

Received April 2, 2020, accepted April 24, 2020, date of publication April 28, 2020, date of current version May 12, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2991049

# Hierarchical Trellis Based Decoder for Total Variation Sequence Detection (TVSD) in Space-Time Trellis Coded (STTC) Wireless Image/ Video Communication

ANKIT KUDESHIA<sup>1</sup>, (Student Member, IEEE), MOHD ABBAS ZAIDI<sup>2</sup>,  
ADITYA K. JAGANNATHAM<sup>1</sup>, (Member, IEEE), AND CHANDRA PRAKASH<sup>3</sup>

<sup>1</sup>Department of Electrical Engineering, Indian Institute of Technology Kanpur, Kanpur 208016, India

<sup>2</sup>Next AI Solutions Lab, Samsung Research, Samsung Electronics, Seoul 06765, South Korea

<sup>3</sup>Space Applications Center (SAC) - Indian Space Research Organization, Ahmedabad 380015, India

Corresponding author: Ankit Kudeshia (ankitkud@iitk.ac.in)

This work was carried out at Indian Institute of Technology (IIT) Kanpur and was supported in part by the TCS Research Scholarship Program, IIT Kanpur, and in part by the ISRO-IITK Space Technology Cell under Grant STC/EE/2017204.

**ABSTRACT** This paper proposes a robust reconstruction scheme for image/ video transmission in space-time trellis coding (STTC) based MIMO wireless multimedia sensor networks (WMSN). The information bits of the image/ video stream are modulated using STTC prior to transmission over the MIMO wireless channel between the multimedia sensors and the cluster head. At the receiver, a novel total variation sequence detector (TVSD) is developed that employs an anisotropic total variation (TV) regularized cost function to leverage the bounded variation property of the image/ video stream, in addition to the coding and diversity advantages of the STTC, for improved image/ video recovery. For reconstruction, a novel bi-layered hierarchical trellis based modified Viterbi decoder is developed to decode the optimal bit sequence corresponding to the least TV cost function using appropriately modified node and branch metrics. Simulation results using several test images and video sequences demonstrate the improved reconstruction performance of the proposed scheme, especially in the low SNR regime (0 – 10 dB), in comparison to the conventional maximum likelihood (ML) decoder, which makes it ideally suited for implementation in practical WMSN.

**INDEX TERMS** Image/video communication, MIMO, space-time trellis codes, total variation, Viterbi algorithm, wireless multimedia sensor networks.

## I. INTRODUCTION

Wireless multimedia sensor networks (WMSN) have gained popularity due to their applicability in a diverse spectrum of applications such as surveillance, law enforcement, industrial monitoring, health care, agriculture, UAVs, among others [1], [2]. The fabrication of multimedia capable sensor nodes has been made possible by the recent advances in the development of low-cost and sophisticated CMOS-based imaging and audio sensors [3]. In contrast to the traditional wireless sensor nodes [4], [5] that sense, process and

transmit measurements of parameters like temperature, pressure, humidity etc, WMSN generate and process multimedia content comprising of audio, images and videos streams. A typical WMSN is shown in Fig. 1. Given the stringent computational and battery constraints [6]–[8] of the miniature sensor nodes, it is imperative to develop low complexity image/ video recovery techniques at the fusion center, to achieve high quality image/ video reconstruction at the base station/fusion center. This is especially important in WMSN since the fading nature of the wireless channel can lead to a significant degradation in the quality of reconstruction [9]. A brief review of the existing research in this context is presented below.

The associate editor coordinating the review of this manuscript and approving it for publication was Xingwang Li<sup>1</sup>.

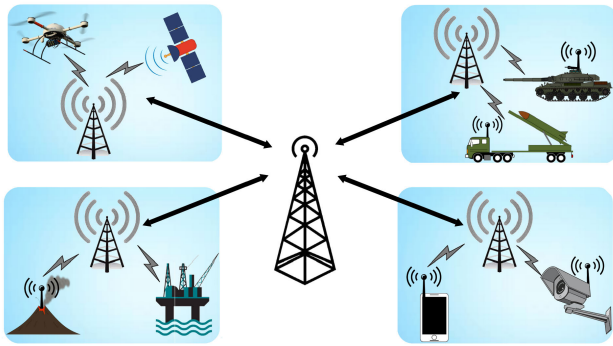


FIGURE 1. Wireless multimedia sensor network architecture.

### A. REVIEW OF WORKS IN EXISTING LITERATURE

Several techniques have been proposed in the existing literature for reliable image/ video reconstruction in wireless sensor networks (WSN). For instance, the authors in [10] proposed a network adaptive image/ video transmission scheme for sensor networks based on quality-scalable coding for multipath transmission. The work in [11] proposes a real time error correction scheme for WSN based on predictive models that exploit the temporal correlation in sensor data. A significant advantage of the scheme proposed therein is that it neither requires any additional resources nor increases the computational complexity at the sensor nodes. The multiple description (MD) source coding based approach in [12]–[14] uses a correlating block transform at the application layer, thus introducing controlled redundancy in the transmitted data to enhance the performance of multimedia streaming in wireless networks. However, such approaches are not suitable for the WMSN due to the increased bandwidth overhead caused by the redundancy in the bit stream, which can be significant for a WMSN with a large number of transmitting sensor nodes. Traditional schemes such as automatic repeat request (ARQ) [15] that deal with packet losses in sensor networks also lead to a decrease in the effective spectral efficiency. Interestingly, the work in [16] demonstrates that coding schemes over multiple parallel channels e.g. frequency bands, antennas or time, which leverage the diversity properties of the fading channel, are more efficient for Rayleigh fading channels in comparison to the source coding schemes.

The development of multiple-input and multiple-output (MIMO) technology that supports increased data rates for the same transmit power, coupled with its diversity advantages [17]–[19] for increased reliability, has led to a radical improvement in the quality and efficiency of wireless communication. Its energy and spectral efficiency naturally renders it attractive for implementation in WMSN [20]. MIMO wireless systems have also been employed for efficient and reliable video transmission in various recent works as in [21]–[23]. Further, STTC [24] is a popular technique that can yield additional coding and diversity benefits in such systems [25], in turn leading to a decrease in the error rate of communication. An STTC typically employs a low

complexity generator matrix for encoding the information bits [26] at the transmitter, which makes it a better choice for use in WMSN. Decoding at the receiver can be performed via the efficient Viterbi algorithm [27] based trellis decoder. Hence, STTC has been popular in literature for robust communication in multi-antenna systems. However the performance of STTC specifically for image/ video transmission can be significantly enhanced by leveraging the bounded variation property [28] that is a fundamental characteristic of multimedia signals. TV regularization based image/ video reconstruction schemes that exploit the BV property have gained popularity for several image/ video processing applications such as noise removal [29], deblurring [30]–[32], interpolation [33], inpainting [34], [35], super resolution (SR) based reconstruction [36]–[39], structure decomposition [40] etc. This is due to the edge preserving property [41] of the  $l_1$  norm based anisotropic TV regularization term that can lead to a significant improvement in reconstruction quality [42]. Previous works [43], [44] have described error resilient techniques for image/ video transmission in wireless networks that exploit their bounded variation property employing a novel Viterbi algorithm based TVSD technique for reconstruction. Furthermore, this is well suited for sensor networks since it does not increase the processing complexity at the sensors. This work therefore proposes a MIMO-STTC encoder for multimedia data followed by a novel joint STTC-TV decoder at the cluster head for improving the quality of image/ video recovery. The various contributions of this work are described next.

### B. CONTRIBUTIONS OF THIS PAPER

This work develops a framework for STTC-based transmission and detection of image/ video frames. At the receiver, a TV regularized cost function is proposed for improved reconstruction of the received frames. Another contribution of this work is a novel hierarchical trellis structure, comprising of two layers, which is developed for optimal decoding of the received signal stream. The coded component (CC) layer of the trellis comprises of nodes that correspond to various possible quantization levels and pixel intensities for entropy coded and uncompressed streams, respectively. The other layer termed the STTC symbol (ST) layer has nodes corresponding to the encoder states of the conventional STTC trellis. A novel modified Viterbi decoder, with suitably defined branch and state metrics, is developed to obtain the optimal decoded stream for the above hierarchical trellis which also incorporates an  $l_1$  norm based TV regularization term. Simulation results demonstrate the improved STTC-TV reconstruction in comparison to the conventional STTC-ML decoder.

### C. ORGANIZATION OF THE PAPER

The outline of the paper is as follows. Section II presents an STTC coded wireless communication framework for image/ video stream transmission in a WMSN followed by a brief description of the conventional STTC-ML image/ video

decoder. The proposed TV regularization based STTC-TV scheme for robust image/ video recovery and the hierarchical trellis based Viterbi decoder for joint STTC and TV decoding are proposed in section III. Section IV describes the simulation setup for the proposed image/ video reconstruction scheme together with the various visual and numerical results and also compares them with those obtained from conventional recovery schemes. This is followed by the conclusion in section V.

**D. SYMBOLS AND NOTATIONS**

The various notations used in this paper are as follows. Vectors and the matrices are denoted by lower case boldface letters  $\mathbf{x}$ , upper case boldface letters  $\mathbf{X}$  respectively. Scalars are denoted by lowercase letters  $x$ .  $\mathbb{E}\{X\}$  denotes the expectation operator over the random variable  $X$  that computes its stochastic average with respect to its probability density function (PDF). The estimate of the signal vector  $\mathbf{x}$  is denoted by  $\hat{\mathbf{x}}$ .  $\mathbb{R}$ ,  $\mathbb{C}$  and  $\mathbb{Z}$  denote the sets of real numbers, complex numbers and integers, respectively. An identity matrix of size  $N \times N$  is represented by  $\mathbf{I}_N$ . The cardinality of a set  $\mathcal{A}$  is denoted by  $|\mathcal{A}|$ .  $\mathcal{N}[\boldsymbol{\mu}, \mathbf{C}]$  denotes the the multivariate Gaussian distribution with mean  $\boldsymbol{\mu}$  and covariance matrix  $\mathbf{C}$ .

