TELKOMNIKA, Vol. 11, No. 10, October 2013, pp. 6051 ~ 6059
ISSN: 2302-4046

Region of Interest Coding Based on Support Vector Machine in Transmission System

Li Zhu*, Xiaoguang Wu, Rong Li
School of Mechanical Engineering and Automation/Wuhan Textile University, China
Wuhan Textile University, 430073/ Wuhan Textile University, China
*Corresponding author, e-mail: zlcurie@gmail.com

Abstract

The development of network technology and multimedia technology has promoted the development of the integration, digitalization and individualization of communication technology, and has realized multiple functions based on wireless channel network, including voice, data and image. On the basis of classification by Support Vector Machine (SVM), this paper has proposed a low bit-rate coding method for segmenting the Regions of Interest (ROI) of objects through reasonably selecting initial class center and dynamically adjusting the number of categories, and designs its 3-dimension model method. In low bit-rate coding, the ROI is coded first using SPIHT algorithm, and then the background or the regions insensitive are coded using H.264 code, analyzing the experimental data and designing the structure, we can get the relation among related parameters under conditions of different permanent-magnet thickness and air gaps. It is proved that this method can achieve a better recognition effect, greatly improve the coding efficiency and satisfy the visual effects of human eyes, which can provide criteria in theory and application for designing on new driving principle, control system.

Keywords*: Region of Interest, suspension-type selecting, Transmission System*

1. Introduction

Low bit-rate transmission and detection system uses computer detection technology to process, analyze and resolve signals, and controls the transmission system to enable the detection system have the automatic tracking function as human eyes.

When processing signals based on ROI, the coding part is usually after the region segmentation module, and the study of ROI coding is generally in the context of the segmented ROI. Using different compression strategies for different regions can ensure the storage and transmission of important information prior. The studies on static image coding started earlier, on which the current research achievements on ROI coding are focusing. The research on ROI coding compression for sequence has never ceased. To meet the security needs, the demand for detection systems is rising day by day, so it is necessary to reduce the dependence on cumbersome manual work, and make it implement real-time detection of the targets in complex environment and the behavior analysis and transmission of ROI [1-4]. This paper studies on the target tracking algorithms of the low bit rate object detection system with the key computer visual techniques applied in low bit rate coding system, and the practical parallel processing technology for DSP architectures.

At present, the stream generated by traditional coding techniques has been difficult to adapt to the fluctuations in network bandwidth and the demand of network heterogeneity, it is easy to miss some detail information in the low bit rate compression. ROI coding can flexibly adjust the coding rate and receive the stream partly or completely according to the channel capacity and visual needs so as to restore the quality of different signals.

2. Preprocessing of ROI Target

The noise of acquisition devices, such as sensors, variation of texture roughness (resolution) in the background area and the serious imbalance of pixel-scale of edge-region will worsen the performance of the low bit-rate coding algorithms. Therefore, filtering and denoising the original sequence is necessary [5-7].

Received February 19, 2013; Revised July 12, 2013; Accepted July 22, 2013

2.1. Median Filtering

Median filtering is simple and fast in operation, and is excellent in filtering out superimposed white noise and superimposed heavy-tailed noise, while protecting the signal details soundly.

W is the filter window with a length of *L* = 2*N*+1, and the sample points of the input signal sequence within the window at the time n are *x*(*n* - *N*), … ,*x*(*n*), …, *x*(*n* + *N*). The output of median filtering is defined as:

$$y(n) = med[x(n - N),\ldots,x(n),\ldots,x(n+N)] \quad (1)$$

med[] expresses the median of all the numbers in the window arrayed in an order from small to big [8].

2.2. Anisotropic Filtering

The basic idea of anisotropic diffusion method is to spread the original image u0, defined in the convex domain $\Omega \subset R*R$ at a variable rate to acquire a series of gradually increased smooth image *u*(*x*,*y*,*t*) [9].

