ECTI Transactions on Computer and Information Technology
Journal homepage: https://ph01.tci-thaijo.org/index.php/ecticit/
Published by the ECTI Association, Thailand, ISSN: 2286-9131

BCS-Based Encoding Schemes for Monoview and Multiview Visual Configurations in WVSN Data-Gathering: A Survey

G. L. Priya[1] and Debashis Ghosh[2]

ABSTRACT

Wireless Visual Sensor Network (WVSN) has become a valuable tool in addressing the evolving needs of modern monitoring systems. Encoding in WVSNs is a multifaceted process that involves compressing visual data, optimizing energy consumption, ensuring error resilience, and adapting to various network and application requirements. The associated lightweight encoders and the demand for less storage space make block compressive sensing (BCS) techniques suitable for WVSN applications where energy, bandwidth, and storage resources are limited. Based on the number of visual perspectives or camera angles available within a network for data capture, there are two primary configurations: monoview and multiview. This paper provides a comprehensive survey of different BCS-based encoding schemes used for data-gathering in both monoview and multiview scenarios within WVSNs. A comparative study of these algorithms based on compression level, computational complexity, relative gain in encoder energy, and reconstruction quality is performed. A BCS-based joint encoding scheme for multiview configuration that ensures a relatively high compression level is also proposed in this paper.

Article information:

Keywords: Wireless Visual Sensor Network, Block Compressive Sensing, Monoview, Multiview, Sparsity

Article history:
Received: April 1, 2024
Revised: May 30, 2024
Accepted: July 4, 2024
Published: September 28, 2024 (Online)

DOI: 10.37936/ecti-cit.2024184.256278

1. INTRODUCTION

Wireless Visual Sensor Networks (WVSNs) integrate visual sensors with wireless communication to monitor and capture visual information from remote environments. The visual configuration of a network for data capture can be categorized into two types based on the number of visual perspectives or camera views available: Monoview and multiview. In both configurations, transmitting raw data from the sensor node to the sink node demands many network resources. The encoding process helps optimize the use of limited resources, such as bandwidth and energy, while maintaining the integrity and quality of the visual data to meet various network and application requirements [1]. Research is ongoing in developing efficient, secure, and less complex encoding algorithms that exploit intra-sensor and inter-sensor correlations in WVSN data, thereby minimizing the demand for network resources and increasing network lifetime [2]–[4].

Video signals often exhibit sparsity in spatial and temporal domains, and compressive sensing (CS) leverages this inherent sparsity during the acquisition and reconstruction of video signals, leading to more efficient energy utilization. In addition, the lightweight encoder associated with CS, which minimizes computational complexity and resource requirements, makes it preferred over other traditional compression schemes in WVSN applications [5]. The adoption of the block compressive sensing (BCS) technique, in which the sensing of each frame is performed in a block-wise manner, is proposed as a further refinement to reduce computational complexity and memory requirements [6], [7].

This paper presents a comparative study of various BCS-based encoding schemes for both monoview and multiview WVSN scenarios. The comparative analysis is based on compression level, computational complexity, relative gain in encoder energy, and reconstruction quality. A new BCS-based joint encoding scheme for multiview configuration is also proposed. The subsequent sections of this paper are structured in the following manner. Section 2 gives a basic overview of visual configurations and the BCS technique. Different BCS-based encoding schemes for

[1,2]The authors are with the Research Scholar, Department of ECE, Indian Institute of Technology Roorkee, India, E-mail: priya_g@ec.iitr.ac.in and debashis.ghosh@ece.iitr.ac.in

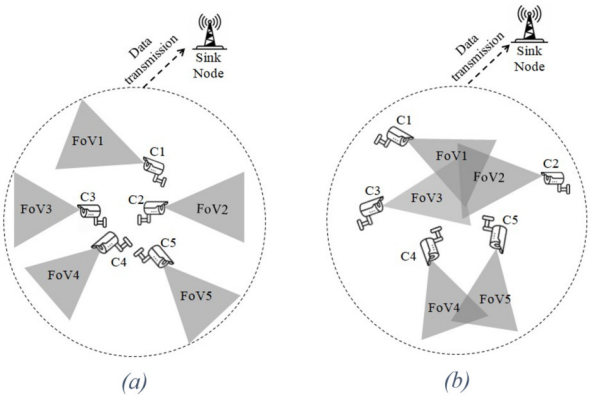

Fig.1: *Visual configurations (a) Monoview (b) Multiview.*

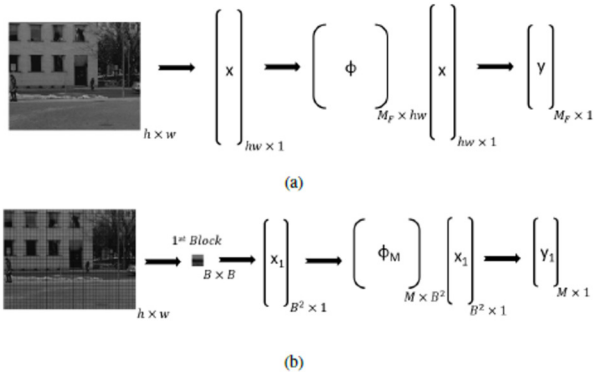

Fig.2: *Illustration of (a) Traditional compressive sensing (b) Block compressive sensing.*

monoview configuration are detailed in Section 3, and that of multiview configuration in Section 4. Subsequently, Section 5 showcases simulation results and a comparative analysis. Our concluding remarks are then presented in Section 6.

2. PRELIMINARIES

2.1 Visual Configurations

WVSNs consist of intelligent cameras serving as sensor nodes that capture images from a designated area and a sink node that collects the data from these sensor nodes for further processing. Basically, two visual configurations exist in WVSNs, namely, monoview and multiview. In a monoview configuration, as depicted in Fig. 1(a), each sensor node within the WVSN captures visual information from a distinct single viewpoint. Monoview setups are characterized by simplicity in terms of hardware and data processing requirements. In a multiview environment, the WVSN incorporates multiple visual perspectives by deploying cameras with overlapped fields of vision (FoVs) across the monitored area. A multiview configuration is illustrated in Fig. 1(b) where sensors C1, C2, and C3 with overlapped FoVs capture views of a target area, and the same is the case with sensors C4 and C5. The choice between monoview and multiview environments in WVSNs depends on the specific application requirements, cost considerations, and the level of detail and redundancy needed for effective monitoring and analysis in the given context.

2.2 An Overview of Block Compressive Sensing

Compressive Sensing (CS) is a signal processing technique that allows for the efficient acquisition and reconstruction of sparse signals using significantly fewer samples than traditional methods require [8]. In CS, the signal is acquired in a compressed form through a linear measurement process, often represented as a matrix-vector multiplication. The reconstruction is performed by solving an underdetermined system of equations, exploiting the sparse nature of the signal.

Consider frame $\boldsymbol{X}$ in a video sequence. In BCS, an image is divided into non-overlapping blocks, and compressive measurements are taken for each block separately [8]. Consider the vector $\boldsymbol{x}_i \in R^N$ representing the vectorized i^{th} block with a block size of $B \times B = N$. The corresponding M-dimensional measurement vector ($M \ll N$) obtained using measurement basis matrix $\boldsymbol{\phi}_M$ of dimension $M \times N$ is expressed as

$$\boldsymbol{y}_i = \boldsymbol{\phi}_M \boldsymbol{x}_i \tag{1}$$

The recovery of the image block vector $\boldsymbol{x}_i$ from a reduced set of measurement samples $\boldsymbol{y}_i$ is possible when $\boldsymbol{x}_i$ exhibits sparsity in a specific domain with an orthonormal sparsity basis $\boldsymbol{\psi}$ of dimensions $N \times N$. This is expressed as:

$$\boldsymbol{x}_i = \boldsymbol{\psi} \boldsymbol{s}_i \tag{2}$$

where $\boldsymbol{S}_i$is the sparse data vector corresponding to $\boldsymbol{x}_i$.
Therefore, (1) may be written as

$$\boldsymbol{y}_i = \boldsymbol{\phi}_m \boldsymbol{x}_i = \boldsymbol{\phi}_M \boldsymbol{\psi} \boldsymbol{s}_i = \boldsymbol{A} \boldsymbol{s}_i \tag{3}$$

where $\boldsymbol{A}$ is the $M \times N$ dictionary matrix. To reconstruct the image block vector $\boldsymbol{x}_i$ from the measurement vector, obtain the $\boldsymbol{s}_i$ vector using (3) while incorporating sparsity constraints. Subsequently, determine $\boldsymbol{x}_i$ using (2). Various CS reconstruction algorithms are Orthogonal Matching Pursuit (OMP) [9], Simultaneous Orthogonal Matching Pursuit (SOMP) [10], CoSaMP [11], Gradient Projection for Sparse Reconstruction (GPSR) [12], Smooth Projected Landweber (SPL) [13], etc.

