2012 Eighth International Conference on Intelligent Information Hiding and Multimedia Signal Processing

An Improved Video Mosaic Block Detection Method Considering the Boundary of Macroblock

Quanwu Dong
Institute of Computer Science and Technology, Peking University
Beijing 100871, China
Email: dqw@pku.edu.cn

Tong Zhou
Zhongguancun Haidian Sci. Park Postdoctoral Workstation
Peking University Founder Group Co., Ltd
Institute of Computer Science and Technology, Peking University
Beijing 100085, China
Email: zhou.tong@founder.com

Jianguo Xiao
Institute of Computer Science and Technology, Peking University
Beijing 100871, China
Email: jgxiao@pku.edu.cn

***Abstract*—This passage proposes a new method to detect mosaic not only using the Y (luminance) component in YUV color space of videos, but also using the U (chrominance) and V component. The mosaic effect is measured by the boundary pixel difference from the neighbor macroblock. Instead of detecting the existence and position of mosaic blocks as traditional methods do, this method focuses on the statistics of the number of suspected mosaic blocks so that the quality of the whole frame affected by mosaic can be assessed. Experimental results show that the new method has good performance on fallout ratio, omission factor and computational complexity.**

Keywords-mosaic;video;TV;macroblock.

I. INTRODUCTION

With the rapid development of Digital TV Systems [1], some video defects include blockiness [2]-[4], blur, noise [5], et al. become the main factors that degrade the experience of the views. Mosaic is a special kind of blockiness. While watching TV, sometimes we may find several squares and lattices in the picture, accompanied with 'kaka' sounds. This is a typical scene of mosaic defects.

Mosaic usually occurs when videos cannot be decoded correctly due to the loss of information during transmitting. Then the content of some macroblocks is missing, and the video frame is wrongly displayed. Such information loss may be caused by weak digital signal, attenuation of satellite signal, electromagnetic interference, low wiring standard, fault of set-top boxes, et al. Some mosaic blocks are added artificially. For example, to protect the privacy of a person, the face may be blurred deliberately by mosaic blocks.

It is important to detect the mosaic blocks in images and videos accurately in digital video production and broadcasting organizations. As [6] mentions, the mosaic block detection methods fall into two categories: edge domain and spatial domain.

Edge detection is the first step of the mosaic detection methods in edge domain. The most widely used edge detection operator is Canny operator, which is considered the best edge detection operator until now. After the edges have been recognized, they can be used in the following two ways: one is based on the frequency of edge occurrence. In this way, the frequency of edge occurrence of every row and every line is calculated by statistical techniques. The rows and lines that have the peak frequency of edge occurrence are marked as candidates. The intersection points of the candidate lines and rows are chosen as the mosaic position [6]. The other is based on pattern recognition. The templates are used to match the result of edge detection, as shown in Fig. 1. If the four images in the template of Fig. 1 can be found, and they can form a square or rectangle exactly, then the position of the mosaic block is determined.

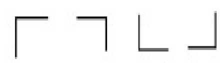

Figure 1. The templates for pattern recognition methods [6],[7]

The methods in spatial domain are based on the assumption that the mosaic block is very different from its neighbor blocks. To measure the difference between the neighbor blocks, the means of Y (luminance) component of the blocks are calculated. If the difference of the mean of a block between its neighbor block exceeds the threshold, it is supposed to be a suspected mosaic block. A spatial domain method based on SVM is proposed in [7]. A spatial method based on all phase DCT / IDCT interpolation filter color space is proposed to eliminate mosaic in [8].

However, the methods above are not suitable for the workflow of TV industry due to their high computational complexity, fallout ratio or omission factor. Therefore, we propose an improved spatial domain method of mosaic detection. It makes full use of the chrominance information and the variance of the pixels in the blocks in addition to the luminance information and the mean of the pixels. The algorithm has already been integrated to the video defect detection application systems used in the workflow of TV stations.

The rest of the paper is organized as follows. Section 2 introduces The details and improvements of algorithm. Then

978-0-7695-4712-1/12 $26.00 © 2012 IEEE
DOI 10.1109/IIH-MSP.2012.53

the experimental results are shown in Section 3. Section 4 concludes this paper finally.

II. ALGORITHM OF THE NEW METHOD

The workflow of digital TV systems is very complex. Video programs are produced by TV stations or other production organizations, then broadcasted and transmitted through wireless channel, cable or satellite, in the end we can receive and watch them by TV sets. Errors in any step of the chain may degrade the final quality of videos.

