Motion-adaptive adjacent-reference skipping for distributed video compressive sensing with general decoders*

YUAN Wenye and LIU Hao**
College of Information Science and Technology, Donghua University, Shanghai 201620, China

(Received 23 April 2022; Revised 8 September 2022)

On an internet of video things (IoVT), an encoder needs to collect a large number of signal samples to improve the reconstruction quality. It is challenging to some occasions where the resources of an encoder are extremely limited. The distributed video compressive sensing (DVCS) can save a lot of resources for the encoder. For the skip-block coding at such an encoder, this paper proposes a motion-adaptive adjacent-reference skipping (MAS) algorithm for DVCS with general decoders. The proposed algorithm makes full use of the spatial-temporal correlation between consecutive frames, and the reconstruction quality can be improved significantly. What's more, the skipping ratio of non-keyframes is adaptive to the difference of their motion-speeds. The proposed algorithm does not need to change any decoder, so it can be easily applied to general decoders. The simulation results show that under different skipping ratios, the proposed algorithm can achieve better reconstruction quality than other existing algorithms, and thus improve the energy-efficiency of the encoder.

Document code: A **Article ID:** 1673-1905(2022)12-0755-8
DOI https://doi.org/10.1007/s11801-022-2069-7

With the emergence of new applications such as wearable multimedia sensing and micro-unmanned aerial vehicle monitoring, the internet of video things (IoVT) with encoder and decoder has shown increasingly important strategic value in the new generation of visual communication mainly for video signals[1,2]. Because the coding mode prediction, motion estimation (ME), and motion compensation (MC) will bring a huge computational burden, the traditional video encoders such as MPEG-x and H.26x cannot be applied to the encoder of a typical IoVT system[3,4]. Distributed video coding (DVC) can reduce the encoding complexity by finely modeling the correlations between consecutive frames only at a decoder[5], but it is still constrained by the Nyquist sampling theorem. The compressive sensing (CS) theory shows that at a sampling frequency far lower than the Nyquist sampling frequency, sparse signals can be recovered with high probability[6]. CS greatly improves the energy-efficiency of an encoder by synchronizing the sampling and compression, where a subrate is a ratio between the number of measurements and the signal dimension. Further, the video compressive sensing (VCS) extends the image CS, and gradually acquires each frame. For increasing the resolution, GAN[7] proposed block compressive sensing (BCS) to reduce the storage overhead of an encoder. Distributed video compressive sensing (DVCS) has the dual advantages of distributed video coding and BCS, so it has become a research hotspot in recent years[8].

An IoVT system enables real-time vision monitoring, but its monitoring time is severely limited by the insufficient battery capacity of visual sensor. It is a challenge for long-term surveillance. Video coding consumes the most energy from visual sensors, where a low-power video encoder is important to extend the monitoring time. The skip-block coding is used to save energy consumption for an encoder. Each frame performs a block-by-block observation process. When any sensor is off, a block is not observed and no energy is consumed, where these non-observed blocks are called skip-blocks. The skip-block coding might lead to the degradation of reconstruction quality. By utilizing the spatial-temporal correlation, the adaptive selection of skip-blocks can effectively improve its reconstruction quality.

COSSALTER et al[9] utilized the three-dimensional (3D) sparse transformation to reduce the temporal redundancy information of multiple frames, which is the first application of spatial-temporal correlation in VCS. MUN et al[10] proposed the MC-based residual reconstruction algorithm to improve the reconstruction quality of DCVS. To improve the reconstruction quality

* This work has been supported by the National Natural Science Foundation of China (No.62001099).
** E-mail: liuhao@dhu.edu.cn

of predicted images by motion estimation and motion compensation techniques, an optimal multihypothesis prediction algorithm was proposed according to multi-reference frames[11]. Based on the temporal correlation of consecutive frames, ZHAO et al[12] proposed the re-weighting of hypothesis sets after each iteration. In a DCVS system, each group-of-pictures (GOP) includes multiple non-keyframes, so the inter-frame correlation decreases significantly when the distance between keyframe and non-keyframe increases. To solve this problem, ZHENG et al[13] proposed a local quadratic reconstruction algorithm by multi-reference frames. Based on the spatial-temporal correlation of different blocks, a new hypotheses acquisition scheme is applied to multihypothesis VCS[14], where the temporal correlation-based mechanism is used to acquire the matching block in the key-reference-frame, and the spatial correlation-based mechanism is used for the non-key-reference-frame, and this kind of complex motion estimation is difficult to be applied to a low-power video encoder. YUE et al[15] utilized the structural similarity (SSIM) as the matching criterion to find the optimal matching block for the current frame. However, the SSIM-based matching criterion significantly increases the computational complexity, which makes it unrealistic to be applied to any low-power encoder.

