

# Compressive Video Sensing Based on Intra-Inter-Frame Constraints and Genetic Algorithm

Yuchen YUE<sup>a</sup>, Hua LI<sup>a,1</sup> and Jianhua LUO<sup>b</sup>

<sup>a</sup>*Department of Armament and Control, Army Academy of Armored Forces, Beijing, China*

<sup>b</sup>*Center of Maneuver and Training, Army Academy of Armored Forces, Beijing, China*

**Abstract.** Establishing structured reconstruction models and efficient reconstruction algorithms according to practical engineering needs is of great concern in the applied research of Compressed Sensing (CS) theory. Targeting problems during high-speed video capture, the paper proposes a set of video CS scheme based on intra-frame and inter-frame constraints and Genetic Algorithm (GA). Firstly, it employs the intra-frame and inter-frame correlation of the video signals as the priori information, creating a video CS reconstruction model on the basis of temporal and spatial similarity constraints. Then it utilizes overcomplete dictionary of Ridgelet to divide the video frames into three structures, smooth, single-oriented, or multijointed. Video frames cluster according to the structure using Affinity Propagation (AP) algorithm, and finally clusters are reconstructed using evolutionary algorithm. It is proved efficient in terms of reconstruction result in the experiment.

**Keywords.** Compressed sensing, genetic algorithm, AP cluster, Ridgelet, overcomplete dictionary

## 1. Introduction

Compressed sensing theory [1-2] is a new information acquisition theory proposed by D. Donoho, E. Candès, and Chinese scientists T. Tao, which states that as long as a signal is compressible or sparse in some domain of mutation, it can be projected into a low-dimensional space by a measurement matrix, and the measurement data can be reconstructed by solving an optimal model. The measurement data can be reconstructed into the original signal by solving an optimization model. Currently, compressed perception is widely used in the field of video processing [3], and the existing video compressed perception reconstruction algorithms [4-7] are mainly divided into two categories: frame-by-frame reconstruction, i.e., distributed video compressive sensing algorithm, divides the video sequence into group of pictures (GOP), in which the first frame of each GOP is the Key frame and the rest are the CS frames. Key frames are sampled with high sampling rate to obtain better reconstruction quality through intra-frame reconstruction, while CS frames are sampled with lower sampling rate and

---

<sup>1</sup>Corresponding author: Hua Li, Department of Armament and control, Army Academy of armored Forces, Feng tai, Beijing; E-mail: Gonzo123@sina.com.

reconstructed with the aid of Side Information generated from Key frames through motion compensation to obtain high quality reconstructed video, but the disadvantage of this method is that when there are frames with abrupt scene changes in the video, the fixed-length GOP does not produce better results when there are frames with abrupt scene changes in the video, in addition, the algorithm does not make use of the structural prior information of the video. The overall reconstruction class algorithm does not distinguish between Key frames and CS frames at the coding end, but uses different methods for different types of measurements for reconstruction. The reference [8] points out that the structural a priori information of the signal plays a key role in reducing the number of measurements required for accurate signal reconstruction and reducing the uncertainty of the reconstruction model, and for high-speed videos with drastic scene changes, to ensure the efficiency of video acquisition, a structured compressed sensing framework to ensure the efficiency of video acquisition, the sparse representation of the video signal is achieved by constructing a structured perfect dictionary, and a structured measurement scheme suitable for it is formed by using the structural prior information of the video signal, and the reconstruction of high-speed video is achieved based on the structural prior.

In this paper, a genetic algorithm video compression-aware reconstruction algorithm based on video frame motion classification is proposed using the characteristic of relatively single block structure information [9-10]. The algorithm first divides the video sequence into a Data Group (DG) consisting of several consecutive video frames. Then divides all video frames into blocks, and determined the blocks as change class blocks and non-change class blocks based on the judgment of whether the blocks change within adjacent video frames, and then, constructs a Ridgelet over-complete dictionary to determine the block structure type of all video frames in a single DG, retaining only the non-change class block measurements in the previous frame of two adjacent video frames, and retaining all change class block measurements in two adjacent frames. Next, all the retained block measurements are treated as a Data Unit (DU), based on which the DU is clustered using the AP clustering algorithm to make full use of the intra-frame correlation and inter-frame correlation of the video sequence, and the clusters are reconstructed using the genetic evolution algorithm. Finally, the video reconstruction method is trained by video sequence samples, and the reconstruction performance of the proposed algorithm is further improved by training.