The existence of mosaic in digital video greatly reduces the experience of the views. So it is a daily work for the digital video production and broadcasting organizations, like TV stations, to find out such video defects as mosaics. Reviewing videos by human resource is time consuming and of low efficiency. It is very necessary to find out a method to detect mosaic automatically. To be applicable in TV workflow, the method should also meet the following requirements:

- Fallout ratio (the ratio of the number of those non-mosaic frames which be mistaken for mosaic frames to the total frames) should be low enough so that the reviewer will not have to cope with those false warnings, and can focus on those true mosaic frames.
- Omission factor (the ratio of the number of those mosaic frames which be mistaken for non-mosaic frames to the total frames) should be low enough to make sure that true mosaic frames will not be omitted.
- Computational complexity should be low enough so that the system can process the videos as fast as possible to achieve real-time or super real-time ability since it is a daily job.

For the defect detection applications in TV workflow, the methods in edge domain have two drawbacks:

- Edge detection is time-consuming. Experiments show that it takes more than 60 ms to find out the edges for every frame of a 720x576 video using Canny edge detection operator with the CPU being Intel core E7200. The time will be much longer when using SVM (support vector machine) in pattern recognition method. That is why none of these methods can be used in daily workflow of TV stations. Because Canny edge detection is a very complex and time-consuming algorithm[9].
- The fallout ratio is high for some images since these methods only use the grayscale information. It is also found that the fallout ratio is high when detecting those images with many vertical and horizontal lines in them as shown in Fig. 2 (a) and (b). They both come from our test dataset. Fig. 2 (a) contains several Chinese characters. Fig. 2 (b) is a program about the delicious food.

The advantage of the spatial method is its low computational complexity. The fallout ratio is also low, since they are mostly based on macroblocks, and can match mosaic well. Based on the spatial method in [6], we propose a new mosaic detection method with three improvements.

First, the difference between a block and its neighbor is measured by U (chrominance) component and V (chrominance) component in addition to Y (luminance) component. The chrominance information improves the accuracy of mosaic detection effectively because that the chrominance difference between the mosaic block and its normal neighbor is often much greater than the luminance difference. This is proved by the experimental results.

(a)

(b)

Figure 2. The video frames with high fallout ratio: (a) Chinese characters, (b) entertainment program

Second, the difference between a block and its neighbor is measured by the difference of the pixels in the boundary of the block. The mosaic block often appears obviously different from neighbor block in the aspect of pixels along boundary.

Third, this method focuses on the quality of the whole video frame affected by mosaic blocks by counting the number of the suspected mosaic blocks. This method is optimized to find out which frame contains mosaic. The position of the mosaic in the image is not important in the defection detection workflow of TV industry.

The details of the algorithm are as the following. There are two kinds of mosaic block now. One is caused by decoding error, the other is caused by artificially added mosaic. This algorithm only detects the mosaic blocks caused by decoding error which greatly reduces MOS. The artificially added mosaic doesn't reduce MOS at all, which will not make the viewers uncomfortable. This algorithm also considers the component U and V. The mosaic detection algorithm will be applied for every frame of the video.

A slide window of macroblock is used during the process. The window starts at the macroblock located in the 2nd row and the 2nd colume. When the mosaic detection algorithm for a macroblock finishes, the slide window moves one macroblock right horizontally to find the next macroblock to process. On reaching the right boundary of the frame, the slide window moves one macroblock down vertically and then moves to 2nd macroblock in the row. The process continues until the slide window reaches the right-bottom corner of the frame.

For the macroblock in the window, First the difference between the maximum and the minimum the of Y component, U component, V component are computed respectively as(Y_m, U_m, V_m). If Y_m is below the threshold $Y_{m\text{-gate}}$, the difference between the pixels of upper boundary in the macroblock and the pixels of the bottom boundary in the above macroblock are computed respectively as $Y_v[16]$, the difference between the pixels of left boundary in the macroblock and the pixels of the right boundary in the left macroblock are computed respectively as $Y_h[16]$,if every member in $Y_v[16]$ and $Y_h[16]$ is above the threshold Ypixel, the macroblock will be marked as Y type mosaic macroblock.

The same process also is applied to U and V component. Of course U and V only have 8 pixels per side. U and V has its own $U_{m\text{-gate}}$, U_{pixel}, $V_{m\text{-gate}}$, V_{pixel}.

For a mosaic block, the content inside the macroblock will be very flat, no serious change will be found. So the difference of Y,U,V component won't be very large. Y_m, U_m, V_m will be small enough and won't exceed the threshold $Y_{m\text{-gate}}$, $U_{m\text{-gate}}$, $V_{m\text{-gate}}$.