$$\frac{\partial u}{\partial t} = div(c|\nabla u|\nabla u) \quad (2)$$

c(•)is the diffusion coefficient, satisfying *c*(0) = 1, *c(s*) ≥ 0 (*s* ≥ 0), and the algorithm calculation is as follows:

$$I(x,y,t+\Delta t) = I(x,y,t) + \Delta t(d_n \cdot C_n + d_s \cdot C_s + d_e \cdot C_e + d_w \cdot C_w) \quad (3)$$

$$d_n = I(x,y-1,t) - I(x,y,t),\ C_n = \frac{1}{1+\frac{d_n^2}{k^2}} \quad (4)$$

$$d_s = I(x,y+1,t) - I(x,y,t),\ C_s = \frac{1}{1+\frac{d_s^2}{k^2}} \quad (5)$$

$$d_e = I(x-1,y,t) - I(x,y,t),\ C_e = \frac{1}{1+\frac{d_e^2}{k^2}} \quad (6)$$

$$d_w = I(x,y+1,t) - I(x,y,t),\ C_w = \frac{1}{1+\frac{d_w^2}{k^2}} \quad (7)$$

Where *n*, *s*, *e*, *w* represent four directions(north, south, east and west) of the center pixel.

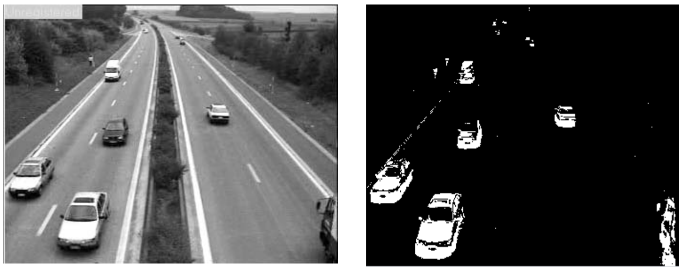
2.3. Mean Filtering

Using the average or weighted average of neighborhood value can effectively suppress noises. Mean filtering allows the main low-frequency components of the signal pass through, and cuts off high frequency noise signals. While reducing the impact of random noise points, the smooth process will lead to edge blur due to the edges of images are also within high-frequency part [10].

2.4. Region Filling

The filling algorithms for random enclosed regions are mainly two: Seed filling algorithm and polygon scan conversion algorithm. The basic idea of the seed filling algorithm is: First, assume a point within the contour is known, then begin searching the points within the contour adjacent to the seed. If the adjacent point is not within the contour, then reach the contour of the boundary; if the adjacent point is inside the contour, it becomes the new seed point, and then continue search.

Through the above pretreatment, the basic structure of the transmission target detection system can be obtained. Using the image frame at t - 1 as the background model of image frame at t, capture the test sequence by using USB camera to detect the moving target. As in the actual testing process, the edge of the target object is often not smooth, there is the phenomenon of noise and defects, and voids exist within the target area, which will influence the subsequent object segmentation algorithm. Therefore it is necessary to use the algorithm of the morphological region filling and edge extraction, smooth edge pixels on the target and fill the inside of target to reach for a appropriate results of segmentation. The results shown in Figure 1.



Original image Region filling

Figure 1. Results of Morphological Detection

Conduct target segmentation to the image after region filling and determine the edge position. The structural information of moving targets can be detected in the context of relatively simple conditions. There are thousands of kinds of image segmentation algorithms currently in the world, but the universality of current segmentation algorithms is limited due to the complexity of scenes, so the silhouette of detected moving object is not smooth and there are a few of voids inside. According to the characteristics of ROI, This paper has proposed a new moving target extraction algorithm based on SVM classification and neural network clustering, discovered the Region of Interest (ROI) to extract the target, as a basis of follow-up ROI coding.

3. Segmentation Based on SVM Classification

The detection and classification of moving target is a process separating the moving target from the sequential image, extracting the features of the moving target, and conducting the image classification and recognition with trained classifier [11].

We have outlined a SVM approach with good generalization ability first, and introduced an idea of classifying the moving target based on SVM according to the characteristics of monitored target, then conducted ROI process on the target acquired by SVM-based segmentation and optimized coding to realize a ROI coding algorithm design under the conditions of low bit rate transmission.

3.1. Brief Review of Support Vector Machine

This paper has put forward a number of classification and recognition methods according to classified objects and applications[12-15]. Some applications prove that using SVM turns out a better adaptability and promotion ability. In this paper, a SVM-based approach of classifying the moving targets on static background has been proposed on the basis of SVM algorithm.

