Computers and Electrical Engineering 44 (2015) 51–66

Contents lists available at ScienceDirect

Computers and Electrical Engineering

journal homepage: www.elsevier.com/locate/compeleceng

Video Compressed Sensing framework for Wireless Multimedia Sensor Networks using a combination of multiple matrices ☆

Aasha Nandhini Sukumaran *, Radha Sankararajan, Kishore Rajendiran

Department of ECE, SSN College of Engineering, Chennai, India

ARTICLE INFO

Article history:
Received 14 May 2014
Received in revised form 7 February 2015
Accepted 9 February 2015
Available online 12 March 2015

Keywords:
WMSN
VCS
DWT–DCT
Combination matrix
Hybrid matrix
Orthogonal matching pursuit

ABSTRACT

Wireless multimedia sensor networks (WMSNs) have been used for sensitive applications such as video surveillance and monitoring applications. In a WMSN, storage and transmission become complicated phenomena that can be simplified by the use of compressed sensing, which asserts that sparse signals can be reconstructed from very few measurements. In this paper, memory-efficient measurement matrices are proposed for a discrete wavelet transform (DWT)–discrete cosine transform (DCT) hybrid approach based video compressed sensing (VCS). The performance of the framework is evaluated in terms of PSNR, storage complexity, transmission energy and delay. The results show that the proposed matrices yield similar or better PSNR and consume less memory for generating the matrix when compared with a Gaussian matrix. The DWT–DCT based VCS yields better quality and compression when compared with DCT and DWT approaches. The transmission energy is 50% less and the average delay is 52% less when compared to raw frame transmission.

1. Introduction

A WMSN consists of simple and low-cost sensor nodes that are used for a variety of applications, such as environmental monitoring, healthcare monitoring and surveillance applications. In the case of surveillance applications, sensor nodes with multimedia capability are deployed in the area of interest to detect anomalies. These sensor nodes will capture and transmit the surveillance video wirelessly to the network operator. Transmission of multimedia data that range from several megabytes to a few gigabytes is a challenging task because it imposes requirements for large amounts of storage and high bandwidth for transmission. These challenges can be overcome by using an emerging technique called compressed sensing (CS), which asserts that the original signal can be reconstructed from a small number of measurements [1]. CS can be applied to sparse signals or compressible signals. The original signal is made sparse using a transform basis, and a measurement matrix is applied to the sparse signal to obtain the measurements. These measurements are transmitted for reconstruction at the receiver side. There are many reconstruction algorithms, such as basis pursuit, greedy algorithms and iterative algorithms, for reconstructing the original signal from the measurement vector.

The objective of this paper is to implement VCS based on a DWT–DCT hybrid approach and to propose efficient measurement matrices for VCS. This VCS framework is based on a block-based approach; CS is applied to all of the blocks to obtain

☆ Reviews processed and recommended for publication to the Editor-in-Chief by Associate Editor Dr. S. Rajavelu.
* Corresponding author. Tel.: +91 9962089611.
E-mail addresses: aasha.nandhu@gmail.com (S. Aasha Nandhini), radhas@ssn.edu.in (R. Sankararajan), kishorer@ssn.edu.in (K. Rajendiran).

http://dx.doi.org/10.1016/j.compeleceng.2015.02.008
0045-7906/

the measurements. The performance of the system with the proposed measurement matrices is evaluated by analysing the PSNR, percentage of reduction in samples, storage, energy complexity, transmission energy and end-to-end delay.

The rest of the paper is organized as follows. Section 2 provides a brief survey of related works. Section 3 provides a brief description of the CS technique, and Section 4 provides details about VCS based on the DCT, DWT and DWT–DCT hybrid approach. This section also explains in detail the proposed measurement matrices and orthogonal matching pursuit (OMP) algorithm, whereas Section 5 discusses the simulation and experimental results in detail; finally, Section 6 gives the conclusions and proposed future work.

2. Related works

This section provides a brief discussion of works related to VCS along with the advantages and limitations of each technique.

Akyildiz et al. [2] provided a survey of techniques used in the algorithms, hardware and protocols for WMSNs. Different architectures of WMSNs are explored along with their advantages and disadvantages. Different varieties of low-resolution and medium-resolution image sensor nodes and their collaborative processing are also explained. The authors have also discussed the working of wireless nodes in different layers of transmission. The network management and security used for multimedia transmission purposes are also emphasized.

In [3,4], the authors have explained in detail the basics, theorems and mathematics involved in CS and how the mathematics can be implemented in real time. They have motivated the design of new sampling schemes and devices that provide information required for signal recovery using the smallest possible representation.

Gan [5] has proposed block-based CS for natural images, in which image acquisition is conducted in a block-by-block manner. The image reconstruction algorithm includes both linear and nonlinear operations such as Wiener filtering, projection onto the convex set and hard thresholding in the transform domain. The results show comparable performances between block-based CS systems and current CS schemes, but the former have much lower implementation costs.

Han et al. [6] have proposed an image representation scheme using CS because it reduces the computational complexity of the image/video encoder used in the compression process. The encoder first divides the image into two parts, dense and sparse components, where the sparse component alone is encoded using CS. The number of random measurements needed and the decoding complexity are reduced considerably. Yu et al. [7] proposed a method to construct a Toeplitz-structured matrix with chaotic sequences for CS and proved that the Toeplitz-structured Chaotic Sensing Matrix retains the Restricted Isometry Property (RIP), which guarantees exact recovery. The results show similar or better performance compared with the Chaotic Sensing Matrix and the Gaussian Sensing Matrix.

Gleichman and Eldar [8] stated that the necessity of prior knowledge of the sparsity of the signal can be avoided by using blind CS. They have proven that under some unique constraints, their methods could produce comparable results to those of conventional CS techniques. Jain and Tewari [9] slightly modified OMP to generate Orthogonal Matching Pursuit with Replacement (OMPR), which replaces an existing support element with a new one. They have explained different methods for faster implementation of OMPR algorithms, such as using a hash function. Additionally, performance for noisy samples was also analysed.

Schaeffer et al. [10] presented a method to predict the optimal real-time compression rate for videos. During the acquisition of input data, the video is spatially compressed. Simple filters and polynomial fittings are used so that the hardware implementation is easy. The quality of the reconstructed video can also be improved without increasing the amount of video stored.

Schenkel et al. [11] proposed a scheme for wireless video multicast based on CS. When multicasting video signals, CS is applied to generate measurements of equal importance from a video such that a receiver with a better channel will naturally have more information to reconstruct the content effectively. They have also examined how the properties of the natural image can be exploited to improve the reconstruction performance by adding small side information. They experimented with different matrices to improve performance, not only on the receiver side but also on the transmitter side.

Jiang et al. [12] proposed a method to separate moving objects from the background in a surveillance video. Low rank and sparse decomposition of the matrix was used to reconstruct the video acquired by compressive measurements. The low rank matrix was used to identify the background, and the sparse matrix represented any moving objects present in the video. The experimental results showed that very few measurements were needed to extract moving objects in the surveillance video.

In [13], Balouchestani et al. provided the most recent survey of CS theory applied to wireless sensor networks (WSNs). The authors also proposed new sampling methods for information, data networks and a new recovery algorithm. In [14], Pudlewski et al. explained compressive video streaming for WMSN. They found that new cross-layered network protocols combined with video encoders based on CS can provide solutions to the problems caused by encoder complexity and limited resiliency to channel errors. The surveillance image when CS is used shows better reconstruction results compared to the same surveillance image reconstructed using JPEG for bit error rates (BERs) to the powers of -5, -4 and -3.

In [15] and [16], Zhang et al. proposed a new framework for image CS recovery via collaborative sparsity and structural group sparse representation (SGSR) modelling, respectively. An efficient augmented Lagrangian-based technique and an iterative shrinkage/thresholding algorithm-based technique were also developed to solve the optimization problem. The experimental results showed that this novel CS recovery strategy increased the image quality by a large margin compared with current state-of-the-art schemes.

Sungkwang and Fowler [17] proposed block-based CS sampling with a smooth projected Landweber (BCS-SPL) reconstruction framework for video. The authors have incorporated the motion estimation/motion compensation (ME/MC) framework into the reconstruction process of BCS-SPL. The results showed that the independent frame by frame reconstruction was fast but that the reconstruction achieved by MC-BCS-SPL had significantly higher quality.

