IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received January 20, 2020, accepted February 9, 2020, date of publication February 12, 2020, date of current version February 20, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2973392

High-Performance Distributed Compressive Video Sensing: Jointly Exploiting the HEVC Motion Estimation and the $\ell_1 - \ell_1$ Reconstruction

RUIFENG ZHANG, SHAOHUA WU, (Member, IEEE), YE WANG, AND JIAN JIAO, (Member, IEEE)
Harbin Institute of Technology at Shenzhen, Shenzhen 518055, China
Pengcheng Laboratory, Shenzhen 518055, China

Corresponding author: Shaohua Wu (hitwush@hit.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61871147, Grant 61831008, and Grant 91638204, in part by the Shenzhen Municipal Science and Technology Plan under Grant JCYJ20170811160142808, Grant JCYJ20170811154309920, and Grant ZDSYS201707280903305, in part by the project The Verification Platform of Multi-Tier Coverage Communication Network for Oceans under Grant PCL2018KP002, and in part by the Guangdong Science and Technology Planning Project under Grant 2018B030322004.

ABSTRACT The distributed compressive video sensing (DCVS) system combines the advantages of compressed sensing (CS) and distributed video coding (DVC), suitable for the limited-resource video sensing and transmission environment. In this paper, we propose a comprehensive high performance DCVS system. First, we introduce the BM3D-AMP algorithm reconstruct key (K) frames. Second, we propose a new high efficiency video coding (HEVC) motion estimation (ME) algorithm with motion vector (MV) prediction method. By integrating the segmentation idea and motion estimation, this algorithm can gets more accurate side information (SI). Finally, we propose the $\ell_1 - \ell_1$ minimization model to achieve non-key (NK) frames joint high-quality reconstruction. We utilize the alternating direction method of multipliers (ADMM) algorithm to solve it. With the idea of dividing and conquering, the general problem is decomposed into several smaller pieces. Experimental results demonstrate that the proposed system has significant improvement over its counterparts.

INDEX TERMS Distributed compressive video sensing, side information, HEVC motion estimation, $\ell_1 - \ell_1$ minimization.

I. INTRODUCTION

With the development of Internet of Multimedia Things (IoMT) and unmanned aerial vehicle (UAV) technology, how to transmit a large number of video signals quickly and safely has become a research hotspot. In these cases, the video channel conditions are asymmetrical and resource environments are constrained. The widely used video compression algorithms MPEG-2/4 and h.26x adopt the ideas of nonlinear compression and linear decompression, which bring great pressure to the encoding end [1], [2]. Therefore, the traditional video coding technology is not suitable for these applications, and it is necessary to seek for a new coding compression method.

The associate editor coordinating the review of this manuscript and approving it for publication was You Yang.

Distributed video coding (DVC) [3] has good noise immunity and low coding complexity. Through ''intra-frame encoding and inter-frame decoding'', the system implements efficient compression on encoding side and relatively high-quality reconstruction on decoding side. Compared with the traditional method, such as the video decoding code standard H.264 [4], the workload of the encoding end is transferred to the decoding end. In order to further improve the rate distortion (RD) performance and coding efficiency in DVC, a new decoder is proposed in [5]. The decoder can recursively carry out the decoding process on the augmented factor graph, and in each frame, the recursive message passing algorithm is used for decoding the bit-planes.

In 2006, Donoho et al. proposed the theory of compressed sensing (CS) [6], [7], which means the sparse original signal can be reconstructed from low-dimensional observations.

Due to the similarity redundancy of the video frames, video image data can be regarded as sparse signals, which theoretically provides the possibility of entirely refactoring from the compressed value. However, traditional CS algorithms, such as convex optimization algorithms, do not make full use of prior information, resulting in certain deficiencies in the computational complexity and reconstruction accuracy. In [8], the authors first combined the CS principle with DVC scheme, designed the CS-based video coder. Subsequently, the author in [9] systematically combined CS with DVC, and proposed the Distributed Compressive Video Sensing (DCVS) system.

The DCVS system groups the video sequence frames at the encoding end. Generally, the first frame of group is considered as key (K) frame, and the rest are set to non-key (NK) frames. Then the compression processing is conducted respectively. Compared with traditional coding methods, the DCVS scheme firstly reduces the complexity of the encoding. Motion estimation and correlation calculations between video frames are transfered to the decoder. Secondly, the encoding end of the DCVS does not need to encode or propagate the prediction information, which can bring better anti-noise and fault-tolerance performance. On the decoding side, the reconstruction K frames are obtained by compression values and the corresponding CS reconstruction algorithm. Then the reconstructed one will be used to provide a side information (SI) for NK frames reconstruction.

In [9], the authors noted that the side information can be regarded as the non-key frames with noise, and generated by the motion compensation interpolation method. On this basis, the authors in [10] presented an enhanced side information algorithm for obtain precise motion vectors (MV) [11] through log search algorithm. In recent years, the motion estimation algorithms [12] are widely used to generate accurate SI frames. With the idea of difference fusion and redundant dictionary, the bidirectional motion estimation (BME) [13], [14] side information generation algorithm can be effectively used to reconstruct NK frames. In [15], the initial reconstruction K frames are used to improve the secondary reconstruction value. And based on motion estimation, a hypotheses set acquisition algorithm is proposed to further improve the overall performance. The author in ref. [16] utilized the motion similarity of the reconstructed non-key frames, key frames and the current frame to perform position cross-reconstruction algorithm. An adaptive prediction scheme is proposed to create SI. When the motion of the pixel block is weak, the linear combination of the key frames before and after the current frame are used as the prediction result, without motion estimation and compensation. In [17], the non-key frames compression values are used to correct SI errors, and use the motion compensated frame interpolation (MCFI) techniques to realize the compromise between computational complexity and estimation. Moreover, Fowler et al. proposed the MH-BCS-SPL [18], [19] algorithm, which integrate bidirectional motion estimation and multiple hypothesis (MH) prediction to generate SI. Based on MH prediction, the re-weighted residual sparsity (RRS) model [20] was proposed to further enhances iteratively reconstruction quality. However, these methods do not fully utilize the motion correction between the contiguous video frames in time and space domain.

In order to solve the above questions, we propose a High Efficiency Video Coding motion estimation (HEVC-ME) algorithm, with motion vector prediction techniques, make full use of the motion associations between adjacent video frames. The high performance motion estimation algorithm adopts quadtree image partition method and realizes multi-size estimation. HEVC-ME is the further development of the advanced video coding standard H.265/HEVC [21], [22]. It incorporates the new division ideas and prediction methods in motion estimation process. To further improve quality of the restored video, we introduce the $\ell_1 - \ell_1$ [23] minimization compression sensing reconstruction algorithm. The model weight coefficient β is dynamically adjusted by the NK frames compression rate. In other words, the model relies more on side information at low NK frames compression rates, and relies more on compressed NK frames at high compression rate. Moreover, we use alternating direction method of multipliers (ADMM) [24] algorithm to solve this kind of optimization problem. Through decomposition-coordination process, the overall issue is divided into multiple small local pieces that are easy to solve.

This paper was previously presented at the conference [25], with the following extensions. We propose a comprehensive high-performance video frame reconstruction framework, which outperforms the previous structure in terms of K and NK reconstruction algorithms. Through the in-depth analysis of the existed DCVS decoder, we select the BM3D-AMP algorithm for key frames reconstruction, and introduce the implementation process and denoising principle. More importantly, build on the $\ell_1 - \ell_1$ minimization, we propose a new non-key frames reconstruction model. The algorithm can make full use of the side information and compressed NK frames. With the change of compression rate, the relative accuracy of the two kinds of information changes. This algorithm can dynamically adjust the weight coefficients of information, and lean towards the more one. In brief, the main contributions of this paper are summarized as follows:

- We first introduce the BM3D-AMP into DCVS system for key frames reconstruction.
- We propose the new HEVC-ME algorithm with motion vector (MV) prediction method to improve the accuracy of side information.
- Based on $\ell_1 - \ell_1$ minimization, we propose a new NK frames reconstruction model, which realize the dynamic reliance on information.