The advantage of using BCS over traditional CS is illustrated in Fig. 2. In traditional CS, an entire frame of dimension $h \times w$ is vectorized to form $\boldsymbol{x} \in$

Table 1: *Comprehensive Analysis of Monoview Uniform Encoding Schemes.*

Research Works	Encoding Scheme	Decoding Scheme	Correlation	Reconstruction Quality	Compression	Encoder Complexity	Trans. Cost	Remarks
Mun et al. [13]	Uniform BCS	BCS_SPL (IFR)	Spatial	Good	Less	Less	High	At lower bit rates, reconstruction quality is less.
S. Mun [14]	Uniform BCS	MC_BCS_SPL	Spatio-temp.	Moderate	Moderate	Less	Moderate	Motion estimation/compensation is employed at the decoder.
Azghani et al. [15]	Uniform BCS	MH-Tikhonov	Spatio-temp.	Moderate	Moderate	Less	Moderate	Multihypothesis strategy is incorporated at the decoder.
Chen et al. [16]	Uniform BCS	MH-RTIK	Spatio-temp.	Moderate	Moderate	Less	Moderate	Robust to imprecise hypotheses.
Banerjee et al. [17]	Partial DCT + Uniform BCS	OMP + IDCT	Spatial	Moderate	Moderate	High	Moderate	Performing DCT prior to CS necessitates the availability of input frame in spatial domain.
Hao-quan et al. [18]]	Uniform BCS	Discriminative K-SVD	Spatio-temp.	Moderate	Moderate	Less	Moderate	Performance analysis is limited.
Nezhad et al. [19]	Uniform BCS	Deep GAN	Spatio-temp.	Good	Moderate	Less	Moderate	
Zhong et al. [20]	Uniform BCS	LLR-VCSNet	Spatio-temp.	Moderate	Moderate	Less	Moderate	
Gu et al. [21]	Uniform BCS	Temporal shift rec. net	Spatio-temp.	Moderate	Less	Less	High	Requires data to train the netwo
Shi et al. [22]	Uniform BCS	VCSNet	Spatio-temp.	Moderate	Moderate	Less	Moderate	
Wei et al. [23]	Uniform BCS	STM-Net	Spatio-temp.	Good	Moderate	Less	Moderate	
Yang et al. [24]	Uniform BCS	MCNet	Spatio-temp.	Moderate	Moderate	Less	Moderate	
Pei et al. [25]	Uniform BCS	FCN+DCT	Spatio-temp.	Moderate	Less	Less	High	

$R^{hw\times 1}$ and is then sensed using $\phi \in R^{M_F\times hw}$, where M_F is the required number of measurements corresponding to a frame. Whereas, in the case of BCS, individual image block is sensed using $\phi \in R^{M\times B^2}$, the dimension of which is much less than that required in traditional CS. Hence, block-wise sending reduces the computational complexity and memory requirement compared to the conventional CS approach.

3. MONOVIEW ENCODING SCHEMES

3.1 Uniform Sampling

The encoders based on the conventional BCS technique detailed in Section 2.2, which use a fixed sampling rate to sense all the image blocks, form the foundation of most of the compression algorithms used for WVSN. Computational simplicity is the principal advantage of this uniform sampling scheme. Many researchers have developed efficient reconstruction algorithms that leverage the spatiotemporal correlations in video frames to jointly reconstruct uniformly sampled measurements. Prominent algorithms among those are BCS_SPL [13], in which the Smooth Projected Landweber algorithm is employed for the reconstruction of BCS measurements, MC_BCS_SPL [14], in which motion estimation/ compensation is integrated into BCS-SPL, and algorithms in which multihypothesis (MH) strategy is incorporated with SPL for reconstruction [15], [16]. In [17], Banerjee et al. proposed a uniform sampling scheme in which partial Discrete Cosine Transform (DCT) is evaluated before BCS. They used orthogonal matching pursuit followed by inverse DCT for image reconstruction. The computation of DCT before BCS demands the availability of an input frame in the spatial domain at the encoder. The reconstruction scheme based on discriminative KSVD, as proposed in [18], would benefit from a more thorough study and detailed performance analysis. Another class of CS reconstruction algorithms generally employed nowadays is based on deep learning techniques [19]-[25]. The lack of training data in the case of WVSN systems limits the use of those algorithms in WVSN systems. Among the existing state-of-the-art algorithms, SPL reconstruction stands out for its superior quality. It eliminates blocking artifacts using a Wiener filter and also ensures fast reconstruction. A comprehensive analysis of all monoview uniform encoding schemes, highlighting their pros and cons, is presented in Table 1.

3.2 Adaptive Sampling

When an image is subdivided into smaller blocks, blocks with minimal information can be reconstructed with fewer observations than those with more details. Consequently, there is potential to enhance BCS reconstruction performance by adaptively assigning different sampling rates to distinct blocks. The related works emphasizing their advantages and disadvantages are listed in Table 2. In [26], the adaptive measurement rate is assigned based on the spar-

Table 2: *Comprehensive Analysis of Monoview Adaptive Encoding Schemes.*

Research Works	Encoding Scheme	Adaptive Rate Param.	Correlation	Compression	Encoder Comp.	Trans. Cost	Remarks
Zhang et al. [26]	Adaptive BCS	Sparseness in DCT	Spatial	Good	High	Less	Transforming to DCT domain increases the encoder computational complexity.
Wang et al. [27]	Adaptive BCS	Texture info	Spatial	Moderate	Moderate	Moderate	Temporal correlation is not addressed.
Hadizadeh et al. [28]	Adaptive BCS	Texture complexity	Spatio-temp.	Good	High	Less	Encoder computational complexity increases due to adaptive sensing rate.
Zhang et al. [29]	Adaptive BCS	Statistical texture	Spatial	Moderate	High	Moderate	High encoder complexity due to statistical texture analysis.
Li et al. [30]	Adaptive BCS	Spatial entropy	Spatial	Moderate	High	Moderate	Impractical because spatial domain data is not available in a CS framework.
Yang et al. [32]	Adaptive BCS	Temporal correlativity	Spatio-temp.	Moderate	Moderate	Moderate	
Yang et al. [33]	Adaptive BCS	ROI detection	Spatio-temp.	Moderate	Moderate	Moderate	Additional computations are involved at encoder for ROI detection.
Zhao et al. [34]	Adaptive BCS	Channel state	Spatio-temp.	Good	High	Less	Compression rate and reconstruction quality are low for videos with fast-moving objects.
Zhang et al. [35]	Adaptive BCS	Std. deviation	Spatial	Good	Less	Less	Relatively less computational complexity.

sity of each image block in the Discrete Cosine Transform (DCT) domain. Another approach in [27]–[29] proposes adaptive image BCS algorithms based on texture information. The algorithm requires image reconstruction during sampling to achieve adaptive sampling rate allocation, which increases computational complexity on the encoding side. In [30], Li et al. proposed adaptive sampling rate allocation based on spatial entropy at the encoder and a linear reconstruction strategy at the decoder to reduce decoder complexity. In [31], an image CS algorithm is introduced, utilizing an adaptive learning sparse matrix. While the algorithm achieves high image reconstruction quality, the formation of its sparse groups demands extensive training, thereby limiting its practical deployment. Other similar features used to evaluate the adaptive sensing rate include temporal correlativity [32], region of interest (ROI) [33], channel state [34] etc.