Liu et al. [18] proposed a sparsity-aware motion-accounting decoder for video streaming systems with plain CS encoding. The decoder performs sliding-window interframe decoding that adaptively estimates Karhunen–Loève bases from adjacent previously reconstructed frames. This decoder enhances the sparse representation of each video frame, improving the reconstruction quality of video at any fixed CS rate. The sparsity-aware motion-accounting decoder outperforms the conventional fixed-basis intraframe and interframe and K-SVD decoders.

Jha et al. [19] proposed a CS-based sparse representation of the directional empirical mode decomposition (DEMD) residue of the frame. The wavelet transform and the noiselet transform are used as a sparsifying matrix and sensing matrix, respectively, for CS measurements. The experimental results show that the proposed approach provides better compression results compared with the standard wavelet transform in CS schemes.

3. Compressed sensing

Compressed sensing is a new approach in which sensing and compression are performed simultaneously, resulting in a significant reduction in sampling and computation cost at the sensor. CS theory states that the original signal can be reconstructed from very few measurements; hence, it is also called sub-Nyquist sampling. This technique ultimately reduces the complexity of the process, requires much less storage capacity and uses less bandwidth. CS provides accurate recovery of the signal represented by a vector of length N from $M \ll N$ measurements [1].

An important requirement of the CS technique is the sparsity of the input signal. Every signal, $Z \in R^N$ is supposed to have K non-zero coefficients. The ultimate goal of CS is to construct a set of M samples so that the unknown signal Z can be recovered from the sample measurements. Thus an $M \times N$ matrix generated for obtaining M measurements is called the measurement matrix Φ. These measurements are transmitted to the receiver for reconstruction. Any compressible signal $Z \in R^N$ can be represented as

$$Z = \Psi r \tag{1}$$

where Ψ represents the sparsifying basis and r is the $N \times 1$ sparse vector. The compressible signal Z can be represented as a linear combination of K vectors with $K \ll N, K$ non-zero coefficients and $N - K$ zero coefficients. In many applications, signals have a small number of large coefficients that can be approximated by K. The CS method offers M measurements ($K < M \ll N$) and enough information to reconstruct Z. The measurement matrix Φ is used to obtain the measurement vector y:

$$y = \Phi Z \tag{2}$$

Using (1), the measurement vector can be represented as:

$$y = \Phi Z = \Phi \Psi r \tag{3}$$

$$y = \Theta r \tag{4}$$

where $\Theta = \Phi\Psi$ represents an $M \times N$ matrix and the resulting measurement vector y is a $M \times 1$ vector. The measurement process for M is non-adaptive, and hence, Φ is independent of the signal Z. To recover the original signal, a good reconstruction algorithm and M measurements are needed. For the reconstruction process to be perfect, the measurement matrix must obey certain properties, such as the incoherence property and the restricted isometry property. In [20], Candes and Tao explained that the restricted isometric property constrains the measurement matrices such that the reconstruction of the measurement vector will be flawless. The idea of CS is explained by the block diagram shown in Fig. 1.

The diagram explains how CS is performed to obtain the measurements. The input signal is made sparse because CS can only be applied to sparse signals. The signal can be made sparse using any transform domain, such as DCT or DWT. After being sparsified, the signal is converted into a single vector called the sparse vector. The sparse vector consists of K non-zero values that represent the sparsity level of the vector. The measurement matrix of dimension $M \times N$ is generated by calculating the minimum number of measurements M required for reconstructing the original signal. The measurement matrix

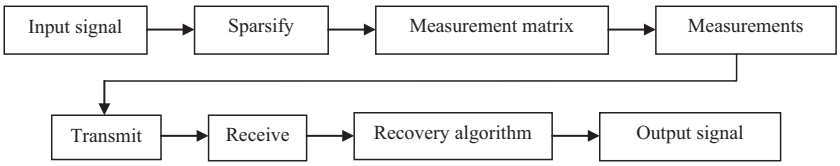

Fig. 1. Block diagram of compressed sensing.

is applied to the sparse vector to obtain the measurements. These measurements are then transmitted for reconstruction at the receiver end.

4. Video compressed sensing framework

This section discusses the implementation of the proposed VCS based on the DWT–DCT hybrid approach and compares it with the VCS based on DCT and DWT approaches. Efficient measurement matrices are proposed and implemented that are further compared with the Gaussian measurement matrix. In this framework, measurement matrices are proposed using a combination of multiple matrices. The block-based approach is adopted for VCS in order to reduce the size of the measurement matrix that is used to obtain the measurements. Considering the constraints of WMSN, a small measurement matrix is used to reduce the storage capacity required. The measurements that are obtained from the framework are transmitted and reconstructed using the OMP recovery algorithm.

Consider a video sequence that consists of frames represented as $z_j \in R^N$ formed from the pixels of frame j of the video sequence, for $j = 1, 2, \ldots, J$, where J is the total number of frames and N is the total number of pixels in a frame. Let $Z = [z_1, z_2, \ldots, z_j] \in R^{N_1 \times J}$ be a video formed from video frames and the total number of pixels in a video be $N_1 J$. Each frame is made sparse by using an orthonormal basis of $N \times 1$ vectors $\{\Psi_i\}_{i=1}^{N}$. Using a basis matrix of dimension $N \times N$ with $\{\Psi_i\}$ as columns, the signal can be expressed as

$$z_j = \Psi_i r_j \tag{5}$$

where r_j is a $N \times 1$ sparse vector with K non-zero coefficients. The signal Z_j is said to be sparse if it has very few non-zero coefficients. This signal is multiplied with the measurement matrix of dimension $\Phi_j \in R^{M \times N}$ to obtain the measurement vector y_j using (6)

$$y_j = \Phi_j z_j \tag{6}$$

The measurement vector y_j is of length M, which is much smaller than the total number of pixels in the frame. The measurement matrix of dimension $M \times N$ is generated by calculating the minimum number of measurements required to reconstruct the signal using (7)

$$M \geqslant K \log(N/K) \tag{7}$$

where K denotes the sparsity level and N is the total number of samples in a frame [1]. Fig. 2 shows the detailed block diagram of the VCS framework.

The input is the video sequence from which the frames of size $N \times N$ are separated, and each frame is divided into blocks of size $n \times n$. These blocks are sparsified, and CS is applied to each block to obtain the measurements. The CS process is explained in detail below.

4.1. Sparse representation

Sparsity is an underlying concept in most of today's filtering and compression applications. A signal is said to be sparse or compressible when the transform coefficient vector has a small number of large-amplitude coefficients and a large number of small-amplitude coefficients. In this paper, the DWT–DCT hybrid approach is used for a sparse representation that combines the advantages of both DCT and DWT.

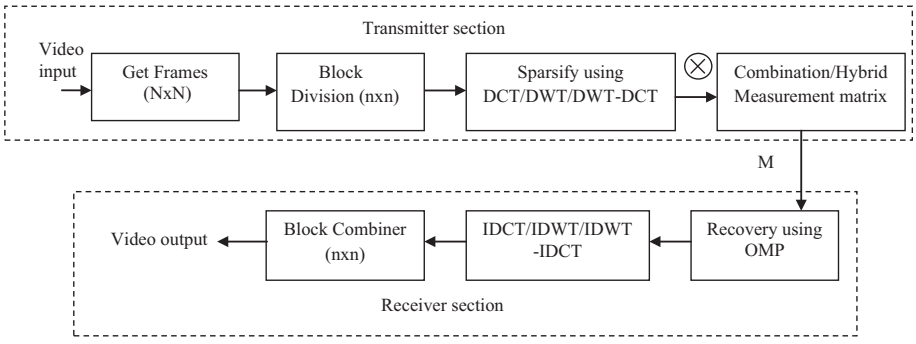

Fig. 2. Video compressed sensing framework.

