International Journal of Signal Processing, Image Processing and Pattern Recognition
Vol.7, No.5 (2014), pp.93-104
http://dx.doi.org/10.14257/ijsip.2014.7.5.09

A Spatiotemporal Attention-Based Method for Geo-Referenced Video Coding

Jiangfan Feng[1, 2], Yi Zhu[2] and Haibin Hu[3]

[1]*Key Lab of Instrument Science and Dynamic Test, North University of China, Taiyuan 030051, China*
[2]*School of Computer Science and Technology, Chongqing University of Posts and Telecommunications, Chongqing 400065, China*
[3]*Experiment Center, China West Normal University, Nanchong 637002, China*
Correspondence should be addressed to Jiangfan Feng, fengjf@cqupt.edu.cn

Abstract

This paper focuses on the problem that video-GIS has a huge amount of data, which leads to high transmission resource consumption, and introduced attention calculation model to video GIS coding to optimize the coding method so the compression efficiency will be improved as well. Specifically, on the one hand, it will use optical flow technique to calculate the specific location and motion vector of the movement point set of each frame, and reduce the amount of the movement point set integrate with the features of video GIS, then calculate the mask matrix of the foreground movement, and finally compute the Discrete Cosine Transform (DCT) compression integrate with the mask matrix to realize the optimization of the frame macro block level coding. On the other hand, it will calculate the attention of the video frame using the specific location and the motion vector of the movement point set as the primary indicator, and optimize the coding of the video of frame-level according to the calculation result. By experimental verification, the efficiency of video GIS compression is enhanced by the introduction of the features of the video GIS and the cognitive attention theory.

Keywords: *Geographic Information Systems, Video-GIS, Video Streaming, Video Coding, Attention*

1. Introduction

With the growth of geographic information system (GIS) whose major growth area is the convergence between GIS and multimedia technology, a new paradigm named video-GIS emerged [1-3].

Video-GIS can provide users with intuitive and detailed spatial information via video data. But video data has the problem of having huge amount of data which leads to high transitions resource consumption. To solve this problem, the effective way is to focus on the optimizing of the coding method, in order to achieve the goal that the video redundancy is reduced while the useful information of the video can still be retained [4]. The most widely used way in the present study is to optimize the coding method with the features of visual attention in order to improve the video compression efficiency. But the video-GIS have its own characteristic that only with a better use of it can the compression efficiency be improved.

Geo-Referenced video is fundamental processes in video-GIS development. Prior research activities on Geo-Referenced video technologies and applications have been conducted. Most

ISSN: 2005-4254 IJSIP

of them make use of video and GPS sensors. In [5][6] Stefanakis and Klamma proposed a unified framework for hypermedia and GIS. Pissinou [7] explored topology and direction under the proposed Geo-Referenced video. The work of Hwang and Joo [8] defined the metadata of Geo-Referenced video, which support interoperability between GIS and video images. In the field of Geo-Referenced video search, Liu [9] presented a sensor-enhanced video annotation system, which search video clips for the appearance of particular objects. Arslan proposed the use of geographical properties of videos [10] while Wang gave a method of time-spatial images to extract the basic movement information [11]. Although single media have been studied extensively, its semantics in geographic space are poorly understood. This paper will introduce the attention theory combined with the features of geographic space into video coding, to optimize the coding method based on the existing coding standards, and improve the compression efficiency of GIS video.

2. The Features of Attention-based GIS Video

Different from the information express method of traditional "to see the world from the bird's view", GIS video provides a more easily accepted "to see the world from the side view" perspective [12]. Because this perspective is the same with human's perspective, the attention mechanism of GIS video to information is very similar to the attention mechanism of people's view to the world. The similarities are as follows:

1) Continuity of contents. The contents of GIS video are usually continuous scene-change and angle-switching. This is the same with human's visual information.
2) Instability of background. Like human sight, GIS video's visual transition transits as itself as the reference. During the transition, the background and foreground both moves while the human visual system can quickly distinguish between foreground and background to determine its attention center instantly

At the same time, GIS video has the features different from human visual attention. The differences are as follows:

1) The primary factor which causes human visual attention is the movements of foreground. For instance, the object which suddenly appeared in sight will be specially noticed when people moves. However, the main information in GIS video is the geographic coordinate information corresponding to the video pixel, and this geographic coordinate information directly relate to the background with the video camera moves.
2) Human visual attention center is relatively unstable. As people's subjective selection and scene changes, the attention center will transfer disorderly. However, GIS video's attention center is mainly decided by the video's use. And because of the strong purpose of video-recording, the video attention center is usually the center of the video.

