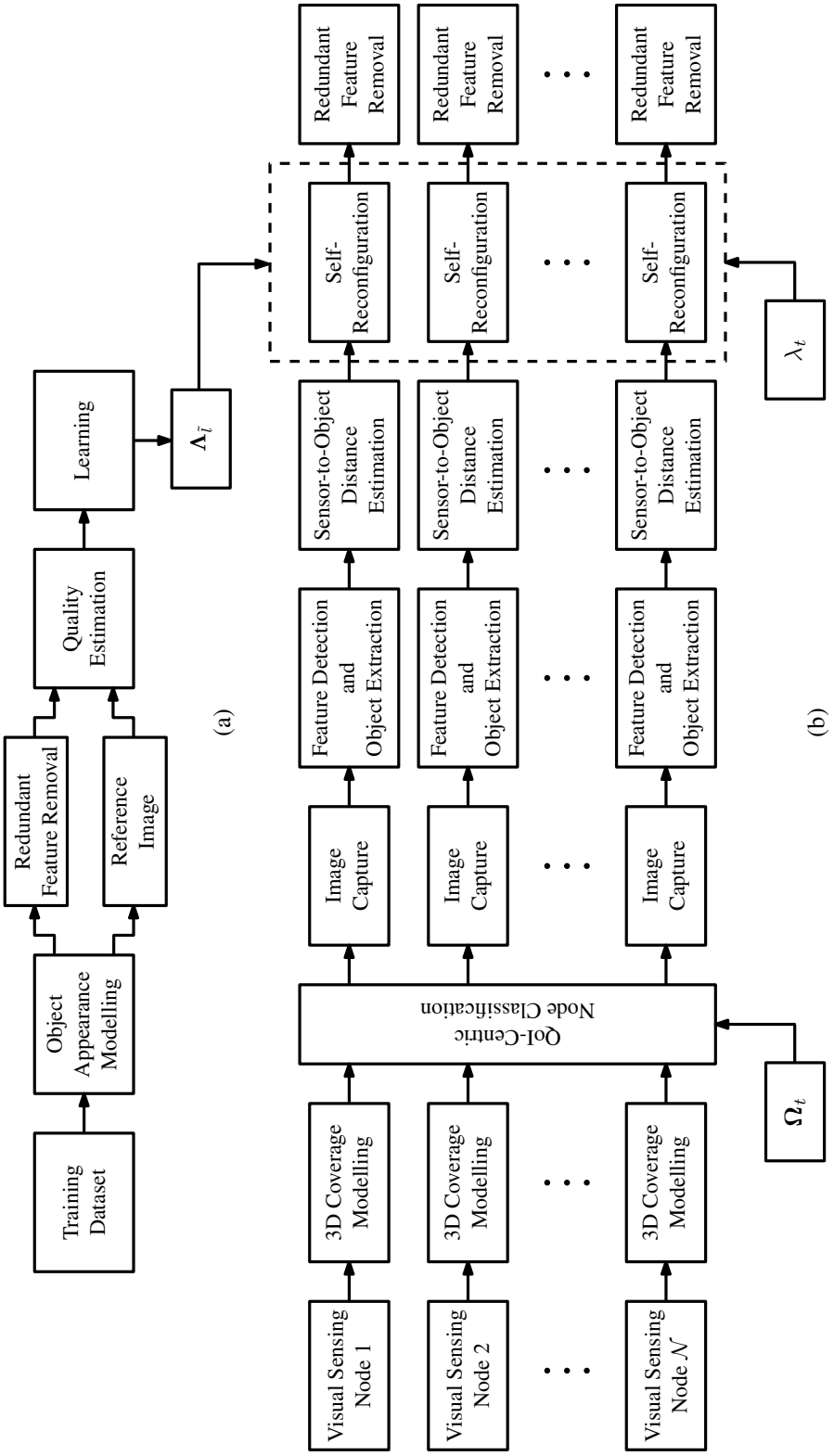

FIGURE 5.1: (a) Training and calibration process in the pre-deployment phase (b) Proposed unified framework for node classification and self-reconfiguration in resource constrained VSNs to achieve targeted threshold based optimisation

5.2.1 3D Coverage Modelling

Consider the visual sensing node's 3D projection model shown in Figure 3.1 where sensing node VS is located at the origin of the cartesian coordinate system i.e. $(0, 0, 0)$ and the sensor's optical axis overlaps onto the y-axis with $X = 0$ and $Z = 0$. Within the context of a VSN, where $\mathcal{N}$ sensing nodes are present, each sensing node $\text{VS}_{\tilde{l}}$ ($\tilde{l} = \{1, 2, 3, \ldots, \mathcal{N}\}$) is identified by its location which is described by the local cartesian coordinates $\mathbf{s} = [X_{\tilde{l}}, Y_{\tilde{l}}, Z_{\tilde{l}}, 1]^{\mathrm{T}}$, azimuth angle ϕ_a and elevation angle θ_e. In the simplest scenario where the sensor's optical axis is assumed to be parallel with the y-axis, thus the azimuth angle ϕ_a and elevation angle θ_e are given by,

$$\phi_a = \frac{\pi}{2} - \frac{\theta_h}{2} \tag{5.1}$$

$$\theta_e = \frac{\pi}{2} - \frac{\theta_v}{2} \tag{5.2}$$

where ϕ_a is measured clockwise and θ_e is measured counter-clockwise.

For a particular sensing range R, suppose $\mathbf{p} = [X_o, Y_o, Z_o, 1]^{\mathrm{T}}$ represent the coordinates of O_2 which are calculated by,

$$\mathbf{p} = [X_o, Y_o, Z_o, 1]^{\mathrm{T}} = [X_{\tilde{l}}, Y_{\tilde{l}} + R, Z_{\tilde{l}}, 1]^{\mathrm{T}} \tag{5.3}$$

Consider a more complex scenario where the sensor's optical axis is not parallel to the y-axis. Let θ_{ya} be the horizontal angle from y-axis to the sensor's optical axis measured clockwise. For a particular azimuth angle ϕ_a, θ_{ya} is calculated by,

$$\theta_{ya} = \begin{cases} \phi_a + \dfrac{\theta_h}{2} + \dfrac{3\pi}{2}, & 0 < \left(\phi_a + \dfrac{\theta_h}{2}\right) < \dfrac{\pi}{2} \\ \phi_a + \dfrac{\theta_h}{2} - \dfrac{\pi}{2}, & \dfrac{\pi}{2} \leq \left(\phi_a + \dfrac{\theta_h}{2}\right) \leq 2\pi \end{cases} \tag{5.4}$$

Similarly, let θ_{ye} be the vertical angle from y-axis to the sensor's optical axis measured counter-clockwise. For a particular elevation angle θ_e, θ_{ye} is calculated by,

$$\theta_{ye} = \begin{cases} \theta_e + \dfrac{\theta_v}{2} + \dfrac{3\pi}{2}, & 0 < \left(\theta_e + \dfrac{\theta_v}{2}\right) < \dfrac{\pi}{2} \\ \theta_e + \dfrac{\theta_v}{2} - \dfrac{\pi}{2}, & \dfrac{\pi}{2} \leq \left(\theta_e + \dfrac{\theta_v}{2}\right) \leq 2\pi \end{cases} \tag{5.5}$$

By adjusting the azimuth and elevation angles to fit a particular region of interest within the sensing node's FoV, the coordinates of O_2 are expected to change. The problem of calculating the new coordinates of O_2 can be classified into three cases. In the first case, suppose ϕ_a is adjusted to capture a particular region of interest within the sensing node's FoV and θ_e is kept constant i.e. $\theta_e = \frac{\pi}{2} - \frac{\theta_v}{2}$. This results in $0 < \theta_{ya} < 2\pi$ whereas θ_{ye} remains consistent i.e. 0. In the second case, suppose θ_e is adjusted to capture a particular region of interest within the sensing node's FoV and ϕ_a is kept constant i.e. $\phi_a = \frac{\pi}{2} - \frac{\theta_h}{2}$. This results in $0 < \theta_{ye} < 2\pi$ whereas θ_{ya} remains consistent i.e. 0. In the third case, suppose both ϕ_a and θ_e are adjusted to capture a particular region of interest within the sensing node's FoV. This results in $0 < \{\theta_{ya}, \theta_{ye}\} < 2\pi$. There are two possibilities for changing the azimuth and elevation angles in this case: (a) azimuth angle is adjusted first followed by the elevation angle, (b) elevation angle is adjusted first followed by the azimuth angle.

Based on the above discussion, let $\hat{c} \in \{1, 2, 3a, 3b\}$ denote Case 1, Case 2, Case 3a and Case 3b respectively. The new coordinates of point O_2 are represented by $\mathbf{p}'_{\hat{c}} = [X'_o, Y'_o, Z'_o, 1]^{\mathrm{T}}$ and derived as,

$$\mathbf{p}'_{\hat{c}} = \mathbf{T}^{-1}\Psi_{\hat{c}}\mathbf{T}\mathbf{p} \tag{5.6}$$

In Equation 5.6, $\mathbf{T}$ is the translation matrix given by,

$$\mathbf{T} = \begin{bmatrix} \mathbf{J} & \mathbf{t} \\ \mathbf{z} & 1 \end{bmatrix} \tag{5.7}$$

where $\mathbf{J}$ is a 3×3 identity matrix, $\mathbf{t}$ represent the transformation coordinates given by $\mathbf{t} = [-p_{(1)}, R - p_{(2)}, -p_{(3)}]^{\mathrm{T}}$ and $\mathbf{z}$ denote a 1×3 all-zeros vector.

$\mathbf{\Psi}_{\hat{c}}$ is expressed as,

$$\mathbf{\Psi}_{\hat{c}} = \begin{bmatrix} \mathbf{\Theta}_{\hat{c}} & \mathbf{z}^{\mathrm{T}} \\ \mathbf{z} & 1 \end{bmatrix} \tag{5.8}$$

where $\mathbf{\Theta}_{\hat{c}}$ denote the rotation matrices $\mathbf{\Theta}_1$, $\mathbf{\Theta}_2$, $\mathbf{\Theta}_{3\mathrm{a}}$ and $\mathbf{\Theta}_{3\mathrm{b}}$ given by [200, 201],

$$\mathbf{\Theta}_1 = \begin{bmatrix} \cos\theta_{ya} & \sin\theta_{ya} & 0 \\ -\sin\theta_{ya} & \cos\theta_{ya} & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{5.9}$$

$$\mathbf{\Theta}_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_{ye} & -\sin\theta_{ye} \\ 0 & \sin\theta_{ye} & \cos\theta_{ye} \end{bmatrix} \tag{5.10}$$

$$\mathbf{\Theta}_{3\mathrm{a}} = \begin{bmatrix} \cos\theta_{ya} & \sin\theta_{ya} & 0 \\ -\cos\theta_{ye}\sin\theta_{ya} & \cos\theta_{ye}\cos\theta_{ya} & -\sin\theta_{ye} \\ -\sin\theta_{ye}\sin\theta_{ya} & \sin\theta_{ye}\cos\theta_{ya} & \cos\theta_{ye} \end{bmatrix} \tag{5.11}$$

$$\mathbf{\Theta}_{3\mathrm{b}} = \begin{bmatrix} \cos\theta_{ya} & \sin\theta_{ya}\cos\theta_{ye} & -\sin\theta_{ya}\sin\theta_{ye} \\ -\sin\theta_{ya} & \cos\theta_{ya}\cos\theta_{ye} & -\cos\theta_{ya}\sin\theta_{ye} \\ 0 & \sin\theta_{ye} & \cos\theta_{ye} \end{bmatrix}$$

Let $\mathcal{S}$ be the sampling interval, the coordinates of s_y points on the y-axis within the sensing node's FoV are represented by $\mathbf{p}_y$ as,

$$\mathbf{p}_y = \{\mathcal{S}, 2\mathcal{S}, 3\mathcal{S}, \ldots, R\} \tag{5.12}$$

For a particular point $p_{y(\cdot)}$ on the y-axis, the coordinates of s_x points on the x-axis and s_z points on the z-axis are denoted by $\mathbf{p}_x$ and $\mathbf{p}_z$ respectively, and given by,

$$\mathbf{p}_x = \left\{\gamma_x^-, \gamma_x^- + \mathcal{S}, \gamma_x^- + 2\mathcal{S}, \ldots, \gamma_x^+\right\} \tag{5.13}$$

$$\mathbf{p}_z = \left\{\gamma_z^-, \gamma_z^- + \mathcal{S}, \gamma_z^- + 2\mathcal{S}, \ldots, \gamma_z^+\right\} \tag{5.14}$$

where γ_x^-, γ_x^+, γ_z^- and γ_z^+ represent the x-axis lower bound, x-axis upper bound, z-axis lower bound and z-axis upper bound respectively which are expressed as,

$$\left.\begin{matrix}\gamma_x^+ \\ \gamma_x^-\end{matrix}\right\} = X_o' \pm p_{y(\cdot)} \tan\left(\frac{\theta_h}{2}\right) \tag{5.15}$$

$$\left.\begin{matrix}\gamma_z^+ \\ \gamma_z^-\end{matrix}\right\} = Z_o' \pm p_{y(\cdot)} \tan\left(\frac{\theta_v}{2}\right) \tag{5.16}$$

and,

$$s_x = \left\lfloor \frac{\gamma_x^+ - \gamma_x^- + \mathcal{S}}{\mathcal{S}} \right\rfloor \tag{5.17}$$

$$s_z = \left\lfloor \frac{\gamma_z^+ - \gamma_z^- + \mathcal{S}}{\mathcal{S}} \right\rfloor \tag{5.18}$$

The total number of points in cartesian coordinates within a sensing node's FoV are derived as,

$$t_p = \sum_{\breve{i}=1}^{s_y} \left\lfloor \frac{2p_{y(\breve{i})} \tan\left(\frac{\theta_h}{2}\right) + \mathcal{S}}{\mathcal{S}} \right\rfloor \left\lfloor \frac{2p_{y(\breve{i})} \tan\left(\frac{\theta_v}{2}\right) + \mathcal{S}}{\mathcal{S}} \right\rfloor \tag{5.19}$$

where $s_y = \left\lfloor \frac{R}{\mathcal{S}} \right\rfloor$.

Suppose $\mathbf{c}_x$, $\mathbf{c}_y$ and $\mathbf{c}_z$ represent the set of 3D coordinates within a visual sensing node's FoV and are defined as,

$$\begin{aligned} \mathbf{c}_x &= \left\{c_{x(1)}, c_{x(2)}, c_{x(3)}, \ldots, c_{x(t_p)}\right\} \\ \mathbf{c}_y &= \left\{c_{y(1)}, c_{y(2)}, c_{y(3)}, \ldots, c_{y(t_p)}\right\} \\ \mathbf{c}_z &= \left\{c_{z(1)}, c_{z(2)}, c_{z(3)}, \ldots, c_{z(t_p)}\right\} \end{aligned} \tag{5.20}$$

where each respective pair $\left(c_{x(\cdot)}, c_{y(\cdot)}, c_{z(\cdot)}\right)$ denote the 3D cartesian coordinates of a point within a sensing node's FoV.

Algorithm 5.1 proposes a 3D coverage modelling scheme for visual sensing nodes to calculate $\mathbf{c}_x$, $\mathbf{c}_y$ and $\mathbf{c}_z$.

Algorithm 5.1 Proposed 3D coverage modelling scheme for visual sensing nodes

Input:

The sensing node's coordinates $[X_{\tilde{l}}, Y_{\tilde{l}}, Z_{\tilde{l}}, 1]^{\mathrm{T}}$, the azimuth angle ϕ_a, the elevation angle θ_e, the horizontal FoV θ_h, the vertical FoV θ_v, the sensing range R and the sampling interval $\mathcal{S}$.

Output:

$\mathbf{c}_x$, $\mathbf{c}_y$ and $\mathbf{c}_z$ representing the 3D coverage coordinates of a visual sensing node.

