Video Mosaic Block Detection Based on Template Matching and SVM

Xiaodong Huang, Huadong Ma, Haidong Yuan
Beijing Key Laboratory of Intelligent Telecommunications Software and Multimedia,
Beijing University of Posts and Telecommunications, Beijing 100876, China
dawn_hxd@yahoo.com.cn, mhd@bupt.edu.cn, authoryhd@msn.com

Abstract

Mosaic block is easily occurred in the TV signals, which will cause the degraded video. In this paper, we propose an automatic degraded video detection approach based on support vector machine and template matching to detecting mosaic block in video. We develop a detection algorithm based on a cascade of processing steps as follows. First, we use DWT to remove the vertical high frequency of a frame and use the Canny edge algorithm to detect the frame. Second, we use the template matching to find the candidate mosaic block coordinate. Third, a SVM classifier is employed to distinguish mosaic block from the candidate ones. When the number of mosaic block is higher than a threshold, we can regard the frame as the frame with mosaic block. Experimental results show that the proposed algorithm has good performance.

Keywords: mosaic blocking, degraded video, template matching, SVM

1. Introduction

Within the last five years, there has been an explosion in the exploitation and availability of digital visual media. However, the video quality is typically degraded due to physical problems in repeated projection or playback or simply the chemical decomposition of the original material [1]. It becomes important to detect degraded video with the increasing application of visual digital media. In order to preserve and exploit this material, these defects must be detected so that the picture quality can be restored. Because of the large amount of video resource, manual detection is impractical. Therefore, automated techniques have become important.

There have been some efforts by the BRAVA consortium (http://brava.ina.fr) during 2000–2002 to attempt to standardize the community about the names and types of defects that can occur in archived video and film. An exhaustive description is given at *http://brava.ina.fr/ brava_public_impairments_list.en.htm.*

The defects detection in degraded video has been studied for a long time. The earliest work on designing an automatic system to detect dirt and sparkle was implemented by Storey at the BBC [2] as early as 1983. The design was incorporated directly into hardware which was subsequently used in-house for video restoration before broadcast. The idea was to flag a pixel as missing if the forward and backward pixel difference was high. This idea was, of course, beset with problems in parts of the image where motion occurred. The natural extension of this idea was presented by Kokaram around 1993 [3] which allowed for motion compensated differences. That type of detector was called a "Spike Detection Index" (SDI).

In 1996, Nadenau and Mitra [4] presented another scheme which used a spatio-temporal window for inference: the Rank Order Detector (ROD). It is generally more robust to motion estimation errors than any of the SDI detectors although it requires the setting of three thresholds. It uses some spatial information in making its decision. The essence of the detector is the premise that blotched pixels are outliers in the local distribution of intensity.

The image histogram has been used to detect abrupt changes in image sequences. This idea can be extended to detect very large areas of missing data in degraded archive frames. It is totally unsuitable for detection of all but the largest defects since otherwise the changes in consecutive histograms are not noticeable. In [5], Kokaram et al. present a mechanism for isolating the image quadrant that contains the defect. The attractive aspect of this proposal is that no motion estimation is necessary, hence the computational cost is extremely low.

In the one-dimensional case, Paisan and Crise [6] were the first to spot that one could use a median filtered signal as a rough estimate of the original signal before corruption by impulsive noise. The authors of [7, 8] have implemented these types of techniques for film restoration.

Mosaic block is one of the defects in the degraded video. As a result, mosaic block detection is critical to the degraded video detection. Mosaic block of the video is caused by the data loss. During the video frame compression and transmission, the data loss will occur. When the video is decoded, the block will appear as shown in Fig.1(a). We refer to the mosaic block appeared in the frame as MB (Mosaic Block).

However, most of degraded video defects detection techniques [2-7] are not suited to deal with mosaic block

978-0-7695-3398-8/08 $25.00 © 2008 IEEE
DOI 10.1109/ICYCS.2008.69

detection. So far no method, to our best knowledge, of automatically treating mosaic block detection has been proposed. This is challenging because mosaic block are relatively uncontrolled with a wide variability of scale, color, texture. Thus, we try to study an automatic degraded video detection approach based on support vector machine and template matching to detect mosaic block in video.

The rest of this paper is organized as follows. In section 2, we first introduce the flowchart of MB detection. Then the frame preprocessing and feature extraction based on template matching are described in section 3 and section 4, respectively. SVM based mosaic detection and the experimental results are described in detail in section 5. We conclude our paper in section 6.