With the features of GIS video and human visual attention, the writer will build GIS video attention model which should solve the following problems:

1) How to reflect the movement characteristics of the video effectively?
2) How to distinguish between foreground and background effectively based on the psychological principles of the attention mechanism?
3) How to ensure the attention adapts to the changes of the background?
4) How to quantify attention and try to match the actual situation as much as possible?

3. The Attention Calculation Model of GIS Video

3.1. Motion Feature Modeling

To set up an attention calculation model on GIS video we should first extract the motion feature in the video. A good motion feature model relates to the fact whether the final established attention model can reflect the actual situation or not. In addition, a good motion feature model is also directly relates to the accuracy of the background separation and the attention quantification. Therefore, on the one hand, the motion feature model should be fully integrated with the features of attention-based GIS video; on the other hand, it should provide effective support to the follow-up work.

Video motion feature modeling needs to extract the motion feature, the main work is to detect moving objects. Generally, there are four current detection methods: optical flow calculation method [13], adjacent frame motion analysis method [14], motion energy detection method [15] and background subtraction method [16]. Optical flow method is suitable in the case of the moving camera motion detection, and it is sensitive to the moving position and direction of the moving target, but the calculation is more complex than others. Adjacent frame motion analysis method can detect complex motion better, but cannot divide moving objects effectively. Motion energy detection method can eliminates the interference of the background noise, but it is only sensitive to particular direction motions. Background subtraction method has best real-time simple computing, but is only applicable to the stationary camera situation.

The camera position of GIS video is always in motion, and usually the foreground and background move at the same time, while the background moves more frequently. Therefore this paper extracts the moving features of the video frames cyclically as follows:

1) Using matrix segmentation method to divide the i-th frame of the video image into several sub-blocks. This ensures that the corners are uniformly distributed on the image when the features of the image are being scanned (This paper uses the 4×4 matrix considering the symmetry of video image and the complexity of calculation).
2) Detect known corner scales of all the sub-blocks in proper order. When the corner scale is smaller than threshold N_{min}, re-scan the corners of this sub-block and the newly-get corners will be added into the corner sets of this sub-block. This can solve the problem that features got lost easily because the background content in GIS video updates frequently.
3) Merge the corner sets in all sub-blocks and generates the corner sets of the whole image as in formula:

$$D_i\left\{\left(x_{d1},y_{d1}\right),\left(x_{d2},y_{d2}\right).....\left(x_{dj},y_{dj}\right)\right\} \tag{1}$$

4) Calculate set D_i, the i-th frame and the i+1th frame using optical flow method to get the location set as in formula :

$$P_i\left\{\left(x_{p1},y_{p1}\right),\left(x_{p2},y_{p2}\right).....\left(x_{pj},y_{pj}\right)\right\} \tag{2}$$

5) Get the motion vector set from set D_i and set P_i as in formula :

$$V_i\left\{\left(x_{p1}-x_{d1},y_{p1}-y_{d1}\right).....\left(x_{pj}-x_{dj},y_{pj}-y_{dj}\right)\right\} \tag{3}$$

6) If i is not the last frame, then $D_{i+1} = P_i$, and executes from step 2); if is the last frame, then end the cycle.

From the experiment it is known that if the corner scale threshold N_{min} is bigger than 50 it will get a lot of useless corners while the number of the actual useful corners is not greatly improved. Considering the complexity and precision of the algorithm, the value of N_{min} is set at 50.

