Advanced Science and Technology Letters
Vol.43 (Multimedia 2013), pp.136-139
http://dx.doi.org/10.14257/astl.2013.43.29

A Temporal Redundancy Removal Method for Video Sequences using Feature Information

Jae-Hyun Jun[1], Min-Jun Kim[2], Yong-Suk Jang[3], Cheol-Woong Ahn[4], Sung-Ho Kim[5*]

[1,3,5*]School of Computer Science and Engineering, Kyungpook National University, jhjun@mmlab.knu.ac.kr[1], ysjangii@gmail.com[3], shkim@knu.ac.kr[5]
[2]Dept. Of R&D, A2Tec, mjkim@a2tec.co.kr
[4]Division of Digitial Contents, Keimyung College University, homepig@kmcu.ac.kr

Abstract. In this paper, we propose a feature information-based temporal redundancy removal method. The proposed method uses scenes, memorial/non-memorial blocks, and SIFT parameters as compression structure, compression units, and a compression method, respectively. This electronic document is a "live" template. Experimental results showed that the proposed method has better performance than H.264/AVC main profile in objective quality, compression ratio, and compression time.

Keywords: *Temporal redundancy; Compression; H.264/AVC; Video context; feature*

1 Introduction

Video is one of the most aesthetic valuable data in computer industries. Nowadays, people using computers satisfy their aesthetic demands through video sequences. This tendency is predicted to be accelerated more and more [6].

Generally, video data rate is huge because video sequences include dozens of still images per a second. Therefore, needs and characteristics of video, video compression methods must be necessary for storing or transmission of video sequences effectively.

Traditional temporal redundancy removal methods [4, 5] are processed by the syntax without consideration of video contents. In other words, formal compression process based on a standard is performed. These methods are influenced easily by change of illumination, rotating of camera, zoom in/out, or etc. As a result, temporal redundancy is cancelled out, and global compression performance is reduced.

In this paper, we propose a temporal redundancy method using feature information in order to solve the compression performance degradation by changing of environments. The proposed method consists of 4 processes, detecting scenes, detecting energy points, block shaping, error processing. Through each process, the proposed method can remove temporal redundancy based on video contents effectively.

ISSN: 2287-1233 ASTL

2 Feature-based Temporal Redundancy Removal Method

We use the Scale-Invariant Feature Transform (SIFT) feature descriptor [1]. SIFT consist of scale-space extrema detection, keypoint localization, orientation assignment and keypoint descriptor, that are robust changing environments as temporal redundancy removal. First, we must divide a video sequence according to scenes are criteria of changing context in order to feature-based temporal redundancy removal method.

Changing a scene means a changing context of video through movements of objects, illumination changing, or etc., is categorized as a cut boundary or a gradual. These scenes are detected by [2], the first frames of scenes are used as key frames.

A detected key frame is handled as I frames in traditional compression standards, is used as a reference for other frames by detecting local features. We define the mentioned process is a detecting energy points process, and the process uses the SIFT feature descriptor which are robust to external environments. Figure 1 shows result of energy point process.

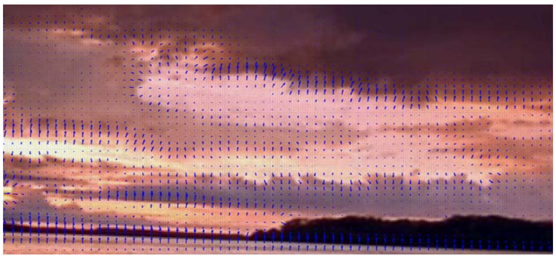

Fig. 1. Detect result of energy point process

Detected features have their location, scale, and rotation information, and have robust characteristic for illumination. These properties of features can be used as elements of video compression.

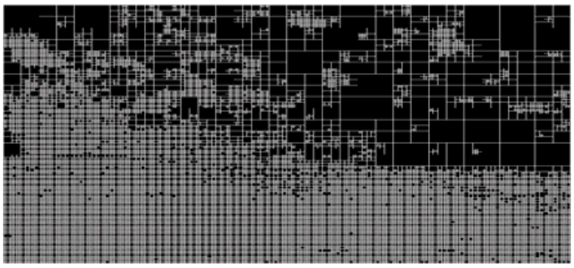

Fig. 2. Result of non-memorial and memorial blocks

In the block shaping process, each frame is divided as memorial/non-memorial blocks whether there is/are energy points or not. We use quad-tree [7] in order to find as memorial blocks and non-memorial blocks. Because non-memorial blocks which have energy points are more import than memorial blocks, extra coding process is necessary in the final step. Figure 2 shows the result of non-memorial and memorial blocks.

