Multi-hypothesis distributed video compression sensing based on key frame secondary reconstruction

Yue Yuchen¹, Luo Jiahua² and Li Hua³

¹PhD student, Army Armoured Force Academy, Beijing, China
² Professor, Army Armoured Force Academy, Beijing, China
³ Associate Professor, Army Armoured Force Academy, Beijing, China

E-mail: gonzo123@sina.com

Abstract. Reconstruction algorithms are the key technology of distributed video compressed sensing. The research focus of traditional distributed video compressed sensing reconstruction algorithms is mostly on improving the reconstruction quality of non-key frames, ignoring the reconstruction quality of key frames, and the information of key frames are not underutilized. In view of the above problems, a distributed video compression sensing algorithm based on secondary reconstruction of key frames is proposed. Firstly, the fractional order total variation algorithm is used for the initial reconstruction of the key frame, and the reconstructed frame is used as the reference frame to assist the secondary reconstruction of the key frame, which improves the reconstruction quality and reduces the calculation complexity. Then, a multi-reference frame bidirectional prediction hypothesis set optimization algorithm is proposed to increase the number of reference frames and improve the quality of the hypothesis set through optimization without expanding the size of the hypothesis set. Experimental results show that the overall performance of the proposed algorithm is better than the most advanced methods.

1. Introduction

The emergence of Compressed Sensing [1] (CS) technology breaks through the limitation of Nyquist sampling. It can be used to undersample the sparse or compressible signal at a sampling rate far below the Nyquist criterion and recover with a high probability and low sampling rate. The lower limit of sampling rate is no longer limited by bandwidth. The CS framework processes signal acquisition and compression in parallel, and saving a lot of sampling resources compared to traditional Nyquist sampling. Compressed sensing greatly reduces signal storage costs, transmission costs, and signal processing computational complexity. Compressed Video Sensing (CVS) [2-5] uses CS technology in the video signal encoding and decoding process, and transfer the computational pressure of video encoding to the decoder. CVS provides an effective solution for application scenarios where video signals are limited in resources at the encoder.

Distributed Video Coding (DVC) system reduces the computational complexity of the encoding side by moving the steps with high computational complexity such as motion estimation and motion
compensation to the decoding side. Distributed Compressive Video Sensing (DCVS) was proposed by Do [6]. Compressed sensing was introduced into the distributed video coding framework. The video frame was divided into multiple image groups at the coding end. The first frame of each Group of Picture (GOP) is a key frame, and the remaining frames are CS frames. Key frames with a high sampling rate use traditional intra-frame coding and CS frames with a low sampling rate use compressed sensing coding. The key frame is decoded independently at the decoding end, and then the side information is generated to assist the decoding of the CS frame. The distributed compressed video sensing scheme proposed by Kang [7] uses CS encoding for both key frames and CS frames, which further reduces the computational complexity of the encoding end. Chen [8] proposed a multi-hypothesis (MH) distributed compressed video sensing algorithm, which uses reference frames to obtain the optimal hypothesis set to assist the CS frame reconstruction. However, not only the calculation complexity of obtaining the optimal hypothesis set is large, but also the algorithm runs slow. Fowler [9] et al. proposed the MH-BCS-SPL algorithm that combines the advantages of multiple hypotheses and Landweber smooth projection technology. This algorithm adds constraints to the hypothesis weight vectors through the Tikhonov regularization method to improve the quality of multiple hypothesis prediction. Li [10] proposed a distributed adaptive video compression perception reconstruction algorithm based on Landweber smooth projection. After obtaining the edge information of the CS frame by interpolation, the block containing large distortion was detected, and the measurement number was adaptively allocated and performed. The residual reconstruction results in a higher quality CS frame. The hypothesis set optimization technique proposed by Kuo [11] interpolates reference frames to obtain more hypotheses, which are then used to replace the hypothesis with a small contribution rate in the original hypothesis set.