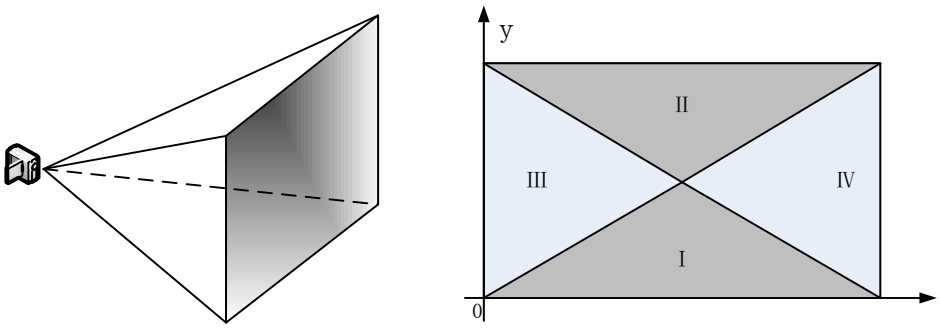
From the above process the specific location set D_i and motion vector set V_i of all the frames' motion points can be got. However, human visual system's judgment to movement characteristics is often based on the closed complete targets. Background targets in the GIS video images are often much larger than the foreground targets. Once the background moves, there will be both large detected motion point set scale and large proportion of background features. So if the motion point set scale in the GIS video is larger, then there is greater likelihood of background motion; otherwise there is greater likelihood that the movement occurred in the foreground. Due to the differences of filming conditions and video qualities, motion point set scale is a relative concept. Thus this paper calculates the scale of the motion point set in each frame of the video based on the vector set element normalization method, the scale of the motion point set in the i-th frame will be recorded as $Scale_i$, and $Scale_i \in (0,1)$.

3.2. Macro-block Level Attention Calculation Model

One of the features of GIS video attention is the focus of background information; foreground targets often do not have the value of attention. Therefore, we need to separate the foreground from the background, and then label the foreground region in the video image frames, in order to do macro-block level attention calculation to the video frames.

The two main ways for human attention mechanism to distinguish between advertent content and non-advertent content are data-driven and concept-driven [17]. When there is a specific attention target in a person's mind, the brain will compare the visual information with the transcendental template one by one to identify the advertent region, for example, to search for a particular word from a paragraph. When there is only one advertent object the brain has to pay attention to, it will classify the visual information in accordance with the possible expectations, and thus find the object with equal expectation, such as to find a figure in this paper.

The currently used background separation method is usually data-driven, which uses the known pre-built template matching the background image information. However, various as the GIS video background images are, and also without fixed content, it may result in large amount of calculation and low accuracy by only using templates to match and separate the content of the GIS video. In this paper, concept-driven method and data-driven method are combined to separate the background, which may decrease the calculation complexity while ensuring that the foreground motion and background motion will be effectively separated.



(a)Schematic Diagram of Camera FOV

(b)Schematic Diagram of FOV Segmentation

Figure 1. Schematic Diagram of FOV Law

The camera position movement has certain regularity in video GIS, also the background motion generally follow the symmetry rule. As shown in Figure 1 (a), the field of the camera view is gradually spread into the distance, but the shot of video GIS is usually a process both front and back, the rule show in video images as the background motion of region I, II, III in Figure 1 (b) frequently are symmetrical. Therefore as the background change, the motion rule can be foreseen, and according to that divide the video region to nine sub-regions as show in Figure 2, the background motion occurs in each sub-region is relatively stable.

