Research on Conditional Random Field Video Segmentation Technology Based on Block Processing

Jing WANG1*, Feng XU2
1Ningbo Ploytechnic, Ningbo, China
2Ningbo fastener LTD, Ningbo, China

*Corresponding author(E-mail: wangjing16@163.com)

Abstract
This paper begins with the image compression coding, and then studies the key point video segmentation of video compression and processing. In order to solve the problem of time complexity of video segmentation in traditional conditional random fields, this paper proposed a conditional random field segmentation algorithm based on video block processing. Firstly, the pixel spatial correlation is adopted, to make 3*3 block processing of the original video frames and to form a new video frame. And then, the energy function is reconstructed, adaptively selecting the conditional random field segmentation parameters, and carrying on the segmentation. At last, according to the initial segmentation, segmentation result of previous frame, and the current frame conditional random field’s segmentation results, the segmentation video sequence in the original size is recovered. Finally, according to the framework, the computer simulation is conducted. According to the experimental results, the performance of image coding and video segmentation is compared. The experimental results showed that the image coding algorithm can obtain the improvement in different degrees in the PSNR performance indicators. However, the video segmentation algorithm can effectively reduce the time complexity of the algorithm without obviously increasing the false segmentation rate.

Keywords: Particle Swarm Optimization, conditional random fields, video segmentation, correlation of adjacent pixels

1. INTRODUCTION

With the rapid development of computer technology, the pace of human being into the information age is gradually speeding up, and the information level of the whole society is gradually deepening. Network, multimedia and other sources of information is increasingly rich and diverse. Among them, the image and video information are the very important information sources. According to statistics, about 70% of the information are received from the human visual. It can be seen that the development of video technology is of great significance. To meet the people's growing demand for more real representation of the scene and more natural description, the computer viewpoint video technology is continuously developing. Recognized as important and basic technology, multi view video technology is paid more and more attention to by the academic and industrial field, and has become a hot topic in the field of video research.

The bottleneck of the technology is that the amount of raw data generated by multi view video is large. In order to realize the storage and transmission of data, the demand for bandwidth and storage space will be higher. The corresponding cost of production and application will be more expensive. In order to improve its efficiency, it is necessary to compress video information efficiently. The key to the application of digital video technology lies in the video compression technology. Because there is a strong correlation between frames in multi view video stream, it is possible for video compression. The key of video compression is to make full use of the temporal and spatial information to remove the redundant information between the frame or in the frame, so as to achieve the purpose of saving bandwidth, efficient transmission and storage.

In this paper, firstly, from the field of image compression, the codebook design algorithm, which is one of the key technologies of vector quantization, is studied. The particle swarm algorithm (PSO) and vector quantization (VQ) encoding are effectively combined, and an improved particle swarm algorithm based on codebook design is put forward, effectively improving the codebook performance. In addition, the vector quantization of image segmentation is studied. Then, starting from the key technologies of video segmentation based on content encoding, the conditional random fields (CRF) video segmentation algorithm based on block processing is proposed. What’s more, the context and frame redundant information are efficiently applied, adaptively selecting the threshold, and reducing the time complexity of traditional CRFs video segmentation algorithm, which effectively segments the objects that are interested in (usually foreground target).

2. VECTOR QUANTIZATION CODEBOOK DESIGN BASED ON IMPROVED PSO
PSO algorithm is a new swarm intelligence optimization algorithm. It not only preserves the position information of the intermediate state of the particles, but also preserves the velocity information, and provides evolutionary formulas of velocity and position. As a result, compared with the genetic algorithm and other algorithms, it has many advantages, such as fast and simple usage, fast searching speed, and easy to be implemented. However, in the process of particle swarm pursuing the optimal particles, it is easy to fall into local optimum and occur “premature”. To overcome this problem, an improved PSO algorithm is proposed in this paper. The algorithm uses the method of double particle swarm paralleling pursuing the optimal particles. When a population falls into local optimum, it can jump out of the local optimum under the help of another population, as described below.