If the GOP in DCVS is larger, the number of CS frames is greater. Due to the low sampling rate of CS frames, the reconstruction quality is poor. Improving the reconstruction quality of CS frames plays an important role in improving the overall reconstruction quality of the video sequence. Therefore, the reconstruction of non-key frames is improved. Quality is the research focus of DCVS. The reference [10-14] all use key frame reconstruction frames as CS reference frame auxiliary reconstruction. When the GOP is large, the correlation between the key frame and the far (GOP distance) CS frame is poor, making the reconstruction quality of the CS frame difficult to guarantee, and does not make use of the reconstructed neighboring CS frames with high correlation with the CS frame to be reconstructed. Reference [15] proposes an algorithm for the secondary reconstruction of key frames, which generates side information (SI) from the reconstructed CS frames to assist the secondary reconstruction of the key frames to improve the reconstruction quality of the key frames. However, the reconstruction quality of CS frames is poor under the low sampling rate. The resulting SI does not ensure the improvement of the quality of the secondary reconstruction of the key frame. When processing high-speed video sequences, the quality of the secondary reconstruction of the key frame may even decline [16]. References [17] and [18] use multi-hypothesis to predict the secondary reconstruction of key frames, respectively, by expanding the search window and reference frame difference to obtain more alternative hypotheses, and then performing optimal hypothesis solving and hypothesis set update to improve the hypothesis set quality. However, the above two operations for expanding the size of the alternative hypothesis set instead increase the computational complexity.
To solve the above problems, a multi-hypothesis distributed video compressed sensing reconstruction algorithm based on the secondary reconstruction of key frames is proposed. First, Fractional-order Total Variation (FrTV) is used to achieve the initial reconstruction of key frames to ensure the quality of the initial reconstruction of key frames. Then a multi-reference frame bidirectional prediction hypothesis set optimization selection algorithm is proposed. The algorithm uses the initial reconstruction frame of the key frame as the reference frame to synchronize the bidirectional reconstruction of the CS frame with the highest correlation with the current key frame, that is, the second frame of the current GOP and the last frame of the previous GOP. After that, the reconstructed CS frame is also used as a reference frame to reconstruct other CS frames, thereby expanding the set of multi-hypothesis candidate prediction vectors. Next, select the high-quality hypothesis and eliminate those poor quality hypothesis which based on the assumption of contribution rate. Finally, a second reconstruction of the key frame is performed to further improve the quality of the key frame reconstruction.

The rest of the paper is organized as follows: Section 2 briefly introduces the fractional order total variation algorithm and multi-hypothesis prediction. Section 3 introduces the overall framework and detailed information of the algorithm. Section 4 carries out experiments and analyzes the experimental results. Section 5 draws conclusions and summarizes future work ideas.

2. Related knowledge

2.1 Fractional order total variation sparse model and reconstruction algorithm

The TV model has the advantages of strong robustness and no need to design sparse basis. Only the sparseness of the video image itself can be used as a priori information to complete the reconstruction. However, the TV model assumes that the video image is segmented and smooth, so it is inevitable that some texture and details will be lost in the reconstructed image, and the excessively smooth area of the reconstructed image will also produce a staircase effect that affects the visual sense [19]. A fractional order total variation model is introduced to solve the above problems. For the size of the image, the elements representing the rows and columns of the image, the image is observed using the two-dimensional projection method:

\[ Y = AXB^T \]  

(1)

In equation (1), the Gaussian random matrix \( A \in R^{M \times N} \) and \( B \in R^{M \times N} \) are both observation matrices. The observation matrix of \( M \times M \) is obtained by the two-dimensional projection of Image \( X \) on \( A \) and \( B \). Therefore the compressed sensing observation of the image is achieved.