An adaptive sampling scheme with relatively less computational overhead is proposed in [35], which relies on standard deviations of image blocks to assign an adaptive sampling rate. Standard deviation acts as a statistical measure to identify whether the image block is a smooth or non-smooth. A non-smooth block contains more texture information than smooth blocks, requiring more measurements for effective reconstruction. Adaptive rates are assigned based on the percentage standard deviation. Consider a frame of size $h \times w$, subjected to BCS-based adaptive sampling with block size $B \times B$ and sensing rate R. Then, the overall sampling frequency is

$$M_{overall} = Rhw \tag{4}$$

A fixed sampling frequency determined by (5) is allotted to every image block to ensure a minimum quality while reconstructing the image.

$$M = \frac{W \times R \times M_{overall}}{\text{No. of image blocks}} = WR^2B^2 \tag{5}$$

where $W(0 \leq W \leq 1)$ is the fixed sampling frequency distribution parameter and No.of image blocks $= \frac{hw}{B^2}$. Block-wise sampling frequency allocation is performed by evaluating the percentage standard deviation and sampling frequency using (6) and (7).

$$P_{\sigma_i} = \frac{\sigma_{x_i}}{\sum_{i=1}^{n} \sigma_{x_i}} \tag{6}$$

where σ_{x_i} is the block-wise standard deviation.

$$\begin{aligned} M_i &= P_{\sigma_i}(M_{overall} - \text{No. of image blocks} \times M) \\ &= P_{\sigma_i}(R \times B^2 - M)\frac{hw}{B^2} \end{aligned} \tag{7}$$

An upper limit of $0.4 \times B^2$ is set for the adaptive sampling frequency to ensure it remains lower than the total number of pixels in an image block. If the block-wise sampling frequency M_i exceeds the upper limit, then the cumulative excessive frequency is calculated as per (8), and its average is assigned to all blocks.

$$M_{excess} = M_{excess} + (M_i - 0.4 \times B^2) \tag{8}$$

Once the block-wise frequency assignment is completed, the adaptive measurements are evaluated and concatenated with fixed measurements to form the fi-

Table 3: *Comprehensive Analysis of Multiview Encoding Schemes.*

Research Works	Encoding Scheme	Decoding Scheme	Correlation	Reconstruction Quality	Compression	Encoder Complexity	Trans. Cost	Remarks
Ebrahim et al. [37]	Uniform	Multi-phase joint reconstruction	Intra and Inter-sensor	Moderate	Less	Less	High	Exploits inter-view and spatio-temporal correlations.
Liu et al. [38]	Uniform	Disparity comp. total variation minim. joint decoder	Spatial, Inter-sensor	Good	Less	Less	High	Better reconstruction quality.
Cen et al. [39]	Uniform	Inter-view motion comp. + GPSR	Spatial, Inter-sensor	Moderate	Less	Less	High	Blind video quality estimation is proposed.
Zhu et al. [40]	Uniform	Joint optimization using ARTV and MNLRT	Spatial, Inter-sensor	Good	Less	Less	High	Better reconstruction at low sensing rates.
Liu et al. [41]	Uniform	Disparity and motion comp. total variation minim.	Intra and Inter-sensor	Good	Less	Less	High	Recon. scheme is robust to corrupted data.
Fei et al. [42]	Uniform	Joint recon. based on spatial correlation and low-rank background constraints	Spatial, Inter-sensor	Moderate	Less	Less	High	Reduces the effect of noise.
Song et al. [43]	Uniform	Local and nonlocal constraint-based CS recovery model	Intra and Inter-sensor	Moderate	Less	Less	High	Less decoding time.
Yang et al. [33]	Adaptive	Region of interest detection	Spatial, Inter-sensor	Good	Moderate	Moderate	Moderate	Additional computations at encoder for ROI detection.
Cen et al. [44]	Adaptive	Independent LASSO	Spatial, Inter-sensor	Good	Less	Less	High	Adaptive rate estimation is performed by a centralized controller at the decoder end.

nal measurements, which are used to reconstruct the image in the sink node.

3.3 Joint Encoding

Given that the energy needed to transmit a bit is higher than the energy required for computation, it's important to reduce the encoder's computational complexity. Additionally, emphasis should be placed on enhancing the compression offered by various encoding schemes. Although the adaptive sampling scheme improves reconstruction quality, it still offers the same level of compression as the uniform sampling scheme. A joint encoding scheme proposed in [36] overcomes this limitation by performing a second-level compression in the measurement domain thereby ensuring a better compression. The better compression level is achieved by exploiting the correlation between measurement vectors of successive frames on account of intra-sensor temporal correlation.

In this method, joint encoding is performed by considering G consecutive frames of a video $\{\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots\ldots, \boldsymbol{X}_G\}$ as a group of pictures (GOP). The first frame of every GOP, i.e., $\boldsymbol{X}_1, \boldsymbol{X}_{G+1}, \boldsymbol{X}_{2G+1}, \ldots\ldots$is taken as the key frame (KF) and the remaining frames as non-key frames (NKFs). The first compression phase is BCS with block size B×B, performed independently on every frame. In the second phase of compression, the measurement vectors corresponding to the NKFs are subjected to joint encoding with reference to the corresponding KF measurement, i.e., $\boldsymbol{Y}_1, \boldsymbol{Y}_{G+1}, \boldsymbol{Y}_{2G+1}, \ldots\ldots$ as expressed below.

$$\boldsymbol{y}_{comp_i}^{(k)} = \boldsymbol{y}_i^{(j)} \quad \text{for} \quad \begin{cases} i = 2 : (\text{centre index}) \\ \text{corr}(\boldsymbol{y}_i^{(j)}, \boldsymbol{y}_1^{(j)}) < T \\ i = (\text{centre index}+1) : (G) \\ \text{corr}(\boldsymbol{y}_i^{(j)}, \boldsymbol{y}_{G+1}^{(j)}) < T \end{cases}$$

where average correlation is set as the threshold T. Finally, the compressed measurement vectors, along with an index vector containing the indices of the least correlated blocks, are transmitted to the sink node.

4. MULTIVIEW ENCODING SCHEMES

4.1 Uniform Sampling

Up to now, most BCS-based encoders designed for multiview data streaming in WVSNs have relied on conventional uniform block sensing, as discussed in Section 2.2. Inter-view correlation is exploited in the decoder, and researchers have developed many algorithms for efficient reconstruction of uniform sampled multiview data. A multi-phase joint reconstruction scheme exploiting intra and inter-sensor correlations is proposed in [37]. In [38], Liu et al. proposed a total variation (TV) minimization joint reconstruction scheme based on disparity compensation (DC), which resulted in improved reconstruction quality. In [39], a gradient projection for sparse reconstruction (GPSR) based joint decoder is proposed that relies on side information generated through inter-view motion compensation. A joint optimization model (JOM) exploiting intra and inter-view correlation is proposed in [40], which jointly optimizes adaptive disparity

compensated residual total variation (ARTV) and a multi-image non-local low-rank tensor (MNLRT). Research paper [41] proposes a joint reconstruction approach based on disparity and motion-compensated total variation (DC/MC-TV) minimization. To an extent, the proposed scheme is robust to corrupted data. Furthermore, [42] introduces a novel CS joint reconstruction method for multiview images, guided by spatial correlation and low-rank background constraints, which helped to reduce noise in the reconstructed image. Another joint reconstruction scheme that could provide a fast convergence is proposed by Song *et al.* in [43], which relied on local and non-local constraints for CS reconstruction. An analysis of the advantages and disadvantages of these related works is presented in Table 3.

4.2 Adaptive Sampling

In an image, different blocks have varying levels of sparsity. Allocating sensing rates at the block level, based on a given frame-level sensing rate, helps improve reconstruction quality. One such adaptive sampling scheme proposed in [33] performs rate allocation based on region of interest (ROI) detection. Since the encoder carries out ROI detection, it increases the encoder's complexity. The distinctive BCS-based encoder system introduced in [44] represents a novel approach for multiview scenarios, incorporating a variable block sensing rate. In their work, Cen *et al.* [44] presented an encoding and decoding framework for multiview video systems using BCS comprising a collaborative sparsity-aware block-level rate-adaptive encoder, feedback channel, independent decoder, and a centralized controller located at the decoder.

The key element at the receiver is the centralized controller, primarily responsible for choosing the reference view (R-view) and assessing the necessary sparsity and adaptive sampling frequency to be transmitted back to the transmitter. Shifting the process of adaptive rate estimation to the decoder side ensures that the encoder is not burdened with additional computational complexity. The controller evaluates the Pearson correlation coefficient γ_{ij} among the v views, and the view with the highest coefficient γ^* is chosen as the R-view, as expressed below.

$$\begin{aligned} \gamma_{ij} &= \mathrm{corr}(y_i, y_j) \quad for \quad \forall i \neq j,\ i,j = 1,\ldots,v \\ \gamma^* &= \underset{i=1,\ldots,v}{\mathrm{argmax}}\, \tilde{\gamma}_i \end{aligned} \tag{9}$$

where $\tilde{\gamma}_i \triangleq \frac{1}{v-1}\sum_{i\neq j}\gamma_{ij}$ denotes the average Pearson coefficient of the $\boldsymbol{i}^{th}$ view.

