An Adaptive Distributed Compressed Video Sensing Algorithm Based on Normalized Bhattacharyya Coefficient for Coal Mine Monitoring Video

YONGGANG XU1,2, YONGZHI XUE1, GANG HUA1, AND JIANWEI CHENG3
1School of Information and Control Engineering, China University of Mining and Technology, Xuzhou 221116, China
2Xuzhou Engineering Research Center of Intelligent Industry Safety and Emergency Collaboration, Xuzhou 221116, China
3School of Safety Engineering, China University of Mining and Technology, Xuzhou 221116, China
Corresponding author: Yongzhi Xue (ts18060053A31@cumt.edu.cn)

ABSTRACT Compared with traditional video surveillance systems, wireless video sensor systems are more suitable for emergency application scenarios, such as underground coal mine disaster rescue, due to their low power consumption and rapid deployment. Considering the limited computing power and transmission bandwidth of video sensor nodes, we propose an adaptive compression and hybrid multiple hypothesis based residual reconstruction algorithm based on normalized Bhattacharyya coefficient (NBCAC-MHRR) to solve the high efficiency video coding (HEVC) problem in underground coal mines. First, a low-complexity adaptive sampling rate allocation method is performed on the encoding side. Second, by integrating the background subtraction idea, we combine the high-quality reconstruction of the foreground with the multi-hypothesis residual reconstruction of the background to improve the overall reconstruction effect of the video sequence. Simulation results show that the proposed algorithm can achieve higher reconstruction quality and efficiency than previous methods, especially in underground coal mine application scenarios.

INDEX TERMS Distributed compressed video sensing, normalized Bhattacharyya coefficient, adaptive sampling, multi-hypothesis.

I. INTRODUCTION

Because of the limited transmission bandwidth of wireless communication in underground coal mines, how to reconstruct high-quality video sequences from a limited number of samples has become the main research direction. Traditional video compression algorithms, such as MPEG-2/4 and h.26x, require a lot of additional calculations on the encoding side, which bring a huge energy burden to the wireless video sensor node. In addition to the design of micro-power consumption multiple-input and multiple-output devices [1], transferring the amount of calculation from the encoding end to the decoding end is another feasible solution. Distributed compressed video sensing (DCVS) [2] combines the characteristics of “independent coding and joint decoding” of distributed video coding [3] and the low coding complexity of compressed sensing (CS) [4], which makes low coding complexity and high reconstruction quality possible. In DCVS, an important research direction is to obtain more accurate side information (SI) from the reference frame. Tramel and Fowler [5] obtained a higher reconstruction quality by using iterative pixel-domain motion estimation (ME)/motion compensation (MC) and Tikhonov regularization. Chen et al. [6] proposed a hybrid hypothesis prediction method combining single-hypothesis and multiple-hypotheses (MH). Ou et al. [7] developed a two-stage MH reconstruction scheme, this scheme successively performs ME/MC iteration and residual reconstruction in measurement-domain and pixel-domain, improving the reconstruction quality at the expense of a certain reconstruction efficiency. Li et al. [8] presented a multihypothesis-based residual reconstruction scheme (MR-MHRR) that...
generates hypothesis blocks in the residual domain and calculates the linear prediction weights in measurement-domain. Chen et al. [9] presented an iterative reweighted Tikhonov-regularized scheme for MH prediction reconstruction and a Bhattacharyya coefficient-based stopping criterion to avoid over-iteration.

References [5]–[9] mainly use the pixel domain/measurement-domain ME/MC method to generate the hypothesis as accurately as possible, thereby improving the performance of residual reconstruction.

Besides, the HEVC standard introduces more compression and coding techniques, such as variable-size block transformation, multi-directional intra prediction, multi-frame MC, and so on. Li et al. [10] proposed a new ME method called High Efficiency Video Coding Motion Estimation (HEVC-ME), which uses coding units of different sizes, the calculated residuals are subjected to Hardman transform first and then sum the absolute values as the rate-distortion function to generate more accurate SI. Zhang et al. [11] used the 3D block matching approximate message passing [12] to reconstruct key frames, and then integrated the motion vector prediction method to propose a new HEVC-ME algorithm, which combines the concept of segmentation with ME to obtain more accurate SI, finally the l1 − l1 minimization model is employed for joint high-quality reconstruction of non-key frames.

In general application scenarios, reconstruction quality is widely concerned as an important performance indicator. Tian et al. [13] presented a dictionary learning based reconstruction method to achieve the joint optimization of sparse representation and signal reconstruction. Oishi and Kuroki [14] used l1-norm error instead of l2-norm to increase the robustness against outliers, and then applied an alternating direction method of multipliers to minimize the cost function. However, the actual situation has slight differences when applied to underground coal mines. Surveillance videos in the underground environment usually have the characteristics of poor illumination and low resolution, and their frames are closer to the gray image. Jiang et al. [15] indicated that the underground coal mine monitoring requires rapid response to various anomalies in a complex environment.

The idea of distinguishing the foreground and background in target detection algorithms [16] is also applicable to video compression. However, most background difference methods will increase the computational burden and energy consumption of video sensor nodes, which is contrary to the DCVS theoretical framework.

In terms of adaptive sampling, Stanković et al. [17] proposed a sparseness prediction method based on previous reference frames which are sampled conventionally. The image blocks are simply divided into sparse and non-sparse blocks, and the sparse blocks are compressively sampled and the non-sparse blocks are fully sampled. Wu and Zhu [18] proposed a dynamic measurement rate allocation method that can adaptively adjust measurement rates by estimating the sparsity of each block via feedback information.