4.1.1. CS based on the DCT approach

Each block is sparsified using DCT, as described in (8), and converted into a single vector $n^2 \times 1$ called a sparse vector because it has only a few non-zero coefficients. The number of non-zero elements in the sparse vector represents the sparsity level K_1. Given an image block $f(i_1, i_2)$ of size $n \times n$, the forward and inverse DCT are defined as [21]

$$F(u,v) = \frac{2}{n} c_u c_v \sum_{i_1=0}^{n-1} \sum_{i_2=0}^{n-1} f(i_1,i_2) \cos\left[\frac{\pi(2i_1+1)u}{2n}\right] \cos\left[\frac{\pi(2i_2+1)v}{2n}\right] \tag{8}$$

$$f(i_1,i_2) = \frac{2}{n} c_u c_v \sum_{u=0}^{n-1} \sum_{v=0}^{n-1} F(u,v) \cos\left[\frac{\pi(2i_1+1)u}{2n}\right] \cos\left[\frac{\pi(2i_2+1)v}{2n}\right] \tag{9}$$

where $u, v = 0, 1, \ldots, n-1$ and $c_u, c_v = \begin{cases} \frac{1}{\sqrt{2}}, & u, v = 0 \\ 1, & u, v \neq 0 \end{cases}$

When DCT is applied to the block, most of the significant information is concentrated in a few non-zero coefficients only; it thereby achieves a higher compression rate.

4.1.2. CS based on the DWT approach

DWT is applied to each block as described in (10) and (11), which results in four bands (LL, LH, HL, and HH). Given an image block $f(i_1, i_2)$ of size $n \times n$, the forward and inverse DWT are defined as [22]

$$W_\varphi(j_0,k_1,k_2) = \frac{1}{\sqrt{n^2}} \sum_{i_1=0}^{n-1} \sum_{i_2=0}^{n-1} f(i_1,i_2) \varphi_{j_0,k_1,k_2}(i_1,i_2) \tag{10}$$

$$W_\psi^q(j_1,k_1,k_2) = \frac{1}{\sqrt{n^2}} \sum_{i_1=0}^{n-1} \sum_{i_2=0}^{n-1} f(i_1,i_2) \psi_{j_1,k_1,k_2}^q(i_1,i_2) \tag{11}$$

$$f(i_1,i_2) = \left\{ \frac{1}{\sqrt{n^2}} \sum_{i_1} \sum_{i_2} W_\varphi(j_0,k_1,k_2) \phi_{j_0,k_1,k_2}(i_1,i_2) + \frac{1}{\sqrt{n^2}} \sum_{q=H_1,V_1,D_1} \sum_{j_1=0}^{\infty} \sum_{k_1} \sum_{k_2} W_\psi^q(j_1,k_1,k_2) \psi_{j_1,k_1,k_2}^q(i_1,i_2) \right\} \tag{12}$$

where $W_\varphi(j_0,k_1,k_2)$, $W_\psi^q(j_1,k_1,k_2)$ denote the approximation coefficients and detail coefficients, respectively, $\varphi_{j_0,k_1,k_2}(i_1,i_2)$, $\psi_{j_1,k_1,k_2}^q(i_1,i_2)$ denote the scaling and wavelet functions, respectively, $j_0 = 0, j_1 \geqslant j_0$ represents the scale and $q = \{H_1, V_1, D_1\}$ indicates the index of LH, HL, HH bands. The LL band contains the approximation coefficients, and the other three bands contain the detail coefficients. The LH, HL and HH bands will be sparse, that is, converted into a sparse vector. DWT has multi-resolution capability, thus achieving better quality.

4.1.3. CS based on the hybrid DWT–DCT approach

This approach combines the advantages of both DCT and DWT, achieving better quality with a higher compression rate. DWT is applied to each block using (10) and (11), which results in four bands, and then DCT is applied to the LH, HL and HH bands using (8), achieving more sparsity. These bands with few non-zero coefficients are converted into a single vector called the sparse vector to which the measurement matrix is applied to obtain the measurements.

4.2. Measurement matrix

The measurement matrix is generated based on the number of measurements required for reconstruction as in (7). Two measurement matrices are proposed using multiple matrices, such as the Gaussian matrix, the Binary matrix and the Toeplitz matrix. The first proposed matrix is called the Combination matrix, which is defined as the Kronecker product of the Gaussian matrix and the Toeplitz matrix. The second matrix, called the Hybrid matrix, is generated by combining the Toeplitz matrix and the Binary matrix.

4.2.1. Combination matrix

This matrix is obtained by taking the Kronecker product of the Gaussian matrix and the Toeplitz matrix. The Gaussian matrix is used universally, as it can be paired with any sparse matrix. It also yields better results in terms of PSNR. The Toeplitz matrix is highly preferred because the number of elements used in the measurement matrix is less. The Toeplitz matrix is a square matrix, and the elements are constant along its diagonal. The Kronecker product of matrix 'A' of size $m_1 \times n_1$ and matrix 'B' of size $p \times q$ yields a matrix of size $m_1 p \times n_1 q$ that is given as

$$A \otimes B = \begin{bmatrix} a_{11}B & \cdots & a_{1n_1}B \\ \vdots & \ddots & \vdots \\ a_{m_1 1}B & \cdots & a_{m_1 n_1}B \end{bmatrix}$$

where A is a Gaussian matrix and B is a Toeplitz matrix. For example, let A be a 2×2 Gaussian matrix represented as

$$A = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix}$$

and B be a 2×2 Toeplitz matrix represented as

$$B = \begin{bmatrix} t_{11} & t_{12} \\ t_{21} & t_{11} \end{bmatrix}$$

the Combination matrix of size 4×4 is generated by taking the Kronecker product of A and B represented as

$$C = A \otimes B = \begin{bmatrix} g_{11}t_{11} & g_{11}t_{12} & g_{12}t_{11} & g_{12}t_{12} \\ g_{11}t_{21} & g_{11}t_{11} & g_{12}t_{21} & g_{12}t_{11} \\ g_{21}t_{11} & g_{21}t_{12} & g_{22}t_{11} & g_{22}t_{12} \\ g_{21}t_{21} & g_{21}t_{11} & g_{22}t_{21} & g_{22}t_{11} \end{bmatrix}$$

4.2.2. Hybrid matrix

The Hybrid measurement matrix of size $M_1 \times n^2$ is generated by combining the Toeplitz matrix of size $M_1 \times n^2/2$ and the Binary matrix of size $M_1 \times n^2/2$. The Toeplitz matrix is generated with -1 and +1 entries, and the Binary matrix is generated with 0 and 1 values. For example, let A be a 2×2 Toeplitz matrix represented as

$$A = \begin{bmatrix} t_{11} & t_{12} \\ t_{21} & t_{11} \end{bmatrix}$$

and B be a 2×2 Binary matrix represented as

$$B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$$

then the Hybrid matrix of size 2×4 is given as

$$H = \begin{bmatrix} t_{11} & t_{12} & b_{11} & b_{12} \\ t_{21} & t_{11} & b_{21} & b_{22} \end{bmatrix}$$

These proposed matrices are memory efficient, as they require fewer elements to generate the matrix when compared with the existing Gaussian matrix. The measurement matrix is applied to the sparse vector to obtain the measurement vector y that is transmitted to the receiver side, where it is then reconstructed using the OMP algorithm.

4.3. Recovery algorithm

The OMP algorithm is one of the greedy algorithms designed to reconstruct the sparse vector from very few measurements by exploiting an iterative procedure. The iterations required to obtain the estimated sparse vector are the same as the sparsity level. This algorithm goes through an iterative procedure to minimize the residual error [23]. The OMP algorithm has five major steps, as shown in Table 1

The OMP algorithm yields better recovery and has less computational complexity. The inverse transform is applied to the estimated sparse vector using (9) and (12) to reconstruct the original frame. The DWT–DCT hybrid approach-based VCS with

Table 1
OMP algorithm.

Input:
- An $M_1 \times n^2$ measurement matrix Φ
- An $M_1 \times 1$ measurement vector y
- Number of iterations K_1

Output:
- Estimated signal es_i of size $n^2 \times 1$
- Index set IND_t containing t elements from $\{1, \ldots, n^2\}$
- An $M_1 \times 1$ residual vector res

Procedure:
(1) Set the residual value initially as $res = y$ where y is the measurement vector.
(2) Determine the index I_t by taking the maximum correlation between the measurement matrix and res.
(3) Update the index $IND_t = IND_{t-1} \cup I_t$ and the measurement matrix $\Phi_i = \Phi_{i-1} \cup \Phi_i$.
(4) Determine the estimated signal es_i using the least squares method.
(5) Calculate the new residual $res = z_{11} - es_i$ and go back to step (2) and repeat the steps.

proposed measurement matrices yields better PSNR and requires less storage capacity, which makes it more suitable for video surveillance applications in WMSNs.