The isotropic fractional order total variation model of natural imagine from the order of any real number \( \alpha \in R^+ \) is defined as below.

\[ FrTV_2^\alpha(X) = \sum_{i=1}^{N} \sum_{j=1}^{N} \sqrt{(D_i^\alpha X)_{i,j}^2 + (XD_i^\alpha)_{i,j}^2} \]  

(2)

The anisotropic fractional order total variation model is defined as:

\[ FrTV_1^\alpha(X) = \|D_i^\alpha X\|_{\alpha} + \|XD_i^\alpha\|_{\alpha} \]  

(3)
where $D_h^\alpha, D_v^\alpha \in \mathbb{R}^{N \times N}$ in equations (2) and (3) are the horizontal and vertical gradients, respectively, which are defined as:

$$
D_h^\alpha = \begin{bmatrix}
W_0^{(a)} & W_1^{(a)} & \cdots & W_{N-1}^{(a)} \\
0 & W_0^{(a)} & \cdots & W_{N-2}^{(a)} \\
\vdots & \vdots & \ddots & \vdots \\
0 & \cdots & 0 & W_0^{(a)}
\end{bmatrix}
$$

$$
D_v^\alpha = \begin{bmatrix}
W_0^{(a)} & 0 & \cdots & 0 \\
W_1^{(a)} & W_0^{(a)} & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
W_{N-1}^{(a)} & W_{N-2}^{(a)} & \cdots & W_0^{(a)}
\end{bmatrix}
$$

(4)

$$
W_i^{(a)} = (-1)^i \frac{\Gamma(\alpha + 1)}{\Gamma(i + 1)\Gamma(\alpha - i + 1)} = (-1)^i \binom{\alpha}{i}, i = 0, 1, \ldots, N - 1
$$

(5)

Where $\binom{\alpha}{i}$ is the generalized binomial coefficient. Combining equations (2) and (3), the fractional order total variation compressed sensing reconstruction model can be designed as:

$$
\min_X FrTV_p^\alpha(X) \quad \text{s.t.} \quad Y = AXB^T
$$

(6)

When $p = 1$ is the anisotropic fractional order total variation compression reconstruction model, and when $p = 2$ is the isotropic fractional order total variation model. The objective function FrTV of equation (6) is not smooth. In order to solve the optimization problem of equation (6), FrTV is transformed into a series of relatively simple iterative convex optimization problems based on the Majorization-Minimization (MM) algorithm in [19].

2.2 MH-BCS-SPL

BCS-SPL [20] is an efficient and high-quality reconstruction algorithm. Reference [21] adds multi-hypothesis prediction technology to the MH-BCS-SPL proposed in BCS-SPL to better use inter-frame correlation, making the reconstruction quality further improved, so this paper chooses this algorithm for subsequent key frame secondary reconstruction and CS frame residual reconstruction.

The original video frame image $X$ is divided into block $b$, and the block $x_i$ and its prediction $\tilde{x}_i$ satisfies:

$$
x_i = \tilde{x}_i + r_i, i = 1, 2, \ldots, b
$$

(7)

Where $r_i$ is the residual of $x_i$ and $\tilde{x}_i$. Through independent measurement of residuals $r_i$ through measurement matrix $\Phi$:

$$
q_i = \Phi r_i = y_i - \Phi \tilde{x}_i
$$

(8)
Where \( y_i \) is the measured value of the block \( x_i \), and \( q_i \) is the measured value of \( r_i \).

Combining equations (7) and (8), the final reconstruction \( \hat{x}_i \) obtained \( x_i \) is:

\[
\hat{x}_i = \tilde{x}_i + \text{reconstruct}(q_i, \Phi)
\]  

(9)

In equation (9), \( \text{reconstruct}(\bullet) \) represents the process of residual reconstruction. The quality of \( \hat{x}_i \) is directly related to the ability to rebuild from \( r_i \).