Finally, memorial and non-memorial blocks are compressed in the error processing process. Memorial blocks are estimated from a pixel at the same position in the previous frame. Non-memorial blocks are estimated more accurately from the SIFT parameters that are location, scale, and rotation information.

3 Experimental Results

To estimate the performance of the proposed method, we measured the PSNR (Peak-Signal-to-Noise Ratio) and compression ratio (compressed size / original size * 100). Each result was compared to the H.264/AVC main profile [3]. Figure 3 and Figure 4 represent the experimental results for the PSNR, and compression ratio, respectively.

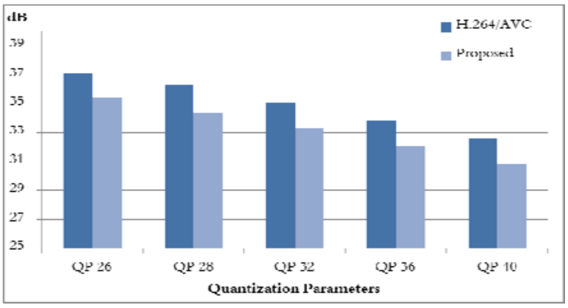

Fig. 3. PSNR

Figure 3 represent the experimental result for the PSNR. In this figure, the PSNR of H.264/AVC measured 37 db, 36.3 db, 35 db, 34 db and 32.8 db on QP26, QP28, QP32, QP36 and QP 40, respectively. And the PSNR of proposed method showed 35.3 db, 34.4 db, 33.5 db, 32.4 db, 30.8 db on QP26, QP28, QP32, QP36 and QP 40, respectively.

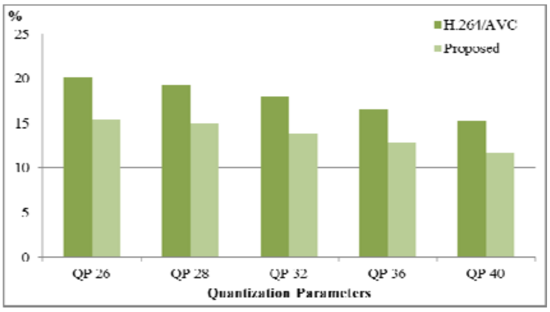

Fig. 4. Compression ratio

Figure 4 represent the compression ratio. In this figure, the compression ratio of H.264/AVC measured 20%, 19.1%, 17.6%, 16.7% and 15.4% on on QP26, QP28, QP32, QP36 and QP 40, respectively. And the compression ratio of proposed method showed 15.35%, 15%, 14.03%, 13.2% and 12.5% on QP26, QP28, QP32, QP36 and QP 40, respectively.

The proposed method showed 3.57 ~ 4.65% better than the H.264/AVC in compression ratio. Although the proposed method had lower performance in PSNR, it was very weak and was hardly noticeable in human visual system. The average spent time in the proposed method was about 11.66 % less than in the H.264/AVC.

4 Conclusion

In compression structure (scene), compression units (memorial/non-memorial blocks), and compression parameters (SIFT parameters), we propose the new temporal redundancy removal method from the traditional method based on syntax. The performance of our method was estimated on the PSNR, compression ratio, and compression time comparing to the H.264/AVC main profile, and the proposed method showed better performance than the H.264/AVC in compression ratio and compression time. The proposed method is expected to achieve more accurate performance by verification of each functional module.

Acknowledgments. This research was financially supported by the Ministry of Knowledge Economy (MKE) and Korea Institute for Advancement of Technology (KIAT) through the Research and Development for Regional Industry (A001100462).

References

1. D. G. Lowe, “Distinctive Image Features from Scale-invariant Keypoints,” International Journal of Computer Vision, Vol. 60, No. 2, pp. 91-110, Jan. 2004.
2. S. Lian, “Automatic Video Temporal Segmentation based on Multiple Features,” Soft Computing – A Fusion of Foundations, Methodologies and Applications – Special Issue on Digital Information Forensics, Vol. 15, No. 3, pp. 469-482, Mar. 2011.
3. ITU-T SG16 and ISO/IEC JCT1/SC29/WG11, “Information Technology – Coding of Audio-Visual Objects – Part 10: Advanced Video Coding,” ISO/IEC 14496-10, 2012.
4. ITU Telecommuncation Standardization Sector, “www.itu.int/en/ITU-T/pages/default.aspx”.
5. The International Organization for Standardization, “www.iso.org/iso/home.html”.
6. Cisco Systems, Inc., www.cisco.com
7. H. Samet, "Application of Spatial Data Structure," Addison-Wesley, New York, 1990