5. Results and discussion

The DWT–DCT approach-based VCS using the proposed Combination matrix and Hybrid matrix is implemented in MATLAB and the performance of the measurement matrix is analysed in terms of PSNR, storage and energy complexity. The measurements obtained from the VCS framework are transmitted in real time using TelosB motes. Experimental analysis is carried out by evaluating the quality of reconstruction, transmission energy and end-to-end delay.

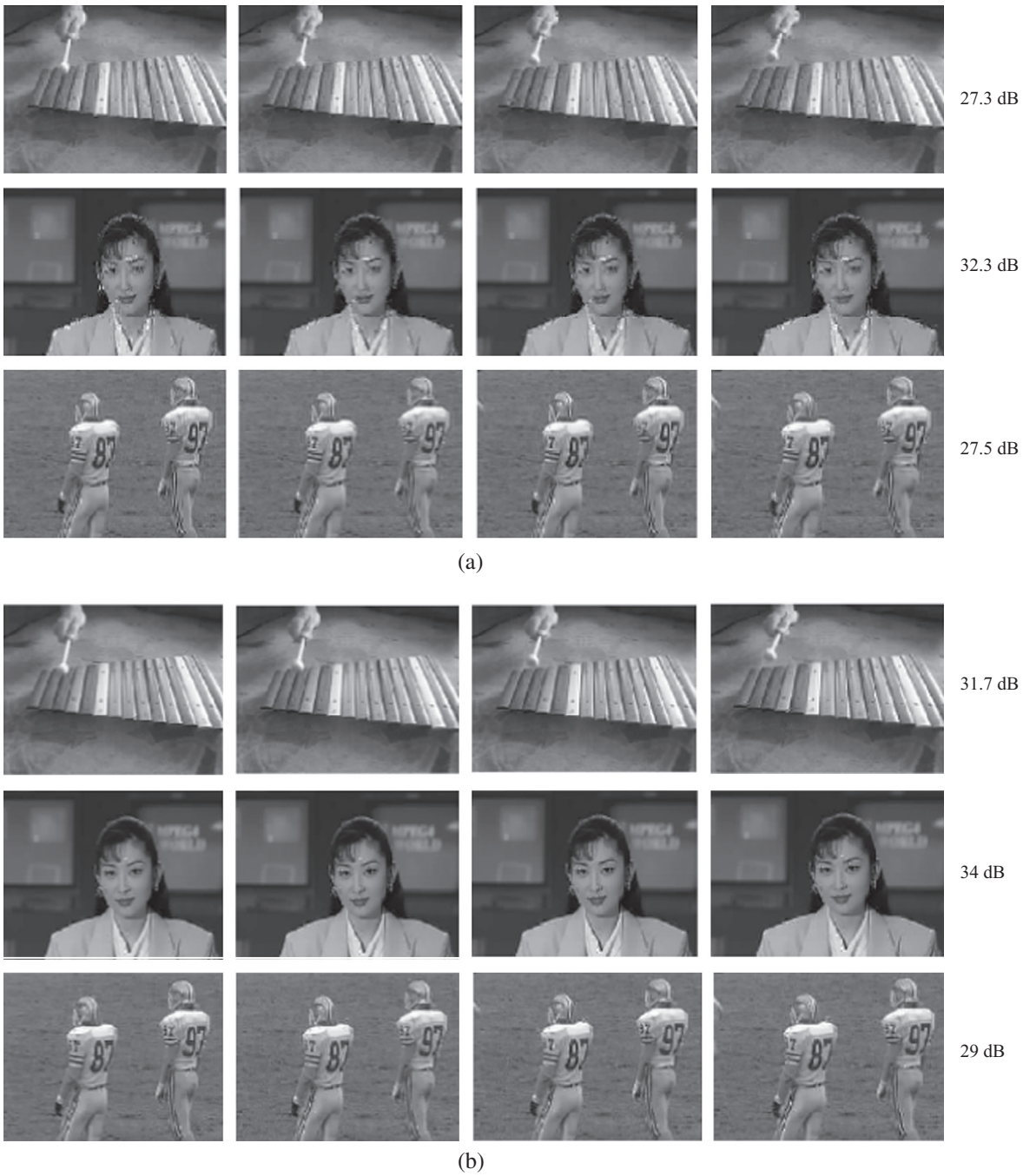

Fig. 3. Results for DCT approach-based VCS using (a) Combination matrix and (b) Hybrid matrix.

5.1. Simulation results

The video sequences considered for implementation are xylophone, akiyo and football. The first four video frames of size 240×320 are considered for implementation. In this VCS framework, the proposed Combination matrix of size $M_1 \times 64$ is generated by taking the Kronecker product of the Gaussian matrix of size $m_1 \times 8$ and the Toeplitz matrix of size $p \times 8$. The proposed Hybrid matrix of size $M_1 \times 64$ is generated by combining the Toeplitz matrix of size $M_1 \times 32$ and the Binary matrix of size $M_1 \times 32$. An analysis of the quality of the reconstruction for CS based on the DCT, the DWT and the hybrid DWT–DCT approaches is presented in detail.

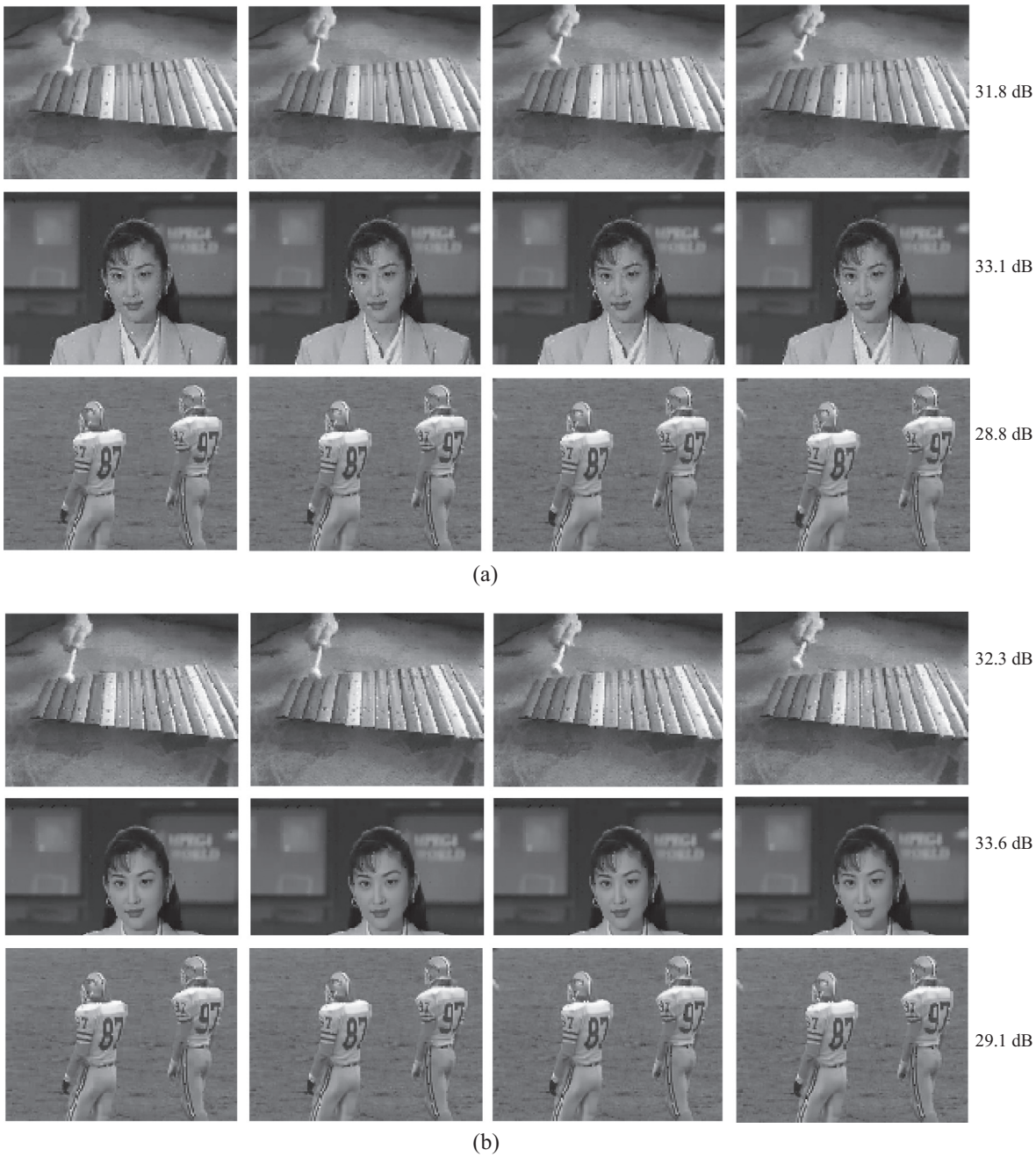

Fig. 4. Results for DWT approach-based VCS using (a) Combination matrix and (b) Hybrid matrix.