The multi-hypothesis (MH) of measurement domain prediction can better retain the structural information of the block \( x_i \) to be reconstructed [18] and improve the algorithm's residual reconstruction ability. In the sequence of video to be reconstructed, the solution process in the measurement domain for the image block \( x_{ik} \), which is the number \( i \) for the image \( X_k \) at the frame \( k \), and is as big as \( B \times B \) is as below:

\[
\tilde{x}_{ik}^{mh} = H_{k,i} w_{k,i}^{mh}
\]  

(10)

\[
w_{k,i}^{mh} = \arg \min_w \left\| y_{k,i} - \Phi_B H_{k,i} w \right\|_2^2 + \lambda \left\| \Gamma_{k,i} w \right\|_2^2
\]  

(11)

Where \( \tilde{x}_{ik}^{mh} \) is the estimated vector of \( x_{ik} \) in equation (10). \( H_{k,i} \) is the set of hypothesis vectors for \( x_{ik} \). Where \( w_{k,i}^{mh} \) is the weight vector for each hypothesis in the hypothesis set [21] \( y_{k,i} \) is the measured value of the original image. \( \lambda \) is the scale factor for controlling the degree of influence of the Tikhonov regular term. The Tikhonov regularization matrix \( \Gamma_{k,i} \) reflects the distance between the current block and each hypothetical block in the measurement domain, \( \Gamma_{k,i} = \text{diag} \left( \| y_{k,i} - \Phi_B h_1 \|_2, \| y_{k,i} - \Phi_B h_2 \|_2, \ldots, \| y_{k,i} - \Phi_B h_K \|_2 \right) \). \( h_j \) corresponds to the column \( j \) of \( H_{k,i} \), and \( j = 1, 2, \ldots, K \) \( \| y_{k,i} - \Phi_B h_j \|_2 \) is the Euclidean distance of \( y_{k,i} \) from the corresponding hypothesis \( h_j \). The larger the value of \( \| y_{k,i} - \Phi_B h_j \|_2 \) is, the lower the contribution of the corresponding hypothesis \( h_j \). The method of solving the weights \( w_{k,i}^{mh} \) is:

\[
w_{k,i}^{mh} = \left( (\Phi_B H_{k,i})^T (\Phi_B H_{k,i}) + \lambda^2 \Gamma_{k,i}^T \Gamma_{k,i} \right)^{-1} (\Phi_B H_{k,i})^T y_{k,i}
\]  

(12)

Equation (12) requires only matrix operations for the solution of \( w_{k,i}^{mh} \), and it is fast. The operation
using Tikhonov regularization matrix $\Gamma$ to hypothesis and filter is relatively simple, and the hypothesis quality is high. The image block $x_{i,j}$ reconstructed by the solution is expressed as:

$$x_{i,j} = H_{i,j}w_{i,j} + r_{i,j}$$  \hspace{1cm} (13)

### 3. The proposed algorithm

Existing multi-hypothesis algorithm, taking MC-BCS-SPL and MH-BCS-SPL as an example, when predicting multiple hypothesis sets, the reference frame reconstructed from each non-key frame comes from two adjacent key frames, that is, non-reference frames. The key frame of the current GOP and the key frame of the next GOP, the process of obtaining the hypothesis set is shown in figure 1.

![Figure 1. Schematic diagram of multiple hypothesis acquisition based on search window.](image)

First, centering on the image block to be reconstructed, the search window is constructed by expanding the width outward at the corresponding position in the reference frame. Then, the pixel-by-pixel sliding in the search window is used to obtain hypotheses in sequence, forming a hypothesis set. Although key frames with a high sampling rate can obtain good reconstruction results, adjacent key frames are less relevant for some non-key frames in the GOP that are far away from each other, making the key frames function as reference frames. Reduced, there will be a large number of hypotheses in the multi-hypothesis set that have a low contribution, and even will reduce the quality of non-key frame reconstruction; reference [18] adopts the method of expanding the search window to obtain more hypotheses, but it also causes a large number of hypotheses with a low contribution rate are added to the hypothesis set, and a large number of correlation calculations must be performed to eliminate the inferior hypotheses, resulting in an increase in reconstruction time and a significant increase in the amount of calculation. Reference [23] searched part of the CS reference frames to increase the number of alternative hypotheses, but this method also made a large number of hypotheses with low contribution rates added to the hypothesis set, which not only increased the algorithm delay, when the CS frame sampling rate was low, it will even reduce the quality of reconstruction. Therefore, this paper proposes a multi-reference frame bidirectional prediction hypothesis set optimization algorithm, which introduces a high contribution rate hypothesis as much as possible without increasing the size of the hypothesis set to improve the reconstruction quality.

