IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received November 2, 2016, accepted November 26, 2016, date of publication December 9, 2016, date of current version January 23, 2017.

Digital Object Identifier 10.1109/ACCESS.2016.2635941

QoI-Aware Unified Framework for Node Classification and Self-Reconfiguration Within Heterogeneous Visual Sensor Networks

ANAS AMJAD, (Student Member, IEEE), ALISON GRIFFITHS, AND MOHAMMAD PATWARY, (Senior Member, IEEE)
Faculty of Computing, Engineering and Sciences, Staffordshire University, Stoke-on-Trent ST4 2DE, U.K.
Corresponding author: A. Amjad (anasamjad@ieee.org)

ABSTRACT Due to energy and throughput constraints of visual sensing nodes, in-node energy conservation is one of the prime concerns in visual sensor networks (VSNs) with wireless transceiving capability. To cope with these constraints, the energy efficiency of a VSN for a given level of reliability can be enhanced by reconfiguring its nodes dynamically to achieve optimal configurations. In this paper, a unified framework for node classification and dynamic self-reconfiguration in VSNs is proposed. The proposed framework incorporates quality-of-information (QoI) awareness using peak signal-to-noise ratio-based representative metric to support a diverse range of applications. First, for a given application, the proposed framework provides a feasible solution for the classification of visual sensing nodes based on their field-of-view by exploiting the heterogeneity of the targeted QoI within the sensing region. Second, with the dynamic realization of 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. The robustness of the proposed framework is evaluated under various scenarios by considering: 1) target QoI thresholds; 2) degree of heterogeneity; and 3) compression schemes. From the simulation 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.

INDEX TERMS 3D field-of-view modelling, dynamic reconfiguration, energy optimization, node classification, quality-of-information, reliability analysis, visual sensor networks.

I. INTRODUCTION

Visual Sensor Networks (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 Wireless Sensor Network (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 such as surveillance [1], [2], object detection and tracking [3], [4], health care monitoring [5], [6], and many others. 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. In order to facilitate such intelligent sensing within a visual sensing node, a dynamic coverage modelling approach can be employed to obtain the 3D Field-of-View (FoV) information.

Energy is a scarce resource in VSNs due to the resource constrained nature of its nodes and the possibility of deployment in inadequately resourced areas to support complex algorithms [7], [8]. Therefore, the main challenge in designing VSNs is to utilize resources optimally while maintaining a certain degree of reliability, as per the given application. Efficient utilization of network resources and optimization of processing algorithms lead to the conservation of energy resulting in increased lifetime. Furthermore, the performance and lifetime of a network is characterized 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 of static visual sensing nodes. In addition to these, the orientation and FoV parameters are also considered for the configuration of networks comprising of Pan-Tilt-Zoom (PTZ) capable visual sensing nodes [9].

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 optimized energy consumption configuration. This paper is focused on the quality-of-information (QoI) aware node classification and self-reconfiguration of visual sensing nodes to select optimized parameters for resource constrained scenarios. QoI is defined in literature as the degree to which the data is suitable for a given application or a decision making process [10]–[12]. 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 [13]. 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 wide range of heterogeneity within target QoI is a challenging task. In this paper, QoI is characterized by the quality of the visual data provided by a sensing node and quantified by a peak signal-to-noise ratio (PSNR) based metric, as in [13]. Furthermore, the configuration of a visual sensing node for optimization 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 [14]–[16]. Therefore, optimizing 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.

The contributions of this paper are summarized as follows:

1) 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. Utilizing the proposed in-node processing model, the scheme yields optimized configurations for visual sensing nodes resulting in substantial energy savings. Subsequently, providing the system design engineers with a trade-off model between reliability and energy efficiency.
2) A 3D coverage modelling scheme is proposed for visual sensing nodes to dynamically obtain their FoV information. The proposed scheme can be utilized in the network initialization phase to support intelligent sensing by making the sensing nodes aware of the targeted sensing environment.
3) By the heterogeneous realization 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 optimized energy utilization in resource constrained VSNs.
4) 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 the system design engineers with the confidence bounds for fine-tuning to the required QoI thresholds while attaining the desired reliability.

The rest of the paper is organized as follows: Section II reviews the state of the art in visual sensor networks, focusing on network reconfiguration and limitations of the existing schemes. Section III provides the system model. Section IV presents the proposed framework. Section V provides the performance analysis of the proposed framework. Section VI presents an analytical model to evaluate the proposed framework's performance reliability. Finally, Section VII concludes the paper and discusses future scope of the work.

II. RELATED WORK

In the recent years, a significant number of research studies have been conducted on network reconfiguration. The interests in this context lie in the optimization of: sensing nodes' spatial coverage, task allocation, resource utilization and visual data for transmission. Such optimizations aim to provide energy efficient solutions for the lifetime maximization of VSNs. Some notable schemes from the existing literature for optimization in resource constrained scenarios are discussed here.

An optimization technique based on particle swarm optimizer is proposed in [18] 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. In [19], an occlusion-aware method is proposed to dynamically obtain the optimal configuration of cameras' pan, tilt and zoom parameters in the network for coverage maximization. Dieber *et al.* [20] 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. Authors in Kyrkou *et al.* [21] proposed an adaptive market-based solution for the energy-aware assignment of tasks to cameras in resource constrained networks resulting in

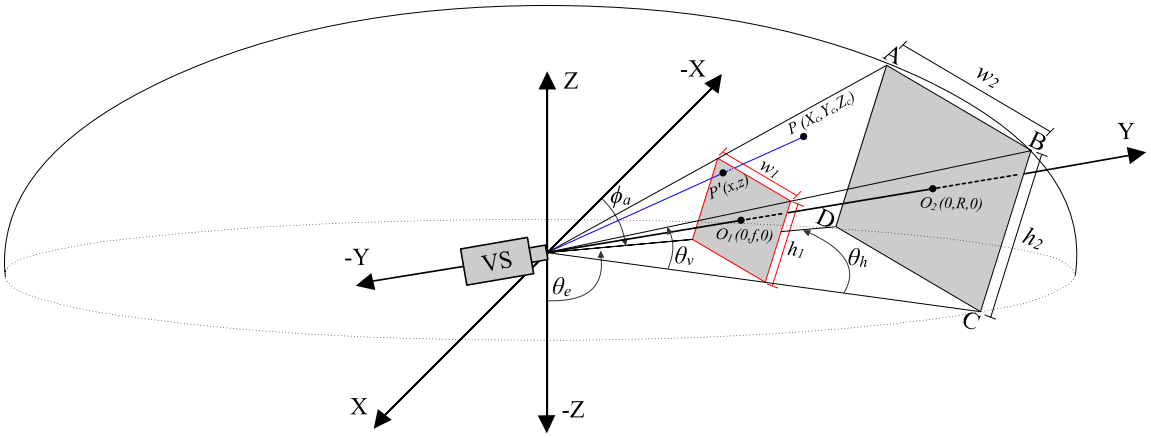

FIGURE 1. Visual sensing node's 3-dimensional projection model [17].

lifetime maximization. Considering 2D environment in Halder and Ghosal [22], authors proposed a scheme which employs a location-wise strategy for pre-determined deployment of sensing nodes attaining energy balancing and leading to optimized network lifetime.

An energy efficient image prioritization framework is proposed in [23] 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. Chow *et al.* [16] 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 to transmit images upon request to the mobile sink which leads to reduced energy consumption in the network. An energy efficient architecture for image processing and a protocol for communication in VSNs are proposed in [15]. Employing the proposed object detection architecture with DWT processing, the proposed approach leads to a reduction in image transmission cost. In [24], 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.

Although the existing work in literature provides solutions to optimize the energy consumption in resource constrained scenarios, many existing schemes assume simplified 2D camera model and sensing environment. Moreover, the existing schemes do not consider the heterogeneous realization of targeted QoI within the sensing region and dynamic target threshold based optimization of visual data prior to transmission. In contrast, the unified framework proposed in this paper incorporates 3D FoV modelling with dynamic realization of visual data to achieve heterogeneous target QoI thresholds. To the best of our knowledge, this is the first formulation which jointly considers the utilization of heterogeneous target QoI thresholds in sensing sub-regions with dynamic reconfiguration of visual sensing nodes to achieve targeted threshold based optimization. The proposed framework leads to reduced energy consumption within the network while maintaining an acceptable degree of reliability.

III. SYSTEM MODEL

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 $\{\mathrm{VS}_{\tilde{\ell}} | \tilde{\ell} = 1, 2, 3, \ldots, \mathcal{N}\}$. 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. The 3D sensing model of a visual sensing node [17] is depicted in Fig. 1; where R is the sensing range, θ_h and θ_v are the horizontal and vertical FoVs respectively; θ_e and ϕ_a are the elevation and azimuth angles respectively; h_2 and w_2 are the height and width of the ABCD-plane respectively; h_1 and w_1 are the height and width of the image plane respectively; and f is the focal length. 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. Moreover, the effective sensing range R of a visual sensing node for a given application can be estimated from [17] and is assumed to be known.