5.1.1. DCT approach-based video compressed sensing

In the DCT approach, frames are divided into blocks of size 8×8 and sparsified by applying DCT to each block. After being sparsified, the blocks are converted to a single vector, which is called the sparse vector. The sparsity level of the sparse vector is fixed at $K_1 = 5$, and the minimum number of measurements is calculated to be $M_1 = 21$ using (7). The Combination matrix and the Hybrid matrix are applied to each sparse vector to obtain the measurement vector. These measurement vectors are transmitted to the receiver side for reconstruction. The OMP algorithm is used to obtain the estimated vector, and further blocks are reconstructed by applying the inverse discrete cosine transform (IDCT). Fig. 3 shows the reconstructed frames along with average the PSNR using DCT for the xylophone, akiyo and football sequences for $M_1 = 30$.

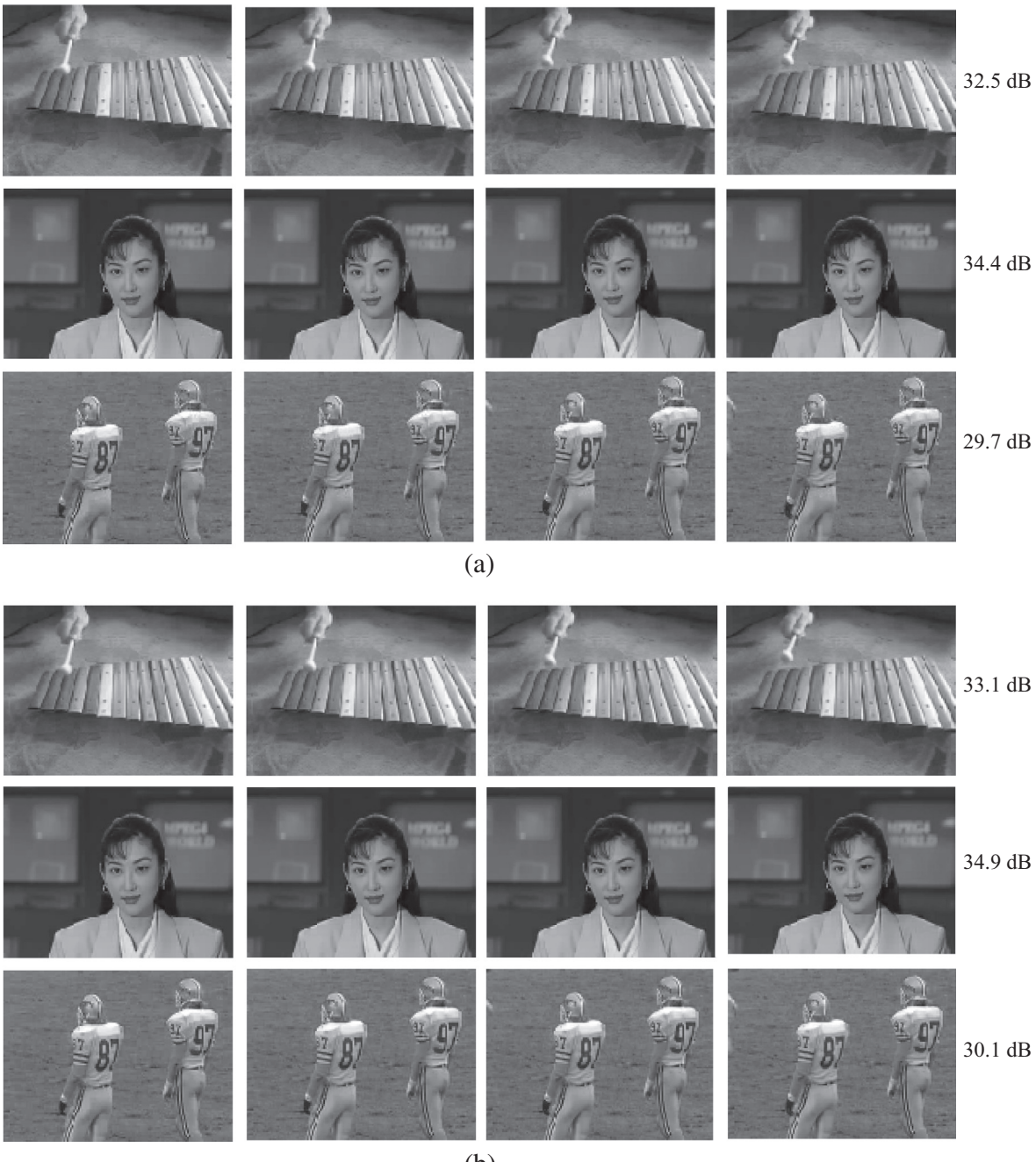

Fig. 5. Results for DWT–DCT approach-based VCS using (a) Combination matrix and (b) Hybrid matrix.

Table 2
Comparison of Gaussian matrix, Hybrid matrix and Combination matrix in terms of PSNR for DCT, DWT and DWT–DCT hybrid approach based VCS.

Video sequence	M_1	Average PSNR (dB)								
		Gaussian matrix			Hybrid matrix			Combination matrix		
		DCT	DWT	DWT–DCT	DCT	DWT	DWT–DCT	DCT	DWT	DWT–DCT
Xylophone	20	24.8	28.6	30.1	27.5	29.6	30.8	25.6	29.3	30.6
	30	27.3	30.2	32.3	31.7	32.3	33.1	27.3	31.8	32.5
	40	28.9	32.4	34.4	32.6	33.5	35.1	28.6	32.5	34.7
Akiyo	20	28.1	31.5	32.6	29.6	32.2	33.4	28	32.1	32.8
	30	31.4	32.9	34.2	34	33.6	34.9	32.3	33.1	34.4
	40	34.8	35	35.3	35	35.2	35.8	34.9	35.1	35.4
Football	20	23.1	25.8	27.8	25.9	27.5	28.6	24.8	27	28.5
	30	26.3	28.9	30.1	29	29.1	30.1	27.5	28.8	29.7
	40	29.5	29.7	31.5	30.3	30.4	32	30.1	30.8	31.7

5.1.2. DWT approach-based video compressed sensing

For the DWT approach the frames are divided into blocks of size 16 × 16, and DWT is applied to each block, yielding four bands of size 8 × 8. After sparsifying, it is converted to a sparse vector, and the measurements are obtained by applying the Combination matrix and the Hybrid matrix. The blocks are finally recovered by applying the inverse discrete wavelet transform (IDWT) to the estimated sparse vector. Fig. 4 shows the reconstructed frames along with average PSNR using DWT for the xylophone, akiyo and football sequences for $M_1 = 30$.

5.1.3. DWT–DCT hybrid approach-based video compressed sensing

In the DWT–DCT hybrid approach, the frames are divided into blocks of size 16 × 16, and DWT is applied to each block, yielding four bands of size 8 × 8. DCT is applied to each band, yielding a higher compression rate. After sparsifying, it is converted to a sparse vector to which the proposed measurement matrices are applied to obtain the measurements. The blocks are finally recovered by applying IDCT and IDWT to the estimated sparse vector. Fig. 5 shows the reconstructed frames along with the average PSNR using DWT–DCT for the xylophone, akiyo and football sequences for $M_1 = 30$.

The PSNR values for the DCT, the DWT and the DWT–DCT hybrid approach using the Combination matrix and the Hybrid matrix are given in Table 2.

