

Characterisation of Field-of-View for Energy Efficient Application-Aware Visual Sensor Networks

Anas Amjad, *Student Member, IEEE*, Mohammad Patwary, *Senior Member, IEEE*, Alison Griffiths
and Abdel-Hamid Soliman, *Member, IEEE*

Abstract—Energy consumption is one of the primary concerns in a resource constrained Visual Sensor Network. The existing Visual Sensor Network design solutions under particular resource constrained scenarios are application specific; whereas the degree of sensitivity of any of these resource constrains (e.g. energy etc) varies from one application to another. This limits the implementation of the existing energy efficient solutions within a Visual Sensor Network node which may be considered to be a part of a heterogeneous network. The heterogeneity of image capture and processing within a Visual Sensor Network can be adaptively reflected with a dynamic Field-of-View realisation. This is expected to allow the implementation of a generalised energy efficient solution to adapt with the heterogeneity of the network. In this paper, an energy efficient Field-of-View characterisation framework is proposed which can support a diverse range of applications. The context of adaptivity in the proposed Field-of-View characterisation framework is considered to be: a) sensing range selection; b) maximising spatial coverage; c) adaptive task classification and d) minimising the number of required nodes. Soft decision criteria is exploited and it is observed that for a given detection reliability, the proposed framework provides energy efficient solutions which can be implemented within heterogeneous networks. It is also found that the proposed design solution for heterogeneous networks leads to 49.8% energy savings compared to the trivial design solution.

Index Terms—Energy optimisation, field-of-view characterisation, resource optimisation, sensing range estimation, task classification, visual sensor networks.

I. INTRODUCTION

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 that limits their suitability to many applications [1]. In order to enhance WSN's suitability for a wider range of applications, its 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) and it collaborates with other nodes in the network [2]. VSNs are used in surveillance [3, 4], environmental monitoring [5, 6], assisted living and tele-healthcare [7, 8] applications. Visual sensors within a

VSN employ directional sensing to provide pixel based measurements as a 2D dataset and they require a large bandwidth to transmit image data. The 3D viewing volume of a visual sensor is known as its Field-of-View (FoV) [2]. During the VSN design phase, some image processing algorithms require precise knowledge of the FoV. The fundamental differences between a traditional WSN and a VSN make the deployment of the latter 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.

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. Therefore, the characterisation of FoV and task classification to provide an energy efficient design solution is an important and challenging problem in VSNs.

This paper is focused on the FoV characterisation and task classification to obtain an optimised VSN configuration for resource constrained scenarios. The configuration of a VSN considered in this paper 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. The contributions of this paper are summarised as follows:

- 1) 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 sensor node for any given application.
- 2) 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.
- 3) In the context of heterogeneous VSNs, a comparison of hard decision and proposed soft decision based tech-

Manuscript received April 24, 2015; revised January 20, 2016; accepted January 21, 2016.

The authors are with the Faculty of Computing, Engineering and Sciences, Staffordshire University, Stoke-on-Trent, UK (e-mail: a.amjad@staffs.ac.uk; m.n.patwary@staffs.ac.uk; a.l.griffiths@staffs.ac.uk; a.soliman@staffs.ac.uk).

niques is presented. The proposed framework, when employed with the proposed task classification and soft decision based sensing range selection schemes resulted 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

The rest of the paper is organised as follows; Section II explores the existing solutions for VSN design and their limitations, Section III presents the visual sensor 3D projection model. Section IV introduces the proposed FoV characterisation framework. Section V describes the experimental setup and presents the results. Section VI provides an analysis of the proposed framework's energy efficiency. Section VII presents an analysis of system failure and finally, Section VIII provides the concluding remarks.

II. RELATED WORK

In the last decade, researchers have been actively engaged in VSN coverage, design and optimisation problems. In [9], an unsupervised scheme is proposed to identify the overlapping FoV for the estimation of network topology in large VSNs. A mathematical model is proposed in [10] to solve VSN coverage problem by deploying each node sequentially and removing the overlapping nodes. Closed form solutions for VSN coverage estimation problem are proposed in [11] and [12] for homogeneous and heterogeneous VSNs respectively. The latter approach also considers the visual occlusions and boundary effect. A visual feature extractor, BRISKOLA (Binary Robust Invariant Scalable Keypoints Optimized for Low-power ARM architectures), is proposed in [13] by optimising BRISK [14] for ARM architectures. The proposed approach can also be used in resource constrained VSNs. Ye et al. [15] 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 [16], 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 [17] proposed camera scheduling and energy allocation schemes to maximise a VSN's lifetime. Kim et al. [18] proposed an energy efficient scheme to maximise the data quality and lifetime of solar-powered VSNs. A VSN lifetime maximisation strategy is proposed in [19] that optimises the source rates, encoding powers and routing schemes to prolong the network's lifetime. Authors in [20] proposed an energy efficient relaying scheme for data packets transmission within the VSN to increase its lifetime.

Although the existing work in literature addresses several key issues relating to VSNs such as FoV identification, coverage estimation, feature extraction, camera scheduling and transmission; 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. In this paper, an energy efficient FoV characterisation framework is proposed that exploits the heterogeneity of the network to obtain a design solution suitable for a diverse range of applications.

III. VISUAL SENSOR 3D PROJECTION MODEL

The 3D projection model of a visual sensor VS within a spherical sector is shown in Fig. 1. VS employs directional sensing and transforms a projection of the 3D scene from \mathbb{R}^3 to a 2D image plane in \mathbb{R}^2 . 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 in front of the sensor's centre and provides the same image with respect to the scene within the FoV. The visual sensor covers a certain part of the spherical area of interest. The region within the sensor'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.