All of the above algorithms are the research results of the spatial-temporal correlation mainly for improving the reconstruction quality of a decoder, but there are relatively few research results about the energy-efficiency of an encoder. For a low-power encoder, FOWLER et al[16] firstly applied the temporal correlation between consecutive frames. Based on the idea of downsampling, no observation of some blocks at intervals will not add coding distortion for highly similar non-keyframes. The CS-based terminal-to-cloud video transmission is proposed for a DVCS system, and the skip-block coding is also introduced, where the sum of absolute difference (SAD) is calculated to measure the similarity between the current frame and adjacent keyframe[13], and a threshold is used to mark some blocks with high similarity without observation. These marked blocks are called skip-blocks. Further, UNDE et al[17] combined the above-mentioned skip-block coding with rate-adaptive sampling, and utilized the mean of absolute difference (MAD) to measure the similarity between the current frame and the adjacent keyframe, and then divided the block coding mode into three types, namely, SKIP, LOW, and HIGH. Based on statistical characteristics estimation, WANG et al[18] proposed an adaptive rate BCS method for vision monitoring, where the BCS measurements of each block are used to estimate its mean and variance, and then all blocks are classified into four categories by using the Chebyshev inequality, but the rate-adaptive method must change its decoder at the same time, which makes it unsuitable for general decoders. LI et al[19] proposed a context-based allocation method for energy-efficient compressive video sensing, where the context information of each frame is extracted from its measurements instead of original pixels, and blocks are re-sampled into new measurements according to the contextual features, but the twice sampling of original signal increases the energy consumption of the encoder.

In a DVCS system, the previous keyframe is widely used as the reference-frame of non-keyframes. Because the position of a non-keyframe goes backward in the same GOP, its correlation with the previous keyframe becomes smaller, and the previous keyframe may not be enough to measure the similarity between the current block and those reference blocks, which will affect the quality of video reconstruction at general decoders. When the similarity between the current block and the reference block reaches a threshold, the existing skip-block coding algorithms generally determine whether the current block belongs to a skip-block. Based on the existing algorithms, the number of skip-blocks is unstable and insufficient, and thus the encoder cannot achieve high energy-efficiency. The motion-speed of a video sequence is a very important indicator for the selection of skip-blocks. Therefore, this paper proposes a motion-adaptive adjacent-reference skipping (MAS) algorithm, which has the advantage that a DVCS system does not change any decoder. In brief, the major contributions of this work can be summarized as follows.

For the skip-block coding, an adjacent-reference-frame analyzing module is designed at an encoder of a DVCS system. By introducing adjacent frames as reference-frames, the encoder can accurately eliminate redundant information between frames. The new module significantly reduces the coding distortion and improves the reconstruction quality of a video sequence under the same decoder.

A reference-frame similarity sorting module is also designed at our encoder. The new module solves the problem that the number of skip-blocks is unstable in different video sequences, and thus the number of skip-blocks can be adaptively determined according to the low-power conditions of an encoder.

A motion-adaptive setting module is further designed at our encoder. According to the motion-speeds of different non-keyframes in the same GOP, the module adaptively adjusts the skipping ratio of non-keyframes. This new module can enhance the reconstruction quality of a video sequence under the same skipping ratio, which further improves the energy-efficiency of the encoder.

Although the high correlation often occurs between adjacent frames in a video sequence, those blocks with extremely high similarity are still sent to the decoder in a DVCS system. This operation is a waste of limited encoder energy. It is necessary to design a simple yet effective method to remove redundant information between frames. In the case of limited resources at an encoder, to further improve the energy-efficiency of the encoder, the observation process can be skipped for some blocks in similar non-keyframes. These blocks are called skip-blocks. Tab.1 lists the mathematical notations used in this paper.

Fig.1 illustrates a typical DVCS system, where the blue shaded part is the main research content of this paper. In the figure, the blocker is used to segment a frame into many blocks. A skip-block coding algorithm will determine the similarity value of each block. At an encoder, the blocks of a frame are classified into skip-blocks and CS-blocks (non-skip-blocks). The location marker is used to convey the location information of skip-blocks. The measurements compensation module involves the completion of skip-block measurements, which will be mentioned after. Our encoder has the advantage that it can be flexibly applied to a DVCS system with general decoders.

If an encoder utilizes the key-reference-frame for analyzing and thresholding[13], the main disadvantage is that the threshold is difficult to select, and the motion-speed of a video sequence has a great impact on the selection results of skip-blocks. In the above CS-based terminal-to-cloud video transmission, the skipping ratios of three video sequences (Mother-daughter, Foreman, and Soccer) with increasing motion-speed are 40%, 10%, and 1%, respectively.

To solve the above problem, we propose a novel skip-block coding algorithm, i.e., MAS. Under a novel skipping framework, the proposed algorithm consists of three modules, as shown in Fig.2. At an encoder, a skip-block means no energy consumption without