The rest of this paper is organized as follows. In Section II, we briefly introduce the CS theory and the DCVS system framework. The detailed HEVC-ME algorithm and $\ell_1 - \ell_1$ minimization reconstruction algorithm are presented

in Section III. In Section IV, we provide the simulation results to verify the effectiveness of the proposed HEVC-ME and $\ell_1 - \ell_1$ minimization algorithms. Conclusions and future works are given in Section V.

II. RELATED WORK

In this section, we introduce the basics of the CS theory and the DCVS system framework, and then briefly describe the sate-of-art key frames reconstruction algorithm.

A. THE BASIC CS THEORY

According to compressed sensing theory, video image signals can be sparse represented or approximately sparse represented in a sparse domain. First, the video signals are splited to continuous video image. Then these image are expressed as one-dimensional by signal vectorization treatment, and these vectors are projected onto an incoherent compression matrix to obtain the compression value. At the decoding end, with these CS values, the sparse representation coefficient of the original signals can be obtained. Finally, the reconstructed video frames signal is obtained by sparse inverse transform. The block diagram of the CS system is shown in Fig. 1.

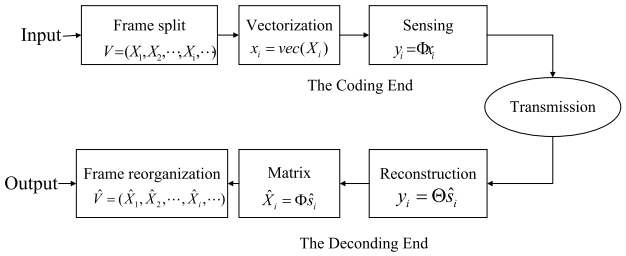


FIGURE 1. Block diagram of the video CS system.

The CS process can be described as follows:

$$y = \Phi x = \Phi \Psi s = \Theta s, \tag{1}$$

where $x_{N\times 1}$ denotes the signals in time domain, $y_{M\times 1}$ is the compressed result that be transmitted into decoder, and $s_{N\times 1}$ is the sparse signal reconstruction in the sparse domain. $\Phi_{M\times N}$ $(M \ll N)$ is the compression matrix, $\Psi_{N\times N} = [\Psi_1, \Psi_2, \cdots, \Psi_N]$ is the sparse representation basis, and $\Theta = \Phi\Psi$ is the sensing matrix. In addition, N is the dimension of the original vector, and M is the dimension of compressed elements obtained by sampling.

The essence of the decoding process is calculating N elements in the sparse representation coefficients s by the compression result y and the compression matrix Φ. Because $M \ll N$, the reconstruction process is a low-dimensional to high-dimensional process, equivalent to an infinite solution underdetermined equations:

$$\hat{s} = \arg\min_s \|s\|_0, \quad s.t.\ y = \Phi\Psi s \equiv \Theta s. \tag{2}$$

Eq. (2) is a NP-hard problem, then we convert it to l_1-norm problem to solve:

$$\hat{s} = \arg\min_s \|s\|_1, \quad s.t.\ y = \Phi\Psi s \equiv \Theta s. \tag{3}$$

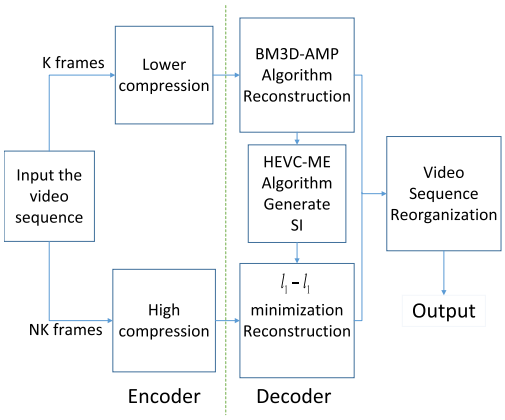


FIGURE 2. The proposed DCVS framework.

The problem in Eq. (3) can be solved by well-developed approaches, including the convex optimization algorithms (such as classical GPSR algorithm [26]) and the greedy algorithms.

B. THE DCVS SYSTEM FRAMEWORK

The DCVS system introduces the CS theory on the basis of DVC technology, reducing the complexity and computational pressure on the encoder side. The encoding end directly compresses and samples the signals without any other operations. The framework of the proposed high-performance DCVS system is illustrated in Fig. 2. The workflow of DCVS system can be summarized as follows:

At the encoding end, the video is first divided into frames of video images, then these frames are gathered together to form the group of pictures (GOP) frames. The first one of GOP frame is often considered to be the key frames, and the others are non-key frames. Video frames in each GOP are compressed in the order of K frames and NK frames respectively, and then transmitted to the decoding end through the channel. We use the scrambled block Hadamard ensemble (SBHE) matrix as the compression matrix Φ, which only contains 0, $+1$, -1 three elements. The compression matrix has universality, ensures the incoherence of video images and sparse basis. And the Hadamard diagonal structure also ensures the lowest computational complexity. It should be noted that during the compression process, the compression rate of the K frames must be greater than or equal to the NK frames. And the compression ration is defined as $CR = M/N$.

At the decoding end, two kinds of frames are respectively restored. First, K frames are directly obtained according to the compression value and the corresponding CS reconstruction algorithm. Then the reconstructed K frames are used to provide side information. With the prior information frames and compressed NK frames, we can use $\ell_1 - \ell_1$ minimization algorithm to reconstruct the NK frames.

III. THE PROPOSED HIGH EFFICIENT DCVS FRAMEWORK

In this section, we innovatively propose a efficient DCVS framework. We mainly focus on three points: how to

efficient reconstruct K frames, how to produce high-quality side information and how to optimal combine SI with compressed NK frames for decoding. For the first question, we introduce the BM3D-AMP algorithm for key frames reconstruction. For the second one, we proposed the HEVC-ME algorithm, which combines the motion estimation prediction methods with video coding standard HEVC. For the last one, we propose the $\ell_1-\ell_1$ minimization compression sensing algorithm to make high-performance reconstruction.

A. HIGH QUALITY K FRAMES RECONSTRUCTION ALGORITHM

Under the basic of compression sensing algorithms, we incorporate the Block Matching 3D (BM3D) [27] filtering noise removal model. This model utilizes the non-local self-similarity of images, uses the similar pixel blocks to get more accurate information. With the idea of non-local mean, the BM3D model divides video frames into several overlapping or non-overlapping small blocks. Each block will search the smallest Euclidean distance as a similar block. Then the similar pieces are grouped to perform the 3D transformation. Finally, the reconstructed K frames are obtained through the collaborative filtering, inverse transformation, coefficient aggregation and the final estimations.

The filtering process is composed of two phases. First, we perform hard threshold collaborative filtering on coefficients. Then the wiener filter using the basic estimated energy spectrum as the real one. Through two phases of filter estimate, BM3D can better preserves image clarity. The CS reconstruction model based on BM3D denoising model can be summed up as algorithm 1:

Algorithm 1 CS Reconstruction Model Based on BM3D

Require: sensing matrix $\Theta = \Phi\Psi$, compressed result y, BM3D denoising model

Ensure: reconstruction video frames $\hat{x}$

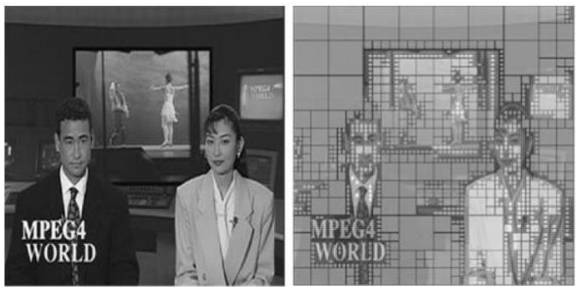
1: initialize the residuals: $z^0 = y, x^0 = 0, n = 0$
2: $n = n + 1$
3: calculate the current estimate of the signal:$x^n \leftarrow \eta_\tau(\Theta^T Z^t + \alpha^t)$
4: correct the results: $\varepsilon = STD(x^s), x \leftarrow D_{BM3D}(x^s, \varepsilon)$
5: update the residual: $Z^{t+1} = y - \Theta x^{t+1} + \frac{1}{\delta}\left\langle \eta'_\tau(\Theta^T Z T + x^t)\right\rangle$
6: judge the stop condition
7: **return** $\hat{x}$=$\Psi(x)$

B. HEVC-ME BASED MV PREDICTION ALGORITHM

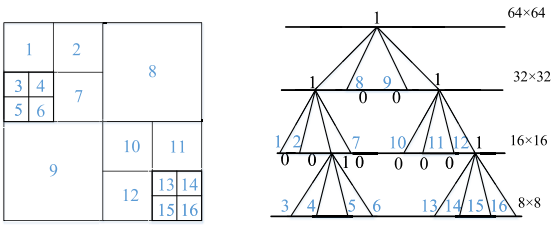
In DCVS system, non-key frames are usually compressed with high compression ratio at the encoding end. So few transmitted compressed NK frames can not be directly reconstruction with high quality. The side information algorithms compensate can greatly solve this question. When the video sequence motion intensity is large, SI precision is low, the encoding end needs to transmit more compressed NK frames. When SI has a high similarity with the frame to be reconstructed, the required decoding check value is small, then the NK frames can be compressed to a great extent. Therefore, accurate SI generation algorithm can greatly relieve the video frames compression rate.