1: $l_c \leftarrow 1$

2: $\mathbf{p}_{\mathrm{y}} = \{\mathcal{S}, 2\mathcal{S}, 3\mathcal{S}, \ldots, R\}$

Algorithm 5.1 (continued) Proposed 3D coverage modelling scheme for visual sensing nodes

$s_y \leftarrow \left\lfloor \frac{R}{\mathcal{S}} \right\rfloor$

for $\breve{i} \leftarrow 1$ **to** s_y **do**

Set $\phi_a = \frac{\pi}{2} - \frac{\theta_h}{2}$ and $\theta_e = \frac{\pi}{2} - \frac{\theta_v}{2}$ for $\theta_{ya} = 0$ and $\theta_{ye} = 0$.

Calculate the coordinates of O_2 by substituting $R = p_{y(\breve{i})}$ in Equation 5.3 as,
$[X_o, Y_o, Z_o, 1]^{\mathrm{T}} = [X_{\tilde{l}}, Y_{\tilde{l}} + p_{y(\breve{i})}, Z_{\tilde{l}}, 1]^{\mathrm{T}}$

As $\theta_{ya} = 0$ and $\theta_{ye} = 0$,
$[X'_o, Y'_o, Z'_o, 1]^{\mathrm{T}} = [X_o, Y_o, Z_o, 1]^{\mathrm{T}}$

Calculate $\gamma_x^{\pm}$ and $\gamma_z^{\pm}$ using Equation 5.15 and Equation 5.16 respectively.

$\mathbf{p}_x = \{\gamma_x^-, \gamma_x^- + \mathcal{S}, \gamma_x^- + 2\mathcal{S}, \ldots, \gamma_x^+\}$

$\mathbf{p}_z = \{\gamma_z^-, \gamma_z^- + \mathcal{S}, \gamma_z^- + 2\mathcal{S}, \ldots, \gamma_z^+\}$

$s_x \leftarrow \lfloor (\gamma_x^+ - \gamma_x^- + \mathcal{S}) / \mathcal{S} \rfloor$

$s_z \leftarrow \lfloor (\gamma_z^+ - \gamma_z^- + \mathcal{S}) / \mathcal{S} \rfloor$

for $\breve{j} \leftarrow 1$ **to** s_x **do**

for $\breve{q} \leftarrow 1$ **to** s_z **do**

Based on ϕ_a and θ_e for targeted sensing within a 3D plane, calculate θ_{ya} and θ_{ye} using Equation 5.4 and Equation 5.5 respectively.

Substitute $\mathbf{p} = [p_{x(\breve{j})}, Y_{\tilde{l}} + p_{y(\breve{i})}, p_{z(\breve{q})}, 1]^{\mathrm{T}}$ in Equation 5.6 to obtain the coordinates after rotation $\mathbf{p}'_{(\cdot)} = [c_{x(l_c)}, c_{y(l_c)}, c_{z(l_c)}, 1]^{\mathrm{T}}$

$l_c \leftarrow l_c + 1$

end for

end for

end for

return $\mathbf{c}_x$, $\mathbf{c}_y$, $\mathbf{c}_z$

In the proposed 3D coverage modelling scheme for visual sensing nodes presented in Algorithm 5.1, initially l_c is defined as 1 which will act as a counter. Afterwards, depending on a particular sampling interval $\mathcal{S}$, the coordinates of points within the visual sensing node's FoV on the y-axis are obtained as $\mathbf{p}_{\mathrm{y}} = \{\mathcal{S}, 2\mathcal{S}, 3\mathcal{S}, \ldots, R\}$. In order to calculate the number of points s_y on the y-axis, $s_y = \left\lfloor \frac{R}{\mathcal{S}} \right\rfloor$ is used. Subsequently, for each point on the y-axis within the visual sensing node's FoV, considering $\theta_{ya} = 0$ and $\theta_{ye} = 0$, ϕ_a is set to $\left(\frac{\pi}{2} - \frac{\theta_h}{2}\right)$ and θ_e is set to $\left(\frac{\pi}{2} - \frac{\theta_v}{2}\right)$. The coordinates of O_2 are calculated by substituting $R = p_{y(\check{i})}$ in Equation 5.3 and the new coordinates of point O_2 are also obtained as $[X'_o, Y'_o, Z'_o, 1]^{\mathrm{T}} = [X_o, Y_o, Z_o, 1]^{\mathrm{T}}$. Next, utilising Equation 5.15 the x-axis lower and upper bounds γ_x^- and γ_x^+ respectively are calculated. Furthermore, z-axis lower and upper bounds γ_z^- and γ_z^+ respectively are obtained from Equation 5.16. Thereafter, $\mathbf{p}_x$ and $\mathbf{p}_z$ are calculated utilising Equation 5.13 and Equation 5.14 to obtain the coordinates of points on the x-axis and z-axis respectively for a particular point on the y-axis. Moreover, the number of points on the x-axis and z-axis for a particular point on the y-axis are calculated from Equation 5.17 and Equation 5.18 respectively. Afterwards, for each point on the x-axis and each point on the z-axis, depending on the azimuth and elevation angles, Equation 5.4 and Equation 5.5 are used to calculate θ_{ya} and θ_{ye} respectively. The post-rotation coordinates $\mathbf{p}'_{(\cdot)} = [c_{x(l_c)}, c_{y(l_c)}, c_{z(l_c)}, 1]^{\mathrm{T}}$ are obtained from Equation 5.6 by substituting $\mathbf{p} = [p_{x(\check{j})}, Y_{\check{l}} + p_{y(\check{i})}, p_{z(\check{q})}, 1]^{\mathrm{T}}$. The algorithm terminates once the coverage coordinates $\mathbf{c}_x$, $\mathbf{c}_y$ and $\mathbf{c}_z$ of a visual sensing node are obtained.

5.2.2 QoI-Centric Node Classification

In resource constrained scenarios, the utilisation of a homogeneous target QoI threshold may reduce the energy efficiency of a VSN. This is due to the fact that compared to others, some regions under VSN coverage may offer relaxation in the requirement of maintaining a particular QoI threshold. Furthermore, due to the time-varying nature of the targets' characteristics monitored within the visual

sensing nodes' FoV, heterogeneous QoI realisation in VSNs is more suitable as compared to the homogeneous QoI realisation. Thus, dividing the overall region of interest within the sensing environment into smaller sub-regions and realising heterogeneity in the target QoI thresholds is expected to preserve energy leading to improved network lifetime. In order to optimise the network's energy consumption based on such realisation, each node within the network needs to be assigned a local target QoI threshold to be achieved in accordance with its 3D spatial coverage coordinates. Let $\mathbf{C}_r^t = [\mathbf{C}_r^{\tilde{i}} \big| \tilde{i} = 1, 2, 3, \cdots, \hat{s}]$ be the matrix denoting the 3D coordinates of $\hat{s}$ sub-regions and $\mathbf{\Omega}_t = [\lambda_t^{\tilde{i}} \big| \tilde{i} = 1, 2, 3, \cdots, \hat{s}]$ be the set of their respective heterogeneous target QoI thresholds; where $\mathbf{C}_r^{\tilde{i}} = [\mathbf{c}_{x_r}^{\tilde{i}}, \mathbf{c}_{y_r}^{\tilde{i}}, \mathbf{c}_{z_r}^{\tilde{i}}]^{\mathrm{T}}$ represents the 3D coordinates of the $\tilde{i}th$ region of interest. Suppose $\mathbf{C}_s^{\tilde{l}} = [\mathbf{c}_{x_s}^{\tilde{l}}, \mathbf{c}_{y_s}^{\tilde{l}}, \mathbf{c}_{z_s}^{\tilde{l}}]^{\mathrm{T}}$ denote the 3D coordinates of the region within the $\tilde{l}th$ visual sensing node's FoV. Let the total number of points in cartesian coordinates within the FoV of $\tilde{l}th$ visual sensing node be represented by $t_p^{\tilde{l}}$ and suppose $t_p^{\tilde{i}}$ denote the number of points in $\tilde{i}th$ region of interest. The degree of overlap between $\tilde{l}th$ sensing node's FoV and $\tilde{i}th$ region of interest is derived as,

$$\mathcal{O}_{\tilde{i}\tilde{l}} = \frac{1}{\min\{t_p^{\tilde{l}}, t_p^{\tilde{i}}\}} \sum_{\check{i}=1}^{t_p^{\tilde{l}}} \sum_{\check{j}=1}^{t_p^{\tilde{i}}} \left\{ \prod_{\hat{q} \in \{x,y,z\}} 1 - \left[\mathrm{sgn}\left(c_{\hat{q}_s(\check{i})}^{\tilde{l}} - c_{\hat{q}_r(\check{j})}^{\tilde{i}} \right) \right]^2 \right\} \tag{5.21}$$

where sgn is the signum function. The value of $\mathcal{O}_{\tilde{i}\tilde{l}}$ for $t_p^{\tilde{l}} < t_p^{\tilde{i}}$ is categorised as,

$$\mathcal{O}_{\tilde{i}\tilde{l}} = \begin{cases} 0, & (\mathbf{c}_{x_s}^{\tilde{l}} \nsubseteq \mathbf{c}_{x_r}^{\tilde{i}}) \cap (\mathbf{c}_{y_s}^{\tilde{l}} \nsubseteq \mathbf{c}_{y_r}^{\tilde{i}}) \cap (\mathbf{c}_{z_s}^{\tilde{l}} \nsubseteq \mathbf{c}_{z_r}^{\tilde{i}}) \\ (0,1), & (\mathbf{c}_{x_s}^{\tilde{l}} \subset \mathbf{c}_{x_r}^{\tilde{i}}) \cap (\mathbf{c}_{y_s}^{\tilde{l}} \subset \mathbf{c}_{y_r}^{\tilde{i}}) \cap (\mathbf{c}_{z_s}^{\tilde{l}} \subset \mathbf{c}_{z_r}^{\tilde{i}}) \\ 1, & (\mathbf{c}_{x_s}^{\tilde{l}} \subseteq \mathbf{c}_{x_r}^{\tilde{i}}) \cap (\mathbf{c}_{y_s}^{\tilde{l}} \subseteq \mathbf{c}_{y_r}^{\tilde{i}}) \cap (\mathbf{c}_{z_s}^{\tilde{l}} \subseteq \mathbf{c}_{z_r}^{\tilde{i}}) \end{cases} \tag{5.22}$$

In case of $t_p^{\tilde{l}} > t_p^{\tilde{i}}$, Equation 5.22 is modified as,

$$\mathcal{O}_{\tilde{i}\tilde{l}} = \begin{cases} 0, & (\mathbf{c}_{x_r}^{\tilde{i}} \nsubseteq \mathbf{c}_{x_s}^{\tilde{l}}) \cap (\mathbf{c}_{y_r}^{\tilde{i}} \nsubseteq \mathbf{c}_{y_s}^{\tilde{l}}) \cap (\mathbf{c}_{z_r}^{\tilde{i}} \nsubseteq \mathbf{c}_{z_s}^{\tilde{l}}) \\ (0,1), & (\mathbf{c}_{x_r}^{\tilde{i}} \subset \mathbf{c}_{x_s}^{\tilde{l}}) \cap (\mathbf{c}_{y_r}^{\tilde{i}} \subset \mathbf{c}_{y_s}^{\tilde{l}}) \cap (\mathbf{c}_{z_r}^{\tilde{i}} \subset \mathbf{c}_{z_s}^{\tilde{l}}) \\ 1, & (\mathbf{c}_{x_r}^{\tilde{i}} \subseteq \mathbf{c}_{x_s}^{\tilde{l}}) \cap (\mathbf{c}_{y_r}^{\tilde{i}} \subseteq \mathbf{c}_{y_s}^{\tilde{l}}) \cap (\mathbf{c}_{z_r}^{\tilde{i}} \subseteq \mathbf{c}_{z_s}^{\tilde{l}}) \end{cases} \tag{5.23}$$

where $\mathcal{O}_{\tilde{i}\tilde{l}} = 0$, $0 < \mathcal{O}_{\tilde{i}\tilde{l}} < 1$ and $\mathcal{O}_{\tilde{i}\tilde{l}} = 1$ refer to no overlapping, partial overlapping and complete overlapping respectively between a sensing node's FoV and particular sensing sub-region.

The $\tilde{l}th$ sensing node is assigned a local target QoI threshold $\lambda_t^{\tilde{i}}$ based on the value of $\tilde{i}$ that satisfies the following criterion,

$$\xi_{\tilde{l}} = \arg\max_{\tilde{i}} \left[\frac{1}{\min\{t_p^{\tilde{l}}, t_p^{\tilde{i}}\}} \sum_{\check{i}=1}^{t_p^{\tilde{l}}} \sum_{\check{j}=1}^{t_p^{\tilde{i}}} \left\{ \prod_{\hat{q}\in\{x,y,z\}} 1 - \left[\text{sgn}\left(c_{\hat{q}_s(\check{i})}^{\tilde{l}} - c_{\hat{q}_r(\check{j})}^{\tilde{i}} \right) \right]^2 \right\} \right] \tag{5.24}$$

Suppose $\mathcal{H}$ denote the degree of heterogeneity, such that $\mathcal{H} < \hat{s}$. Let $\hat{m} = \{1, 2, 3, \ldots, \mathcal{H}+1\}$ represent a particular heterogeneity level and $\mathcal{N}_{\hat{m}}$ be the number of nodes classified within the $\hat{m}th$ heterogeneity level given by,

$$\mathcal{N}_{\hat{m}} = \mathcal{N}\left(1 - f_{\hat{m}}\right) \tag{5.25}$$

where,

$$f_{\hat{m}} = 1 - \frac{\sum_{\tilde{l}=1}^{\mathcal{N}} 1 - \left[\text{sgn}\left(\xi_{\tilde{l}} - \hat{m}\right)\right]^2}{\mathcal{N}} \tag{5.26}$$

such that $\sum_{\hat{m}=1}^{\mathcal{H}+1} f_{\hat{m}} = \mathcal{H}$; therefore,

$$\mathcal{N}_{\hat{m}} = \sum_{\tilde{l}=1}^{\mathcal{N}} 1 - \left[\text{sgn}\left(\xi_{\tilde{l}} - \hat{m}\right)\right]^2 \tag{5.27}$$

such that $\sum_{\hat{m}=1}^{\mathcal{H}+1} \mathcal{N}_{\hat{m}} = \mathcal{N}$.

5.3 Energy Model

The performance of a visual sensor network is characterised by the energy conservation within the network. Figure 5.2 shows the energy dissipation model of a radio transceiver [202], which is characterised by the energy cost of the transmitter and receiver units. Let E_{tx} and E_{rx} denote the energy consumed for transmitting and receiving one bit respectively. Suppose $\hat{k}_{\hat{l}}^{\hat{m}}$ and $\tilde{k}_{\hat{l}}^{\hat{m}}$ represent the number of bits transmitted and received respectively by $\hat{l}th$ visual sensing node belonging to $\hat{m}th$ level of heterogeneity. Let $\mathcal{N}_t$ and $\mathcal{N}_r$ denote the total number of image frames transmitted and received by a visual sensing node. Suppose $e_{\hat{l}}^t$ and $e_{\hat{l}}^r$ represent the number of control signal bits transmitted and received by a visual sensing node respectively. Within the context of the proposed framework, each visual sensing node transmits its location, azimuth and elevation angles to the sink node; and the sink node sends control signals to each visual sensing node for classification within a suitable level of heterogeneity. Let n_t and n_r be the total

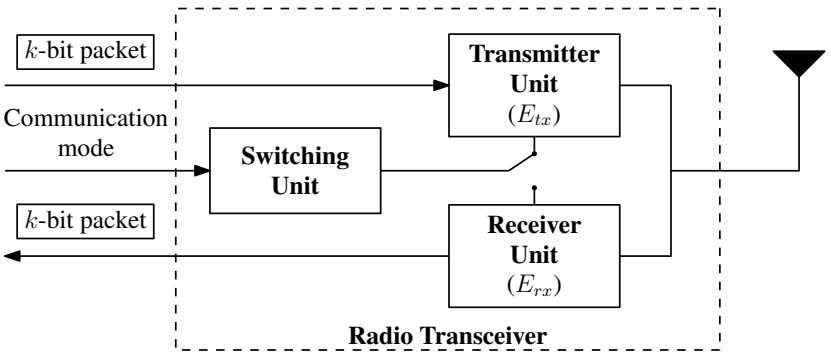

FIGURE 5.2: Communication energy dissipation model [202]

number of control signals transmitted and received by a visual sensing node. The total energy consumption within the communication phase of a VSN is denoted by $\tilde{E}_c$ and obtained as,

$$\tilde{E}_c = \sum_{\hat{m}=1}^{\mathcal{H}+1} \sum_{\hat{l}=1}^{\mathcal{N}_{\hat{m}}} \left[E_{tx} \left(\mathcal{N}_t \hat{k}_{\hat{l}}^{\hat{m}} + n_t e_{\hat{l}}^t \right) + E_{rx} \left(\mathcal{N}_r \tilde{k}_{\hat{l}}^{\hat{m}} + n_r e_{\hat{l}}^r \right) \right] \tag{5.28}$$

As in Chapter 3 and Chapter 4, in order to calculate the energy consumption of the visual sensing nodes, the energy model of [183] is used. The parameters used to model the energy consumption of the radio transceiver are given in Table 3.8. In VSNs, the energy cost incurred for visual data transmission and reception is significantly higher than the data processing cost [27–29]. Therefore, only the communication energy cost is considered to model the energy dissipation of a visual sensing node, as in Chapter 4. Moreover, it is assumed that adequate resources are available at the sink node and it is not constrained by limited energy, data storage and computational capability. This is a widely adopted assumption in the literature; consequently, the energy cost for reception at the sink node does not influence lifetime of the VSN and can be ignored.