Vijayanagar et al. [19] introduced an adaptive measurement rate allocation algorithm using rate-distortion optimization and utilized a simple block classification and frame-differencing module to reduce encoding complexity. The methods proposed in [17]–[19] have limitations such as high coding complexity or relying on feedback information, which are usually not suitable for practical application scenarios.

With insight into the adaptive sampling strategy which can reduce background redundancy by well exploiting temporal correlation, an adaptive compression and hybrid multiple hypothesis residual reconstruction algorithm based on normalized Bhattacharyya coefficient (NBCAC-MHRR) is presented. There are two contributions in the NBCAC-MHRR: 1) By using the normalized Bhattacharyya coefficient to shift the center of gravity of the video compression from frame to block, which provides convenience for adaptive compression and reconstruction. 2) A hybrid reconstruction framework that combines traditional reconstruction methods with multihypothesis-based residual reconstruction is used for the subsequent reconstruction, which improves the overall reconstruction performance.

II. RELATED WORK

In this chapter, we briefly introduce the basic MH prediction and residual reconstruction framework and describe the research on adaptive sampling.

A. MH PREDICTION AND RESIDUAL RECONSTRUCTION

In DCVS, the block-based compressed sensing method is used for sampling video frames, the initial frames are divided into b × b blocks, then each block is independently sampled using measurement matrix Φb, \( \hat{x}_{i,j}^k = \Phi_b x_{i,j}^k \), where \( x_{i,j}^k \) denotes a rasterized vector of the block in row i, column j and frame k, \( \Phi_b \) is an \( m \times b^2 \) orthonormal Gaussian random measurement matrix such that the block sampling rate is \( S = m/b^2 \).

How to get a hypothesis that is as close as possible to the original image block is a key issue in the MH prediction scheme. The reconstructed frames are used as a priori information to participate in the subsequent reconstruction of non-key frames in the form of SI, and the hypothesis set \( H_{i,j}^k \) will be generated from the obtained SI, then the MH prediction is conducted by:

\[
\hat{x}_{i,j}^k = H_{i,j}^k \hat{w}_{i,j}^k
\]

where \( \hat{x}_{i,j}^k \) is an approximate estimate of \( x_{i,j}^k \), \( H_{i,j}^k \) is a \( b^2 \times P \) matrix whose columns \( h_1, h_2, \ldots, h_P \) are the rasterizations of hypothesis blocks generated from the SI and \( \hat{w}_{i,j}^k \) is a linear prediction weight vector obtained by the following optimization method:

\[
\hat{w}_{i,j}^k = \arg \min_w \| x_{i,j}^k - H_{i,j}^k w \|_2^2
\]
in (2) is as follows:

$$\hat{w}_{i,j}^k = \arg\min_w \left\{ \Phi_B \hat{x}_{i,j}^k - \Phi_B H_{i,j}^k w \right\}^2_2$$

$$= \arg\min_w \left\| \hat{x}_{i,j}^k - \Phi_B H_{i,j}^k w \right\|^2_2$$

(3)

For the ill-posedness of the solution in (3), the most common method is to use Tikhonov regularization:

$$\hat{w}_{i,j}^k = \arg\min_w \left\{ \hat{x}_{i,j}^k - \Phi_B H_{i,j}^k w \right\}^2_2 + \lambda \| w \|^2_2$$

(4)

where $\lambda$ is a factor and $\Gamma$ is the Tikhonov matrix defined as:

$$\Gamma = \text{diag}(y_{i,j}^k, \ldots, y_{i,j}^k, \ldots, y_{i,j}^k)$$

(5)

In fact, for each block of the video frame, $\hat{w}_{i,j}^k$ can be directly calculated by the following formula:

$$\hat{w}_{i,j}^k = \left( \Phi_B H_{i,j}^k \Gamma^{-1} (\Phi_B H_{i,j}^k)^T + \lambda^2 \Gamma^{-1} \right)^{-1} (\Phi_B H_{i,j}^k)^T y_{i,j}^k$$

(6)

Since $\hat{x}_{i,j}^k$ is an approximate estimate of $x_{i,j}$ satisfying $\hat{x}_{i,j}^k \approx x_{i,j}$, then the residual $r_{i,j}^k$ is defined as: $r_{i,j}^k = x_{i,j}^k - \hat{x}_{i,j}^k$ as an alternative to directly reconstructing $y_{i,j}$, the final result can be obtained from (7) and (8):

$$d_{i,j}^k = \Phi_B r_{i,j}^k = \Phi_B x_{i,j}^k - \Phi_B \tilde{x}_{i,j}^k = y_{i,j}^k - \Phi_B \tilde{x}_{i,j}^k$$

(7)

$$\tilde{x}_{i,j}^k = \hat{x}_{i,j}^k + \text{reconstruct}(d_{i,j}^k)$$

(8)

where $\text{reconstruct}(\cdot)$ represents a certain CS image reconstruction algorithm.

In recent years, scholars have conducted a lot of research on reconstruction algorithms for compressed sensing. Wen et al. [20] proved sufficient conditions for the minimization of $\|x\|_1 - \|x\|_2$ based on mutual coherence. Li et al. [21] presented a sufficient condition for the exact support recovery of $K$-row sparse matrices in the noisy case. The smoothed projected Landweber reconstruction algorithm is widely used because of its simplicity and efficiency. On this basis, Fowler et al. [22] used a method called SPL-DDWT that applied dual-tree DWT (DDWT) [23] as the sparsity transform $\Psi$ with bivariate shrinkage [24] to enforce sparsity. Compared with the method of directly using reconstruction algorithms, the residual reconstruction can usually achieve higher accuracy.