With the aid of a BCS reconstruction algorithm, a sparse representation $\mathcal{S}_{R^*}$ of R-view is obtained. From $\mathcal{S}_{R^*}$, the centralized controller assesses the sparsity and required adaptive sampling rate. Sparsity $\mathcal{K}$ corresponds to the count of non-zero elements in $\mathcal{S}_{R^*}$. However, natural images are typically not entirely sparse in the transform domain. Under the assumption that $\rho_s hw$ measurements (where ρ_s is a pre-determined empirical percentile) are sufficient to recover a view of size $h \times w$, a threshold T is identified using (10) to make $\mathcal{S}_{R^*}$ precisely sparse.

$$\| \max(|\mathcal{S}_{R^*}| - T, 0)\|_0 = \rho_s hw \tag{10}$$

where $\|\cdot\|$ corresponds to L0 norm, which is a count of non-zero elements. An exhaustive search algorithm is employed to solve this L0 norm problem to estimate T.

Then, block sparsity $\mathcal{K}^i$ is estimated as in (11) by applying T to every $B \times B$ block of $\mathcal{S}_{R^*}$ followed by evaluation of the number of adaptive measurements using (12).

$$\mathcal{K}^i = \| \max(|\mathcal{S}^i_{R^*}| - T, 0)\|_0 \tag{11}$$

$$M_i = \frac{\mathcal{K}^i \log_{10}\left(\frac{B^2}{\mathcal{K}^i}\right)}{\sum_{i=1}^{\frac{hw}{B^2}} \mathcal{K}^i \log_{10}\left(\frac{B^2}{\mathcal{K}^i}\right)} hwR \tag{12}$$

where R is the frame sensing rate.

Subsequently, mean value subtraction is performed at the encoder to further increase compression and reduce bandwidth demand. The required mean value is estimated by the centralized controller from the reconstructed R-view $\hat{\boldsymbol{X}}_{\boldsymbol{R}}$ as

$$\mu = \frac{1}{hw}\sum_{i=1}^{hw} \hat{\boldsymbol{X}}_{\boldsymbol{R}}(i) \tag{13}$$

The adaptive measurement frequency and the estimated mean value are transmitted to the encoder via a feedback channel to aid adaptive sensing of subsequent frames. Based on the information received from the centralized controller, the encoder initiates block-wise adaptive sensing of subsequent frames as

$$\boldsymbol{y}^i_v = \boldsymbol{\phi}^i_M \boldsymbol{x}^i_v \tag{14}$$

where $\boldsymbol{y}^{\boldsymbol{i}}_{\boldsymbol{v}} \in R^{M_i \times 1}$ is the measurement vector corresponding to the i^{th} block, and $\boldsymbol{\phi}^i_M \in R^{M_i \times B^2}$ is the corresponding sensing matrix.

Thereafter, mean value reduction is performed in the measurement domain for further compression. For that, first, a vector $\boldsymbol{\mu} \in R^{B^2 \times 1}$ in which every element is the same as μ is constructed, and then it is sampled using $\boldsymbol{\phi}^i_M$. The sampled version of $\boldsymbol{\mu}$ vector is then subtracted from $\boldsymbol{y}^i_v$ to generate the final measurement vector to be transmitted to the decoder as in (15).

$$\tilde{\boldsymbol{y}}^i_v = \boldsymbol{y}^i_v - \boldsymbol{\phi}^i_M \boldsymbol{\mu} \tag{15}$$

Table 4: *Frame-wise CR of Monoview Encoding Schemes.*

Frame #	Uniform/ Adaptive	Joint (CR)				
	All Datasets	Cont	MD	News	FB	CG
1	3.32	3.32	3.32	3.32	3.32	3.32
2	3.32	15.49	18.29	17.32	16.25	16.67
3	3.32	20.25	22.31	28.01	9.4	16.67
4	3.32	21.24	21.58	27.43	9.54	17.79
5	3.32	20.9	20.25	24.84	8.03	14.96
6	3.32	20.9	21.24	25.82	8.84	15.86
7	3.32	20.57	23.51	28.01	9.02	17.1
8	3.32	19.95	14.79	18.04	11.76	15.31
9	3.32	3.32	3.32	3.32	3.32	3.32
10	3.32	14.79	18.54	17.32	19.95	16.88
11	3.32	20.9	19.08	26.33	9.61	16.25
12	3.32	21.24	16.67	23.51	8.9	17.32
13	3.32	23.1	18.29	21.58	8.72	16.25
14	3.32	22.7	23.1	22.7	9.61	16.67
15	3.32	22.7	23.94	23.51	9.27	16.88
16	3.32	21.94	18.04	21.58	11.76	16.25
17	3.32	3.32	3.32	3.32	3.32	3.32
18	3.32	18.29	15.67	18.04	19.08	14.63
19	3.32	23.51	17.55	16.67	9.68	14.96
20	3.32	23.94	18.04	22.7	9.61	14.47
21	3.32	25.82	21.94	22.31	8.78	16.06
22	3.32	23.94	21.58	22.7	9.34	16.25
23	3.32	23.1	23.1	21.94	9.54	16.25
24	3.32	23.51	23.94	23.94	12.42	16.25

Table 5: *Frame-wise CR of Multiview Encoding Schemes.*

Frame #	Uniform/Adaptive	Joint (CR)		
	All Datasets	Ballroom	Vassar	Exit
1	3.32	3.32	3.32	3.32
2	3.32	24.01	20.96	12.74
3	3.32	21.84	23.26	15.72
4	3.32	19.65	24.16	18
5	3.32	21.41	24.5	19.51
6	3.32	23.75	23.61	17.52
7	3.32	26.48	20.51	13.13
8	3.32	3.32	3.32	3.32
9	3.32	3.32	3.32	3.32
10	3.32	25.2	19.78	13.88
11	3.32	22.97	22.35	18.4
12	3.32	20.73	23.35	20.17
13	3.32	17.88	21.62	20.2
14	3.32	19.03	20.15	18.64
15	3.32	20.07	18.2	14.02
16	3.32	3.32	3.32	3.32
17	3.32	3.32	3.32	3.32
18	3.32	19	17.69	14.02
19	3.32	18.09	19.21	19.35
20	3.32	17.6	20.22	19.6
21	3.32	17.18	18.39	17.97
22	3.32	17.59	17.15	15.79
23	3.32	19.56	16.37	12.19
24	3.32	3.32	3.32	3.32

4.3 Joint Encoding

1) Problem Formulation: Determining the adaptive sampling frequency in [44] occurs at the sink node, necessitating a feedback channel from the sink node to the source node. The total compression achieved by the adaptive scheme for an image frame remains equal to that of a uniform sampling scheme. Hence, it does not reduce transmission cost or energy gain at the encoder. Additionally, the encoding process does not account for intra-sensor temporal correlation. These limitations can be addressed by implementing a joint encoding scheme for multiview data compression that takes advantage of both intra-sensor and inter-sensor correlations.

2) Proposed Method: In our proposed scheme, the correlations are addressed in two phases of compression. Phase-I is performed at individual sensor nodes, which involves addressing the intra-sensor spatial correlation by sensing the frames using the BCS technique as detailed in Section 2.2. Phase-II compression occurs at a cluster head (CH) or in a different sensor in the multi-hop path. Intra-sensor temporal and inter-sensor spatial correlations are addressed by performing joint encoding of measurements of multiview frames. The first step of Phase-II involves interleaving the measurement vectors of views corresponding to a particular frame to create an interleaved matrix $\mathcal{J}$. Interleaving operation is performed so that the adjacent columns of the matrix created have a high correlation. $\{\mathcal{J}_1, \mathcal{J}_2, \ldots\ldots, \mathcal{J}_G\}$ forms the interleaved matrices corresponding to a GOP of size G. While we consider the j^{th} column of i^{th} interleaved matrix $\mathcal{J}_i$, it may possess a significant correlation with the j^{th} column of $\mathcal{J}_{i-1}$, which is on account of intra-sensor temporal correlation. Also, $\mathcal{J}_i(:, j)$ may hold substantial similarity with $\mathcal{J}_{i-1}(:, j \pm 1)$ due to inter-sensor correlation. These correlations are taken advantage of to compress the data further.