## **2. Block structure classification based on Ridgelet overcomplete dictionary**

Video frames have self-similarity. After the non-overlapping blocking operation of a single video frame using the blocking strategy, the obtained blocks contain only a limited number of different structural classes, and the measurement vectors of structurally similar blocks also have similarity. Therefore, the similarity of the measurement vectors of the blocks can be used for metrics and classification can be achieved based on the structural properties of the blocks. Each class of blocks can be represented by the same set of atoms in the overcomplete dictionary. This imposes two requirements on the overcomplete dictionary. First, the overcomplete dictionary should have a redundant orientation structure in order to be able to adaptively and sparsely represent texture and edge information in an image with arbitrary orientation. Second, the overcomplete dictionary should have an efficient search and optimization algorithm that can find the

appropriate combination of atoms and corresponding combination coefficients for the sparse representation of the signal to be reconstructed in the shortest possible time.

The overcomplete dictionary  $\mathbf{D} \in \mathbb{R}^{B \times N}$  used to represent the block is noted as  $\mathbf{D} = (\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_N)$ ,  $i = 1, 2, \dots, N$ , where  $\mathbf{d}_i$  is the atom in dictionary  $\mathbf{D}$  with ordinal number  $i$ , generated in the manner of equation (1).

$$\mathbf{d}_i(\mathbf{z}) = \frac{1}{W} \left[ e^{-(a_i \mathbf{u}_i^T \mathbf{z} - b_i)^2 / 2} - \frac{1}{2} e^{-(a_i \mathbf{u}_i^T \mathbf{z} - b_i)^2 / 8} \right] \quad (1)$$

Where  $\mathbf{d}_i(\mathbf{z}) \in \mathbb{R}^{\sqrt{B} \times \sqrt{B}}$  is the atom with the same size as the block, and  $\mathbf{d}_i \in \mathbb{R}^B$  is the vectorized atom.  $\mathbf{z} = (\mathbf{z}_1, \mathbf{z}_2) \in [0, 1, 2, \dots, \sqrt{B} - 1]^2$  is the position vector of the atom, and  $1/W$  is the normalization factor of the atom. The atom  $\mathbf{d}_i(\mathbf{z})$  in the dictionary  $\mathbf{D}$  corresponds to the parameter set  $\gamma_i = (\theta_i, a_i, b_i)$ , where  $a_i$  is the scale parameter,  $b_i$  is the displacement parameter,  $\theta_i$  is the orientation parameter,  $\mathbf{u}_i = (\cos \theta_i, \sin \theta_i)^T$ .

After the prototype atoms are selected, the atomic size of the dictionary and the sparse representation capability of the video frames also depend on the range of values of  $a_i$ ,  $b_i$  and  $\theta_i$ , their respective discrete intervals. Following the Ridgelet dictionary construction method of the reference [11], the parameter space is set as equation (2).

$$\Gamma_\gamma = \{(\theta, a, b) \mid \theta \in [0, \pi), a \in [0, 3], b \in \Gamma_b\} \quad (2)$$

Where the displacement parameter  $b_i$  is taken in relation to the orientation parameter  $\theta_i$ , as shown in equation (3).

$$\Gamma_b = \begin{cases} [0, \sqrt{B} (\sin \theta + \cos \theta)], & \text{if } \theta \in [0, \pi/2) \\ [\sqrt{B} \cos \theta, \sqrt{B} \sin \theta], & \text{otherwise} \end{cases} \quad (3)$$

When using an overcomplete dictionary for the sparse representation of video frames, the individual atoms in the dictionary respond only to those blocks that are consistent in shape, scale, position and orientation information. Meanwhile, the subsequent algorithm wants to determine the block structure characteristics and classify them based on the characteristics of the sub-dictionaries. Therefore, the discrete interval of the orientation parameters must be small enough when discretizing the Ridgelet overcomplete dictionary.

Since two block sizes,  $8 \times 8$  and  $16 \times 16$ , are used in the experiment, two Ridgelet overcomplete dictionaries are constructed. Setting  $a \in [0, 3]$ ,  $\theta \in [0, 179]$ , scale interval and displacement interval to 0.2 and 1, respectively, and directional interval to 5. Therefore, the constructed Ridgelet overcomplete dictionary for  $8 \times 8$  blocks have 6271 atoms, and the Ridgelet overcomplete dictionary for  $16 \times 16$  blocks have 12032 atoms. Ridgelet overcomplete dictionary for  $16 \times 16$  blocks have a total of 12032 atoms.