Assume a sample set (x_i, y_i), i = 1, 2,…,n, $x_i \in R_d$, and $y_i \in$ {+1, -1}express the category, Class 1 and Class 2 are linear separable, then there is (w, b) to make:

$(w \cdot x_i) + b > 0,\ x_i \in \text{Class1} \qquad (w \cdot x_i) + b < 0,\ x_i \in \text{Class2}$

$g(x) = w \times x + b$, the linear expression of decision function g(x), analysis (w, b) as the best threshold to segment Class 1 and Class 2 by discriminant calculation. As shown in Figure 2, H means the classification interface separating the classification plane of the two classes. The distance between H1 and H2 are the classification intervals of two classes. The SVM-based optimal classification plane is able to separate the two types of interface to the maximum extent, and maximize the interval of two classification planes. When the structure of the needle selecting device is determined, the coefficient of force is only related to the air gap h which is the displacement of needle movement. Therefore, the loading mode of the driving current can be obtained according to the movement of needles, and their relation can be defined as:

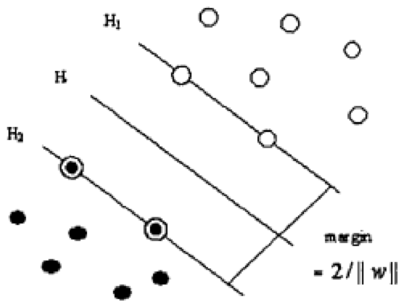


Figure 2. Optimal Classification Interface

Let *w* set for a certain number of arbitrarily defined a set of *x*, classified according to the following decision function:

$$f(x) = \operatorname{sgn}(w \cdot x) = \begin{cases} 1 & class\ 1 \\ -1 & class\ 2 \end{cases} \quad (8)$$

x_w is a point in set w nearest to x, it can be deduced from the knowledge of geometry that the distance between set x and set w is $\|x-x_w\|$.

SVM-based optimal hyperplane *H*: $w \cdot x = 0$, with introduced bias item h, then the decision function is:

$$f(x) = sgn(w \cdot x) + b \quad (9)$$

To maximize the boundary distance, you need to minimize $\frac{1}{2}\|w\|^2 = \frac{1}{2}w^T w$. The objective function is strict convex quadratic form. Based on the solution for convex quadratic program in optimization theory, solve the function by translating it into a dual problem in the form as below:

$$\Phi(w,b,a) = \frac{1}{2}\|w\|^2 - \sum_{i=1}^{N} a_i \left[y_i (w \cdot x_i + b) - 1 \right] \quad (10)$$

Seek the partial derivatives of *w* and *b* respectively, as the expressions given by:

$$\max_a W(a) = \max_a \left(\min_{w,b} \Phi(w,b,a) \right) \quad (11)$$

When α_i is greater than zero, x_i is the support vector to determine the classification boundary. Substitute Equation (12) into the decision function (10):

$$f(x) = \mathrm{sgn}\left(\sum_{i=1}^{N}(y_i\alpha_i(x_i \cdot x) + b)\right) \tag{12}$$

The classification function obtained by SVM is similar to the solution of the neural network in the form, as shown in Figure 3.

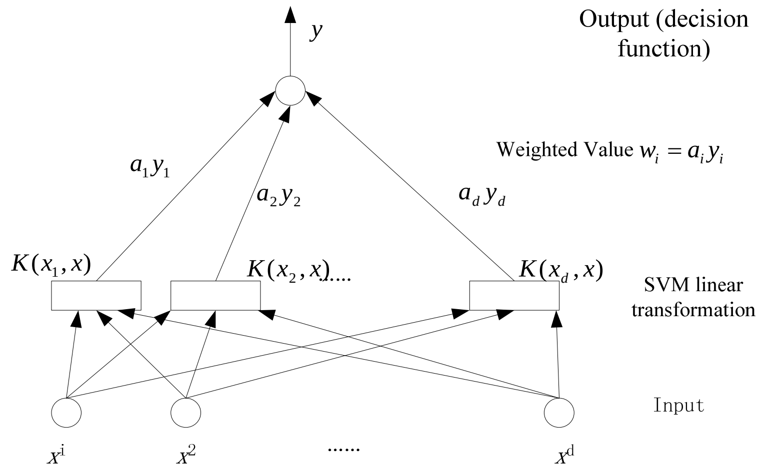


Figure 3. SVM Classification Diagram

3.2. SVM-based Classification of Moving Targets

It is found from a number of experiments that the image acquired by traditional method of target detection contains a small amount of noise and voids. In order to improve the accuracy of target analysis, we must eliminate this interference factor[16]. In order to effectively solve the problem of noise and voids, we should fill the voids of the image first by using smoothing filter and threshold filter on the binary image acquired after of the moving target detection by the motion vector model. In the clustering algorithm of target classification, the algorithm process is as follows:

Step 1: Traverse the template $B_k^{'}$;

Step 2: Traverse all the target pixels, determine whether there are other pixels within the square area threshold × threshold on the upper left of the pixel (departing from the pixel). If any, lump the scanned pixels as the target class, update the pixel values of this area and the 8 neighborhoods surrounding it, and set them all as the target; if not, create a new target class according to the scanning area, and classify the current target pixel into the target class;

Step 3: Scan the generated m target classes, set the regions, C_{ki} is the region of target class *i*;

Step 4: Calculate the number of the target regions and the threshold of each target region. If the threshold of target region is less than the initial threshold, this region can be regarded as the noise area and need to be filtered out; Retain the target regions whose threshold greater than the initial threshold and get the ultimate target class $C_k^{'}$, representing the location of pixel spaces of different target in the image.

As shown in Figure 4, the test image is the test sequence of vehicles in the road, and the noise points in the image are filtered out through a clustering target template, such as the noise in the lower right corner of the image is filtered out through clustering; the target template in the image is divided into three parts by clustering segmentation, and the white pixels in the red box represent independent target templates in the image. The effect of classification is shown in the Figure 4.

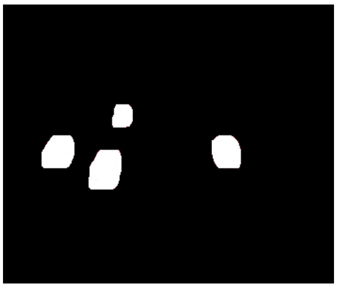
For segmented target, because of its adaptability and the uncertainty of segmented shapes, its position can not be directly targeted, so it should be processed by region filling, filling the target with blocks as follows:

Step 1: By using the eight-neighborhood counter-clockwise track scan mode.

Step 2: Scan the image from top to bottom and from left to right to find the first nonzero point (the pixel value of to-be-filled region is non-zero), then starting from this point as the center, scan the pixel points in Direction 1. If the pixel value in this direction is zero, then turn to Direction 2; otherwise, assign the direction code value to the center point as 1, and rotate the scanning direction two grids clockwise around the pixel in Direction 1.

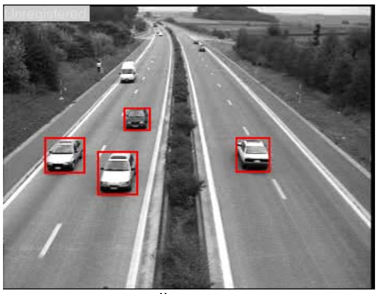
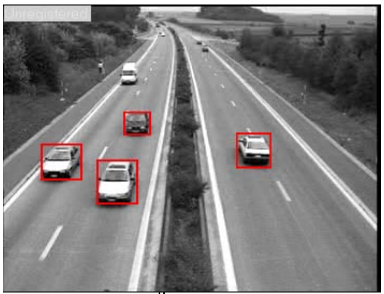
Step 3: Rotate the scanning direction a grid counter-clockwise to scan the pixels in direction 2. If the pixel value is nonzero, assign the direction code value to the center point.

Summing up the above, sample 4 frames consecutively from the images of road vehicle test sequence starting from the 10th frame, the effect observed is as shown in Figure 5.

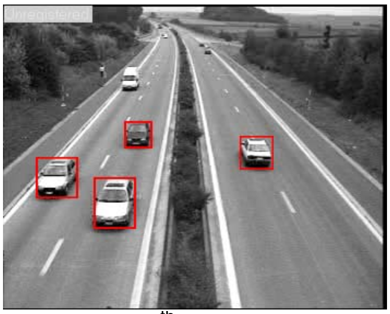
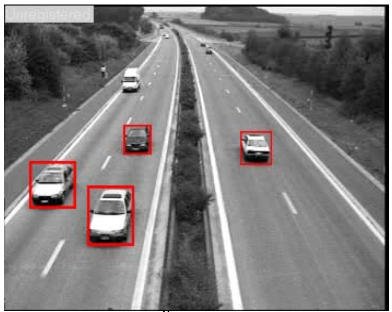



Classification Effect Binary of Segmentation

Figure 4. SVM-based ROI Segmentation

20th frame 21th Frame

22th Frame 23th Frame

Figure 5. Target Detection Effect of Transmission Based on Wireless Channel

The accuracy of this segmentation is unstable. The target may be identified as the background (missed), and the background may be identified as the target (false alarm). The detection probability and false alarm probability are the most important parameters for target detection. This paper has selected 200 consecutive images of the highway buses for moving target detection and extracted the moving targets based on SVM classification, ensuring a correct detection probability of 98.3%.