### B. ADAPTIVE SAMPLING SCHEME

Considering the limited storage and computing power at the encoding end, CS-based video encoding methods usually directly compress the captured video data without temporarily storing the original data. Therefore, it is difficult to allocate the sampling rate without the support of the original data.

The method proposed in [17] performs a full sampling of key frames and then applies DCT on each block. The sparse blocks are selected in the following manner. Let $C$ denotes a small positive constant, $T$ represents the average number of non-significant DCT coefficients over all blocks, and $D_{ij}$ represents the number of DCT coefficients in the block $x_{ij}$.

If $T < \text{abs}(D_{ij}) < C$, the block $x_{ij}$ is selected as a reference for compressive sampling. This method of using DCT coefficients to determine the sparsity of image blocks requires full sampling of references and additional transformation operations, and the sampling rate allocation method that determines only partial sampling or full sampling is relatively rough.

In response to the problems in [17], an adaptive sampling rate allocation method based on feedback information was proposed in [18]. This method performs sparsity estimation at the decoding end and sends it back to the encoding end in the form of feedback information. The detailed description is shown in Fig. 1.

Although this method of using feedback information can reduce the calculation and storage burden of the encoding end, it does not consider the complex transmission structure and is not suitable for non-point-to-point application scenarios, such as wireless video sensor networks.

Reference [19] proposed a measurement rate computation method based on rate-distortion optimization. The method first performs DCT transformation on the original signal, then calculates the difference between consecutive DC measurements, and binarizes the residual using a $K$th-order Exp-Golomb code. Let the size of the compressed DC measurements be denoted by $R_{\text{DC}}$, then the maximum size for encoding the AC measurements can be calculated by:

$$R_{\text{AC}} = R_{\text{total}} - (R_{\text{DC}} + R_{\text{map}})$$

where $R_{\text{total}}$ represents the bit-budget available for compressing, $R_{\text{map}}$ is the size of the classification map.

Assume that the set of all CS blocks is defined as $X = \{X_1, X_2, \ldots, X_P\}$, and $L = \{L_1, L_2, \ldots, L_P\}$ represents the set of all measurement rates, then the problem can be transformed into the following constrained optimization problem:

$$L^* = \min \limits_L D(X, L)$$

subject to $R(X, L) \leq R_{\text{AC}}$
calculating the sum of the mean square error between the original CS blocks and the reconstructed blocks.

Equation (10) can be rewritten as an unconstrained Lagrangian formulation as:

$$L^* = \min_{L} \left( \sum_{p=1}^{P} (D(X_p, L_p) + \lambda R(X_p, L_p)) \right)$$

(11)

where $\lambda \geq 0$ is the Lagrangian multiplier, $D(X_p, L_p)$ and $R(X_p, L_p)$ represent the distortion and total sample number of the block $X_p$, respectively.

This rate-distortion optimization method provides an idea for adaptive sampling. However, it is noted that the calculation of the distortion $D(X, L)$ requires the reconstructed information on the decoding end, which is difficult to achieve for practical application scenarios such as coal mine video surveillance systems.

### III. ADAPTIVE COMPRESSION AND MULTIHYPOTHESIS-BASED RESIDUAL RECONSTRUCTION ALGORITHM

Although many researchers have proposed various methods to optimize the accuracy of SI, there are still differences between MH predictions and actual frames. From a block perspective, these differences will decrease between background blocks and increase between foreground blocks. Especially, it will be reduced to zero for completely still background blocks.

The NBCAC-MHRR algorithm is an adaptive sampling method that uses normalized Bhattacharyya coefficient (NBC) to simply reduce background redundancy and make full use of limited sampling. It uses an additional codebook to associate the encoding and decoding end, and then uses a hybrid scheme for adaptive reconstruction. By integrating the advantages of adaptive sampling and multihypothesis-based residual reconstruction, the proposed method solves the problem of inaccurate SI to a certain extent.

#### A. NORMALIZED BHATTACHARYYA COEFFICIENT

The Bhattacharyya coefficient is widely used in similar image recognition because of its characteristic of intuitively reflecting the degree of similarity between images. In the proposed NBCAC-MHRR, the NBC is used to distinguish between foreground and background.

For each block of the current frame, the calculation method of the NBC is as follows:

**Step1:** Initialize a histogram array of size 256 for the current frame and the preceding frame respectively, and then perform frequency distribution statistics based on pixel values.

**Step2:** Use the obtained histogram array for calculation according to the following formula:

$$N_{i,j}^k = 1 - \frac{\sum_{h=1}^{256} I_{i,j}^k(h) \times I_{i,j}^{k-1}(h)}{\sqrt{\sum_{h=1}^{256} I_{i,j}^k(h) \times \sum_{h=1}^{256} I_{i,j}^{k-1}(h)}}$$

(13)

where $N_{i,j}^k$ represents the NBC between $x_{i,j}^k$ and $x_{i,j}^{k-1}$, $I_{i,j}(h)$ is the frequency of the pixel-value $h$ in the given block.

The closer the NBC value to zero, the smaller the inter-frame variation of the given block. According to this characteristic, an adaptive compression method will be described in detail later.