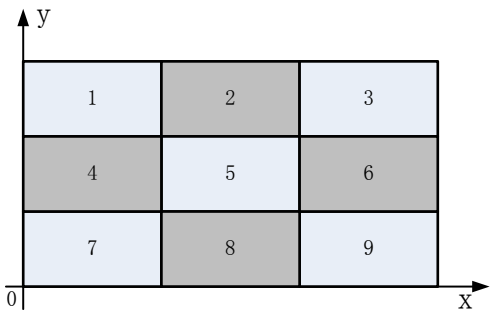


Figure 2. Schematic Diagram of Image Segmentation

The size of the motion points set D_i in the video GIS directly relate to the possibility of background motion, according to this concept; dispose variously to various size of motion points sets. When the size of motion points set is greater than the threshold value N_{Scale}, decide the background is moving and calculate the mean vector of the motion points in the sub-region as function:

$$\vec{V}_{avg} = (\bar{x}, \bar{y}) = (\frac{\sum_{i=0}^{n} X_i}{n}, \frac{\sum_{i=0}^{n} Y_i}{n}) \quad (4)$$

Each sub-region mean vector $\vec{V}_{avg}$ can represent the travel direction and displacement of the background motion in this region, hence by using value vector to decrease the background feature points in the motion points sets, that lead to narrow the size of motion points sets. The expression aims at decreasing each motion vector shown as:

$$\vec{V}_i = \left(f\left(x_i, \bar{x}\right), f\left(y_i, \bar{y}\right) \right) \tag{5}$$

$f(x)$ is cutting function which as:

$$f\left(x_i, \bar{x}\right) = \begin{cases} 0 \quad , & x * x_i > 0 \text{ and } |\bar{x}| > |x_i| \\ x_i - \bar{x}, & \text{others} \end{cases} \tag{6}$$

As the motion vector value $f(x_i,x)$ and $f\left(y_i, \bar{y}\right)$ are equal to 0, decrease the point (x_i, y_i) from the motion points set. The process as the decreasing function is shown as Figure 3. In the picture $\bar{x}$ is the mean value of a, b,c,d,e,f which are separately as the motion vectors of background. Therefore, by the decreasing process it small the background feature points; on the other side enlarge the motion of the foreground objects.

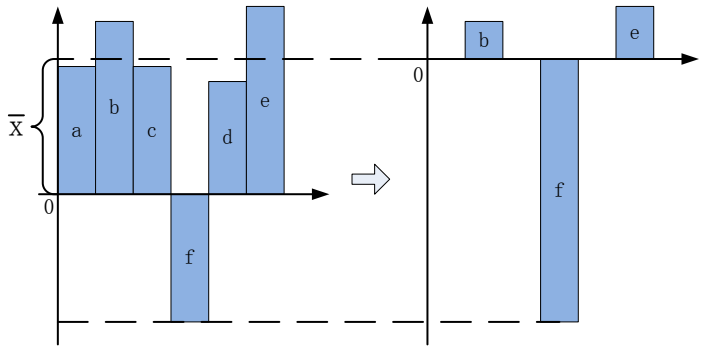


Figure 3. Schematic Diagram of Cutting Function

After decreasing all sub-regions in the video image re-calculate the size of the motion points sets, if the new $Scale_i$ is less than N_{Scale} end the decreasing process; otherwise the new $Scale_i$ is still larger than N_{Scale} re-start the decreasing process to the new sets, and halve the $\bar{x}$ parameter value of the decreasing function, then end the process.

It gets the new motion points set S_i after decreasing, and then consider the coordinates of the points in set S_i locate on the foreground objects. Because the foreground objects are a closed region, in this paper it adopts the general size of the image processing block（8X8）to block the images, when the motion point locates in one block record the one as 1 otherwise as 0, by this way it can build a label matrix which as :

$$Mask_i = \begin{bmatrix} a_{11} & \cdots & a_{1w} \\ \vdots & \ddots & \vdots \\ a_{h1} & \cdots & a_{hw} \end{bmatrix} \tag{7}$$

The height of the matrix is as 1/8 of the image pixel height and the width as 1/8 image pixel width, each element corresponding to a 8X8 pixel block of the image, as the value is equal to 1 the block consider as foreground block.

The process to establish marking matrix is shown in Figure 4.