In a GOP, considering $\mathcal{J}_1$ and $\mathcal{J}_G$ as the reference matrices, compression is performed on the remaining matrices as detailed. As the first step, evaluate intra-sensor and inter-sensor correlation measures denoted as γ_{intra} and γ_{inter} respectively.

$$\gamma_{\text{intra}} = \text{corr}(\mathcal{J}_i(:,j), \mathcal{J}_{\text{ref}}(:,j)) \forall j \in 1:C \quad (16)$$

$$\gamma_{\text{inter}} = \text{corr}(\mathcal{J}_i(:,j), \mathcal{J}_{\text{ref}}(:,j \pm 1)) \forall j \in 1:C \quad (17)$$

where $\mathcal{J}_{\text{ref}} = \begin{cases} \mathcal{J}_1 & \text{if } 2 \leq i \leq \frac{G}{2} \\ \mathcal{J}_G & \text{if } \frac{G}{2}+1 \leq i \leq G-1 \end{cases}$
and C denoted the number of columns in $\mathcal{J}$.
Based on the intra and inter-sensor correlation measures, a threshold T is evaluated as

$$T = \gamma_{\text{avg}} + \alpha(\gamma_{\text{max}} - \gamma_{\text{avg}}) \quad (18)$$

where α is the compression level control parameter,γ_{avg} = Average(γ_{intra}, γ_{inter}) and γ_{max}= Maximum(γ_{intra}, γ_{inter}). Now, based on the threshold T, compression is carried out as

$$\mathcal{J}_{\text{comp}_i} = \mathcal{J}_i(\gamma_i < T) \quad (19)$$

where $\gamma_i = \{\gamma_{\text{intra}}, \gamma_{\text{inter}}\}$

Table 6: *Storage Requirement of Monoview Encoders for GOP Size=8.*

Encoding Scheme	Cont	MD	News	FB	CG
Uniform	2.06	2.06	2.06	2.06	2.06
Adaptive	2.11	2.11	2.11	2.11	2.11
Joint	0.68	0.66	0.64	0.95	0.75

Table 7: *Storage Requirement of Multiview Encoders for GOP Size=8, Views=6.*

Encoding Scheme	Ballroom	Vassar	Exit
Uniform/Adaptive	34.8	34.8	34.8
Joint	12.71	12.56	14.26

Finally, for each GOP, the set of compressed measurements $\{\mathcal{J}_1, \mathcal{J}_{\text{comp}_2}, \mathcal{J}_{\text{comp}_3}, \ldots\ldots, \mathcal{J}_G\}$, along with a vector containing the indices of strongly correlated vectors, is transmitted to the decoder. The suggested approach does not require a feedback channel. Experimental results in Section 5 reveal that it's superior compression and reduced transmission costs, compared to other methods, lead to efficient energy usage at the encoder, making the proposed method well-suited for WVSN.

5. SIMULATION RESULTS AND COMPARISON

The appropriateness of an encoding scheme in a WVSN scenario mainly depends on two factors: (i) the compression level and (ii) the encoder computational complexity. In subsequent subsections, these two factors are thoroughly analyzed for different encoding schemes employed in monoview and multiview visual configurations. All the encoding schemes under consideration are BCS-based with block size set to 16 × 16 using random Gaussian matrix as sensing matrix and sensing rate $R = 0.3$. Simulations of monoview visual configuration were conducted on 24 frames each of size 288 × 352 extracted from five standard videos, namely Container (Cont), Mother-Daughter (MD), News, Football (FB), and Coastguard (CG). In the case of multiview configuration, simulations were conducted on six views and 24 frames with dimensions 480×640 of each view extracted from three standard multiview datasets, namely Ballroom, Vassar, and Exit. A comparative analysis of the reconstruction quality is also performed and presented in Section 5.4.

5.1 Compression Ratio and Storage Requirement

The compression level offered by different schemes is compared based on the compression ratio (CR), evaluated as the size ratio of an uncompressed image to that of a compressed one. Table 4 and 5 show the CR values of monoview and multiview configurations for 24 frames of all datasets under consideration. It is observed that uniform and adaptive encoding schemes yield a low, fixed CR value that depends solely on the sensing rate. The joint encoding scheme gave a much better CR, and in addition to sensing rate, CR depends on correlations existing in video frames. The encoder storage space demanded by different algorithms under consideration is compared in Table 6 and 7. Results reveal that the joint encoding scheme demands less storage space compared to others.

Fig. 3(a) and Fig. 4(a) present the frame-wise CR variations averaged across all datasets. In joint encoding, Phase-II compression is performed on NKFs alone, keeping KFs as reference, and hence, NKFs have higher CR. Variations of CR for GOP size are also analyzed and shown in Fig. 3(b) and Fig. 4(b). In both visual configurations, GOP size does not impact CR values of uniform/adaptive encoding. In the case of the joint encoding scheme, as GOP size increases, CR also increases. The behaviour of compression for various sensing rates is depicted in Fig. 3(c) and Fig. 4(c). In all encoding schemes, CR decreases with an increase in sensing rate, but CR of joint encoding scheme dominates other schemes. The relatively high CR offered by the joint encoding scheme for higher GOPs and lower sensing rates makes it the right choice for WVSN applications.

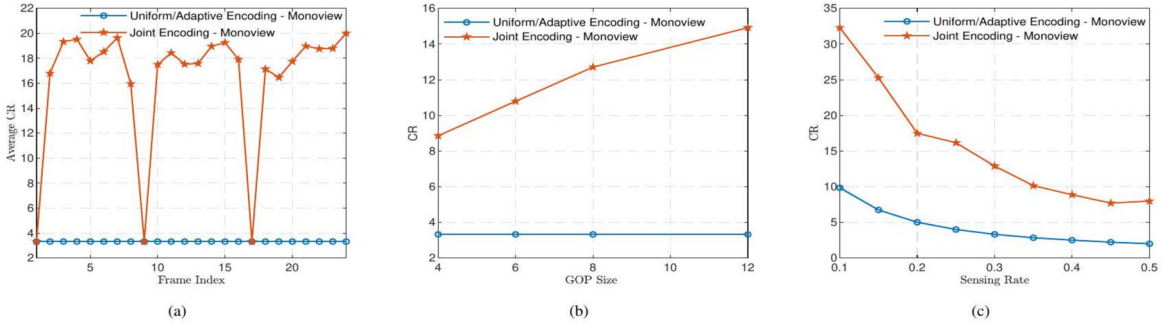

Fig.3: *CR variations in monoview encoding schemes (a) Average CR against frame index (b) CR vs GOP size (c) CR vs sensing rate.*

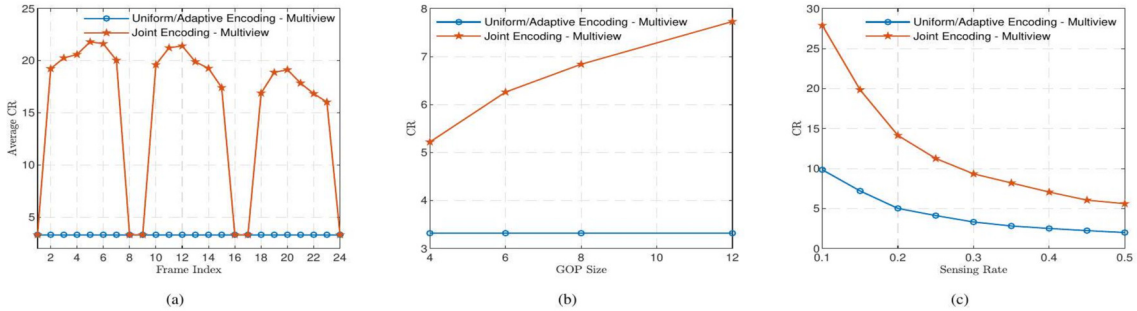

Fig.4: *CR variations in multiview encoding schemes (a) Average CR against frame index (b) CR vs GOP size (c) CR vs sensing rate.*

5.2 Computational Complexity

The computation complexities summarized in Table 8 and Table 9 are estimated by considering the number of views as v, frame size as $h \times w$, block size as $B \times B$, number of fixed measurements as M, and number of adaptive measurements as M_i. Among the three encoding schemes, the uniform sampling scheme always offers the least complexity encoder, and adaptive/joint encoding schemes involve additional computations. However, in the case of the multiview adaptive encoder proposed in [44], the additional computations are pushed to the decoder with the demand of a feedback channel to convey the adaptive measurements required by the encoder. Even though there is an increase in demand for computational energy, there is a relative gain in energy at sensor nodes due to the high CR offered by the joint encoding scheme, which is deduced in Section 5.3.

5.3 Analysis of Gain in Energy

The relative gain in energy at the sensor node for both visual configurations is evaluated based on the analysis proposed in [36]. The same for our experimental setup with a GOP size G, frame size $h \times w$, frame sensing rate R, and block size $B \times B$ is deduced in subsequent sections assuming $N_{b_i} = 8$ bits are used to encode each measurement, E_T is the average transmission energy required per bit, and E_C is the average computation energy per computation.