2.1 Initial codebook design

The selection of the initial codebook affects the efficiency of the algorithm and the performance of the coding to a certain extent. In order to improve the coding performance, compared with the standard algorithm, this paper adopts the more ergodic initial codebook. The experimental results showed that the better the ergodicity of codebook is, the better it will be for the convergence of the iterative algorithm and the global optimal solution. Accordingly, this paper generates the initial codebook according to the following rules:

1) Make an ascending arrangement of \( N \times d \) training vector sets \( \{X_1, X_2, \ldots, X_N\} \) according to its sum;
2) According to the size of codebook, \( N_c \) is divided into \( N_c \) sets with the length of \( C \), \( C = N / N_c \), as shown in Table 1:

| \( X_1, X_2, \ldots, X_C \) | \( X_{C+1}, X_{C+2}, \ldots, X_{2C} \) | \( \ldots \) | \( X_{(N_c-1)*C+1}, X_{(N_c-1)*C+2}, \ldots, X_{N_c*C+C} \) |

| Table 1. Division of initial codebook |

3) Each time, take one from the sets with the size of \( C \), a total of \( N_c \) constitute a codebook. Repeated in this way for \( L \) times, we can get \( L \) initial particles.

It is not difficult to see that the initial codebook generated according to the above method is uniform distributed in the entire training vector space. At the same time, the initial codebook generated according to this method, the code vector is sequenced in accordance with the sum value in an ascending order from small to large, having the consistency, which is conducive to the subsequent iterative optimization searching [4]. Compared with the randomly generated initial codebook, the codebook obtained by this method has ergodicity, which increases the diversity of search and it can better search for the global optimal solution.

2.2 Double particle swarm optimization

In order to overcome the disadvantage that the PSO is easy to fall into local optimum, this paper presents a method of parallel optimization using two populations. The basic idea is shown as follows:

As shown in Figure 1, there is only one global optimal solution in the region. Set the two ellipses represent the population 1 and 2 in the region. \( A_{best} \) and \( B_{best} \) refer to the global optimization (codebook) founded by population 1 and 2 in an optimization searching process. \( d \) represents the distance between \( A_{best} \) and \( B_{best} \), using Euclidean distance for the measurement. As shown in (1), \( f(P_{A_{best}}) \) and \( f(P_{B_{best}}) \) indicate the fitness of the two codebooks, respectively (negatively proportional to the class dispersion).

\[
f(X_i) = K / \sum_{j=1}^{n} J_j
\]

In (1), \( K \) refers to the adjustment consonant that is determined by depending on specific cases. According to this, the smaller the class dispersion, the greater \( f(X_i) \), indicating the better the performance of codebook.
The two populations are carried out in different initial states for the evolution. Preset a threshold that is dependent on the specific circumstances. In an iterative process, if the distance between \( A_{best} \) and \( B_{best} \) is less than (or equal to) the threshold, it is considered that they are in the same optimal solution; if \( d > \) threshold, it is believed that they are at different optimum ranges.

When \( d \) is less than or equal to the threshold, there are two situations: 1) they are in the same local optimum; 2) they are in the global optimum, because there is only one global optimum in the same region, and there are many local optimal solutions [6]. Two populations make the optimum searching from different directions, the optimal solution they find is unlikely to be the same local optimal solution. In consequence, the first possibility is ignored, and it is considered that the optimal solution they find is the global optimum. In fact, in the process of the experiment, this paper has judged the first case that may occur, and introduced into the judgment algebra. As a result, it is able to jump out when the two populations are trapped into the local optimum, but it was found that this step did not make a better result. Therefore, it is abandoned, and only the second case is considered.

On the other hand, when \( d \) is greater than the threshold, there are also two possibilities: 1) one of them is the global optimum, and the other is the local optimum; 2) two different local optimums. In either case, at least one of the two is not in the global optimum, or at least one is a local optimum rather than a global optimum. If \( f(P_{Ag-best}) > f(P_{Bg-best}) \), then the performance of codebook \( A_{best} \) is better than that of \( B_{best} \), so \( A \) is more close to the global optimum. In this case, the algorithm updates the global optimum of the two populations according to the following formulas:

\[
\begin{align*}
A_{best} &= \left( f(P_{Bg-best}) / f(P_{Ag-best}) \right) A_{best} + \left( 1 - f(P_{Bg-best}) / f(P_{Ag-best}) \right) B_{best} \\
B_{best} &= A_{best}
\end{align*}
\]

(2)

When \( f(P_{Ag-best}) \leq f(P_{Bg-best}) \),

\[
\begin{align*}
A_{best} &= \left( 1 - f(P_{Ag-best}) / f(P_{Bg-best}) \right) A_{best} + \left( f(P_{Ag-best}) / f(P_{Bg-best}) \right) B_{best} \\
B_{best} &= A_{best}
\end{align*}
\]