For an ordinary macroblock, the content between macroblocks will not change dramatically, so the difference between the boundary of blocks such as Y_v, Y_h, U_v, $U_h$will be very low. But if there is a mosaic block, its luminance, chrominance or flatness will be very different from its neighbor blocks along the boundary, so that Y_v, Y_h, U_v, U_h, Y_v, Y_h, V_v, will be extraordinarily greater than usual. We set three threshold constants Y_{pixel} , U_{pixel} and V_{pixel} , which denote the threshold value of boundary difference of Y component, U component and V component respectively. If one of the three metrics above exceeds its threshold, this group is marked as a suspected macroblock group with its corresponding metrics type.

Three counters, N_y, N_u and N_v are used to get the number of suspected mosaic groups whose corresponding metric exceeds its threshold. Small amount of suspected mosaic macroblocks will not affect the subjective quality of the video too much, and can be omitted. But if there are too many suspected mosaic groups in one frame, N_y, N_u and N_v will be larger than usual. In the same way as the difference metrics Y_{pixel} , U_{pixel} and V_{pixel}, we set three threshold constants: $N_{y\text{-gate}}$, $N_{u\text{-gate}}$ and $N_{v\text{-gate}}$ to measure which amount should be regarded as mosaic frame. While computing N_y, N_u and N_v for each macroblock group, if all of the three counters N_y, N_u and N_v reach their threshold values $N_{Y\text{-gate}}$, $N_{u\text{-gate}}$ and $N_{v\text{-gate}}$ respectively, the frame will be marked as a mosaic frame. If the slide window moves to the end of the frame and one of N_y, N_u and N_v does not exceed its threshold, the frame will be marked as a non-mosaic frame.

The flowchart of the algorithm is shown in Fig. 3.

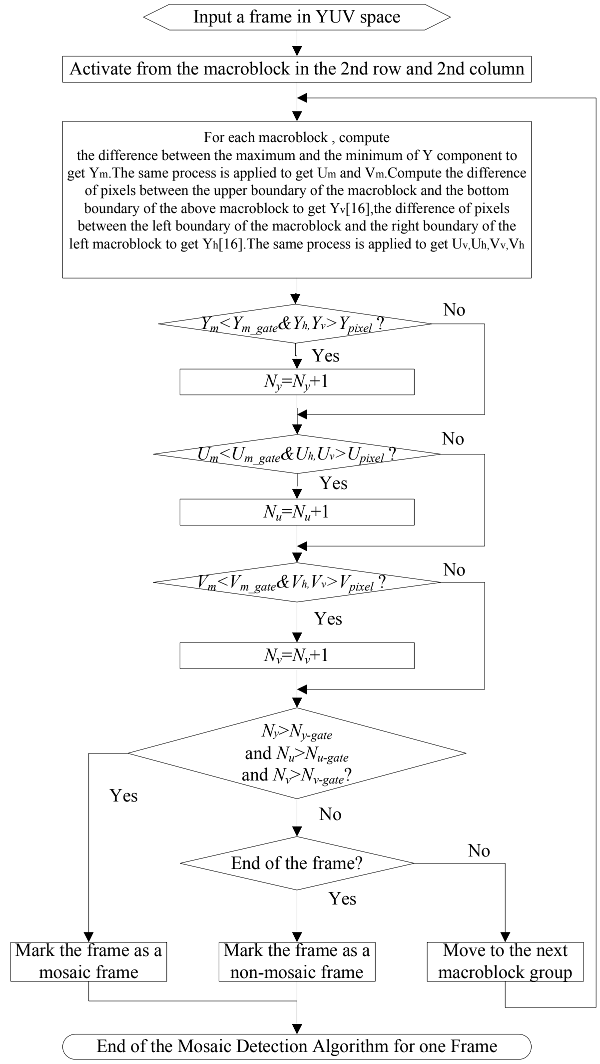

Figure 3. The flowchart of the proposed method

III. Experimental Results

To compare the performance of different mosaic detection methods, we use four video clips as test dataset. Three of them contain no mosaic frames, which are referenced as basketball, cartoon and Jiangsu TV. The fourth video comes from a program of the phoenix TV received by satellite, which contains mosaic blocks in every frame. The image resolution is 720x576. Fig. 5 lists these video clips: (a) is a cartoon with rich chrominance information and square or triangle shapes. (b) is a basketball match of NBA, which contains rapid movement. (c) is an advertisement of Jiangsu TV which contains rich chrominance information and big Chinese characters. (d) is a news clip from phoenix TV which usually contains a big background.