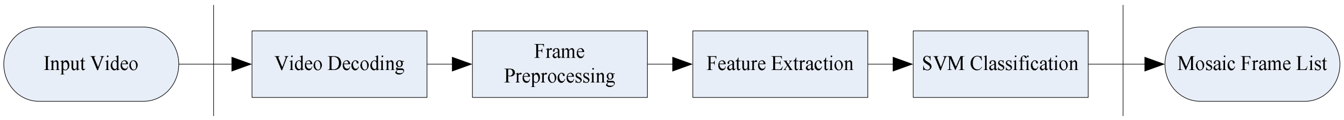
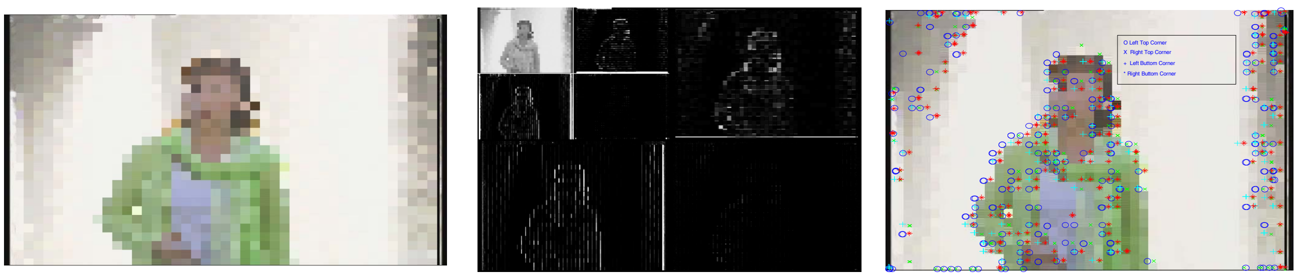
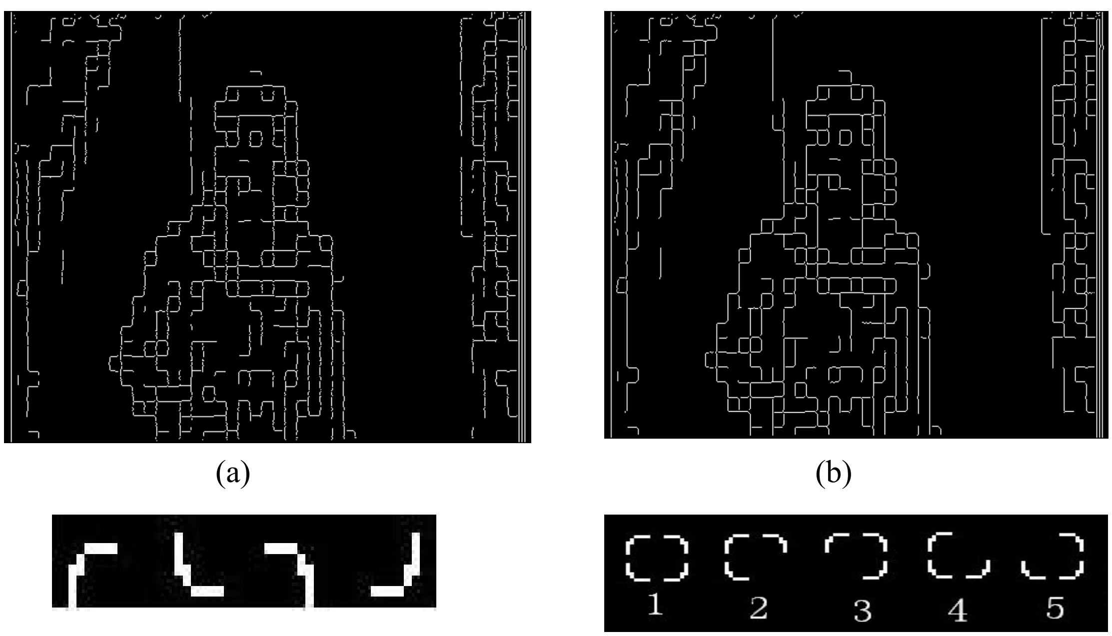


Fig.1. the Flowchart of the Mosaic Detection



(a) (b) (c)

Fig.2. (a) Mosaic Frame. (b) Mosaic 2-D Wavelet Transform Result. (c) the Template Matching Result.