5.4 Results and Analysis

This section evaluates the robustness of the proposed unified framework compared to existing state-of-the-art techniques. An analysis of the proposed framework's performance reliability is also presented. As in Chapter 4, in order to evaluate the performance of the proposed unified framework, MATLAB simulation platform was used to conduct the experiments on an Intel Core i5-4200U CPU @ 1.60GHz 2.30GHz with 8 GB RAM and 3 MB cache.

5.4.1 Energy Efficiency of the Unified Framework

A 3D sensing environment of size $50 \times 50 \times 10$ m^3 is considered for visual sensing nodes deployment and the origin of the coordinate system is assumed to be at $(1, 1, 1)$. The number of sensing nodes $\mathcal{N}$ within the network is considered to be 100 and the sampling interval $\mathcal{S}$ is 0.1 m. The horizontal and vertical FoVs are considered to be 48.39° and 37.25° respectively, as in Chapter 3. The azimuth and elevation angles are bounded within $[0^\circ, 360^\circ]$ and $[0^\circ, 155^\circ]$ respectively. It is assumed that the visual sensing nodes are static; therefore, θ_e and ϕ_a are assumed to be constant for each sensing node. Degree of heterogeneity up to 2 is considered, therefore $\mathcal{H} \in \{0, 1, 2\}$ appears in the simulation model. $\mathcal{H} = 0$ refers to homogeneous target QoI realisation; $\mathcal{H} = 1$ and $\mathcal{H} = 2$ refer to heterogeneous realisation of target QoI. The number of control signal bits transmitted ($e_{\hat{l}}^t$) and received ($e_{\hat{l}}^r$) each time the node classification takes place are 115 and 5 respectively. In the simulation model, the number of control signals transmission $n_t = 1$ and reception $n_r = 1$ are considered for simplicity. However, the proposed framework can support scenarios with dynamically changing QoI by updating the visual sensing nodes' target QoI thresholds for reclassification within a suitable level of heterogeneity. A detailed analysis of the proposed self-reconfiguration scheme to achieve different homogeneous target QoI thresholds has been presented in Chapter 4.

Consider a heterogeneous realisation of target QoI i.e. $\mathcal{H} > 0$. Modelling the simulations using the parameters presented in the earlier discussion and incorporating them within the proposed 3D coverage modelling and QoI-centric node classification schemes, $\{f_1, f_2\} = \{0.48, 0.52\}$ and $\{f_1, f_2, f_3\} = \{0.64, 0.67, 0.69\}$ are obtained, as given in Table 5.1, for target QoI with first and second degree of heterogeneity respectively. Table 5.1 also presents the number of nodes classified in $\hat{m}th$ level of heterogeneity i.e. $\{\mathcal{N}_1, \mathcal{N}_2\} = \{52, 48\}$ and $\{\mathcal{N}_1, \mathcal{N}_2, \mathcal{N}_3\} = \{36, 33, 31\}$ for first ($\mathcal{H} = 1$) and second ($\mathcal{H} = 2$) degree of heterogeneity respectively with $\hat{m} = \{1, 2, 3, \ldots, \mathcal{H} + 1\}$.

TABLE 5.1: Node classification for target QoI with first and second degree of heterogeneity

$\mathcal{H}$	f_1	f_2	f_3	$\sum_{\hat{m}=1}^{\mathcal{H}+1} f_{\hat{m}}$	$\mathcal{N}_1$	$\mathcal{N}_2$	$\mathcal{N}_3$	$\sum_{\hat{m}=1}^{\mathcal{H}+1} \mathcal{N}_{\hat{m}}$
1	0.48	0.52	–	1	52	48	–	100
2	0.64	0.67	0.69	2	36	33	31	100

The distribution of visual sensing nodes within the network and classification based on two different target QoI thresholds, i.e. λ_t^1 and λ_t^2, is shown in Figure 5.3. Moreover, Figure 5.4 illustrates the visual sensing node distribution and classification for second degree of heterogeneity. It is expected that increasing

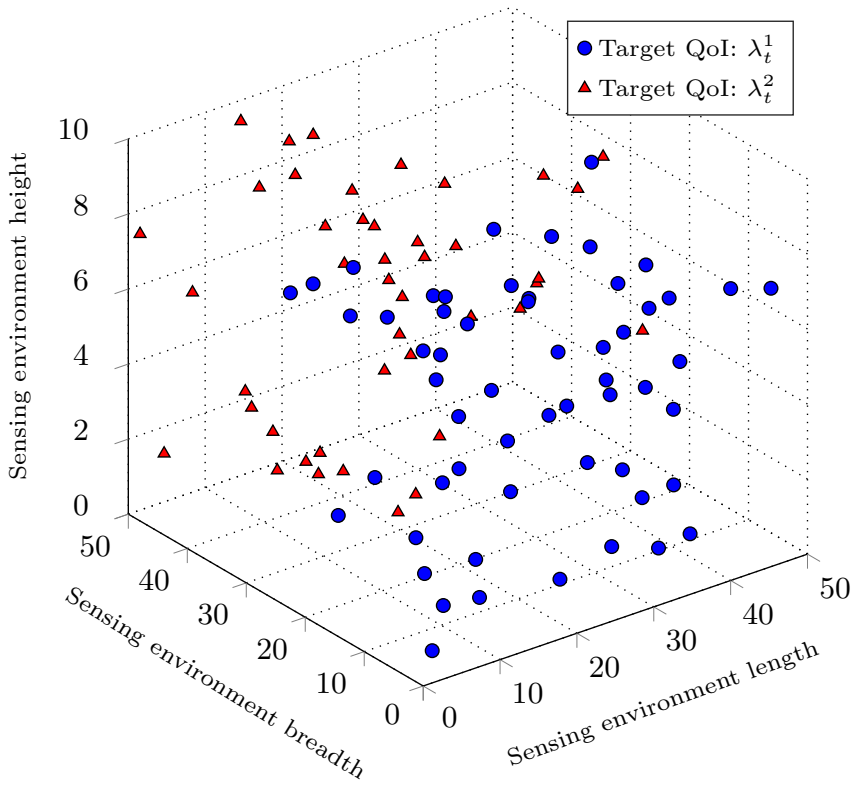

FIGURE 5.3: Visual sensing nodes distribution within the network for first degree of heterogeneity

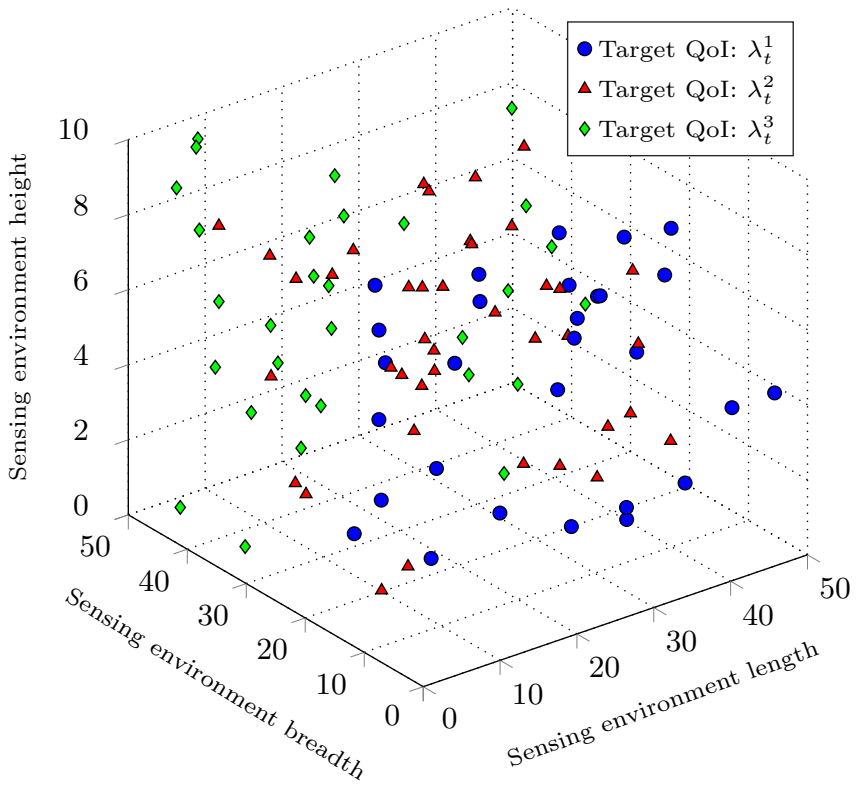

FIGURE 5.4: Visual sensing nodes distribution within the network for second degree of heterogeneity

the degree of heterogeneity will enhance the energy efficiency of a VSN. The energy consumption analysis utilising the proposed unified framework with different degree of heterogeneity will demonstrate such behaviour.

In order to analyse the efficiency of the proposed scheme, its energy consumption for homogeneous and heterogeneous target QoI realisations is compared with the energy cost of existing 2D-DWT and 2D-DCT based state-of-the-art approaches proposed in [28, 29, 55] and the results are summarised in Table 5.2. These approaches are chosen for comparison as they utilise 2D-DWT or 2D-DCT, which are suitable for visual sensing nodes [191] and are employed for the proposed unified framework's performance analysis. Moreover, these approaches show promising performance, therefore, the selection of these approaches will lead to a rigorous comparative analysis.

TABLE 5.2: Comparison of the proposed scheme with existing state-of-the-art techniques for homogeneous (HM) and heterogeneous (HT) QoI realisations

Compression	**Scenario**	**Scheme**	**Degree of Heterogeneity**	**Target QoI** (dB)	**Energy Consumption** (mJ)	**Energy Savings with Proposed** (%)	**Minimum QoI Achievement Reliability** (%)
2D-DWT	1a	Aziz & Pham [28]	0	40	37.18	–	–
		Proposed with HM-QoI			29.36	21.03	97.52
	1b	Proposed with HT-QoI	1	{39, 40}	23.02	38.08	97.37
		Proposed with HT-QoI	2	{38, 39, 40}	19.10	48.63	97.21
	1c	Proposed with HT-QoI	1	{30, 40}	14.16	61.92	94.84
		Proposed with HT-QoI	2	{30, 35, 40}	10.43	71.95	95.09
2D-DCT	2a	Halder & Ghosal [55]	0	39	48.25	–	–
		Proposed with HM-QoI			39.78	17.55	97.22
	2b	Proposed with HT-QoI	1	{38, 39}	34.87	27.73	97.05
		Proposed with HT-QoI	2	{37, 38, 39}	28.49	40.95	96.87
	2c	Proposed with HT-QoI	1	{30, 39}	19.24	60.12	94.69
		Proposed with HT-QoI	2	{30, 35, 39}	14.83	69.26	94.99
2D-DCT	3a	Chow *et al.* [29]	0	35	7.05	–	–
		Proposed with HM-QoI			4.88	30.78	95.59
	3b	Proposed with HT-QoI	1	{34, 35}	3.59	49.08	95.32
		Proposed with HT-QoI	2	{33, 34, 35}	2.81	60.14	95.03
	3c	Proposed with HT-QoI	1	{30, 35}	2.49	64.68	93.88
		Proposed with HT-QoI	2	{30, 33, 35}	2.10	70.21	94.07

Scenarios 1a, 2a and 3a present the performance analysis for homogeneous target QoI realisation; scenarios 1b, 2b and 3b show the performance analysis for heterogeneous target QoI realisation with tight thresholds; and scenarios 1c, 2c and 3c provide the performance analysis for heterogeneous target QoI realisation with relaxed thresholds. Table 5.2 also presents the minimum QoI achievement reliability for such scenarios, which is discussed in the following section.

It is observed from scenarios 1a, 2a and 3a that with homogeneous QoI realisation i.e. 40 dB, 39 dB and 35 dB target QoI respectively, the proposed scheme results in 21.03%, 17.55% and 30.78% energy savings compared to the schemes presented in [28], [55] and [29] respectively. The results demonstrate that lower target QoI threshold of 35 dB achieves higher energy efficiency compared to higher QoI target thresholds. Similarly, for the first degree of heterogeneity with tight QoI thresholds i.e. $\{39\text{ dB}, 40\text{ dB}\}$, $\{38\text{ dB}, 39\text{ dB}\}$ and $\{34\text{ dB}, 35\text{ dB}\}$, it is found from scenarios 1b, 2b and 3b that the proposed unified framework leads to 38.08%, 27.73% and 49.08% energy conservation compared to [28], [55] and [29] respectively. Moreover, it is noticed from scenarios 1b, 2b and 3b that, for the second degree of heterogeneity with tight QoI thresholds i.e. $\{38\text{ dB}, 39\text{ dB}, 40\text{ dB}\}$, $\{37\text{ dB}, 38\text{ dB}, 39\text{ dB}\}$ and $\{33\text{ dB}, 34\text{ dB}, 35\text{ dB}\}$, the proposed unified framework leads to 48.63%, 40.95% and 60.14% energy savings compared to [28], [55] and [29] respectively. The results demonstrate that by increasing the level of heterogeneity in target QoI, superior performance in terms of energy savings is achieved. Moreover, as discussed earlier, the results signify the suitability of the proposed scheme for resource constrained scenarios by providing significant energy savings at lower QoI target thresholds.

The experiments are also conducted for first degree of heterogeneity with relaxed QoI thresholds i.e. $\{30\text{ dB}, 40\text{ dB}\}$, $\{30\text{ dB}, 39\text{ dB}\}$ and $\{30\text{ dB}, 35\text{ dB}\}$, it is observed from scenarios 1c, 2c and 3c that the proposed unified framework provides 61.92%, 60.12% and 64.68% energy savings compared to [28], [55] and [29] respectively. Furthermore, results obtained by considering second degree of heterogeneity

with relaxed QoI thresholds i.e. $\{30\text{ dB}, 35\text{ dB}, 40\text{ dB}\}$, $\{30\text{ dB}, 35\text{ dB}, 39\text{ dB}\}$ and $\{30\text{ dB}, 33\text{ dB}, 35\text{ dB}\}$, in scenarios 1c, 2c and 3c demonstrate 71.95%, 69.26% and 70.21% energy savings with the proposed unified framework compared to [28], [55] and [29] respectively. It is observed from the results that relaxed QoI thresholds result in higher energy savings compared to the tight QoI thresholds. Moreover, the scenarios where second degree of heterogeneity is considered with relaxed QoI thresholds lead to substantial energy conservation compared to other scenarios.