Tab.1 Mathematics symbol

Symbol	Definition
i	Number of blocks
j	Number of frames
k	Size of GOP
B	Size of a block
m, n	Size of observation matrix
$\boldsymbol{X}_j$	Matrix of j^{th} frame
$\boldsymbol{x}_j^i$	Vector of i^{th} block of j^{th} frame
$\boldsymbol{R}_{j\text{-}1}$	j^{th} frame as reference-frame
$\boldsymbol{r}_{j\text{-}1}^i$	Vector of i^{th} block of j^{th} reference-frame
N	Total number of blocks
$m_{j,j\text{-}1}^i$	Similarity of i^{th} block between j^{th} and $(j-1)^{\text{th}}$ frames
$\boldsymbol{M}_{j,j\text{-}1}$	Frame similarity between j^{th} frame and $(j-1)^{\text{th}}$ frame
$\boldsymbol{y}_j^i$	Vector of i^{th} block measurements
$\boldsymbol{y}_j^{i,\text{skip}}$	Vector of i^{th} block measurements of j^{th} frame, which is skip-block
$\boldsymbol{y}_j^{i,\text{CS}}$	Vector of i^{th} block measurements of j^{th} frame, which is CS-block
$\boldsymbol{y}_j$	Matrix of j^{th} frame measurements
$\boldsymbol{\Phi}$	Observation matrix
SP	Specified skipping ratio
MS_j	Motion-speed of j^{th} frame compared to $(j+1)^{\text{th}}$ frame
FP_j	Fixed skipping ratio of j^{th} frame
P_j	Percentage of motion-speed contribution of j^{th} frame
AP_j	Adaptive skipping ratio of j^{th} frame
TP_j	Total skipping ratio of j^{th} frame
UB	Upper bound of TP_j

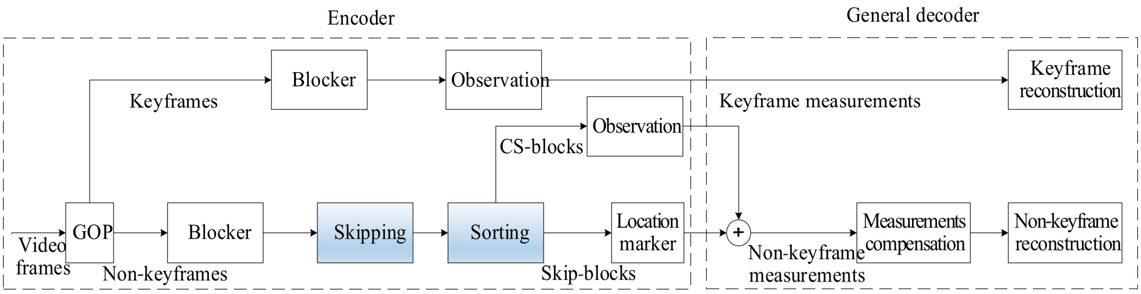

Fig.1 DVCS system with general decoders

observation. By taking the block with a small similarity value in the current frame $\boldsymbol{X}$ as a skip-block, the reference-frame similarity sorting can meet the specific energy-efficiency requirement of an encoder. The skipping ratio depends on the current low-power conditions of an encoder. When a DVCS system works in such scenarios as a battery-driven encoder, it may be required to reduce the energy consumption by half. The key-reference-frame analyzing and thresholding will not meet the requirements. However, the novel skipping framework can save half the energy by controlling every non-keyframe to skip the observation process of half blocks. Under the novel skipping framework, the proposed skip-block coding algorithm will be described after.

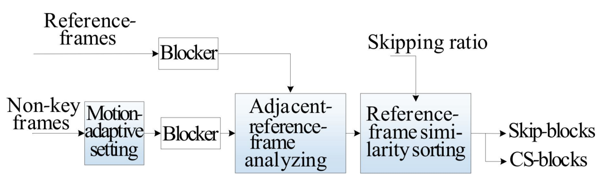

Fig.2 Novel skipping framework

Let $\boldsymbol{X}_j=\{\boldsymbol{x}_j^1,\boldsymbol{x}_j^2,\cdots,\boldsymbol{x}_j^N\}$ be a non-keyframe, and $\boldsymbol{R}_{j-1}=\{\boldsymbol{r}_{j-1}^1,\boldsymbol{r}_{j-1}^2,\cdots,\boldsymbol{r}_{j-1}^N\}$ be the reference-frame of the non-keyframe. $\boldsymbol{x}_j^i$ and $\boldsymbol{r}_{j\text{-}1}^i$ respectively denote the i^{th} block of the non-keyframe and its reference-frame, and N is the number of blocks. Then the similarity value of the current frame is as follows:

$$\boldsymbol{M}_{j,j-1}=[m_{j,j-1}^{1},\ldots,m_{j,j-1}^{N}]=\\ [SAD(\boldsymbol{x}_{j}^{1},\boldsymbol{r}_{j-1}^{1}),\ldots,SAD(\boldsymbol{x}_{j}^{N},\boldsymbol{r}_{j-1}^{N})], \quad (1)$$

where $m_{j,j-1}{}^{i}$ is a constant that denotes the similarity between the block $\boldsymbol{x}_j{}^i$ and $\boldsymbol{r}_{j-1}{}^i$. The definition of $SAD(\cdot)$ is

$$SAD(\boldsymbol{x}_j^i,\boldsymbol{r}_j^i)=\left\|\boldsymbol{x}_j^i-\boldsymbol{r}_j^i\right\|_1 . \quad (2)$$

The existing skip-block coding algorithms use the previous keyframe as the reference of a non-keyframe. If the non-keyframe located in the back of a GOP refers to the previous keyframe, the similarity value will be inaccurate. That is, the keyframe may not be enough to measure the similarity between the current block and the reference block. When the skip-blocks are improperly selected, the reconstruction quality of a video sequence will decrease significantly.