4. Coding Based on Segmentation Target
4.1. Encoding Algorithm

According to the method mentioned in the previous section, the SVM-based classification algorithm can accurately split the target by the vector machine, and give priority to encode it as the ROI for the next step. A ROI codec is designed based on wavelet transform and SPIHT algorithm on the framework of H.264 motion image encoder, as shown in Figure 7 and Figure 8, combing the intra-frame coding with the predictive inter-frame coding. Compared with the general sequence compression, the ROI coding to realize the region of interest is distinguished in several points[16-20]:

1) ROI bit-plane shifting, making the data of the ROI be at a higher bit-plane, and encoding ROI data first in the subsequent embedded SPIHT_ROI coding method.
2) ROI tracking area, seeking matched ROI in the current image by using motion vector V_{ij} and original ROI mask data, and transferring the ROI mask information to the ROI bit-plane shift mask module.

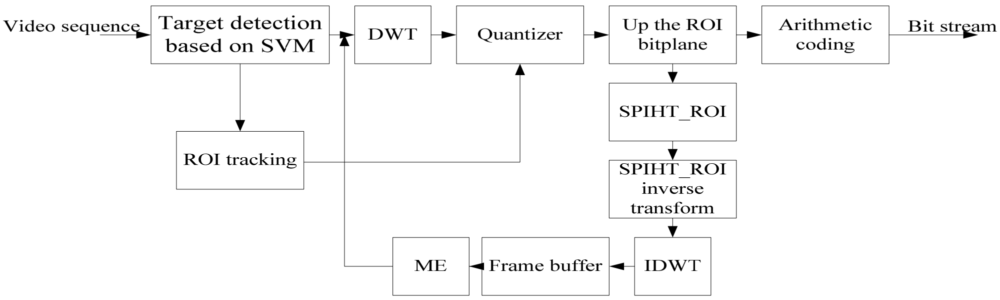


Figure 6. SVM-based ROI Segmentation

The encoder consists of two parts: "Predictive Coding" and "Rate-Control Feedback". Predictive Coding is mainly composed by DWT, quantization, ROI bit-plane shifting, SPIHT_ROI, ROI bit-plane recovery, inverse quantization, wavelet transform and coupler. The loop of Rate-Control Feedback is composed by the buffer, rate controller, DWT, quantization, SPIHT_ROI coding, adaptive arithmetic coding, etc. to control the stream output at a constant rate.

Send an image processed by wavelet transform into the quantizer for uniform quantization, conduct bit-plane shifting on background coefficients of ROI by the quantized wavelet coefficients and realize the embedded bit plane coding by the ROI-based SPIHT coding algorithm. SPIHT algorithm is a process sorting the wavelet coefficients by amplitude, setting all the coefficients of transform sub-band outside the boundaries of target to be zero. SPIHT algorithm itself does not need much change.

The ROI-SPIHT algorithm flow is as follows:

1. Segment the ROI of the input image sequence by SVM, and mark the regions of target and background;
2. Implement wavelet decomposition to the image, pad zero in the non-ROI area in the decomposition, encode the target region with SPIHT coding;
3. Mark the corresponding ROIs and their borders in the low-frequency sub-band, then back to step 2;
4. Repeat above steps to achieve the preset decomposition levels and make it adapt to the SPIHT coding;
5. Encode the segmented area with SPIHT after filling.

Since SPIHT algorithm itself has a progressive quantitative strategy, the step size of quantizer can be set to $n/2$, therefore, the relative quality of ROI and background can be flexibly controlled according to different needs.

The encoded image of current frame couples with the motion vector of current frame, and they together go into the encoder for adaptive arithmetic coding, to achieve the compression rate control and delivery the stream to the encoding buffer. Rate control is implemented by directly controlling the stop bit of SPIHT coding. In some encoding cases with higher requirements, you can close the wavelet transform coding, and use the method of frame-skipping to significantly reduce rate and improve the coding efficiency.

Working Principle of ROI Decoder: The working principle of the decoder is relatively simple. On the contrary, it directly decodes the encoded information of the encoder.