#### B. ADVANTAGES OF THE PROPOSED METHOD USING NBC

For grayscale images, the method of using NBC for sampling rate allocation has the following advantages compared to the methods proposed in [17]–[19]:

1) **STORAGE BURDEN**

For each block of the preceding frame, a histogram array of size 256 needs to be stored. Compared with the method of storing entire key frames, the method using NBC can save about three-fourths of the storage space.

2) **COMPUTATIONAL BURDEN**

Compared with the complicated transformation, quantization, and coding operations, the proposed method using NBC is relatively simple.

3) **DATA TRANSMISSION**

A fixed-size codebook used to record the number of adaptive samples needs to be additionally transmitted. The size of the codebook depends on the number of blocks, and usually does not bring additional transmission burden. Besides, the one-way transmission method is more suitable for practical application scenarios than the two-way transmission structure that requires feedback information from the decoder.

#### C. ADAPTIVE COMPRESSION AND RECONSTRUCTION BASED ON NBC

In the proposed scheme, a major constraint is to ensure that the actual total number of samples is consistent with the target, and the number of samples per block does not exceed the block sampling limit. Therefore, the method for sampling rate allocation can be described as the following constraint optimization problem:

$$L_{i,j} = \max_{s \in \{1, \cdots, s_{i,j}, \cdots, s_{p,Q}\}} \frac{\sum_{h=1}^{256} I_{i,j}^k(h)}{\sum_{h=1}^{256} I_{i,j}^{k-1}(h)}$$

subject to $L_{i,j} = \cdots = L_{p,Q} = L$

$$L_{i,j} \leq b^2$$

(14)

where $L_{i,j}$ represents the number of samples in block $x_{i,j}$, $\mathcal{L}$ represents the total number of target samples, $f$ is a function between the numbers of samples and the ratio of NBC, $V$ is the maximum block sampling rate used to prevent oversampling, $b$ is the block size, and $s$ is the NBC proportion whose calculation is as follows:

$$s_{i,j} = N_{i,j} / \sum_{m=1}^{r_{\text{max}}} \sum_{n=1}^{c_{\text{max}}} N_{m,n}$$

(15)

where $r_{\text{max}}$ and $c_{\text{max}}$ represent the number of rows and columns respectively.
The specific adaptive compression scheme is as follows: 

**Step 1:** Calculate the NBC between \( x_{ij}^k \) and \( x_{ij}^{k-1} \). 

**Step 2:** Initialize the number of samples for each block according to the NBC proportion by the following formula: 

\[
L_{ij}^k = L \times x_{ij}^k 
\]  

(16) 

**Step 3:** For each block \( x_{ij}^k \), if \( L_{ij}^k > V \times b^2 \), then: 

\[
L_{ij}^k = V \times b^2 
\]  

(17) 

\[
W = W + (L_{ij}^k - V \times b^2) 
\]  

(18) 

where \( W \) is the accumulation of the over-sampling number whose initial value is zero. 

**Step 4:** For all blocks whose \( L_{ij}^k < V \times b^2 \), distribute the over-sampling number \( W \) according to the NBC proportion. Repeat steps 3 to 4 until \( W \) becomes zero. 

For a completely static background block, the number of adaptive samples will be zero, and the reconstructed image block at the corresponding position of the previous frame is used as the reconstruction result. Also, since blocks with large inter-frame variations are assigned a higher sampling rate, a special reconstruction framework is used for the reconstruction of blocks whose adaptive sampling number reaches a preset upper limit. First, we use an orthogonal matching pursuit [25] algorithm with a discrete wavelet transform as the sparse basis for the initial reconstruction of the image block. Then, a total variance denoising algorithm [26] is used to improve the visual effect of the reconstructed blocks. The multi-hypothesis residual reconstruction method for blocks with small inter-frame motion will be described in the next section. 

Through the above steps, we allocate more samples to the blocks with a larger inter-frame motion to obtain better reconstruction quality. Compared with the traditional uniform sampling method, the proposed method is particularly suitable for underground coal mine monitoring application scenarios due to its uneven sampling characteristics and is more conducive to reconstructing relatively clear images with lower sampling rates. 

**D. MULTIHYPOTHESIS-BASED RESIDUAL RECONSTRUCTION** 

The method based on multiple reference frames trades a lot of extra calculations for better reconstruction quality, which is not suitable for underground coal mine monitoring scenarios. A compromise method is to use the reconstructed frames that contain more SI as the reference. In this section, we use a measurement domain multi-hypothesis residual reconstruction scheme to reconstruct background blocks. The detailed process is as follows: 

**Step 1:** Extend the reference frame. The block to be reconstructed at the edge of the image will result in the inability to generate a complete search window, a feasible method is to extend the reference frame. For all blocks at the edge position, a mirror projection step is conducted to obtain an extended reference. 

**Step 2:** In the extended reference, the search window is generated with the same position as the current block. The size of the search window is generally related to the degree of motion between frames. 

**Step 3:** In the search window, all blocks are obtained by rasterizing traversal and then vectorized to generate the search matrix. Afterward, select the best matching block by using the following method: 

**Algorithm 1 NBCAC Algorithm** 

**Input:** Current frame \( x^k \), the histogram arrays \( f^k \) and \( f^{k-1} \), the number of samples available \( L \) 

**Output:** Adaptive sampling number matrix \( L^k \) 

1: **Step 1:** Calculate the normalized Bhattacharyya coefficient for each block. 