Let $\boldsymbol{X}_1=\{\boldsymbol{x}_1^1,\boldsymbol{x}_1^2,\cdots,\boldsymbol{x}_1^N\}$, $\boldsymbol{X}_5=\{\boldsymbol{x}_5^1,\boldsymbol{x}_5^2,\cdots,\boldsymbol{x}_5^N\}$, and $\boldsymbol{X}_6=\{\boldsymbol{x}_6^1,\boldsymbol{x}_6^2,\cdots,\boldsymbol{x}_6^N\}$ be the first frame, the fifth frame and the sixth frame in the same GOP of a video sequence, where the first frame is a keyframe. For $\boldsymbol{X}_6$, when taking $\boldsymbol{X}_1$ as a reference, the similarity value of its i^{th} block is

$$m_{1,6}^i=SAD(\boldsymbol{x}_1^i,\boldsymbol{x}_6^i). \quad (3)$$

When taking $\boldsymbol{X}_5$ as a reference, the similarity value of its i^{th} block is

$$m_{5,6}^i=SAD(\boldsymbol{x}_5^i,\boldsymbol{x}_6^i). \quad (4)$$

If $m_{1,6}{}^i \ll m_{5,6}{}^i$, the key-reference-frame analyzing and thresholding will regard the i^{th} block as a skip-block, while the adjacent-reference-frame analyzing will not. Fig.3 demonstrates the diagram of adjacent-reference-frame analyzing. A skip-block means that the block is not observed, and only the position information of the block is sent. Before reconstructing the video sequence, the decoder needs to complete the measurements of skip-blocks. In order not to change any decoder, the adjacent-reference-frame analyzing module will compensate the measurements of the skip-block with the measurements of the adjacent frame. If the block corresponding to the adjacent frame is also a skip-block, the corresponding block of the previous frame will be taken. Fig.4 illustrates our skip-block measurements compensation.

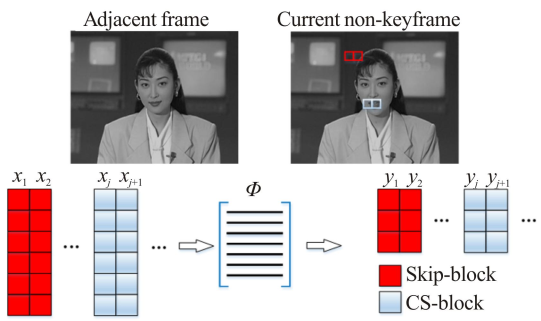

Fig.3 Diagram of adjacent-reference-frame analyzing

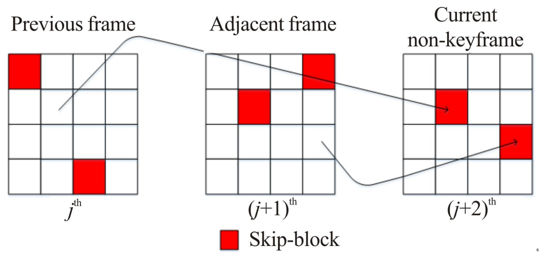

Fig.4 Skip-block measurements compensation of adjacent-reference-frame analyzing

The motion-speed difference of various video sequences causes the number of skip-blocks by the threshold-based sorting will be extremely unstable. Our motion-adaptive setting module is based on the motion-speed calculation of each non-keyframe. A slow motion-speed of a non-keyframe means that it is highly similar to the previous frame, and then more skip-blocks will be extracted from this particular frame. Hence, different numbers of skip-blocks can be allocated to different non-keyframes according to their motion-speeds. Fig.5 illustrates the flow chart of the motion-adaptive setting. The frames in a GOP are divided into reference part and adaptive part. The skipping ratio of frames in the reference part is a specified ratio, and the skipping ratio of frames in the adaptive part consists of fixed and adaptive ratios.

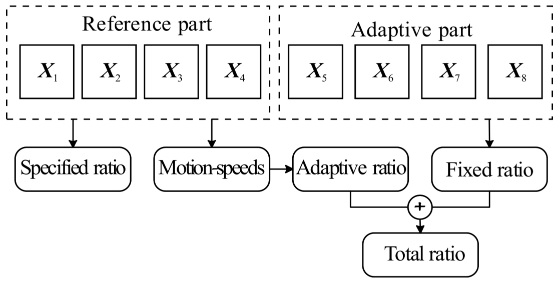

Fig.5 Flow chart of the motion-adaptive setting

Under the novel skipping framework, Tab.2 gives the main pseudo code of the proposed MAS algorithm, where the specific procedures or some other details are explained as follows.