1) Gain in Energy - Monoview Encoding Schemes: When comparing uniform and adaptive encoding schemes, both provide the same level of compression. However, the adaptive scheme requires more energy due to the additional computations involved in adaptive measurement processing. The adaptive scheme provides better reconstruction quality compared to uniform encoding, and if energy should be minimized by compromising on reconstruction quality, then uniform encoding is preferred over adaptive. Now, if we consider the joint encoding scheme, due to high CR, the reduction in transmission energy ($TE_{\text{Redn.}}$) is much higher compared to the additional demand for computational energy ($CE_{\text{Inc.}}$) as deduced below.

$$TE_{\text{Redn.}} = N_{b_i}\left[R - \frac{1}{CR_{\text{Joint}}}\right] GhwE_T \tag{20}$$

$$CE_{\text{Inc.}} = (G-1)hw\left(2R - \frac{1}{B^2}\right)E_C \tag{21}$$

$$TE_{\text{Redn.}} = \mu_1 N_{b_i}\frac{E_T}{E_C} \times CE_{\text{Inc.}} \tag{22}$$

where $\mu_1 = \frac{\left[R - \frac{1}{CR_{\text{Joint}}}\right]G}{\left[2R - \frac{1}{B^2}\right](G-1)}$.

The resulting range of μ_1 is 0.38 to 0.48 in our experimental setup, and average CR variations are ob-

Table 8: *Computational Complexities of Monoview Encoding Schemes.*

Encoding Scheme	Operations	Frame Type	Multiplications	Additions
Uniform Sampling	$\boldsymbol{y_i} = \boldsymbol{\phi_M x_i}$	KF/NKF	$\mathcal{O}(Mhw)$	$\mathcal{O}((\frac{B^2-1}{B^2})Mhw)$
Adaptive Sampling	Percentage standard deviation	KF/NKF	$\mathcal{O}((\frac{4}{B^2}+1)hw)$	$\mathcal{O}((\frac{3B^2-1}{B^2})hw-1)$
	Adaptive sampling frequency		$\mathcal{O}(\frac{hw}{B^2})$	1
	Fixed measurement		$\mathcal{O}(Mhw)$	$\mathcal{O}((\frac{B^2-1}{B^2})Mhw)$
	Adaptive measurement		$\mathcal{O}\left(B^2\sum_{i=1}^{hw/B^2} M_i\right)$	$\mathcal{O}\left((B^2-1)\sum_{i=1}^{hw/B^2} M_i\right)$
Joint Encoding	$\boldsymbol{y_i} = \boldsymbol{\phi_M x_i}$	KF	$\mathcal{O}(Mhw)$	$\mathcal{O}((\frac{B^2-1}{B^2})Mhw)$
	Correlation	NKF	$\mathcal{O}(M \times \frac{hw}{B^2})$	$\mathcal{O}((M-1) \times \frac{hw}{B^2})$

Table 9: *Computational Complexities of Multiview Encoding Schemes.*

Encoding Scheme	Operations	Frame Type	Multiplications	Additions
Uniform Sampling	$\boldsymbol{y_i} = \boldsymbol{\phi_M x_i}$	KF/NKF	$\mathcal{O}(Mhwv)$	$\mathcal{O}((\frac{B^2-1}{B^2})Mhwv)$
Adaptive Sampling	$\boldsymbol{y_i} = \boldsymbol{\phi_{Mi} x_i}$	KF/NKF	$\mathcal{O}\left(B^2v\sum_{i=1}^{hw/B^2} M_i\right)$	$\mathcal{O}\left((B^2-1)v\sum_{i=1}^{hw/B^2} M_i\right)$
Joint Encoding	$\boldsymbol{y_i} = \boldsymbol{\phi_M x_i}$	KF	$\mathcal{O}(Mhwv)$	$\mathcal{O}((\frac{B^2-1}{B^2})Mhwv)$
	Correlation	NKF	$\mathcal{O}(\frac{3Mhwv}{B^2})$	$\mathcal{O}(\frac{3(M-1)hwv}{B^2})$

served in Table 4. Considering the fact that $E_T > E_C$ and $N_{b_i} = 8$, from (22) it is evident that $TE_{\text{Redn.}} > CE_{\text{Inc.}}$.

2) Gain in Energy - Multiview Encoding Schemes: A comparative study of the adaptive encoding scheme for multiview proposed in [44] with uniform encoding shows that both schemes offer the same level of compression. The computational cost remains the same because, in the adaptive scheme, the additional computations are performed at the decoder. Hence, to obtain better reconstruction quality, an adaptive scheme may be preferred over a uniform, but at the cost of the feedback channel. The joint encoding scheme leads to a relative energy gain due to the relatively high CR it offers, which is analyzed below for a set of views from v multiview cameras.

$$TE_{\text{Redn.}} = N_{b_i}\left[1 - \frac{1}{CR_{\text{Joint}}}\right] GRhwvE_T \tag{23}$$

$$CE_{\text{Inc.}} = 3(G-2)hwv\left(2R - \frac{1}{B^2}\right)E_C \tag{24}$$

$$TE_{\text{Redn.}} = \mu_2 N_{b_i}\frac{E_T}{E_C} \times CE_{\text{Inc.}} \tag{25}$$

where $\mu_2 = \frac{\left[R - \frac{1}{CR_{\text{Joint}}}\right]GR}{3\left[2R - \frac{1}{B^2}\right](G-2)} \approx 0.21$. for our experimental setup. From (25), it is clear that $TE_{\text{Redn.}} > CE_{\text{Inc.}}$ since $E_T > E_C$ and $N_{b_i} = 8$ bits.

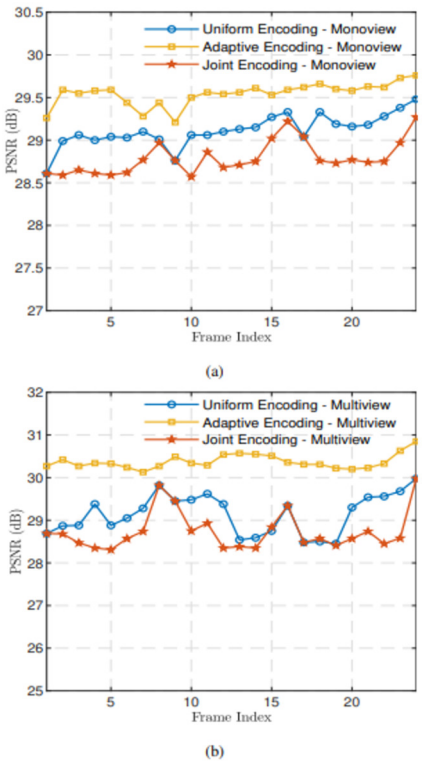

Fig.5: *Comparison of reconstruction quality based on PSNR (a) Average PSNR against frame index - Monoview encoding schemes (b) PSNR against frame index for view 6 of Ballroom sequence - Multiview encoding schemes.*

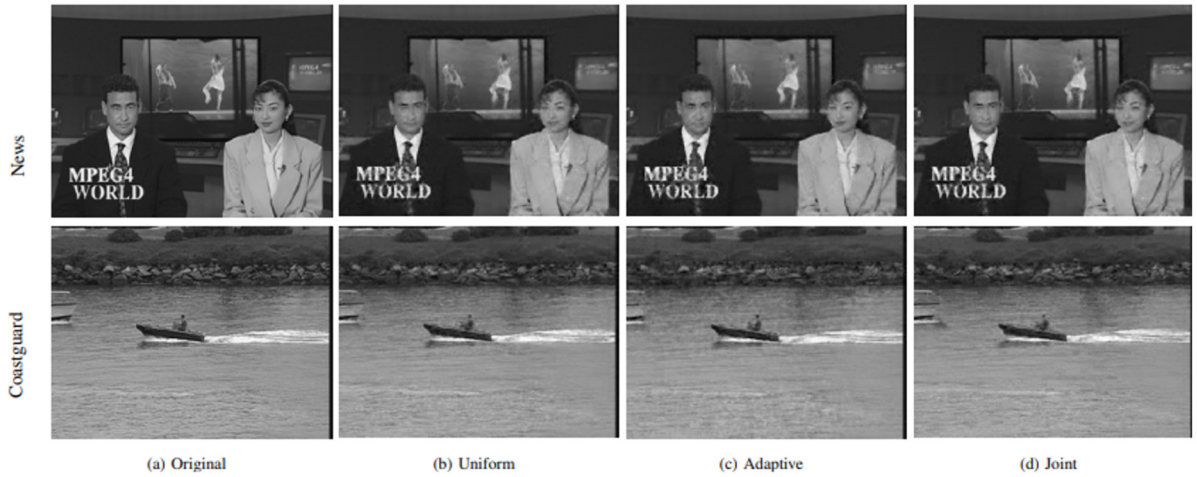

Fig.6: *Original and reconstructed center frames of the first GOP of various sequences for subjective comparison.*