In light of the aforementioned analysis, it can be concluded that the proposed unified framework results in substantial energy savings compared to its existing counterparts and thus leads to an improved network lifetime. The reason for these energy savings is the dynamic nature of the proposed scheme, where visual sensing nodes are dynamically assigned a suitable target QoI threshold and reconfigured accordingly. In contrast, the schemes in [28, 29, 55] are static for particular target QoI thresholds. Moreover, the results demonstrate that heterogeneous QoI realisations result in significantly higher energy savings as compared to their homogeneous counterparts. Hence, the proposed scheme provides a feasible solution to enhance the energy efficiency of individual visual sensing nodes and it is found to be suitable for VSNs with strict constraints on available energy. The following section formulates an analytical model to analyse the proposed framework's performance reliability.

5.4.2 Analysis of Proposed Framework's Performance Reliability

Reliable reporting of visual information to the sink node is an important requirement in VSNs [1, 203]. Within the context of the proposed unified framework for node classification and self-reconfiguration, reliability can be defined as the ability of the visual sensing nodes to deliver the visual information while satisfying the target QoI requirements. In order to analyse the robustness of the proposed

framework, an analytical model is developed to calculate the performance reliability as a function of: the target QoI threshold (λ_t) and the QoI delivered by a visual sensing node (β). Due to the fact that the fidelity of the compressive calibration matrix is quantified within a confidence bound $[e^-, e^+]$ (as shown in Equation 4.9), the performance reliability of the proposed framework is bounded between $[\mathcal{P}_r^-, \mathcal{P}_r^+]$, where $\mathcal{P}_r^{\pm} \in [0, 1]$, and is derived as,

$$\mathcal{P}_r^{\pm} = \begin{cases} 1 - \xi_p^{\pm}, & \lambda_t \geq \beta + e^{\pm} \\ 1, & \text{otherwise} \end{cases} \tag{5.29}$$

$$\xi_p^{\pm} = \left[Q\left(\frac{10^{2\log(m) - \frac{\lambda_t}{20}} - 10^{2\log(m) - \frac{\beta + e^{\pm}}{20}}}{\sum_{k=0}^{\hat{\alpha}_b - 1} 2^{2k}} \right) - 0.50 \right] \times \left[Q\left(\frac{10^{2\log(m) - \frac{\zeta^u}{20}} - 10^{2\log(m) - \frac{\zeta^l}{20}}}{\sum_{k=0}^{\hat{\alpha}_b - 1} 2^{2k}} \right) - 0.50 \right]^{-1} \tag{5.30}$$

where $\xi_r^{\pm}$ denote the probability of failure to ensure the target QoI satisfaction and it is calculated from Equation 5.30 with $m = 2^{\hat{\alpha}_b - 1}$ ($\hat{\alpha}_b$ denote the number of bits per pixel), $e^{\pm}$ is obtained from Equation 4.9 and $Q(\cdot)$ is the Q-function given by,

$$Q(\hat{x}) = \frac{1}{\sqrt{2\pi}} \int_{\hat{x}}^{\infty} e^{-\frac{1}{2}q^2} dq \tag{5.31}$$

The robustness of a system employing image processing algorithms is directly proportional to the PSNR, and $Q(\cdot)$ is a monotonically decreasing function; thus the realisation with $(1 - \xi_p^{\pm})$ leads to a measure of robustness in terms of reliability

to dynamically ensure that the targeted QoI is achieved. Using Equation 4.9, e^- and e^+ for the proposed framework are found to be -0.78 dB and 0.87 dB respectively. Substituting the parameters in Equation 5.29 and Equation 5.30, the system reliability to achieve targeted threshold based optimisation is shown in Figure 5.5 and Figure 5.6. The lower bound $\mathcal{P}_r^-$ obtained with respect to e^-, which denotes the minimum reliability offered by the system for a range set of target QoI thresholds (λ_t) and delivered QoI (β) is shown in Figure 5.5. It is observed from Figure 5.5 that the reliability increases with the increase in delivered QoI and attains a maximum value when $\beta + e^- \geq \lambda_t$.

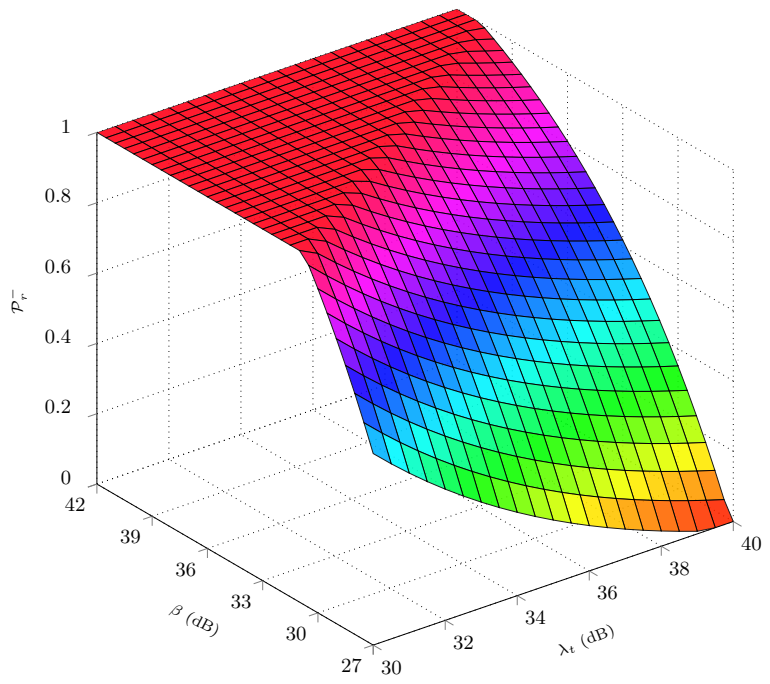

FIGURE 5.5: The minimum level of reliability for a range set of target QoI threshold (λ_t) and delivered QoI (β).

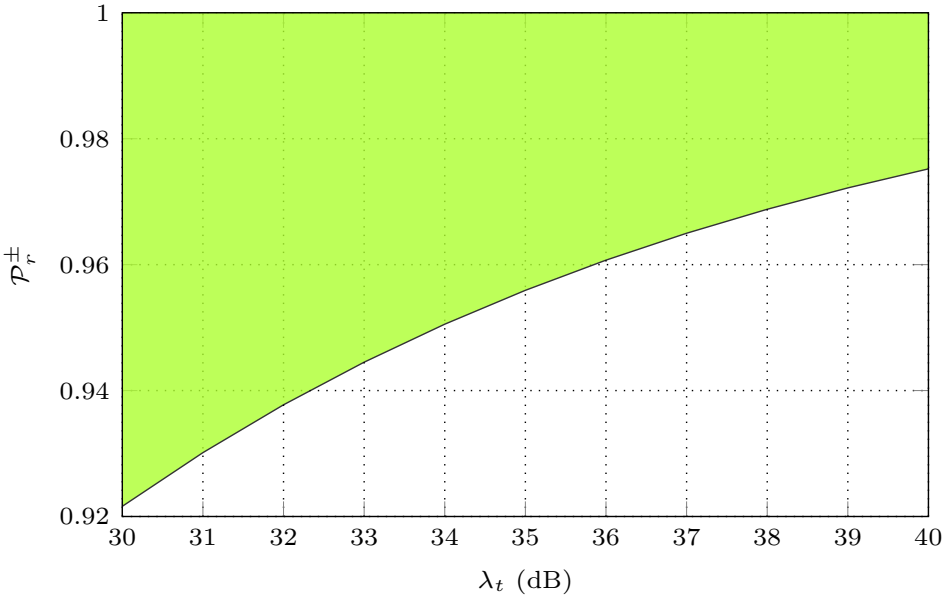

FIGURE 5.6: Proposed framework's confidence bound for retrieving the information.

In order to demonstrate the reliability of the QoI ensured by the proposed framework, incorporating both e^- and e^+, the region between the upper and lower confidence bounds, representing $\mathcal{P}_r^{\pm}$, is shown in Figure 5.6. It is noticed that the reliability, bounded between $\mathcal{P}_r^-$ and $\mathcal{P}_r^+$ (shown in the shaded region), increases with the increase in target QoI threshold. This is due to the fact that the impact of the fidelity of the compressive calibration matrix decreases with the increase in target QoI threshold and thus enhances the confidence bound for retrieving the information.

As the fidelity of the compressive calibration matrix is quantified within a confidence bound $[e^-, e^+]$, it is pertinent to study the impact of different values of the parameter e^- on the minimum QoI achievement reliability. Figure 5.7 illustrates such impact by considering the values of e^- from -1 dB to 0 dB for different target QoI thresholds i.e. 40 dB, 35 dB and 30 dB.

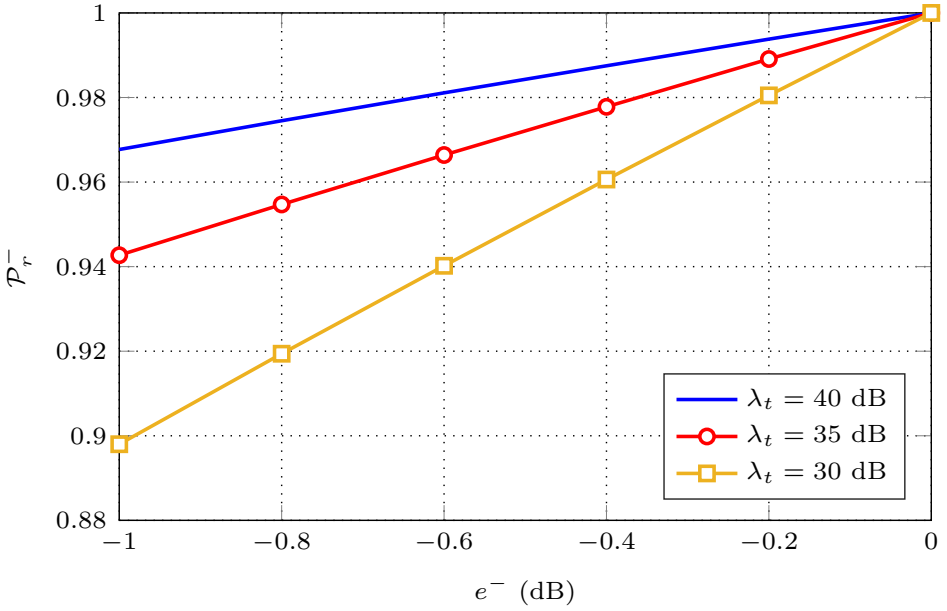

FIGURE 5.7: Proposed framework's information retrieval confidence bound for varying fidelity of the compressive calibration matrix.

The results demonstrate that the reliability level is minimum at the lowest value of e^- and it increases as the value of e^- increases. Moreover, as mentioned earlier, the impact of e^- on the reliability is higher at lower target QoI thresholds. As an example, the results show that the information retrieval reliability at $\lambda_t = 40$ dB is higher than the reliability at $\lambda_t = 35$ dB and $\lambda_t = 30$ dB for $e^- = -1$ dB.

Due to the dependency of the reliability level on various parameters, it is required to develop a methodology which can assist the design engineers in selecting a suitable target QoI threshold λ_t that can satisfy a desired level of reliability $\mathcal{P}_r^t$. Suppose $\mathbf{t}_\lambda$ be the set of target QoI thresholds and $\mathbf{a}_\beta$ be the set of respective delivered QoI. Algorithm 5.2 presents a method for calculating an optimal target QoI threshold λ_t to be employed within a visual sensing node such that it operates at a desired reliability level $\mathcal{P}_r^t$.

Algorithm 5.2 Optimal target QoI threshold calculation for sensing node operation at a desired level of reliability

Input:

The set of target QoI thresholds $\mathbf{t}_\lambda$, the set of respective delivered QoI $\mathbf{a}_\beta$, the lower bound that quantifies the fidelity of the compressive calibration matrix e^-, the number of bits per pixel $\hat{\alpha}_b$, the system's dynamic PSNR range (ζ^l, ζ^u) and the desired reliability level $\mathcal{P}_r^t$.

Output:

The target QoI threshold λ_t to be employed by a visual sensing node such that it operates at a desired reliability level.

$\hat{\mathbf{p}}^- \leftarrow \varnothing$
for $\breve{k} \leftarrow 1$ **to** $\|\mathbf{t}_\lambda\|_0$ **do**
    $\lambda_t \leftarrow t_{\lambda(\breve{k})}$
    $\beta \leftarrow a_{\beta(\breve{k})}$
    if $\lambda_t \geq \beta + e^-$ **then**
        Calculate ξ_p^- from Equation 5.30
        $\mathcal{P}_r^- \leftarrow 1 - \xi_p^-$
    else
        $\mathcal{P}_r^- \leftarrow 1$
    end if
    $\hat{p}_{(\breve{k})}^- \leftarrow \mathcal{P}_r^-$
end for
$\sigma \leftarrow \arg\min_{\breve{k}} \left[\left| \mathcal{P}_r^t - \hat{p}_{(\breve{k})}^- \right| \right]; \breve{k} = \{1, 2, \ldots, \|\mathbf{t}_\lambda\|_0\}$
$\lambda_t \leftarrow t_{\lambda(\sigma)}$
return λ_t

In the method presented in Algorithm 5.2, initially a matrix $\hat{\mathbf{p}}^-$ is defined. Afterwards, each element from the vector $\mathbf{t}_\lambda$ and the respective value from vector $\mathbf{a}_\beta$ is extracted and the criterion $\lambda_t \geq \beta + e^-$ is employed. In case if the criterion is true, ξ_p^- is calculated from Equation 5.30 and $\mathcal{P}_r^-$ is obtained using

$\mathcal{P}_r^- = 1 - \xi_p^-$. Whereas, if the criterion is not satisfied, $\mathcal{P}_r^- = 1$ is obtained. Next, the value of $\mathcal{P}_r^-$ is stored at the $\breve{k}$th index of the vector $\hat{\mathbf{p}}^-$. Afterwards, function $\arg\min_{\breve{k}} \left[\left| \mathcal{P}_r^t - \hat{p}_{(\breve{k})}^- \right| \right]$ is utilised and the resulting value of $\breve{k}$ is stored in σ; where $\breve{k} = \{1, 2, \ldots, \|\mathbf{t}_\lambda\|_0\}$ and $\|\cdot\|_0$ is the l_0-norm. The value at index σ is extracted from $\mathbf{t}_\lambda$ and stored in λ_t, which can be utilised within a visual sensing node for satisfying a desired reliability level.

A comparison of the minimum QoI achievement reliability for homogeneous and heterogeneous QoI realisations is presented in Table 5.2. In the case of homogeneous target QoI thresholds of 40 dB, 39 dB and 35 dB, the reliability of the proposed framework to ensure that the targeted QoI is achieved is as low as 97.52%, 97.22% and 95.59% respectively. On the other hand, for heterogeneous target QoI thresholds, the proposed unified framework guarantees as low as 93.88% reliability. Although, compared to the homogeneous scenarios, the heterogeneous realisations result in up to 2.68% degradation in the reliability. However, the latter lead to substantial energy savings and justify the robustness of the proposed unified heterogeneous framework of node classification and self-reconfiguration for resource constrained VSNs. Moreover, the tight and relaxed heterogeneous target QoI thresholds provide a trade-off between energy efficiency and reliability.

5.5 Summary

This chapter considered the adoption of heterogeneity in the reconfiguration criteria to maximise the global efficiency of achieving the reconfiguration objective. Since, the manual allocation of a suitable level of heterogeneity to each visual sensing node within the VSN is not feasible, a classification scheme is developed to dynamically categorise visual sensing nodes into suitable heterogeneity levels. In the proposed node classification scheme, visual sensing nodes are categorised depending on the QoI requirements of the region within their FoV. In order to equip visual sensing nodes with their FoV information, a 3D coverage modelling

scheme is proposed which is capable of dynamically obtaining visual sensing nodes' FoV information. Once the visual sensing nodes obtain their FoV information, the node classification scheme categorises them into various heterogeneity levels.