(1) k frames (j = 1, 2, 3, …, k) in a GOP (k is even) are considered. The 1$^{\text{st}}$ to $(k/2)^{\text{th}}$ frames are classified as the reference part, and the rest are classified as the adaptive part. With the specified skipping ratio SP, the skipping ratio of frames in the reference part is SP, and the motion-speed MS_j (j = 1, 2, 3, … k/2) of each frame in the reference part is calculated by

$$MS_j=\left\|\boldsymbol{X}_j-\boldsymbol{X}_{j+1}\right\|_1 . \quad (5)$$

(2) The percentage P_j (j=1, 2, 3, …, k) of the motion-speed contribution of each frame in the reference part is

$$P_j=\frac{MS_j-\max(\boldsymbol{MS})}{\sum_{j=1}^{k/2}(MS_j-\max(\boldsymbol{MS}))}, \quad (6)$$

where $\boldsymbol{MS}$ is the vector that consists of MS_j, and max($\boldsymbol{MS}$) is the maximum MS_j in $\boldsymbol{MS}$.

Tab.2 Pseudo code of the proposed MAS algorithm

Algorithm 1 MAS

Input: A GOP $\boldsymbol{X}_j$ (j=1, 2, 3, …k) with the GOP size k,
the specified skipping ratio SP,
the fixed-skipping allocation parameter W,
the block size B,
the total number N of blocks of $\boldsymbol{X}_j$,
Output: The skipping ratio of each frame TP_j.

for j = 1 to k/2 **do**
  $TP_j = SP$;
  MS_j = abs ($\boldsymbol{X}_j - \boldsymbol{X}_{j+1}$);
end for
for j = 1 to k/2 **do**
  $P_j = (MS_j - \max(\boldsymbol{MS})) / (\mathrm{sum}(\boldsymbol{MS} - \max(\boldsymbol{MS})))$
end for
for $j = k/2 + 1$ to k **do**
  $FP_j = W * SP$
  $AP_j = P_{(j-k/2)} * k/2 * (SP - FP_j)$
  $TP_j = FP_j + AP_j$
end for
for $j = k/2+1$ to k **do**
  if $TP_j > UB$ **do**
    for $j = k/2+2$ to k **do**
      $TP_j = TP_j + (TP_j - UB)/(k-j)$
    end for
    $TP_j = UB$
  end if
end for
$\boldsymbol{x}_i^j$ = im2col ($\boldsymbol{X}_j$, B)
for i = 1 to N **do**
$m_{j,j-1}^i = SAD(\boldsymbol{x}_j^i, \boldsymbol{x}_{j-1}^i)$
end for
sort($\boldsymbol{M}_{j,j-1}$)
$\boldsymbol{y}_j^{i,\mathrm{skip}} = \boldsymbol{y}_{j-1}^i$
$\boldsymbol{y}_j^{i,\mathrm{CS}} = \boldsymbol{\Phi x}_j^i$
$\boldsymbol{y}_j = \boldsymbol{y}_j^{i,\mathrm{skip}} + \boldsymbol{y}_j^{i,\mathrm{CS}}$

(3) To ensure better reconstruction quality for slow motion-speed frames, the fixed skipping ratio FP_j (j = k/2+1, k/2+2, k/2+3, …, k) for each frame in the adaptive part is calculated by

$$FP_j = W \times SP, \tag{7}$$

where W is a fixed-skipping allocation parameter.

(4) The adaptive skipping ratio AP_j (j = k/2+1, k/2+2, k/2+3, …,k) of each frame in the adaptive part is calculated by

$$AP_j = P_{(j-k/2)} \times \frac{k}{2} \times (SP - FP_j). \tag{8}$$

(5) The total skipping ratio TP_j of each frame in the adaptive part is calculated by

$$TP_j = FP_j + AP_j. \tag{9}$$

(6) To ensure better reconstruction quality for fast motion-speed frames, we set an upper bound UB for frame in the adaptive part. If TP_j>UB, the extra skipping ratio is evenly distributed to other frames in the adaptive part.

The two main parameters of the MAS algorithm are the fixed-skipping allocation parameter W and the upper bound UB. Theoretically, the skipping ratio of a frame is from 0 to 100%. A 100% skipping ratio means that the current frame is not observed at all, which is unreasonable. To avoid the extreme skipping ratio, the MAS algorithm robustly set the two parameters. W is the minimum proportion of the specified skipping ratioper frame, typically W=0.5. In the case that the specified skipping ratio SP<0.5, UB is set to twice the specified skipping ratio, and 0.9 in other cases. The principle of the MAS algorithm is that the motion-speed of the frame in the first half of the GOP (frames in the reference part) is used as a reference for the frame in the second half of the GOP (frames in the adaptive part). The skipping ratio of the frame in the adaptive part is estimated according to the motion-speed of the frame in the reference part. Hence, the frame with slow motion-speed will be assigned a high skipping ratio precisely, which can improve the reconstruction quality.

The proposed skip-block coding algorithm is compared with some state-of-the-art algorithms on MATLAB R2020b. The simulation platform is 64-bit Windows 10, IntelCore(TM) i3-10500 CPU, 3.60 GHz, 16 G RAM. The experiments are based on the DVCS system with general decoders, and the Gaussian random matrix is used as the observation matrix. The first 25 frames of the standard test sequences (Coastguard, Soccer, Foreman, and Football) are used as experimental videos. The motion-speeds of these sequences are from slow to fast, and almost all motion features are included. Similar to the existing algorithms, the keyframe subrate is set to 0.7, and the non-keyframe subrate varies from 0.2 to 0.4. The GOP size is 8 and the block size is 16×16 in all experiments, and these parameters can well balance the computational overhead of an encoder and the reconstruction quality at a decoder.