2: for \( i = 1 \) to \( r_{\text{max}} \), \( j = 1 \) to \( c_{\text{max}} \) do 

3: \( N_{ij} = 1 - \frac{\sum_{h=1}^{I^k_{ij}(h)} \sqrt{\sum_{l=1}^{L^k_{ij}(h)}}}{\sum_{h=1}^{I^k_{ij}(h)} \sqrt{\sum_{l=1}^{L^k_{ij}(h)}}} \) 

4: **Step 2:** Assign the initial number of samples to each block. 

5: for \( i = 1 \) to \( r_{\text{max}} \), \( j = 1 \) to \( c_{\text{max}} \) do 

6: \( S_{ij} = N_{ij} \times \sum_{m=1}^{r_{\text{max}}} \sum_{n=1}^{c_{\text{max}}} N_{m,n} \) 

7: \( L_{ij}^k = L \times S_{ij} \) 

8: **Step 3:** Distribute the number of samples exceeding the preset upper limit to the blocks that have not reached the upper limit. 

9: for \( i = 1 \) to \( r_{\text{max}} \), \( j = 1 \) to \( c_{\text{max}} \) do 

10: if \( L_{ij}^k > V \times b^2 \) then 

11: \( L_{ij}^k = V \times b^2 \) 

12: \( W = W + (L_{ij}^k - V \times b^2) \) 

13: \( N_{ij} = 0 \) 

14: if \( L_{ij}^k = V \times b^2 \) then 

15: \( N_{ij} = 0 \) 

16: for \( i = 1 \) to \( r_{\text{max}} \), \( j = 1 \) to \( c_{\text{max}} \) do 

17: \( S_{ij} = N_{ij} \times \sum_{m=1}^{r_{\text{max}}} \sum_{n=1}^{c_{\text{max}}} N_{m,n} \) 

18: \( L_{ij}^k = L_{ij}^k + W \times S_{ij} \) 

19: **Step 4:** Repeat Step 3 until \( W = 0 \) or the number of iterations reaches the upper limit. 

20: return \( L^k \) .

**match_block = min \| \Phi_{ij}^k - s_k \|_k** 

(19) 

where \( s_k \) represents the \( k \)th column of the search matrix. 

**Step 4:** Generate a hypothesis set generation window centered on the position corresponding to \( \text{match_block} \). Then the hypothesis set \( H_1 \) is obtained by rasterizing traversal. 

**Step 5:** Obtain the hypothesis set \( H_2 \) from the other reference. Then combine \( H_1 \) and \( H_2 \) to get the final hypothesis set \( H_{ij} \). 

Then, the measurement-domain linear prediction weight is calculated by (6) and the MH prediction is obtained by (1).
FIGURE 2. Adaptive compression and hybrid multihypothesis-based residual reconstruction scheme.

FIGURE 3. Reference frame extension. M and N represent the height and width of the original video frame, respectively.

FIGURE 4. Rasterized traversal method.

The SPL-DDWT algorithm is used for the final residual reconstruction.

IV. EXPERIMENTAL RESULTS AND ANALYSIS

In this section, extensive experimental evaluations based on peak signal to noise ratio (PSNR) and reconstruction time are conducted to demonstrate the superiorities of the proposed NBCAC-MHRR.

A. CONFIGURATION AND SETTINGS

Experimental environment: We use an Intel(R) Core(TM) 4 CPU (~3.6 GHz), with 8GB of memory. The programming software is Matlab-R2019b, and the operating system is Microsoft Windows 10. The formula for calculating the PSNR is as follows:

\[
PSNR = 10 \times \log_{10} \frac{f_{\text{max}}^2}{MSE}
\]

where \(f_{\text{max}}\) represents the maximum gray value and the default value is 255.

One set of standard test video images and three sets of underground coal mine surveillance video images are used to verify the actual performance of the algorithm. The size of all video frames used for testing is 288 \times 352, and the block size is 32 \times 32, DDWT decomposition level is 3, the maximum sampling rate of all blocks is set to 0.8. All schemes use the Gaussian random matrix as the measurement matrix and SPL-DDWT [22] as the reconstruction algorithm.

Besides, we use a 99-frame underground coal mine surveillance video to verify the performance of the proposed algorithm when applied to video sequences. The size of the GOP is set to 9, where the first frame is regarded as a key frame and the rest are CS frames. The sampling rate of key frames is set to 0.6, and BCS-SPL [27] is used as the compression and reconstruction scheme.

B. PERFORMANCE OF NBCAC-MHRR

In this subsection, we verify the superiority of the proposed NBCAC-MHRR through multiple sets of comparative experiments. We compare it with some other DCVS schemes: 1) MH-BCS-SPL [5], which uses intra-frame multiple hypotheses to iteratively optimize the initial reconstruction results. 2) MR-MHRR [8], which uses a multi-hypothesis residual reconstruction framework based on multiple reference frames. 3) HEVC-ME [10], which generates more accurate SI by performing motion estimation with coding units of different sizes. Fig. 5 shows the original images used for experiment and analysis. We select the test image sets from video sequences “Akiyo” and “Walk”, which include...
an eye-opening motion and a leg-raising motion, respectively. Fig. 6 shows the reconstructed images of different algorithms at the sampling rate of 0.1. By comparison, it can be found that the overall reconstruction effects of the four schemes are relatively close, but there are some differences in the areas containing motion details. In these areas, MR-MHRR and HEVC-ME lose part of the detailed information, resulting in the reconstructed image being closer to the reference image rather than the original image. The proposed algorithm retains relatively complete motion information due to its non-uniform sampling characteristics. For a completely still background area, the adaptive samples number is zero, and for regions with large inter-frame motion, the adaptive sampling number will be much larger than the fixed sampling number to solve the uneven performance of SI. In this way, the problem of poor reconstruction quality of the image detail due to the limitation of the total number of samples has been solved to a certain extent, while the reconstruction quality of the background part will not be significantly reduced due to the relatively accurate SI obtained from the reference frame.