Smoothed blocks are insensitive to the orientation parameter, so the atoms of each direction have good reconstruction quality for this type of blocks, and only the sub-

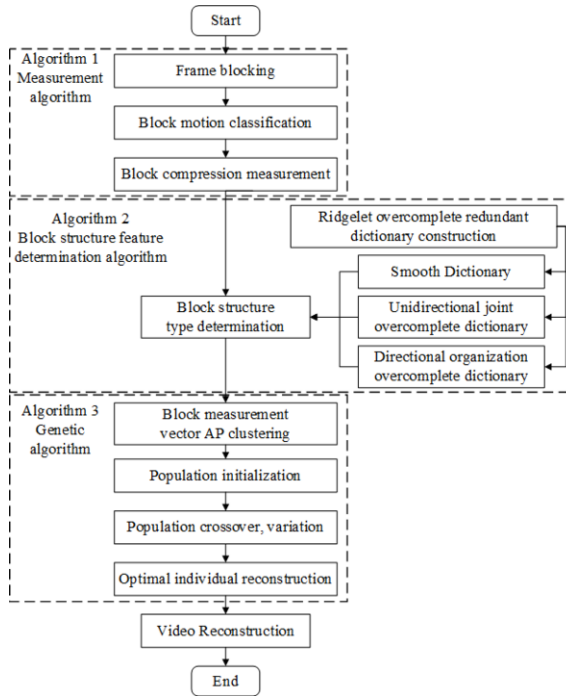
dictionaries of individual directions have high reconstruction quality for single-direction blocks and multi-direction blocks. The process of reconstructing blocks undoubtedly increases the difficulty of block structure feature determination, so the analytical formula-based method proposed in reference [12]. Firstly, we determine whether the block is smooth by the relationship between the variance of the block measurements and the mean variance of the block, and then we find the 10 most relevant atoms in the directional sub-dictionary over the complete dictionary along the same lines and solve the block measurement residuals by the least squares' method. If the block has the largest measurement value residual on a single direction sub-dictionary, the corresponding block is marked as a single-direction block, and vice versa as a multi-direction block. To ensure the fast convergence of the algorithm and to facilitate the optimal direction statistics of multi-directional class blocks, only the three directions with the smallest residuals are marked for all blocks marked as multi-directional blocks with the number of directions greater than or equal to 3. This method is inaccurate in the determination of block structure features at low sampling rate, but the purpose of adopting this method in this paper is block structure feature classification, not exact reconstruction, so it can be used for determination.

Smoothed blocks are insensitive to the orientation parameter, and the atoms of each direction have good reconstruction quality for this type of blocks. Only the sub-dictionaries of individual directions have high reconstruction quality for single-direction blocks and multi-direction blocks. The process of reconstructing blocks undoubtedly increases the difficulty of block structure feature determination, so the analytical formula-based method proposed in the reference [12]. Firstly, we determine whether the block is smooth by the relationship between the variance of the block measurements and the mean variance of the block, and then we find the 10 most relevant atoms in the directional sub-dictionary over the complete dictionary along the same lines and solve the block measurement residuals by the least square's method. If the block has the largest measurement value residual on a single direction sub-dictionary, the corresponding block is marked as a single-direction block, and vice versa as a multi-direction block. To ensure the fast convergence of the algorithm and to facilitate the optimal direction statistics of multi-directional class blocks, only the three directions with the smallest residuals are marked for all blocks marked as multi-directional blocks with the number of directions more than or equal to 3. Even though this method is inaccurate in the determination of block structure features at low sampling rate, but the purpose of adopting this method in this paper is block structure feature classification, not exact reconstruction, so it can be used for determination.

### **3. Genetic Algorithm-based Compressive Video Sensing Algorithm**

#### *3.1. The overall framework of the algorithm*

The flow of the genetic algorithm-based compressive video sensing algorithm is shown in Figure 1.



**Figure 1.** Genetic algorithm-based compressive video sensing algorithm flow block diagram.

The proposed genetic algorithm-based compressive video sensing algorithm consists of three parts: measurement algorithm, block structure type determination algorithm and genetic algorithm. Firstly, a blocking operation is implemented for each frame of the input video and a plurality of consecutive video frames are set as one DG, and all the blocks in a single DG are compressed and sampled to obtain measurement values. The blocks at the same position in two adjacent video frames are classified by motion through calculation and divided into change blocks and non-change blocks, and a single non-change block and all change blocks are retained. Then the measurements of all the retained blocks in DG are treated as a DU and subjected to AP clustering to maximize the intra- and inter-frame correlation of the video sequences, and the reconstruction is performed uniformly for the same clusters. The reconstruction algorithm uses Genetic Evolutionary Algorithm (GA).