4.2. Experimental Results and Analysis

This ROI coding experiment system for sequence images is realized based on VC++6.0 software platform, running on the hardware environment of a common PM2.2G/256M/40G microcomputer, with 2G frequency and 512M memory. The test sequences are CIF sequence specified by MPEG-4. Frame rate: 30 frames/sec. Resolution: 352 × 288 pixels. Color space: Y: Cb: Cr. Encoder software: JM9.6. Reference frame digit: I. Format: 4:2:0.

In the experiment, the following program is used for testing targets on static background. First, process the bus objects on the highway with SVM segmentation to get the regions of interest as the moving targets, then encode these regions of interest with above-mentioned ROI coding method and background region with coarse coding. Using the open source of H.264 TMN9.8 for comparison, in TMN9.8, the first frame is I frame, followed by all P frames or B frames. This paper has realized SPIHT coding of the regions of interest by using ROI coding, and using H.264 coding for other regions. Figure 8 is the second test program, for ROI coding on static background.

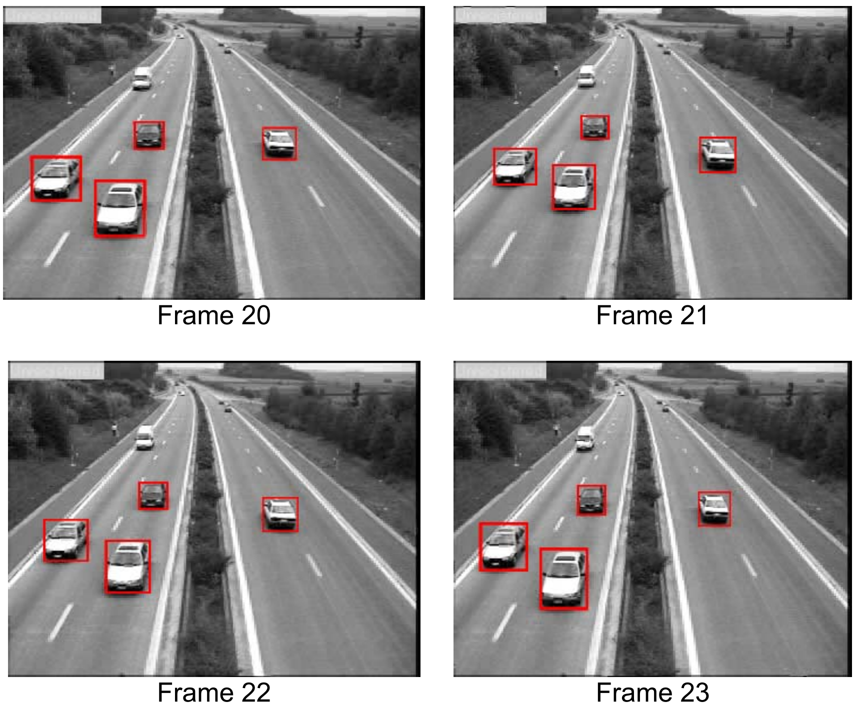


Figure 7. Target ROI Coding on Static Background Based on SVM Segmentation

Figure 7 is the result of target ROI coding on static background based on SVM segmentation. After ROI is selected, encode the region using SPIHT elaborated code to get better image quality. For non-sensitive visual areas, the peak signal-to-noise ratio has decreased, but this change is based on perception characteristics of human, and is valuable in practical application. Especially when the channel bandwidth or storage capacity is limited, it will give priority to ensure the quality of the ROI-based and the quality of encoding will be further improved. In aspect of coding efficiency, it requires a relatively long time if using the coding in this paper and needs to be optimized further with DSP.

5. Conclusion

Traditional ROI optimization algorithm is classified as bit-plane lifting and regional qualification. Among them, selective bit-plane lifting for ROI will increase the number of overall bit-planes and decrease the overall coding efficiency. And because the coding part and the decoding part need to increase bit-planes and restore modules respectively, the complexity in computation will increase. Regarding ROI as a separate slice through FMO technology and using a smaller quantization parameter (QP) can also improve ROI quality. It should be noted that, if the variation in QP is too small, the improvement in ROI may be not obvious; if the variation in QP is great, it will lead to a decrease in overall performance in rate-distortion.