200 frames are extracted randomly from each of the non-mosaic videos. They are all normal frames without any mosaic blocks. 100 frames are extracted randomly from the phoenix TV clip. They are all mosaic frames. Four methods mentioned

above are applied on this test frames. They are Edge-FE (edge domain methods based on the frequency of edge occurrence), Edge-PR (edge domain methods based on pattern recognition), Spatial-Y (spatial domain methods using Y component only) and the proposed method.

The fallout ratio and omission factor of these methods are shown in table 1and 2 respectively.

(a) (b) (c) (d)

Figure 4. Test dataset: (a) basketball, (b) cartoon, (c) Jiangsu TV, (d) phoenix TV

TABLE I. FALLOUT RATIO OF THE MOSAIC DETECTION METHODS

Test video	Total frames	Our Method	Edge-FE	Edge-PR	Spatial-Y
basketball	200	0 (0%)	2 (1%)	2 (1%)	1 (0.5%)
cartoon	200	0 (0%)	5 (2.25%)	6 (3%)	3 (1.5%)
Jiangsu TV	200	0 (0%)	12 (6%)	20 (10%)	10 (5%)

Data show the number of frames. Data in parentheses are fallout ratios.

TABLE II. OMISSON FACTOR OF THE MOSAIC DETECTION METHODS

Test video	Total frames	Our Method	Edge-FE	Edge-PR	Spatial-Y
phoenix TV	100	2 (2%)	5 (5%)	8 (8%)	4 (4%)

Data show the number of frames. Data in parentheses are omission factors.

Table 3 shows the experimental results on the video clips as a whole without frame extraction using the proposed method. The fallout ratio for Jiangsu TV and the omission factor for phoenix TV are both about 1%. The fallout ratio of basketball and cartoon are only about 0.2%.

TABLE III. FALLOUT AND OMISSION FRAMES FOR ALL FRAMES IN THE TEST VIDEO CLIPS

Test video	Total frames	fallout ratio / omission factor
basketball	193130	fallout: 342frames (0.17%)
cartoon	134632	fallout: 203 frames (0.15%)
Jiangsu TV	4525	fallout: 40 frames (0.9%)
phoenix TV	2400	Omission: 24 frames (1%)

Data in parentheses are fallout ratios or omission factors.

The experimental results show that the proposed method has superior fallout ratio, omission factor compared with existing mosaic detection methods. It is also found that the processing time of this algorithm for one frame is less than 1 micro seconds. This proves that the algorithm is very fast.

IV. CONCLUSION

This passage proposed a new mosaic detection method considering the boundary of macroblock. It makes full use of the difference of Y,U,V along boundary of the macroblock to achieve better accuracy. Its fallout ratio, omission factor and computational complexity are low enough to be integrated into the defection detection workflow of TV industry. This is very helpful to improve the subjective quality of videos when watching TV. The algorithm has been put into use in a few TV stations. New model with more metrics will be employed to improve the fallout ratio and omission factor in the future.

ACKNOWLEDGMENT

This work was supported by National Natural Science Foundation of China under contract No.61101078, Doctoral Fund of Ministry of Education of China under contract No. 20110001120117 and National Key Technology R&D Program of China under Grant 2012BAH18B03.

REFERENCES

[1] Jianhua Tong, Bangjun Ding, Technology in Digital TV, Publish House of Electronics Industry, 2008

[2] Tao Su, “The research on eliminating blockiness of video communication”, Thesis of Shan Dong Univesity, 2007.

[3] K.T.Tan, Mohammed Ghanbari,”Blockiness Detection for MPEG2-Coded Video”, IEEE Signal Processing Letters, vol. 7, no. 8, August 2000.

[4] G.A. Triantafillidis, D. Tzovaras, M. G. Strintzis, “Blockiness Detection in Compressed Data”,vol.3, IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2001,pp.1533-1536.

[5] Ping Yi, Fusheng Chen, “The method to eliminate noise in image”, computer application and software, vol.5, 2003.

[6] Shuifa Sun, Shuheng Han, Gang Wang, Yichun Xu, Bangjun Lei, “Mosaic defect detection in digital video”, Chinese Conference on Pattern Recognition (CCPR) 2010.

[7] Xiaodong Huang, Huadong Ma, Haidong Yuan, “Video Mosaic Block Detection Based on Template Matching and SVM”, The 9th International Conference for Young Computer Scientists (ICYCS) 2008, pp.1082-1086.

[8] Li Li, Zhengxin Hou, Chengyou Wang, Fan He, “A mosaic elimination method based on all phase DCT / IDCT interpolation filter color space”, Optical engineering, vol.35, No.12, 2008, pp.13-17.

[9] John Canny, “A Computational approach to edge detection”, IEEE transaction on Pattern analysis and machine intelligence ,vol pami-8,NO.6, Nov. 1986, pp.679-698.