### 3.2. Measurement Algorithm

A new video sampling scheme is proposed to make the proposed algorithm have better reconstruction quality in the changing part of the video. After the video frames are divided into multiple  $B \times B$ -sized blocks, a Data Unit (DU) is set with  $n$  consecutive frames, and the  $l_2$ -norm of the difference between the blocks at the same position in two adjacent frames in the DU is calculated, and a threshold  $\sigma$  is set. The blocks with the difference  $l_2$ -norm value of the blocks at the same position in two adjacent frames less than the threshold  $\sigma$  are marked as non-changing blocks, and then the blocks at the same position in the previous video frame in two adjacent frames are kept and recorded; the blocks with the difference  $l_2$ -norm value greater than the threshold  $\sigma$  are marked as changing blocks. The blocks in the same position in two adjacent frames are marked as

non-changing blocks, and then the blocks in the same position in the previous two frames are kept and recorded; the blocks with the difference  $l_2$ -norm value greater than the threshold  $\sigma$  are marked as changing blocks, and all the changing blocks in the DU are recorded and kept.

### 3.3. Population Initialization Algorithm

After obtaining the measurement vectors of all the reserved blocks in the DU, the block structure classification is determined, which lays a solid foundation for the population initialization in the subsequent evolutionary algorithm solving. Different types of blocks adopt different population initialization strategies. The block structure category judgment method adopts the judgment method proposed in section 2:

#### ① Smooth blocks

Dividing the Ridgelet over-complete dictionary into 16 sub-dictionaries according to the scale, and selecting the first 7 scale to form the sub-dictionary of the smoothing block, and using this to randomly generate the initial population corresponding to the smoothing block.

#### ② Single directional blocks

Divide the Ridgelet overcomplete dictionary into 36 sub-dictionaries according to the direction, select the forward 2 directions and backward 2 directions including the dominant direction of the corresponding single direction block, a total of 5 directions to form the single direction sub-dictionary, and use this to randomly generate the initial population corresponding to the single direction block.

#### ③ Multi-directional blocks

In the process of judging multi-directional blocks, all the most relevant directions in the measurement vector of the block are first counted, and then the most dominant 3 directions are selected, and each direction is selected in the way of single-directional blocks in the forward 2 directions and the backward 2 directions, a total of 15 directions constitute the sub-dictionary of multi-directional blocks, and the initial populations corresponding to multi-directional blocks are randomly generated.

### 3.4. Reconstruction based on Genetic Algorithm

The reconstruction algorithm uses Genetic Algorithm (GA), and GA needs to consider the following five issues in the design phase.

#### ① Individual Codes

An  $N \times N$ -dimensional Ridgelet overcomplete dictionary is constructed, and a real number encoding is adopted to ensure that the atoms of this Ridgelet overcomplete dictionary represent the signal  $\mathbf{x}$  as optimally as possible, while each atom in the dictionary is numbered. Let the individual population be  $P = [\rho_1, \rho_2, \dots, \rho_K]$ , each gene position  $\rho_i$  ( $1 \leq i \leq N$ ) in  $P$  is the number of the over-complete dictionary atoms, and  $K$  is the signal sparsity.

#### ② Adaptation function

The fitness function is an index to search for feasible solutions, and its mathematical expression is shown in equation (4).

$$f(Y_m) = \frac{1}{\sum_{i=1}^n \|\mathbf{y}_i - \Phi \text{dec}(Y_m) \alpha_i\|_2^2} \quad (4)$$

where  $f(Y_m)$  is the fitness of the same clusters,  $\mathbf{y}_i$  is the measurement vector of the signal,  $Y_m$  is the set of measurement vectors of the same clusters,  $\text{dec}(Y_m)$  is the optimal set of atoms of  $Y_m$ ,  $n$  is the capacity of  $Y_m$ ,  $i$  is the number of individuals within  $Y_m$ , and  $\Phi$  is the measurement matrix.

Because of the property that there are a large number of inter- and intra-frame similar blocks in the video. Using the AP clustering is used to cluster the measurement vectors of the blocks separately by the structural characteristics of the blocks  $Y_m$ , and the inverse of the residual sum of the block measurements and the clustered measurements is used as the fitness function. The sum of the fitness functions is found to take advantage of the constraints between similarly structured measurement data, while effectively avoiding possible deviations in the measurement vector search process, resulting in a blockier and more accurate search for the optimal combination atoms of  $Y_m$ .

### ③Decoding process

The GA uses a non-convex  $l_0$ -model for compressed sensory reconstruction. First, search for the  $\text{dec}(Y_m)$  that maximizes the  $f(Y_m)$  in the reconstruction process. Then, solves the sparse representation coefficient  $\alpha$  for each measurement corresponding to the dictionary atom using a least-squares algorithm. The mathematical expression is shown in equation (5).

$$\alpha = (\Phi \text{dec}(Y_m))^+ \mathbf{y} \quad (5)$$

Finally, using Ridgelet overcomplete dictionary  $\mathbf{D}$  to reconstruct the original signal. The mathematical expression is shown in equation (6).

$$\hat{\mathbf{x}} = \mathbf{D}\alpha \quad (6)$$

### ④Population initialization

According to Table 1 to initialization the population.