**II. SPACE-TIME TRELLIS CODING BASED WIRELESS IMAGE/WWW/VIDEO TRANSMISSION SYSTEM**

Consider a digital video stream consisting of  $P$  frames where the  $p$ th frame denoted by  $A^p \in \mathbb{Z}^{R \times C}$  for  $1 \leq p \leq P$ , comprises of  $R$  rows and  $C$  columns of pixels. Note that this can also be used to model image transmission by setting  $P = 1$ . The quantity  $A^p(i, j)$  represents the intensity level of the corresponding pixel in the  $i$ th row and the  $j$ th column of the  $p$ th frame. The frames can potentially be encoded using a pertinent source coding scheme such as JPEG for images, MPEG for videos etc [45], [46] to compress the stream in order to meet the limited bandwidth constraint at the sensor node. Towards this end, each frame is divided into blocks of size  $b \times b$  pixels resulting in  $C_s = \frac{C}{b}$  and  $R_s = \frac{R}{b}$  blocks in each row and column, respectively. Typically  $b = 8$  for various popular schemes such as JPEG and MPEG. Setting  $b = 1$  corresponds to the transmission of the uncompressed image/ video frame. The  $(r, l)$ th block of the video frame, where  $1 \leq r \leq R_s$  and  $1 \leq l \leq C_s$ , is encoded into  $Q_{r,l}^p$  coded components (e.g. DC, AC components). For digital modulation, prior to transmission in a digital wireless system, the  $k$ th coded component of the block is further subsequently binary coded with  $B_{r,l}^{p,k}$  information bits. Let the sequence of  $B_{r,l}^{p,k}$  information bits obtained above be denoted by the vector

$$\mathbf{x}_{r,l}^{p,k} = \left[ \mathbf{x}_{r,l}^{p,k}(1), \mathbf{x}_{r,l}^{p,k}(2), \dots, \mathbf{x}_{r,l}^{p,k}(B_{r,l}^{p,k}) \right]^T, \quad (1)$$

where each  $\mathbf{x}_{r,l}^{p,k}(i) \in \{0, 1\}$ . The coded bits corresponding to the  $k$ th coded components of all the blocks in the  $r$ th row are

further concatenated as the vector

$$\mathbf{x}_r^{p,k} = \begin{bmatrix} \mathbf{x}_{r,1}^{p,k} \\ \mathbf{x}_{r,2}^{p,k} \\ \vdots \\ \mathbf{x}_{r,C_s}^{p,k} \end{bmatrix}. \quad (2)$$

The next section describes the STTC for encoding the information bit vector sequence  $\mathbf{x}_r^{p,k}$  in a MIMO wireless communication system.

**A. STTC FOR IMAGE/WWW/VIDEO TRANSMISSION**

Consider a MIMO system with  $N_t$  transmit and  $N_r$  receive antennas [47] employing an  $M$ -ary digital symbol constellation for modulation of the symbols. For instance,  $M = 2, 4$  represent BPSK, QPSK respectively. Let  $m = \log_2 M$  denote the number of bits per symbol. STTC is subsequently applied on the information bits using an appropriate generator matrix to obtain the space-time coded symbols that are transmitted over the MIMO wireless channel, the procedure for which can be described as follows. The coded video bit stream  $\mathbf{x}_r^{p,k}$  is partitioned into blocks of  $m + s$  bits, where  $s$  denotes the number of memory bits, to obtain the input bit vector  $\mathbf{u}_{r,t}^{p,k} \in \{0, 1\}^{(m+s) \times 1}$  defined below

$$\mathbf{u}_{r,t}^{p,k} = \left[ \mathbf{x}_r^{p,k}(mt), \dots, \mathbf{x}_r^{p,k}(m(t-1) - s + 1) \right]^T, \quad (3)$$

which is acted on by the encoder at time instant  $t$ . Hence, the number of encoder memory states is  $V = 2^s$ . Let  $\mathbf{G} \in \mathbb{R}^{(m+s) \times N_t}$  denote the generator matrix for the STTC. Using the procedure given in [48], the coded symbol index vector  $\mathbf{d}_{r,t}^{p,k} \in \mathbb{Z}_M^{N_t \times 1}$  corresponding to the  $M$ -ary constellation at time instant  $t$  is generated as

$$\mathbf{d}_{r,t}^{p,k} = \left( \left( \mathbf{u}_{r,t}^{p,k} \right)^T \mathbf{G} \pmod{M} \right)^T, \quad (4)$$

where  $\mathbb{Z}_M = \{0, \dots, M - 1\}$  is the set of possible  $M$ -ary symbol indices. The initial encoder state is assumed to be zero i.e.  $\mathbf{x}_r^{p,k}(t) = 0$  for  $t \leq 0$ . Thus the process above transforms the  $m + s$  coded image/ video bit stream at time  $t$  into  $N_t$  trellis coded symbol indices. Let the  $i$ th element  $\xi_{r,t}^{p,k}(i)$  of the transmit symbol vector  $\boldsymbol{\xi}_{r,t}^{p,k} \in \mathcal{M}^{N_t \times 1}$  be obtained from the corresponding coded symbol index  $\mathbf{d}_{r,t}^{p,k}(i)$  via the modulation operation

$$\xi_{r,t}^{p,k}(i) = \mathcal{F}_M \left( \mathbf{d}_{r,t}^{p,k}(i) \right); \quad 1 \leq i \leq N_t, \quad (5)$$

where  $\mathcal{F}_M : \{0, 1, \dots, M - 1\} \rightarrow \mathcal{M}$  is the modulation operator. For instance, this is defined as  $\mathcal{F}_M(y) = \sqrt{P_s} \exp \left( \frac{2\pi y}{M} \right)$  for the  $M$ -ary PSK modulation scheme, where  $P_s$  denotes the symbol power at the transmitter. The set of symbols in the  $M$ -ary constellation is denoted by  $\mathcal{M} = \{u_1, u_2, \dots, u_M\}$ , with each  $u_i \in \mathcal{M} \subset \mathbb{C}$ . Let the transmit

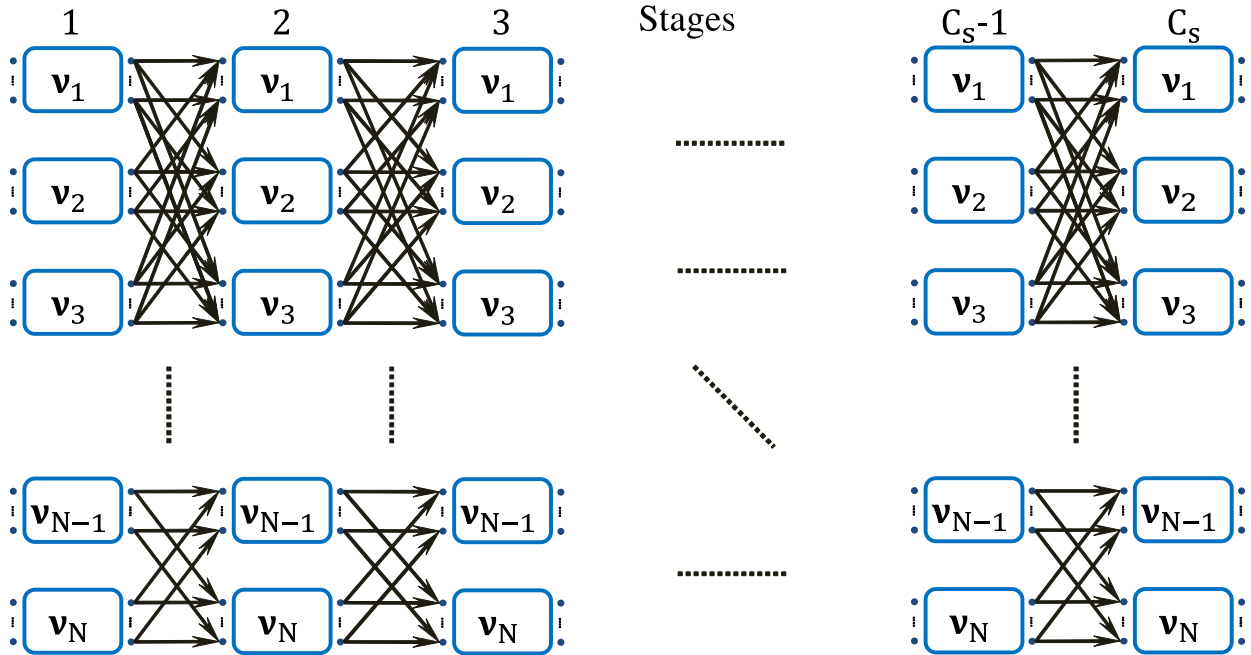


FIGURE 2. CC-Layer of the trellis structure for the proposed STTC-TV Viterbi decoder.

symbol vector sequence denoted by  $\Xi_r^{p,k} \in \mathcal{M}^{N_t \times T}$  be defined as

$$\Xi_r^{p,k} = [\xi_{r,1}^{p,k}, \xi_{r,2}^{p,k}, \dots, \xi_{r,T}^{p,k}], \quad (6)$$

where  $T$  is the STTC block length corresponding to the transmission of the information bits in  $\mathbf{x}_r^{p,k}$ . The MIMO wireless channel model for STTC coded wireless transmission of image/ video frames is detailed next.

### B. MIMO WIRELESS COMMUNICATION SYSTEM MODEL

The transmit symbol vector  $\xi_{r,t}^{p,k}$  generated as shown in (5) is subsequently transmitted over an  $N_r \times N_t$  MIMO wireless channel. Let  $\mathbf{y}_{r,t}^{p,k}$  denote the corresponding received symbol vector, which is related to the transmit symbol vector  $\xi_{r,t}^{p,k}$  as [49]

$$\mathbf{y}_{r,t}^{p,k} = \mathbf{H}_{r,t}^{p,k} \xi_{r,t}^{p,k} + \boldsymbol{\eta}_{r,t}^{p,k}, \quad (7)$$

where  $\mathbf{H}_{r,t}^{p,k} \in \mathbb{C}^{N_r \times N_t}$  represents the MIMO wireless channel matrix with  $\mathbf{H}_{r,t}^{p,k}(i,j)$  denoting the wireless channel fading coefficient between the  $j$ th transmit and the  $i$ th receive antenna at time instant  $t$ . The quantity  $\boldsymbol{\eta}_{r,t}^{p,k} \in \mathbb{C}^{N_r \times 1}$  represents the symmetric complex additive white Gaussian noise vector at the receiver with zero mean and covariance matrix  $\sigma_\eta^2 \mathbf{I}_{N_r}$ , where  $\sigma_\eta^2$  denotes the noise variance. The conventional trellis based ML decoder for STTC coded image/ video transmission over the MIMO wireless channel is described next, followed by the proposed scheme for improved decoding.

Stages

.....

.....

.....

.....

.....