Unlike traditional standards, in DCVS, motion estimation is only used to generate side information at the decoding end. As a prerequisite for generating side information, the motion vector (MV) is a very important parameter that determines the effect of motion estimation. Through bidirectional motion compensation, accurate MV can obtain more precise side information. H.265/HEVC video coding standard has made many improvements to the motion estimation part of predictive coding. H.265 still use block-based hybrid coding framework. With the quadtree method, coding block partition structure ensures the prediction of multi-size. Based on the new idea of H.265/HEVC in motion estimation, we propose an HEVC-ME side information generation algorithm in this section. We introduce the motion vector prediction to determine the accurate search starting point and range, realize the effective combination of HEVC and MV.

In the H.265/HEVC standard, one video frame is firstly divided into a number of no overlapping large coding units (LCU), with the size of 64×64. Each LCU is recursively divided into multiple coding units (CU) of size 32 × 32, 16 × 16 and 8 × 8 by quadtree method. The optimal division scheme is made according to the charateristics of image. As indicated in Fig.3 (a), small CUs partitions are applied to predict the detail regions with large luminance variations, and large CUs are used to predict areas with small background changes. Specifically, each CU performs a full search match in reference frame. Each LCU first using the coordinate position in the reference frame as the search center, searches in up, down, left, and right directions. Then the optimal matching



(a) CU division of different characteristic regions



(b) Specific quadtree partition map

FIGURE 3. HEVC quadtree partition CU schematic.

coding unit with the smallest SATD function value is selected. Then each LCU is divided into four coding units by quadtree method, with 32×32 size. Repeat the search method, sum the four optimal matches values and compare with LCU. If the SATD value of LCU scheme is smaller, it will be preserved. Otherwise, the divided scheme will be accepted. Each 32×32 coding unit is recursively divided into 16×16, 8×8 to judgment. The specific quadtree process is shown in Fig.3 (b).

The essence of motion estimation is to find the best matching block for each current frame pixel block in the reference frame, so the search result of block matching directly determines the effect of motion estimation. Motion estimation in traditional video coding standards usually occurs at the encoding end, and the effect of it is usually measured by the rate distortion (RD) function. That mean RD function needs to maintain a compromise between the subjective quality and the compressed. The DCVS system transfer motion estimation to the decoding end, and only to minimize the distortion of the predicted frame. Therefore, the rate-distortion function can be replaced by the cost function. The widely used sum of absolute differences (SAD) function function calculates and selects the smallest distance between the current pixel block and the reference frame. Based on this, the standard H.265/HEVC proposed the SATD function as judge means. The SATD function can simply expressed as the sum of absolute values after hadamard transformation. The algorithm performs hadamard transformation on the matrix Q obtained by the difference between the matching block found in the reference and the current block A, just multiplying the hadamard matrix left and right of Q. The sum of absolute values of the elements in matrix HDH is used as the basis for similarity determining.

Block matching search is another important step in motion estimation. The first step of block matching search is to determine the starting point in the reference frame and set the search range according to its location. Therefore, different starting points correspond to different search scopes, and greatly affects the search performance. To obtain more accurate SI, we propose the motion vector prediction algorithm, which hugely exploits the correlation of motion between adjacent coding units. This algorithm first establishes the candidate list for the current coding unit, with the MV of these nearby CU that have obtained the prediction results. By traversing the candidate MV and calculating each cost, the minimum one is selected as the optimal MV. The schematic diagram of the motion vector prediction algorithm is shown in Fig. 4. As we can see, the coding unit M_n is the current CU to be estimated, and the obtained MV from its left, upper left, upper and upper right are utilized to compensate the search starting point. Together with the scope determined by the traditional corresponding method, this results in six search scopes. Then block matching is performed in six ranges, and the MV corresponding to the minimum rate-distortion function value is taken as the optimal match.

After obtaining the motion vector of the current CU, we obtain the optimal prediction block by bi-direction motion

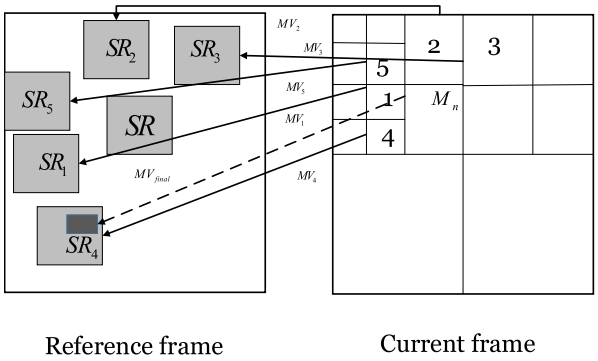


FIGURE 4. Motion vector prediction algorithm.

estimation and compensation. Taking GOP= 2 as an example, the previous and next of NK frame are K frames, and only the adjacent two reconstructed K frames exist at the decoding end. As shown in Fig. 5, the coding unit to be estimated in the unknown NK (t) frame is M_n. In the previous frame K ($t-1$) and the next frame K ($t+1$), the encoding unit with the same positions are M_{n-1} and M_{n+1} respectively. The coding unit M_{n-1} performs motion estimation in the K ($t+1$) frame, and the result deviates from the initial coordinate position by (i, j). The NK (t) frame is located at the intermediate equidistant position. According to the proportional scaling property of MV, it can be inferred that the M_n coding unit will move to M_n' in K (t + 1) frame, with $(i/2, j/2)$ offset. Similarly, we can obtain the motion vector $(i/2, j/2)$ and the estimated block M_n'' through backward motion estimation. By averaging M_n' and M_n'', the final estimation result of one unit can be obtained. By traversing all the coding units in one video frame, the predicted one is obtained.

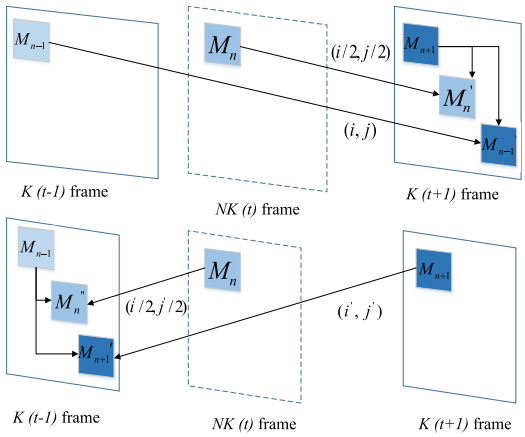


FIGURE 5. Motion estimation and compensation processes.

The complete process of our proposed HEVC-ME is shown in Algorithm 2.