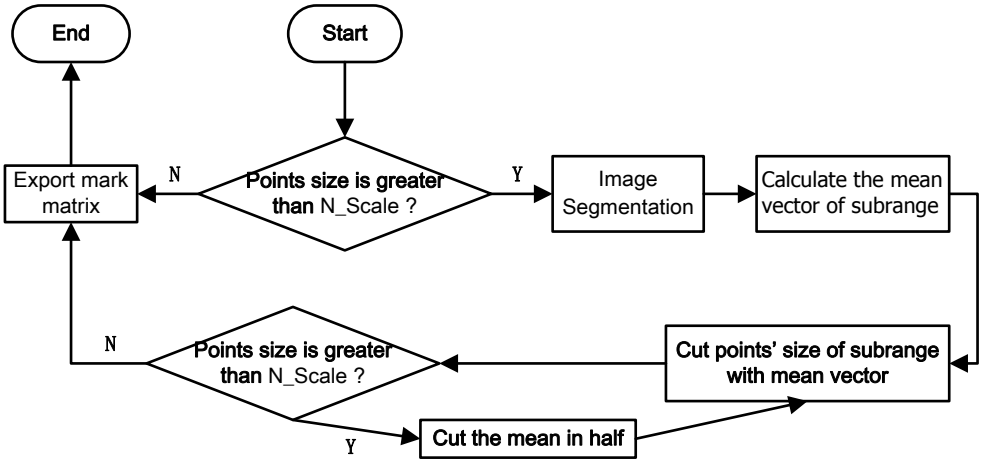


Figure 4. Building Process of Marking Matrix

Though the label matrix is not always accuracy reflect the exact profile of the foreground objects, the foreground itself is non-attentive objects. The purpose to label is narrow the coding bit of the region not remove them from the image. Besides, this method adapts well in the camera uniform motion and stationary state, usually the video GIS is quite stable.

3.3 Frame Level Attention Calculation Model

To combine the feature of video GIS and the psychological theory of human visual attention mechanism, that may acquire the background motion which the video GIS rather greater concern close to the middle of the image, also the motion state of the video directly lead to the update frequency of the background [18]. Therefore, in this paper it chooses the size of the motion points sets φ_{Scale}、the foreground size φ_{noise} and the mean value of visual center distance φ_{dis}, and the three rules are as index to calculate the frame level attention, and the function to approach the attention between frames $\varphi_{attention}$ shown as:

$$\varphi_{attention} = \frac{\varphi_{scale} \times (1 - \varphi_{noise})}{1 + \varphi_{dis}} \tag{8}$$

The size of motion points set: it is the global size of the motion points set acquires from preceding context, defined as:

$$\varphi_{scale} = Scale_i \tag{9}$$

Foreground size: the specific value between the amounts of value 1 in the foreground label matrix $Mask_i$ and element amounts defined as:

$$\varphi_{noise} = \frac{\sum_{j=1}^{h} \sum_{k=1}^{w} a_{jk}}{h \times w} \tag{10}$$

The mean value of visual center distance: the mean value of the distance that from each motion points to the central point defined as:

$$\varphi_{dis} = \frac{\sum_{i=1}^{n} \left[(X_{center} - x_i)^2 + (Y_{center} - y_i)^2 \right]^{1/2}}{n} \tag{11}$$

After analyzing all the video frames, unitization process $\varphi_{attention}$ and $\varphi_{attention} \in (0,1)$. $\varphi_{attention}$ Represents the global attention of video frames, as the size of background motion gets greater and closes to the center, it means the frame attention values great.

4. GIS Video Coding Method Based on Attention

Large amounts of communication bandwidth and memory space are occupied to transmit and store GIS video on account of its enormous data volume. Video compression can be achieved by a specific video coding method which can remove the video data redundancy. H.261/AVC [19] is a main video coding standard; it has high network adaptability and high compression-rate. Therefore, h. 261 / AVC standard is suitable for used as GIS video coding method. X264 is an open source encoder which that has the advantages of low encoding complexity and good adaptability. Considering the application environment and performance requirement of GIS, and combining with the Attention Calculation Model of GIS, this paper suggests an improved X264 coding method, which has a series of optimize rate control strategy for GIS video, to increase compression ratio.