(3)

After several experiments, it is found that, the smaller the class dispersion, the larger the fitness function, the better the performance of codebook, and the closer to the global optimum. When \( f(P_{Ag-best}) > f(P_{Bg-best}) \), \( A_{best} \) is closer to the global optimum. As a result, in the above formulas of global optimal solutions, \( A_{best} \) accounted the larger proportion; on the contrary, when \( f(P_{Ag-best}) \leq f(P_{Bg-best}) \), \( B_{best} \) accounted the larger proportion.

2.3 Vector quantization image segmentation

Because of the consistency of vector quantization and image segmentation, vector quantization can be used for image segmentation. Image segmentation method based on generalized image and vector quantization is used here. The algorithm uses the generalized image as the training sequence, and then uses LBG algorithm for the training and iteration to obtain the optimal segmentation of the generalized image, so as to realize the segmentation of the original image. Because of the generation of general images makes use of smoothing filter, the robustness of algorithm to noise is quite good, which can effectively suppress noise and reduce the segmentation error. It specifically can be divided into the following steps:

1) Generalized image acquisition: in order to obtain the segmentation vector, we first of all need to obtain the generalized image. Set there is an original image \( X \) with the size of \( M \times N \), which is carried out with 5 * 5 dot matrix smoothing (in this paper, 5 x 5 median filter is used to verify) to obtain the smooth image \( Y \) with the same size. \( X \) and \( Y \) constitute a generalized image, and the pixel combination \((x_i, y_i)\) of the corresponding position constitutes a two-dimensional vector.

2) Initial codebook selection: because it is only the validation of binary images, split method is used to select the initial codebook. Let \( X \) and \( Y \) refer to the gray level mean value of the original image and the smooth image, respectively, and \( \sigma_x^2 \) and \( \sigma_y^2 \) are the variances, respectively. The initial codebook \( Y_1^0 = \left( y_{11}^0, y_{12}^0 \right) \) and \( Y_2^0 = \left( y_{21}^0, y_{22}^0 \right) \) are selected as follows:

\[
\begin{align*}
y_{11}^0 &= x + \alpha \sigma_x \\
y_{12}^0 &= y + \alpha \sigma_y \\
y_{21}^0 &= x - \alpha \sigma_x \\
y_{22}^0 &= y - \alpha \sigma_y
\end{align*}
\]

(4)
\[ \begin{align*}
    y_{21}^0 &= x - \alpha \partial_x \\
    y_{22}^0 &= y - \alpha \partial_y
\end{align*} \]

In (4) and (5), the parameter \( \alpha \) is selected 0.5.

3) Algorithm iteration and termination: the process of training and iteration is performed by LBG algorithm. Firstly, after selecting the initial codebook, the generalized image is divided into \( N \) independent regions according to the nearest neighbor criterion. The distance measurement is conducted by using European distance formula. The average distortion is calculated, and the iterative process is aborted if the average distortion satisfies the preset stop threshold; otherwise, carry out the next round iteration.

Since the generalized image contains the original image, it is easy to get the segmentation result from the final iteration results. It can be seen that the advantage of the algorithm is not sensitive to noise. And the disadvantage is that the scope of adaptation is limited, and it can get good segmentation results for the simple two value image, but it is not ideal for the complex image [10]. However, for the case of multiple targets, it is necessary to consider the selection of initial codebook. Therefore, the possibility of practical application is not large. Then, because this CRF video segmentation algorithm has not high requirements for the segmentation of the input pixel level, it is possible to still use the algorithm for pixel level segmentation of video frame; and the algorithm is novel that it can be used as reference.

3. CONDITION RANDOM FIELDS VIDEO SEGMENTATION BASED ON BLOCK PROCESSING

3.1 Limitations of CRF video segmentation

Video segmentation is the process of making use of the prior knowledge of the context and neighbourhood information to construct the model (probabilistic graph model, etc.), and then to solve the model and finally get the segmentation target. Traditional segmentation algorithms, for various reasons (such as no good use of the context), and bring some, brings “wormhole”, “empty” error classification and so on defects. It is also not good for the adaptability of the background movement.