This paper has presented a ROI coding method based on regarding SVM segmentation targets as the regions of interest. It uses SVM to segment the ROI targets, and the algorithm has a high resistance to distortion. Even if some slight disturbances exist in the image, such as image distortion, noise, partially blocking, etc., it can still maintain a high probability and accuracy of segmentation. Based on segmentation target, this paper has suggested that using ROI encoder to realize ROI coding regarding the region of interest as the basis for coding, making the image quality of ROI better than the background, and a better decoding quality. It can solve the conflict between coding and the channels under the conditions of low bit-rate, limited channels and low storage space, and has a good application prospect in remote sensing image analysis, object tracking, telephony, conferencing and other areas.

References

[1] Wren CR, Azarbayejani A, et al. Pfinder: Real-time tracking of the human body. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1997; 19(7): 780-785.
[2] Thakoor N, Gao J. *Automatic object shape extraction and its classification with camera in motion*. Proceedings of 2005 IEEE International Conference on Image Processing. 2005, pp. 437-440.
[3] Rajagopalan R, Orchard MT and Brandt RD. Motion field modeling for sequences. *IEEE Transactions on Image Processing.* 1997; 6(11): 1503-1516.
[4] Hao Hu, Ming Zhao Wang, Jie Yang. Adaptive Fuzzy Weighted Mean Filter. *Systems Engineering and Electronics*. 2002; 24(2): 15-17.
[5] Oten R, Figueiredo R. Adaptive alpha-trimmed mean filters under deviations from assumed noise model. *IEEE Transactions on Image processing*. 2004; 13(5): 627-639.
[6] Qun sheng Yang, Ji Wu Huang, Xian_gui Kang. Weighted mean filter histogram Based on Weighted mean. *Chinese Journal of Electronics*. 2004; 32(7): 1108-1111.
[7] Nieminen A, Neuvo Y. Comments on Theoretical analysis of the max/median filter. *IEEE Transactions on Acoustics, Speech and Signal Processing*. 1998; 36(5): 826-827.
[8] Wang X. Adaptive multistage median filter. *IEEE Transactions on Signal Processing*. 1992; 40(4): 1015-1017.
[9] Hwang H, Haddad RA. Adaptive median filters: new algorithms and results. *IEEE Transactions on Image Processing*. 1995; 4(4): 499-502.
[10] Lei Jian Liu, Jing Yu Yang. Object recognition Based on Integration of Information. *Pattern Recognition and Aritficial Intelligence*. 1993; 5(3): 28-33.
[11] Jian cheng Luo, Cheng hu Zhou, etc. Support Vector Machine for Spatial Feature Extraction and Classification of Remotely Sensed Imagery. *Journal of Remote Sensing*. 2002; 6(1): 125-131.
[12] Cha Pelleo, Hanffer P and Va Pink VN. Support vector Machines for Histogram-based Image Classification. *IEEE Trans. on Neural Network.* 1999; 10(5): 1055-1064.
[13] De Hua Peng, Rui Ming Shen, Tong zhen Zhang. The Segmentation Techniques and Advances in Content-Based Retrieval. *Computer Engineering and Applications*. 2003; 13(3): 94-97.
[14] Cannon RL, Dave J, Bezdek JC. Efficient implementation of the fuzzy c-means clustering algorithms. *IEEE Trans. on Pattern Analysis and Machine Intelligence*. 1986; 8(2): 248-255.
[15] S Han and J Woods, Adaptive coding of moving Objects for very low bitrates. *IEEE Journal on Selected Areas in Conununieations, issue on very low bit-rate coding*. 1998; 5(1): 576-583.
[16] Jaralikar SM, Aruna M. Case study of a hybrid (wind and solar) power plant. *TELKOMNIKA Indonesian Journal of Electrical Engineering*. 2011; 9(1): 19-26.
[17] Daut I, Irwanto M, Suwarno, et al. Potential of Wind Power Generation in Perlis, Northern Malaysia. *TELKOMNIKA Indonesian Journal of Electrical Engineering*. 2011; 9(3): 575-582.
[18] A Cavallaro, O Steiger, and T Ebrahimi. Semantic coding analysis for adaptivce content delivery and automatic description. *IEEE Trans. Circuits Syst. and Technol*. 2005; 15(10): 1200-1209.
[19] TMS320C6000 Image/ Processing Library. Oct. 2002: Texas Instruments Incorporated.
[20] CCS help contents: TMS320C6000 chip support library API help: Texas Instruments Incorporated. 2007.