Energy conservation is a primary issue within resource constrained VSNs which is expected to be achieved by dynamic self-reconfiguration, for example, by the realization 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 Fig. 2. Let $\mathcal{C}_c = \{C_1^c, C_2^c, C_3^c, \ldots, C_{\tilde{n}}^c\}$ be the set of criteria provided by the design engineer for a given application and $\mathcal{C}_s = \{C_1^s, C_2^s, C_3^s, \ldots, C_{\tilde{n}}^s\}$ denotes the set of parameters in the network configuration space. Suppose $\mathcal{C}_o = \{C_1^o, C_2^o, C_3^o, \ldots, C_{\tilde{n}}^o\}$ represents the set of objectives to be achieved, for example, coverage maximization, energy

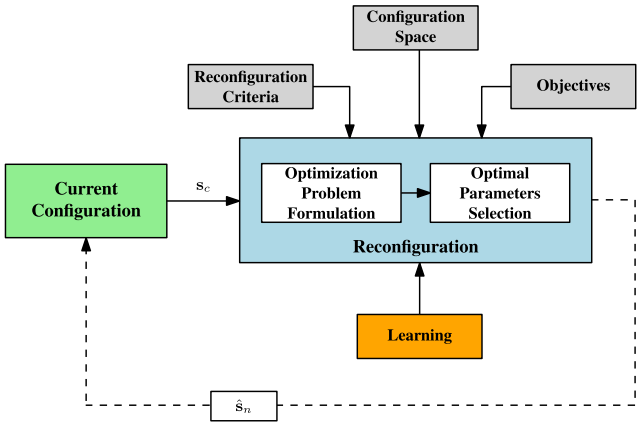

FIGURE 2. Visual sensing node self-reconfiguration model.

preservation, 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 Fig. 2 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 centralized and distributed categories. Distributed algorithms are preferred over the centralized ones due to the complexity and scalability issues of the latter [25]. 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 paper, a distributed decision making strategy for reconfiguration is devised where each node takes the decision independently to accelerate the decision making process. The list of key symbols used in this paper along with their definition is given in Table 1.

IV. PROPOSED FRAMEWORK

The proposed unified framework to enhance the energy efficiency of VSNs by achieving targeted threshold based optimization is presented in Fig. 3. In the proposed framework, training and calibration takes place in the pre-deployment phase which consists of training dataset selection, object appearance modelling, redundant feature removal, quality estimation and learning, as shown in Fig. 3a. In the post-deployment phase, the framework incorporates 3D coverage modelling with QoI-centric node classification, image capture, feature detection and object extraction, sensor-to-object distance estimation, self-reconfiguration and redundant feature removal, as shown in Fig. 3b.

A. TRAINING AND CALIBRATION

In the proposed framework, training and calibration takes 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.

TABLE 1. Key symbols and their definition.

Symbol	Definition
$\mathrm{VS}_{\bar{\ell}}$	$\bar{\ell}th$ visual sensing node
θ_h	Horizontal Field-of-View
θ_v	Vertical Field-of-View
R	Sensing range of a visual sensing node
R_r	Reference distance within close proximity of a sensing node
R_d	Sensor-to-object distance
$\mathbf{I}_{\bar{\ell}}$	Image captured by $\bar{\ell}th$ visual sensing node
$\mathcal{N}$	Number of required sensing nodes to cover area of interest
λ	QoI index
λ_t	Target QoI threshold in dB to be achieved for a given application
$\mathbf{\Lambda}_{\tilde{l}}$	Compressive calibration matrix (CCM)
$\mathbf{\Omega}_t$	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

1) DATASET SELECTION

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

2) OBJECT APPEARANCE MODELLING

In the proposed framework, 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 $\mathbf{I}_r$ denotes 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, it's appearance at the reference distance R_r can be modelled as,

$$\hat{\mathbf{I}}_r = \left(\mathbf{I}_r \downarrow D_d^{\pm}\right) \uparrow D_u^{\pm} \tag{1}$$

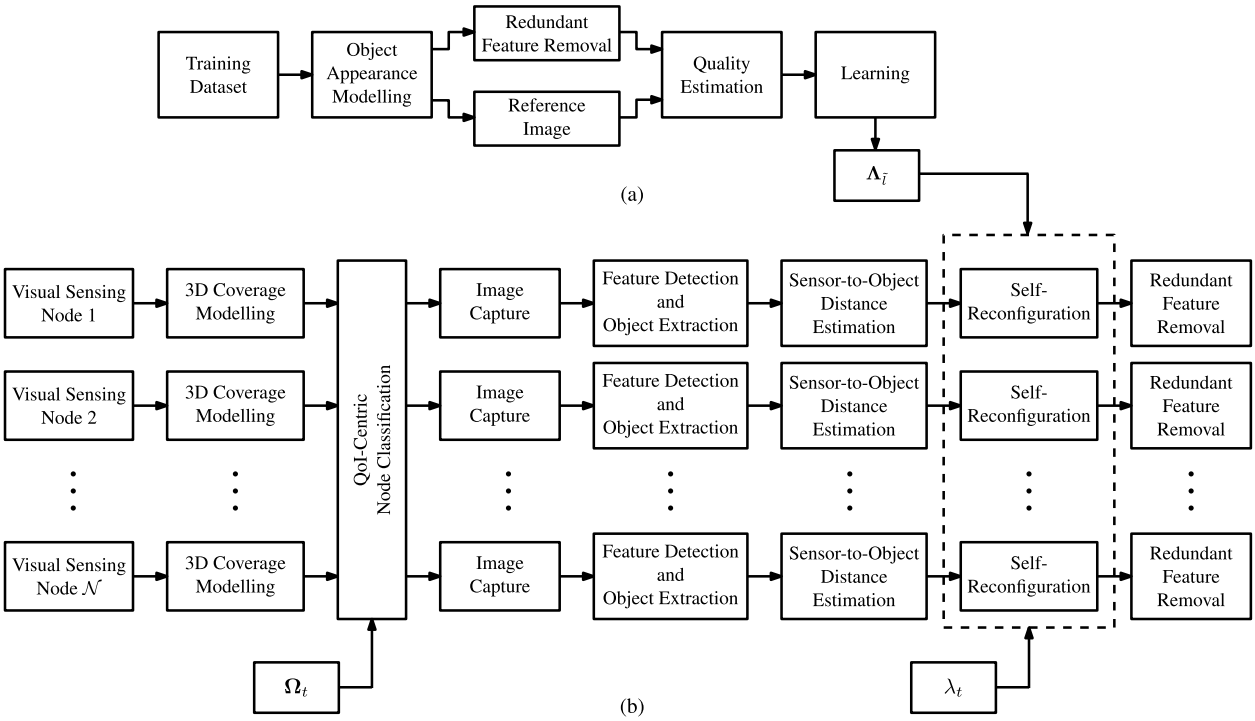

FIGURE 3. (a) Training and calibration process in the pre-deployment phase. (b) Proposed unified framework for node classification and self-reconfiguration in resource constrained VSNs to achieve targeted threshold based optimization.

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

$$D_d^{\pm} = D^{\pm}\left(\frac{R_r}{R_d}\right)$$
$$D_u^{\pm} = D^{\pm}\left(\frac{R_d}{R_r}\right) \tag{2}$$

3) REDUNDANT FEATURE REMOVAL

In resource constrained VSNs, energy conservation can be achieved through redundant feature removal, which leads to the minimization 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 $\mathbf{a}_c = \{\alpha_c^l, \alpha_c^l + \mathcal{S}_c, \alpha_c^l + 2\mathcal{S}_c, \ldots, \alpha_c^u\}$ and the resulting bits per pixel values can be denoted by $\mathbf{a}_b$ as $\mathbf{a}_b = \{\alpha_{b(l)} | l = 1, 2, 3, \ldots, l_m\}$; 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.