To further illustrate the superiority of the adaptive sampling scheme and enhance the contrast effect, we conduct the remaining experiments and analysis at the sampling rate of 0.2. Fig. 7 shows the distribution of the number of samples at the sampling rate of 0.2. Compared with fixed sampling, the adaptive sampling method can make full use of the limited number of samples, thereby improving the reconstruction quality of the moving area.

Table 1 shows the PSNR and reconstruction time of several algorithms at different sampling rates. From the perspective of reconstruction quality, MH-BCS-SPL only uses intra-frame multi-hypothesis prediction without considering inter-frame correlation, resulting in poor performance at low sampling rates. From the perspective of reconstruction efficiency, the coding unit segmentation operation used by HEVC-ME will consume extra time, resulting in lower overall reconstruction efficiency. On the other hand, the amount of calculation required for inter-frame MH prediction will increase as the sampling rate increases. Therefore, the reconstruction time required by the proposed method and
MR-MHRR will increase as the sampling rate increases. It is worth noting that the proposed NBCAC-MHRR has obvious advantages when the sampling rate is low. As shown in Fig. 8, we compare the PSNR of the “Walk” sequence reconstructed by the four algorithms at the sampling rate of 0.2. MH-BCS-SPL performs the same intra-frame multi-hypothesis operation on all video frames, so the reconstruction quality is relatively stable. However, such a method that does not consider the inter-frame correlation is greatly limited by the sampling rate, resulting in lower overall
reconstruction quality. The other three algorithms all use the scheme of high-quality reconstruction of key frames and multi-hypothesis residual reconstruction of non-key frames, this kind of scheme will cause the reconstruction quality of non-key frames in the same GOP to show different degrees of fading. HEVC-ME optimizes motion estimation by calculating the rate-distortion cost of the current frame and the reference frame. This method is largely limited by the accuracy of the reference frame, which causes its reconstruction quality to lag behind the multi-reference-based schemes. The proposed NBCAC-MHRR has a significant improvement in the overall reconstruction quality due to the adaptive sampling scheme. The relatively slow fading curve shows that this scheme can effectively reduce the influence caused by error transmission.

| Sequences | Methods      | 0.1      | 0.2      | 0.3      | 0.4      | 0.5      |
|-----------|--------------|----------|----------|----------|----------|----------|
|           | PSNR  | TIME     | PSNR  | TIME     | PSNR  | TIME     | PSNR  | TIME     | PSNR  | TIME     |
| Akiyo     |            |          |        |          |        |          |        |          |        |          |
| MH-BCS-SPL| 31.45  | 19.28    | 35.09  | 14.39    | 37.74  | 8.27     | 39.96  | 9.23     | 42.12  | 9.12     |
| MR-MHRR   | 37.85  | 8.24     | 38.38  | 9.76     | 39.26  | 11.13    | 39.67  | 12.00    | 41.18  | 14.71    |
| HEVC-ME   | 38.74  | 17.79    | 40.89  | 17.45    | 41.80  | 17.54    | 42.17  | 17.47    | 42.30  | 17.40    |
| NBCAC-MHRR| 39.57  | 4.96     | 40.91  | 7.08     | 42.79  | 17.73    | 44.34  | 24.91    | 44.65  | 27.41    |
| Walk      |            |          |        |          |        |          |        |          |        |          |
| MH-BCS-SPL| 34.50  | 12.99    | 37.21  | 17.21    | 38.53  | 12.34    | 40.61  | 13.40    | 42.37  | 8.88     |
| MR-MHRR   | 38.12  | 8.05     | 39.00  | 9.73     | 39.93  | 10.53    | 40.73  | 12.17    | 42.06  | 15.70    |
| HEVC-ME   | 37.78  | 17.76    | 40.06  | 17.35    | 41.27  | 17.53    | 41.57  | 17.39    | 41.72  | 17.47    |
| NBCAC-MHRR| 41.28  | 4.29     | 42.97  | 9.91     | 43.51  | 14.91    | 43.92  | 20.08    | 44.09  | 22.57    |
| Miners    |            |          |        |          |        |          |        |          |        |          |
| MH-BCS-SPL| 30.27  | 14.46    | 34.00  | 13.12    | 36.51  | 12.26    | 38.25  | 9.40     | 39.76  | 9.77     |
| MR-MHRR   | 37.16  | 8.08     | 37.95  | 9.91     | 38.58  | 10.88    | 39.45  | 11.90    | 40.19  | 16.46    |
| HEVC-ME   | 35.48  | 17.32    | 38.11  | 17.30    | 39.49  | 17.37    | 39.92  | 17.54    | 40.35  | 17.45    |
| NBCAC-MHRR| 38.20  | 5.17     | 39.54  | 7.53     | 41.44  | 16.55    | 42.49  | 23.29    | 42.85  | 27.26    |
| Digger    |            |          |        |          |        |          |        |          |        |          |
| MH-BCS-SPL| 30.98  | 17.89    | 34.27  | 15.60    | 35.93  | 19.26    | 37.34  | 9.04     | 38.73  | 9.03     |
| MR-MHRR   | 35.27  | 8.06     | 35.97  | 9.77     | 36.65  | 10.60    | 37.94  | 11.86    | 39.11  | 15.28    |
| HEVC-ME   | 34.46  | 17.53    | 36.71  | 17.39    | 37.84  | 17.46    | 38.22  | 17.48    | 38.25  | 17.54    |
| NBCAC-MHRR| 36.50  | 5.51     | 37.91  | 7.13     | 39.55  | 13.92    | 40.77  | 23.43    | 41.44  | 29.61    |