C. $\ell_1 - \ell_1$ MINIMIZATION NK FRAMES RECONSTRUCTION ALGORITHM

Due to the high power compression of non-key frames under the DCVS framework, the compressed signal transmitted to the decoding end is relatively small. It's not reliable to

Algorithm 2 HEVC-ME Algorithm

Require: current video frame imgP, reference video frame;
Ensure: motion vector (MV)

1: assume (i, j) as the starting point with a length of *CUSize*
2: **for** m,n $= -\frac{\text{CUSize}}{2} + 1$ to $\frac{\text{CUSize}}{2} + 1$ **do**
3: $Q_{ij} = \text{imgP}(i, j) - \text{imgI}(i + m, j + n)$
4: $S_{ij} = HQ_{ij}H$
5: **end for**
6: $S = \min \sum_{i,j} S_{i,j}$
7: $\text{MV}_{i,j} = (i, j)$
8: $S_{\text{sum}} = \sum_{n=1}^{4} S_n$
9: length increase to $2 \times CUSize$, got $S_{2\text{CUSize}}$
10: **if** $(S_{\text{sum}} < S_{2\text{CUSize}})$ **then**
11: division
12: **else**
13: retain
14: **end if**
15: **return** The motion vector MV.

directly use the compressed value for reconstruction. In order to combine the side information with the NK frames compression value effectively, we first introduce the $\ell_1 - \ell_1$ minimization reconstruction model. And then we focus on the specific steps of weight coefficient selection and algorithm implementation.

As described in section II-*A*, the compressed video frames can obtain the reconstruction result of original signal by solving ℓ_1 norm problem. Assuming that $x \in R^n$ is an unknown original sparse signal with s non-zero values, and the side information $w \in R^n$ is highly similar to x, which can be regarded as the prior information of it. We aim to reconstruct the signal x successfully using the compression value $b = Ax$, A is the compression matrix and w is the prior information. We introducing a function $g(\cdot)$ to characterize the degree of similarity between the original signal x and the prior information w, and the more similar they are, the smaller the value of the function. An additional weight coefficient β is also introduced, which value depends on the relative accuracy of the compression signal and the prior information. When the compression rate of the encoding end is low, the NK frames existing at the decoding end have less check information and poor check performance. But the side information obtained by the HEVC-ME algorithm are more accurate. Therefore, the reconstruction process should rely more on the information provided by the SI frames, and the value of $\beta > 1$ is greater than 1. As the compression rate of the NK frames increase, the compression value of the NK frames obtained by the decoding end gradually increases. The information provided by the NK sampling result are equivalent to the accuracy of the SI frames, i.e. the NK sample value and the SI are equally important in the reconstruction process, and $\beta = 1$. When the NK frames compression rate continues to increase, the compression value at the decoding end are more reliable than the SI frames, and the β value is further reduced.

So the compressed sensing reconstruction problem of joint prior information can be described as:

$$\underset{x}{\text{minimize}} \|x\|_1 + \beta g(x - w), \quad \text{s.}t.\ Ax = b. \tag{4}$$

The equation can be further modified to:

$$\underset{x}{\text{minimize}} \|x\|_1 + \beta \|x - w\|_1, \quad \text{s.}t.\ Ax = b. \tag{5}$$

We use alternating direction method of multipliers (ADMM) algorithm to solve Eq. (5). Through the decomposition and coordination process, the ADMM algorithm decomposes large global problem into several small local sub-problems that are easy to be solved. And the solution of the global problem is obtained by coordinating the solution of each sub-problem. First, we convert Eq. (5) as two problems:

$$\underset{x,y}{\text{minimize}} f(x) + g(y), \quad \text{s.t. } x = y, \tag{6}$$

where

$$f(x) = \|x\|_1 + \beta \|x - w\|_1, \quad g(y) = i_{x:b=Ax}(y),$$
$$i_s(x) = \begin{cases} 0 & x \in S \\ +\infty & x \notin S \end{cases}$$

Then the augmented Lagrangian form can written as:

$$L_\rho(x, y; \lambda) = f(x) + g(y) + \lambda^T(x - y) + \frac{\rho}{2}\|x - y\|^2, \tag{7}$$

where λ is the Lagrange coefficients, ρ is a penalty factor that usually takes a value of 1. The ADMM algorithm constantly updates the value of x, y, λ by iterating over the solution. The interation is stopped until the optimization condition and the stop criterion are satisfied, and the optimal solution is obtained.

In step 2, x^k is updated by assigning each element in the vector. Let $v = \lambda^k - \rho y^k$, then $x^{k+1} = \arg\min_x \|x\|_1 + \beta\|x - w\|_1 + v^T x + \frac{\rho}{2}\|x\|^2$. For each element of the vector x_k, there is $x_i^{k+1} = \arg\min_{x_i} |x_i| + \beta|x_i - w_i| + v_i x_i + \frac{\rho}{2}x_i^2$, and its value is closely related to the corresponding position pixel value in side information w. Therefore, x_i can be updated according to the positive and negative of each element in the side information w and the value of the variable v in each iteration. We then get:
while $W_i > 0$:

$$x_i^* = \begin{cases} \frac{1}{\rho}(-\beta - 1 - v_i) & v_i < -\rho w_i - \beta - 1 \\ w_i & -\rho w_i - \beta - 1 \le v_i \le -\beta w_i + \beta - 1 \\ \frac{1}{\rho}(\beta - 1 - v_i) & -\beta w_i + \beta - 1 < v_i < \beta - 1 \\ 0 & \beta - 1 \le v_i \le \beta + 1 \\ \frac{1}{\rho}(\beta + 1 - v_i) & v_i > \beta + 1 \end{cases}$$

Algorithm 3 Video CS Reconstruction Algorithm Based on $\ell_1 - \ell_1$ Minimization Model

Require: compression matrix $\Theta = \Phi\Psi$, compression value b, side information w

Ensure: reconstructed original signal $\hat{x}$

1: initialization: $\lambda \leftarrow 0$, $\rho \leftarrow 1$, $r_{prim} \leftarrow 0$, $s_{dual} \leftarrow 0$, $k \leftarrow 0$
2: fix y^k, λ^k, update x^k : $x^{k+1} \leftarrow \arg\min_x \|x\|_1 + \beta\|x - w\|_1 \ (\lambda^k - \rho y^k)^T x + \frac{\rho}{2}\|x\|^2$
3: fix x^k, λ^k, update y^k $z = \frac{1}{\rho}(\lambda^k + \rho x^{k+1})$ $y^{k+1} \leftarrow z - A^T(AA^T)^{-1}(Az - b)$
4: update λ^k : $\lambda^{k+1} \leftarrow \lambda^k + pr\_prim$
5: update ρ
6: **if** $\|r\_prim\| > \tau\|s\_dual\|$, **then**
7: $\quad \rho \leftarrow \mu\rho$
8: **else**
9: $\quad \rho \leftarrow \frac{\rho}{\mu}$
10: **end if**
11: **if** $\|r\_prim\| \le 10^{-3}$ or $\|s\_dual\| \le 10^{-3}$ **then**
12: $\quad$ stop iteration
13: **else**
14: $\quad$ skip back to step 2
15: **end if**
16: **return** reconstructed original signal: $\hat{x} \leftarrow x^{k+1}$, $\hat{x} = \Psi x^*$

while $W_i < 0$,

$$x_i^* = \begin{cases} \frac{1}{\rho}(-\beta - 1 - v_i) & v_i < -\beta - 1 \\ 0 & -\beta - 1 \le v_i \le -\beta + 1 \\ \frac{1}{\rho}(-\beta - 1 - v_i) & -\beta + 1 < v_i < -\beta w_i - \beta + 1 \\ w_i & -\beta w_i + \beta - 1 \le v_i \le -\beta w_i - \beta + 1 \\ \frac{1}{\rho}(\beta + 1 - v_i) & v_i > \beta w_i - \beta + 1 \end{cases}$$

IV. EXPERIMENTAL RESULTS

In this section, we conduct extensive experimental evaluations for the proposed high quality DCVS framework. For evaluating performance of reconstruction K frames, we compare our approach with two conventional methods: the GPSR [26] method and the AMP [31] method. For evaluating performance of generated SI, we compare our approach to the widely used BME [13] and latest MH-BCS-SPL [18]. For evaluating performance of reconstruction NK frames, we compare our approach to the state-of-art BM3D-SAPCA-AMP [28] and the frequently-used BM3D-AMP, GPSR methods [26]. In Section IV-A, we introduce the configuration and settings of our experiments. In section IV-B, IV-C and IV-D, we compare the performance for K, SI and NK methods, respectively.