4) QUALITY ESTIMATION

As discussed earlier, redundant features can be removed to optimize 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 framework, PSNR is used as a quality metric, which is realized with the system's dynamic PSNR range in dB (ζ^l to ζ^u) to obtain the QoI index λ as,

$$\lambda = \frac{10\log_{10}\left(\frac{\mathcal{A}^2}{\mathcal{M}}\right) - \zeta^l}{\zeta^u - \zeta^l} \tag{3}$$

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 $\mathcal{M}$ represents the Mean Squared Error (MSE) given by,

$$\mathcal{M} = \frac{1}{3(D^+D^-)}\sum_{x=1}^{D^+}\sum_{z=1}^{D^-}\left[I_r(x,z) - \hat{I}_r(x,z)\right]^2 \tag{4}$$

Suppose the set of possible sensor-to-object distances is given by $\mathbf{r} = \{R_r, R_r + \mathcal{S}_r, R_r + 2\mathcal{S}_r, \ldots, R_u\}$ and let $\boldsymbol{\lambda} = \{\lambda_p \mid p = 1, 2, 3, \ldots, p_m\}$ be the respective quality metrics calculated using (3); 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,

$$\boldsymbol{\Lambda} = \begin{bmatrix} \lambda_{1,1} & \lambda_{1,2} & \ldots & \lambda_{1,p_m} \\ \lambda_{2,1} & \lambda_{2,2} & \ldots & \lambda_{2,p_m} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{l_m,1} & \lambda_{l_m,2} & \ldots & \lambda_{l_m,p_m} \end{bmatrix} \tag{5}$$

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

5) 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 $\mathbf{\Lambda}_{\tilde{l}}$ denotes 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,

$$\mathbf{\Lambda}_{\tilde{l}} = \frac{1}{\tilde{t}} \sum_{i=1}^{\tilde{t}} \mathbf{\Lambda}_i \tag{6}$$

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 utilization during its self-reconfiguration phases. The fidelity of CCM is quantified within a confidence bound $[e^-, e^+]$, measured in dB, given by,

$$e^{\pm} = \mathfrak{D}^{\pm}(\zeta^u - \zeta^l) \tag{7}$$

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 $(\bar{\cdot})$ represents the mean.

B. 3D COVERAGE MODELLING

Consider the visual sensing node's 3D projection model shown in Fig. 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 $\text{VS}_{\tilde{\ell}}$ ($\tilde{\ell} = \{1, 2, 3, \ldots, \mathcal{N}\}$) is identified by its location which is described by the local cartesian coordinates $\mathbf{s} = [X_{\tilde{\ell}}, Y_{\tilde{\ell}}, Z_{\tilde{\ell}}, 1]^{\mathrm{T}}$, azimuth angle ϕ_a and elevation angle θ_e. In the simplest scenario, the sensor's optical axis is assumed to be parallel with the y-axis, thus the azimuth and elevation angles can be given by $\phi_a = \frac{\pi}{2} - \frac{\theta_h}{2}$ and $\theta_e = \frac{\pi}{2} - \frac{\theta_v}{2}$ respectively; where ϕ_a is measured clockwise and θ_e is measured counter-clockwise.

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

$$\mathbf{p} = [X_o, Y_o, Z_o, 1]^{\mathrm{T}} = [X_{\tilde{\ell}}, Y_{\tilde{\ell}+R}, Z_{\tilde{\ell}}, 1]^{\mathrm{T}} \tag{8}$$

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,

$$\theta_{ya} = \begin{cases} \phi_a + \dfrac{\theta_h}{2} + \dfrac{3\pi}{2}, & 0 < \left(\phi_a + \dfrac{\theta_h}{2}\right) < \dfrac{\pi}{2} \\ \phi_a + \dfrac{\theta_h}{2} - \dfrac{\pi}{2}, & \dfrac{\pi}{2} \leq \left(\phi_a + \dfrac{\theta_h}{2}\right) \leq 2\pi \end{cases} \tag{9}$$

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,

$$\theta_{ye} = \begin{cases} \theta_e + \dfrac{\theta_v}{2} + \dfrac{3\pi}{2}, & 0 < \left(\theta_e + \dfrac{\theta_v}{2}\right) < \dfrac{\pi}{2} \\ \theta_e + \dfrac{\theta_v}{2} - \dfrac{\pi}{2}, & \dfrac{\pi}{2} \leq \left(\theta_e + \dfrac{\theta_v}{2}\right) \leq 2\pi \end{cases} \tag{10}$$

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 $\hat{c} \in \{1, 2, 3\text{a}, 3\text{b}\}$ denotes 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,

$$\mathbf{p}'_{\hat{c}} = \mathbf{T}^{-1} \Psi_{\hat{c}} \mathbf{T} \mathbf{p} \tag{11}$$

In (11), $\mathbf{T}$ is the translation matrix given by,

$$\mathbf{T} = \begin{bmatrix} \mathbf{J} & \mathbf{t} \\ \mathbf{z} & 1 \end{bmatrix} \tag{12}$$

where $\mathbf{J}$ is a 3×3 identity matrix, $\mathbf{t}$ represents the transformation coordinates given by $\mathbf{t} = [-p_{(1)}, R - p_{(2)}, -p_{(3)}]^{\mathrm{T}}$ and $\mathbf{z}$ denotes a 1×3 all-zeros vector.
$\mathbf{\Psi}_{\hat{c}}$ is expressed as,

$$\mathbf{\Psi}_{\hat{c}} = \begin{bmatrix} \mathbf{\Theta}_{\hat{c}} & \mathbf{z}^{\mathrm{T}} \\ \mathbf{z} & 1 \end{bmatrix} \tag{13}$$

where $\mathbf{\Theta}_{\hat{c}}$ denotes the rotation matrices $\mathbf{\Theta}_1$, $\mathbf{\Theta}_2$, $\mathbf{\Theta}_{3\text{a}}$ and $\mathbf{\Theta}_{3\text{b}}$ given by [26], [27],

$$\mathbf{\Theta}_1 = \begin{bmatrix} \cos\theta_{ya} & \sin\theta_{ya} & 0 \\ -\sin\theta_{ya} & \cos\theta_{ya} & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{14}$$

$$\mathbf{\Theta}_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_{ye} & -\sin\theta_{ye} \\ 0 & \sin\theta_{ye} & \cos\theta_{ye} \end{bmatrix} \tag{15}$$

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

$$\mathbf{\Theta}_{3\text{b}} = \begin{bmatrix} \cos\theta_{ya} & \sin\theta_{ya}\cos\theta_{ye} & -\sin\theta_{ya}\sin\theta_{ye} \\ -\sin\theta_{ya} & \cos\theta_{ya}\cos\theta_{ye} & -\cos\theta_{ya}\sin\theta_{ye} \\ 0 & \sin\theta_{ye} & \cos\theta_{ye} \end{bmatrix} \tag{16}$$

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,

$$\mathbf{p}_y = \{\mathcal{S}, 2\mathcal{S}, 3\mathcal{S}, \ldots, R\} \tag{17}$$

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,

$$\mathbf{p}_x = \left\{\gamma_x^-, \gamma_x^- + \mathcal{S}, \gamma_x^- + 2\mathcal{S}, \ldots, \gamma_x^+\right\} \tag{18}$$
$$\mathbf{p}_z = \left\{\gamma_z^-, \gamma_z^- + \mathcal{S}, \gamma_z^- + 2\mathcal{S}, \ldots, \gamma_z^+\right\} \tag{19}$$

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,

$$\left.\begin{matrix}\gamma_x^+ \\ \gamma_x^-\end{matrix}\right\} = X_o' \pm p_{y(\cdot)} \tan\left(\frac{\theta_h}{2}\right) \tag{20}$$

$$\left.\begin{matrix}\gamma_z^+ \\ \gamma_z^-\end{matrix}\right\} = Z_o' \pm p_{y(\cdot)} \tan\left(\frac{\theta_v}{2}\right) \tag{21}$$

and

$$s_x = \left\lfloor \frac{\gamma_x^+ - \gamma_x^- + \mathcal{S}}{\mathcal{S}} \right\rfloor \tag{22}$$

$$s_z = \left\lfloor \frac{\gamma_z^+ - \gamma_z^- + \mathcal{S}}{\mathcal{S}} \right\rfloor \tag{23}$$

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

$$t_p = \sum_{\check{i}=1}^{s_y} \left\lfloor \frac{2p_{y(\check{i})} \tan\left(\frac{\theta_h}{2}\right) + \mathcal{S}}{\mathcal{S}} \right\rfloor \left\lfloor \frac{2p_{y(\check{i})} \tan\left(\frac{\theta_v}{2}\right) + \mathcal{S}}{\mathcal{S}} \right\rfloor \tag{24}$$

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,

$$\begin{aligned}\mathbf{c}_x &= \left\{c_{x(1)}, c_{x(2)}, c_{x(3)}, \ldots, c_{x(t_p)}\right\} \\ \mathbf{c}_y &= \left\{c_{y(1)}, c_{y(2)}, c_{y(3)}, \ldots, c_{y(t_p)}\right\} \\ \mathbf{c}_z &= \left\{c_{z(1)}, c_{z(2)}, c_{z(3)}, \ldots, c_{z(t_p)}\right\}\end{aligned} \tag{25}$$

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

Algorithm 1 proposes a 3D coverage modelling scheme for visual sensing nodes to calculate $\mathbf{c}_x$, $\mathbf{c}_y$ and $\mathbf{c}_z$.