The video segmentation based on CRF can use the contextual features to solve the labelling bias, and the global optimal value can be obtained by global normalization of all the features. Compared with the Markov random field and the hidden Markov random field, it is discriminant and has no strict independence assumption condition, so it can accommodate any contextual information, and overcome the problem of label bias.

But it also has some disadvantages: high complexity of algorithm, large amount of computation and high training cost. The application prospects are correspondingly limited. In order to reduce the time complexity of conditional random fields video segmentation algorithm, this paper applies the block processing technique to CRF video segmentation, as described below.

3.2 Video block processing

3.2.1 Video frame block

The traditional conditional random fields video segmentation algorithm is to use the video pixels as the segmentation unit, which needs to construct the energy function for the pixels one by one, and at last solve the maximum of the energy function. Although more accurate segmentation results can be obtained, the high complexity and the large amount of computation limit its application in the process of training and model deduction.

Because the pixel in the same spatial characteristic neighbourhood has great similarity, the block used here is in the size of 3 * 3 after several experiments, so as to make full use of the characteristics of the video frame, but also guarantee the correctness of segmentation. For the input video sequence of \( N \times M \), the video frame size becomes about \( (N \times M) / 9 \) after the 3 * 3 block segmentation, which is 1/9 of the original size of the original video frame. The process of getting the pixel value \( g \) of the new pixel from the block \( B \) of 3 * 3 is in accordance with the following formula:

\[ g = \mu \sum_{i=1}^{8} g_i d_i \]

In (6), \( \mu \) refers to the experience parameter, so as to ensure the new pixel value \( g \) calculated in a reasonable range. \( g_1 \sim g_8 \) indicates the 8 neighbors of \( g_0 \), and \( d_1 \sim d_8 \) suggests the distance measurement between the two pixels. The pixel value of \( g \) is obtained through the pixel values \( g_i \) in the block \( B \) multiplying the corresponding distance measurement \( d_i \). And \( d_i \) is obtained by the following formula:

\[ d_i = \frac{\exp(x_i / \text{avg}(B))}{e + \varepsilon} \]

(7)
In (7), \( \text{avg}(B) \) represents the pixel average value of the pixel block \( B \), that is to say, it is the sum of each pixel divided by the number of pixels in the block. And \( x_i \) refers to the pixel value of the \( i \)-th pixel. It can be seen that, the smaller the difference between the pixel value of a pixel and the pixel average value of the entire block, the larger the proportion of it in the new pixel; on the contrary, it is smaller.

3.2.2 CRF segmentation

After the new video sequence is obtained by the sub block processing technique, the pixel level is initially segmented to obtain the initial label value.

In this paper, the energy function is redefined as follows:

\[
E(L; X) = \alpha U^G(L; X) + \beta U^C(L; X) + \eta V^S(L; X)
\]

(8)

Among them, the meaning of the components of each energy function is as follows:

a) \( U^G (L; U) \) represents the local energy characteristics of pixels, defined as follows:

\[
U^G (L; X) = \sum_i \delta(l_i, l_{i,k}) \log P(x_i | l_i)
\]

(9)

In (9), \( x_i \) and \( l_i \) indicate the pixel value and the new label value of the pixel point \( i \), respectively, \( \delta \) refers to the Kronecker delta function, and \( l_{i,k} \) suggests the initial label of the pixel point \( i \), which is obtained according to the segmentation in the pixel level. And is the probability of the background, shadow, and foreground calculated in accordance with mixed Gauss distribution model.

b) \( U^C (L; X) \) indicates the color feature energy function, defined as follows:

\[
U(L; X) = \sum_i \log P(x_i | l_i)
\]

(10)

Among them, the probability \( P(x_i | l_i) \) of the node's colour feature is estimated by using the UV histogram in the YUV colour space.

c) \( V^S (L; X) \) refers to the neighbor relationship energy item. For the pixel \( x \), it is defined as follows:

\[
V(L; X) = \sum_{j \in N(i) \setminus M(i)} v(l_i, l_j)
\]

(11)

In (11), \( l_i \) is the neighbor node of \( l_i \), and \( v(l_i, l_j) \) is often defined as follows:

\[
v(l_i, l_j) = \begin{cases} 
\lambda_1 & \text{if } l_i = l_j = 0 \\
\lambda_2 & \text{if } l_i = l_j = 1 \\
\lambda_3 & \text{if } l_i = l_j = 2 \\
\lambda_4 & \text{if } l_i = l_j 
\end{cases}
\]