In Fig. 1, sensor 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 homogeneous VSN, where N sensor nodes are present, each sensor node VS_u ($u = \{1, 2, 3, \dots, N\}$) is identified by its location which is described by the cartesian coordinates (X_u, Y_u, Z_u), azimuth angle ϕ_a and elevation angle θ_e . These parameters define sensor distribution within the network and are tuned to fit the respective areas of interest within each sensor's FoV. 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 sensor's coverage area due to the change in ABCD-plane dimensions. Therefore, R is a key parameter for FoV characterisation. The width and height of the physical image plane are represented by w_1 and h_1 respectively. The physical distance $f \in \mathbb{R}^+$ between the sensor's optical centre and the image plane is known as the *focal length*. Each sensor maps $P \times Q$ pixels onto the image plane, where $P \times Q$ is known as the *resolution* of the sensor. High resolution sensors are capable of observing a large area within their FoVs and result in reduction of the number of sensors required for full coverage. However, such sensors may increase the overall network design cost. Therefore, the selection of sensors for VSN design requires careful consideration of all the aforementioned parameters.

IV. PROPOSED FOV CHARACTERISATION FRAMEWORK

The proposed FoV characterisation framework to design and calibrate energy efficient VSNs is presented in Fig. 2.

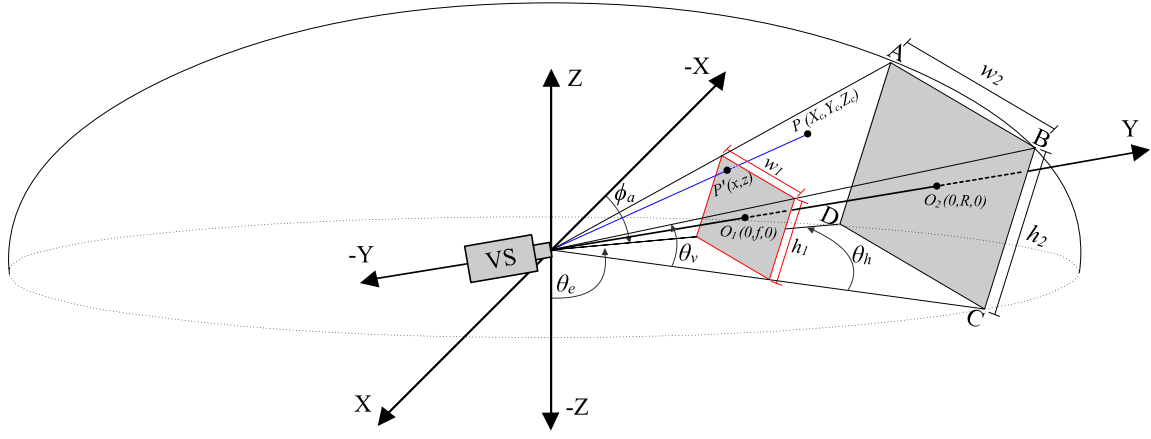


Fig. 1: Visual sensor's 3D projection model.

Based on the type of VSN, there are two approaches for its design and calibration: (i) Approach I for homogeneous sensor networks shown in Fig. 2a (ii) Approach II for heterogeneous sensor networks shown in Fig. 2b. 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.

A. Image Capture

Each sensor VS_u captures an image I_u of dimension $P \times Q$ which is a function of the following parameters: the range R , the horizontal FoV θ_h and the vertical FoV θ_v .

$$I_u = \mathcal{F}(R, \theta_h, \theta_v) \quad (1)$$

B. Projection Modelling

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

$$V = \frac{4R^3 \sin \theta_h \sin \theta_v}{3(1 + \cos \theta_h)(1 + \cos \theta_v)} \quad (2)$$

In order to characterise sensor'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 [21], FoV (θ_h, θ_v) is calculated with prior knowledge of the sensor specifications.

$$\theta_h = 2 \arctan \left(\frac{w_1}{2f} \right) \quad (3)$$

$$\theta_v = 2 \arctan \left(\frac{h_1}{2f} \right) \quad (4)$$

(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 tasks 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 sensor node using the following equations,

$$\theta_h = 2 \arctan \left(\frac{w_2}{2R} \right) \quad (5)$$

$$\theta_v = 2 \arctan \left(\frac{h_2}{2R} \right) \quad (6)$$

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

C. ABCD-plane Modelling

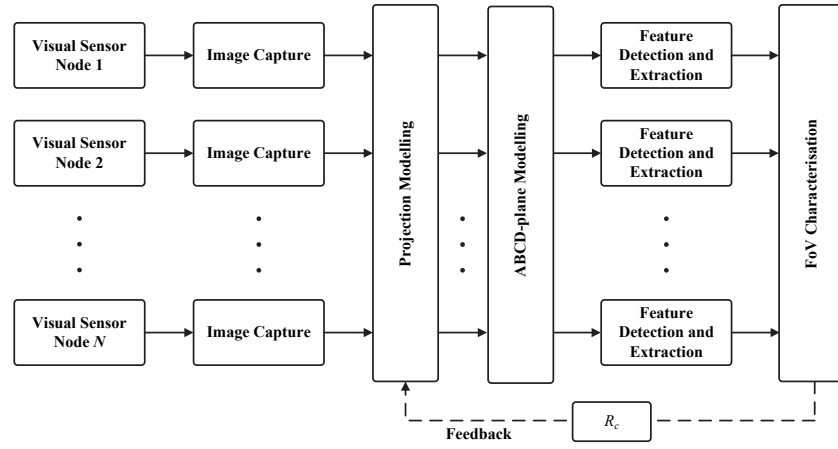
Once θ_h and θ_v are known, the dimension of the ABCD-plane is calculated for a range of values of R i.e. R_{min} to R_{max} using the following equations,