From Table 2, it is evident that the average PSNR increases with the number of measurements. It can also be seen that the average PSNR for the Combination matrix and the Hybrid matrix is similar or better when compared with that of the Gaussian matrix [17,18]. It is clear that the DWT–DCT hybrid approach yields better PSNR for fewer measurements than DCT and DWT, which proves that the DWT–DCT hybrid approach has better quality at higher compression rates.

5.2. Experimental results

Consider an experimental setup for transmitting the compressed measurements of video between an end-to-end PC. The DWT–DCT hybrid-based VCS is implemented in a PC, and compressed measurements are relayed wirelessly via the deployed network to the destination where the video is reconstructed. The deployed network consists of TelosB nodes operated under the Contiki OS platform. The TelosB consists of a low-power MSP430 microcontroller, a Chipcon CC2420 transceiver, 10 KB RAM, 48 KB flash and 1 MB external flash memory [24]. The Contiki OS is an open source embedded operating system that was designed for WSN [25]. The experimental analysis is carried out in terms of quality of reconstruction, transmission energy and end-to-end delay. The source node attached to the PC at the transmitter side broadcasts the measurements in the form of packets to the sink node via relay nodes. The sink node, which was deployed at a distance of 5 m from the source node, receives the measurements from which the video is recovered in a PC. Fig. 6 shows the reconstructed frames along with the average PSNR obtained using the Combination matrix and the Hybrid matrix for the xylophone, akiyo and football sequences from the measurements received by the sink node.

The experimental results of the DWT–DCT hybrid approach in terms of PSNR are shown in Table 3. This table shows the comparison between the Gaussian matrix, the Hybrid matrix and the Combination matrix in terms of PSNR.

It is observed from the table that, on average, the PSNR of the frames is approximately 1.5 dB less when compared to the simulation results given in Table 2. This reduction in PSNR is due to the packet loss that occurred during transmission.

5.2.1. Transmission energy

To compute the energy consumed for transmission in real time, the powertrace tool in the Contiki OS was used [26]. The computed energy was compared with a theoretical energy calculation. The transmission energy per bit (E_{tx}) is computed theoretically by considering the current and voltage specifications of the TelosB node and the time to transmit a 128 byte packet on a 250 kbits/s IEEE 802.15.4 link using (13)

$$E_{tx} = (t * I * V)/1024\ \text{J} \tag{13}$$

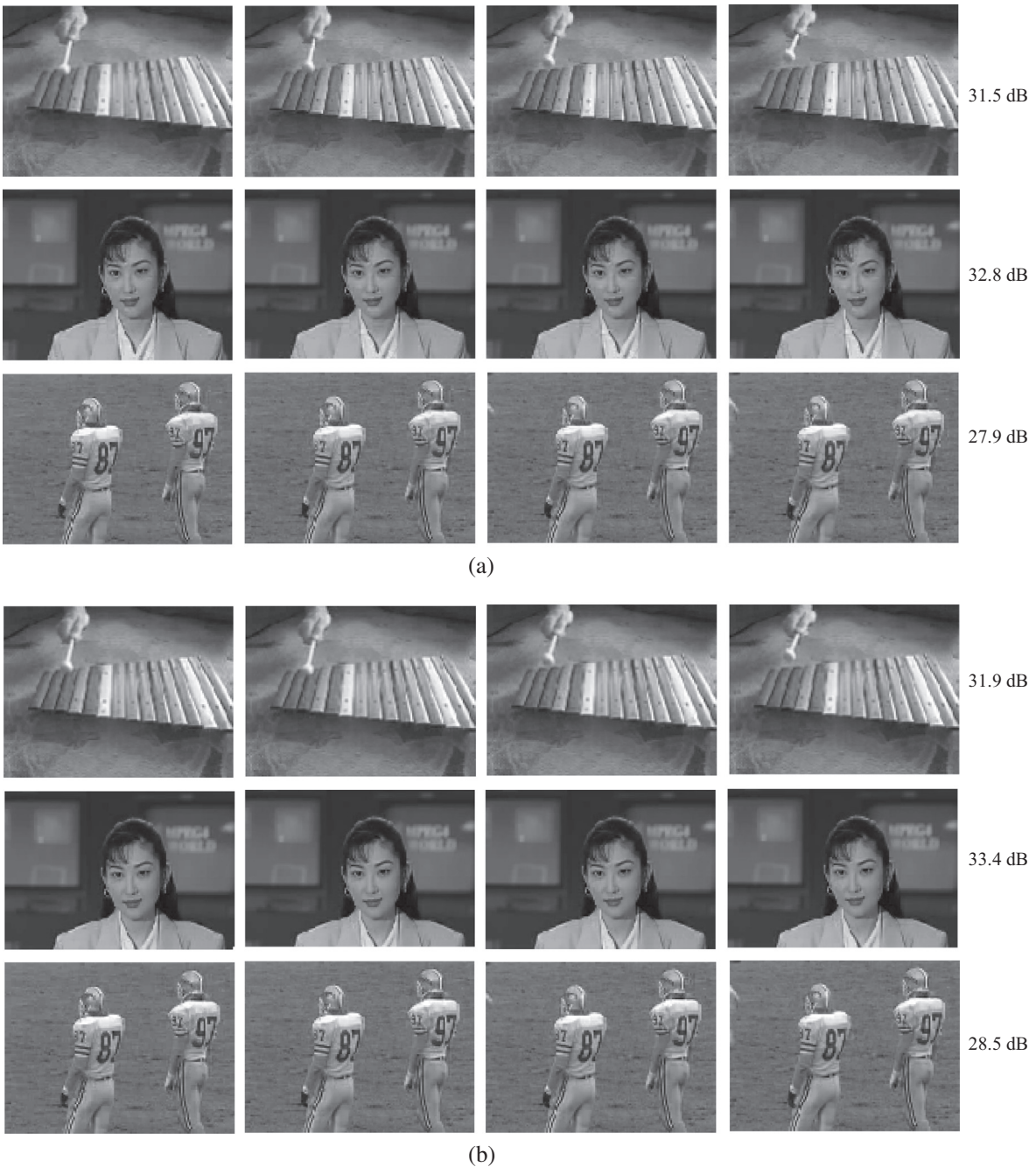

Fig. 6. Experimental results for DWT–DCT approach-based VCS using (a) Combination matrix and (b) Hybrid matrix.

where t represents the time to transmit a 128 byte packet, which is approximately 1/250 s, I represents the current, which is 19.5 mA taken from the TelosB datasheet, and V represents the voltage, which is 3 V. The computed transmission energy per bit is 0.23 μJ. The transmission energy per frame (E_f) was calculated theoretically using (14).

$$E_f = B_M * E_{tx} \tag{14}$$

where B_M represents the number of bits to be transmitted. Table 4 shows the theoretical and practical values of the transmission energy for different measurements.

The practical energy computed with the help of the powertrace tool is similar to the theoretical energy. The energy consumed for transmitting a frame of size 240 × 320 is calculated using (13) and (14) as

$$E_f = (240 * 320 * 8) * 0.23\ \mu J \approx 141\ mJ \tag{15}$$

Table 3
Experimental results for DWT–DCT hybrid based VCS framework in terms of PSNR.

Video sequence	Measurements	Experimental results Average PSNR (dB)		
		Gaussian matrix	Hybrid matrix	Combination matrix
Xylophone	20	28.9	29.4	29.2
	30	31	31.9	31.5
	40	32.5	33.8	32.6
Akiyo	20	31.3	31.6	31.4
	30	32.5	33.4	32.8
	40	33.6	33.7	33.9
Football	20	26.1	26.4	27.1
	30	28.8	28.5	27.9
	40	29.6	31.2	30.5

Table 4
Transmission energy for various numbers of measurements.