Fig.7: *Original and reconstructed frames of Vassar (Frame 12 View 5), Ballroom (Frame 24 View 3), and Exit (Frame 12 View 1) for subjective comparison.*

5.4 Analysis of Reconstruction Quality

The reconstruction quality achieved from the compressed measurements of various encoding schemes is analyzed based on the peak signal-to-noise ratio (PSNR). To ensure a fair comparison, the same reconstruction scheme is used for all encoders, and it can be any of the prominent CS reconstruction techniques. In the case of monoview configuration, a multihypothesis reconstruction strategy based on Smooth Projected Landweber (MH_BCS_SPL) [15], [16] with a window size of 3 is applied. The choice of window size is based on the fact that a broader search window may lead to the generation of numerous incorrect hypotheses. The significance of SPL-based schemes lie in their enhanced reconstruction quality, which is achieved by removing blocking artifacts through the use of a Wiener filter, along with their capability for rapid reconstruction [13], [45].

A comparative study of reconstruction quality is performed by evaluating frame-wise PSNR for all datasets under consideration. The average PSNR values for monoview encoders are graphically represented in Fig. 5(a). The frame-wise average PSNR measure reveals that the adaptive encoding scheme ensures relatively better reconstruction quality with approximately 0.5dB higher than the uniform encoding. The KFs of the joint encoding scheme (frame

no.: 1, 9 & 17) possess the same PSNR as that of a uniform encoder, whereas the NKFs have a PSNR of roughly 0.5dB less owing to the second level compression performed on NKFs at the joint encoder. The visual quality of the reconstruction can be subjectively evaluated for monoview configuration by examining Fig. 6, which shows the original and reconstructed frame 5 of the News sequence. The reconstruction quality of the joint encoding scheme can be improved further by iterative multihypothesis strategy as proposed in [36].

In the case of multiview configuration, at the decoder end, the signal is assumed to be sparse in the DCT domain, and the Least Absolute Shrinkage and Selection Operator (LASSO) algorithm [46] is utilized for reconstruction, inspired by its low complexity. From the PSNR variation of view 6 of the Ballroom sequence illustrated in Fig. 5(b), it is evident that the adaptive encoding scheme resulted in a relatively better reconstruction. When comparing uniform and joint encoding schemes, the PSNR of NKFs in the joint encoder was observed to be slightly lower than that of the uniform encoder. The visual quality of the reconstruction for multiview configuration may be evaluated by examining Fig. 7, in which the original and reconstructed view 5 frame 12 of the Vassar sequence is shown.

6. CONCLUSION

The energy limitations of intelligent camera sensor nodes in WVSNs necessitate the implementation of efficient data compression algorithms with less encoder complexity and higher compression levels to enhance the lifespan of sensor networks. Video compression schemes based on BCS techniques are widely used in WVSNs because of their lightweight encoders and minimal storage space requirements. A comparative study of various BCS-based encoding schemes in the case of monoview and multiview configurations is carried out in this paper. A joint encoding scheme that provides a much higher compression level is also proposed in this paper. Comparative studies reveal that uniform encoding schemes ensure the least complex encoder, while adaptive encoders provide relatively better reconstruction quality. The joint encoding scheme became prominent since it can provide a higher level of compression, thereby reducing the energy demand. The joint encoder offers high compression at higher GOPs and lower sensing rates, making it well-suited for WVSN applications. Further work related to data-gathering in WVSN may be focused on (i) developing an exhaustive and iterative reconstruction strategy to improve the reconstruction quality of the joint encoding scheme, (ii) optimization of the joint encoding scheme to reduce its computational complexity without compromising compression efficiency and reconstruction quality, (iii) developing an optimal camera node selection algorithm to improve network lifetime, and (iv) validation of simulation results by performing real-world testing to ensure applicability and reliability of proposed schemes.

AUTHOR CONTRIBUTIONS

Conceptualization, G. L. Priya and D. Ghosh; methodology, G. L. Priya; software, G. L. Priya; validation, G. L. Priya and D. Ghosh; formal analysis, G. L. Priya; investigation, G. L. Priya; data curation, G. L. Priya; writing – original draft preparation, G. L. Priya; writing – review and editing, D. Ghosh; visualization, G. L. Priya and D. Ghosh; supervision, D. Ghosh; funding acquisition, D. Ghosh. All authors have read and agreed to the published version of the manuscript.

References

[1] S. Eleuch, N. Khouja, S. Milani, T. Erseghe, and F. Tlili, "A distributed rate-control approach to reduce communication burdens in VSNs," *IEEE Access*, vol. 11, pp. 1011–1022, 2022.

[2] I. K. Abbood and A. K. Idrees, "Data reduction techniques for wireless multimedia sensor networks: a systematic literature review," *The Journal of Supercomputing*, pp. 1–46, 2023.

[3] J. Koteich, C. Salim, and N. Mitton, "Image processing-based data reduction technique in WVSN for smart agriculture," *Computing*, vol. 105, no. 12, pp. 2675–2698, 2023.

[4] T. Pal and S. Das Bit, "An improved low-overhead secure image compression over wireless multimedia sensor network," *Wireless Personal Communications*, vol. 132, no. 2, pp. 1049–1081, 2023.

[5] M. Ebrahim, S. H. Adil, T. Gul, and K. Raza, "Comparative analysis: Conventional video codecs v/s compressive sensing video codecs," in *Proc. 3rd International Conference on Emerging Trends in Engineering, Sciences and Technology (ICEEST)*, Karachi, Pakistan, 21-22 Dec. 2018.

[6] R. Monika, S. Dhanalakshmi, R. Kumar, and R. Narayanamoorthi, "Coefficient permuted adaptive block compressed sensing for camera enabled underwater wireless sensor nodes," *IEEE Sensors Journal*, vol. 22, no. 1, pp. 776–784, 2021.

[7] X. Chai, J. Fu, Z. Gan, Y. Lu, and Y. Zhang, "An image encryption scheme based on multi-objective optimization and block compressed sensing," *Nonlinear Dynamics*, vol. 108, no. 3, pp. 2671–2704, 2022.

[8] B. Lal, R. Gravina, F. Spagnolo and P. Corsonello, "Compressed Sensing Approach for Physiological Signals: A Review," in *IEEE Sensors Journal*, vol. 23, no. 6, pp. 5513-5534, March 2023.

[9] S. Zhang, J. Wu, D. Chen, S. Li, B. Yu, and J. Qu, "Fast frequency-domain compressed sensing

analysis for high-density super-resolution imaging using Orthogonal Matching Pursuit," *IEEE Photonics Journal*, vol. 11, no. 1, pp. 1–8, 2018.

[10] Z. Li, W. Xu, Y. Tian, Y. Wang, and J. Lin, "Compressed sensing reconstruction algorithms with prior information: Logit Weight Simultaneous Orthogonal Matching Pursuit," in *Proc. IEEE 79th Vehicular Technology Conference (VTC Spring)*, Seoul, Korea (South), 18-21 May 2014.

[11] M. A. Davenport, D. Needell, and M. B. Wakin, "Signal space CoSaMP for sparse recovery with redundant dictionaries," *IEEE Transactions on Information Theory*, vol. 59, no. 10, pp. 6820–6829, 2013.

[12] S. Li, H. Wang, T. Liu, Z. Cui, J. N. Chen, and Z. Xia, "A fast Barzilai-Borwein gradient projection for sparse reconstruction algorithm based on 3D modeling: Application to ERT imaging," *IEEE Access*, vol. 9, pp. 152 913–152 922, 2021.

[13] S. Mun and J. E. Fowler, "Block compressed sensing of images using directional transforms," in *Proc. IEEE 16th international conference on image processing (ICIP)*, Snowbird, UT, USA, pp. 3021–3024, 24–26 March 2010.

[14] S. Mun, "Residual reconstruction for block-based compressed sensing of video," in *Proc. IEEE Data Compression Conference*, Snowbird, UT, USA, pp. 183 –192, 29-31 March 2011.

[15] M. Azghani, M. Karimi, and F. Marvasti, "Multihypothesis compressed video sensing technique," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 26, no. 4, pp. 627–635, 2016.