4.1. Macro-block Level Rate Control Strategy

X264 does not has default macro-block level rate control strategy, the improved method realizes macro-block level rate control strategy by compress the frames' foreground area of marked. The process is as follows:

1) Separating frame into 8X8 macro-block, thus each macro-block can has a corresponding element of mask matrix.
2) Scan every element of mask matrix in turn. If there has a value of 1 at matrix edge or near the value of 0, mark the macro-block which corresponding to that element, and set the value of element to 0. If there is no such element, this process is completed.
3) Denoising marc-block of marked based on DCT, which can concentrate image signal energy in the low frequency coefficient.
4) Concentrate the marked area signal energy in the low frequency coefficient with DCT [20], then set the high frequency coefficient, which is less than Qp_{block}, to 0.
5) Doubled the value of Qp_{block}, and go to step 2).

The above process can effectively compress the foreground area, and it also can reduce the blocking effects and mask error.

4.2. Frame Level Rate Control Strategy

X264 encoder controls the video rate with the average rate control strategy. The rate of current frame was controlled using the feedback information of latest rate statistics, so the rate of current frame was based on latest average rate. Based on that, this paper suggests a quantitative parameter calculation method with the definition as

$$Qp_{frame} = Qp_{frame} + (1 - \varphi_{attention}) \times \varepsilon \quad (12)$$

Among them, Qp_{frame} represents the default value of quantitative parameter whose value is usually 26. $\varphi_{attention}$ means the calculation result of frame level attention model. And ε is the regulation parameter. This quantification method can effectively keep the geographic

information of video data and reduce rate.

5. Experiment and Analysis

In order to verify the feasibility and availability of the proposed method, attention value and foreground matrix are extracted from a test video of on-board recording. The video clip falls into five phases, as show in Figure 5.

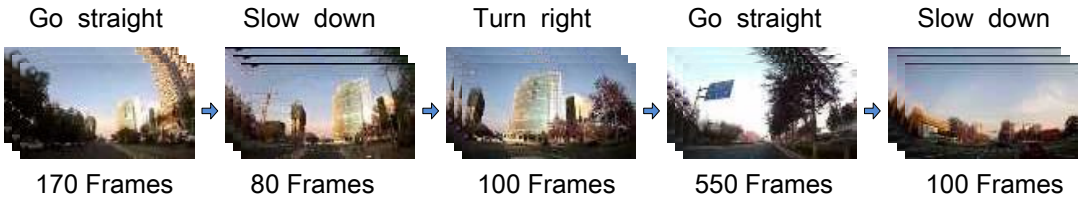


Figure 5. Schematic Diagram of Video Clip

According to the proposed method, the test video is transcoded into a new video stream which meets demand of GIS applications and services.

5.1. Video Analysis

With the video analysis based on macro-block level attention calculation model, the number of moving points and foreground points of each frame are achieved. Using the test video data, the moving region size and foreground size statistics of each frame are shown in Figure 6.

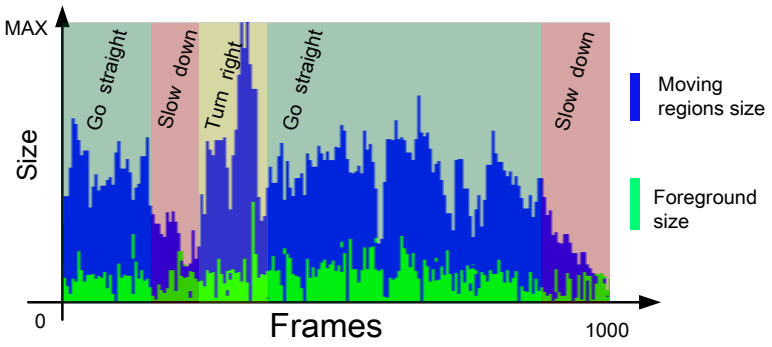


Figure 6. Schematic Diagram of Motion Scale

As can be seen from the above illustration, the moving region size has obvious change, but the change of foreground size is relatively stable. It means the number of moving foreground objects is stable, and the background motion follows the camera. In fact, this is consistent with the reality environment.