**Table 1.** Population initialization algorithm

<b>Step 1 :</b>	Determining the structural category of video blocks, marking smooth blocks and single/multi-directional video blocks respectively, with different population initialization generation methods for different categories of video blocks.
<b>Step 2 :</b>	The Ridgelet overcomplete dictionary is divided by scale, and the first seven scales are selected to form a smooth block sub-dictionary, from which the corresponding initial populations are randomly generated.
<b>Step 3 :</b>	The Ridgelet overcomplete dictionary is divided by direction, and after marking the dominant direction $i$ of the unidirectional block, a total of five directions $\{i-2, i-1, i, i+1, i+2\}$ in the forward and backward adjacent directions are selected to form the unidirectional block sub-dictionary, and the population is initialized.
<b>Step 4 :</b>	Count the most dominant three directions $\{j, k, l\}$ among the measurement vectors marked as multi-directional blocks, and record only the most dominant three when there are more than 3. Each dominant direction ( $j$ ) is selected as a single-directional block in the way of 5 directions $\{j-2, j-1, j, j+1, j+2\}$ respectively to form a single-directional block sub-dictionary and then combined into a multi-directional block sub-dictionary with a total of 15 directions.

---

Generating individuals for all directions  $(j,k,l)$  according to the above operation, and finally combining all individuals to become the initialized population of the multi-directional block class.

---

⑤ Genetic operators

There are three types of genetic operators, crossover, mutation and selection, and the specific operators are designed as follows.

- Crossover operators

A single-point crossover operation is used, in which a gene locus is randomly selected in two randomly paired individuals, and the lower half of the two individuals are swapped at the midpoint of the gene locus to create two new individuals. Crossover occurs randomly according to a set crossover probability  $P_c$ .

- Mutation operator

Using mutation operator based on directional statistics, mutation occurs randomly according to a set probability of mutation  $P_m$ , all genetic positions in individuals smaller than  $P_m$  are mutated, and the mutation value used to replace the original value of the genetic position comes from the sub-dictionary of individuals with the least directional genetic position.

- Selection operator

The fitness of the two new individuals generated from the crossover of two paired parent individuals is solved by the fitness function respectively, and if the fitness of one of the two parent individuals is smaller than that of the individual with larger fitness in the new individual, this parent individual is removed from the parent population and the new individual is added, otherwise, no operation is performed.

Finally, the individual with the largest fitness within the population is computationally selected as the optimal atomic group for the reconstructed signal measurement vector clustering, and the inverse least squares operation is used to complete the signal reconstruction.

The pseudo-code of the genetic algorithm-based compressive video sensing reconstruction algorithm is shown in Table 2.

**Table 2.** Genetic algorithm-based compressive video sensing reconstruction algorithm

---

<b>Input:</b>	video Sequence $\mathbf{X}$ , random Gaussian measurement matrix $\Phi$ , number of iterations, Ridgelet overcomplete dictionary $\mathbf{D}$
	Measurement of $\mathbf{X}$ ;
	Determining the blocks as smooth, single directional, multi-directional;
	Extracting structurally identical blocks and AP clustering;
	Set the population size and sparsity. The population initialization is performed using the corresponding initialization scheme according to the clustering of the block measurement data;
<b>Iteration:</b>	Crossover and mutation operations on individuals in the population;
	Selecting the best individuals in the population for reconstruction $\hat{\mathbf{x}}$ :
	$\alpha = (\Phi \text{dec}(\mathbf{x}_m))^+ \mathbf{y}$ ;
	$\hat{\mathbf{x}} = \mathbf{D}\alpha$ ;
	Reconstruct the video sequence $\hat{\mathbf{x}}$ by recombining the estimated signal $\hat{\mathbf{X}}$ according to the chunking and coding strategy;
<b>Output:</b>	Reconstructed video $\hat{\mathbf{X}}$

---



#### 4. Performance Analysis of Proposed Algorithm

In this section, several experiments are designed to verify the performance of the proposed evolutionary algorithm-based video compression perception algorithm and to optimize the parameter settings involved in the algorithm. The basic configuration of the experiments includes: the experimental platform is a laptop-type computer with Inter Core i7 2.8GHz CPU, 16GB RAM, and 64-bit Win7 operating system installed. All experiments were conducted in MATLAB R2017a.

##### 4.1. Analysis of the impact of sampling rate on algorithm performance

The following experiments analyze the performance of the proposed algorithm at different sampling rates.