**FIGURE 9.** Reconstruction time of the “Walk” sequence at the sampling rate of 0.2.
On the other hand, the adaptive sampling scheme allocates more samples to the foreground area, which will inevitably lead to the degradation of the reconstruction quality of the background area. In most cases, this decline will not affect the visual quality of the reconstructed image, but it will still cause a decrease in PSNR. As we can see in Fig. 8, there are a few points in the PSNR curve of MR-MHRR that are better than the proposed NBCAC-MHRR, but the generally attenuated curve shows that the scheme cannot maintain this advantage. Fig. 9 shows the reconstruction time curve of different algorithms. Compared with the other three algorithms, the proposed NBCAC-MHRR has obvious advantages.

In general, our simulation experiments show that the proposed NBCAC-MHRR can bring a certain improvement in reconstruction quality and reconstruction efficiency. This non-uniform sampling scheme is particularly suitable for monitoring application scenarios.

V. CONCLUSION

Herein, an adaptive sampling rate allocation and hybrid multi-hypothesis-based residual reconstruction algorithm based on normalized Bhattacharyya coefficient is proposed. The proposed algorithm uses NBC for adaptive sampling at the encoding side and then combines the pixel-domain ME technology with the measurement-domain linear prediction weight calculation to perform the residual reconstruction. In particular, a method combining OMP reconstruction and TV denoising is used to reconstruct blocks whose number of samples reaches the preset upper limit. The simulation results show that the proposed algorithm improves the reconstruction performance of video compression in complex surveillance scenarios, and achieves relatively high-efficiency and high-quality reconstruction at low sampling rates. The algorithm is of great significance for the realization of high-performance compression and reconstruction of surveillance video and the establishment of wireless video surveillance systems in underground coal mines.

REFERENCES

[1] J. Wen, L. Li, X. Tang, and W. H. Mow, “An efficient optimal algorithm for the successive minima problem,” IEEE Trans. Commun., vol. 67, no. 2, pp. 1424–1436, Feb. 2019, doi: 10.1109/TCOMM.2018.2874746.

[2] 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, doi: 10.1109/ICASSP.2009.4959797.

[3] J. Prades-Nebot, Y. Ma, and T. Huang, “Distributed video coding using compressive sampling,” in Proc. Picture Coding Symp., May 2009, pp. 1–4, doi: 10.1109/PCS.2009.5167431.

[4] D. L. Donoho, “Compressed sensing,” IEEE Trans. Inf. Theory, vol. 52, no. 4, pp. 1289–1306, Apr. 2006, doi: 10.1109/TIT.2006.871582.

[5] E. W. Tramel and J. E. Fowler, “Video compressed sensing with multi-hypothesis,” in Proc. Data Compress. Conf., Mar. 2011, pp. 193–202, doi: 10.1007/978-3-642-19373-2_24.

[6] J. Chen, Y. Chen, D. Qin, and Y. Kuo, “An elastic net-based hybrid hypothesis method for compressed video sensing,” Multimedia Tools Appl., vol. 74, no. 6, pp. 2085–2108, Mar. 2015, doi: 10.1007/s11454-013-1743-y.

[7] W.-F. Ou, C.-L. Yang, W.-H. Li, and L.-H. Ma, “A two-stage multi-hypothesis reconstruction scheme in compressed video sensing,” in Proc. IEEE Int. Conf. Image Process. (ICIP), Sep. 2016, pp. 2494–2498, doi: 10.1109/ICIP.2016.7532008.

[8] W.-H. Li, C.-L. Yang, and L.-H. Ma, “A multi-hypothesis-based residual reconstruction scheme for compressed video sensing,” in Proc. IEEE Int. Conf. Image Process. (ICIP), Sep. 2017, pp. 2766–2770, doi: 10.1109/ICIP.2017.8296786.

[9] C. Chen, C. Zhou, P. Liu, and D. Zhang, “Iterative reweighted tvonov-regularized multi-hypothesis prediction scheme for distributed compressive video sensing,” IEEE Trans. Circuits Syst. Video Technol., vol. 20, no. 1, pp. 146–158, Feb. 2010, doi: 10.1109/TCSVT.2018.2886310.

[10] Z. Li, S. Wu, M. Ma, J. Jiang, 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, doi: 10.1109/ICCCChina.2018.8641167.

[11] R. Zhang, S. Wu, Y. Wang, and J. Jiao, “High-performance distributed compressive video sensing: Jointly exploiting the HEVC motion estimation and the reconstruction,” IEEE Access, vol. 8, pp. 31306–31316, 2020, doi: 10.1109/ACCESS.2020.2973392.

[12] 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, doi: 10.1109/ICIP.2015.7351377.

[13] F. Tian, J. Guo, B. Song, H. Liu, and H. Qin, “Distributed compressed video sensing with joint optimization of dictionary learning and 1-analysis based reconstruction,” IEICE Trans. Inf. Syst., vol. 10,E99.D, no. 4, pp. 1202–1211, 2016, doi: 10.1587/transinf.2015EDP7373.