### C. CONVENTIONAL STTC-ML IMAGE/VIDEO/VIDEO DECODER

Conventionally, the symbol vector sequence estimate  $\hat{\Xi}_r^{p,k} \in \mathcal{M}^{N_t \times T}$ , defined as

$$\hat{\Xi}_r^{p,k} = [\hat{\xi}_{r,1}^{p,k}, \hat{\xi}_{r,2}^{p,k}, \dots, \hat{\xi}_{r,T}^{p,k}], \quad (8)$$

for the STTC coded video stream  $\Xi_r^{p,k}$ , is obtained using the Viterbi algorithm based ML decoder [50] that is described as follows. Consider a trellis structure with  $V = 2^s$  states and  $T$  stages. The  $i$ th state in each stage is denoted by  $s_i$ , which corresponds to one of the  $V$  encoder states of the STTC. Each transition from state  $s_i$  to state  $s_j$  corresponds to the transmission of  $m$  bits of the coded image/ video bit stream at time  $t$  and the corresponding STTC encoded symbol vector is determined using the generator matrix  $\mathbf{G}$ . The Viterbi algorithm at the receiver [51] is employed to decode the state sequence that has the maximum likelihood property through the trellis corresponding to the received signal vector block  $\mathbf{Y}_r^{p,k} \in \mathbb{C}^{N_r \times T}$  defined as

$$\mathbf{Y}_r^{p,k} = [\mathbf{y}_{r,1}^{p,k}, \mathbf{y}_{r,2}^{p,k}, \dots, \mathbf{y}_{r,T}^{p,k}]. \quad (9)$$

Let the STTC encoded symbol vector corresponding to the transition from state  $s_i$  to state  $s_j$  be denoted by  $\tilde{\xi}_{i,j}$  and the set of valid next states from the memory state  $s_i$  for different input symbols be denoted by  $\mathcal{S}_i$ . It can be noted that the vector  $\tilde{\xi}_{i,j}$  and the set  $\mathcal{S}_i$  can be determined using either the generator matrix or state diagram of the STTC code. Example generator matrices and state transitions in a trellis for the STTC used in simulation are shown in Fig. 2 and Fig. 3 respectively. The branch metric corresponding to the

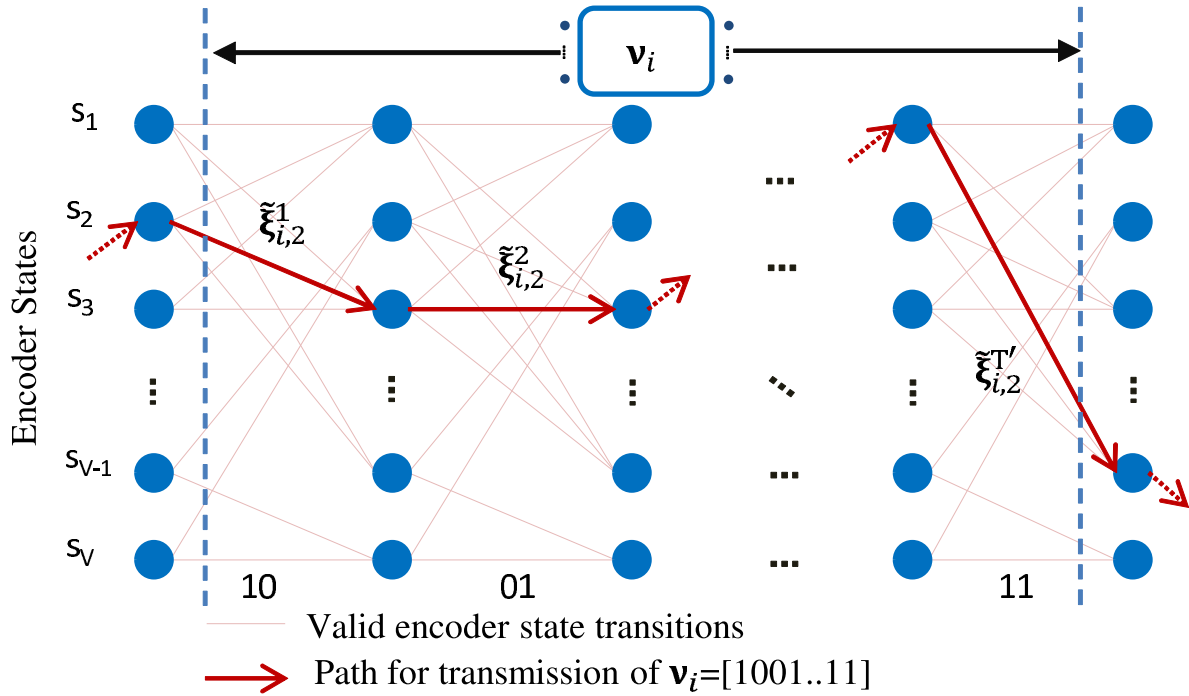


FIGURE 3. ST-layer: Illustration of procedure for finding transmit symbol vector sequence corresponding to transmission of information bits  $\mathbf{v}_i$  with initial encoder state  $s_2$ .

state transition from state  $s_i$  to state  $s_j$  at instant  $t$  is defined as

$$\mu_{r,t}^{p,k}(s_i, s_j) = \begin{cases} \left\| \mathbf{y}_{r,t}^{p,k} - \mathbf{H}_{r,t}^{p,k} \tilde{\xi}_{i,j} \right\|_2^2, & \text{if } (t \neq 1, s_j \in \mathcal{S}_i) \\ & \text{or } (t = 1, i = 1) \\ \infty, & \text{otherwise.} \end{cases} \quad (10)$$

The initial state of the STTC encoder at  $t = 1$  is assumed to be  $s_1$ . The accumulated path metric is initialized as  $\alpha_{r,1}^{p,k}(s_j) = \mu_{r,1}^{p,k}(s_1, s_j), 1 \leq j \leq n$ , and is subsequently computed recursively for  $t > 2$  by minimizing the sum of the accumulated metric at  $t - 1$  and the branch metric at time  $t$  as given in the equation below

$$\alpha_{r,t}^{p,k}(s_j) = \min_{1 \leq i \leq n} \left( \alpha_{r,t-1}^{p,k}(s_i) + \mu_{r,t}^{p,k}(s_i, s_j) \right). \quad (11)$$

Let the survivor state index corresponding to state  $s_j$  at stage  $t$  be denoted by  $\mathbf{v}_r^{p,k}(j, t)$ , which is initialized as  $\mathbf{v}_r^{p,k}(j, t) = 1, \forall j$  for  $t = 1$ . Further, for stages  $t \geq 2$ , it is determined for each state  $s_j$  at stage  $t$  of the trellis as

$$\mathbf{v}_r^{p,k}(j, t) = \arg \min_{1 \leq i \leq n} \left( \alpha_{r,t-1}^{p,k}(s_i) + \mu_{r,t}^{p,k}(s_i, s_j) \right). \quad (12)$$

The ML decoded sequence corresponds to the trellis path terminating at the state with corresponding to the minimum accumulated metric at  $t = T$ . The state at  $t = T$  is determined as shown below

$$\mathbf{q}_r^{p,k}(T) = \arg \min_{1 \leq i \leq n} \alpha_{r,T}^{p,k}(s_i). \quad (13)$$

The decoded state sequence is obtained by back tracking through the survivor states in the trellis using the survivor state index matrix  $\mathbf{v}_r^{p,k}$  as

$$\mathbf{q}_r^{p,k}(t) = \mathbf{v} \left( \mathbf{q}_r^{p,k}(t+1), t+1 \right) \quad \forall 1 \leq t \leq T-1. \quad (14)$$

The optimal state sequence estimate  $\hat{\mathbf{s}}_r^{p,k}$  is in turn obtained from the index sequence  $\mathbf{q}_r^{p,k}$  as

$$\hat{\mathbf{s}}_r^{p,k} = \left[ s_{\mathbf{q}_r^{p,k}(1)}, s_{\mathbf{q}_r^{p,k}(2)}, \dots, s_{\mathbf{q}_r^{p,k}(T)} \right]^T. \quad (15)$$

The decoded video bit stream  $\hat{\mathbf{x}}_r^{p,k}$  for the  $k$ th coded component in the  $r$ th row and the  $p$ th frame is recovered from  $\hat{\mathbf{s}}_r^{p,k}$  using the generator matrix  $\mathbf{G}$ , by repeating steps in (3) and (4) with  $\mathbf{x}_r^{p,k}$  replaced by  $\hat{\mathbf{x}}_r^{p,k}$ . This process is repeated for all the rows of the frame to obtain the frame pixel intensity estimate  $\hat{\mathbf{A}}^p$  in turn for each of the  $P$  frames of the video stream. A significant shortcoming of the conventional method described here is that it does not exploit the bounded variation property [29] of the image/ video stream towards data detection, which has the potential to substantially reduce the decoding errors at the receiver, thereby leading to a significant improvement in the quality of reconstruction. Motivated by this fact, a novel TV regularization based STTC decoder is proposed in the following section for robust image/ video reconstruction in MIMO wireless systems.

### III. PROPOSED STTC-TV DECODER FOR IMAGE/VIDEO RECOVERY

The proposed STTC-TV decoder computes the transmit symbol vector sequence estimate  $\hat{\Xi}_r^{p,k}$  and the information bit sequence estimate  $\hat{\mathbf{x}}_r^{p,k}$  from the received signal vector sequence  $\mathbf{Y}_r^{p,k}$  employing the maximum *a posteriori* probability (MAP) rule that is simplified below

$$\begin{aligned} \hat{\Xi}_r^{p,k} &= \underset{\Xi_r^{p,k}}{\text{argmax}} P(\Xi_r^{p,k} | \mathbf{Y}_r^{p,k}) \\ &= \underset{\Xi_r^{p,k}}{\text{argmax}} P(\mathbf{Y}_r^{p,k} | \Xi_r^{p,k}) P(\Xi_r^{p,k}) \\ &= \underset{\Xi_r^{p,k}}{\text{argmax}} \prod_{t=1}^T P(\mathbf{y}_{r,t}^{p,k} | \xi_{r,t}^{p,k}) P(\Xi_r^{p,k}). \end{aligned} \quad (16)$$

The likelihood function  $P(\mathbf{y}_{r,t}^{p,k} | \xi_{r,t}^{p,k})$  is given by the standard Gaussian distribution corresponding to the MIMO wireless channel model in (7). The *prior* probability density function (PDF)  $P(\Xi_r^{p,k})$  for an image/ video stream is modeled using the standard Gibbs distribution [52], which has been shown to be well suited for such purposes, that is expressed as

$$P(\Xi_r^{p,k}) = \frac{1}{\sigma_z} e^{-U(\Xi_r^{p,k})}, \quad (17)$$

where  $U(\Xi_r^{p,k})$  is the prior energy functional and  $\sigma_z$  is a suitably chosen normalization factor. On substituting the likelihood function and the *prior* probability density in the MAP rule determined in (16), the decoder optimization problem simplifies to

$$\hat{\Xi}_r^{p,k} = \underset{\Xi_r^{p,k}}{\text{argmax}} \sum_{t=1}^T \left( \|\mathbf{y}_{r,t}^{p,k} - \mathbf{H}_{r,t}^{p,k} \xi_{r,t}^{p,k}\|_2^2 \right) + \gamma U(\Xi_r^{p,k}) \quad (18)$$

where the regularization parameter  $\gamma$  controls the weight ratio of the regularization energy functional  $U(\Xi_r^{p,k})$  vis-a-vis the approximation error in the first term inside the sum. The regularization term in the cost function above exploits the prior information about the smoothness of the reconstructed image/ video stream obtained from the problem in (18). The use of an  $l_2$  norm based isotropic TV regularization functional  $U(\Xi_r^{p,k})$  was first introduced by Rudin *et al.* in [29] for noise removal in images. Subsequently, a scheme employing the  $l_1$  norm based anisotropic TV norm was developed in [53] for restoration of color/vector valued images. The TV regularization cost function therein was shown to be rotational invariant as well as better suited for preserving edges, thereby leading to superior image recovery. Motivated by these advantages of the  $l_1$  norm based regularization methods, the TV regularization functional is set as  $U(\Xi_r^{p,k}) \triangleq \|\mathbf{x}_r^{p,k}\|_{TV}$ , where the TV norm is defined as

$$\|\mathbf{x}_r^{p,k}\|_{TV} = \sum_{l=1}^{C_s} \|\mathbf{f}_D(\mathbf{x}_{r,l}^{p,k})\|_1. \quad (19)$$