Algorithm 1 Proposed 3D Coverage Modelling Scheme for Visual Sensing Nodes

Input:

The sensing node's coordinates $[X_{\tilde{\ell}}, Y_{\tilde{\ell}}, Z_{\tilde{\ell}}, 1]^{\mathrm{T}}$, 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.

$l_c \leftarrow 1$
$\mathbf{p}_y = \{\mathcal{S}, 2\mathcal{S}, 3\mathcal{S}, \ldots, R\}$
$s_y \leftarrow \left\lfloor \frac{R}{\mathcal{S}} \right\rfloor$
for $\check{i} \leftarrow 1$ *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 $R = p_{y(\check{i})}$ in (8) as,
  $[X_o, Y_o, Z_o, 1]^{\mathrm{T}} = [X_{\tilde{\ell}}, Y_{\tilde{\ell}} + p_{y(\check{i})}, Z_{\tilde{\ell}}, 1]^{\mathrm{T}}$
  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 (20) and (21) respectively.
  $\mathbf{p}_x = \left\{\gamma_x^-, \gamma_x^- + \mathcal{S}, \gamma_x^- + 2\mathcal{S}, \ldots, \gamma_x^+\right\}$
  $\mathbf{p}_z = \left\{\gamma_z^-, \gamma_z^- + \mathcal{S}, \gamma_z^- + 2\mathcal{S}, \ldots, \gamma_z^+\right\}$
  $s_x \leftarrow \left\lfloor \left(\gamma_x^+ - \gamma_x^- + \mathcal{S}\right)/\mathcal{S} \right\rfloor$
  $s_z \leftarrow \left\lfloor \left(\gamma_z^+ - \gamma_z^- + \mathcal{S}\right)/\mathcal{S} \right\rfloor$
  for $\check{j} \leftarrow 1$ *to* s_x **do**
    for $\check{q} \leftarrow 1$ *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 (9) and (10) respectively.
      Substitute $\mathbf{p} = [p_{x(\check{j})}, Y_{\tilde{\ell}} + p_{y(\check{i})}, p_{z(\check{q})}, 1]^{\mathrm{T}}$ in (11) 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 $\mathbf{c}_x$, $\mathbf{c}_y$, $\mathbf{c}_z$

C. QoI-CENTRIC NODE CLASSIFICATION

In resource constrained scenarios, the utilization 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 realization in VSNs is more suitable as compared to the homogeneous QoI realization. Thus, dividing the overall region of interest within the sensing environment into smaller sub-regions and realizing heterogeneity in the target QoI thresholds is expected to preserve energy leading to improved network lifetime. In order to optimize the network's energy consumption based on such realization, 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 $\mathbf{C}_r^t = [\mathbf{C}_r^{\tilde{i}} | \tilde{i} = 1, 2, 3, \cdots, \hat{s}]$ be the matrix denoting the 3D coordinates of $\hat{s}$ sub-regions and $\mathbf{\Omega}_t = [\lambda_t^{\tilde{i}} | \tilde{i} = 1, 2, 3, \cdots, \hat{s}]$ 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 $\mathbf{C}_s^{\tilde{l}} = [\mathbf{c}_{x_s}^{\tilde{l}}, \mathbf{c}_{y_s}^{\tilde{l}}, \mathbf{c}_{z_s}^{\tilde{l}}]^{\mathrm{T}}$ denotes 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 is represented by $t_p^{\tilde{l}}$ and suppose $t_p^{\tilde{i}}$ denotes 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,

$$\mathcal{O}_{\tilde{i}\tilde{l}} = \frac{1}{\min\{t_p^{\tilde{l}}, t_p^{\tilde{i}}\}} \sum_{\check{i}=1}^{t_p^{\tilde{l}}} \sum_{\check{j}=1}^{t_p^{\tilde{i}}} \left\{ \prod_{\hat{q}\in\{x,y,z\}} 1 - \left[\mathrm{sgn}\left(c_{\hat{q}_s(\check{i})}^{\tilde{l}} - c_{\hat{q}_r(\check{j})}^{\tilde{i}} \right) \right]^2 \right\} \tag{26}$$

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 categorized as,

$$\mathcal{O}_{\tilde{i}\tilde{l}} = \begin{cases} 0, & (\mathbf{c}_{x_s}^{\tilde{l}} \nsubseteq \mathbf{c}_{x_r}^{\tilde{i}}) \cap (\mathbf{c}_{y_s}^{\tilde{l}} \nsubseteq \mathbf{c}_{y_r}^{\tilde{i}}) \cap (\mathbf{c}_{z_s}^{\tilde{l}} \nsubseteq \mathbf{c}_{z_r}^{\tilde{i}}) \\ (0,1), & (\mathbf{c}_{x_s}^{\tilde{l}} \subset \mathbf{c}_{x_r}^{\tilde{i}}) \cap (\mathbf{c}_{y_s}^{\tilde{l}} \subset \mathbf{c}_{y_r}^{\tilde{i}}) \cap (\mathbf{c}_{z_s}^{\tilde{l}} \subset \mathbf{c}_{z_r}^{\tilde{i}}) \\ 1, & (\mathbf{c}_{x_s}^{\tilde{l}} \subseteq \mathbf{c}_{x_r}^{\tilde{i}}) \cap (\mathbf{c}_{y_s}^{\tilde{l}} \subseteq \mathbf{c}_{y_r}^{\tilde{i}}) \cap (\mathbf{c}_{z_s}^{\tilde{l}} \subseteq \mathbf{c}_{z_r}^{\tilde{i}}) \end{cases} \tag{27}$$

In case of $t_p^{\tilde{l}} > t_p^{\tilde{i}}$, (27) is modified as,

$$\mathcal{O}_{\tilde{i}\tilde{l}} = \begin{cases} 0, & (\mathbf{c}_{x_r}^{\tilde{i}} \nsubseteq \mathbf{c}_{x_s}^{\tilde{l}}) \cap (\mathbf{c}_{y_r}^{\tilde{i}} \nsubseteq \mathbf{c}_{y_s}^{\tilde{l}}) \cap (\mathbf{c}_{z_r}^{\tilde{i}} \nsubseteq \mathbf{c}_{z_s}^{\tilde{l}}) \\ (0,1), & (\mathbf{c}_{x_r}^{\tilde{i}} \subset \mathbf{c}_{x_s}^{\tilde{l}}) \cap (\mathbf{c}_{y_r}^{\tilde{i}} \subset \mathbf{c}_{y_s}^{\tilde{l}}) \cap (\mathbf{c}_{z_r}^{\tilde{i}} \subset \mathbf{c}_{z_s}^{\tilde{l}}) \\ 1, & (\mathbf{c}_{x_r}^{\tilde{i}} \subseteq \mathbf{c}_{x_s}^{\tilde{l}}) \cap (\mathbf{c}_{y_r}^{\tilde{i}} \subseteq \mathbf{c}_{y_s}^{\tilde{l}}) \cap (\mathbf{c}_{z_r}^{\tilde{i}} \subseteq \mathbf{c}_{z_s}^{\tilde{l}}) \end{cases} \tag{28}$$

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,

$$\xi_{\tilde{l}} = \arg\max_{\tilde{i}} \left[\frac{1}{\min\{t_p^{\tilde{l}}, t_p^{\tilde{i}}\}} \sum_{\check{i}=1}^{t_p^{\tilde{l}}} \sum_{\check{j}=1}^{t_p^{\tilde{i}}} \left\{ \prod_{\hat{q}\in\{x,y,z\}} 1 - \left[\mathrm{sgn}\left(c_{\hat{q}_s(\check{i})}^{\tilde{l}} - c_{\hat{q}_r(\check{j})}^{\tilde{i}} \right) \right]^2 \right\} \right] \tag{29}$$

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

$$\mathcal{N}_{\hat{m}} = \mathcal{N}\,(1 - f_{\hat{m}}) \tag{30}$$

where,

$$f_{\hat{m}} = 1 - \frac{\sum_{\tilde{l}=1}^{\mathcal{N}} 1 - \left[\mathrm{sgn}\left(\xi_{\tilde{l}} - \hat{m}\right)\right]^2}{\mathcal{N}} \tag{31}$$

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

$$\mathcal{N}_{\hat{m}} = \sum_{\tilde{l}=1}^{\mathcal{N}} 1 - \left[\mathrm{sgn}\left(\xi_{\tilde{l}} - \hat{m}\right)\right]^2 \tag{32}$$