[14] T. Oishi and Y. Kuroki, “Dictionary learning on 1-norm fidelity for non-key frames in distributed compressed video sensing,” in Proc. Int. Symp. Intell. Signal Process. Commun. Syst. (ISPACS), Dec. 2019, pp. 1–2, doi: 10.1109/ISPACS48206.2019.8996278.

[15] J. Dai-Hong, D. Lei, L. Dan, and Z. San-You, “Moving-object tracking algorithm based on PCA-SIFT and optimization for underground coal mines,” IEEE Access, vol. 7, pp. 35556–35563, 2019, doi: 10.1109/ACCESS.2019.2899362.

[16] N. S. Sakpal and M. Sabnis, “Adaptive background subtraction in images,” in Proc. Int. Conf. Adv. Commun. Comput. Technol. (ICACCT), Feb. 2018, pp. 439–444, doi: 10.1109/ICACCT.2018.8529323.

[17] V. Stankovi, L. Stankovi, and S. Cheng, “Compressive video sampling,” in Proc. Eur. Signal Process. Conf., 2008, pp. 2–6.

[18] H.-W. Chen, L.-W. Kang, and C.-S. Lu, “Dynamic measurement rate allocation for distributed compressive video sensing,” in Proc. Vis. Commun. Image Process., Jul. 2010, pp. 62–67, doi: 10.1117/12.863094.

[19] K. V. Y. Anag, Y. Liu, and J. Kim, “Adaptive measurement rate allocation for block-based compressed sensing of depth maps,” in Proc. IEEE Int. Conf. Image Process. (ICIP), Oct. 2014, pp. 1307–1311, doi: 10.1109/ICIP.2014.7025261.

[20] J. Wen, J. Weng, C. Tong, C. Ren, and Z. Zhou, “Sparse signal recovery with minimization of 1-Norm minus 2-Norm,” IEEE Trans. Veh. Technol., vol. 68, no. 7, pp. 6847–6854, Jul. 2019, doi: 10.1109/TVT.2019.2919612.

[21] H. Li, L. Wang, X. Zhan, and D. K. Jain, “On the fundamental limit of orthogonal matching pursuit for multiple measurement vector,” IEEE Access, vol. 7, pp. 48860–48866, 2019, doi: 10.1109/ACCESS.2019.2907684.

[22] J. E. Fowler, S. Mun, and E. W. Tramel, “Multiscale block compressed sensing with smooth projected Landweber reconstruction,” in Proc. Eur. Signal Process. Conf., 2011, pp. 564–568.

[23] N. Kingsbury, “Complex wavelets for shift invariant analysis and filtering of signals,” Appl. Comput. Harmon. Anal., vol. 10, no. 3, pp. 234–253, May 2001, doi: 10.1006/acha.2000.0343.

[24] L. Sendur and I. W. Selesnick, “Bivariate shrinkage functions for wavelet-based denoising exploiting interscale dependency,” IEEE Trans. Signal Process., vol. 50, no. 11, pp. 2744–2756, Nov. 2002, doi: 10.1109/TSP.2002.804091.

[25] J. A. Tropp and A. C. Gilbert, “Signal recovery from random measurements via orthogonal matching pursuit,” IEEE Trans. Inf. Theory, vol. 53, no. 12, pp. 4655–4666, Dec. 2007, doi: 10.1109/TIT.2007.909108.

[26] G. Gilboa, “Nonlinear band-pass filtering using the TV transform,” Eur. Signal Process. Conf., no. 3, pp. 1699–1700, 2014.

[27] S. Mun and J. E. Fowler, “Block compressed sensing of images using directional transforms,” in Proc. Data Compress. Conf., 2010, p. 547, doi: 10.1109/DCC.2010.90.
YONGGANG XU was born in Pingchang, Sichuan, China, in September 1977. He received the B.S., M.S., and Ph.D. degrees in electronic information from the School of Information and Control Engineering, China University of Mine and Technology, Xuzhou, Jiangsu, China, in 2001, 2006, and 2013, respectively. He is currently an Associate Professor. He joined the Department of Information and Control Engineering, China University of Mine and Technology, in 2001. He mainly researches in information fusion, the mine Internet of Things, coalmine communication, and big data processing. His research interests include coalmine signal processing and its applications in enterprises, the mine Internet of Things, compressed sensing, and image processing.

Dr. Xu won the NSFC CAREER Award in 2012, and the Jiangsu Province Science and Technology Progress Award in 2017.

YONGZHI XUE was born in Nanyang, Henan, China, in 1997. He received the B.S. degree in electronic information science and technology from the Xuzhou University of Technology, Xuzhou, Jiangsu, China, in 2018. He is currently a Graduate Student in the Department of Information and Control Engineering, China University of Mine and Technology, Xuzhou. His main research interests include video compression and reconstruction.

GANG HUA received the B.S. degree from Southeast University, Nanjing, and the M.S. and Ph.D. degrees from the China University of Mine and Technology, Xuzhou, China, in 2000 and 2004, respectively. He is currently a Professor with the Department of Information and Electronic Engineering, China University of Mine and Technology. His research interests include the control and supervision of mining safety, signal processing, compressed sensing, and network technology.

JIANWEI CHENG received the Ph.D. degree from West Virginia University, USA. He is currently an Associate Professor with the China University of Mining and Technology and an Australia “Endeavour” Research Fellow with Curtin University. He is teaching and researching in the field of underground mining safety.