Foreground matrix is consistent with moving foreground objects or noise area. The features of moving points and foreground points from frame 891 are show in Figure 7. As we can see, the blue rectangle marked the foreground area, and most feature points of the background have been ruled out. Therefore, each frame size can be compressed by reducing the image quality of the marked area.

Figure 7. Schematic Diagram of Mask Matrix

With the analysis based on frame level attention calculation model, we can get the attention value of each frame. The attention statistics of each frame from the test video clip are shown in Figure 8.

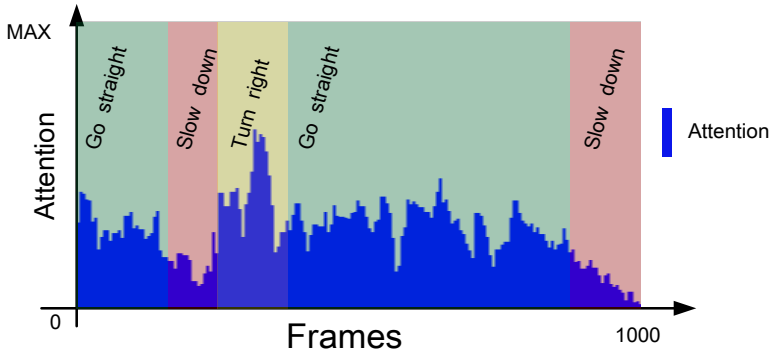


Figure 8. Schematic Diagram of Frame Attention

Since the attention value is based on each frame's background motion and the barycenter of moving points, the video phases with complex background motion have higher value. So, it can guarantee the background motion with high resolution and foreground motion or noise with high compression ratio.

5.2. Video Coding

To validate the proposed coding optimization method of GIS video, we compare the attention coding method and default coding method. Table 1 lists the data comparison of two video files with different coding methods:

Table 1. Comparison of Recoded Files with Different Methods

Coding Method	Size(KB)	Average Rate(kbps)	FPS
Attention Method	2055.4	215.2	20
Default Method	3588.3	375.1	20

From Table 1, we can see that the video file using attention recoding method has smaller size and rate than the file recoded by default method, but they have no obvious difference in FPS and video quality. With the attention recoding method, we can efficiently reduce

communication bandwidth and memory space.

Table 2 lists the comparison of average time consumption between the attention recoding process and the default process.

Table 2. Comparison of Recoding Time Consumption with Different Methods

Coding Method	Analysis(ms)	Block-Quantify(ms)	Frame-Quantify(ms)	Encode(ms)
Attention Method	132.7	6.2	10.1	10.1
Default Method	0	0	8.0	9.6

From Table 2, we observe that the two methods have similar encoding time consumption, but the attention method needs extra analysis and block-quantify time. Therefore, attention coding method can completely meet the requirements of the general GIS video service without real-time performance. Even in the personal computer, it also can realize the recoding of 20 frames per second after analyzing.

6. Conclusion

In summary, the present studies are all based on the function of GIS video, but there is not an effective method to solve the problem of transmitting GIS video. It is necessary to establish a coding method with high compression of GIS video. In this paper, an improved video coding method is proposed based on the human visual attention. Experimental results with the attention coding method of Geo-Referenced video can completely meet the requirements of the general GIS video service, but it cannot work in the real-time environment.

As for the future work, we should propose a way to improve the real-time performance of attention coding method, and then the method can meet more requirements of GIS video service.

Acknowledgments

The work is supported by the National Nature Science Foundation of China (41101432, 41201378), the Natural Science Foundation Project of Chongqing CSTC (cstc2011jjA20005).