The experimental object is a 96-frame Traffic video sequence of  $256 \times 256$ , with video frame chunk size  $B = 8$ . Construct an overcomplete dictionary based on Ridgelet, and set the direction range  $[0, 179)$ , direction interval as 5, total 36 directions, scale range  $[0, 3]$ , scale interval as 0.2, total 16 scales. The dictionary size as 6172. Set the population size of the genetic algorithm is 36, the individual size is 12. Set the crossover probability  $P_c = 0.6$  and the mutation probability  $P_m = 0.01$ . Let 8 consecutive frames are used as a DG, and the number of smooth block iterations is set to 20, the number of single-direction block iterations to 50, and the number of multi-direction block iterations to 100. the sampling rates are set to 0.125, 0.2, and 0.25, respectively, and the algorithm performance evaluation criterion is the average PSNR of the reconstructed frames.

The comparison experiments are reconstructed using the Gaussian mixture model proposed in the reference [13], and to ensure fairness, the algorithm in the reference [13] is trained offline in advance to obtain a set of mean and variance as the initial parameters of the Gaussian mixture model. When the sampling rate is 0.125, the Gaussian mixture model reconstructs the sum of measurement data generated by making  $\{0, 1\}$  measurements with probability 0.5 for 8 consecutive frames; when the sampling rate is 0.2, the Gaussian mixture model reconstructs the sum of measurement data generated by making  $\{0, 1\}$  measurements with probability 0.5 for 5 consecutive frames; when the sampling rate is 0.25, the Gaussian mixture model reconstructs the sum of measurement data generated by making  $\{0, 1\}$  measurements with probability 0.5 for 4 consecutive frames.

The first 8 frames of the Traffic sequence are used as an example for analysis. Figure 2 shows the consecutive images of the original sequence, and Figure 3 shows the enlarged image of the 3rd frame of the first 8 frames.



**Figure 2.** The first 8 original frames of Traffic sequence.

The proposed algorithm and the Gaussian mixture model proposed in the reference [13] are used to reconstruct the Traffic sequences respectively. When the sampling rate is 0.125, the reconstruction results of both for the first 8 frames of Traffic sequence are shown in Figure 4, where Figure 4(a) is the reconstruction result of the algorithm of reference [13], and Figure 4(b) is the reconstruction result of the proposed algorithm; when the sampling rate is 0.2, the reconstruction results of both for the first 8 frames of

Traffic sequence are shown in Figure 5, where Figure 5(a) is the reconstruction result of the algorithm of reference [13], and Figure 5(b) is the reconstruction result of the proposed algorithm. Figure 5(b) is the reconstruction result of the proposed algorithm; when the sampling rate is 0.25, the reconstruction results of both for the first 8 frames of the Traffic sequence are shown in Figure 6, where Figure 6(a) is the reconstruction result of the algorithm of the reference [13] and Figure 6(b) is the reconstruction result of the proposed algorithm.



Figure 3. The 3<sup>rd</sup> original frames of Traffic sequence.



(a) The results of reference [13] algorithm



(b) The results of proposed algorithm

Figure 4. The results of first 8 reconstruction frames when sample rate is 0.125.



(a) The results of reference [13] algorithm



(b) The results of proposed algorithm

Figure 5. The results of first 8 reconstruction frames when sample rate is 0.2.



(a) The results of reference [13] algorithm



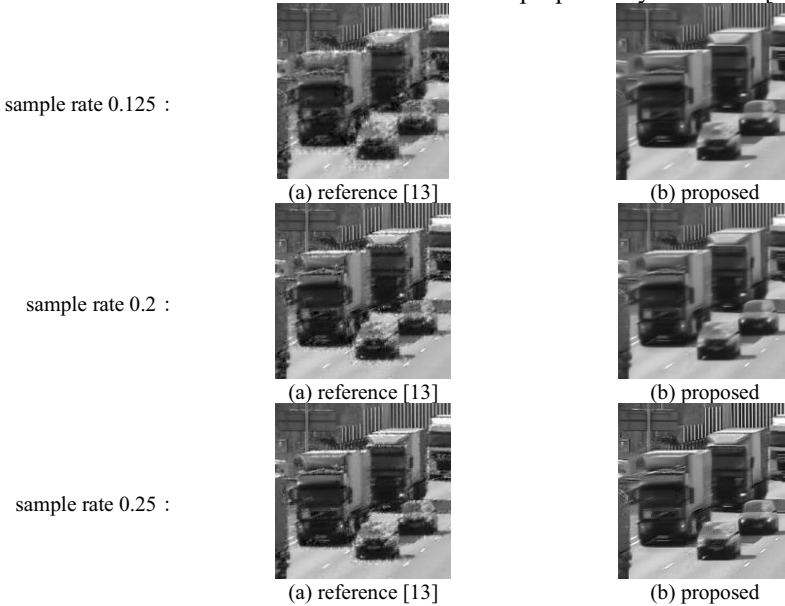
(b) The results of proposed algorithm

Figure 6. The results of first 8 reconstruction frames when sample rate is 0.25.