We will compare the performance of the proposed MAS algorithm with equal skipping (ES)[16], random skipping (RS)[12], key-reference-frame analyzing and thresholding (KAT)[17]. To test the ablation performance, we delete the motion-adaptive setting module in the proposed MAS algorithm, and name it as MAS-. The skipping ratio of an encoder is 30%, 50%, and 70%. To ensure fairness, all algorithms are performed under the same skipping ratio, and the reconstruction algorithm is always the reweighted residual sparsity algorithm at general decoders[12].

The comparisons of the GOP-wise average peak signal-to-noise ratio ($PSNR$) under 50% skipping ratio are shown in Tab.3 for instance. The keyframe subrate and non-keyframe subrate are 0.7 and 0.4. The maximum value is marked in bold. In most cases, the MAS has the best GOP-wise average $PSNR$ in all video sequences and all GOPs. Under different non-keyframe subrates and different skipping ratios, the average $PSNR$ results of MAS, MAS- and KAT are shown in Fig.6. It can be seen that compared with KAT, the performance of MAS- is improved at each subrate and each skipping ratio, and the largest improvement is about 1 dB. And for MAS, compared with KAT, the largest improvement is about 1.5 dB. For a video sequence, the higher the skipping ratio is, the higher the improvement of MAS and MAS- compared with KAT. For those video sequences with fast motion and complex texture, such as Football, many non-information-redundant blocks are selected as skip-block by using the existing skip-block coding algorithms. Correspondingly, our MAS algorithm utilizes adjacent frames as a reference, which can more accurately find the information-redundant blocks and then mark them as skip-blocks.

The addition of the motion-adaptive setting module makes the performance of MAS better. It can be seen from Fig.6 that as compared with MAS-, the largest improvement of MAS is about 0.2 dB. Similarly, for the same sequence, compared with MAS-, the higher the skipping ratio is, the higher the improvement of MAS is. Because the larger the specified skipping ratio is, the greater the difference between the skipping ratio of each frame within the same GOP becomes. What's more, compared with MAS-, the faster the motion-speed of a sequence is, the higher the improvement of MAS is. Compared with a fixed skipping ratio to all non-keyframes, the adaptive skipping ratio can bring better reconstruction quality.

Tab.3 GOP-wise average *PSNR* (dB) under the 50% skipping ratio

GOP	Algorithm	Video sequence			
		Coastguard	Soccer	Foreman	Football
I	ES	29.09	22.41	28.31	21.64
	RS	29.29	23.19	28.06	21.85
	KAT	30.94	30.61	34.99	26.15
	MAS-	**31.00**	31.19	**36.11**	26.43
	MAS	30.99	**31.26**	36.01	**26.60**
II	ES	28.59	22.84	27.30	20.64
	RS	28.83	23.57	27.00	21.47
	KAT	30.84	29.74	29.92	24.35
	MAS-	30.95	30.13	31.07	25.00
	MAS	**31.00**	**30.32**	**31.40**	**25.17**
III	ES	27.72	20.74	27.30	20.61
	RS	28.13	21.79	27.29	20.89
	KAT	30.35	28.90	31.15	24.45
	MAS-	30.53	29.14	32.77	**25.09**
	MAS	**30.65**	**29.75**	**32.94**	25.01