[16] C. Chen, C. Zhou, P. Liu, and D. Zhang, "Iterative reweighted Tikhonov-regularized multihypothesis prediction scheme for distributed compressive video sensing," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 30, no. 1, pp. 1–10, 2020.

[17] R. Banerjee and S. Das Bit, "Low-overhead video compression combining partial discrete cosine transform and compressed sensing in WMSNs," *Wireless Networks*, vol. 25, no. 8, pp. 5113–5135, 2019.

[18] H.-q. Wang, Q.-n. Tang, Y.-h. Wang, and S.-l. Ren, "Application of D-KSVD in compressed sensing-based video coding," *Optik*, vol. 226, p. 165917, 2021.

[19] V. A. Nezhad, M. Azghani, and F. Marvasti, "Compressed video sensing based on deep generative adversarial network," *Circuits, Systems, and Signal Processing*, pp. 1–17, 2024.

[20] Y. Zhong, C. Zhang, X. Yang and S. Wang, "Video Compressed Sensing Reconstruction via an Untrained Network with Low-Rank Regularization," in *IEEE Transactions on Multimedia*, vol. 26, pp. 4590-4601, 2024.

[21] Z. Gu, C. Zhou, and G. Lin, "A temporal shift reconstruction network for compressive video sensing," *IET Computer Vision*, vol. 18, no. 4, pp. 448–457, 2024.

[22] W. Shi, S. Liu, F. Jiang, and D. Zhao, "Video compressed sensing using a convolutional neural network," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 31, no. 2, pp. 425–438, 2020.

[23] Z. Wei, C. Yang, and Y. Xuan, "Efficient video compressed sensing reconstruction via exploiting spatial-temporal correlation with measurement constraint," in *IEEE International Conference on Multimedia and Expo (ICME)*, pp. 1–6, 2021.

[24] X. Yang and C. Yang, "Imrnet: an iterative motion compensation and residual reconstruction network for video compressed sensing," in *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 2350–2354, 2021.

[25] Y. Pei, Y. Liu, and N. Ling, "Deep learning for block-level compressive video sensing," in *IEEE international symposium on circuits and systems (ISCAS)*, pp. 1–5, 2020.

[26] S. Zhang, K. Li, J. Xu, and G. Qu, "Image adaptive coding algorithm based on compressive sensing," *J. Tianjin Univ.*, vol. 45, no. 4, pp. 319–324, 2012.

[27] R.-F. Wang, L.-C. Jiao, F. Liu, and S.-Y. Yang, "Block-based adaptive compressed sensing of image using texture information," *ACTA ELECTONICA SINICA*, vol. 41, no. 8, p. 1506, 2013.

[28] H. Hadizadeh and I. V. Bajic´, "Soft video multicasting using adaptive compressed sensing," *IEEE transactions on multimedia*, vol. 23, pp. 12– 25, 2020.

[29] Z. Zhang, H. Bi, X. Kong, N. Li, and D. Lu, "Adaptive compressed sensing of color images based on salient region detection," *Multimedia Tools and Applications*, vol. 79, pp. 14777–14791, 2020.

[30] R. Li, X. Duan, X. Guo, W. He, and Y. Lv, "Adaptive compressive sensing of images using spatial entropy," *Computational Intelligence and Neuroscience*, vol. 2017, 2017.

[31] J. Zhang, C. Zhao, D. Zhao, and W. Gao, "Image compressive sensing recovery using adaptively learned sparsifying basis via L0 minimization," *Signal Processing*, vol. 103, pp. 114–126, 2014.

[32] Y. Yang, D. Zhang, and F. Ding, "Distributed compressive video sensing with adaptive measurements based on temporal correlativity," in *Proc. IEEE 9th International Conference on Wireless Communications and Signal Processing (WCSP)*, Nanjing, China, 11-13 Oct. 2017.

[33] J. Yang, H. Wang, I. Taniguchi, Y. Fan and J. Zhou, "aVCSR: Adaptive Video Compressive Sensing Using Region-of-Interest Detection

in the Compressed Domain," in *IEEE MultiMedia*, vol. 31, no. 1, pp. 19-32, Jan.-March 2024.

[34] Y. Zhao, Q. Zeng, and E. Y. Lam, "Adaptive compressed sensing for real-time video compression, transmission, and reconstruction," in *IEEE 10th International Conference on Data Science and Advanced Analytics (DSAA)*, pp. 1–10, 2023.

[35] J. Zhang, Q. Xiang, Y. Yin, C. Chen, and X. Luo, "Adaptive compressed sensing for wireless image sensor networks," *Multimedia Tools and Applications*, vol. 76, pp. 4227–4242, 2017.

[36] G. L. Priya and D. Ghosh, "An effectual video compression scheme for WVSNs based on block compressive sensing," *IEEE Transactions on Network Science and Engineering*, vol. 11, no. 2, pp. 1542–1552, 2024.

[37] M. Ebrahim and W. C. Chai, "Multi-phase joint reconstruction frame- work for multiview video compression using block-based compressive sensing," in *Visual Communications and Image Processing (VCIP)*, pp. 1–4, 2015.

[38] Y. Liu, C. Zhang, and J. Kim, "Disparity-compensated total-variation minimization for compressed-sensed multiview image reconstruction," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 1458–1462, 2015.

[39] N. Cen, Z. Guan, and T. Melodia, "Interview motion compensated joint decoding for compressively sampled multiview video streams," *IEEE Transactions on Multimedia*, vol. 19, no. 6, pp. 1117–1126, 2017.

[40] J. Zhu, J. Wang and Q. Zhu, "Compressively Sensed Multi-View Image Reconstruction Using Joint Optimization Modeling," *2018 IEEE Visual Communications and Image Processing (VCIP)*, Taichung, Taiwan, pp. 1-4, 2018.

[41] Y. Liu, D. A. Pados, J. Kim, and C. Zhang, "Reconstruction of compressed-sensed multiview video with disparity-and motion-compensated total variation minimization," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 28, no. 6, pp. 1288–1302, 2017.

[42] X. Fei, L. Li, H. Cao, J. Miao, and R. Yu, "View's dependency and low-rank background-guided compressed sensing for multiview image joint reconstruction," *IET Image Processing*, vol. 13, no. 12, pp. 2294–2303, 2019.

[43] Y. Song, D. Zhang, Q. Tang, S. Tang, and K. Yang, "Local and non-local constraints for compressed sensing video and multiview image recovery," *Neurocomputing*, vol. 406, pp. 34–48, 2020.

[44] N. Cen, Z. Guan, and T. Melodia, "Compressed sensing based low-power multiview video coding and transmission in wireless multi-path multihop networks," *IEEE Transactions on Mobile Computing*, vol. 21, no. 9, pp. 3122–3137, 2022.

[45] K. Q. Dinh and B. Jeon, "Iterative weighted recovery for block- based compressive sensing of image/video at a low subrate," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 27, no. 11, pp. 2294–2308, 2017.

[46] C.-K. Wen, J. Zhang, K.-K. Wong, J.-C. Chen, and C. Yuen, "On sparse vector recovery performance in structurally orthogonal matrices via LASSO," *IEEE Transactions on Signal Processing*, vol. 64, no. 17, pp. 4519–4533, 2016.

G. L. Priya received the B.Tech. degree in Electronics and Communication Engineering from the University of Kerala, Thiruvananthapuram, India, in 2006 and the M.E. degree in Applied Electronics from Karunya University, Coimbatore, India, in 2008. Currently, she is working toward the Ph.D. degree with the Department of Electronics and Communication Engineering, Indian Institute of Technology, Roorkee. Her research interests include image/video processing, wireless visual sensor networks, and sparse signal processing.

Debashis Ghosh (Senior Member, IEEE), received the Ph.D. degree from the Indian Institute of Science, Bangalore, India, in 2000. He is currently a Professor in the Department of Electronics and Communication Engineering, Indian Institute of Technology Roorkee, Roorkee, India. From April 1999 to November 1999, he was a DAAD Research Fellow at the University of Kaiserslautern, Kaiserslautern, Germany. In November 1999, he joined the Indian Institute of Technology Guwahati, Guwahati, India, as an Assistant Professor of Electronics and Communication Engineering. He spent the 2003–2004 academic year as a visiting faculty member in the Department of Electrical and Computer Engineering at the National University of Singapore. Between 2006 and 2008, he was a Senior Lecturer with the Faculty of Engineering and Technology, Multimedia University, Malaysia. His teaching and research interests include communication systems, signal processing, image/video processing, computer vision, and pattern recognition.