A. CONFIGURATION AND SETTINGS

All experiments are performed on a desktop computer with Intel core i7-8700K CPU @3.70GHz, 16GB RAM and the Windows 10 Enterprise 64-bit operating system. In order to prove the universality of our method, we select six representative test video sequences, which contained video frames of different image resolution and various change degrees. Each sequence we intercept the first 25 frames, and the size of GOP is set to 2. For sequence "Akiyo", we select relatively redundant part with little difference between frames. For sequence "Bus", we choose the fast changing segment. The pixels at same position in adjacent frames are always changing drastically. And for each frame, the sequence are complicated and difficult to reconstruct. For sequence "bowing" and "Foreman", frames have a varying degree of variation. The sequence "Rush-field-cuts" and "Fourpeople" are of high definition (HD) images. The sparse basis is set to common DWT transform, and the compression matrix is the SBHE matrix. Reconstruction quality is reported as the average Peak Signal-to-Noise Ratio (PSNR) and actually visual effect.

$$PSNR = 10\log_{10}\left(\frac{(2^n - 1)^2}{MSE}\right), \tag{8}$$

B. KEY FRAMES CONSTRUCTION

This section evaluates the performance of our BM3D-AMP method application in K frames. The evaluation mainly focuses on the reconstruction quality and image visual effect. Then two classical algorithms are compared with our method, including [26] and [31].

1) RECONSTRUCTION QUALITY

Fig.6 depicts the "Foreman" K frames reconstruction quality under different algorithms with the change of compression rate. At each rate, we simulate all 25 frames and take the average results. In addition, table 1 describes the exact reconstruction values for four test sequences. Simulation result shows that the BM3D-AMP algorithm always has the best performance.

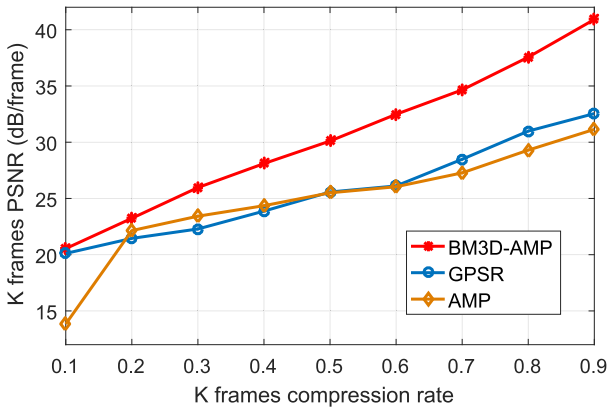


FIGURE 6. PSNR of different algorithms for K frames.

2) VISUAL EFFECT

Due to the characteristics of frame group division in DCVS system, the compression rate of key frames is higher than that of non-key, and is usually set to greater than 0.5. According to Table 1, the GPSR works better at high compression rates.

TABLE 1. PSNR of different algorithms for K frames.

Sequence		K frames compression rate							
		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8
Akiyo	BM3D-AMP	22.34	26.05	29.56	33.27	36.01	38.92	42.25	45.62
	GPSR	21.82	23.64	25.84	28.46	30.86	32.86	34.75	37.05
	AMP	19.44	23.72	25.65	27.97	29.48	31.60	32.97	34.27
Bus	BM3D-AMP	17.24	19.97	21.36	23.61	25.54	27.31	28.91	31.15
	GPSR	17.00	19.29	20.13	21.28	22.09	23.30	24.62	26.94
	AMP	16.87	18.59	19.71	20.54	21.28	21.90	22.44	22.88
Bowing	BM3D-AMP	20.89	24.14	26.33	28.65	31.02	33.57	36.28	38.66
	GPSR	20.62	23.11	24.08	24.79	25.52	26.11	28.69	30.89
	AMP	18.50	22.54	24.57	24.89	25.36	25.88	26.74	28.85
Foreman	BM3D-AMP	21.07	23.24	26.74	28.42	30.84	33.12	35.24	38.19
	GPSR	20.81	22.02	22.96	24.37	25.48	26.05	28.64	30.77
	AMP	14.75	22.68	23.83	24.94	25.45	25.96	27.01	29.41
Rush-field-cuts	BM3D-AMP	19.65	20.39	21.61	22.99	24.44	26.29	28.50	31.01
	GPSR	17.89	19.23	20.44	21.35	22.59	24.25	25.14	25.98
	AMP	16.00	18.08	19.83	20.94	21.45	22.17	23.01	24.21
Fourpeople	BM3D-AMP	18.81	21.33	23.30	25.38	27.47	29.60	31.81	34.52
	GPSR	18.28	20.33	21.27	22.28	24.13	25.07	25.84	27.20
	AMP	18.02	19.65	20.75	21.27	22.08	22.97	24.20	25.44

So we only compare the image visual effect of GPSR and BM3D-AMP. For each test sequence, we select one GOP to analyze. The compression rate is set to 0.7. As shown in Fig.7, BM3D-AMP preserves the details of video frames well, and the brightness of image is significantly clearer and less noise. So we use the BM3D-AMP as the K frames reconstruction algorithm in our subsequent experiments.

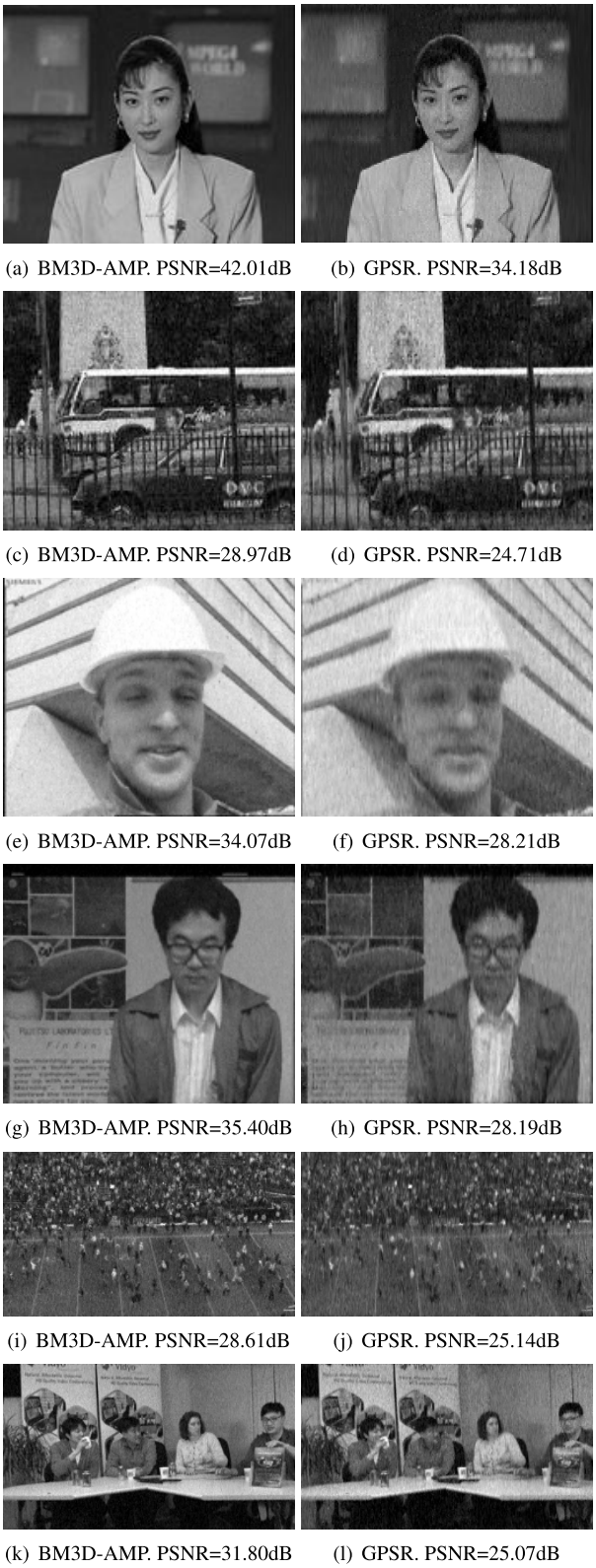
C. SIDE INFORMATION GENERATION

This section evaluates the quality of side information. Through the reconstructed K frames generated by BM3D-AMP, we use the HEVC-ME, BME [13] and MH-BCS-SPL [18], [19] algorithms to generate SI separately. The compression rate of K frames are always set to 0.7.