The operator  $\mathbf{f}_D(\cdot)$  evaluates the difference vector of  $\mathbf{x}_{r,l}^{p,k}$  computed with respect to its causal connected causal neighbors [44] as shown below

$$\mathbf{f}_D(\mathbf{x}_{r,l}^{p,k}) = \begin{bmatrix} f_w(\mathbf{x}_{r,l}^{p,k}) - f_w(\mathbf{x}_{r,l-1}^{p,k}) \\ f_w(\mathbf{x}_{r,l}^{p,k}) - f_w(\hat{\mathbf{x}}_{r-1,l-1}^{p,k}) \\ f_w(\mathbf{x}_{r,l}^{p,k}) - f_w(\hat{\mathbf{x}}_{r-1,l}^{p,k}) \\ f_w(\mathbf{x}_{r,l}^{p,k}) - f_w(\hat{\mathbf{x}}_{r-1,l+1}^{p,k}) \\ \gamma_f (f_w(\mathbf{x}_{r,l}^{p,k}) - f_w(\hat{\mathbf{x}}_{r-1,l-1}^{p-1,k})) \\ \gamma_f (f_w(\mathbf{x}_{r,l}^{p,k}) - f_w(\hat{\mathbf{x}}_{r-1,l}^{p-1,k})) \\ \vdots \\ \gamma_f (f_w(\mathbf{x}_{r,l}^{p,k}) - f_w(\hat{\mathbf{x}}_{r+1,l+1}^{p-1,k})) \end{bmatrix}, \quad (20)$$

where the operator  $f_w(\mathbf{x}_{r,l}^{p,k})$  evaluates the numerical value corresponding to the coded component  $\mathbf{x}_{r,l}^{p,k}$ . The quantity  $\gamma_f$  controls the weight of the gradient of the coded components of the frame along the temporal direction with respect to the gradients in the spatial directions. On substituting (20) and (19) in (18), the proposed MAP rule based optimization problem in (16) simplifies to the one given in (21), as shown at the bottom of the next page, where  $\mathcal{U}(p, r, l)$  and  $\mathcal{U}_d(p, r, l)$  denote the frame, row and column indices of the causal neighbours and causal decoded neighbours as defined in (22) and (24) as shown at the bottom of the next page, respectively, and  $\delta_a(b) = 1$  if  $a = b$  and 0 otherwise is the standard Kronecker delta function.

The solution to the optimization problem in (21) cannot be determined using conventional techniques due to the high computational complexity associated with the large dimensionality of the coded component space that grows exponentially with the sequence length  $C_s$ . Therefore, in order to overcome this challenge of dimensionality and recover the video frame efficiently, a novel modified Viterbi decoder is proposed next, which employs a hierarchical trellis for decoding the sequence  $\mathbf{x}_r^{p,k}$ . This is well suited for signal recovery in practical scenarios since it has a linear complexity with respect to the sequence length  $C_s$ , similar to that of the conventional Viterbi decoder [51].

#### A. HIERARCHICAL TRELLIS BASED OPTIMAL VITERBI DECODER FOR STTC-TV DETECTION

In the proposed hierarchical trellis for TV decoding, the higher layer is termed as the CC-layer and corresponds to the sequence of the coded components, as shown in Fig. 2. The CC-layer also evaluates the TV cost between the coded components, that arises due to the bounded variation property of images/ video. The lower layer is termed the ST-layer, and the corresponding decoder is similar to a regular STTC decoder. The ST-layer evaluates the ML cost corresponding to the STTC symbols. The CC-layer comprises of  $C_s$  stages with  $N$  nodes in each stage. Each  $i$ th node in this trellis corresponds to a valid coded component bit sequence vector

denoted by  $\mathbf{v}_i \in \{0, 1\}^{B_{r,l}^{p,k} \times 1}$ . For instance,  $\mathbf{v}_i = [10010011]^T$  corresponds to the pixel intensity value  $\mathbf{A}^P(r, l) = 147$  for  $B_{r,l}^{p,k} = 8$ . Each node corresponds to transmission of the STTC coded symbols over the duration of  $T' = B_{r,l}^{p,k}/m$  time instants, which is further expanded in the lower layer, termed as the ST layer, as shown in Fig. 3. The ST-layer has  $V$  encoder states with the  $i$ th state denoted by  $s_i$  and  $T'$  ST-layer stages, similar to the conventional decoder in section-II. Each node in the CC-layer of the trellis has  $V$  initial and final encoder states referred to as initial and final subnodes denoted by  $s_i$  for  $1 \leq i \leq V$ , which denote the possible encoder states before and after STTC encoding of the corresponding coded component. The subnodes determine the valid branches between two nodes at successive stages in the trellis, since encoder state continuity is maintained for STTC encoding of the coded components from one CC-layer stage to the next. The final encoder state after encoding of a coded component is identical to the initial encoder state of the coded components for the next CC-layer stage. Therefore, a valid branch between two nodes in the CC-layer exists only between identical subnodes connecting the same encoder states in the ST-layer. The modified node and branch metrics derived from the objective function in the optimization problem in (21) corresponding to the hierarchical trellis are described next.

The node metric  $\Psi_{r,l}^{p,k}(i, q)$  corresponding to the  $q$ th initial subnode of the  $i$ th node at the  $l$ th CC-layer stage is defined in (23) as shown at the bottom of this page. This comprises of the decoding error for the symbol vector sequence

$$\tilde{\mathbf{e}}_{i,q} = \left[ \tilde{\xi}_{i,q}^1, \tilde{\xi}_{i,q}^2, \dots, \tilde{\xi}_{i,q}^{T'} \right], \quad (25)$$

corresponding to STTC encoding of  $\mathbf{v}_i$  with initial encoder state  $s_q$  together with the TV norm regularization term corresponding to the causal decoded neighbors denoted by the

---

**Algorithm 1** STTC-TV Decoder for Evaluation of  $\hat{\mathbf{x}}_{r,l}^{p,k}$ 


---

**Input Data:**  $r, p, \mathbf{Y}_r^{p,k}, \mathbf{H}_{r,1:T}^{p,k}$ .

**Input Parameter:**  $\mathbf{G}, \gamma, N, V, T', R_s, C_s, T, P$ .

1. Set  $l = 1$  and determine  $\mathcal{U}(p, r, 1)$  using (22).

**for**  $l = 1$  to  $C_s$  **do**

**for**  $i = 1$  to  $N$  **do**

**for**  $q = 1$  to  $V$  **do**

            a. Evaluate the transmit symbol vector sequence  $\left[ \tilde{\xi}_{i,q}^1, \tilde{\xi}_{i,q}^2, \dots, \tilde{\xi}_{i,q}^{T'} \right]$  and set  $\mathcal{V}_i^q$  corresponding to

STTC encoding of codeword  $\mathbf{v}_i$  in the CC-layer.

            b. Compute the node metric  $\Psi_{r,l}^{p,k}(i, q)$  as given in (26) for  $l = 1$  and in (23) for  $l \geq 2$  otherwise.

            c. Initialize the accumulated metric  $\Phi_{r,l}^{p,k}(i, q)$  using (31) for  $l = 1$  and in (28) for  $l \geq 2$  otherwise.

**if**  $l \geq 2$  **then**

            I. Evaluate  $\tilde{\xi}_{r,l}^{p,k}(i, q)$  and  $\xi_{r,l}^{p,k}(i, q)$  using  $\Phi_{r,l}^{p,k}(i, q)$  in (c) above as given in (29) and (30) respectively.

**end if**

**end for**

**end for**

**end for**

2. Determine node index  $\tau(C_s)$  and the subnode index  $\tilde{\tau}(C_s)$  corresponding to the least accumulated cost at CC-stage  $l = C_s$  as given in (32).

3. Compute the optimal transmit CC bit sequence estimate for remaining CC-stages recursively using  $\tilde{\xi}_{r,l}^{p,k}(i, q)$  and  $\xi_{r,l}^{p,k}(i, q)$  as given next

**for**  $l = C_s - 1$  to  $1$  **do**

    I. Find the CC transmit bit sequence estimate  $\hat{\mathbf{x}}_{r,l}^{p,k}$  using the steps described in (33).

**end for**

---

$$\hat{\mathbf{e}}_r^{p,k} = \underset{\mathbf{e}_r^{p,k}}{\arg \min} \sum_{t=1}^T \left( \left\| \mathbf{y}_{r,t}^{p,k} - \mathbf{H}_{r,t}^{p,k} \xi_{r,t}^{p,k} \right\|_2^2 \right) + \sum_{l=1}^{C_s} \left( \frac{\gamma}{|\mathcal{U}(p, r, l)|} \sum_{(p', r', l') \in \mathcal{U}(p, r, l)} (\delta_p(p') + \gamma_f \delta_{p-1}(p')) \right. \\ \left. \left| f_w(\mathbf{x}_{r,l}^{p,k}) - \delta_{r,p}(r', p') f_w(\mathbf{x}_{r',l'}^{p',k}) - (1 - \delta_{r,p}(r', p')) f_w(\hat{\mathbf{x}}_{r',l'}^{p',k}) \right| \right) \quad (21)$$

$$\mathcal{U}(p, r, l) = \left\{ (p', r', l') \mid 1 \leq l' \leq C_s; 1 \leq r' \leq R_s; 1 \leq p' \leq P; p' \in \{p-1, p\}; r' \in \{r-1, r, r+1\}; l' \in \{l-1, l, l+1\}; 1 \leq (R_s C_s (p' - 1) + C_s (r' - 1) + l') < (R_s C_s (p - 1) + C_s (r - 1) + l) \right\}. \quad (22)$$

$$\Psi_{r,l}^{p,k}(i, q) = \sum_{t=1}^{T'} \left\| \mathbf{y}_{r,(l-1)T'+t}^{p,k} - \mathbf{H}_{r,t}^{p,k} \tilde{\xi}_{i,q}^t \right\|_2^2 + \left( \frac{\gamma}{|\mathcal{U}(p, r, l)|} \sum_{(p', r', l') \in \mathcal{U}_d(p, r, l)} (\delta_p(p') + \gamma_f \delta_{p-1}(p')) \left| f_w(\mathbf{v}_i) - f_w(\hat{\mathbf{x}}_{r',l'}^{p',k}) \right| \right); \quad (23)$$

$$\mathcal{U}_d(p, r, l) = \mathcal{U}(p, r, l) \setminus \{(p, r, l-1)\}, \forall 1 \leq p \leq P, 1 \leq r \leq R_s, 2 \leq l \leq C_s \quad (24)$$



FIGURE 4. Test images for simulations.