(12)

In the above formula, \( \lambda_i \) is the default parameter. Traditional algorithms have different empirical values for different segmentation targets. It is obvious that the empirical setting method of the parameter needs to find out the optimal value in many experiments in specific operations, and the operation is complex and does not take into account the temporal and spatial correlation between pixels. Therefore, the algorithm makes full use of the correlation between pixels, referring to the normal distribution model, and defines the following formula to adaptively select parameters:

\[
v(l_i, l_j) = \delta(l_i, l_j) \frac{1}{\sqrt{2\pi} \psi_{\text{adj}}} \exp \left\{ -\frac{1}{2} \frac{(x_j - x_i)^2}{\psi_{\text{adj}}} \right\}
\]

(13)

In (13), \( x_i \) indicates the central pixel, \( x_j \) represents its neighbor pixel, and \( \psi_{\text{adj}} \) refers to the variance of the entire neighbor pixel and the central pixel \( x_i \). The calculation formula is shown as follows:

\[
\psi_{\text{adj}} = \sum_{j \in N(i) \setminus M(i)} \frac{\|x_i - x_j\|^2}{\|x_i - x_j\|^2}
\]

(14)

In (14), \( \| \| \) refers to the European distance. It can be seen that, \( v(l_i, l_j) \) is mainly determined by the difference between the two pixels. The smaller the difference, the larger the value of \( v(l_i, l_j) \), and the greater the contribution of the neighbor pixel to the neighbor energy; on the contrary, the contribution to the neighbor energy is smaller.

According to the energy function defined, the segmentation results of conditional random fields for the video is shown as follows:
\( \hat{^L} = \arg \max_P (L | X) \) \hspace{1cm} (15)

For the solution of above formula, generally, there are ICM, graph-cut, Belief propagation, Gibbs sampling and so on.

4. SIMULATION EXPERIMENTS AND ANALYSIS

4.1 Vector quantization simulation system

As a technology of information processing and transmission system, vector quantization coding can be evaluated by the indexes of validity and reliability. There are three commonly used indicators: coding rate, implementation complexity, and distortion measure. Subjective indicators are also used to evaluate the advantages and disadvantages of vector quantizer. Peak signal to noise ratio (PSNR), only measures the degree of distortion of the image, but also reflects the amount of information. As a result, this paper chooses PSNR index as the evaluation index of the performance of the proposed algorithm.

4.1.1 Initial codebook design

In order to verify the effectiveness of the method for the selection of the initial codebook, this paper has done the following experiments: select 256 gray level standard Lena and Cameraman images with size of 256 x 256 as training images. The images are divided into 4096 4 x 4 blocks, 16 pixels of each block as a 16 dimensional vector training. The codebook size is 32, 64, 128 and 256, respectively. And then, the initial codebook randomly generated and the codebook generated by this algorithm are used as the final codebooks, and the Lena image and Cameraman image are encoded and decoded, respectively. The experimental results are shown in the following table:

<table>
<thead>
<tr>
<th>Name of images</th>
<th>The method used</th>
<th>The size of codebook (N)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>32</td>
</tr>
<tr>
<td>Lena</td>
<td>Random codebook</td>
<td>13.86</td>
</tr>
<tr>
<td></td>
<td>The method in this paper</td>
<td>15.92</td>
</tr>
<tr>
<td>Cameraman</td>
<td>Random codebook</td>
<td>13.05</td>
</tr>
<tr>
<td></td>
<td>The method in this paper</td>
<td>14.10</td>
</tr>
</tbody>
</table>

As can be seen, the performance of the codebook of initial codebook generated in the algorithm is better than PSNR index and the codebook randomly generated. And with the increase in the size of the codebook, the effect will be more obvious. Thus, in the subsequent iteration process, the initial codebook generated by this algorithm can reduce the number of iterations, and it is more conducive to converge to the global optimal codebook.