We first performed the simulation on "Foreman" sequence. Fig.8 directly reflects the performance of side information generated by different algorithms. The advantage of our method is that the obtained SI can retain reference frame detail information, which requires the reconstruction K frames are sufficiently clear. BM3D-AMP algorithm introduced in section IV-A greatly meets the requirement. Based on this, HEVC-ME improves the side information quality by $1 \sim 4$ dB. Our algorithm has some limitations in fast changing frame segments. Such as 15 and 17 frames, large displacements occur between the pixel blocks in a single plane, and our advantage is no longer obvious. Then we use BM3D-AMP to jointly reconstruct the non-key frames. Fig. 9 verifies the effectiveness of HEVC-ME algorithm for NK frames reconstruction, with about 2.5 dB improvement than MH-BCS-SPL and 8 dB than BME. Similarly, we perform experiments on all four different characteristic sequence. The SI results are shown in Table 2. In most cases, our algorithm maintains a relatively large lead.

D. CS RECONSTRUCTION ALGORITHM BASED ON JOINT PRIOR INFORMATION

This section evaluates the performance of our $\ell_1 - \ell_1$ minimization model application in joint NK frames reconstruction. Quality and visual effect are our core concerns.



(a) BM3D-AMP. PSNR=42.01dB (b) GPSR. PSNR=34.18dB
(c) BM3D-AMP. PSNR=28.97dB (d) GPSR. PSNR=24.71dB
(e) BM3D-AMP. PSNR=34.07dB (f) GPSR. PSNR=28.21dB
(g) BM3D-AMP. PSNR=35.40dB (h) GPSR. PSNR=28.19dB
(i) BM3D-AMP. PSNR=28.61dB (j) GPSR. PSNR=25.14dB
(k) BM3D-AMP. PSNR=31.80dB (l) GPSR. PSNR=25.07dB

FIGURE 7. Comparison of image visual effects.

The state-of-art algorithm BM3D-SAPCA-AMP [28] and widely-used BM3D-AMP are set to control group. When using the contrast reconstruction algorithms, an algorithm based on the difference fusion idea is used to achieve the

TABLE 2. PSNR of different algorithms for side information.

Sequence		NK frames order											
		1	3	5	7	9	11	13	15	17	19	21	23
Akiyo	HEVC-ME	39.42	29.72	37.80	26.72	36.02	33.73	39.39	36.64	36.30	35.47	39.56	40.99
	MH-BCS-SPL	34.12	25.84	33.40	25.67	31.89	29.76	34.57	32.29	32.11	31.97	34.55	34.90
	BME	30.89	24.80	30.47	22.80	28.48	25.76	30.50	27.44	27.05	26.78	29.87	29.63
Bus	HEVC-ME	22.15	22.27	22.18	21.79	22.25	22.45	21.51	22.08	21.31	21.21	22.09	21.95
	MH-BCS-SPL	22.05	22.16	21.89	21.37	22.00	22.15	21.19	21.50	20.58	20.41	20.94	21.01
	BME	20.00	20.10	19.90	19.80	19.55	19.54	19.44	19.32	19.19	19.11	19.09	19.03
Bowing	HEVC-ME	31.43	29.67	26.32	23.59	23.87	26.22	28.60	34.08	36.88	37.80	34.35	29.17
	MH-BCS-SPL	29.42	27.91	25.43	23.36	23.45	25.79	27.34	30.89	31.12	31.86	30.88	27.64
	BME	26.37	25.33	24.16	22.75	22.98	24.68	25.08	26.43	27.80	27.86	27.55	26.43
Foreman	HEVC-ME	30.92	31.98	28.54	29.60	29.32	24.87	29.56	20.95	19.67	31.68	30.40	28.77
	MH-BCS-SPL	27.64	28.25	26.50	28.87	27.69	23.52	28.11	22.73	22.04	28.76	28.68	26.43
	BME	24.62	25.93	24.47	25.47	25.60	23.48	25.80	19.21	19.17	26.28	26.01	25.76
Rush-field-cuts	$\ell_1 - \ell_1$	28.68	28.63	28.52	28.52	28.44	28.74	28.36	28.38	28.40	28.36	28.45	28.29
	MH-BCS-SPL	26.10	26.07	26.03	26.06	26.07	26.12	26.06	26.06	26.10	26.09	26.07	26.02
	BME	25.44	25.35	25.28	25.13	25.17	25.21	25.13	25.12	25.12	25.17	25.09	25.04
Fourpeople	$\ell_1 - \ell_1$	23.78	27.06	31.05	29.18	31.04	30.02	30.75	30.10	20.00	19.35	30.63	32.80
	MH-BCS-SPL	22.85	25.42	27.28	26.59	27.19	26.87	27.28	26.97	19.72	19.31	27.08	28.40
	BME	22.20	24.98	26.35	25.67	26.02	25.75	26.09	25.88	19.21	18.74	25.83	26.67

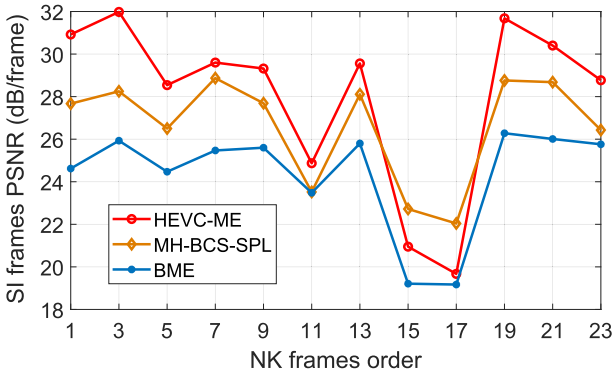


FIGURE 8. PSNR of generated SI by different algorithms.

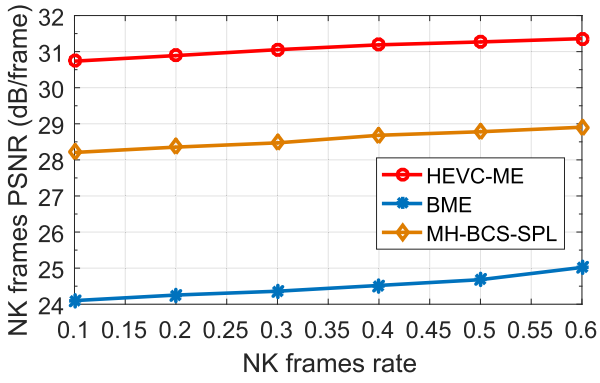


FIGURE 9. Effects of different SI generation algorithms on NK frames.

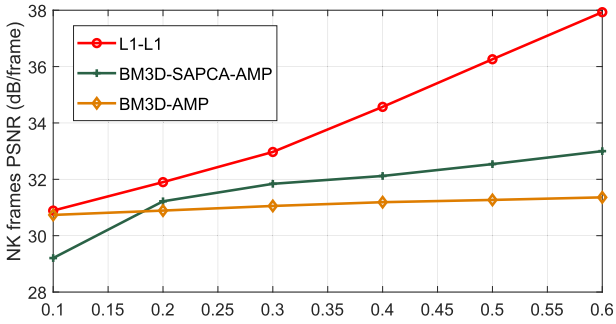


FIGURE 10. PSNR of different resconstruction algorithms for NK frames.