Measurements per frame	Transmission energy (mJ)	
	Theoretical values	Practical values
24,000	44	46
36,000	66	69
48,000	88	92

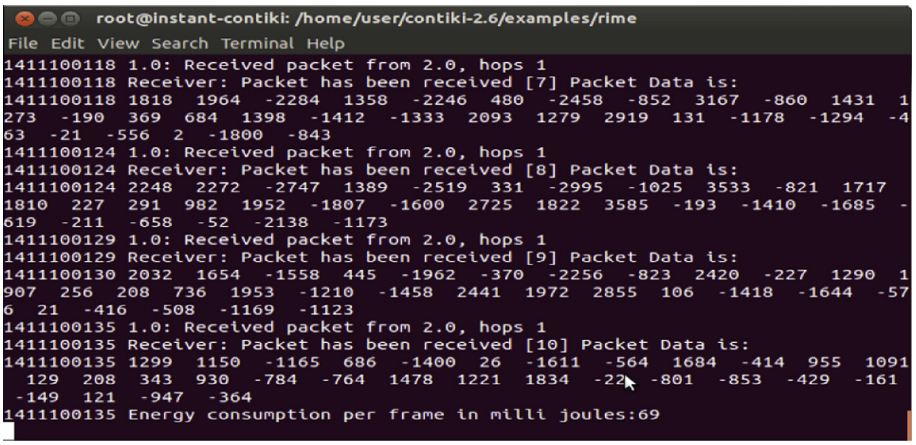

Fig. 7. Screenshot displaying received measurements of a single frame for $M = 36,000$.

On average, the energy consumed for transmitting the measurements of a single frame is 50% less when compared to the theoretical energy consumption for raw frame transmission. Fig. 7 shows the screenshot of the measurements received for a single frame in the Contiki OS platform. The energy consumption per frame calculated using the powertrace tool is also shown in the screenshot.

5.2.2. End-to-end delay

The end-to-end delay in the network is calculated by computing the time taken to transmit the measurements and the propagation delay of the link. The transmission delay is obtained from the powertrace tool. In WSN, the sensor nodes are only a short distance apart, and hence, the propagation delay is negligible.

The end-to-end delay ($d_{end-end}$) for a single hop is calculated using (16)

$$d_{end-end} = d_{trans} + d_{prop} \tag{16}$$

In the case of a multi-hop network, the end-to end-delay in the network is the sum of the delay of each link

$$\text{Delay in multi-hop network} = H * d_{end-end} \tag{17}$$

where H represents the number of hops between the source node and the destination node.

Table 5
End-to-end delay in the network for different measurements.

M_1	Average delay (ms) Distance = 5 m	
	Single hop	Two hop
20	0.64	1.28
30	0.96	1.92
40	1.28	2.56

From Table 5, it is observed that the end-to-end delay increases as the number of bits and hops increases. The average end-to-end delay for transmitting the measurements is 52% less when compared to the raw frame transmission, which has an average delay of 2 ms for a single hop. For timely delivery of data, the delay in the network must be small, which proves that the proposed VCS technique is efficient for WMSN.

5.3. Complexity analysis

The complexity is analysed in terms of storage, the energy for generating the measurement matrix and the percentage of reduction in samples. Storage complexity is analysed by calculating the number of elements to be stored in the WSN node for generating the matrix, and energy complexity is analysed by computing the energy consumed for generating the measurement matrix.

5.3.1. Storage complexity

Generating a Combination matrix of size $M_1 \times 64$ requires $((2 * 8) - 1)$ elements for the Toeplitz matrix and $M_1 * 8$ elements for the Gaussian matrix. Generating a Hybrid matrix of size $M_1 \times 64$ requires two elements for the Toeplitz matrix (with entries -1 & +1) and two elements for the Binary matrix (with entries 0 & 1). Table 6 shows the comparison of Gaussian, Combination and Hybrid matrices in terms of the number of elements required to generate the matrix.

From Table 6, it is evident that the number of elements required to generate the measurement matrix increases with the number of measurements. The Combination matrix requires fewer elements compared with the Gaussian matrix and the Hybrid matrix requires many fewer elements when compared with the Gaussian matrix and the Combination matrix.

5.3.2. Energy complexity

The energy consumed for generating the measurement matrix is computed using (18)–(20).

5.3.2.1. Gaussian matrix. To generate a Gaussian matrix of size $M_1 * 64$, $M_1 * 64$ read and write operations are required.

$$E_G = (M_1 * 64) * e_{read} + (M_1 * 64) * e_{write} \tag{18}$$

5.3.2.2. Combination matrix. To generate a Combination matrix of size 20×64, 47 read and write and 1280 multiplication operations are required, as shown in (19)

$$E_C = 47 * e_{read} + 47 * e_{write} + 1280 * e_{mul} \tag{19}$$

5.3.2.3. Hybrid matrix. To generate a hybrid matrix of size 20×64, four read and write operations are required, as shown in (20)

$$E_H = 4 * e_{read} + 4 * e_{write} \tag{20}$$

The energy consumption analysis is performed using the characteristics of the TelosB node. The energy consumption for the basic operations in TelosB is provided in Table 7.

The energy consumed in generating the measurement matrix is computed by substituting the values given in Table 7 in (18)–(20). The energy consumption of the Gaussian and the proposed sensing matrices for $M_1 = 20, 30, 40$ is given in Table 8.

Table 6
Comparison of Gaussian, Combination and Hybrid matrices in terms of elements to be stored.

Size of measurement matrix	Elements required to generate matrix (Bytes)			Percentage of reduction in elements (%)
	Gaussian matrix	Combination matrix	Hybrid matrix	
20×64	1280	47	4	98
30×64	1920	63	4	98.3
40×64	2560	79	4	98.4

Table 7
Energy consumed by the MSP430 microcontroller [24].

Variable	Operation performed	Energy consumed
e_{mul}	Multiplication over 1 byte	2.92 nJ
e_{read}	Reads 1 byte from the flash memory	8.2 μJ
e_{write}	Writes 1 byte to the flash memory	34.9 μJ

Table 8
Energy consumed for generating the measurement matrix.

Size of measurement matrix	Energy consumed (mJ)			Percentage of reduction in energy (%)
	Gaussian matrix	Combination matrix	Hybrid matrix	
20×64	55	2	0.17	98
30×64	83	2.72	0.17	98.25
40×64	110	3.4	0.17	98.37

From Table 8, it is observed that on average, the energy consumed for generating the Combination matrix and the Hybrid matrix is approximately 98% less compared with the existing Gaussian measurement matrix. The results show that the Combination matrix and Hybrid matrix prove to be memory and energy efficient for VCS, thus making it more suitable for surveillance applications in WMSN.

5.3.3. Percentage of reduction in samples

The percentage of reduction in samples (P_s) is calculated for the DCT, DWT, and DWT–DCT hybrid approaches using (21)

$$P_s = \left(1 - \frac{P}{Q}\right) * 100\% \tag{21}$$

where P represents the number of measurements per 16×16 block and Q represents the number of samples per 16×16 block. Fig. 8 depicts the comparison of the DCT, DWT and DWT–DCT hybrid approaches in terms of PSNR and reduction of samples per frame (P_{sf}).

From Fig. 8(a), it is observed that upon increasing the number of measurements, the PSNR also increases, but the percentage of reduction in the samples decreases. It is also observed that the DCT approach achieves higher compression rate and

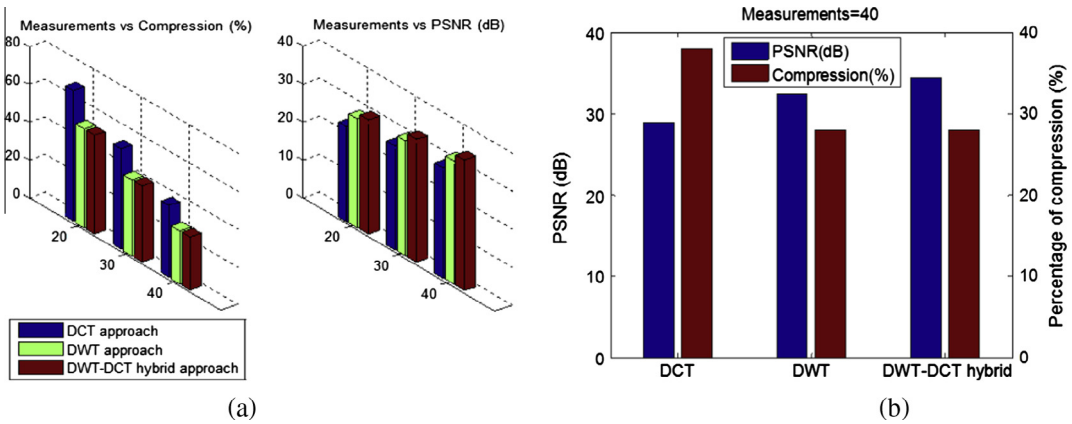

Fig. 8. Comparison of DCT, DWT and DWT–DCT approaches: (a) measurements vs PSNR and measurements vs percentage of compression (b) PSNR vs percentage of compression.