Figure 7 shows the reconstructed 3<sup>rd</sup> frame zoomed image of Traffic sequence for two algorithms with sampling rates of 0.125, 0.2 and 0.25 respectively. The PSNR of the

reconstructed Traffic sequences for the two algorithms at sampling rates of 0.125, 0.2 and 0.25 are respectively shown in Table 3.

Combining the results from Figure 4 to Figure 6 and Table 3, it can be seen that the average rate distortion performance of the proposed algorithm at all three sampling rates is better than that of the Gaussian mixture model proposed by reference [13].



**Figure 7.** Two algorithms to reconstruct the zoomed image of the 3rd frame of Traffic sequence at three sampling rate.

**Table 3.** The PSNR of two algorithms at sampling rates of 0.125, 0.2 and 0.25

DG Serial	Frame Serial	SR=0.125		SR=0.2		SR=0.25	
		Ref. [13]	Proposed	Ref. [13]	Proposed	Ref. [13]	Proposed
1	1-8	20.80	26.40	24.04	26.65	23.38	28.05
2	9-16	21.21	25.88	23.29	26.40	22.37	27.07
3	17-24	22.27	25.29	22.41	25.48	22.19	27.15
4	25-32	23.02	24.61	21.48	25.68	23.20	26.99
5	33-40	22.76	25.45	21.53	24.75	22.20	26.64
6	41-48	22.03	25.25	21.41	25.19	22.24	26.78
7	49-56	21.33	24.37	20.53	25.32	22.52	27.00
8	57-64	21.58	24.81	20.48	25.89	21.59	26.94
9	65-72	21.19	24.89	21.26	25.60	21.67	26.14
10	73-80	22.18	24.29	21.66	26.58	20.89	25.82
11	81-88	21.62	23.38	20.61	25.57	21.60	26.45
12	89-96	20.55	24.07	21.13	24.54	22.60	26.81
Average		21.71	24.89	21.65	25.64	22.20	26.82

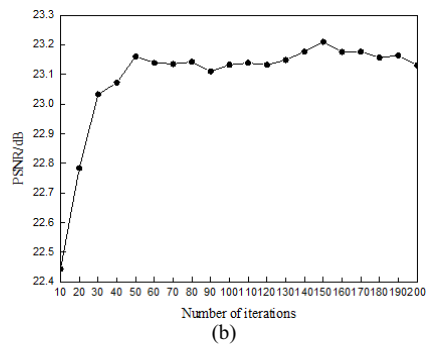
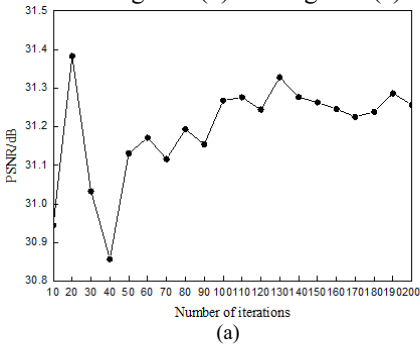
As can be seen from Figure 7, the proposed algorithm not only has a higher PSNR value when the sampling rate is 0.125, but also the reconstructed video frames from the subjective visual point of view have a good reconstruction effect in the edge part has a good reconstruction effect, while the reconstructed image by the Gaussian mixture model of reference [13] shows a large blur in the contour of the change region of the adjacent frames. When the sampling rate is 0.2, it can be seen from Figures 7 that the proposed algorithm demonstrates a better reconstruction detail effect, and the blurring phenomenon in the change region of the reconstructed video frames by the Gaussian

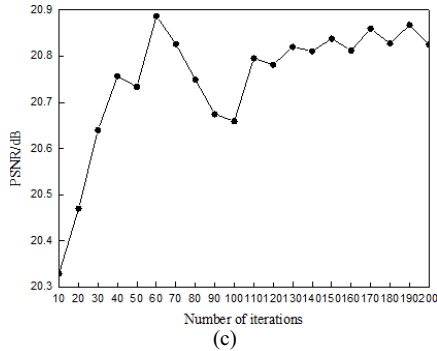
mixture model reconstruction algorithm proposed in the reference [13] is reduced, but the reduction effect is not obvious, and in general the blurring phenomenon in the edge of the change region still exists. When the sampling rate is 0.25, both algorithms improve in image details and visual effects, and the proposed algorithm still has better reconstruction performance.