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

D. IN-NODE PROCESSING MODEL

1) IMAGE CAPTURE

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

2) 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 $\mathbf{I}_{\tilde{\ell}}$ belonging to the object of interest O_b can be given by,

$$\mathcal{D}(x,z) = \begin{cases} 1, & (x,z) \in O_b \\ 0, & \text{otherwise} \end{cases} \tag{33}$$

This can be accomplished by utilizing an appropriate detection method depending on the given application; for example, [28] and [29] can be considered for face detection, [30] can be used for human detection and vehicles can be detected using [31] and [32]. The pixels probabilities from (33) 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,

$$e_d = \begin{cases} 1, & \gamma_{\hat{p}} \geq \gamma_t \\ 0, & \text{otherwise} \end{cases} \tag{34}$$

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

$$\gamma_t = \frac{A \times D^+ \times D^- \times \sin\theta_h \times \sin\theta_v}{4R^2\,(1-\cos\theta_h)\,(1-\cos\theta_v)} \tag{35}$$

The object of interest is extracted from $\mathbf{I}_{\tilde{\ell}}$ by image segmentation using the following equation [33],

$$\mathbf{S}_g = e_d \cdot \mathbf{I}_{\tilde{\ell}} \cdot \mathbf{S}_m \tag{36}$$

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

3) 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, utilizing the output of feature detection and extraction process, R_d is estimated prior to reconfiguration, as shown below,

$$R_d = \frac{1}{2}\sqrt{\frac{A \times D^+ \times D^- \times \sin\theta_h \times \sin\theta_v}{\gamma_{\hat{p}}\,(1-\cos\theta_h)\,(1-\cos\theta_v)}} \tag{37}$$

4) SELF-RECONFIGURATION

Suppose λ_t denotes the target QoI in dB to be achieved for a given application and γ_b represents 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 maximizing 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,

$$M_{l,p} = \begin{cases} 1, & \Lambda_{\tilde{l}(l,p)} \geq \hat{\lambda}_t \\ 0, & \text{otherwise} \end{cases} \tag{38}$$

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,

$$\mathbf{\Lambda}_a = \mathbf{\Lambda}_{\tilde{l}} \cdot \mathbf{M} \tag{39}$$

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

$$\mathbf{E} = \begin{bmatrix} E^t_{1,1} & E^t_{1,2} & \cdots & E^t_{1,p_m} \\ E^t_{2,1} & E^t_{2,2} & \cdots & E^t_{2,p_m} \\ \vdots & \vdots & \ddots & \vdots \\ E^t_{l_m,1} & E^t_{l_m,2} & \cdots & E^t_{l_m,p_m} \end{bmatrix} \tag{40}$$

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

Within the context of the visual sensing node self-reconfiguration model depicted in Fig 2, the compression level γ_b forms the configuration space, target QoI threshold λ_t is the reconfiguration criterion and minimization of the transmission energy cost $E^t_{l,p}$ is the ultimate objective. Hence, the following optimization problem can be stated,

$$\begin{aligned} &\text{minimize } E^t_{l,p} \\ &\text{subject to } \beta = \lambda_t \end{aligned} \tag{41}$$

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 2 to find optimal compression level $\gamma_b \in \mathbf{a}_b$ that solves the optimization problem expressed in (41).

Algorithm 2 Proposed Dynamic Self-Reconfiguration Scheme

Input:

The target QoI threshold λ_t, the set of possible sensor-to-object distances $\mathbf{r}$, 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 optimized 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
$\mathbf{\Lambda}_a \leftarrow \mathbf{\Lambda}_{\tilde{l}} \cdot \mathbf{M}$
$\gamma_1 \leftarrow \arg\min_p \left[|R_d - r_{(p)}|\right]; p = \{1, 2, 3, \ldots, p_m\}$
$\mathbf{t} \leftarrow \mathbf{\Lambda}_{a(1:l_m,\gamma_1)}$
$\gamma_2 \leftarrow \arg\min_{\hat{p}} \left[|\lambda_t - t_{(\hat{p})}|\right]; \hat{p} = \{1, 2, 3, \ldots, l_m\}$
$\beta \leftarrow t_{(\gamma_2)}$
$E^t_{min} \leftarrow \mathcal{L}\left(\alpha_{b(\gamma_2)}\right)$
$\gamma_b \leftarrow a_{b(\gamma_2)}$
$\hat{s}_n \leftarrow \gamma_b$
return $\hat{s}_n$

5) 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,

$$\mathbf{S}_r = \mathcal{G}(\mathbf{S}_g, \hat{s}_n) \tag{42}$$

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,

$$\hat{\mathbf{S}}_g = \mathcal{G}^{-1}(\mathbf{S}_r) \tag{43}$$

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

E. ENERGY MODEL

The performance of a visual sensor network is characterized by the energy conservation within the network. Fig. 4 shows

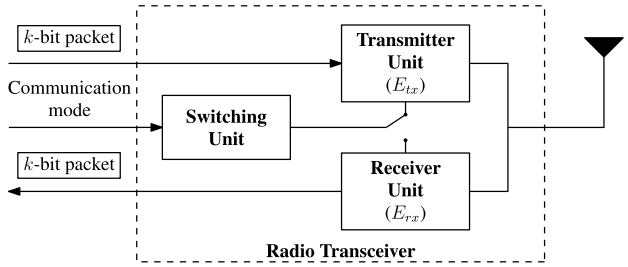

FIGURE 4. Communication energy dissipation model.

the energy dissipation model of a radio transceiver [34], which is characterized 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 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,

$$\tilde{E}_c = \sum_{\hat{m}=1}^{\mathcal{H}+1} \sum_{\hat{l}=1}^{\mathcal{N}_{\hat{m}}} \left[E_{tx} \left(\mathcal{N}_t \hat{k}_{\hat{l}}^{\hat{m}} + n_t e_{\hat{l}}^t \right) + E_{rx} \left(\mathcal{N}_r \tilde{k}_{\hat{l}}^{\hat{m}} + n_r e_{\hat{l}}^r \right) \right] \tag{44}$$

In order to calculate the energy consumption of the visual sensing nodes, the energy model of [35] is used. The parameters used in this paper to model the energy consumption of the radio transceiver are given in Table 2. In VSNs, the energy cost incurred for visual data transmission and reception is significantly higher than the data processing cost [14]–[16]. Therefore, only the communication energy cost is considered in this paper to model the energy dissipation of a visual sensing node. 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 receive energy cost at the sink node does not influence lifetime of the VSN and can be ignored.

V. RESULTS

This section demonstrates the performance of the proposed framework for resource constrained VSNs. The Long Distance Heterogeneous Face (LDHF) dataset [36], [37] is used

TABLE 2. Energy-measurement parameters [35].

Parameter	Value
Data consumption rate	144 kbps
Transmission cost	2.20×10^{-7} J/bit
Receiving cost	2.92×10^{-6} J/bit

TABLE 3. 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

FIGURE 5. An image subset from the LDHF dataset containing cropped faces for $R_r = 1$ m.

for training and calibration and its specification is presented in Table 3. This particular dataset is 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 Fig. 5. 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 $\hat{t}_s$ denotes a subset consisting of 15 test images selected randomly from the dataset, i.e. $\hat{t}_s \subset \hat{t}_m$. Utilizing 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 distance $R_d \in [2, 40]$ and the resulting QoI index is shown in Fig. 6. 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. Moreover, it can be

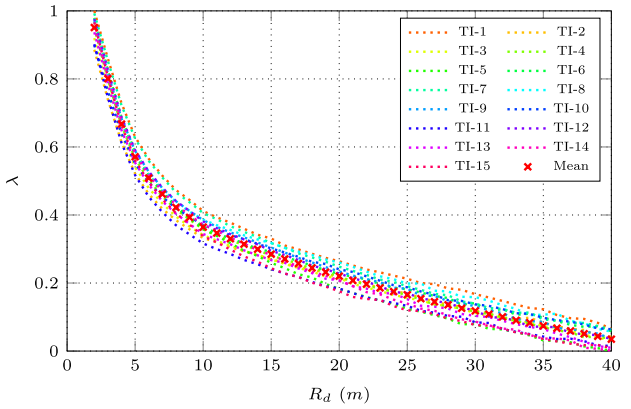

FIGURE 6. QoI index estimation for object appearance modelling utilizing test images (TIs) from the LDHF dataset.

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.

In order to analyze 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 [38] and utilization in many existing schemes [15], [16], [22]. However, 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, (6) is used to obtain the compressive calibration matrix, which is shown in Fig. 7a and Fig. 7b for 2D-DWT and 2D-DCT respectively. It is observed that increasing the level of redundancy removal results in QoI index reduction. Consequently, 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 Fig. 7a and Fig. 7b 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 minimize 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 [39]. Therefore, the overall transmission cost of the DCT-aided scheme is expected to be higher than the DWT-aided scheme.