Table 9
Percentage of reduction in samples per frame (P_{sf}) for DCT, DWT, and DWT–DCT hybrid approaches.

M_1	DCT		DWT		DWT–DCT	
	P_{sf} (%)	PSNR (dB)	P_{sf} (%)	PSNR (dB)	P_{sf} (%)	PSNR (dB)
20	69	24.8	52	28.6	52	30.1
30	53	27.3	40	30.2	40	32.3
40	38	28.9	28	32.4	28	34.4

the DWT–DCT hybrid approach achieves better PSNR compared with the DWT approach for the same percentage of compression. Fig. 8 (b) shows a comparison of the DCT, DWT and DWT–DCT approaches in terms of PSNR and percentage of compression. The values observed from Fig. 8(b) are given in Table 9.

From Table 9 and Fig. 8(b), it is evident that DWT achieves better PSNR when compared with DCT. It is also observed that the DWT–DCT approach achieves a PSNR of approximately 30 dB with 52% reduction in samples, whereas for the DWT approach, it is only a 40% reduction. It is clear that the DWT–DCT hybrid approach achieves better quality even at higher compression rates compared to the DCT and DWT approaches.

6. Conclusions and future work

A VCS framework based on the DWT–DCT hybrid approach is implemented using the proposed Combination matrix and Hybrid matrix. The performance of the measurement matrix is evaluated in terms of PSNR, storage and energy complexity. Transmission energy and delay are analysed experimentally with the help of TelosB nodes. From the results, it is concluded that the proposed matrices are memory and energy efficient while maintaining an acceptable range of PSNR. The DWT–DCT hybrid approach yields better PSNR for fewer measurements when compared with standalone DCT and DWT approaches. The experimental results show that on average, the transmission energy is 50% less and the end-to-end delay is 52% less when compared with raw frame transmission.

Conventional wireless sensor nodes do not have adequate memory, power and bandwidth for multimedia applications. The future work is to develop customized hardware that can incorporate the VCS technique. The transmission energy can be further reduced by encoding the measurements before transmission. The customized hardware should have a camera module to capture the scene, a processor with high processing speed and adequate memory to process the video using the CS technique. To capture and process only the scene of interest or to detect anomalies, the camera module must be associated with a PIR sensor, which triggers the camera only upon interruption. This highly efficient hardware could then be used for video surveillance applications in WMSN.

References

[1] Candès Emmanuel J, Wakin Michael B. An introduction to compressive sampling. IEEE Signal Process Mag 2008;25(2):21–30.
[2] Akyildiz Ian F, Melodia Tommaro, Chowdury Kaushik R. A survey on wireless multimedia sensor networks. IEEE Wireless Commun 2007;51(4):921–60.
[3] Romberg J. Imaging via compressive sampling. IEEE Signal Process Mag 2008;25(2):14–20.
[4] Duarte Marco F, Eldar Yonina C. Structured compressed sensing: from Theory to applications. IEEE Trans Signal Process 2011;59(9):4053–85.
[5] Gan Lu. Block based compressed sensing of natural images. In: 15th international conference on digital signal processing; 2007. p. 403–6.
[6] Han Bing, Wu Feng, Wu Depang. Image representation by compressed sensing. In: 15th IEEE international conference on image processing; 2008. p. 1344–7.
[7] Yu Lei, Barbot Jean-Pierre, Zheng Gang, Sun Hong. Toeplitz-structured chaotic sensing matrix for compressive sensing. In: Communication systems networks and digital signal processing (CSNDSP); 2010. p. 229–33.
[8] Gleichman Sican, Eldar Yonina C. Blind compressed sensing. IEEE Trans Inform Theory 2011;57(10):6958–75.
[9] Jain Prateek, Tewari Ambuj. Orthogonal matching pursuit with replacement. In: Proceedings of advances in neural information processing systems; 2011. p. 1215–23.
[10] Schaeffer Hayden, Yang Yi, Osher Stanley. Real-time adaptive video compressive sensing. In: UCLA CAM, Tech. Rep.; 2013, p. 1–22.
[11] Schenkel Markus B, Luo Chong, Wu Feng, Frossard Pascel. Compressed sensing based video multicast. In: SPIE visual communications and image processing; 2010. p. 77441H-77441H.
[12] Jiang Hong, Deng Wei, Shen Zuowei. Surveillance video processing using compressive sensing. arXiv preprint arXiv:1302.1942; 2013. p. 201–14.
[13] Balouchestani Mohammadreza, Raahenufar Kaamran, Krishnan Sridhar. Compressed Sensing in wireless sensor networks: survey. Can J Multimedia Wireless Networks 2011;2(1):1–4.
[14] Pudlewski Scott, Prasanna Arvind, Melodia Tommaso. Compressed-sensing-enabled video streaming for wireless multimedia sensor networks. IEEE Trans Mobile Comput 2012;11(6):1060–72.
[15] Zhang Jian, Zhao Debin, Zhao Chen, Xiong Ruiqin, Ma Siwei, Gao Wen. Image compressive sensing recovery via collaborative sparsity. IEEE J Emerg Sel Top Circ Syst 2012;2(3):380–91.
[16] Zhang Jian, Zhao Debin, Jiang Feng, Gao Wen. Structural group sparse representation for image compressive sensing recovery. In: Data compression conference (DCC). IEEE; 2013. p. 331–40.
[17] Sungkwang Mun, Fowler James E. Residual reconstruction for block-based compressed sensing of video. In: Data compression conference (DCC). IEEE; 2011. p. 183–92.
[18] Liu Ying, Li Ming, Pados Dimitrios A. Motion-aware decoding of compressed-sensed video. IEEE Trans Circ Syst Video Technol 2013;23(3):438–44.
[19] Jha Mithilesh Kumar, Lall Brejesh, Roy Sumantra Dutta. DEMD-based image compression scheme in a compressive sensing framework. J Pattern Recognit Res 2014:1–14.
[20] Candes E, Tao T. Decoding by linear programming. IEEE Trans Inform Theory 2005;51(12):4203–15.
[21] Haghighat Mohammad Bagher Akbari, Ali Aghagolzadeh, Seyedarabi Hadi. Multi-focus image fusion for visual sensor networks in DCT domain. Comput Electr Eng 2011;37(5):789–97.
[22] Dua S, Acharya R, Ng EYK. Computational analysis of the human eye with applications. World Scientific; 2011.
[23] Tropp Joel A, Gilbert Anna C. Signal recovery from random measurements via orthogonal matching pursuit. IEEE Trans Inform Theory 2007;53(12):4655–66.
[24] Amiri Moslem. Measurements of energy consumption and execution time of different operations on Tmote Sky sensor nodes. 2010.
[25] Available from: <http://www.contiki-os.org>.
[26] Dunkels A, Eriksson J, Finne N, Tsiftes N. Powertrace. network-level power profiling for lowpower wireless networks. Technical Report T2011:05. SICS.

Aasha Nandhini Sukumaran is a Junior Research Fellow at SSN College of Engineering, India. She received the B.E. degree in Electronics and Communication Engineering from Rajalakshmi Engineering College, India, in 2010 and the M.E. degree in Communication Systems from SSN College of Engineering, India, in 2012. Her research interests include security issues and compressive sensing in wireless sensor networks.

Radha Sankararajan is a Professor and Head of the department of ECE, SSN College of Engineering, Chennai, India. She is the recipient of the IETE – S.K. Mitra Memorial Award (2006) from the IETE Council of India and the CTS – SSN Best Faculty Award (2007, 2009) for outstanding academic performance. Her research interests include security, architecture issues of mobile ad hoc networks, WSNs and cognitive radios.

Kishore Rajendiran is working as an Associate Prof., in the department of ECE, SSN College of Engineering, Chennai, India. He received the B.E. degree in Electronics and Communication Engineering from Vellore Engineering College, Madras University, in 1998, the M.Tech. degree in Communication Systems from Pondicherry University, in 2003 and the Ph.D. degree from Anna University, in 2012. His research interests include security issues in wireless networks, compressive sensing and image fusion.