$$w_2 = 2R \tan \left(\frac{\theta_h}{2} \right) \quad (7)$$

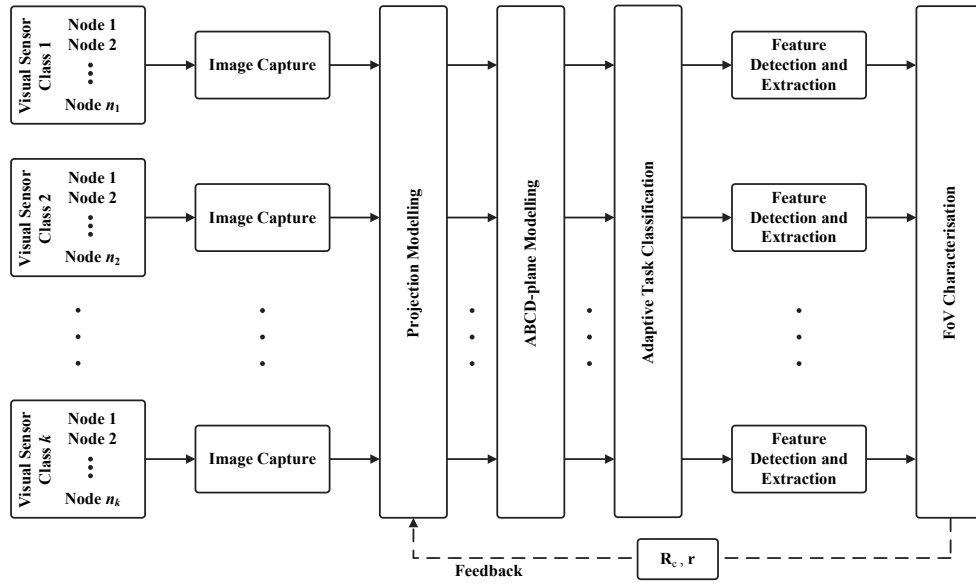
$$h_2 = 2R \tan \left(\frac{\theta_v}{2} \right) \quad (8)$$

D. Adaptive Task Classification

Heterogeneous networks comprise of sensor nodes with different capabilities and their performance is better as compared to their homogeneous counterparts due to the classification of



(a) FoV Characterisation (Approach I) for homogeneous VSNs



(b) FoV Characterisation (Approach II) for heterogeneous VSNs

Fig. 2: Proposed FoV Characterisation Framework for Energy Efficient VSNs.

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 networks. Consider a heterogeneous sensor network with k sensor classes; each sensor is denoted by $VS_{l,j}$ such that $l = \{1, 2, 3, \dots, k\}$ represents the sensor class and $j = \{1, 2, 3, \dots, n\}$ represents n_l sensors belonging to a sensing class l . Let n denotes the maximum number of sensors belonging to a particular sensing class given by $n = \max \{n_l \mid l = 1, 2, 3, \dots, k\}$. The sensors within the VSN can be represented by,

$$\mathbf{VS} = \begin{bmatrix} VS_{1,1} & VS_{1,2} & \dots & VS_{1,n} \\ VS_{2,1} & VS_{2,2} & \dots & VS_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ VS_{k,1} & VS_{k,2} & \dots & VS_{k,n} \end{bmatrix}$$

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. $VS_{l,j}$ is assigned a value based on the following condition,

$$VS_{l,j} = \begin{cases} -1, & \text{if } j > n_l \\ 1, & \text{if the sensor is active} \\ 0, & \text{if the sensor is inactive} \end{cases} \quad (9)$$

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

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

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

$$T_{l,j} = \begin{cases} 1, & \text{if } i\text{th task is assigned to sensor } VS_{l,j} \\ 0, & \text{otherwise} \end{cases} \quad (10)$$

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

$$\mathbf{T}'_i = \left\lfloor \frac{1}{2} [\mathbf{T}_i \cdot \mathbf{VS} + \mathbf{J}] \right\rfloor \quad (11)$$

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

Feedback \mathbf{R}_c and \mathbf{r} are substituted in (2) 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 1 presents the proposed adaptive task classification scheme that calculates \mathbf{T} and then \mathbf{T}' in an optimised way.

E. Feature Detection and Extraction

Global colour histogram is used for object detection and feature extraction. It represents the distribution of colours within each captured image \mathbf{I}_i of size $P \times Q$ and is given by [22],

$$h_c(b) = \sum_{x=1}^P \sum_{z=1}^Q \begin{cases} 1, & \text{if } I_i(x, z) \text{ is in bin } b \\ 0, & \text{otherwise} \end{cases} \quad (12)$$

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,

$$P(E) = \begin{cases} 1, & \text{if } \gamma_{Cb}^l \leq Cb \leq \gamma_{Cb}^u \cap \gamma_{Cr}^l \leq Cr \leq \gamma_{Cr}^u \\ 0, & \text{otherwise} \end{cases} \quad (13)$$

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 [23],

$$\mathbf{S}_g = \mathbf{I}_i \cdot \mathbf{S}_m \quad (14)$$

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

F. FoV Characterisation

The relationship between sensor's resolution and ABCD-plane dimensions for a distance R is given by,

$$d_h = \frac{P}{w_2} \quad ; \quad d_v = \frac{Q}{h_2} \quad (15)$$

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

Algorithm 1 Adaptive task classification scheme for heterogeneous networks

Require:

The number of: sensor classes k , sensors 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  $\forall j \in \{1, 2, 3, \dots, n\}$  and  $\forall i \in \{1, 2, 3, \dots, t\}$ 
   $0 < \sum_{l=1}^k T'(l, j, i) \leq \lceil \sqrt{k} \rceil$ 
1:  $n \leftarrow \max \{n_1, n_2, n_3, \dots, n_k\}$ 
2:  $\mathbf{VS} \leftarrow \emptyset$ 
3:  $\mathbf{s} \leftarrow [-1 \ 1 \ 0 \ 0 \ 0]$ 
4: for  $l \leftarrow 1$  to  $k$  do
5:   for  $j \leftarrow 1$  to  $n$  do
6:     if  $j > n_l$  then
7:        $\mathbf{VS}(l, j) \leftarrow s_1$ 
8:     else if sensor is active then
9:        $\mathbf{VS}(l, j) \leftarrow s_2$ 
10:    else if sensor is inactive then
11:       $\mathbf{VS}(l, j) \leftarrow s_3$ 
12:    end if
13:  end for
14: end for
15:  $\mathbf{T} \leftarrow \emptyset$ 
16:  $\mathbf{T}' \leftarrow \emptyset$ 
17: for  $i \leftarrow 1$  to  $t$  do
18:   for  $l \leftarrow 1$  to  $k$  do
19:      $s_5 \leftarrow k - l + 1$ 
20:     if  $V_c(s_5, i) \geq v(s_5)$  &  $s_4 < \lceil \sqrt{k} \rceil$  then
21:        $T(s_5, 1:n, i) \leftarrow s_2$ 
22:        $s_4 \leftarrow s_4 + 1$ 
23:     else
24:        $T(s_5, 1:n, i) \leftarrow s_3$ 
25:     end if
26:   end for
27:    $\mathbf{T}'_i = \left\lfloor \frac{1}{2} [\mathbf{T}_i \cdot \mathbf{VS} + \mathbf{J}] \right\rfloor$ 
28:    $s_4 \leftarrow 0$ 
29: end for
30: return  $\mathbf{T}'$ 

```

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 utilises either one or a combination of the following parameters: Object Pixel Occupancy (O_{po}), Estimation Error ($|\varepsilon_d|$) and Peak Signal-to-Noise Ratio (PSNR_{dB}) as 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 derived as,

$$O_{po} = A \times \frac{P}{2R \tan\left(\frac{\theta_h}{2}\right)} \times \frac{Q}{2R \tan\left(\frac{\theta_v}{2}\right)} \quad (16)$$

where A is the area of the object in mm^2 . Let ξ_o defines the required minimum pixel occupancy for a particular application, the chosen range R_1 of a visual sensor must guarantee the criteria $O_{po} \geq \xi_o$ and it can be calculated using the following equation,

$$R_1 = \sqrt{\frac{P \times Q \times A}{4 \times O_{po} \times \tan\left(\frac{\theta_h}{2}\right) \times \tan\left(\frac{\theta_v}{2}\right)}} \quad (17)$$

Table I 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 classifiers. The detection accuracy calculated on PETS 2005 data set in [24] with $\xi_o = 25$ for LOTS, SGM and MSM is 91.2%, 86.8% and 85.0% respectively.

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

Detection Method	ξ_o
Viola-Jones face detector [25]	315
Lehigh Omnidirectional Tracking System (LOTS) [26]	25
Single Gaussian Model (SGM) [27]	25
Multiple Gaussian Model (MGM) [28]	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 sensing range 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} denotes the number of pixels representing the detected object; the framework estimates pixels representing the diameter p_d by,

$$p_d = 2\sqrt{\frac{O_{po}}{\pi}} \quad (18)$$

The diameter d_e of the object is estimated by,

$$d_e = \frac{4R \tan\left(\frac{\theta_h}{2}\right)}{P} \sqrt{\frac{O_{po}}{\pi}} \quad (\text{metres}) \quad (19)$$

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

$$|\varepsilon_d| = \frac{|d_a - d_e|}{d_a} \times 100 \quad (\%) \quad (20)$$

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

Substituting (19) in (20), the range R_2 can be calculated using the following equation,

$$R_2 = \left[\sqrt{\frac{\pi}{O_{po}}} \right] \left[\frac{Pd_a}{4 \tan\left(\frac{\theta_h}{2}\right)} \right] \left[1 - \frac{|\varepsilon_d|}{100} \right] \quad (21)$$

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

(iii) *PSNR*: Suppose image 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 I'_1 of dimension $P' \times Q'$. As I'_1 contains the object captured at distance R_p , the dimension $P_s \times Q_s$ of image I_s containing the extracted object at distance R_s (such that $R_s > R_p$) is estimated by,

$$P_s = P' \left(\frac{w_2^{R_p}}{w_2^{R_s}} \right) \quad ; \quad Q_s = Q' \left(\frac{h_2^{R_p}}{h_2^{R_s}} \right) \quad (22)$$

where $w_2^{R_p}$, $w_2^{R_s}$, $h_2^{R_p}$ and $h_2^{R_s}$ are calculated using (7) and (8). 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, I'_1 and I_s are compared to find the Peak Signal-to-Noise Ratio (PSNR) value. As PSNR requires both images to have the same size, I_s is resized to $P' \times Q'$. First, Mean Squared Error (MSE) is calculated and then the PSNR.

$$MSE = \frac{1}{3P'Q'} \sum_{x=1}^{P'} \sum_{y=1}^{Q'} \left[I'_1(x, y) - I_s(x, y) \right]^2 \quad (23)$$

$$PSNR_{dB} = 10 \times \log_{10} \left(\frac{\text{MAX}_{I'_1}^2}{MSE} \right) \quad (24)$$

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

Let ξ_p defines the required minimum $PSNR_{dB}$ for a particular application, the chosen range R_3 of a visual sensor must guarantee the criteria $PSNR_{dB} \geq \xi_p$. 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 V.

Apart from $PSNR_{dB}$, there are many other image quality assessment methods such as [29, 30] that can be used with the proposed FoV characterisation framework based on their respective confidence bounds for sensing range estimation.

The selection of one or more characterisation methods depends on the application and the design criteria. 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.

Let R_c defines the chosen value of sensor's range, the FoV Characterisation Criteria (FoVCC) is proposed as,