#### 3.1. Multi-reference frame bidirectional prediction hypothesis set optimization algorithm
The overall framework of the algorithm is shown in figure 2.

![Figure 2. Multi-reference frame bidirectional prediction hypothesis optimization algorithm framework.](image)

At the encoding end, the video sequence is divided into multiple GOPs, the first frame of each GOP is a key frame, and the remaining frames are CS frames. Two-dimensional observation is used for key frames, and block measurement is performed for CS frames. On the decoding side, the reconstruction steps of the image frame are as follows:

- Reconstruct the key frame based on fractional order total variation, then use the secondary reconstruction algorithm of the key frame to complete the second reconstruction of the current GOP and the next GOP key frame, introduce the concept of long-term reference frame, and put the two reconstruction keys The frame is used as a long-term reference frame for CS frame reconstruction.

- Reconstruct multiple hypothesis residuals based on Tikhonov regularization for all CS frames in the current GOP. The GOP is divided into the first half GOP and the second half GOP. If the GOP length is 8, the second to fourth frames are the first half GOP, the sixth to eighth frames are the second half GOP, and the fifth frame is the middle frame. There are up to 4 reference frames, and the hypothesis set is generated using an improved search window-based multi-hypothesis prediction method. The concept of short-term reference frames is introduced. The reconstructed reference frame of the second frame is the long-term reference frame. Frames 1 and 9 (the key to the next GOP Frame), after completing the reconstruction of the second frame, use it as a short-term reference frame together with two long-term reference frames as the reference frame of the third frame to assist reconstruction. The CS frame reconstruction process of the second half GOP is the same as that of the first half GOP, but the direction is opposite.

- Select the reconstructed CS frame before and after the intermediate frame as the short-term reference frame and the two long-term reference frames as the reference frame to assist the intermediate frame reconstruction. So far, the decoding of one GOP is completed.

The schematic diagram of CS frame reconstruction based on the multi-reference frame bidirectional prediction hypothesis optimization algorithm is shown in figure 3.
Short-term reference frame:

Figure 3. Schematic diagram of CS frame reconstruction.

As the number of reference frames for assisting CS frame reconstruction increases, the size of the initially generated candidate hypothesis set increases accordingly. To keep the hypothesis set size the same as two reference frames were selected as before, need to remove the hypothesis with low contribution rate. The calculation complexity of the Euclidean distance \( \| y - \Phi h \|_2 \) for determining the contribution rate of the hypothesis vector \( h \) is relatively high. Therefore, the contribution rate is determined by calculating the correlation between the hypothesis set of the block to be reconstructed and the projection set in the measurement domain. The closer the correlation value is to 1, the higher the hypothesis contribution rate.

Therefore, in order to obtain a higher-quality multi-hypothesis set, thereby improving the reconstruction quality of non-key frames, a multi-reference frame hypothesis set optimization algorithm is adopted, which uses adjacent key frames and non-key frames to improve without increasing the number of hypothesis sets. The average contribution value of the hypothesis set to the prediction is used to improve the reconstruction quality of the video sequence.

3.2. Secondary reconstruction of key frames

In traditional DCVS, key frames are decoded only once independently. Although the sampling rate of key frames is high, the quality of key frame reconstruction is not improved with the increase of the sampling rate. In order to improve the quality of key frame reconstruction, so that the reconstructed key frame can provide higher quality reference frames for other CS frames, an improved key frame secondary reconstruction algorithm is proposed. Consider that the CS frame adjacent to the key frame has a higher correlation with the key frame. However, due to the low sampling rate of the CS frame, it could not suitable for direct use as a reference frame. So the two adjacent CS frames as the short-term reference frames, and the two key frames of two adjacent GOPs (the previous GOP and the next GOP) are selected as long-term reference frames for multi-hypothesis reconstruction. The schematic diagram of key frame secondary reconstruction is shown in figure 4.
Provide References

Figure 4. Schematic diagram of the key frame secondary reconstruction algorithm.