References

[1] T. Navarrete and J. Blat, "VideoGIS: Segmenting and indexing video based on geographic information," 5th AGILE conference on geographic information science, **(2002),** pp. 1-9.
[2] C. Larouche, C. Laflamme and R. Levesque, "Geo-Referenced Aerial Videography in Erosion Monitoring," The Worldwide Magazine for Geomatics, **(2002),** pp. 46-49.
[3] D. Nigel, C. Keith, M. Keith and E. Alon, "Using and Determining Location in a Context-Sensitive Tour Guide," IEEE Computer, vol. 34, no. 8, **(2001),** pp. 35-41.
[4] H. Sun, X. Chen and T. Chiang, "Digital Video Transcoding for Transmission and Storage", New York: CRC Press, **(2005).**
[5] E. Stefanakis and M. Peterson, "Geographic Hypermedia: Concepts and Systems," Lecture Notes in Geoinformation and Cartography, Springer, **(2006),** pp. 1-21.
[6] R. Klamma, M. Spaniol, M. Jarke, Y. Cao, M. Jansen and G. Toubekis, "A Hypermedia Afghan Sites and Monuments Database," Geographic Hypermedia, **(2006),** pp. 189-209.
[7] N. Pissinou, I. Radev and K. Makki, "Spatio-temporal Modeling in Video and Multimedia Geographic Information Systems," Geoinformatica, vol. 5, no. 4, **(2001),** pp. 375-409.
[8] J. In-Hak, H. Tae-Hyun and C. Kyung-Ho, "Generation of video metadata supporting video-GIS integration," Image Processing, ICIP'04, 2004 International Conference on, vol. 3, **(2004),** pp. 1695-1698.
[9] X. Liu, M. Corner and P. Shenoy, "SEVA: sensor-enhanced video annotation," Proceedings of the 13th annual ACM international conference on Multimedia, **(2005),** pp. 618–627.

[10] A. Ay, S. Kim, S. H. Zimmermann, "Relevance ranking in Geo-Referenced video search," Multimedia systems, vol. 16, no. 2, **(2010),** pp. 105-125.
[11] J. Wang and D. Yang, "A Traffic Parameters Extraction Method Using Time-Spatial Image Based on Multicameras," International Journal of Distributed Sensor Networks, vol. 2013, Article ID 108056, **(2013)**, pp. 17.
[12] J. F. Feng and G. Y. Zhu, "Research of Vehicle Navigation Based Video-GIS", Journal of Korea Spatial Information System, vol. 11, no. 2, **(2009),** pp. 39-44.
[13] J. Barron, D. Fleet and S. Beauchemin, "Performance of optical flow techniques", International Journal of Computer Vision, vol. 12, no. 1, **(1992),** pp. 42-77.
[14] A. Lipton, H. Fujiyoshi and R. Patil, "Moving target classification and tracking from real-time video", In: Proc of WACV'98, **(1998),** pp. 8-14.
[15] R. P. Wildes, "A measure of motion salience for surveillance applications", In: Proc of Image Processing, **(1998)**, pp. 102-121.
[16] C. Stauffer and W. Grimson, "Adaptive background mixture models for real-time tracking", Computer Vision and Pattern Recognition, CV PR'99, Fort Colins, CO, **(1999)** June.
[17] L. Itti, "Models of bottom-up and top-down visual attention", California Institude of Technology, **(2000)**.
[18] "ITU-T VCEG-P07", Draft ITU-T recommendation H.264(a.k.a "H. 26L"),VCEG (SG1 6/Q6)[S],16th Meeting: Fairfax. Virginia, USA, **(2002)** May.
[19] G. Wen, D. B. Hao, S. W. Ma, "Principles of Digital Video Coding Technology", Bei Jing: Science Press .**(2010)**.
[20] K. Fukuda and A. Kawanaka, "Reduction of blocking artifacts by adaptive DCT coefficient estimation in block-based video coding", IEEE CSVT, **(2000),** pp. 969-972.

Authors

JiangFan Feng, he received his B.S. degree from Southwest Agricultural University, and his Ph.D. degree from Nanjing Normal University, in 2002 and 2007. His main research area includes spatial information integration and multimedia geographical information system.

Yi Zhu, he was born in 1987 and studying in Chongqing University of Posts and Telecommunications, Chongqing, China. His research interests include video coding and video-GIS.

HaiBin Hu, he was born in 1980. He works as lecturer of China West Normal University. His main research area includes database and multimedia applications.