$$R_c = \begin{cases} R_1 & \text{if } O_{po} \geq \xi_o \\ R_2 & \text{if } |\varepsilon_d| \leq \xi_d \\ R_3 & \text{if } PSNR_{dB} \geq \xi_p \end{cases} \quad (25)$$

The sensing range for applications where the design engineer utilises more than one characterisation method is selected by $R_c = \min\{R_1, R_2, R_3\}$. The designed VSN's FoV is said to be optimised based on the following criteria,

$$\begin{cases} R_c = R_1 \cup R_c = R_2 \cup R_c = R_3 & \text{Optimised} \\ R_c \neq R_1 \cap R_c \neq R_2 \cap R_c \neq R_3 & \text{Unoptimised} \end{cases} \quad (26)$$

G. Adaptive Range Selection

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

$$R_c = \min \{R_{c(1)}, R_{c(2)}, R_{c(3)}, \dots, R_{c(t)}\} \quad (27)$$

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.

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

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

$$r(l) = \min \{R_{c(l,1)}, R_{c(l,2)}, R_{c(l,3)}, \dots, R_{c(l,t)}\} \quad (28)$$

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

$$\mathbf{r} = [r(1), r(2), \dots, r(k)]^T \quad (29)$$

where $[\cdot]^T$ refers to transpose. The proposed hard decision based scheme is suitable for homogeneous networks as they have identical sensor 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.

(ii) *Soft decision based sensing range selection*: A soft decision based sensing range selection scheme is proposed in Algorithm 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 $(k - \lceil \sqrt{k} \rceil)$ sensor classes.

Algorithm 2 Proposed soft decision based sensing range selection scheme for heterogeneous network design

Require:

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

Ensure:

For $\lceil \sqrt{k} \rceil$ values of $l \in \{1, 2, 3, \dots, k\}$ and $\forall i \in \{1, 2, 3, \dots, t\}$
 $r(l) \leq R_{c(l,i)}$
1: $\mathbf{s}_1 = \emptyset$
2: **for** $l \leftarrow 1$ **to** k **do**
3: $\mathbf{s}_1 \leftarrow \sum_{i=1}^t R_{c(l,i)}$
4: **end for**
5: $\mathbf{s}_2 \leftarrow$ Sort \mathbf{s}_1 in ascending order
6: $\mathbf{s}_3 \leftarrow$ First $\lceil \sqrt{k} \rceil$ values' indices from \mathbf{s}_2
7: **for** $l \leftarrow 1$ **to** k **do**
8: **if** $l \in \mathbf{s}_3$ **then**
9: $r(l) \leftarrow \min \{R_{c(l,1)}, R_{c(l,2)}, R_{c(l,3)}, \dots, R_{c(l,t)}\}$
10: **else if** $l \notin \mathbf{s}_3$ **then**
11: $r(l) \leftarrow \left(\frac{1}{t}\right) \times \sum_{i=1}^t R_{c(l,i)}$
12: **end if**
13: **end for**
14: **return** \mathbf{r}

V. EXPERIMENTAL SETUP AND RESULTS

A. Image Capture

Specification of the visual sensor used for experiments is presented in Table II.

TABLE II: Visual sensor specification

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

B. 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 (3) and (4), the calculated values of horizontal and vertical FoVs are: $\theta_h = 48.39^\circ$, $\theta_v = 37.25^\circ$ respectively.

C. Projection Modelling utilising Approach II

As mentioned in Section IV, in the case of a heterogeneous network with some unknown sensor specifications, (3) and (4) 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 III.

TABLE III: Experiment 1 procedure

Procedure: Experiment 1	
1:	The visual sensor is placed near a wall at a certain height h_t and a known reference distance R_{ref} without any tilt or pan.
2:	A certain portion of the wall is captured within the sensor's FoV.
3:	The width w_2 and height h_2 of the wall's portion within the FoV are measured for R_{ref} .
4:	Using (5) and (6), 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 IV; where each case is distinguished by its reference distance R_{ref} .

TABLE IV: 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_c	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.

D. 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 (7) and (8) for a range of values of R . The simulation results are presented in Fig. 3. 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. In this case, $w_2 > h_2$ for any value of R due to the fact that $\theta_h > \theta_v$.

E. Feature Detection and Extraction

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

$$P(E) = \begin{cases} 1, & \text{if } 43 \leq \text{Cb} \leq 90 \cap 138 \leq \text{Cr} \leq 159 \\ 0, & \text{otherwise} \end{cases} \quad (30)$$

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

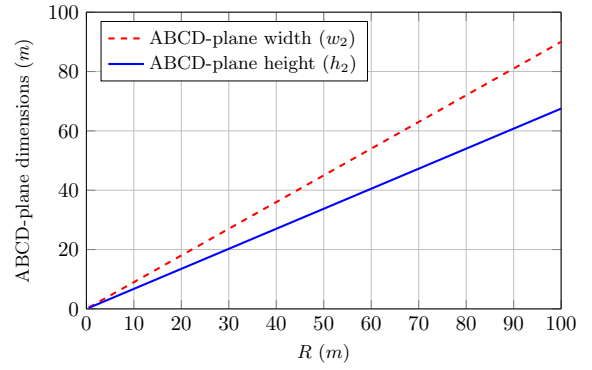


Fig. 3: ABCD-plane modelling.

F. 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 V.

TABLE V: Experiment 2 procedure

Procedure: Experiment 2	
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.

In order to estimate the sensing range for optimised FoV characterisation, a dataset is created by capturing objects 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 error, PSNR. The experimental and simulation results for FoV characterisation are presented and analysed in the following sections.

(i) *Object Pixel Occupancy*: Fig. 4 shows a comparison of the theoretical and experimental object pixel occupancy for a range of values of R with $A = 34\text{cm}^2$ for a homogeneous test object. The theoretical results are obtained from (16) 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 increasing the sensing range decreases the object pixel occupancy. 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.

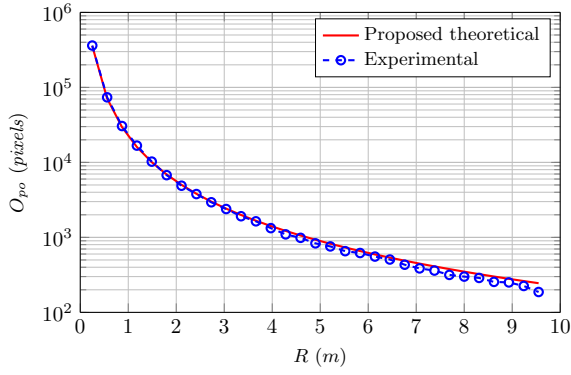


Fig. 4: A comparison of theoretical and experimental values of object pixel occupancy.

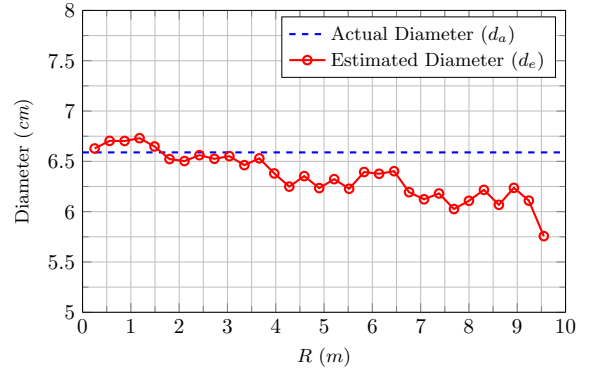


Fig. 5: A comparison of actual and estimated diameter for different sensing range values.