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 [17]. The azimuth and elevation angles are bounded within $[0°, 360°]$ and $[0°, 155°]$ respectively. 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 realization; $\mathcal{H} = 1$ and $\mathcal{H} = 2$ refer to heterogeneous realization 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. The movements of objects within the FoV of visual sensing nodes are modelled randomly in the simulation. As this paper focuses on the network self-reconfiguration, 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 2.

Consider a homogeneous realization of target QoI i.e. $\mathcal{H} = 0$. 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 utilizes 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 Fig. 8. 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. It is observed from Fig. 8a and Fig. 8b 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

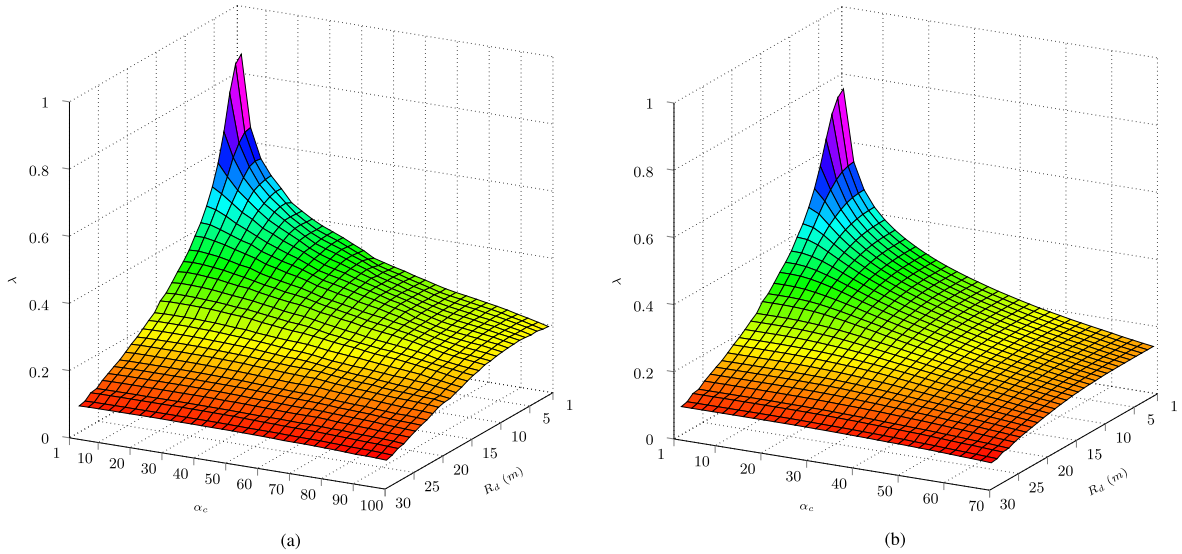

FIGURE 7. QoI index estimation after object appearance modelling and redundant feature removal with: (a) 2D-DWT and (b) 2D-DCT, to obtain compressive calibration matrix.

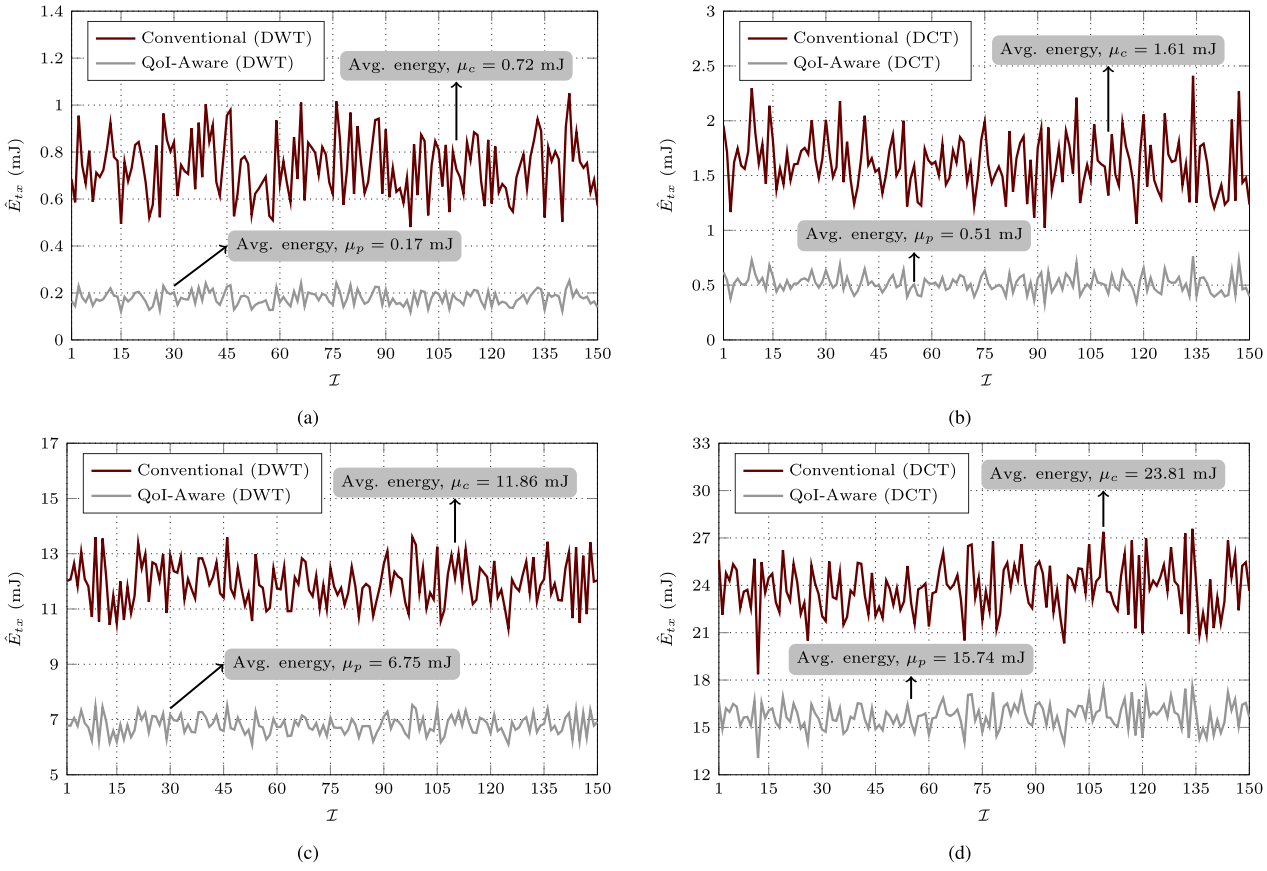

FIGURE 8. Comparison of the average energy cost incurred per node to transmit an image frame with the proposed QoI-aware scheme and conventional scheme for: target QoI $\lambda_t = 31$dB with (a) 2D-DWT and (b) 2D-DCT; target QoI $\lambda_t = 37$dB with (c) 2D-DWT and (d) 2D-DCT.

respectively. Similarly, as shown in Fig. 8c and Fig. 8d, 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.

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 $\lambda_t \in [30, 40]$

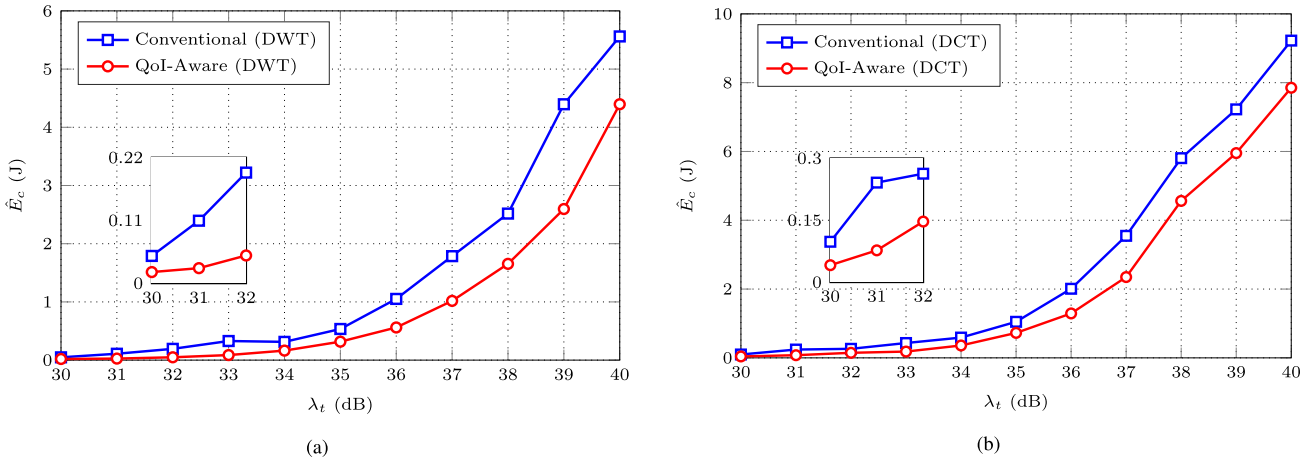

FIGURE 9. Comparison of the average energy cost incurred per node to transmit $\mathcal{N}_t = 150$ image frames with the proposed QoI-aware scheme and conventional scheme to achieve homogeneous target QoI thresholds with: (a) 2D-DWT and (b) 2D-DCT.