set  $\mathcal{U}_d(p, r, l)$  as defined in (24). An illustration of the hierarchical structure of the trellis for each node  $\mathbf{v}_i$  in the CC-layer together with the corresponding trellis in the ST-layer for the transmit vector sequence  $\tilde{\mathbf{e}}_{i,q}$  is given in Fig. 3. Note that for a  $V$ -state STTC, the transmit vector sequence is found using the generator matrix  $\mathbf{G}$ , the initial state  $s_q$  and the information bit sequence  $\mathbf{v}_i$ . Assuming the initial encoder state to be  $s_1$ , for  $l = 1$ , the node metric  $\Psi_{r,l}^{p,k}(i, q)$  is initialized as

$$\Psi_{r,1}^{p,k}(i, q) = \begin{cases} \Psi_{r,1}^{p,k}(i, 1), & \text{if } q = 1 \\ \infty, & \text{otherwise.} \end{cases} \quad (26)$$

The branch metric  $\rho_{r,l}^{p,k}(i, j)$  corresponding to a valid branch from node  $i$  to node  $j$  for  $1 \leq l \leq C_s$ , is defined as

$$\rho_{r,l}^{p,k}(i, j) = \frac{\gamma}{|\mathcal{U}(p, r, l)|} \left( \left| f_w(\mathbf{v}_i) - f_w(\mathbf{v}_j) \right| \right). \quad (27)$$

Further, it can be seen that corresponding to the  $i$ th node, multiple paths emanating from different initial subnodes can terminate at a final subnode  $s_q$ . Let the set of indices of these initial subnodes be denoted by  $\mathcal{V}_i^q$  corresponding to each final subnode  $s_q$  at stage  $i$  in the ST-layer. The accumulated metric  $\Phi_{r,l}^{p,k}(i, q)$ , corresponding to the node  $\mathbf{v}_i$  and final encoder state  $s_q$ , is found as follows. For  $l = 1$ , it is initialized as

$$\Phi_{r,1}^{p,k}(i, q) = \begin{cases} \infty, & \text{if } |\mathcal{V}_i^q| = 0 \\ \min_{q' \in \mathcal{V}_i^q} \Psi_{r,1}^{p,k}(i, q'), & \text{otherwise.} \end{cases} \quad (31)$$

For stages  $l \geq 2$ , the accumulated metric  $\Phi_{r,l}^{p,k}(i, q)$ , the survivor subnode  $\tilde{\zeta}_{r,l}^{p,k}(i, q)$  and the survivor node  $\zeta_{r,l}^{p,k}(i, q)$  of the  $q$ th node in the final stage of the ST-layer corresponding to the  $i$ th node in the  $l$ th CC-layer stage are computed recursively using (28), (29) and (30), as shown at the bottom of the next page respectively. The above metrics and the survivor nodes are computed for all the  $V$  subnodes for each node in the CC-layer. Finally, in order to determine the bit sequence estimate  $\hat{\mathbf{x}}_{r,C_s}^{p,k}$  of the coded component corresponding to the final stage  $l = C_s$  of the CC-layer, the node index  $\tau(C_s)$  and the subnode index  $\tilde{\tau}(C_s)$  corresponding to the least accumulated cost is found as follows

$$\begin{aligned} [\hat{i}, \hat{q}] &= [i, q] \Phi_{r,C_s}^{p,k}(i, q), \\ \tau(C_s) &= \hat{i}; \quad \tilde{\tau}(C_s) = \hat{q}; \quad \hat{\mathbf{x}}_{r,C_s}^{p,k} = \mathbf{v}_{\tau(C_s)}. \end{aligned} \quad (32)$$

The codeword estimates of the coded components corresponding to the rest of the CC-layer stages  $1 \leq l \leq C_s - 1$  are determined from the survivor nodes  $\zeta_{r,l}^{p,k}$  and the survivor subnodes  $\tilde{\zeta}_{r,l}^{p,k}$  recursively as follows

$$\begin{aligned} \tilde{\tau}(l) &= \tilde{\zeta}_{r,l+1}^{p,k}(\tau(l+1), \tilde{\tau}(l+1)) \\ \tau(l) &= \zeta_{r,l+1}^{p,k}(\tau(l+1), \tilde{\tau}(l+1)) \\ \hat{\mathbf{x}}_{r,l}^{p,k} &= \mathbf{v}_{\tau(l)}, \quad l = (C_s - 1), \dots, 2, 1. \end{aligned} \quad (33)$$

A concise summary of the various steps of the procedure for decoding the codewords corresponding to the  $k$ th coded component of the blocks in the  $r$ th row of the  $p$ th frame is given in Algorithm 1. Furthermore, the estimate  $\hat{\mathbf{A}}_{r,l}^{d,p}$  of the  $(r, l)$ th coded block is obtained by decoding the estimate  $\hat{\mathbf{x}}_{r,l}^{p,k}; 1 \leq k \leq Q_{r,l}^p$  of the coded component vector sequence. The estimates of the pixel intensities  $\hat{\mathbf{A}}_{r,l}^p \in \mathbb{R} \in \mathbb{R}^{b \times b}$  for the  $(r, l)$ th block are subsequently obtained via inverse transform (IDCT for JPEG, MPEG) of the coded block estimate as shown in (34), as shown at the bottom of the next page. The  $R_s \times C_s$  size image/ video frame  $\hat{\mathbf{A}}^p$  is finally reconstructed as

$$\hat{\mathbf{A}}^p((r-1)b+i, (l-1)b+j) = \hat{\mathbf{A}}_{r,l}^p(i, j). \quad (35)$$

#### IV. SIMULATION RESULTS AND DISCUSSION

Visual and numerical results are presented in this section followed by discussion regarding the performance of the proposed scheme and its comparison with the conventional reconstruction schemes. The setup for image transmission and reconstruction is described next.

##### A. IMAGE TRANSMISSION AND WIRELESS SYSTEM MODEL

In order to validate the performance of the proposed scheme, simulations are performed on several standard test images such as *Lena*, *House*, *Jetplane*, *Walkbridge* and *Living room* of size  $256 \times 256$  pixels i.e.  $R = C = 256$  as shown in Fig. 4. The images and videos considered for simulations are standard test images/ videos that have been used in various other works in the existing literature such as [54]–[56]. A particularly appealing aspect of these frames that makes them ideally suited for testing the performance of different algorithms is that they represent a wide variety of physical conditions, covering indoor as well as outdoor scenarios, with the objects of interest placed at different locations. Moreover,



a variety of backgrounds and textures are present in each frame. Thus, these images/ videos are considered for comprehensively testing and demonstrating the performance of the proposed scheme for diverse conditions to ensure reliable performance in real-time implementations. For images, the number of frames is set as  $P = 1$  and both uncoded ( $b = 1, N = 256$ ) and JPEG coded ( $b = 8, N = 256$ ) formats are considered. The bits are subsequently STTC coded with different generator matrices [48] that are listed in Table 2. The coded symbols are modulated using QPSK symbols of power  $P_s$  given as

$$\xi_{r,t}^{p,k}(i) \in \left\{ \sqrt{\frac{P_s}{2}} + j\sqrt{\frac{P_s}{2}}, \sqrt{\frac{P_s}{2}} - j\sqrt{\frac{P_s}{2}}, -\sqrt{\frac{P_s}{2}} + j\sqrt{\frac{P_s}{2}}, -\sqrt{\frac{P_s}{2}} - j\sqrt{\frac{P_s}{2}} \right\}, \quad (36)$$

Next, these symbols are transmitted over a  $2 \times 2$  MIMO wireless channel i.e.  $N_t = N_r = 2$  characterized by the channel matrix  $\mathbf{H}_{r,t}^{p,k}$ . Each element  $\mathbf{H}_{r,t}^{p,k}(n_r, n_t)$  of the channel matrix is generated as a zero mean circularly symmetric complex Gaussian random variable  $h_R + jh_I$  where  $j = \sqrt{-1}$  and  $h_R, h_I$  are i.i.d. Gaussian random variables each distributed as  $\mathcal{N}\left[0, \frac{1}{2}\right]$ . This represents a Rayleigh fading channel since the envelope  $\sqrt{h_R^2 + h_I^2}$  is a Rayleigh distributed random variable.

Thus the average gain of each fading channel coefficient is set as unity i.e.  $\mathbb{E}\{|\mathbf{H}_{r,t}^{p,k}(n_r, n_t)|^2\} = 1$ . The noise vector  $\eta_{r,t}^{p,k}$  is considered to be complex Gaussian distributed with i.i.d. noise samples of power  $\sigma_\eta^2 = \frac{P_s}{\text{SNR}}$  where SNR is the signal to noise ratio of the wireless channel that is defined as  $\text{SNR} = \frac{P_s}{\sigma_\eta^2}$ . The parameter SNR refers to the average value of the signal to noise power ratio at the receiver taking the fading into account.

The performance of the proposed scheme has been illustrated using three different metrics: PSNR (Peak signal to noise ratio), bit error rate (BER) and visual quality. The PSNR is a standard quality assessment metric that has been used in many works on image/ video processing such

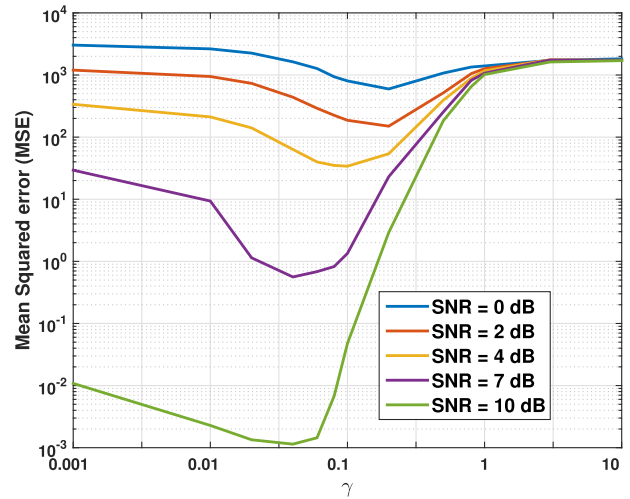


FIGURE 5. Optimal  $\gamma$  determination using lowest mean squared error criterion.