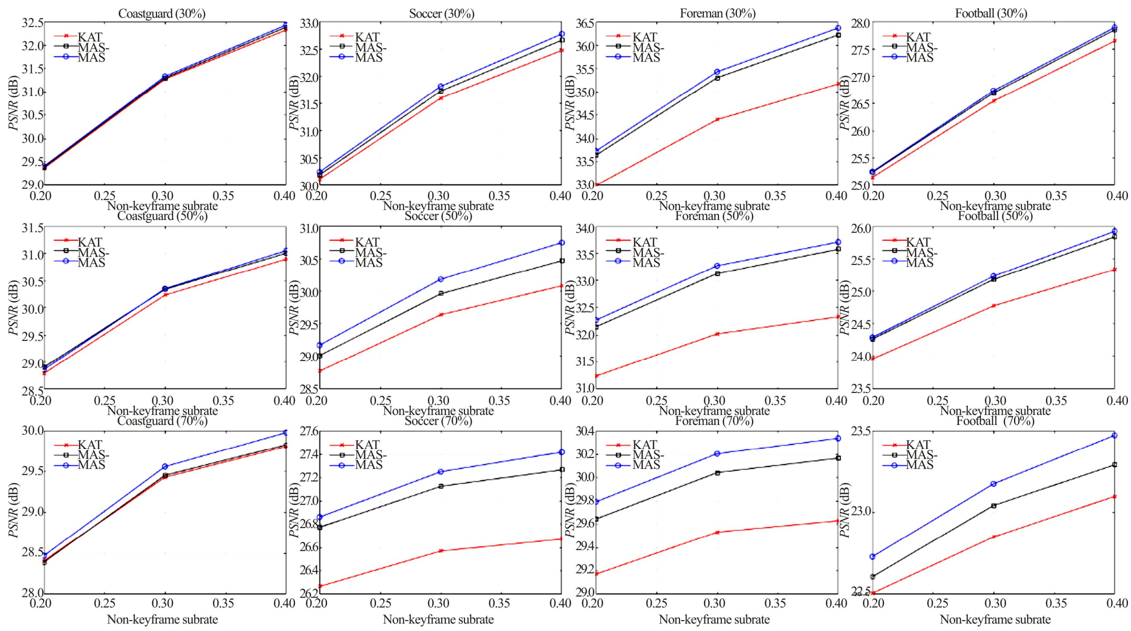

Fig.6 Average *PSNR* results for different algorithms under different non-keyframe subrates and skipping ratios

Fig.7 shows the influence of the skipping ratio on the reconstruction quality. The vertical coordinate is the average *PSNR* for all subrates. It can be seen from Fig.7 that the reconstruction quality almost decreases linearly with the increase of skipping ratio, and the reconstruction quality decreases more rapidly for fast-motion video sequences. The proposed MAS and MAS- algorithms not only improve the reconstruction quality at all skipping ratios, but also reduce the descent speed of reconstruction quality when the skipping ratio increases.

To illustrate the performance advantage of our MAS more clearly, Fig.8 shows the comparison of frame-by-frame *PSNR* results of the KAT, MAS- and MAS under the keyframe subrate of 0.7 and non-keyframe subrate of 0.4. It can be seen that the performance of the MAS algorithm is always competitive especially for highly dynamic video.

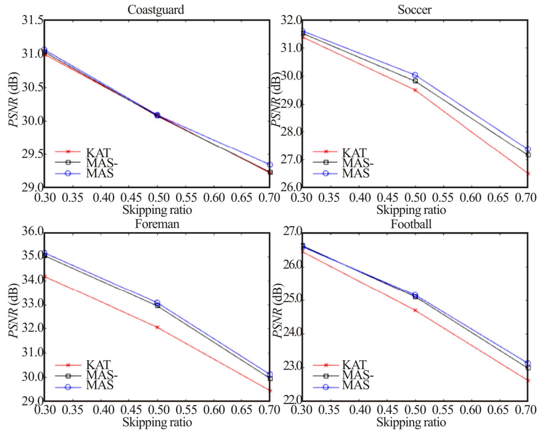

Fig.7 Average *PSNRs* for different skipping ratios

Finally, we evaluate the computational complexity of the proposed algorithm. Compared with existing skip-block coding algorithms, the extra runtime of the MAS algorithm is mainly determined by the computation of adaptive skipping ratios. We choose the Foreman sequence as a typical example. The relative extra runtime is the incremental ratio between the average coding runtime of the MAS algorithm and that of the RS algorithm. For the MAS and RS algorithms, Tab.4 gives the average coding runtime per frame of the Foreman sequence under different skipping ratios. As can be seen from the table, the relative extra runtime only increases by 11.49% at most, which can significantly improve the reconstruction quality of video sequence with such a small increase in computational complexity. Furthermore, the encoder under higher skipping ratio runs faster due to the fact that high skipping ratio savesmore runtime for the block-by-block observation process.

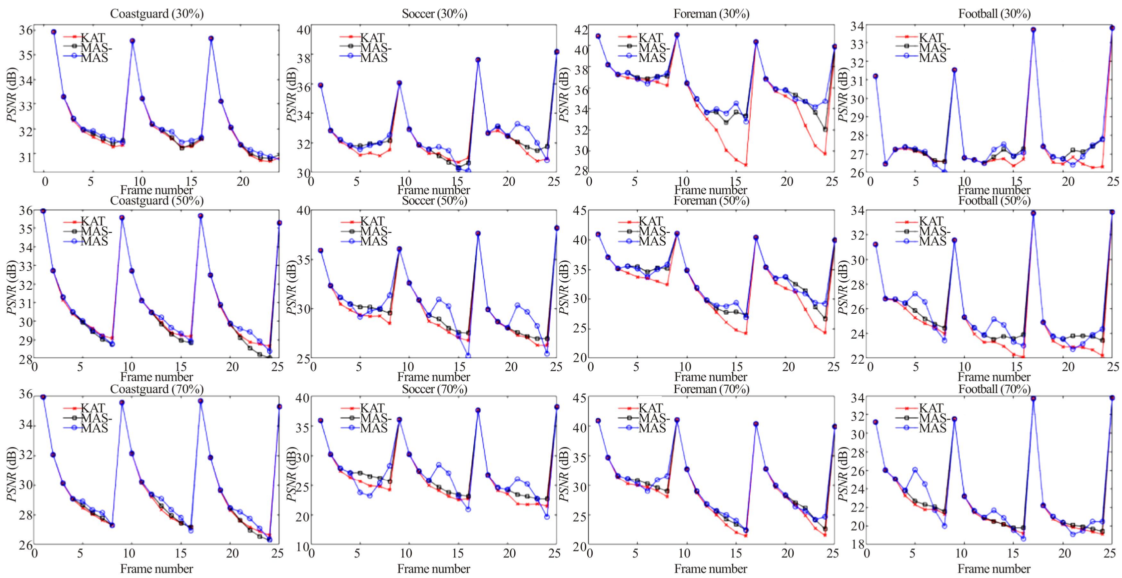

Fig.8 Frame-by-frame *PSNR* results for different algorithms

Tab.4 Average coding runtime per frame of the Foreman sequence for the MAS and RS algorithms