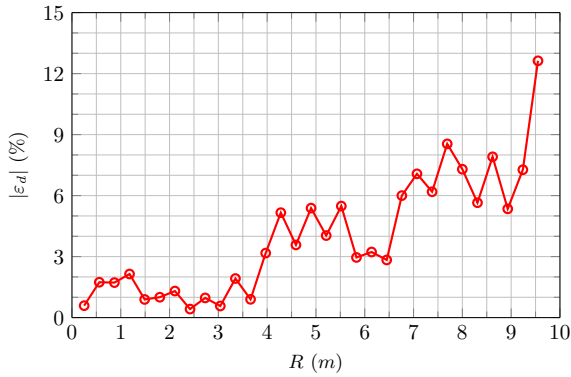


Fig. 6: Estimation error for different sensing range values.

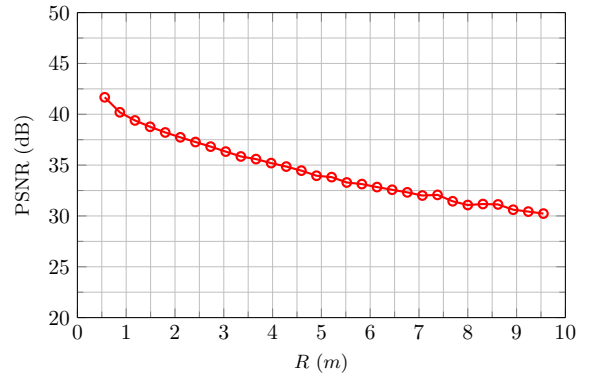


Fig. 7: PSNR estimation for different sensing range values.

As an example, a face detection application [25] requires object pixel occupancy to be atleast 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*: Using (19), Fig. 5 presents a comparison of actual and estimated diameter for images captured for a range of values of R . The absolute percentage estimation error $|\varepsilon_d|$ is shown in Fig. 6. It has been noticed from Fig. 6 that as the range increases, the estimation error increases. 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. 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. Fig. 7 shows the estimated $PSNR_{dB}$ for a range of values of R and it has been noticed that as the range increases, the $PSNR_{dB}$ decreases. 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 [31] requires $PSNR_{dB}$ to be atleast 30dB i.e. $PSNR_{dB} \geq 30dB$, the acceptable sensing range in that case will be $R_3 \leq 9.55m$.

G. Adaptive Range Selection

(i) *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 [25] (Task I) and occluded target surveillance and tracking [26] (Task II). Suppose a medium resolution sensor with the following parameters: $P \times Q = 640 \times 480$, $\theta_h = 48.39^\circ$ and $\theta_v = 37.25^\circ$ is selected for the VSN design. Let the object pixel occupancy based characterisation method is 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 (17), 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 sensor nodes (N_a) required within the VSN to perform the desired tasks.

(ii) *Heterogeneous Networks*: Now consider a heterogeneous network that has to perform the same $t = 2$ tasks described earlier for homogeneous network. 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

TABLE VI: 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	HD (Type 1) Case 1		HD (Type 2) Case 2		SD (Type 1) Case 3		SD (Type 2) 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

$P \times Q = 2304 \times 1728$. The horizontal FoV (θ_h) and the vertical FoV (θ_v) are assumed to be same for $k = 3$ sensor classes and are given by $\theta_h = 48.39^\circ$ and $\theta_v = 37.25^\circ$. 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 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, the chosen range is calculated (in *metres*) to be $\mathbf{r} = [r(1), r(2), r(3)]^T = [4.05, 8.09, 29.12]^T$. The range computed using the soft decision based scheme is $\mathbf{r} = [4.05, 8.09, 175.14]^T$.

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 $(k - \lceil \sqrt{k} \rceil)_{k=3}$ 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.

H. Sensor's 3D Coverage Estimation

The sensor's 3D coverage volume can be calculated from (2) by substituting θ_h and θ_v for a suitable sensing range R .

In the case of homogeneous sensor networks, feedback \mathbf{R}_c is utilised for projection modelling. Considering the homogeneous network design solution presented in the previous section and using (2), 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 earlier and using (2), 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 $\mathbf{v} = [13.41, 106.90, 4.99 \times 10^3]^T$. 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 $\mathbf{v} = [13.41, 106.90, 1.08 \times 10^6]^T$.

I. Adaptive Task Classification

In the proposed adaptive task classification scheme for heterogeneous networks, upto $\lceil \sqrt{k} \rceil$ sensor classes are assigned a certain task. Considering the heterogeneous design solution presented earlier, $\lceil \sqrt{k} \rceil_{k=3}$ 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 $\lceil \sqrt{k} \rceil$ upperbound (case 1), hard decision based approach with $\lceil \sqrt{k} \rceil$ upperbound (case 2), soft decision based approach without $\lceil \sqrt{k} \rceil$ upperbound (case 3) and soft decision based approach with $\lceil \sqrt{k} \rceil$ upperbound (case 4).

Table VI 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.

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 $(k - \lceil \sqrt{k} \rceil)$ 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 $\lceil \sqrt{k} \rceil$ 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.

VI. ENERGY EFFICIENCY OF THE PROPOSED FRAMEWORK

Consider a visual sensor network that requires N_a active nodes to cover an area of size $100 \times 100 m^2$. 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 r of k sensor classes for heterogeneous networks.

To validate the proposed framework which provides optimised energy consumption within certain desired confidence bounds, an energy-measurement testbed employed in [32] 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 VII.

TABLE VII: Energy-measurement testbed's parameters

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

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,

$$\tilde{E}_q = N_a \times E_q \quad ; \quad q \in \{Acq, Tx, Rx\} \quad (31)$$

The total energy consumption within the VSN will be,

$$E_c = N_a \times (E_{Acq} + E_{Tx} + E_{Rx}) \quad (32)$$

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

A. Homogeneous Networks

Consider a homogeneous network for a surveillance application that utilises a sensor with the following parameters: $P \times Q = 320 \times 240$, $\theta_h = 48.39^\circ$ and $\theta_v = 37.25^\circ$. A comparison of image acquisition, transmission and receiving costs for different sensing range values is shown in Fig. 8.