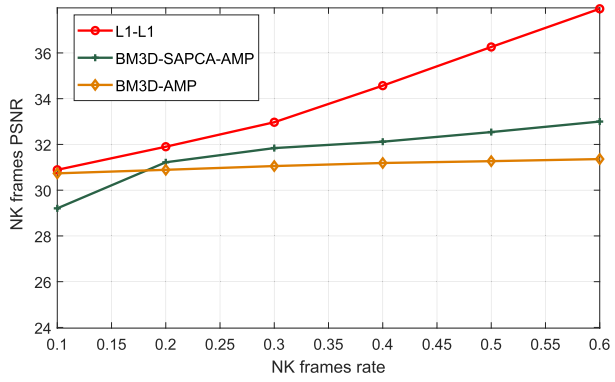
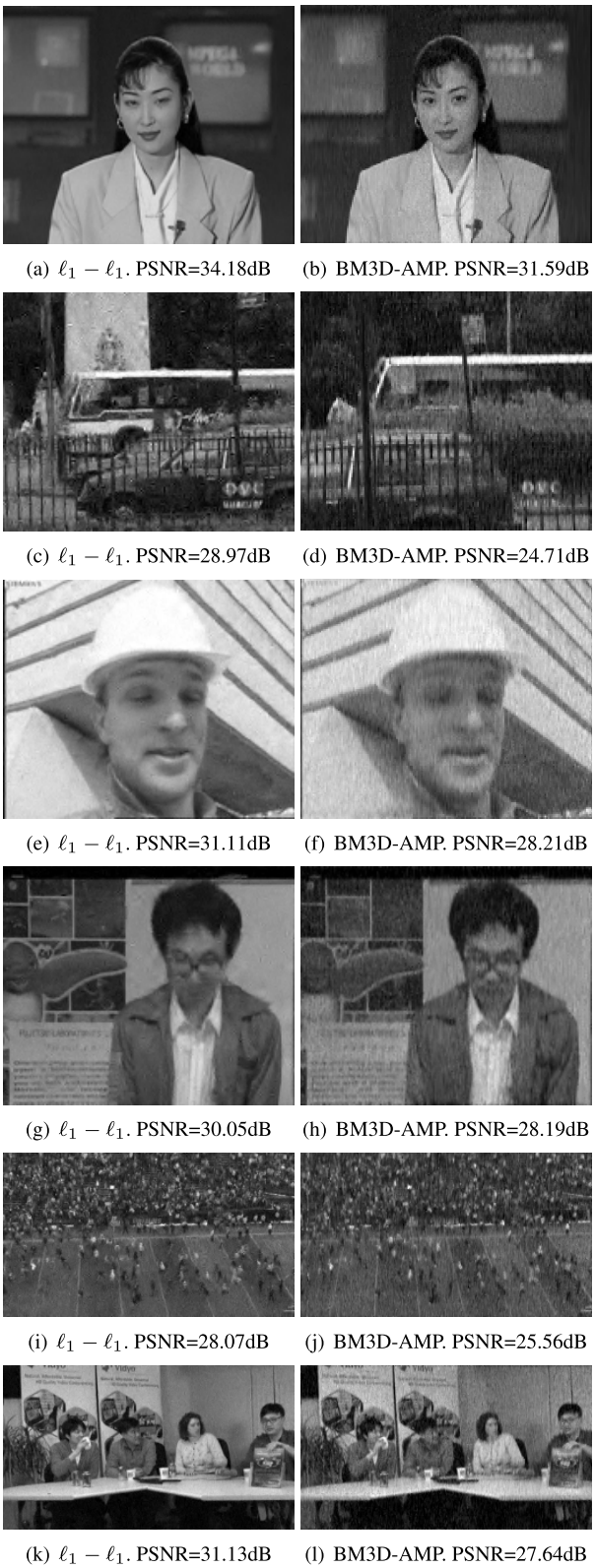


FIGURE 11. Different NK frames algorithms on resconstruction DCVS system.

combination of side information and NK frames. As mentioned above, we use BM3D-AMP, HEVC-ME as the K frames reconstruction and SI generation algorithms. The K frames compression rate are set to 0.7.

1) RECONSTRUCTION QUALITY

As shown in Fig.10, we compare the NK frames of "Foreman" sequence reconstructed by three algorithms. The algorithm based on $\ell_1 - \ell_1$ minimization has the best effect, and the reconstructed PSNR value is $1 \sim 5$ dB higher than the counterparts. When NK frames compression rate is 0.1, our method PSNR value is still greater than 30 dB. The overall reconstruction quality for DCVS system are illustrated in Fig.11. When the NK frames at low compression rate, our method overall PSNR value remains above 32 dB. And as the compression rate increases, its advantages are even more significant. Other sequence NK frames PSNR value are shown in Table 3. In order to adapt the constrained resource environment, the non-key frames should be compressed to the maximum extent, so we only conduct the low compression rate 0.1, 0.2 and 0.3.



(a) $\ell_1 - \ell_1$. PSNR=34.18dB (b) BM3D-AMP. PSNR=31.59dB

(c) $\ell_1 - \ell_1$. PSNR=28.97dB (d) BM3D-AMP. PSNR=24.71dB

(e) $\ell_1 - \ell_1$. PSNR=31.11dB (f) BM3D-AMP. PSNR=28.21dB

(g) $\ell_1 - \ell_1$. PSNR=30.05dB (h) BM3D-AMP. PSNR=28.19dB

(i) $\ell_1 - \ell_1$. PSNR=28.07dB (j) BM3D-AMP. PSNR=25.56dB

(k) $\ell_1 - \ell_1$. PSNR=31.13dB (l) BM3D-AMP. PSNR=27.64dB

FIGURE 12. Comparison of image visual effects.

2) VISUAL EFFECT

Fig. 12 displays the NK frames visual image reconstructed by different algorithms when the compression rate is 0.1.

TABLE 3. PSNR of different algorithms for NK frames.

Sequence		NK frames compression rate		
		0.1	0.2	0.3
Akiyo	$\ell_1 - \ell_1$	35.15	38.42	40.50
	BM3D-SAPCA-AMP	32.47	35.86	37.39
	BM3D-AMP	31.29	34.02	35.89
Bus	$\ell_1 - \ell_1$	23.78	25.64	27.71
	BM3D-SAPCA-AMP	23.26	24.38	25.71
	BM3D-AMP	22.42	23.52	24.77
Bowing	$\ell_1 - \ell_1$	30.68	33.41	34.84
	BM3D-SAPCA-AMP	29.57	31.08	32.77
	BM3D-AMP	28.89	30.46	31.68
Foreman	$\ell_1 - \ell_1$	31.87	33.43	34.77
	BM3D-SAPCA-AMP	29.54	31.65	31.90
	BM3D-AMP	30.14	31.58	31.74
Rush-field-cuts	$\ell_1 - \ell_1$	28.03	28.45	29.87
	BM3D-SAPCA-AMP	25.80	26.12	26.78
	BM3D-AMP	25.44	25.87	26.03
Fourpeople	$\ell_1 - \ell_1$	31.24	32.58	33.15
	BM3D-SAPCA-AMP	27.98	28.61	29.44
	BM3D-AMP	27.74	28.34	28.98

The images reconstructed by $\ell_1 - \ell_1$ mininization preserve the details clear and have less noise.

V. CONCLUSION

In this paper, we design a high-performance DCVS system, which suitable for the asymmetric node resource video transmission scenarios.

Firstly, the BM3D-AMP algorithm is introduced to reconstruct key frames. Secondly, we put forward a new HEVC-ME side information generation algorithm. Combined with the motion vector prediction, the HEVC-ME can greatly improves the quality of the obtained SI. Actually, it provides a possible to improve the quality of NK frames from the prior information aspect. Finally, for effiective integrate SI with NK frames compression results, we proposed a $\ell_1 - \ell_1$ minimization reconstruction model. As the compression rate changes, the model dynamically relies on more accurate information, and the reconstruction accuracy is significantly improved.

There are two promising directions for the future works. Our work, at current stage, mainly focuses on high quality SI generation and joint decoding algorithm. In image deep learning area, some techniques have been proposed to accelerate the reconstruction speed [29]. And how to combine these models with our algorithm is a meaningful work. In addition to directly operating the image, we can further mining the features of video images to improve the recovery level [30].

REFERENCES

[1] J. Ostermann, J. Bormans, P. List, D. Marpe, M. Narroschke, F. Pereira, T. Stockhammer, and T. Wedi, "Video coding with H.264/AVC: Tools, performance, and complexity," *IEEE Circuits Syst. Mag.*, vol. 4, no. 1, pp. 7–28, 1st Quart., 2004.

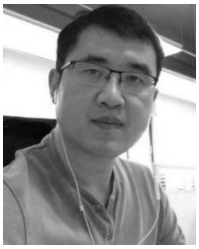
[2] V. Lappalainen, A. Hallapuro, and T. D. Hamalainen, "Complexity of optimized H.26L video decoder implementation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 717–725, Jul. 2003.