At this time, both the two long-term reference frames and the two short-term reference frames can provide high-quality hypotheses. Finally, using the multi-reference frame bidirectional prediction hypothesis set optimization algorithm to optimize the hypothesis set of the secondary reconstruction key frame. Due to the hypothesis with high correlation is selected, the quality of the key frame is improved.

4. Simulation experiment and result analysis
This section verifies the performance of the proposed algorithm in terms of reconstruction quality and computational complexity. Using PSNR as the evaluation standard to measure the reconstruction quality of different schemes. To be fair, each algorithm only runs separately to prevent interference. Video sequences choose standard video sequences in CIF format (352×288), which are download from http://trace.eas.asu.edu/yuv/. The standard test video sequences include Mother-Daughter sequence, Foreman sequence and Soccer sequence. Among them, Mother-Daughter sequence moves slowly, Foreman sequence moves moderately, and Soccer sequence moves violently. Uniformly select the first 88 frames of each sequence. Set the GOP length to 8, the key frame sampling rate is fixed at 0.7, the CS frame sampling rate is 0.1 to 0.5, the block size $B = 16$, the search window radius $W = 15$, and the measurement matrix is a Gaussian orthogonalization matrix. The computer configuration of the simulation experiment is: 64-bit Windows 7 operating system, Intel(R) Core(TM) i7–4790 CPU, 3.60GHz, 32G RAM. The version of Matlab is R2017a.

The performance comparison between the MH-BCS-SPL algorithm and the MC-BCS-SPL algorithm and the proposed multi-hypothesis distributed compressed video perception reconstruction algorithm based on key frame secondary reconstruction is proposed in this paper. First, compare the quality of key frame reconstruction. Since the algorithm proposed in this paper has the CS frame as the reference frame in the secondary reconstruction stage of the key frame, it is more representative to set the CS frame sampling rate to 0.1. The results are shown in table 1.
Table 1. Foreman sequence key frame reconstruction PSNR comparison under three algorithms (sampling rate 0.1).

| Key Frame Number | MH-BCS-SPL, PSNR/dB | Algorithm of [16], PSNR/dB | Proposal, PSNR/dB |
|------------------|----------------------|---------------------------|------------------|
| 1                | 42.1642              | 42.2084                   | 43.2629          |
| 9                | 40.8475              | 41.4219                   | 42.4967          |
| 17               | 40.9022              | 42.1042                   | 43.0297          |
| 25               | 40.9643              | 42.5216                   | 43.4172          |
| 33               | 41.2276              | 42.5873                   | 43.6218          |
| 41               | 41.0461              | 42.2815                   | 43.1421          |
| 49               | 41.1272              | 42.1767                   | 43.3273          |
| 57               | 41.3955              | 42.4389                   | 43.5082          |
| 65               | 40.5727              | 41.2029                   | 42.1211          |
| 73               | 41.0883              | 42.3212                   | 43.4487          |
| 81               | 40.4599              | 41.7264                   | 42.1233          |
| Average          | 41.0723              | 42.0901                   | 43.0454          |