In order to enhance the robustness, a novel unified framework for the classification and self-reconfiguration of sensing nodes in resource constrained VSNs is proposed. The proposed framework incorporates the QoI-centric node classification scheme, the 3D coverage modelling scheme and the in-node processing model by exploiting the heterogeneity of targeted QoI threshold levels within the sensing region. Since, satisfying a particular acceptable reliability level is paramount in VSN applications, an analytical model is formulated to quantify the degree to which the targeted QoI thresholds are achieved by a visual sensing node. The analytical model, depending on the given application, provides a tuneable model to attain suitable QoI thresholds.

An energy model is developed to calculate the energy consumption of the proposed unified framework. Simulations were performed to evaluate the energy efficiency of the proposed unified framework by considering node deployment in a 3D sensing environment and degree of heterogeneity up to 2. For given target thresholds of QoI, it is observed that the proposed unified framework resulted in significant amount of energy savings compared to the existing state-of-the-art techniques, as shown in the results, thus enhancing the lifetime of the network. The energy efficiency of the proposed unified framework demonstrated its feasibility to assist the system design engineers for speedy deployment of VSNs in scenarios with strict resource constraints.

The next chapter briefly summarises the research challenges and contributions of the thesis. Moreover, it provides several directions to extend the proposed schemes for future research.

Chapter 6

Conclusions and Future Directions

This chapter presents a brief summary of the thesis and provides a strategy for future research utilising the proposed schemes.

6.1 Conclusions

A summary of the research challenges, the proposed FoV characterisation framework, the proposed QoI-aware self-reconfiguration scheme and the proposed unified framework is provided in the following sections.

6.1.1 Research Challenges

VSNs have attracted the attention of both the research community and the industry for over a decade. By embedding a visual sensor, processor and a wireless transceiver within a tiny low-powered sensing node, VSNs are capable of autonomously sensing multi-dimensional signals i.e. images, and implementing

complex signal processing algorithms. Compared to a traditional WSN, the visual sensing in a VSN significantly enhances the level of detail in the acquired data and consequently increases feasibility for a diverse range of applications. However, visual data requires high bandwidth for transmission and thus, results in high energy consumption. In order to guarantee suitable levels of reliability within VSNs, the need to maintain targeted QoI thresholds is paramount. Due to the resource constrained nature of VSNs, achieving targeted QoI thresholds is a challenging task. Therefore, the algorithms utilised during the design and implementation phases of VSNs must minimise the energy consumption for an improved network lifetime. It is pertinent to characterise visual sensing nodes' FoV for optimised sensing range estimation and reliable monitoring of the targeted sensing environment. Exploiting the diversity of sensing and processing tasks within a VSN, an efficient task classification scheme can be employed for energy conservation. Furthermore, in-node processing and self-reconfiguring for optimised feature redundancy removal can also contribute towards energy savings resulting in network lifetime maximisation.

A significant number of research studies are present in the existing literature to address issues relating to VSNs such as FoV identification, coverage estimation and enhancement, feature detection and extraction, visual data transmission, task classification, camera scheduling and self-reconfiguration etc. However, many existing solutions are found to be application-specific under particular resource constrained scenarios. Furthermore, many schemes have been developed for homogeneous VSNs which cannot be directly implemented within their heterogeneous counterparts. In most of the existing schemes, a simplified 2D sensing model is assumed along with a 2D sensing environment; whereas, many applications require 3D sensing within a 3D targeted sensing environment. Moreover, the existing schemes do not jointly consider the heterogeneous target QoI thresholds within the sensing environment and dynamic QoI-aware optimisation of visual data prior to transmission. In order to resolve the aforementioned issues within resource constrained VSNs, this thesis provided design solutions for energy efficient sensing

node characterisation with feature utilisation.

6.1.2 FoV Characterisation Framework

In this thesis, a novel approach for characterising a visual sensing node's FoV for both homogeneous and heterogeneous networks is presented. It is observed that an image captured by a visual sensing node is a function of the sensing range, the horizontal FoV and the vertical FoV. Therefore, it is pertinent to estimate the FoV parameters accurately and the sensing range optimally. In the case of static visual sensing nodes, horizontal and vertical FoV are constant; whereas, FoV parameters change in PTZ capable sensing nodes. Sensing range maximisation results in the reduction of total sensing nodes required to provide full coverage and consequently, minimises the overall energy consumption within the network. The proposed FoV characterisation method selects a suitable approach during the projection modelling phase to calculate the FoV parameters of homogeneous and heterogeneous networks. Global colour histogram-based features are employed for feature detection and extraction in YCbCr colour space. FoV characterisation criteria is proposed to estimate optimal range of visual sensing nodes based on the minimum object pixel occupancy, maximum allowable error tolerance and desired image quality. Within the context of heterogeneous VSN design, the proposed FoV characterisation criteria is supported by a soft decision based sensing range selection scheme. Furthermore, an adaptive task classification scheme is also proposed to efficiently distribute tasks between heterogeneous visual sensing nodes. Exploiting heterogeneity, the proposed soft decision based sensing range selection and the adaptive task classification schemes provide an energy efficient solution for visual sensing node's spatial coverage maximisation leading to an improved network lifetime without compromising on the desired reliability.

The performance of the proposed FoV characterisation framework is evaluated through experiments and simulations. Assuming static visual sensing nodes, the

horizontal and vertical FoV are calculated by experiments utilising the projection model and the estimation accuracy has been verified. Afterwards, the probability of a pixel within the captured image belonging to an object of interest has been estimated and utilised for feature detection and extraction. Once the object of interest is extracted from the captured image with feature utilisation, the accuracy of the proposed method to calculate the object pixel occupancy is verified through an experiment. It is observed that an increase in the sensing range results in the reduction of object pixel occupancy. Therefore, it is pertinent to estimate an optimal sensing range based on the application's object pixel occupancy requirement for visual sensing node's FoV characterisation. Furthermore, an experiment was conducted to evaluate the estimation error based FoV characterisation method. For this purpose, an application which estimates the diameter of an object extracted from the visual data was considered. It was noticed from the experiment that as the sensing range increases, the inaccuracy in feature detection and extraction increases and thus results in a higher estimation error. Therefore, the estimation error based FoV characterisation method can be utilised for applications demanding a particular level of estimation reliability. Another method for FoV characterisation based on the desired image quality is evaluated through an experiment. PSNR is used as an image quality assessment technique and it is observed that as the sensing range increases, the resulting image quality decreases. This method is suitable for applications where image quality is paramount for FoV characterisation.

The robustness of the proposed FoV characterisation framework for homogeneous networks has been analysed by considering various applications. It is found that the proposed framework provides an application-aware solution to the FoV characterisation problem by optimising the visual sensing node's spatial coverage leading to a substantial reduction in the energy consumption. Heterogeneous networks provide much more flexibility to the design engineer compared to the homogeneous networks due to the presence of different kinds of sensing nodes within the network.

The performance of the proposed heterogeneous network design solution has been evaluated by considering a surveillance application with two tasks: face detection and occluded target tracking. Utilising the feedback provided by the proposed soft decision based sensing range selection scheme with the proposed task classification scheme, optimised node configurations were obtained for a resource constrained heterogeneous network. It has been observed that the proposed solution leads to 49.8% energy savings compared to the trivial design solution and prolongs the heterogeneous network's lifetime by two-fold. In order to enhance the reliability of the proposed FoV characterisation framework, an analysis of system failure has been presented to predict and minimise the network failure probability. The energy efficiency of the proposed FoV characterisation framework demonstrates that it can be utilised during the network design and calibration phase to achieve an application-aware solution.

6.1.3 QoI-Aware Self-Reconfiguration Scheme

Since visual sensing nodes provide rich information of the scene within their FoV, redundant features can be removed prior to transmission in VSNs. However, such removal may affect the quality of the visual data. Therefore, the selection of an optimised visual sensing node's configuration for feature redundancy removal is required. Due to the dynamic nature of information within the visual data acquired by a sensing node, static node configurations are energy inefficient for resource constrained scenarios. Therefore, the implementation of self-reconfiguration schemes within visual sensing nodes to dynamically select optimised configurations can result in substantial energy savings. This thesis presented a dynamic self-reconfiguration scheme to enhance the energy efficiency of visual sensing nodes with target QoI threshold based optimisation. In the proposed scheme, QoI is characterised by the quality of the visual data and quantified by a PSNR based

representative metric. In VSNs, the energy consumption during the communication phase is significantly higher than the processing phase. Therefore, the configuration of a visual sensing node for optimisation is considered to be the amount and quality of the visual data for transmission. Visual sensing node management algorithms are classified into centralised and distributed categories; nevertheless, the latter are preferred due to the scalability and complexity issues of the former. Moreover, distributed algorithms are suitable for bandwidth constrained visual sensor networks.

Hence, considering benefits of the distributed approach, a decision making strategy is devised for visual sensing nodes to select optimised configuration independently, consequently, accelerating the decision making process. The distributed nature of the proposed scheme is supported by the training and calibration of visual sensing nodes which take place in the network pre-deployment phase. For a given application, the training and calibration process is initiated by the selection of a suitable dataset. The feature redundancy within visual data is characterised by the appearance of objects within visual sensing node's FoV. Since changing the sensor-to-object distance modifies the number of redundant features. Therefore, in the training and calibration process, the appearance of objects within visual sensing node's FoV is modelled. Afterwards, employing a suitable dynamic compression range, redundant features are removed from the acquired visual data. As discussed earlier, feature redundancy removal may introduce undesirable degradation in QoI; therefore, the resulting quality must be taken into account to achieve an acceptable level of reliability for a given application. In the proposed approach, the quality is quantified by a QoI index which incorporates the system's dynamic PSNR range. For a range set of dynamic sensor-to-object distances and compression ratios, QoI index is utilised to obtain the quality metrics for each test image from the dataset. Compressive calibration matrix is calculated which serves as a reference to gauge the impact of feature redundancy removal on the quality of the acquired visual data. Thereupon, each visual sensing node learns the compressive calibration

matrix at the termination of the training and calibration process.

An in-node processing model is proposed to support visual sensing nodes' self-reconfiguration in the post-deployment phase. After the acquisition of visual data, a suitable feature detection scheme can be employed depending on the given application to extract the object of interest. A target-driven approach is considered where the extracted object of interest is processed further prior to transmission only if its pixel occupancy is greater than a given threshold. The level of detail in the visual data is a function of the sensor-to-object distance. Hence, utilising the output of the feature detection and extraction process, the sensor-to-object distance is estimated. A self-reconfiguration scheme is proposed to support the in-node processing model. The proposed reconfiguration scheme is a function of the target QoI threshold to be achieved based on the application design criteria. In order to minimise energy consumption, the spatial location of an object within a visual sensing node's FoV is exploited. Each time an object is detected, the respective visual sensing node utilises the knowledge acquired from the learning process to calculate the active compressive calibration matrix and the transmission energy cost matrix. Thereupon, the visual sensing node is dynamically reconfigured by maximising the level of feature redundancy removal leading to minimum possible energy consumption while achieving the target QoI threshold.

Simulations were carried out to evaluate the performance of the proposed self-reconfiguration scheme. The LDHF dataset is used during the training and calibration process as it provides a variety of facial images captured with various sensor-to-object distances; thus, making it suitable to evaluate the proposed scheme. An image subset is extracted randomly from the data set. The proposed object modelling approach is employed to model the appearance of faces contained in the test images for a range set of sensor-to-object distances. The resulting QoI index is observed and it is found that as the sensor-to-object distance increases, the QoI index decreases. It is also noticed that the decaying characteristics of all the test images are identical. 2D-DWT and 2D-DCT schemes are considered to support

the process of redundant feature removal due to their suitability for visual sensing nodes. Compressive calibration matrix is obtained after object appearance modelling and redundant feature removal with 2D-DWT and 2D-DCT. It is observed that an increase in the level of redundancy removal leads to a reduction in QoI index. Therefore, selecting inappropriate sensing node configuration for feature redundancy removal can degrade the performance reliability; hence, the need for a self-reconfiguration scheme to select optimised node configurations is justified. It is found from the resulting QoI that the scheme supported by DWT is much more efficient as compared to the DCT based scheme. The energy efficiency of the proposed QoI-aware self-reconfiguration scheme as compared to the conventional scheme is quantified in terms of the transmission energy cost. The average transmission energy cost incurred per node to achieve various homogeneous QoI target thresholds with the proposed and conventional schemes supported by 2D-DWT and 2D-DCT are calculated and compared. It is observed from the comparison that for any given target QoI threshold, the proposed QoI-aware self-reconfiguration scheme minimises the transmission energy cost and thus, enhances the energy efficiency of visual sensing nodes within a resource constrained network.

6.1.4 Unified Framework with Heterogeneous QoI Realisation

In order to support intelligent sensing within VSNs, visual sensing nodes are required to be aware of the sensing environment within their FoV. This thesis presented a 3D coverage modelling scheme for visual sensing nodes to dynamically obtain their FoV information. Due to the time-varying nature of the targets' characteristics monitored within a visual sensing node's FoV, heterogeneous QoI realisation in VSNs is more suitable as compared to its homogeneous counterpart. Energy efficiency within the network can be achieved by dividing the overall sensing environment into smaller sub-regions and utilising heterogeneous target

QoI thresholds. Based on its 3D spatial coverage coordinates, each visual sensing node within the network needs to be allocated a suitable local target QoI threshold. Therefore, a QoI-centric scheme is proposed in this thesis for the classification of visual sensing nodes into a suitable level of heterogeneity. A unified framework is developed by incorporating the 3D coverage modelling scheme, QoI-centric node classification scheme and the in-node processing model. The proposed unified framework enhances the sensing and processing intelligence of VSNs by self-reconfiguring the sensing nodes based on their 3D FoV and given heterogeneous target QoI thresholds. Since maintaining a particular acceptable reliability level is paramount in many VSN applications, an analytical model is formulated to gauge the performance reliability of the proposed unified framework as a function of the targeted and delivered QoI thresholds. Depending on the given application, the analytical model assists the system design engineers to attain the desired level of reliability by fine-tuning to the required QoI thresholds.

The energy efficiency and reliability of the proposed unified framework is evaluated through experiments. A 3D sensing environment is considered for visual sensing nodes deployment and the degree of heterogeneity up to 2 is considered. The proposed 3D coverage modelling and QoI-centric node classification schemes are employed to assign each visual sensing node a local target QoI threshold. The energy consumption of the proposed unified framework is compared with the existing state-of-the art techniques utilising 2D-DWT and 2D-DCT based compression schemes. A comparison with tight and relaxed heterogeneous target QoI thresholds is presented and it is observed that the proposed framework results in significant amount of energy savings compared to its counterparts leading to an improved network lifetime. From the analysis of the proposed framework's performance reliability, it is noticed that higher QoI thresholds lead to an improved reliability. Moreover, it is found that the proposed unified framework provides suitable information retrieval reliability. Hence, the utilisation of the proposed framework in resource constrained scenarios is justified due to its energy efficiency

and tuneable reliability.

6.2 Research Limitations

Although the proposed frameworks result in substantial energy savings and demonstrate their suitability for utilisation in resource constrained VSNs, the following aspects of this research require further investigation.

- Colour features are used in the proposed FoV characterisation framework for object detection and feature extraction due to their implementation simplicity and low storage requirement. However, colour features are sensitive to lighting changes in the environment. Therefore, the utilisation of low-level colour features with suitable high-level features can be considered to enhance the efficiency of the feature extraction process.

- The self-reconfiguration scheme proposed in this thesis provides a trade-off model between reliability and energy efficiency. It results in optimisation based on the QoI target thresholds given by the system design engineers. However, further investigation is required to automate the process of QoI target thresholds calculation depending on the application and criticality of events being monitored by the sensing nodes.

- The performance of the proposed self-reconfiguration scheme is evaluated by using the LDHF dataset which contains facial images only. Therefore, datasets containing different types of objects can be used during the training and calibration process in the pre-deployment phase to enhance the intelligence of the visual sensing nodes.