as [32], [57], [58]. This has been shown to be best suited to study the performance of image/ video reconstruction. Furthermore, the performance of the proposed scheme is also characterized in terms of the BER in Fig. 7, which is once again a well-established measure to demonstrate the improvement in the accuracy of symbol detection for wireless communication. Finally, individual reconstructed images and video frames are also shown in Fig. 8, Fig. 9, Fig. 10 and Fig. 11 to demonstrate the improvement in terms of visual quality for various values of SNR. The average PSNR metric is computed for the reconstructed image  $\hat{\mathbf{A}}$  with respect to the original image  $\mathbf{A}$  as

$$\text{PSNR}(\mathbf{A}, \hat{\mathbf{A}}) = 10 \log_{10} \left( \frac{255^2}{\text{MSE}(\mathbf{A}, \hat{\mathbf{A}})} \right), \quad (37)$$

where  $\text{MSE}(\mathbf{A}, \hat{\mathbf{A}})$  is computed as

$$\text{MSE}(\mathbf{A}, \hat{\mathbf{A}}) = \frac{1}{RC} \sum_{r=1}^R \sum_{c=1}^C (\mathbf{A}(r, c) - \hat{\mathbf{A}}(r, c))^2. \quad (38)$$

$$\Phi_{r,l}^{p,k}(i, q) = \begin{cases} \infty, & \text{if } |\mathcal{V}_i^q| = 0 \\ \min_{q' \in \mathcal{V}_i^q} \left( \Psi_{r,l}^{p,k}(i, q') + \min_{1 \leq j \leq N} \left( \Phi_{r,l-1}^{p,k}(j, q') + \rho_{r,l}^{p,k}(i, j) \right) \right), & \text{otherwise.} \end{cases} \quad (28)$$

$$\tilde{\zeta}_{r,l}^{p,k}(i, q) = \begin{cases} 0, & \text{if } |\mathcal{V}_i^q| = 0 \\ \min_{q' \in \mathcal{V}_i^q} \left( \Psi_{r,l}^{p,k}(i, q') + \min_{1 \leq j \leq N} \left( \Phi_{r,l-1}^{p,k}(j, q') + \rho_{r,l}^{p,k}(i, j) \right) \right), & \text{otherwise.} \end{cases} \quad (29)$$

$$\zeta_{r,l}^{p,k}(i, q) = \min_{1 \leq j \leq N} \left( \Phi_{r,l-1}^{p,k}(j, \tilde{\zeta}_{r,l}^{p,k}(i, q)) + \rho_{r,l}^{p,k}(i, j) \right) \quad (30)$$

$$\hat{\mathbf{A}}_{r,l}^p(i, j) = \sum_{u=0}^{b-1} \sum_{v=0}^{b-1} \alpha(u) \alpha(v) \hat{\mathbf{A}}_{r,l}^{d,p}(u, v) \cos\left(\frac{\pi(i + \frac{1}{2})u}{b}\right) \cos\left(\frac{\pi(j + \frac{1}{2})v}{b}\right); \quad \alpha(x) = \begin{cases} \frac{1}{8}, & \text{if } x = 0 \\ \frac{2}{8}, & \text{otherwise} \end{cases} \quad (34)$$

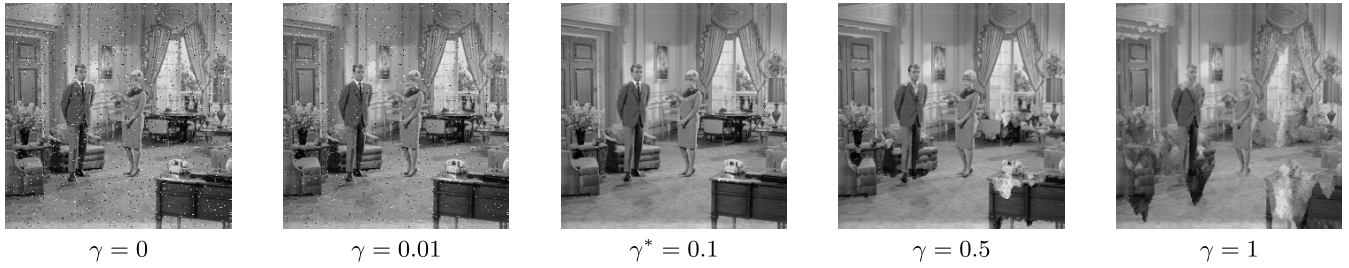


FIGURE 6. Effect of change in  $\gamma$  on test image *Living room* reconstruction at SNR = 4 dB.

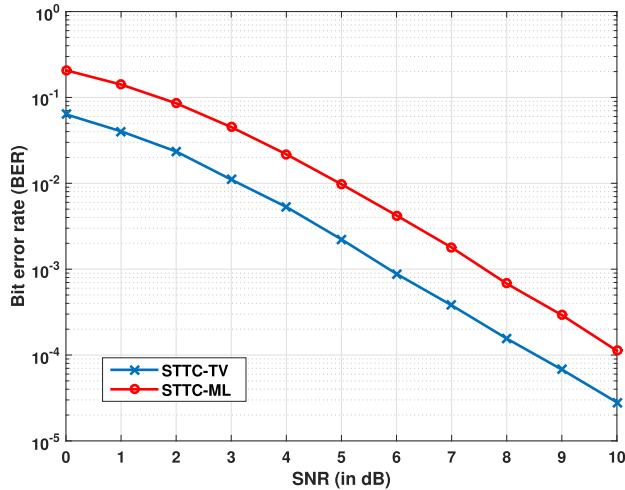


FIGURE 7. Bit error rate performance of the proposed scheme.

**B. OPTIMAL REGULARIZATION PARAMETER  $\gamma$**

The optimal regularization parameter  $\gamma$  is determined using the minimum mean squared error (MSE) criterion for different channel SNR conditions. The average MSE of image reconstruction for different SNRs in the range [0, 10] dB is plotted with respect to  $\gamma$  values in the range  $10^{-3}$ , 10 as shown in Fig. 5 averaged over 100 Monte-Carlo simulations. It is seen that the lowest MSE is achieved for regularization parameter values belonging to the range [0.05, 0.5] for the SNR range considered. The effect of changing the value of  $\gamma$  in the range [0, 1] on the visual quality of image reconstruction at SNR= 4 dB is demonstrated in Fig. 6. It can be observed therein that the lowest reconstruction error and the best visual quality of the reconstructed image is indeed achieved at  $\gamma = \gamma^*$  where  $\gamma^* = 0.1$  is the optimal value of the regularization parameter determined as above from Fig. 5. For  $\gamma < \gamma^*$ , the reconstructed image obtained is noisy whereas further increase in the value of  $\gamma$  value beyond the optimal value degrades the reconstruction performance due to over smoothing.

**C. IMAGE RECONSTRUCTION PERFORMANCE**

For the optimal values of  $\gamma$ , Monte-carlo simulations are carried out to demonstrate the average PSNR performance of the proposed STTC-TV decoder for different channel SNR



FIGURE 8. Test image *Lena*, *Jetplane* and *Living room* recovered at SNR = 0 dB (STTC-ML), SNR = 0 dB (STTC-TV), SNR = 4 dB (STTC-ML), SNR = 4 dB (STTC-TV) (left to right).



FIGURE 9. Test images *Lena*, *House* and *Walkbridge* recovered at SNR = 4 dB (STTC-ML), SNR = 4 dB (STTC-TV), SNR = 7 dB (STTC-ML), SNR = 7 dB (STTC-TV) (left to right) for JPEG coded image transmission.

conditions and also to compare the performance with that of the conventional STTC-ML decoder. The DC components are recovered using the STTC-TV procedure whereas the AC coded-components are estimated using the conventional STTC-ML technique. The corresponding results for average

**TABLE 1.** Average PSNR (in dB) performance of the proposed scheme for image reconstruction.

SNR		0 dB		2 dB		4 dB		7 dB		10 dB	
Test image	Coding	STTC-ML	STTC-TV	STTC-ML	STTC-TV	STTC-ML	STTC-TV	STTC-ML	STTC-TV	STTC-ML	STTC-TV
Lena	Raw	12.95	25.05	16.63	31.60	22.45	37.11	33.11	46.03	42.65	59.59
	JPEG	10.43	13.65	11.74	14.57	13.61	16.51	22.89	27.34	32.19	34.01
House	Raw	13.70	24.90	16.88	30.86	22.44	38.09	33.26	50.31	44.98	59.30
	JPEG	11.23	13.23	12.23	15.24	13.21	17.17	22.62	26.87	31.65	34.12
Jetplane	Raw	11.99	22.01	15.80	27.41	21.39	34.34	32.51	45.60	43.68	58.05
	JPEG	11.15	13.91	11.87	14.89	13.98	17.78	22.65	27.38	32.62	35.29
Walkbridge	Raw	12.76	21.41	16.49	26.48	22.19	33.28	32.91	43.21	44.55	57.42
	JPEG	10.24	12.93	11.13	14.29	13.40	16.29	21.95	26.63	31.19	34.38
Living room	Raw	13.12	23.41	16.79	28.54	22.37	35.16	33.50	47.43	44.59	57.05
	JPEG	10.49	13.92	11.22	14.72	13.91	17.33	22.48	28.16	32.03	34.49

**TABLE 2.** Average PSNR performance for different generator matrices for test image: Lena at SNR = 4dB.

Generator matrix	Coding	STTC-ML	STTC-TV
$G_1 = \begin{bmatrix} 0 & 2 \\ 0 & 1 \\ 2 & 0 \\ 1 & 0 \end{bmatrix}$	Raw	20.50 dB	33.14 dB
	JPEG	10.84 dB	13.44 dB
$G_2 = \begin{bmatrix} 2 & 2 \\ 0 & 1 \\ 2 & 0 \\ 1 & 0 \\ 2 & 2 \end{bmatrix}$	Raw	22.15 dB	35.64 dB
	JPEG	12.48 dB	15.26 dB
$G_3 = \begin{bmatrix} 1 & 2 \\ 2 & 0 \\ 2 & 1 \\ 0 & 2 \\ 0 & 2 \\ 2 & 0 \end{bmatrix}$	Raw	22.45 dB	37.11 dB
	JPEG	13.61 dB	16.51 dB
$G_4 = \begin{bmatrix} 2 & 2 \\ 0 & 2 \\ 1 & 0 \\ 3 & 1 \end{bmatrix}$	Raw	20.62 dB	34.60 dB
	JPEG	12.77 dB	15.86 dB

PSNR are shown in Table 1. The average PSNR values for the proposed scheme using different standard STTC generator matrices reported in literature [48], obtained using the optimal code design criterion, are shown in Table 2. It can be observed from the simulation results shown in Table 2 that the image/ video reconstruction performance of the proposed STTC-TV scheme improves with the number of states in the STTC codes. Thus, the best performance among the four STTC generator matrices in Table 2 is obtained for  $G_3$  that corresponds to a 16-state trellis. The simulations are performed for both uncoded and JPEG coded test image *Lena* at SNR= 4dB. The numerical PSNR results in the table as well as visual comparison of raw and JPEG compressed images



**FIGURE 10.** Reconstructed frame–53 of the test videos *Foreman*, *Coastguard* and *Akiyo* recovered at SNR = 0 dB (STTC-ML), SNR = 0 dB (STTC-TV), SNR = 4 dB (STTC-ML), SNR = 4 dB (STTC-TV) (left to right) for raw video transmission.