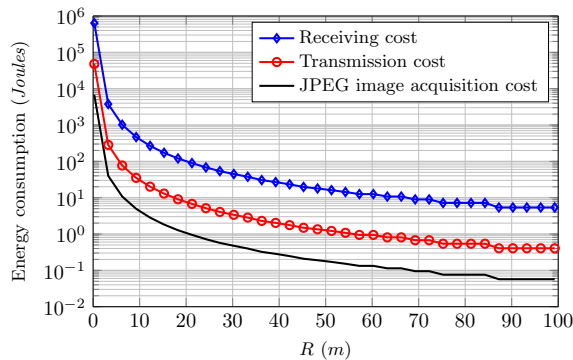


Fig. 8: A comparison of JPEG image acquisition, transmission and receiving cost for different sensing range values .

It is 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 VIII 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. 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 2.78 – 112.92 times 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 [26] 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.

B. 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 sensor 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} represents the number of tasks allocated to a sensing class; the task classification solutions obtained from four different cases given in Table VI are used to analyse the energy efficiency of the proposed framework and the results are presented in Table IX.

It is found from the results that the proposed soft decision based sensing range selection scheme with $\lceil \sqrt{k} \rceil$ 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 VIII: Application-aware sensing range estimation and energy consumption

Application	Characterisation method	Criteria	Range and energy	Gain	Network lifetime
Face detection [25]	object pixel occupancy based	$O_{po} \geq 315$	$R_1 = 4.05m$ $N_a = 1355$ $E_c = 2.64 \text{ kJ}$	—	LT
Feature extraction & size estimation	estimation error based	$ \varepsilon_d \leq 6\%$	$R_2 = 6.76m$ $N_a = 487$ $E_c = 948.78 \text{ J}$	64.06%	$LT \times 2.78$
Image transmission using IEEE 802.15.4a [31]	PSNR based	$PSNR_{dB} \geq 30dB$	$R_3 = 9.55m$ $N_a = 244$ $E_c = 475.36 \text{ J}$	81.99%	$LT \times 5.55$
Occluded targets surveillance & tracking [26]	object pixel occupancy based	$O_{po} \geq 25$	$R_1 = 44.60m$ $N_a = 12$ $E_c = 23.38 \text{ J}$	99.11%	$LT \times 112.92$

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

Sensor class	Sensor type	HD (Type 1) Case 1	HD (Type 2) Case 2	SD (Type 1) Case 3	SD (Type 2) Case 4
1	low resolution	$r(1) = 4.05m$ $\hat{t} = 2$ $N_a = 1355$ $E_c = 5.28 \text{ kJ}$	$r(1) = 4.05m$ $\hat{t} = 0$ $N_a = 0$ $E_c = 0 \text{ J}$	$r(1) = 4.05m$ $\hat{t} = 2$ $N_a = 1355$ $E_c = 5.28 \text{ kJ}$	$r(1) = 4.05m$ $\hat{t} = 1$ $N_a = 1355$ $E_c = 2.64 \text{ kJ}$
2	medium resolution	$r(2) = 8.09m$ $\hat{t} = 2$ $N_a = 340$ $E_c = 5.26 \text{ kJ}$	$r(2) = 8.09m$ $\hat{t} = 2$ $N_a = 340$ $E_c = 5.26 \text{ kJ}$	$r(2) = 8.09m$ $\hat{t} = 2$ $N_a = 340$ $E_c = 5.26 \text{ kJ}$	$r(2) = 8.09m$ $\hat{t} = 2$ $N_a = 340$ $E_c = 5.26 \text{ kJ}$
3	high resolution	$r(3) = 29.12m$ $\hat{t} = 2$ $N_a = 27$ $E_c = 5.40 \text{ kJ}$	$r(3) = 29.12m$ $\hat{t} = 2$ $N_a = 27$ $E_c = 5.40 \text{ kJ}$	$r(3) = 175.14m$ $\hat{t} = 1$ $N_a = 27$ $E_c = 100 \text{ J}$	$r(3) = 175.14m$ $\hat{t} = 1$ $N_a = 27$ $E_c = 100 \text{ J}$
Overall energy consumption		15.94 kJ	10.66 kJ	10.64 kJ	8 kJ
Network lifetime		LT	$LT \times 1.5$	$LT \times 1.5$	$LT \times 2$

VII. 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 denotes 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,

$$P(\lambda_t > \beta) = \begin{cases} 0 & \text{if } \lambda_t < \beta \\ \frac{\delta_2}{\delta_1} & \text{if } \lambda_t \geq \beta \end{cases} \quad (33)$$

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

Fig. 9 shows an analysis of system failure for several values of λ_t . It can be observed from the graph that the system failure probability is maximum when λ_t and β are at the opposite ends of the dynamic PSNR range. The system failure probability reduces when λ_t and β lies between the dynamic range and minimises to zero when $\lambda_t < \beta$.

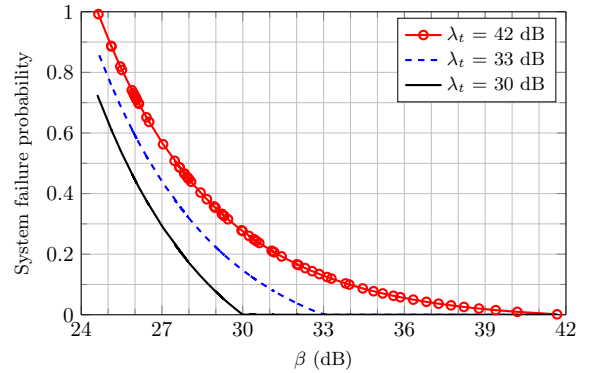


Fig. 9: Analysis of system failure.

VIII. CONCLUSION

In this paper, the issues of FoV characterisation and task classification for VSNs are addressed. A novel framework for the FoV characterisation of homogeneous and heterogeneous networks is proposed. For a given reliability, the proposed framework defines the criteria, referred to as the “Field-of-View Characterisation Criteria”, which estimates the optimal

sensing range of a visual sensor node. For any given application, the proposed solution for FoV characterisation enhances the spatial coverage, optimises the energy consumption and increases the lifetime in homogeneous networks. This paper also proposes adaptive task classification and soft decision based sensing range selection schemes for heterogeneous networks. 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 49.8% energy savings compared to the trivial design solution. Based on the required and achieved quality of the captured image by a visual sensor node, an analysis of system failure is 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. Furthermore, the proposed framework provides simplified direction to future research within the context of homogeneous and heterogeneous VSN design. For the future extension of this work, the authors intend to develop generalised adaptation models of feature detection and extraction schemes for realisation with the proposed framework.

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, April 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] 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.
- [6] N. Ahmad, K. Khurshed, 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.
- [7] 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.
- [8] 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.