TABLE 4. Comparison of the proposed scheme with existing state-of-the-art techniques for homogeneous (HM) and heterogeneous (HT) QoI realizations.

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 [15]	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 [22]	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.* [16]	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

for $\mathcal{N}_t = 150$ image frames with 2D-DWT and 2D-DCT is depicted in Fig. 9. The results demonstrate that for any given target QoI threshold, the proposed scheme minimizes the transmission energy consumption and thus enhances the energy efficiency of the visual sensing node.

Consider a heterogeneous realization 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 for target QoI with first and second degree of heterogeneity respectively. The energy consumption of the proposed scheme for homogeneous and heterogeneous target QoI realizations is compared with the existing state-of-the-art approaches proposed in [15], [16] and [22] and the results are summarized in Table 4. Scenarios 1a, 2a and 3a present the performance analysis for homogeneous target QoI realization; scenarios 1b, 2b and 3b show the performance analysis for heterogeneous target QoI realization with tight thresholds; and scenarios 1c, 2c and 3c provide the performance analysis for heterogeneous target QoI realization with relaxed thresholds. It is observed from the comparison that the proposed scheme 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 due to the

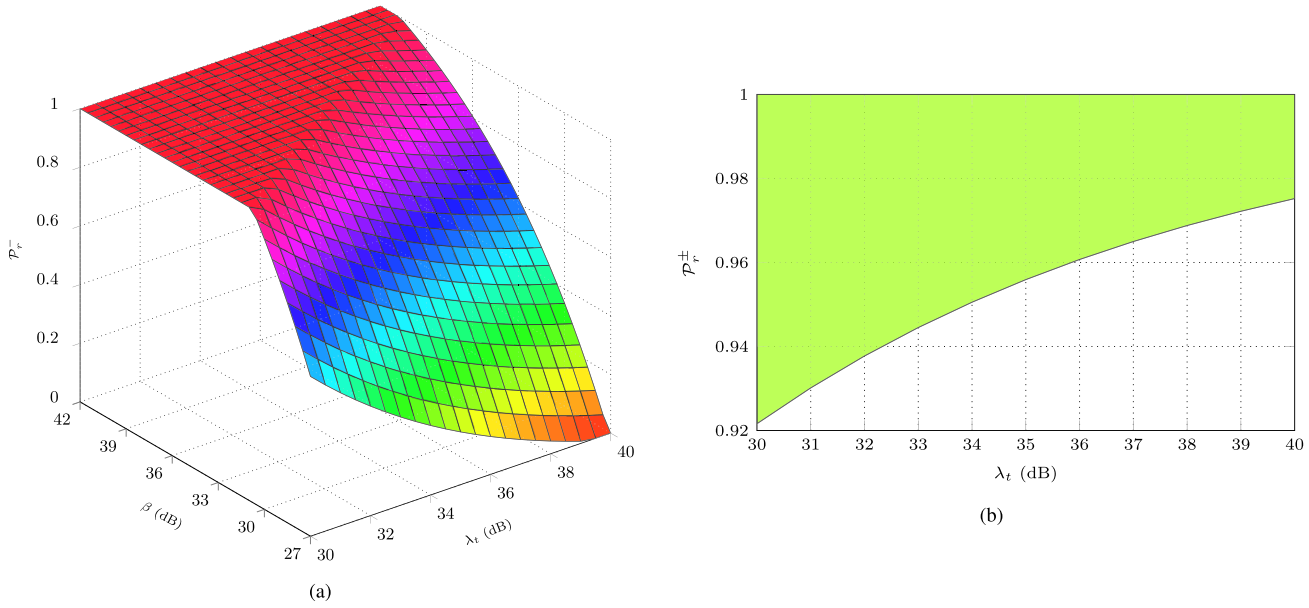

FIGURE 10. System reliability to achieve targeted threshold based optimization. (a) The minimum level of reliability for a range set of target QoI threshold (λ_t) and delivered QoI (β). (b) Proposed framework's confidence bound for retrieving the information.

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 [15], [16] and [22] are static for particular target QoI threshold. 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 visual sensor networks with strict constraints on their available energy.

VI. ANALYSIS OF PROPOSED FRAMEWORK'S PERFORMANCE RELIABILITY

In order to analyze 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 $[e^-, e^+]$, 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,

$$\mathcal{P}_r^{\pm} = \begin{cases} 1 - \xi_p^{\pm}, & \lambda_t \geq \beta + e^{\pm} \\ 1, & \text{otherwise} \end{cases} \tag{45}$$

where $\xi_r^{\pm}$ denotes the probability of failure to ensure the target QoI satisfaction and it is calculated from (46), as shown at the bottom of this page with $m = 2^{\hat{\alpha}_b - 1}$ ($\hat{\alpha}_b$ denotes the number of bits per pixel), $e^{\pm}$ is obtained from (7) and $Q(\cdot)$ is the Q-function given by,

$$Q(\hat{x}) = \frac{1}{\sqrt{2\pi}} \int_{\hat{x}}^{\infty} e^{-\frac{1}{2}q^2} dq \tag{47}$$

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 realization 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 (7), e^- and e^+ for the proposed framework are found to be -0.78 dB and 0.87 dB respectively. Substituting the parameters in (45) and (46), the probability of ensuring the targeted QoI thresholds is shown in Fig. 10. 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 Fig. 10a. It is observed from Fig. 10a that the reliability increases with the increase in delivered QoI and attains a maximum value when $\beta + e^- \geq \lambda_t$.

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 Fig. 10b. It is observed that the reliability, bounded between $\mathcal{P}_r^-$ and $\mathcal{P}_r^+$ shown in the shaded

$$\xi_p^{\pm} = \left[Q\left(\frac{10^{2\log(m) - \frac{\lambda_t}{20}} - 10^{2\log(m) - \frac{\beta + e^{\pm}}{20}}}{\sum_{k=0}^{\hat{\alpha}_b - 1} 2^{2k}} \right) - 0.50 \right] \cdot \left[Q\left(\frac{10^{2\log(m) - \frac{\zeta^u}{20}} - 10^{2\log(m) - \frac{\zeta^l}{20}}}{\sum_{k=0}^{\hat{\alpha}_b - 1} 2^{2k}} \right) - 0.50 \right]^{-1} \tag{46}$$

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. A comparison of the minimum QoI achievement reliability for homogeneous and heterogeneous QoI realizations is presented in Table 4. 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 realizations 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.

VII. CONCLUSION

A novel unified framework for the classification and self-reconfiguration of nodes in resource constrained VSNs is proposed. The proposed framework incorporates a QoI-centric node classification scheme along with a 3D coverage modelling scheme by exploiting the heterogeneity of the targeted QoI threshold levels within the sensing region. This provides the flexibility to dynamically classify visual sensing nodes based on their FoV. This paper also proposes a QoI-aware dynamic self-reconfiguration scheme to obtain suitable configurations of visual sensing nodes for visual data optimization prior to transmission. 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 decentralized, which accelerates the decision making process. An analytical model is formulated to quantify the degree to which the targeted QoI thresholds are achieved by a visual sensing node. For given target thresholds of QoI, it is observed that the proposed heterogeneous framework resulted in significant amount of energy savings compared to the existing state-of-the-art techniques, as reported in the simulation 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. For the future extension of this work, the authors intend to utilize the proposed 3D coverage modelling scheme to develop reconfiguration models for coverage and redundancy optimization of PTZ capable visual sensing nodes. 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.

REFERENCES

[1] X. Wang, S. Wang, and D. Bi, "Distributed visual-target-surveillance system in wireless sensor networks," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 39, no. 5, pp. 1134–1146, Oct. 2009.

[2] 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.

[3] A. O. Ercan, A. El Gamal, and L. J. Guibas, "Object tracking in the presence of occlusions using multiple cameras: A sensor network approach," *ACM Trans. Sensor Netw.*, vol. 9, no. 2, p. 16, 2013.

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

[5] 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*. Cham, Switzerland: Springer, 2015, pp. 26–38.