It can be seen from Table 2 that despite the extremely low sampling rate of CS frames, the quality of the secondary reconstruction of key frames can be steadily improved by using the key frame secondary reconstruction algorithm proposed in this paper, because the key frames are fixed as each CS frame. The improvement of the reconstruction quality of the long-term reference frame improves the quality of the multi-hypothesis prediction vectors provided to the CS frame, thereby improving the reconstruction quality of the CS frame. Reference [16] uses the initial reconstructed key frame as the reference frame for the secondary reconstruction of the key frame, and then uses MH-BCS-SPL to perform the secondary reconstruction of the key frame, which has a limited effect on improving the quality of the key frame reconstruction. As in the 57th frame, the PSNR of the proposed algorithm after the second reconstruction of the key frame is 2.18dB higher than that of the MH-BCS-SPL, and 1.07dB higher than the algorithm in [16]. Finally, the average PSNR of the second reconstruction of the key frame of the proposed algorithm is improved by 1.97dB and 0.9553dB respectively compared with the algorithm of MH-BCS-SPL and the reference [16].

Figures 5 to 7 show the performance curves of the proposed algorithm and MH-BCS-SPL and MC-BCS-SPL.
Figure 5. Mother-Daughter sequence average performance comparison curve.

Figure 6. Foreman sequence average performance comparison curve.

Figure 7. Soccer sequence average performance comparison curve.

In the figure, the abscissa is the sampling rate of the CS frame, and the ordinate is the average peak signal-to-noise ratio of the reconstructed image. As can be seen from figure 5, compared with the algorithm, the reconstruction performance of the proposed algorithm is improved by about 1dB compared with the MH-BCS-SPL, and the video reconstruction effect of the proposed algorithm increases monotonically with the increase of the sampling rate. The reconstruction quality is robust to the sampling rate. In figures 6 and 7, the reconstruction performance of the algorithm in this paper is improved by about 1.5 dB compared to the MH-BCS-SPL algorithm. The reason for the greater improvement in reconstruction performance compared to the algorithm in figure 5 is that the mother-daughter sequence has a slower motion and the reference frame. It does not bring a large number of high-quality assumptions, so the quality of reconstruction is improved slightly. Foreman sequence and Soccer sequence move violently. The algorithm in this paper can introduce more high-quality hypotheses to replace the inferior hypotheses in the original hypothesis set by increasing the number of reference frames.

It is worth noting that the reconstruction quality of the algorithm in this paper is also affected by the GOP size. The larger the GOP, the lower the correlation between the key frame and the CS frame, and the more hypothetical scheme of the proposed algorithm improves the reconstruction quality. When the GOP size is smaller, the number of CS frames becomes smaller. Although the correlation between Key frames and CS frames increases, the proposed multi-reference frame scheme improves the reconstruction quality.
5. Conclusions
This paper proposes a multi-distributed video compressed sensing scheme based on secondary reconstruction of key frames. The algorithm uses key frames with higher sampling rates as the long-term reference frames and the reconstructed frame of CS frame with good inter-frame correlation as the short-term reference frame. Using these 4 reference frames to assist CS frame reconstruction. The number of reference frames increasing and the selection range of hypothesis vectors are expanded while not changing the size of the hypothesis vector set, and the overall performance of multi-hypothesis reconstruction is improved. Under the condition of set size, the overall performance of multi-hypothesis reconstruction is improved. The algorithm also uses multi-reference frame reconstruction for the secondary reconstruction of key frames, which improves the reconstruction quality of CS frames while improving the reconstruction quality of key frames. Through experimental verification, this paper proposes that the quality of video reconstruction with severe motion is improved more. Under the same GOP length, the reconstruction effect of this algorithm is significantly better than the MH-BCS-SPL and MC-BCS-SPL algorithms, and has the robustness and effectiveness.