4.1.2 IPSO-VQ algorithm experiment

In order to verify the effectiveness of the proposed improved algorithm in vector quantization codebook design, we use 256 gray level standard Lena and Cameraman images with the size of 256 x 256 as training images, respectively. In addition, we use the standard PSO algorithm and the algorithm proposed in this paper for vector quantization encoding and decoding. In the experiment, the images are divided into 4096 4 x 4 blocks, 16 pixels of each block as a 16 dimensional vector training. The codebook sizes are 32, 64, 128 and 256 respectively, and then the Lena is encoded by codebook in different sizes obtained by the training. The size of the particle swarm is 10, the number of iterations are taken 40, and \( c_1, c_2, \) and \( w \) are valued as described by the algorithm. When the codebook is 32, 64, 128 and 256, the threshold takes 5 times of its corresponding values, namely 160, 320, 640, and 1280.

And then, the same operation is done for the Cameraman. The difference is that when the codebook is 32, 64, 128 and 256, the threshold takes 3 times of its corresponding values, that is, 96, 192, 384, and 768, and another group of data is obtained. According to the above method, this paper carries on the experimental simulation on the Matlab R2009b platform, and the results obtained are shown as follows:

<table>
<thead>
<tr>
<th>Name of images</th>
<th>The method used</th>
<th>The size of codebook (N)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>32</td>
</tr>
<tr>
<td>Lena</td>
<td>PSO</td>
<td>24.71</td>
</tr>
<tr>
<td></td>
<td>The method in this paper</td>
<td>24.92</td>
</tr>
<tr>
<td>Cameraman</td>
<td>PSO</td>
<td>21.85</td>
</tr>
<tr>
<td></td>
<td>The method in this paper</td>
<td>22.10</td>
</tr>
</tbody>
</table>
Since that the codebook generated by the clustering algorithm will have a certain randomness, the final experimental results will have a certain fluctuation. The PSNR on the above table is the average of the 10 experimental results. As can be seen from the table, compared to the standard PSO vector quantization image compression algorithm, regardless of the size of the codebook, the PSNR of the algorithm has been improved in different degrees.

In order to further verify the effectiveness of the algorithm, this paper has done the following experiments:

![Figure 2. PSNR with codebook size of 32](image)

When the codebook size is 32, the PSNR of the algorithm and the standard PSO vector quantization algorithm is compared with the results of encoding and decoding of the Lena image. It can be seen from the figure that the algorithm is superior to the standard PSO algorithm in the compression effect, and the PSNR amplitude is smaller than that of the standard PSO algorithm.

4.2 Video segmentation simulation system

The segmentation results are obtained by using CRF segmentation method and adaptive parameter model segmentation. Experimental video sequences are Sample Video and Hall_monitor.

The following Figure 3 showed the segmentation results of the Sample Video experimental video sequence 30th frame by using the traditional method and the method proposed here, respectively, and Figure 4 showed the result of Hall_monitor segmentation:

In the segmentation results, the white and black regions represent the foreground and background, respectively. From the experimental results, we can see that in the MRF segmentation, due to the lack of full use of the context features, some of the foreground are mistakenly classified into the background, resulting in a hole. In particular, Figure 3-4d is more obvious because the color of the arm is not significantly different from that of the background. In this paper, the algorithm and the traditional CRF did not produce a hole. Moreover, the segmentation results of the proposed algorithm can be better approximated to the segmentation results of the traditional CRF algorithm without block processing, as shown in the following figures:

![a) Initial video frame](image)  ![b) Segmentation results of the algorithm proposed in the paper](image)
Through the subjective observation and objective data realization (PSNR and false segmentation rate), the experimental results are analyzed. In addition, the performance of the algorithm in each link is obtained, and it is compared with existing mature algorithms. Finally, the advantages of the proposed algorithm PSO-VQ compared with the standard algorithm and the advantages of the CRF segmentation framework are listed. From the experimental results, it is seen that, the algorithm has been improved and optimized to a certain extent. But the results are not ideal enough, and for the video sequences with large background motion, the algorithm is not significantly improved compared with the traditional CRF algorithm, and the application will be limited accordingly.

5. CONCLUSION

In this paper, we first of all study the video coding from the image coding. However, only the video segmentation of image coding and content-based video coding is completed. Firstly, in image encoding, based on the deficiencies that the traditional PSO encoding is easy to fall into local optimum, a kind of improved idea is put forward. Two populations parallel optimization is adopted. When a population is trapped into a local optimum, the other population can “pull out” it. In addition, the validity of the algorithm is verified by experiments and the PSNR index. Secondly, in the aspect of video segmentation, this paper uses CRF video segmentation algorithm to build the framework and construct the potential energy function.
REFERENCES