[6] S. Fleck and W. Strasser, "Smart camera based monitoring system and its application to assisted living," *Proc. IEEE*, vol. 96, no. 10, pp. 1698–1714, Oct. 2008.

[7] 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 Trans. Commun.*, vol. 57, no. 11, pp. 3406–3414, Nov. 2009.

[8] R. Zhu, M. Ma, Y. Zhang, and J. Hu, "Collaborative wireless sensor networks and applications," *Int. J. Distrib. Sensor Netw.*, vol. 2015, 2015, Art. no. 352761.

[9] C. Piciarelli, L. Esterle, A. Khan, B. Rinner, and G. L. Foresti, "Dynamic reconfiguration in camera networks: A short survey," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 26, no. 5, pp. 965–977, May 2016.

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

[11] Z. Song, C. H. Liu, J. Wu, J. Ma, and W. Wang, "QoI-aware multitask-oriented dynamic participant selection with budget constraints," *IEEE Trans. Veh. Technol.*, vol. 63, no. 9, pp. 4618–4632, Nov. 2014.

[12] C. H. Liu, J. Fan, J. W. Branch, and K. K. Leung, "Toward QoI and energy-efficiency in Internet-of-Things sensory environments," *IEEE Trans. Emerg. Topics Comput.*, vol. 2, no. 4, pp. 473–487, Dec. 2014.

[13] E. Gelenbe and L. Hey, "Quality of information: An empirical approach," in *Proc. 5th IEEE Int. Conf. Mobile Ad Hoc Sensor Syst.*, Sep./Oct. 2008, pp. 730–735.

[14] 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*. The Netherlands: Springer, 2008, pp. 520–525.

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

[16] K.-Y. Chow, K.-S. Lui, and E. Y. Lam, "Efficient on-demand image transmission in visual sensor networks," *EURASIP J. Appl. Signal Process.*, vol. 2007, no. 1, p. 225, 2007.

[17] A. Amjad, M. Patwary, A. Griffiths, and A.-H. Soliman, "Characterization of field-of-view for energy efficient application-aware visual sensor networks," *IEEE Sensors J.*, vol. 16, no. 9, pp. 3109–3122, May 2016.

[18] K. R. Konda, N. Conci, and F. De Natale, "Global coverage maximization in PTZ-camera networks based on visual quality assessment," *IEEE Sensors J.*, vol. 16, no. 16, pp. 6317–6332, Aug. 2016.

[19] C. Piciarelli, C. Micheloni, and G. L. Foresti, "Occlusion-aware multiple camera reconfiguration," in *Proc. 4th ACM/IEEE Int. Conf. Distrib. Smart Cameras*, Aug. 2010, pp. 88–94.

[20] B. Dieber, C. Micheloni, and B. Rinner, "Resource-aware coverage and task assignment in visual sensor networks," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 21, no. 10, pp. 1424–1437, Oct. 2011.

[21] C. Kyrkou, C. Laoudias, T. Theocharides, C. G. Panayiotou, and M. Polycarpou, "Adaptive energy-oriented multitask allocation in smart camera networks," *IEEE Embedded Syst. Lett.*, vol. 8, no. 2, pp. 37–40, Jun. 2016.

[22] S. Halder and A. Ghosal, "A location-wise predetermined deployment for optimizing lifetime in visual sensor networks," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 26, no. 6, pp. 1131–1145, Jun. 2016.

[23] I. Mehmood, M. Sajjad, W. Ejaz, and S. W. Baik, "Saliency-directed prioritization of visual data in wireless surveillance networks," *Inf. Fusion*, vol. 24, pp. 16–30, Jul. 2015.

[24] Y. Wang, D. Wang, X. Zhang, J. Chen, and Y. Li, "Energy-efficient image compressive transmission for wireless camera networks," *IEEE Sensors J.*, vol. 16, no. 10, pp. 3875–3886, Mar. 2016.
[25] 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. 10506–10523, 2010.
[26] F. Dunn and I. Parberry, *3D Math Primer for Graphics and Game Development*. Boca Raton, FL, USA: CRC Press, 2015.
[27] E. Lengyel, *Mathematics for 3D Game Programming and Computer Graphics*. Boston, MA, USA, Cengage Learning, 2012.
[28] P. Viola and M. J. Jones, "Robust real-time face detection," *Int. J. Comput. Vis.*, vol. 57, no. 2, pp. 137–154, May 2004.
[29] K. Lin, X. Wang, S. Cui, and Y. Tan, "Heterogeneous feature fusion-based optimal face image acquisition in visual sensor network," in *Proc. IEEE Int. Instrum. Meas. Technol. Conf. (I2MTC)*, May 2015, pp. 1078–1083.
[30] 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," *Proc. IEEE*, vol. 89, no. 10, pp. 1382–1402, Oct. 2001.
[31] X. Li and X. Guo, "Vision-based method for forward vehicle detection and tracking," in *Proc. IEEE Int. Conf. Mech. Autom. Eng. (MAEE)*, Jul. 2013, pp. 128–131.
[32] H. Hajimolahoseini, R. Amirfattahi, and H. Soltanian-Zadeh, "Robust vehicle tracking algorithm for nighttime videos captured by fixed cameras in highly reflective environments," *IET Comput. Vis.*, vol. 8, no. 6, pp. 535–544, 2014.
[33] A. Amjad, A. Griffiths, and M. N. Patwary, "Multiple face detection algorithm using colour skin modelling," *IET Image Process.*, vol. 6, no. 8, pp. 1093–1101, Nov. 2012.
[34] V. Lecuire, C. Duran-Faundez, and N. Krommenacker, "Energy-efficient transmission of wavelet-based images in wireless sensor networks," *J. Image Video Process.*, vol. 2007, no. 1, p. 15, 2007.
[35] A. Redondi, D. Buranapanichkit, M. Cesana, M. Tagliasacchi, and Y. Andreopoulos, "Energy consumption of visual sensor networks: Impact of spatio-temporal coverage," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 24, no. 12, pp. 2117–2131, Dec. 2014.
[36] 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*. Berlin, Germany: Springer, 2012, pp. 708–721.
[37] D. Kang, H. Han, A. K. Jain, and S.-W. Lee, "Nighttime face recognition at large standoff: Cross-distance and cross-spectral matching," *Pattern Recognit.*, vol. 47, no. 12, pp. 3750–3766, 2014.
[38] L. W. Chew, L.-M. Ang, and K. P. Seng, "Survey of image compression algorithms in wireless sensor networks," in *Proc. IEEE Int. Symp. Inf. Technol.*, vol. 4. Aug. 2008, pp. 1–9.
[39] 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 Commun. Surveys Tut.*, vol. 15, no. 3, pp. 963–972, 3rd Quart., 2013.

ANAS AMJAD (S'15) received the B.Eng. degree (Hons.) in electronic engineering from Staffordshire University, Stafford, U.K., in 2011. He is currently pursuing the Ph.D. degree in electronic engineering with Staffordshire University, Stoke-on-Trent, U.K. He is also a member of Sensing, Processing and Communication Research Group, Staffordshire University. His current research interests include resource optimization for wireless visual sensor networks, feature detection and extraction, visual content compression, and quality assessment for future generation of wireless networks.

ALISON GRIFFITHS received the B.Eng., (Hons.) M.Eng., and Ph.D. degrees from Staffordshire University, Stoke-on-Trent, U.K., in 1998, 1999, and 2004, respectively. She was with several companies on industrial collaborations. She is currently a full time Senior Lecturer with the Faculty of Computing, Engineering and Sciences, Staffordshire University. Her research interests include control system technologies, estimation and optimization techniques, wireless sensor networks, and battery life optimization techniques.

MOHAMMAD PATWARY (SM'11) received the B.Eng. degree (Hons.) in electrical and electronic engineering from the Chittagong University of Engineering and Technology, Bangladesh, in 1998, and the Ph.D. degree in telecommunication engineering from the University of New South Wales, Sydney, Australia, in 2005. He was with General Electric Company, Bangladesh, from 1998 to 2000, and Southern-Poro Communications, Sydney, Australia, from 2001 to 2002, as a Research and Development Engineer. He was with the University of New South Wales, Sydney, Australia, as a Lecturer, from 2005 to 2006, Staffordshire University, U.K., as a Senior Lecturer, from 2006 to 2010. He is currently a Professor of Wireless Systems & Digital Productivity and the Chair of the Center of Excellence on Digital Productivity with Connected Services (DiPConS), Staffordshire University, U.K. His current research interests of DiPConS include future generation of cellular network architecture; signal detection, estimation and optimization for future generation of wireless receivers; joint source-channel characterization for wireless sensor networks (heterogeneous networks & Internet of Things), and spatial diversity schemes for wireless communications.