**FIGURE 11.** Reconstructed frame–53 of the test videos *Foreman*, *Coastguard* and *Akiyo* recovered at SNR = 4 dB (STTC-ML), SNR = 4 dB (STTC-TV), SNR = 7 dB (STTC-ML), SNR = 7 dB (STTC-TV) (left to right) for *Ouni* compressed video transmission.

in Fig. 8 and in Fig. 9 respectively, demonstrate the improved image recovery in comparison to that of the conventional ML based STTC decoder. For instance, the proposed scheme

**TABLE 3.** Average PSNR (in dB) performance of the proposed scheme for video reconstruction.

SNR		0 dB		2 dB		4 dB		7 dB		10 dB	
Test video	Coding	STTC-ML	STTC-TV	STTC-ML	STTC-TV	STTC-ML	STTC-TV	STTC-ML	STTC-TV	STTC-ML	STTC-TV
Foreman	Raw	12.51	26.70	16.20	31.48	21.92	36.81	32.65	46.65	45.91	54.40
	Ouni	9.41	12.75	12.78	15.94	19.56	24.24	30.34	34.57	35.52	38.77
Coastguard	Raw	12.28	24.98	16.00	30.79	21.98	35.34	32.69	45.76	45.82	54.92
	Ouni	8.96	11.06	11.13	13.24	18.66	22.81	29.16	32.34	34.61	39.22
Akiyo	Raw	12.96	24.09	15.80	29.41	22.30	34.13	32.51	44.60	43.68	51.05
	Ouni	9.33	13.13	12.59	15.89	20.10	25.85	31.29	36.49	41.02	42.74

can be seen to yield a PSNR improvement that is close to 14dB and 3dB corresponding to Raw and JPEG compression respectively, for SNR = 4dB. The effect of change of average SNR on the overall bit error rate (BER) performance of the proposed scheme has been shown for image recovery for the SNR range [0, 10] dB in Fig.7. The performance for each SNR value is obtained by averaging over several instantiations of the fading wireless channel. The consistent PSNR improvement to the tune of 2dB demonstrates the improved ability of STTC-TV for various values of average SNR considering also the effects of fading in the WMSN. The numerical and visual results above demonstrate the improved image/ video reconstruction of the proposed scheme over the conventional STTC-ML scheme arising due to the exploitation of the bounded variation property of the transmitted stream. The BER performance of wireless communication also shows a similar improvement in decoding the transmitted symbols that get severely distorted due to the fading nature of the wireless channel. This boost in performance demonstrates the ability of STTC-TV to combat the adverse effects of fading for various SNR conditions in the WMSN.

#### D. VIDEO RECONSTRUCTION

Simulations are also performed for video reconstruction using the standard test video sequences: *Foreman*, *Coastguard* and *Akiyo*. The different parameters for the videos under consideration are set as  $R = 176$ ,  $C = 144$  and  $P = 300$ . The scheme is applied for raw as well as on the coded video streams using the low complexity DCT-based video compression technique described by Ouni et.al. in [59] referred herein as *Ouni* for brevity. The procedure for the same is briefly described as follows. A temporal decomposition is performed on the 3D video frames transforming them into 2D image components using the popular Accordion representation. Specifically, each  $8 \times 8 \times 8$  cube of video pixels is transformed into an image frame of size  $8 \times 64$ . Thus, the number of such frames generated for the test video sequences above is  $\tilde{P} = 38$ . These frames are then compressed using JPEG coding. The proposed scheme is employed at the receiver for reconstruction of the video frames with its performance once again assessed using the

average PSNR metric that is defined as

$$\text{PSNR} = \frac{1}{P} \sum_{p=1}^P 10 \log_{10} \left( \frac{255^2}{\text{MSE}(\mathbf{A}^p, \hat{\mathbf{A}}^p)} \right). \quad (39)$$

Using a procedure similar to the one described for image reconstruction above, the optimal regularization parameter  $\gamma$  is observed to lie in the range [0.1, 0.5]. The parameter  $\gamma_f$  corresponding to (20) is set as  $\gamma_f = 1.25$ , similar to [32]. The average PSNR values of the proposed scheme and that for the conventional reconstruction are shown in Table-3 which clearly demonstrates the enhanced quality of video recovery obtained using the proposed approach. For instance, the PSNR improvement is approximately 15dB and 4dB for raw and *Ouni* compressed video reconstruction respectively, for SNR = 4dB. In addition to the numerical results, the representative frames of all the reconstructed video sequences are shown in Fig. 10 and Fig. 11 respectively for raw and *Ouni* compression schemes for the proposed and the conventional recovery techniques. Numerical values and visual examination of the frames clearly show that the proposed decoder outperforms the conventional STTC-ML decoder. The performance margin is especially pronounced in the low SNR regime thus making the approach well suited for application in WMSN with power constrained sensor nodes.

#### V. CONCLUSION

An STTC coding based image/ video communication framework has been developed for WMSN. A TV regularization based image/ video decoder is developed for robust recovery, which employs an  $l_1$  norm regularization term to exploit the bounded variation property of the image/ video stream in conjunction with the  $l_2$  norm based ML decoder for the STTC-MIMO wireless system. A novel STTC-TV decoder is developed based on the Viterbi algorithm to minimize the TV regularization cost function using a bi-layered hierarchical trellis structure for image/ video reconstruction. The modified state, branch and accumulated metrics are defined for the trellis nodes and branches derived from the proposed TV optimization problem. Simulation results, both visually and numerically, demonstrate the improved reconstruction

performance of the proposed framework in comparison to the conventional STTC-ML decoder.

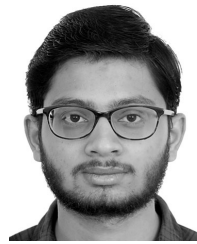
## ACKNOWLEDGMENT

The authors would like to thank Dr. Amrita Mishra, Assistant Professor at the Department of Electronics & Communication Engineering, Dr. SPM International Institute of Information Technology Naya Raipur, India for providing insight and expertise that greatly assisted the research. The authors are also grateful for support from the TCS research scholarship program at IIT Kanpur.

## REFERENCES

- [1] I. F. Akyildiz, T. Melodia, and K. R. Chowdhury, "A survey on wireless multimedia sensor networks," *Comput. Netw.*, vol. 51, no. 4, pp. 921–960, Mar. 2007.
- [2] T. Tomic, K. Schmid, P. Lutz, A. Domel, M. Kassecker, E. Mair, I. Grixia, F. Ruess, M. Suppa, and D. Burschka, "Toward a fully autonomous UAV: Research platform for indoor and outdoor urban search and rescue," *IEEE Robot. Autom. Mag.*, vol. 19, no. 3, pp. 46–56, Sep. 2012.
- [3] I. Akyildiz, T. Melodia, and K. Chowdhury, "Wireless multimedia sensor networks: A survey," *IEEE Wireless Commun.*, vol. 14, no. 6, pp. 32–39, Dec. 2007.
- [4] K. Yang, *Wireless Sensor Networks: Principles, Design and Applications*. London, U.K.: Springer, 2014.
- [5] F. L. Lewis, "Wireless sensor networks," *Smart Environ., Technol., Protocols Appl.*, vol. 11, p. 46, Aug. 2004.
- [6] I. F. Akyildiz, W. Su, Y. Sankarasubramanian, and E. Cayirci, "Wireless sensor networks: A survey," *Comput. Netw.*, vol. 38, no. 4, pp. 393–422, 2002.
- [7] I. F. Akyildiz and I. H. Kasimoglu, "Wireless sensor and actor networks: Research challenges," *Ad Hoc Netw.*, vol. 2, no. 4, pp. 351–367, Oct. 2004.
- [8] D. Ganesan, R. Govindan, S. Shenker, and D. Estrin, "Highly-resilient, energy-efficient multipath routing in wireless sensor networks," *ACM SIGMOBILE Mobile Comput. Commun. Rev.*, vol. 5, no. 4, pp. 11–25, Oct. 2001.
- [9] D. Tse and P. Viswanath, *Fundamentals Wireless Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2005.
- [10] Y. Charfi, N. Wakamiya, and M. Murata, "Network-adaptive image and video transmission in camera-based wireless sensor networks," in *Proc. 1st ACM/IEEE Int. Conf. Distrib. Smart Cameras*, Sep. 2007, pp. 336–343.
- [11] S. Mukhopadhyay, C. Schurgers, D. Panigrahi, and S. Dey, "Model-based techniques for data reliability in wireless sensor networks," *IEEE Trans. Mobile Comput.*, vol. 8, no. 4, pp. 528–543, Apr. 2009.
- [12] V. K. Goyal, "Multiple description coding: Compression meets the network," *IEEE Signal Process. Mag.*, vol. 18, no. 5, pp. 74–93, 2001.
- [13] M. Kazemi, K. H. Sadeghi, and S. Shirmohammadi, "A mixed layer multiple description video coding scheme," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 2, pp. 202–215, Feb. 2012.
- [14] M. Biswas, M. R. Frater, and J. F. Arnold, "Multiple description wavelet video coding employing a new tree structure," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 18, no. 10, pp. 1361–1368, Oct. 2008.
- [15] M. Darnell, "Error control coding: Fundamentals and applications," *IEE Proc. F Commun., Radar Signal Process.*, vol. 132, no. 1, p. 68, 1985.
- [16] J. N. Laneman, E. Martinian, G. W. Wornell, and J. G. Apostolopoulos, "Source-channel diversity for parallel channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 10, pp. 3518–3539, Oct. 2005.
- [17] A. J. Paulraj, D. A. Gore, R. U. Nabar, and H. Bolcskei, "An overview of MIMO Communications—A key to gigabit wireless," *Proc. IEEE*, vol. 92, no. 2, pp. 198–218, Feb. 2004.
- [18] D. Gesbert, M. Shafi, D.-s. Shiu, P. J. Smith, and A. Naguib, "From theory to practice: An overview of MIMO space-time coded wireless systems," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 3, pp. 281–302, Apr. 2003.
- [19] D. Gesbert, H. Bolcskei, D. Gore, and A. Paulraj, "MIMO wireless channels: Capacity and performance prediction," in *Proc. IEEE. Global Telecommun. Conference. Conf. Rec.*, 2000, pp. 1083–1088.
- [20] W. Liu, X. Li, and M. Chen, "Energy efficiency of MIMO transmissions in wireless sensor networks with diversity and multiplexing gains," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2005, p. 897.
- [21] C. Zhou, C.-W. Lin, X. Zhang, and Z. Guo, "A novel JSCC scheme for UEP-based scalable video transmission over MIMO systems," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 25, no. 6, pp. 1002–1015, Jun. 2015.
- [22] D. Song and C. Wen Chen, "Scalable H.264/AVC video transmission over MIMO wireless systems with adaptive channel selection based on partial channel information," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 9, pp. 1218–1226, Sep. 2007.
- [23] H. Kim, P. C. Cosman, and L. B. Milstein, "Motion-compensated scalable video transmission over MIMO wireless channels," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 23, no. 1, pp. 116–127, Jan. 2013.
- [24] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Trans. Inf. Theory*, vol. 44, no. 2, pp. 744–765, Mar. 1998.
- [25] T. M. Duman and A. Ghrayeb, *Coding for MIMO Communication Systems*. Hoboken, NJ, USA: Wiley, 2008.
- [26] K.-T. Shr, H.-D. Chen, and Y.-H. Huang, "A low-complexity viterbi decoder for space-time trellis codes," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 57, no. 4, pp. 873–885, Apr. 2010.
- [27] J. G. Proakis, *Compenders*. Hoboken, NJ, USA: Wiley, 2001.
- [28] J.-F. Aujol, G. Aubert, L. Blanc-Feraud, and A. Chambolle, "Image decomposition into a bounded variation component and an oscillating component," *J. Math. Imag. Vis.*, vol. 22, no. 1, pp. 71–88, Jan. 2005.
- [29] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D, Nonlinear Phenomena*, vol. 60, nos. 1–4, pp. 259–268, Nov. 1992.
- [30] C. R. Vogel and M. E. Oman, "Fast, robust total variation-based reconstruction of noisy, blurred images," *IEEE Trans. Image Process.*, vol. 7, no. 6, pp. 813–824, Jun. 1998.
- [31] T. F. Chan and C.-K. Wong, "Total variation blind deconvolution," *IEEE Trans. Image Process.*, vol. 7, no. 3, pp. 370–375, Mar. 1998.
- [32] S. H. Chan, R. Khoshabeh, K. B. Gibson, P. E. Gill, and T. Q. Nguyen, "An augmented lagrangian method for total variation video restoration," *IEEE Trans. Image Process.*, vol. 20, no. 11, pp. 3097–3111, Nov. 2011.
- [33] M. Nikolova, "An algorithm for total variation minimization and applications," *J. Math. Imag. Vis.*, vol. 20, nos. 1–2, pp. 89–97, Jan. 2004.
- [34] J. Shen and T. F. Chan, "Mathematical models for local nontexture inpaintings," *SIAM J. Appl. Math.*, vol. 62, no. 3, pp. 1019–1043, Jan. 2002.
- [35] T. F. Chan, A. M. Yip, and F. E. Park, "Simultaneous total variation image inpainting and blind deconvolution," *Int. J. Imag. Syst. Technol.*, vol. 15, no. 1, pp. 92–102, 2005.
- [36] S. Farsiu, M. Elad, and P. Milanfar, "Multiframe demosaicing and super-resolution of color images," *IEEE Trans. Image Process.*, vol. 15, no. 1, pp. 141–159, Jan. 2006.
- [37] Y. Bing Han and L. Nan Wu, "Super resolution reconstruction of video sequence based on total variation," in *Proc. Int. Symp. Intell. Multimedia, Video Speech Process.*, 2004, pp. 575–578.
- [38] S. Farsiu, M. D. Robinson, M. Elad, and P. Milanfar, "Fast and robust multiframe super resolution," *IEEE Trans. Image Process.*, vol. 13, no. 10, pp. 1327–1344, Oct. 2004.
- [39] C. Vazquez, H. Aly, E. Dubois, and A. Mitiche, "Motion compensated super-resolution of video by level sets evolution," in *Proc. Int. Conf. Image Process. (ICIP)*, 2004, pp. 1767–1770.
- [40] J. Song, H. Cho, J. Yoon, and S. M. Yoon, "Structure adaptive total variation minimization-based image decomposition," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 28, no. 9, pp. 2164–2176, Sep. 2018.
- [41] D. Strong and T. Chan, "Edge-preserving and scale-dependent properties of total variation regularization," *Inverse Problems*, vol. 19, no. 6, p. S165, 2003.
- [42] J. M. Bioucas-Dias, M. A. T. Figueiredo, and J. P. Oliveira, "Total variation-based image deconvolution: A majorization-minimization approach," in *Proc. IEEE Int. Conf. Acoust. Speed Signal Process. Proc.*, May 2006, p. 2.
- [43] A. Kudeshia and A. K. Jagannatham, "Optimal Viterbi based total variation sequence detection (TVSD) for robust image/video decoding in wireless sensor networks," *IEEE Signal Process. Lett.*, vol. 21, no. 6, pp. 722–726, Dec. 2014.
- [44] A. Kudeshia and A. K. Jagannatham, "Optimal viterbi algorithm based 3-D total variation sequence decoding (3-D TVSD) for robust video reconstruction in multimedia wireless sensor networks," in *Proc. IEEE 81st Veh. Technol. Conf. (VTC Spring)*, May 2015, pp. 1–5.
- [45] A. K. Jain, *Fundamentals of Digital Image Processing*. Upper Saddle River, NJ, USA: Prentice-Hall, 1989.