[3] B. Girod, A. M. Aaron, S. Rane, and D. Rebollo-Monedero, "Distributed video coding," *Proc. IEEE*, vol. 93, no. 1, pp. 71–83, Jan. 2005.

[4] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the scalable video coding extension of the H. 264/AVC standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 9, pp. 1103–1120, Sep. 2007.

[5] Y. Mohammad Taheri, M. O. Ahmad, and M. N. S. Swamy, "A joint correlation noise estimation and decoding algorithm for distributed video coding," *Multimedia Tools Appl.*, vol. 77, no. 6, pp. 7327–7355, Mar. 2018.
[6] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.
[7] E. J. Candes and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 21–30, Mar. 2008.
[8] J. Prades-Nebot, Y. Ma, and T. Huang, "Distributed video coding using compressive sampling," in *Proc. IEEE Picture Coding Symp.*, May 2009, pp. 1–4.
[9] L.-W. Kang and C.-S. Lu, "Distributed compressive video sensing," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Apr. 2009, pp. 1169–1172.
[10] V. Angayarkanni, V. Akshaya, and S. Radha, "Distributed compressive video coding using enhanced side information for WSN," in *Proc. IEEE Int. Conf. Wireless Commun., Signal Process. Netw.*, Mar. 2016, pp. 1133–1136.
[11] A. Zheng, Y. Yuan, H. Zhang, H. Yang, P. Wan, and O. C. Au, "Motion vector fields based video coding," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2015, pp. 2095–2099.
[12] S. M. Basha and M. Kannan, "Design and implementation of low-power motion estimation based on modified full-search block motion estimation," *J. Comput. Sci.*, vol. 21, pp. 327–332, Jul. 2017.
[13] T. Chen, "Adaptive temporal interpolation using bidirectional motion estimation and compensation," in *Proc. Int. Conf. Image Process.*, Jun. 2003, vol. 2, no. 3, p. 2.
[14] N. Van Thang and H.-J. Lee, "An efficient non-selective adaptive motion compensated frame rate up conversion," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, May 2017, pp. 1–4.
[15] J. Chen, F. Xue, and Y. Kuo, "Distributed compressed video sensing based on key frame secondary reconstruction," *Multimedia Tools Appl.*, vol. 77, no. 12, pp. 14873–14889, Jun. 2018.
[16] H. Dong, B. Zhuang, F. Su, and Z. Zhao, "A novel distributed compressive video sensing based on hybrid sparse basis," in *Proc. Vis. Commun. Image Process. Conf.*, Dec. 2014, pp. 320–323.
[17] J. Xu, Y. Qiao, and Q. Wen, "Rate-distortion optimized distributed compressive video sensing," *IEICE Trans. Fundam.*, vol. E99.A, no. 6, pp. 1272–1276, 2016.
[18] J. E. Fowler, S. Mun, and E. W. Tramel, "Block-based compressed sensing of images and video," *Found. Trends Signal Process.*, vol. 4, no. 4, pp. 297–416, 2012.
[19] E. W. Tramel and J. E. Fowler, "Video compressed sensing with multihypothesis," in *Proc. Data Compress. Conf.*, 2011, pp. 193–202.
[20] C. Zhao, S. Ma, J. Zhang, R. Xiong, and W. Gao, "Video compressive sensing reconstruction via reweighted residual sparsity," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 27, no. 6, pp. 1182–1195, Jun. 2017.
[21] G. J. Sullivan, J.-R. Ohm, W.-J. Han, and T. Wiegand, "Overview of the high efficiency video coding (HEVC) standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1649–1668, Dec. 2012.
[22] J.-R. Ohm, G. J. Sullivan, H. Schwarz, T. K. Tan, and T. Wiegand, "Comparison of the coding efficiency of video coding standards—Including high efficiency video coding (HEVC)," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1669–1684, Oct. 2012.
[23] T. Oishi and Y. Kuroki, "An $\ell_1 - \ell_1$-norm minimization solution using ADMM with FISTA," in *Proc. Int. Conf. Intell. Inform. Biomed. Sci. (ICIIBMS)*, vol. 3, 2018, pp. 200–203.
[24] Y.-J. Chiu, L. Xu, W. Zhang, and H. Jiang, "Decoder-side motion estimation and Wiener filter for HEVC," in *Proc. IEEE Global Conf. Signal Inf. Process. (GlobalSIP)*, Nov. 2013, pp. 512–516.
[25] Z. Li, S. Wu, M. Ma, J. Jiao, W. Wu, and Q. Zhang, "Improved distributed compressive video sensing based on HEVC motion estimation," in *Proc. IEEE/CIC Int. Conf. Commun. China (ICCC)*, Aug. 2018, pp. 819–823.
[26] M. A. T. Figueiredo, R. D. Nowak, and S. J. Wright, "Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems," *IEEE J. Sel. Topics Signal Process.*, vol. 1, no. 4, pp. 586–597, Dec. 2007.
[27] M. Lebrun, "An analysis and implementation of the BM3D image denoising method," *Image Process. Line*, vol. 2, pp. 175–213, Sep. 2012.
[28] C. A. Metzler, A. Maleki, and R. G. Baraniuk, "BM3D-AMP: A new image recovery algorithm based on BM3D denoising," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2015, pp. 3116–3120.
[29] Y. Yang, J. Sun, H. Li, and Z. Xu, "ADMM-CSNet: A deep learning approach for image compressive sensing," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 3, pp. 521–538, Mar. 2020.
[30] M. Xu, T. Li, Z. Wang, X. Deng, R. Yang, and Z. Guan, "Reducing complexity of HEVC: A deep learning approach," *IEEE Trans. Image Process.*, vol. 27, no. 10, pp. 5044–5059, Oct. 2018.
[31] D. L. Donoho, A. Maleki, and A. Montanari, "Message-passing algorithms for compressed sensing," *Proc. Nat. Acad. Sci. USA*, vol. 106, no. 45, pp. 18914–18919, Nov. 2009.

RUIFENG ZHANG was born in Shanxi, China, in 1996. He received the B.S. degree in electronics and information engineering from the Harbin Institute of Technology at Weihai, Weihai, China, in 2018. He is currently the Graduate Student with the Harbin Institute of Technology at Shenzhen, Shenzhen. His research interests include compressed sensing video image reconstruction and neural network decoding.



SHAOHUA WU (Member, IEEE) received the Ph.D. degree in communication engineering from the Harbin Institute of Technology, in 2009. From 2009 to 2011, he held a postdoctoral position at the Department of Electronics and Information Engineering, Shenzhen Graduate School, Harbin Institute of Technology, where he has been an Associate Professor, since 2012. From 2014 to 2015, he was a Visiting Researcher with the BBCR, University of Waterloo. He has authored or coauthored over 70 articles in these fields. He holds over 30 Chinese patents. His current research interests include wireless image/video transmission, deep space communication, advanced channel coding techniques, and 5G wireless transmission technologies.

YE WANG received the M.S. and Ph.D. degrees in information and communication engineering from the Harbin Institute of Technology (HIT), Shenzhen, China, in 2009 and 2013, respectively. From 2013 to 2014, he was a Postdoctoral Research Fellow with the University of Ontario Institute of Technology, Canada. He is currently an Assistant Professor with HIT at Shenzhen, Shenzhen. His research interests include satellite communications, resource allocation, and mobile internet.

JIAN JIAO (Member, IEEE) received the M.S. and Ph.D. degrees in communication engineering from the Harbin Institute of Technology (HIT) at Shenzhen, Shenzhen, China, in 2007 and 2011, respectively. From 2011 to 2015, he was a Postdoctoral Research Fellow with the Communication Engineering Research Centre, Shenzhen Graduate School, HIT at Shenzhen. From 2016 to 2017, he was with the School of Electrical and Information Engineering, University of Sydney, Sydney, Australia, as a China Scholarship Council Visiting Scholar. Since 2017, he has been an Assistant Professor with the Department of Electrical and Information Engineering, HIT Shenzhen Graduate School. He has authored or coauthored over 60 peer-reviewed articles. He holds over 30 invention patents in his research fields. His current interests include error control codes, space information networks, random multiple access, and machine-to-machine communications.