#### 4.2. Analysis of the Impact of Generation Number on Algorithm Performance

The number of evolutionary iterations set in section 4.1 are: the number of iterations for the smooth block is 20, the number of iterations for the unidirectional block is 50, and the number of iterations for the multi-directional block is 100. Next, the influence of generation number on the performance of the proposed algorithm is analyzed through experiments. In this experiment, the basic experimental parameters are the same as in section 4.1, and the overcomplete Ridgelet dictionary of  $16 \times 16$  from Section 4.1 is used. The dictionary contains 12032 atoms, and the chunk size is  $16 \times 16$ , with 8 frames as a DG. The experimental object is a 96-frame Traffic video sequence of size  $256 \times 256$ .

First of all, there are the experiment of smoothing block evolution generations is performed. The limit number of iterations is set to 200, and the reconstruction data during the evolution process is saved. The average PSNR value of the reconstructed video frames is calculated every 10 iterations as the evaluation criterion of the reconstruction quality. The change of the average rate distortion performance of the smoothing block with the increase of the evolution number is shown in Figure 8(a). The horizontal coordinate in Figure 8 is the number of evolutionary generations, and the vertical coordinate is the average PSNR value of the smoothed block obtained by the proposed algorithm. From the experimental results, the average PSNR value of the reconstructed smoothing block reaches the peak when the evolution proceeds for 20 generations, and then the average PSNR of the reconstructed smoothing block obtains the minimum value when the evolution proceeds for 30 generations, and then the average PSNR value increases continuously with the increase of the evolution generations and eventually tends to level off. Therefore, the number of iterations of 20 is selected as the optimal evolutionary generation of the smoothing block. Similar experiments were also conducted to obtain the relationship curves between the number of iterations and the average PSNR value for the single-directional block and the multi-directional block as shown in Figure 8(b) and Figure 8(c).





**Figure 8.** Number of iterations analysis diagram. (a) smooth blocks (b) single-directional blocks (c) multi-directional blocks.

According to Figure 8(b) and Figure 8(c), it can obtain the optimal evolutionary generations of 50 for the single-direction block and 100 for the multi-direction block. Therefore the values are taken as 20 generations for smooth blocks, 50 generations for single-direction blocks and 100 generations for multi-direction blocks, respectively.

## 5. Conclusion

This paper firstly describes the relevant technique used in the proposed video compressive sensing reconstruction algorithm based on intra-and-frame and evolutionary algorithm - the block structure type determination method based on Ridgelet over-complete dictionary, then describes in detail the implementation process of the proposed algorithm, designs a series of experiments, and analyzes different Then, a series of experiments are designed to analyze the effects of different sampling rates, data unit frames, block sizes and iterations on the performance of the algorithm, and determine the number of block evolution generations, block sizes and data unit frames that make the algorithm perform best.

## References

- [1] Donoho D L. Compressed sensing[J]. IEEE Transactions on Information Theory, 2006, 52(4):1289-1306.
- [2] D L Donoho, Y Tsaig. Extensions of Compressed Sensing [J]. Signal Processing. 2006, 86(3). 533-548.
- [3] G. Yu, G. Sapiro, and S. Mallat, "Solving inverse problems with piecewise linear estimators: From Gaussian mixture models to structured sparsity," IEEE Trans. Image Process., vol. 21, no. 5, pp. 2481-2499, May 2012.
- [4] LIU Y and PADOS D A. Compressed-sensed-domain L1-PCA video surveillance [J]. IEEE Transactions on Multimedia, 2016, 18(3): 351-363.
- [5] 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.
- [6] 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.
- [7] 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.
- [8] Li Wan. Compressed Sensing Observation and Reconstruction Based on Structured Sparse and Convolutional Network [dissertation]. Xi'an: Xidian University, 2018.9.
- [9] M. Wang, A design and research on the compressed sensing measurement matrix, [dissertation] Xidian University, 2014.

- [10] Mun S, Fowler J E. Block compressed sensing of images using directional transforms[C]. International Conference on Image Processing, 2009: 3021-3024.
- [11] Do M N, Vetterli M. The Contourlet transform: an efficient directional multiresolution image representation[J]. IEEE Transactions on Image Processing, 2005, 14(12): 2091-2106.
- [12] Xu Jinghuan. The Design of Ridgelet Redundant Dictionaries and Reconstruction Algorithm[D]. Xi'an: Xidian University, 2011.
- [13] Li X , Lan X , Yang M, et al. A new compressive sensing video coding framework based on Gaussian mixture model[J]. SIGNAL PROCESSING-IMAGE COMMUNICATION, 2017, 55:66-79.