References
[1] Donoho D L. Compressed sensing. IEEE Transactions on Information Theory, 2006, 52(4): 1289–1306.
[2] LIU Y and PADOS D A. Compressed-sensed-domain L1-PCA video surveillance [J]. IEEE Transactions on Multimedia, 2016, 18(3): 351-363. doi: 10.1109/TMM.2016.2514848.
[3] GUO J, SONG B, and DU X. Significance evaluation of video data over media cloud based on compressed sensing[J]. IEEE Transactions on Multimedia, 2016, 18(7): 1297-1304. doi: 10.1109/TMM.2016.2564100.
[4] REHMAN A U, SHAH G A, and TAHIR M. Compressed sensing based adaptive video coding for resource constrained devices[C]. IEEE International Wireless Communications and Mobile Computing Conference, Paphos, Cyprus, 2016: 170-175.
[5] WANG J, GUPTA M, and SANKARANARAYANAN A C. LiSens A scalable architecture for video compressive sensing[C]. IEEE International Conference on Computational Photography, Houston, TX, 2015: 1-9.
[6] DO T T, CHEN Y, NGUYEN D T, et al. Distributed compressed video sensing[C]// Image Processing (ICIP), 2009 16th IEEE International Conference. 2009: 1393-1396.
[7] Kang LW, Lu CS (2009) Distributed compressive video sensing. In: Proceedings of the international conference on acoustics, speech, and signal processing, pp 1169-1172.
[8] CHEN C, TRAMEL E W, FOWLER J E. Compressed-sensing recovery of images and video using multi-hypothesis predictions[C]//The 45th Asilomar Conference. 2011: 1193-1198.
[9] FOWLER J E, MUN S, TRAMEL E W. Block-based compressed sensing of images and video[J]. Foundations and Trends in Signal Processing, 2012, 4(4): 297-416.
[10] Li Ran, Gan Zongliang, Cui Ziguan, Wu Minghu, Zhu Xiuchang. Distributed Adaptive
Compressed Video Sensing Using Smoothed Projected Landweber Reconstruction [J]. China Communications, 2013, 10(11): 58-69.

[11] KUO Y H, WU K, CHEN J. A scheme for distributed compressed video sensing based on hypothesis set optimization techniques [J]. Multidimensional Systems and Signal Processing, 2017, 28(1): 129-148.

[12] CHEN J, CHEN Y Z, QIN D, et al. H. An elastic net-based hybrid hypothesis method for compressed video sensing [J]. Multimedia Tools and Applications, 2015, 74(6): 2085-2108.

[13] KUO Y H, WU K, CHEN J. A scheme for distributed compressed video sensing based on hypothesis set optimization techniques [J]. Multidimensional Systems and Signal Processing, 2017, 28(1): 129-148.

[14] CHEN J, SU K X, YANG X Z, et al. Reconstruction algorithm for block compressed sensing based on variation model [J]. Journal on Communications, 2016, 37(1): 100-109.

[15] Kuo Y, Wang S, Qin D, Chen J (2013) High-quality decoding method based on resampling and rereconstruction. Electron Lett, 49(16):991-992.

[16] Chen J, Xue F, Kuo Y H. Distributed compressed video sensing based on key frame secondary reconstruction [J]. Multimed Tools Appl. doi: 10.1007/s11042-017-5071-5.

[17] Kuo Y, Wu K, Chen J (2015) A scheme for distributed compressed video sensing based on hypotheses set optimization techniques [J]. Multidimensional Systems & Signal Processing, pp 1-20.

[18] Chen J, Wang N, Xue F, Gao YT (2016) Distributed compressed video sensing based on the optimization of hypotheses set update technique. Multimed Tools Appl. doi:10.1007/s11042-016-3866-4.

[19] Chen G(2016). Research on Video/Image Compressive Sensing Reconstruction Based on Fractional-Order Total Variation[D]. Southwest Jiaotong University, 2017.

[20] MUN S, FOWLER J E. Block compressed sensing of images using directional transforms[C]//Image Processing (ICIP), 2009 16th IEEE International Conference. IEEE, Cairo, Egypt, 2009: 3021-3024.

[21] Tramel EW, Fowler JE (2011) Video compressed sensing with multi-hypothesis. Proc Data Compress Conf, pp 193–202.

[22] Yang Chunling, Ou Weifeng. Multi-reference Frame Based Optimal Multi- Hypothesis Prediction Algorithm for Compressed Video Sensing[J]. Journal of South China University of Technology (Natural Science Edition), 2016.44(1): 1-8.