Skipping ratio	Algorithm	Runtime (s)	Relative
30%	RS	9.340	8.09%
	MAS	10.096	
50%	RS	8.884	10.63%
	MAS	9.828	
70%	RS	8.564	11.49%
	MAS	9.548	

Our MAS algorithm can enhance the reconstruction quality of video sequence by mining the spatial-temporal correlations of consecutive frames. The proposed algorithm designs a reference-frame similarity sorting module that can stabilize the number of skip-blocks. In addition, the proposed algorithm is suitable for general decoders without any change. Currently, the proposed algorithm selects the measurements of the previous frame to compensate for the measurements of the skip-blocks, and subsequent work needs to find a more suitable skip-block compensation mechanism to reduce blocking artifact.

Statements and Declarations

The authors declare that there are no conflicts of interest related to this article.

References

[1] CHEN C W. Internet of video things: next-generation IoT with visual sensors[J]. IEEE internet of things journal, 2020, 7(8): 6676-6685.

[2] YANG Y, GUO S, LIU G, et al. Two-layer compressive sensing based video encoding and decoding framework for WMSN[J]. Journal of network and computer applications, 2018, 117：72-85.

[3] ZHANG Y, ZHANG C, FAN R, et al. Recent advances on HEVC inter-frame coding：from optimization to implementation and beyond[J]. IEEE transactions on circuits and systems for video technology, 2020, 30(11)：4321-4339.

[4] YANG H. Subblock-based motion derivation and inter prediction refinement in the versatile video coding standard[J]. IEEE transactions on circuits and systems for video technology, 2021, 31(10)：3862-3877.

[5] SHANG F, YE J J. Review of distributed video coding[C]//13th IEEE International Conference on Electronic Measurement & Instruments, October 20-22, 2017, Yangzhou, China. New York：IEEE, 2017：315-320.

[6] RANI M, DHOK S B, DESHMUKH R B. A systematic review of compressive sensing：concepts, implementations and applications[J]. IEEE access, 2018, 6：4875-4894.

[7] GAN L. Block compressed sensing of natural images[C]//15th International Conference on Digital Signal Processing, 2007, Cardiff, UK. New York：IEEE, 2007：403-406.

[8] BAIG Y, LAI E M K, PUNCHIHEWA A. Distributed video coding based on compressed sensing[C]//2012 International Conference on Multimedia and Expo Workshops, July 9-13, 2012, Melbourne, Australia. New York：IEEE, 2012：325-330.

[9] COSSALTER M, VALENZISE G, TAGLIASACCHI M, et al. Joint compressive video coding and analysis[J]. IEEE transactions on multimedia, 2010, 12(3)：168-183.

[10] MUN S, FOWLER J E. Residual reconstruction for block-based compressed sensing of video[C]//2011 Data Compression Conference, March 23-26, 2011, Snowbird, USA. New York：IEEE, 2011：183-192.

[11] ZHOU J, LIU H. Video compressive sensing reconstruction via long-short-term double-pattern prediction[J]. Optoelectronics letters, 2020, 16(3)：230-236.

[12] ZHAO C, MA S, ZHANG J, et al. Video compressive sensing reconstruction via reweighted residual sparsity[J]. IEEE transactions on circuits and systems for video technology, 2017, 27(6)：1182-1195.

[13] ZHENG S, ZHANG X P, CHEN J, et al. A high-efficiency compressed sensing-based terminal-to-cloud video transmission system[J]. IEEE transactions on multimedia, 2019, 21(8)：1905-1920.

[14] ZHENG S, CHEN J, ZHANG X P, et al. A new multihypothesis-based compressed video sensing reconstruction system[J]. IEEE transactions on multimedia, 2021, 23：3577-3589.

[15] YUE Y, LUO J, HUA L. Distributed video compressed sensing secondary reconstruction based on inter-frames similarity structure[J]. IOP conference series：materials science and engineering, 2021, 1043(5)：052012.

[16] FOWLER J E. Block-based compressed sensing of images and video[J]. Foundations & trends in signal processing, 2012, 4(4)：297-416.

[17] UNDE A S, PATTATHIL D P. Adaptive compressive video coding for embedded camera sensors：compressed domain motion and measurements estimation[J]. IEEE transactions on mobile computing, 2020, 19(10)：2250-2263.

[18] WANG J, WANG W, CHEN J. Adaptive rate block compressive sensing based on statistical characteristics estimation[J]. IEEE transactions on image processing, 2022, 31：734-747.

[19] LI R, YANG Y, SUN F. Green visual sensor of plant：an energy-efficient compressive video sensing in the internet of things[J]. Frontiers in plant science, 2022, 13：849606.