- [9] K. Shafique, A. Hakeem, O. Javed, and N. Haering, "Self calibrating visual sensor networks," in *2008 IEEE Workshop on Applications of Computer Vision (WACV)*. IEEE, 2008, pp. 1–6.
- [10] 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.
- [11] M. Karakaya and H. Qi, "Coverage estimation for crowded targets in visual sensor networks," *ACM Transactions on Sensor Networks (TOSN)*, vol. 8, no. 3, p. 26, 2012.
- [12] M. Karakaya and H. Qi, "Coverage estimation in heterogeneous visual sensor networks," in *2012 IEEE 8th International Conference on Distributed Computing in Sensor Systems (DCOSS)*. IEEE, 2012, pp. 41–49.
- [13] 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.
- [14] S. Leutenegger, M. Chli, and R. Y. Siegwart, "Brisk: Binary obust invariant scalable keypoints," in *2011 IEEE International Conference on Computer Vision (ICCV)*. IEEE, 2011, pp. 2548–2555.
- [15] 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.
- [16] 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.
- [17] 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.
- [18] 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.
- [19] 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.
- [20] D. G. Costa, L. A. Guedes, F. Vasques, and P. Portugal, "Energy-efficient packet relaying in wireless image sensor networks exploiting the sensing relevancies of source nodes and dwt coding," *Journal of Sensor and Actuator Networks*, vol. 2, no. 3, pp. 424–448, 2013.
- [21] 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)*, Oct 2014, pp. 818–822.
- [22] L. Viet Tran, "Efficient image retrieval with statistical color descriptors," *PhD Thesis No. 810, Linkoping University*, 2003.
- [23] A. Amjad, A. Griffiths, and M. Patwary, "Multiple face detection algorithm using colour skin modelling," *IET Image Processing*, vol. 6, no. 8, pp. 1093–1101, Nov. 2012.
- [24] 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.

- [25] P. Viola and M. Jones, “Robust real-time object detection,” in *International Journal of Computer Vision*, 2001.
- [26] T. Boulton, R. Micallef, 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.
- [27] C. Wren, A. Azarbayejani, T. Darrell, and A. Pentland, “Pfnder: real-time tracking of the human body,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 780–785, Jul 1997.
- [28] 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.
- [29] 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.
- [30] 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, March 2013.
- [31] 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.
- [32] 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.



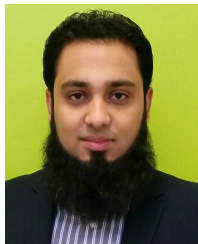
Mohammad Patwary (SM’11) holds a Readership in Wireless Communication Systems since 2011 and leading Sensing, Processing & Communication Research (SPCR) Group at Staffordshire University, UK. He received his Ph.D. degree in Telecommunication Engineering from The University of New South Wales, Sydney, Australia in 2005 and BSc. in Electrical and Electronic Engineering from Chittagong University of Engineering and Technology, Bangladesh in 1998. He was with the General Electric Company of Bangladesh from 1998 to 2000 and with Southern-Poro Communications, Sydney, Australia from 2001–2002 as R&D Engineer. He worked at the University of New South Wales, Sydney, Australia as Lecturer from 2005–2006; then at Staffordshire University, UK as Senior Lecturer from 2006–2010. Current research interests of SPCR Group include future generation of cellular network architecture; signal detection, estimation and optimisation for future generation of wireless receivers; joint source-channel characterisation for wireless sensor networks (Heterogeneous Networks & IoT) and spatial diversity schemes for wireless communications.



Alison Griffiths received the BEng. (Hons.), MEng. and PhD degrees from Staffordshire University, Stoke-on-Trent, UK, in 1998, 1999 and 2004, respectively. She is currently working as full time senior lecturer within the Faculty of Computing, Engineering and Sciences, Staffordshire University, Stoke-on-Trent, UK. She also worked with several companies on industrial collaborations. Her research interests include control system technologies, estimation and optimisation techniques, wireless sensor networks and battery life optimisation techniques.



Abdel-Hamid Soliman received the BSc. degree in electronic and telecommunications, MSc in Smart Data Acquisition Systems and the Ph.D. Image/Video Processing from Staffordshire University. He has a multi-disciplinary academic/research experience in digital signal processing, telecommunications, data acquisition systems, wireless sensor networks (WSN) and image/video processing. In addition to his research activities, he is involved in several enterprise projects and consultancy activities for national and international companies. Now, he is leading and involved in several European funded projects.



Anas Amjad (S’15) received the B.Eng. (Hons.) degree in electronic engineering from Staffordshire University, Stafford, UK in 2011. He is currently pursuing the Ph.D. degree in electronic engineering from Staffordshire University, Stoke-on-Trent, UK. He is a member of Sensing, Processing & Communication Research (SPCR) Group at Staffordshire University, UK. His current research interests include resource optimisation for wireless visual sensor networks, feature detection and extraction, visual content compression and quality assessment

for future generation of wireless networks.