6.3 Future Directions

This thesis contributes to the area of VSN design and implementation by resolving some issues faced due to its resource constrained nature through optimised visual sensing nodes characterisation and feature utilisation. There are several strategies for future extension of the work presented in this thesis which comprise of the following:

6.3.1 Resource-Aware Task Classification and Self-Reconfiguration

Task allocation is a critical issue in VSNs due the complex nature of sensing and processing tasks. Resource-awareness can be introduced in visual sensing nodes to dynamically manage the available resources in an energy efficient manner. Therefore, in addition to the task classification scheme presented in the thesis, the residual energy at a visual sensing node can be considered for energy efficient task allocation. Furthermore, the selection of a suitable visual sensing node configuration for feature redundancy removal can be based on the residual energy. As a visual sensing node's residual energy falls below a pre-determined threshold, the target QoI thresholds can be reduced in order to maximise feature redundancy removal, resulting in transmission energy cost minimisation and network lifetime maximisation.

6.3.2 Node Deployment Technique to Achieve Network Barrier Coverage

Within the context of sensor networks, barrier coverage utilises the sensing nodes to monitor the boundaries of critical infrastructures. The main aim of barrier coverage is to guarantee detection of intruders crossing the barrier of the network.

The majority of the research studies found in literature on barrier coverage consider WSNs and thus, assume sensing nodes having an isotropic sensing model. However, many applications require VSNs to be employed, which comprise of visual sensing nodes having a directional sensing model. Therefore, the existing schemes for barrier coverage assuming an isotropic sensing model are not feasible for such applications. Thus, for applications utilising VSNs, the proposed 3D coverage modelling scheme can be considered for the development of a reliable method to achieve network barrier coverage.

6.3.3 Coverage and Redundancy Management

Coverage maximisation and redundancy management is a challenging task in resource constrained VSNs. Utilisation of the proposed 3D coverage modelling scheme to develop reconfiguration models for coverage and redundancy optimisation of PTZ capable visual sensing nodes can be considered. This is expected to provide energy efficient solutions for collaborative management of visual sensing nodes' orientation and FoV parameters based on the criticality of events to prolong the network lifetime.

6.3.4 Reliable Monitoring for Health Care Applications

Although, a number of research studies have been conducted for health care applications; however, only a few of them consider the joint utilisation of scalar and visual sensors for cooperative and intelligent sensing that leads to an enhanced overall system reliability. Therefore, the proposed FoV characterisation scheme can be employed for health care applications (such as monitoring elderly well-being) to obtain optimised visual sensing node configurations. Furthermore, the proposed self-reconfiguration scheme can be utilised with a Wireless Body Area Network (WBAN) comprising various sensor types such as blood pressure, pulse

etc. The information obtained from WBAN can be considered for reconfiguring the visual sensing nodes such that a particular required level of reliability is maintained.

6.4 Routing Protocol for Real-time Visual Data Delivery

Real-time multimedia delivery is one of the key objectives in next-generation of wireless networks [204, 205]. Although, many research studies are present in the existing literature for data routing in WSNS; however, the conventional WSN routing protocols are not suitable for visual content delivery in VSNs. Therefore, an energy efficient routing protocol can be developed and incorporated within the proposed unified framework for applications where real-time visual data delivery is desirable. Furthermore, the visual data processing time can be characterised for the proposed self-reconfiguration scheme. Such characterisation is expected to support the routing protocol in delivering real-time visual data to the sink node by resolving the synchronisation problems.

6.5 Practical Implementation of the Proposed Schemes

The experiments presented in this thesis to evaluate the performance of the proposed task classification, node classification and self-reconfiguration schemes are conducted utilising the MATLAB simulation platform. One of the strategies for future work can be the practical implementation of the proposed schemes within visual sensing nodes to evaluate the network performance within the context of energy efficiency and reliability. Furthermore, feasible solutions for multi-objective

optimisation to self-reconfigure the visual sensing nodes can be obtained and implemented for VSN performance enhancement.

References

[1] Y. Charfi, N. Wakamiya, and M. Murata, "Challenging issues in visual sensor networks," *IEEE Wireless Communications*, vol. 16, no. 2, pp. 44–49, 2009.

[2] S. Soro and W. Heinzelman, "A survey of visual sensor networks," *Advances in Multimedia*, vol. 2009, 2009.

[3] N. B. Bo *et al.*, "Human mobility monitoring in very low resolution visual sensor network," *Sensors*, vol. 14, no. 11, pp. 20 800–20 824, 2014.

[4] Y. Cho, S. O. Lim, and H. S. Yang, "Collaborative occupancy reasoning in visual sensor network for scalable smart video surveillance," *IEEE Transactions on Consumer Electronics*, vol. 56, no. 3, pp. 1997–2003, Aug 2010.

[5] X. Wang, S. Wang, and D. Bi, "Distributed visual-target-surveillance system in wireless sensor networks," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 39, no. 5, pp. 1134–1146, Oct 2009.

[6] K. Abas, C. Porto, and K. Obraczka, "Wireless smart camera networks for the surveillance of public spaces," *Computer*, vol. 47, no. 5, pp. 37–44, 2014.

[7] A. Filonenko and K.-H. Jo, "Visual surveillance with sensor network for accident detection," in *39th Annual Conference of the IEEE Industrial Electronics Society (IECON)*, Nov 2013, pp. 5516–5521.

[8] N. Ahmad, K. Khursheed, M. Imran, N. Lawal, and M. O'Nils, "Modeling and verification of a heterogeneous sky surveillance visual sensor network," *International Journal of Distributed Sensor Networks*, vol. 2013, 2013.

[9] A. O. Ercan, A. E. Gamal, and L. J. Guibas, "Object tracking in the presence of occlusions using multiple cameras: A sensor network approach," *ACM Transactions on Sensor Networks (TOSN)*, vol. 9, no. 2, pp. 16:1–16:36, 2013.

[10] J. C. SanMiguel and A. Cavallaro, "Cost-aware coalitions for collaborative tracking in resource-constrained camera networks," *IEEE Sensors Journal*, vol. 15, no. 5, pp. 2657–2668, 2015.

[11] S. Fleck and W. Straßer, "Smart camera based monitoring system and its application to assisted living," *Proceedings of the IEEE*, vol. 96, no. 10, pp. 1698–1714, 2008.

[12] M. Brezovan and C. Badica, "A review on vision surveillance techniques in smart home environments," in *2013 19th International Conference on Control Systems and Computer Science (CSCS)*, May 2013, pp. 471–478.

[13] F. Deboeverie, J. Hanca, R. Kleihorst, A. Munteanu, and W. Philips, "A low-cost visual sensor network for elderly care," *SPIE NEWSROOM*, no. 1, 2014.

[14] M. Eldib, F. Deboeverie, W. Philips, and H. Aghajan, "Sleep analysis for elderly care using a low-resolution visual sensor network," in *Human Behavior Understanding*. Springer, 2015, pp. 26–38.

[15] S. Soro, "Application-aware resource management in wireless and visual sensor networks," Ph.D. dissertation, University of Rochester, 2007.

[16] S. M. B. Malek, M. M. Sadik, and A. Rahman, "On balanced k-coverage in visual sensor networks," *Journal of Network and Computer Applications*, vol. 72, pp. 72–86, 2016.

[17] J. Simonjan and B. Rinner, "Autonomous, lightweight calibration of visual sensor networks with dense coverage," in *2016 IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops)*, Mar 2016, pp. 1–4.

[18] Q. Ling and Z. Tian, "Decentralized sparse signal recovery for compressive sleeping wireless sensor networks," *IEEE Transactions on Signal Processing*, vol. 58, no. 7, pp. 3816–3827, Jul 2010.

[19] F. Meshkati, H. V. Poor, S. C. Schwartz, and R. V. Balan, "Energy-efficient resource allocation in wireless networks with quality-of-service constraints," *IEEE Transactions on Communications*, vol. 57, no. 11, pp. 3406–3414, 2009.

[20] R. Zhu, M. Ma, Y. Zhang, and J. Hu, "Collaborative wireless sensor networks and applications," *International Journal of Distributed Sensor Networks*, vol. 2015, 2015.

[21] H. Abid and S. Qaisar, "Distributed video coding for wireless visual sensor networks using low power huffman coding," in *2010 44th Annual Conference on Information Sciences and Systems (CISS)*. IEEE, 2010, pp. 1–6.

[22] C. Piciarelli, L. Esterle, A. Khan, B. Rinner, and G. L. Foresti, "Dynamic reconfiguration in camera networks: a short survey," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 26, no. 5, pp. 965–977, 2016.

[23] J. Edwards, A. Bahjat, Y. Jiang, T. Cook, and T. F. La Porta, "Quality of information-aware mobile applications," *Pervasive and Mobile Computing*, vol. 11, pp. 216–228, 2014.

[24] Z. Song, C. H. Liu, J. Wu, J. Ma, and W. Wang, "QoI-aware multitask-oriented dynamic participant selection with budget constraints," *IEEE Transactions on Vehicular Technology*, vol. 63, no. 9, pp. 4618–4632, 2014.

[25] C. H. Liu, J. Fan, J. W. Branch, and K. K. Leung, "Toward QoI and energy-efficiency in internet-of-things sensory environments," *IEEE Transactions on Emerging Topics in Computing*, vol. 2, no. 4, pp. 473–487, 2014.

[26] E. Gelenbe and L. Hey, "Quality of information: An empirical approach," in *2008 5th IEEE International Conference on Mobile Ad Hoc and Sensor Systems*, Sep 2008, pp. 730–735.

[27] S. S. Rizvi, K. Patel, and C. Patel, "Use of self-adaptive methodology in wireless sensor networks for reducing the energy consumption," in *Novel Algorithms and Techniques In Telecommunications, Automation and Industrial Electronics*. Springer, 2008, pp. 520–525.

[28] S. M. Aziz and D. M. Pham, "Energy efficient image transmission in wireless multimedia sensor networks," *IEEE communications letters*, vol. 17, no. 6, pp. 1084–1087, 2013.

[29] K.-Y. Chow, K.-S. Lui, and E. Y. Lam, "Efficient on-demand image transmission in visual sensor networks," *EURASIP Journal on Applied Signal Processing*, vol. 2007, no. 1, pp. 225–225, 2007.

[30] G. J. Pottie and W. J. Kaiser, "Wireless integrated network sensors," *Communications of the ACM*, vol. 43, no. 5, pp. 51–58, 2000.

[31] F. Shebli, I. Dayoub, A. O. M'foubat, A. Rivenq, and J. Rouvaen, "Minimizing energy consumption within wireless sensors networks using optimal transmission range between nodes," in *IEEE International Conference on Signal Processing and Communications (ICSPC)*. IEEE, 2007, pp. 105–108.

[32] T. Melodia, D. Pompili, and I. F. Akyildiz, "A communication architecture for mobile wireless sensor and actor networks," in *3rd Annual IEEE Communications Society on Sensor and Ad Hoc Communications and Networks (SECON'06)*, vol. 1. IEEE, 2006, pp. 109–118.

[33] M. Y. Mowafi and W. A. Aljoby, "Lightweight target classification for wireless multimedia sensor networks," *International Journal of Image and Data Fusion*, vol. 4, no. 4, pp. 293–307, 2013.

[34] T. Semertzidis, K. Dimitropoulos, A. Koutsia, and N. Grammalidis, "Video sensor network for real-time traffic monitoring and surveillance," *IET intelligent transport systems*, vol. 4, no. 2, pp. 103–112, 2010.

[35] A. Salaheldin, M. ElSayed, A. Alsebai, N. El Gayar, and M. ElHelw, "Change analysis for gait impairment quantification in smart environments," in *2010 International Conference on Autonomous and Intelligent Systems (AIS)*. IEEE, 2010, pp. 1–6.

[36] R. Kumar, *Research methodology: a step-by-step guide for beginners, 3rd ed.* Sage, 2011.

[37] M. Saunders, P. Lewis, and A. Thornhill, *Research methods for business students*. Prentice Hall, 2009.

[38] R. K. Yin, *Case study research: Design and methods*. Sage publications, 2013.

[39] C. Robson and K. McCartan, *Real world research*. John Wiley & Sons, 2016.

[40] M. Easterby-Smith, R. Thorpe, and P. R. Jackson, *Management research*. Sage, 2012.

[41] A. Booth, A. Sutton, and D. Papaioannou, *Systematic approaches to a successful literature review*. Sage, 2016.

[42] J. Jesson, L. Matheson, and F. M. Lacey, *Doing your literature review: Traditional and systematic techniques*. Sage, 2011.

[43] F. G. Yap and H.-H. Yen, "A survey on sensor coverage and visual data capturing/processing/transmission in wireless visual sensor networks," *Sensors*, vol. 14, no. 2, pp. 3506–3527, 2014.

[44] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless sensor networks: a survey," *Computer networks*, vol. 38, no. 4, pp. 393–422, 2002.

[45] J. Boice *et al.*, "Meerkats: A power-aware, self-managing wireless camera network for wide area monitoring," in *Proc. Workshop on Distributed Smart Cameras*. Citeseer, 2006, pp. 393–422.

[46] W.-c. Feng, E. Kaiser, W. C. Feng, and M. L. Baillif, "Panoptes: scalable low-power video sensor networking technologies," *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, vol. 1, no. 2, pp. 151–167, 2005.

[47] M. Zhang and W. Cai, "Vision mesh: A novel video sensor networks platform for water conservancy engineering," in *2010 3rd IEEE international conference on Computer science and information technology (ICCSIT)*, vol. 4. IEEE, 2010, pp. 106–109.

[48] M. Rahimi *et al.*, "Cyclops: in situ image sensing and interpretation in wireless sensor networks," in *Proceedings of the 3rd international conference on Embedded networked sensor systems*. ACM, 2005, pp. 192–204.

[49] M. A. Guvensan and A. G. Yavuz, "On coverage issues in directional sensor networks: A survey," *Ad Hoc Networks*, vol. 9, no. 7, pp. 1238–1255, 2011.

[50] J. O'Rourke, *Art gallery theorems and algorithms*. Oxford University Press Oxford, vol. 57.

[51] H.-H. Yen, "Efficient visual sensor coverage algorithm in wireless visual sensor networks," in *2013 9th International Wireless Communications and Mobile Computing Conference (IWCMC)*. IEEE, 2013, pp. 1516–1521.

[52] E. Yildiz, K. Akkaya, E. Sisikoglu, M. Sir, and I. Guneydas, "Camera deployment for video panorama generation in wireless visual sensor networks," in *2011 IEEE International Symposium on Multimedia (ISM)*. IEEE, 2011, pp. 595–600.

[53] B. Debaque, R. Jedidi, and D. Prevost, "Optimal video camera network deployment to support security monitoring," in *12th International Conference on Information Fusion*. IEEE, 2009, pp. 1730–1736.

[54] Y. E. Osais, M. St-Hilaire, and R. Y. Fei, "Directional sensor placement with optimal sensing range, field of view and orientation," *Mobile Networks and Applications*, vol. 15, no. 2, pp. 216–225, 2010.

[55] S. Halder and A. Ghosal, "A location-wise predetermined deployment for optimizing lifetime in visual sensor networks," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 26, no. 6, pp. 1131–1145, 2016.

[56] L. Hailong, P. Vaibhav, and P. A. Dharma, "Deployment optimization strategy for a two-tier wireless visual sensor network," *Wireless Sensor Network*, vol. 2012, 2012.

[57] J. J. Gonzalez-Barbosa, T. Garcia-Ramirez, J. Salas, J. B. Hurtado-Ramos, and J. d. J. Rico-Jimenez, "Optimal camera placement for total coverage," in *IEEE International Conference on Robotics and Automation (ICRA'09)*. IEEE, 2009, pp. 844–848.

[58] U. M. Erdem and S. Sclaroff, "Automated camera layout to satisfy task-specific and floor plan-specific coverage requirements," *Computer Vision and Image Understanding*, vol. 103, no. 3, pp. 156–169, 2006.

[59] K.-Y. Chow, K.-S. Lui, and E. Y. Lam, "Maximizing angle coverage in visual sensor networks," in *2007 IEEE International Conference on Communications*. IEEE, 2007, pp. 3516–3521.

[60] M. A. Molina-Cabello, E. López-Rubio, R. M. Luque-Baena, E. Domínguez, and K. Thurnhofer-Hemsi, "Neural controller for PTZ cameras based on nonpanoramic foreground detection," in *2017 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2017, pp. 404–411.