- [46] A. C. Bovik, *Handbook of Image and Video Processing*. New York, NY, USA: Academic, 2010.
- [47] E. Biglieri, R. Calderbank, A. Constantinides, A. Goldsmith, A. Paulraj, and H. V. Poor, *MIMO Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2007.
- [48] S. Baro, G. Bauch, and A. Hansmann, "Improved codes for space-time trellis-coded modulation," *IEEE Commun. Lett.*, vol. 4, no. 1, pp. 20–22, Jan. 2000.
- [49] T. S. Rappaport, *Wireless Communications: Principles and Practice*, vol. 2. Upper Saddle River, NJ, USA: Prentice-Hall, 1996.
- [50] B. Vucetic and J. Yuan, *Space-Time Coding*. Hoboken, NJ, USA: Wiley, 2003.
- [51] G. D. Forney, Jr., "The Viterbi algorithm," *Proc. IEEE*, vol. 61, no. 3, pp. 268–278, Mar. 1973.
- [52] V. Y. Panin, G. L. Zeng, and G. T. Gullberg, "Total variation regulated EM algorithm," *IEEE Trans. Nucl. Sci.*, vol. 46, no. 6, pp. 2202–2210, 1999.
- [53] P. Blomgren and T. F. Chan, "Color TV: Total variation methods for restoration of vector-valued images," *IEEE Trans. Image Process.*, vol. 7, no. 3, pp. 304–309, Mar. 1998.
- [54] H. Liu, J. Gu, M. Q.-H. Meng, and W.-S. Lu, "Fast weighted total variation regularization algorithm for blur identification and image restoration," *IEEE Access*, vol. 4, pp. 6792–6801, 2016.
- [55] Z. He, J. Cai, and C. W. Chen, "Joint source channel rate-distortion analysis for adaptive mode selection and rate control in wireless video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 12, no. 6, pp. 511–523, 2002.
- [56] H.-C. Shih, C.-T. Wang, and C.-L. Huang, "Spiral-like pixel reconstruction algorithm for spatiotemporal video error concealment," *IEEE Access*, vol. 6, pp. 6370–6381, 2018.
- [57] S.-M. Tseng and Y.-F. Chen, "Average PSNR optimized cross layer user grouping and resource allocation for uplink MU-MIMO OFDMA video communications," *IEEE Access*, vol. 6, pp. 50559–50571, 2018.
- [58] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Trans. Image Process.*, vol. 16, no. 8, pp. 2080–2095, Aug. 2007.
- [59] T. Ouni, W. Ayedi, and M. Abid, "New low complexity DCT based video compression method," in *Proc. Int. Conf. Telecommun.*, 2009, pp. 202–207.



**MOHD ABBAS ZAIDI** received the B.Tech. degree from IIT Kanpur, in 2019, with a major in electrical engineering and minors in machine learning and cognitive science. He is currently working at the Next AI Solutions Lab, Samsung Research, Seoul, as an Engineer, where he interned with the QA Team, in Summer 2018. His research interests broadly include wireless communication, natural language processing, and cognitive science. Upon his graduation, he was awarded the Prof. PVM Gold Medal and the Proficiency Medal for the best academic performance among all the 4-year undergraduate students. During his stay at IIT Kanpur, he also received the Lalit Narain Das Memorial Scholarship, the Smt Dharm Vati Garg Scholarship, the Anita and Santoshi Mehra Foundation Scholarship, and the consecutive Academic Excellence awards based on his academic performance.



**ADITYA K. JAGANNATHAM** (Member, IEEE) received the bachelor's degree from the Indian Institute of Technology Bombay, and the M.S. and Ph.D. degrees from the University of California at San Diego, USA. From April 2007 to May 2009, he was employed as a Senior Wireless Systems Engineer at Qualcomm Inc., San Diego, CA, USA, where he worked on developing 3G UMTS/WCDMA/HSDPA mobile chipsets as part of the Qualcomm CDMA Technologies Division. Since 2009, he has been a Faculty Member of the Electrical Engineering Department, IIT Kanpur, where he is currently an Associate Professor, and is also associated with the BSNL-IITK Telecom Center of Excellence (BITCOE). His research interests are in the area of next-generation wireless communications and networking, sensor and ad-hoc networks, digital video processing for wireless systems, wireless 3G/4G cellular standards, and CDMA/OFDM/MIMO wireless technologies. He has contributed to the 802.11n high throughput wireless LAN standard. He received the CAL(IT)2 Fellowship for pursuing his graduate studies at the University of California at San Diego, and the Upendra Patel Achievement Award for his efforts towards developing HSDPA/HSUPA/HSPA+ WCDMA technologies at Qualcomm, in 2009. At IIT Kanpur, he received the P.K. Kelkar Young Faculty Research Fellowship, from June 2012 to May 2015, for excellence in research, and the Gopal Das Bhandari Memorial Distinguished Teacher Award for the year 2012 to 2013, for excellence in teaching. He has also delivered a set of video lectures on Advanced 3G and 4G Wireless Mobile Communications for the Ministry of Human Resource Development funded by National Programme on Technology Enhanced Learning (NPTEL).



**ANKIT KUDESHIA** (Student Member, IEEE) received the B.Tech. degree in electronics and communication engineering from the Indian Institute of Information Technology, Design and Manufacturing, Jabalpur, India, in 2010. He is currently pursuing the M.Tech.-Ph.D. dual degree with the Department of Electrical Engineering, Indian Institute of Technology Kanpur, India. He has also worked as an Assistant Professor with the School of Electronics Engineering, KIIT University, Bhubaneswar, India. His research interests include signal processing techniques for wireless communications, smart grids, and wireless multimedia sensor networks. He was awarded the prestigious TCS-RSP Fellowship, from January 2014 to December 2018, for pursuing his Ph.D. degree in computing sciences at IIT Kanpur, and is the Winner of the Idea-Quest-2019 during Intern at Qualcomm India Pvt., Ltd.



**CHANDRA PRAKASH** received the M.Tech. degree in satellite technology and applications from the Aerospace Engineering Department, Indian Institute of Science, Bengaluru, and the Ph.D. degree in wireless communication from DA-IICT, Gandhinagar. He is currently a Scientist/Engineer-SG at the Space Applications Centre, Indian Space Research Organization (ISRO), Ahmedabad, where he heads the Satellite Communication Technology Division. He is also the Associate Project Director for the GSAT-7R Ground Segment and the Deputy Project Director for the MSS Terminal Development and Ground Segment development for GSAT-19 Project. He received the ISRO's Young Scientist Merit Award for the year 2008, the ISRO's Team Excellence Award, in 2007, for contribution to "Implementation of Disaster Warning and Communication System," and the ISRO's Team Excellence Award, in 2015, for "MSS Ground Segment Development and Field Trial for GSAT-6 Satellite."

...