[61] M. Karakaya and H. Qi, "Coverage estimation for crowded targets in visual sensor networks," *ACM Transactions on Sensor Networks (TOSN)*, vol. 8, no. 3, pp. 26:1–26:22, 2012.

[62] F. G. Yap and H.-H. Yen, "Novel visual sensor coverage and deployment in time aware PTZ wireless visual sensor networks," *Sensors*, vol. 17, no. 1, 2016.

[63] Y. Bai, "Feature-based image comparison and its application in wireless visual sensor networks," *EECS Publications*, 2011.

[64] K. Castleman, *Digital Image Processing*. Prentice Hall, 1996.

[65] C. Direkoğlu, "Feature extraction via heat flow analogy," Ph.D. dissertation, University of Southampton, 2009.

[66] A. Webb and K. Copsey, *Statistical pattern recognition*. Wiley, 2011.

[67] I. Fodor, "A survey of dimension reduction techniques," *Technical Report UCRL-ID-148494, Center for Applied Scientific Computing, Lawrence Livermore National Laboratory*, 2002.

[68] D. Masip, L. Kuncheva, and J. Vitrià, "An ensemble-based method for linear feature extraction for two-class problems," *Pattern Analysis and Applications*, vol. 8, no. 3, pp. 227–237, 2005.

[69] L. Jimenez-Rodriguez, E. Arzuaga-Cruz, and M. Vélez-Reyes, "Unsupervised linear feature-extraction methods and their effects in the classification of high-dimensional data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 2, pp. 469–483, 2007.

[70] T. Aribarg, S. Supratid, and C. Lursinsap, "Contemporary classification on medical data based on non-linear feature extraction," in *International Conference on Computational Science and Its Applications (ICCSA)*. IEEE, 2009, pp. 17–23.

[71] J. Yang, A. Frangi, J. Yang, D. Zhang, and Z. Jin, "KPCA plus LDA: a complete kernel fisher discriminant framework for feature extraction and recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 2, pp. 230–244, 2005.

[72] M. Yang, K. Kpalma, and J. Ronsin, "A survey of shape feature extraction techniques," *Pattern Recognition*, pp. 43–90, 2008.

[73] M. Nixon and A. Aguado, *Feature Extraction and Image Processing for Computer Vision*. Academic Press, 2012.

[74] D. Lisin, M. Mattar, M. Blaschko, E. Learned-Miller, and M. Benfield, "Combining local and global image features for object class recognition," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition-Workshops*. IEEE, 2005, pp. 47–47.

[75] K. Mikolajczyk, B. Leibe, and B. Schiele, "Local features for object class recognition," in *Tenth IEEE International Conference on Computer Vision*, vol. 2. IEEE, 2005, pp. 1792–1799.

[76] J. Kovac, P. Peer, and F. Solina, "Human skin color clustering for face detection," in *The IEEE Region 8 EUROCON 2003. Computer as a Tool.*, vol. 2, Sep 2003, pp. 144–148.

[77] S. Zakariya, R. Ali, and N. Ahmad, "Combining visual features of an image at different precision value of unsupervised content based image retrieval," in *IEEE International Conference on Computational Intelligence and Computing Research (ICCIC)*. IEEE, 2010, pp. 1–4.

[78] B. Lei, E. Hendriks, and M. Reinders, "On feature extraction from images," *MCCWS project, Information and Communication Theory Group*, 1999.

[79] B. Bhatt and Z. Shah, "Face feature extraction techniques: A survey," in *National Conference on Recent Trends in Engineering and Technology*, 2011.

[80] M. Magno, D. Brunelli, P. Zappi, and L. Benini, "A solar-powered video sensor node for energy efficient multimodal surveillance," in *2008 11th EUROMICRO Conference on Digital System Design Architectures, Methods and Tools*, Sept 2008, pp. 512–519.

[81] E. O'Connor, A. F. Smeaton, and N. E. O'Connor, "A multi-modal event detection system for river and coastal marine monitoring applications," in *OCEANS 2011 IEEE - Spain*, Jun 2011, pp. 1–10.

[82] D. Xie, T. Yan, D. Ganesan, and A. Hanson, "Design and implementation of a dual-camera wireless sensor network for object retrieval," in *Proceedings of the 7th international conference on Information processing in sensor networks*. IEEE Computer Society, 2008, pp. 469–480.

[83] H. Fu, H. Ma, and L. Liu, "Robust human detection with low energy consumption in visual sensor network," in *2011 Seventh International Conference on Mobile Ad-hoc and Sensor Networks (MSN)*, Dec 2011, pp. 91–97.

[84] Y.-S. Lee and W.-Y. Chung, "Visual sensor based abnormal event detection with moving shadow removal in home healthcare applications," *Sensors*, vol. 12, no. 1, pp. 573–584, 2012.

[85] S. Najjar-Ghabel and S. Yousefi, "Enhancing performance of face detection in visual sensor networks with a dynamic-based approach," *Wireless Personal Communications*, Aug 2017.

[86] E. Eriksson, G. Dán, and V. Fodor, "Real-time distributed visual feature extraction from video in sensor networks," in *2014 IEEE International Conference on Distributed Computing in Sensor Systems*. IEEE, 2014, pp. 152–161.

[87] N. Monteiro, C. Brites, F. Pereira, and J. Ascenso, "Multi-view distributed source coding of binary features for visual sensor networks," in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2016, pp. 2807–2811.

[88] L. Baroffio, A. Canclini, M. Cesana, A. Redondi, and M. Tagliasacchi, "Briskola: Brisk optimized for low-power arm architectures," in *IEEE International Conference on Image Processing*, 2014.

[89] S. Leutenegger, M. Chli, and R. Y. Siegwart, "BRISK: Binary robust invariant scalable keypoints," in *2011 IEEE International Conference on Computer Vision (ICCV)*. IEEE, 2011, pp. 2548–2555.

[90] A. Redondi, M. Cesana, and M. Tagliasacchi, "Rate-accuracy optimization in visual wireless sensor networks," in *2012 19th IEEE International Conference on Image Processing*. IEEE, 2012, pp. 1105–1108.

[91] A. Tharwat, T. Gaber, Y. M. Awad, N. Dey, and A. E. Hassanien, *Plants Identification Using Feature Fusion Technique and Bagging Classifier*. Cham: Springer International Publishing, 2016, pp. 461–471.

[92] J. Yue, Z. Li, L. Liu, and Z. Fu, "Content-based image retrieval using color and texture fused features," *Mathematical and Computer Modelling*, vol. 54, no. 3, pp. 1121 – 1127, 2011.

[93] H. Zhu, F. Meng, J. Cai, and S. Lu, "Beyond pixels: A comprehensive survey from bottom-up to semantic image segmentation and cosegmentation," *Journal of Visual Communication and Image Representation*, vol. 34, pp. 12 – 27, 2016.

[94] R. Beare, "A locally constrained watershed transform," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 7, pp. 1063–1074, Jul 2006.

[95] P. F. Felzenszwalb and D. P. Huttenlocher, "Efficient graph-based image segmentation," *International Journal of Computer Vision*, vol. 59, no. 2, pp. 167–181, Sep 2004.

[96] D. Comaniciu and P. Meer, "Mean shift: a robust approach toward feature space analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 603–619, May 2002.

[97] T. N. A. Nguyen, J. Cai, J. Zhang, and J. Zheng, "Robust interactive image segmentation using convex active contours," *IEEE Transactions on Image Processing*, vol. 21, no. 8, pp. 3734–3743, Aug 2012.

[98] W. Yang, J. Cai, J. Zheng, and J. Luo, "User-friendly interactive image segmentation through unified combinatorial user inputs," *IEEE Transactions on Image Processing*, vol. 19, no. 9, pp. 2470–2479, Sep 2010.

[99] G. Kim and E. P. Xing, "On multiple foreground cosegmentation," in *2012 IEEE Conference on Computer Vision and Pattern Recognition*, Jun 2012, pp. 837–844.

[100] A. Joulin, F. Bach, and J. Ponce, "Multi-class cosegmentation," in *2012 IEEE Conference on Computer Vision and Pattern Recognition*, Jun 2012, pp. 542–549.

[101] P. Kennel, C. Fiorio, and F. Borne, "Supervised image segmentation using q-shift dual-tree complex wavelet transform coefficients with a texton approach," *Pattern Analysis and Applications*, vol. 20, no. 1, pp. 227–237, Feb 2017.

[102] J. Tighe and S. Lazebnik, "Superparsing - scalable nonparametric image parsing with superpixels," *International Journal of Computer Vision*, vol. 101, no. 2, pp. 329–349, Jan 2013.

[103] A. Amjad, A. Griffiths, and M. N. Patwary, "Multiple face detection algorithm using colour skin modelling," *IET Image Processing*, vol. 6, no. 8, pp. 1093–1101, Nov 2012.

[104] C. Bisdikian, "On sensor sampling and quality of information: A starting point," in *Fifth Annual IEEE International Conference on Pervasive Computing and Communications Workshops, 2007. PerCom Workshops' 07.* IEEE, 2007, pp. 279–284.

[105] C. Bisdikian, L. M. Kaplan, and M. B. Srivastava, "On the quality and value of information in sensor networks," *ACM Transactions on Sensor Networks (TOSN)*, vol. 9, no. 4, pp. 48:1–48:26, 2013.

[106] U. Hunkeler and P. Scotton, "A quality-of-information-aware framework for data models in wireless sensor networks," in *2008 5th IEEE International Conference on Mobile Ad Hoc and Sensor Systems*, Sep 2008, pp. 742–747.

[107] C. H. Liu, C. Bisdikian, J. W. Branch, and K. K. Leung, "QoI-aware wireless sensor network management for dynamic multi-task operations," in *2010 7th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON)*, Jun 2010, pp. 1–9.

[108] C. H. Liu, P. Hui, J. W. Branch, and B. Yang, "QoI-aware energy management for wireless sensor networks," in *2011 IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM Workshops)*, Mar 2011, pp. 8–13.

[109] M. A. Hossain, P. K. Atrey, and A. E. Saddik, "Modeling and assessing quality of information in multisensor multimedia monitoring systems," *ACM*

Transactions on Multimedia Computing, Communications, and Applications (TOMM), vol. 7, no. 1, pp. 3:1–3:30, 2011.

[110] H.-H. Yen, "Energy-aware image aggregation in wireless visual sensor networks," in *2011 IEEE Pacific Rim Conference on Communications, Computers and Signal Processing (PacRim)*. IEEE, 2011, pp. 572–577.

[111] A. Kurde and V. Ukani, "Aggregation in wireless multimedia sensor network," in *2013 Nirma University International Conference on Engineering (NUiCONE)*. IEEE, 2013, pp. 1–6.

[112] G. Mali and S. Misra, "Topology management-based distributed camera actuation in wireless multimedia sensor networks," *ACM Transactions on Autonomous and Adaptive Systems*, vol. 12, no. 1, pp. 2:1–2:33, Apr 2017.

[113] S. Colonnese, F. Cuomo, O. Damiano, V. De Pascalis, and T. Melodia, "On the adoption of multiview video coding in wireless multimedia sensor networks," in *2011 Wireless Advanced (WiAd)*. IEEE, 2011, pp. 218–223.

[114] X. Yang, Y. Wen, M. Zhang, and H. Zhao, "3D visual correlation model for wireless visual sensor networks," in *2017 IEEE/ACIS 16th International Conference on Computer and Information Science (ICIS)*. IEEE, 2017, pp. 75–80.

[115] H. ZainEldin, M. A. Elhosseini, and H. A. Ali, "Image compression algorithms in wireless multimedia sensor networks: A survey," *Ain Shams Engineering Journal*, vol. 6, no. 2, pp. 481–490, 2015.

[116] T. Ma, M. Hempel, D. Peng, and H. Sharif, "A survey of energy-efficient compression and communication techniques for multimedia in resource constrained systems," *IEEE Communications Surveys and Tutorials*, vol. 15, no. 3, pp. 963–972, 2013.

[117] A. Mammeri, B. Hadjou, and A. Khoumsi, "A survey of image compression algorithms for visual sensor networks," *ISRN Sensor Networks*, vol. 2012, 2012.

[118] Y. Wang, D. Wang, X. Zhang, J. Chen, and Y. Li, "Energy-efficient image compressive transmission for wireless camera networks," *IEEE Sensors Journal*, vol. 16, no. 10, pp. 3875–3886, 2016.

[119] Z.-y. Xiong, X.-p. Fan, S.-q. Liu, and Z. Zhong, "Low complexity image compression for wireless multimedia sensor networks," in *International Conference on Information Science and Technology*. IEEE, 2011, pp. 665–670.

[120] E. Sun, X. Shen, and H. Chen, "A low energy image compression and transmission in wireless multimedia sensor networks," *Procedia Engineering*, vol. 15, pp. 3604–3610, 2011.

[121] P. Wang, R. Dai, and I. F. Akyildiz, "A spatial correlation-based image compression framework for wireless multimedia sensor networks," *IEEE Transactions on Multimedia*, vol. 13, no. 2, pp. 388–401, 2011.

[122] X. Zhang, Y. Wang, D. Wang, and Y. Li, "Adaptive image compression based on compressive sensing for video sensor nodes," *Multimedia Tools and Applications*, Jul 2017.

[123] Y. Wang, Z. Liu, D. Wang, Y. Li, and J. Yan, "Anomaly detection and visual perception for landslide monitoring based on a heterogeneous sensor network," *IEEE Sensors Journal*, vol. 17, no. 13, pp. 4248–4257, 2017.

[124] I. Mehmood, M. Sajjad, W. Ejaz, and S. W. Baik, "Saliency-directed prioritization of visual data in wireless surveillance networks," *Information Fusion*, vol. 24, pp. 16–30, 2015.

[125] Y. He, I. Lee, and L. Guan, "Distributed algorithms for network lifetime maximization in wireless visual sensor networks," *IEEE Transactions on*

Circuits and Systems for Video Technology, vol. 19, no. 5, pp. 704–718, 2009.

[126] S. Ye, Y. Lin, and R. Li, "Energy-aware interleaving for robust image transmission over visual sensor networks," *IET wireless sensor systems*, vol. 1, no. 4, pp. 267–274, 2011.

[127] R. Dai, P. Wang, and I. F. Akyildiz, "Correlation-aware QoS routing with differential coding for wireless video sensor networks," *IEEE Transactions on Multimedia*, vol. 14, no. 5, pp. 1469–1479, 2012.

[128] D. Sarma, M. K. Tarafder, T. Choudhary, and K. K. Sarma, "An energy efficient image communication framework for bandlimited wireless sensor networks," in *2016 International Conference on Accessibility to Digital World (ICADW)*. IEEE, 2016, pp. 197–201.

[129] O. Alaoui-Fdili, F.-X. Coudoux, Y. Fakhri, P. Corlay, and D. Aboutajdine, "Energy-efficient joint video encoding and transmission framework for WVSN," *Multimedia Tools and Applications*, pp. 1–33, 2017.

[130] B. Dieber, C. Micheloni, and B. Rinner, "Resource-aware coverage and task assignment in visual sensor networks," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 21, no. 10, pp. 1424–1437, 2011.

[131] T. Hu, S. Messelodi, and O. Lanz, "Dynamic task decomposition for decentralized object tracking in complex scenes," *Computer Vision and Image Understanding*, vol. 134, pp. 89–104, 2015.

[132] E. Eriksson, V. Pacifici, and G. Dán, "Efficient distribution of visual processing tasks in multi-camera visual sensor networks," in *2015 IEEE International Conference on Multimedia & Expo Workshops (ICMEW)*. IEEE, 2015, pp. 1–6.

[133] F. Pletzer and B. Rinner, "Distributed task allocation for visual sensor networks: A market-based approach," in *2010 Fourth IEEE International Conference on Self-Adaptive and Self-Organizing Systems Workshop*, Sep 2010, pp. 59–62.

[134] C. Kyrkou, C. Laoudias, T. Theocharides, C. G. Panayiotou, and M. Polycarpou, "Adaptive energy-oriented multitask allocation in smart camera networks," *IEEE Embedded Systems Letters*, vol. 8, no. 2, pp. 37–40, 2016.

[135] A. Cenedese, F. Cerruti, M. Fabbro, C. Masiero, and L. Schenato, "Decentralized task assignment in camera networks," in *49th IEEE Conference on Decision and Control (CDC)*, Dec 2010, pp. 126–131.

[136] C. Ye, Y. Zheng, S. Velipasalar, and M. C. Gursoy, "Energy-aware and robust task (re)assignment in embedded smart camera networks," in *2013 10th IEEE International Conference on Advanced Video and Signal Based Surveillance*, Aug 2013, pp. 123–128.

[137] L. Cui and W. Jia, "Weighted capacitated popular matching for task assignment in multi-camera networks," in *Proceedings of the 2013 25th International Teletraffic Congress (ITC)*, Sep 2013, pp. 1–4.

[138] Y. Jin, S. Vural, A. Gluhak, and K. Moessner, "Dynamic task allocation in multi-hop multimedia wireless sensor networks with low mobility," *Sensors*, vol. 13, no. 10, pp. 13 998–14 028, 2013.

[139] P. Györke and B. Pataki, "Energy-aware measurement scheduling in WSNs used in AAL applications," *IEEE Transactions on Instrumentation and Measurement*, vol. 62, no. 5, pp. 1318–1325, 2013.

[140] M. Szczodrak, O. Gnawali, and L. P. Carloni, "Dynamic reconfiguration of wireless sensor networks to support heterogeneous applications," in *2013 IEEE International Conference on Distributed Computing in Sensor Systems*, May 2013, pp. 52–61.

[141] W. Horr, K. Lee, D. Hughes, S. Michiels, and W. Joosen, "A graph based approach to supporting reconfiguration in wireless sensor networks," in *2009 First International Conference on Networks Communications*, Dec 2009, pp. 326–331.

[142] H. Hoffmann, "Racing and pacing to idle: an evaluation of heuristics for energy-aware resource allocation," in *Proceedings of the Workshop on Power-Aware Computing and Systems*. ACM, 2013, pp. 13:1–13:5.

[143] U. A. Khan and B. Rinner, "A reinforcement learning framework for dynamic power management of a portable, multi-camera traffic monitoring system," in *2012 IEEE International Conference on Green Computing and Communications (GreenCom)*. IEEE, 2012, pp. 557–564.

[144] U. A. Khan and B. Rinner, "Online learning of timeout policies for dynamic power management," *ACM Transactions on Embedded Computing Systems (TECS)*, vol. 13, no. 4, pp. 96:1–96:25, 2014.

[145] C. Yu and G. Sharma, "Camera scheduling and energy allocation for lifetime maximization in user-centric visual sensor networks," *IEEE Transactions on Image Processing*, vol. 19, no. 8, pp. 2042–2055, 2010.

[146] D. R. Karuppiah, R. A. Grupen, Z. Zhu, and A. R. Hanson, "Automatic resource allocation in a distributed camera network," *Machine Vision and Applications*, vol. 21, no. 4, pp. 517–528, 2010.

[147] B. Dieber, L. Esterle, and B. Rinner, "Distributed resource-aware task assignment for complex monitoring scenarios in visual sensor networks," in *2012 Sixth International Conference on Distributed Smart Cameras (ICDSC)*, Oct 2012, pp. 1–6.

[148] P. Fu *et al.*, "An effective and robust decentralized target tracking scheme in wireless camera sensor networks," *Sensors*, vol. 17, no. 3, 2017.

[149] M. Kim, C.-M. Kyung, and K. Yi, "An energy management scheme for solar-powered wireless visual sensor networks toward uninterrupted operations," in *2013 International SoC Design Conference (ISOCC)*, Nov 2013, pp. 023–026.

[150] S. Kim, H. Lee, D. Jeon, and S. Lee, "Reduction in encoding redundancy for overlapped fovs over wireless visual sensor networks," *Digital Signal Processing*, vol. 50, pp. 135–149, 2016.

[151] J. Ai and A. A. Abouzeid, "Coverage by directional sensors in randomly deployed wireless sensor networks," *Journal of Combinatorial Optimization*, vol. 11, no. 1, pp. 21–41, 2006.

[152] Y. Morsly, N. Aouf, M. S. Djouadi, and M. Richardson, "Particle swarm optimization inspired probability algorithm for optimal camera network placement," *IEEE Sensors Journal*, vol. 12, no. 5, pp. 1402–1412, 2012.

[153] S. Indu, S. Chaudhury, N. R. Mittal, and A. Bhattacharyya, "Optimal sensor placement for surveillance of large spaces," in *2009 Third ACM/IEEE International Conference on Distributed Smart Cameras (ICDSC)*. IEEE, 2009, pp. 1–8.

[154] C. Piciarelli, C. Micheloni, and G. L. Foresti, "Occlusion-aware multiple camera reconfiguration," in *Proceedings of the Fourth ACM/IEEE International Conference on Distributed Smart Cameras*. ACM, 2010, pp. 88–94.

[155] K. R. Konda and N. Conci, "Optimal configuration of PTZ camera networks based on visual quality assessment and coverage maximization," in *2013 Seventh International Conference on Distributed Smart Cameras (ICDSC)*. IEEE, 2013, pp. 1–8.

[156] K. R. Konda, N. Conci, and F. D. Natale, "Global coverage maximization in PTZ-camera networks based on visual quality assessment," *IEEE Sensors Journal*, vol. 16, no. 16, pp. 6317–6332, Aug 2016.

[157] M. Rout and R. Roy, "Self-deployment of randomly scattered mobile sensors to achieve barrier coverage," *IEEE Sensors Journal*, vol. 16, no. 18, pp. 6819–6820, 2016.

[158] A. Del Bimbo, F. Dini, G. Lisanti, and F. Pernici, "Exploiting distinctive visual landmark maps in pan–tilt–zoom camera networks," *Computer Vision and Image Understanding*, vol. 114, no. 6, pp. 611–623, 2010.

[159] P. Natarajan, T. N. Hoang, K. H. Low, and M. Kankanhalli, "Decision-theoretic coordination and control for active multi-camera surveillance in uncertain, partially observable environments," in *2012 Sixth International Conference on Distributed Smart Cameras (ICDSC)*. IEEE, 2012, pp. 1–6.

[160] P. Natarajan, T. N. Hoang, Y. Wong, K. H. Low, and M. Kankanhalli, "Scalable decision-theoretic coordination and control for real-time active multi-camera surveillance," in *Proceedings of the International Conference on Distributed Smart Cameras*. ACM, 2014, pp. 38:1–38:6.

[161] C. Ding, A. A. Morye, J. A. Farrell, and A. K. Roy-Chowdhury, "Opportunistic sensing in a distributed PTZ camera network," in *2012 Sixth International Conference on Distributed Smart Cameras (ICDSC)*. IEEE, 2012, pp. 1–6.

[162] F. Z. Qureshi and D. Terzopoulos, "Planning ahead for PTZ camera assignment and handoff," in *Third ACM/IEEE International Conference on Distributed Smart Cameras, 2009. ICDSC 2009*. IEEE, 2009, pp. 1–8.

[163] F. Z. Qureshi and D. Terzopoulos, "Proactive PTZ camera control," in *Distributed Video Sensor Networks*. Springer, 2011, pp. 273–287.

[164] W. Starzyk and F. Z. Qureshi, "A negotiation protocol with conditional offers for camera handoffs," in *Proceedings of the International Conference on Distributed Smart Cameras*. ACM, 2014, pp. 17:1–17:7.

[165] Y. Wang and A. Cavallaro, "Coalition formation for distributed tracking in wireless camera networks," in *2015 12th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*. IEEE, 2015, pp. 1–6.

[166] D. A. Forsyth and J. Ponce, *Computer Vision: A modern approach*. Prentice-Hall, 2002.

[167] R. Dai and I. F. Akyildiz, "A spatial correlation model for visual information in wireless multimedia sensor networks," *IEEE Transactions on Multimedia*, vol. 11, no. 6, pp. 1148–1159, 2009.

[168] Y. Gui, F. Wu, X. Gao, and G. Chen, "Full-view barrier coverage with rotatable camera sensors," in *2014 IEEE/CIC International Conference on Communications in China (ICCC)*. IEEE, 2014, pp. 818–822.

[169] L. Viet Tran, "Efficient image retrieval with statistical color descriptors," 2003.

[170] J. Nascimento and J. Marques, "Performance evaluation of object detection algorithms for video surveillance," *IEEE Transactions on Multimedia*, vol. 8, no. 4, pp. 761–774, Aug 2006.

[171] P. Viola and M. Jones, "Robust real-time object detection," in *International Journal of Computer Vision*, 2001.

[172] T. E. Boult, R. J. Micheals, X. Gao, and M. Eckmann, "Into the woods: Visual surveillance of noncooperative and camouflaged targets in complex outdoor settings," *Proceedings of the IEEE*, vol. 89, no. 10, pp. 1382–1402, Oct 2001.

[173] C. Wren, A. Azarbayejani, T. Darrell, and A. Pentland, "Pfinder: real-time tracking of the human body," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 780–785, Jul 1997.

[174] C. Stauffer and W. E. L. Grimson, "Learning patterns of activity using real-time tracking," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 747–757, 2000.

[175] K. Gu, G. Zhai, X. Yang, and W. Zhang, "Using free energy principle for blind image quality assessment," *IEEE Transactions on Multimedia*, vol. 17, no. 1, pp. 50–63, Jan 2015.

[176] A. Mittal, R. Soundararajan, and A. Bovik, "Making a 'completely blind' image quality analyzer," *IEEE Signal Processing Letters*, vol. 20, no. 3, pp. 209–212, Mar 2013.

[177] A. Moorthy and A. Bovik, "Blind image quality assessment: From natural scene statistics to perceptual quality," *IEEE Transactions on Image Processing*, vol. 20, no. 12, pp. 3350–3364, Dec 2011.

[178] H. ju Park and D.-H. Har, "Subjective image quality assessment based on objective image quality measurement factors," *IEEE Transactions on Consumer Electronics*, vol. 57, no. 3, pp. 1176–1184, Aug 2011.

[179] Z.-X. Xie and Z.-F. Wang, "Color image quality assessment based on image quality parameters perceived by human vision system," in *2010 International Conference on Multimedia Technology (ICMT)*, Oct 2010, pp. 1–4.

[180] Q. Li and Z. Wang, "Reduced-reference image quality assessment using divisive normalization-based image representation," *IEEE Journal of Selected Topics in Signal Processing*, vol. 3, no. 2, pp. 202–211, Apr 2009.

[181] H.-S. Han, D.-O. Kim, and R.-H. Park, "Structural information-based image quality assessment using lu factorization," *IEEE Transactions on Consumer Electronics*, vol. 55, no. 1, pp. 165–171, Feb 2009.

[182] Y.-S. Lee, J.-M. Kwak, S.-E. Cho, J.-W. Kim, and H.-J. Kang, "A study on the medical image transmission service based on ieee 802.15. 4a," in *Advances in Hybrid Information Technology*. Springer, 2007, pp. 159–167.

[183] A. Redondi, D. Buranapanichkit, M. Cesana, M. Tagliasacchi, and Y. Andreopoulos, "Energy consumption of visual sensor networks: Impact of spatio-temporal coverage," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 24, no. 12, pp. 2117–2131, Dec 2014.

[184] D. Bhattacharyya, T.-h. Kim, and S. Pal, "A comparative study of wireless sensor networks and their routing protocols," *Sensors*, vol. 10, no. 12, pp. 10 506–10 523, 2010.

[185] P. Viola and M. Jones, "Robust real-time object detection," *International Journal of Computer Vision*, vol. 4, 2001.

[186] K. Lin, X. Wang, S. Cui, and Y. Tan, "Heterogeneous feature fusion-based optimal face image acquisition in visual sensor network," in *2015 IEEE International Instrumentation and Measurement Technology Conference (I2MTC) Proceedings*. IEEE, 2015, pp. 1078–1083.

[187] X. Li and X. Guo, "Vision-based method for forward vehicle detection and tracking," in *2013 International Conference on Mechanical and Automation Engineering (MAEE)*. IEEE, 2013, pp. 128–131.

[188] H. Hajimolahoseini, R. Amirfattahi, and H. Soltanian-Zadeh, "Robust vehicle tracking algorithm for nighttime videos captured by fixed cameras in highly reflective environments," *IET Computer Vision*, vol. 8, no. 6, pp. 535–544, 2014.

[189] H. Maeng, S. Liao, D. Kang, S.-W. Lee, and A. K. Jain, "Nighttime face recognition at long distance: cross-distance and cross-spectral matching," in *Computer Vision–ACCV 2012*. Springer, 2012, pp. 708–721.

[190] D. Kang, H. Han, A. K. Jain, and S.-W. Lee, "Nighttime face recognition at large standoff: Cross-distance and cross-spectral matching," *Pattern Recognition*, vol. 47, no. 12, pp. 3750–3766, 2014.

[191] L. W. Chew, L.-M. Ang, and K. P. Seng, "Survey of image compression algorithms in wireless sensor networks," in *2008 International Symposium on Information Technology*, vol. 4. IEEE, 2008, pp. 1–9.

[192] C. Yang, J. Lu, W. Yang, and H. Shi, "Sensor scheduling for lifetime maximization in centralized state estimation," *Neurocomputing*, 2017.

[193] A. Redondi *et al.*, "Cooperative image analysis in visual sensor networks," *Ad Hoc Networks*, vol. 28, pp. 38–51, 2015.

[194] A. Redondi, D. Buranapanichkit, M. Cesana, M. Tagliasacchi, and Y. Andreopoulos, *Energy Consumption of Visual Sensor Networks: Impact of Spatio-Temporal Coverage Based on Single-Hop Topologies*. Cham: Springer International Publishing, 2014, pp. 150–165.

[195] K. Li, "Optimal number of annuli for maximizing the lifetime of sensor networks," *Journal of Parallel and Distributed Computing*, vol. 74, no. 1, pp. 1719–1729, 2014.

[196] N. Hema and K. Kant, "Optimization of sensor deployment in wsn for precision irrigation using spatial arrangement of permanent crop," in *2013 Sixth International Conference on Contemporary Computing (IC3)*, Aug 2013, pp. 455–460.

[197] K. Cohen and A. Leshem, "A time-varying opportunistic approach to lifetime maximization of wireless sensor networks," *IEEE Transactions on Signal Processing*, vol. 58, no. 10, pp. 5307–5319, 2010.

[198] Y. Chen, Q. Zhao, V. Krishnamurthy, and D. Djonin, "Transmission scheduling for optimizing sensor network lifetime: A stochastic shortest path approach," *IEEE Transactions on Signal Processing*, vol. 55, no. 5, pp. 2294–2309, 2007.

[199] I. Rekleitis and G. Dudek, "Automated calibration of a camera sensor network," in *2005 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Aug 2005, pp. 3384–3389.

[200] F. Dunn and I. Parberry, *3D math primer for graphics and game development*. CRC Press, 2015.

[201] E. Lengyel, *Mathematics for 3D game programming and computer graphics*. Cengage Learning, 2012.

[202] V. Lecuire, C. Duran-Faundez, and N. Krommenacker, "Energy-efficient transmission of wavelet-based images in wireless sensor networks," *Journal on Image and Video Processing*, vol. 2007, no. 1, 2007.

[203] S. Qaisar and H. Radha, "A reliability framework for visual sensor networks," in *2009 Picture Coding Symposium*, May 2009, pp. 1–4.

[204] A. A. Ahmed, "A real-time routing protocol with adaptive traffic shaping for multimedia streaming over next-generation of wireless multimedia sensor networks," *Pervasive and Mobile Computing*, 2017.

[205] G. Han, J. Jiang, M. Guizani, and J. J. P. C. Rodrigues, "Green routing protocols for wireless multimedia sensor networks," *IEEE Wireless Communications*, vol. 23, no. 6, pp. 140–146, Dec 2016.