(a) (b)

(c) (d)

Fig.3. (a) Canny Detect Result Without DWT. (b) Canny Detect Result With DWT. (c) The Template of the Four Rectangle Corner. (d) Mosaic Block Combination.

2. MB detection

We propose a flowchart to detect the MB in the video frame. The proposed framework is applied in the uncompressed domain of video and consists of four modules, including decoding, frame preprocessing, feature extraction and SVM-based classification, as shown in Fig. 1. The input video is first decoded into video frames. Then, we use the daubechies wavelet to filter the high frequency details of these frames and use the Canny edge detection to get the edge image. In the feature extraction, we use the template matching to get the candidate MB coordinate according to the defined template. The coordinates are calculated as features in the feature extraction module. In the SVM classification module, the features are classified by the trained SVM. Finally, according to the classification results, the frame with MB is detected.

3. Frame preprocessing

We note that the MB shows high activity in the two high-frequency subbands (HL, LH). In the LH subbands, the horizontal lines located in the MB will yield large coefficients in vertical direction, which will result in some edge perturbation in the next edge detection. As a result, we use the DWT (Discrete Wavelet Transform) to retrieve the LH subband coefficients and set them value as zero, then use IDWT (Inverse Discrete Wavelet Transform) to reconstruct two-dimensional image.

We use wavelets to decompose the image because they provide successive approximations to the image by down-sampling and can detect edges during the high-pass filtering. The low-pass filter creates successive approximations to the image while the detailed signal provides a feature-rich representation of textual content. This can be easily seen in the image decomposition as shown in Fig. 2(b). Fig. 2(a) is the original image and Fig. 2(b) is its second-level wavelet decomposition.

In the next step, the Canny edge detection fulfils the detection of the reconstructed image to get the edge image. The results without and with DWT are shown in Fig. 3(a) and Fig.3(b), respectively.

4. Feature extraction based on template matching

4.1. Find the four corners of MB

Using Canny edge detection enables us to attain the edge image. We find that the frame with MB displays the rectangle shape in the edge image. So we can use the template matching to find the four corners of the rectangle. Fig.3(c) is the template of the four corners and Fig.2(c) shows the result of the template matching.

4.2. Extract features for SVM classification

After the template matching, we can get the candidate MB corner coordinate.

According to the observation of the frame with MB, we find the MB size is under the 30×30 pixels. As a result, we use a 50×50 pixels sliding window to pinpoint the coordinates of the corner in every window. We can randomly categorize the corner coordinates into the input vectors of SVM for classification. For every candidate MB, the input vector is described by a four dimensional vector representing the bounding box of mosaic block:

$$x = (TL_x, TL_y, RB_x, RB_y)^T$$

where TL_x and TL_y are the coordinates of top-left point of bounding box, RB_x and RB_y are the coordinates of right-bottom point of bounding box, respectively.

We use the trained SVM to classify the input vectors. When the input vectors fit one of the five combinations, which are shown in Fig.3(d), we can draw the conclusion that it is MB.

After the classification, we consider using the number of positive classification (we will refer to it as Numblock) to detect the frame with MB. If the Numblock is above some threshold, we can draw the conclusion that the frame has MB. The threshold is obtained by choosing the mean of the Numblock, which can be got from the trained samples. According to our experimental results, we set the Numblock to 5.

5. SVM based mosaic detection

5.1. C-Support Vector Classification

Support Vector Machines (SVM) was developed by Vapnikand et al [9,10]. Given a training set $S = \{(x_i, y_i)\}_{i=1}^{l}$ of size l where $x_i \in R^n$ and $y_i \in \{1, -1\}$ for i=1,…,l, we can use C-SVC to solve the following problem:

$$\min_{w,b,\xi} [\tfrac{1}{2} W^T W + C \sum_{i=1}^{l} \xi_i] \tag{1}$$

Subject to $y_i(W^T \phi(X_i) + b) \geq 1 - \xi_{i,}$

$\xi_i \geq 0, i = 1, \ldots, l$

Using the C-SVM classification algorithm, we trained the mosaic detection algorithms over the training data, and evaluate each record in the testing set. We evaluate the algorithm model by computing two statistics variable: the True Positive (TP) rate and the False Positive (FP) rate. The TP is the percentage of target records that are classified correctly. The false positive (FP) rate is the percentage of non-target records, which are mislabelled.

LIBSVM (developed by Chang and Lin) is an integrated tool for support vector classification and regression. The LIBSVM is available at: http://www.csie.ntu.edu.tw/~cjlin/libsvm.

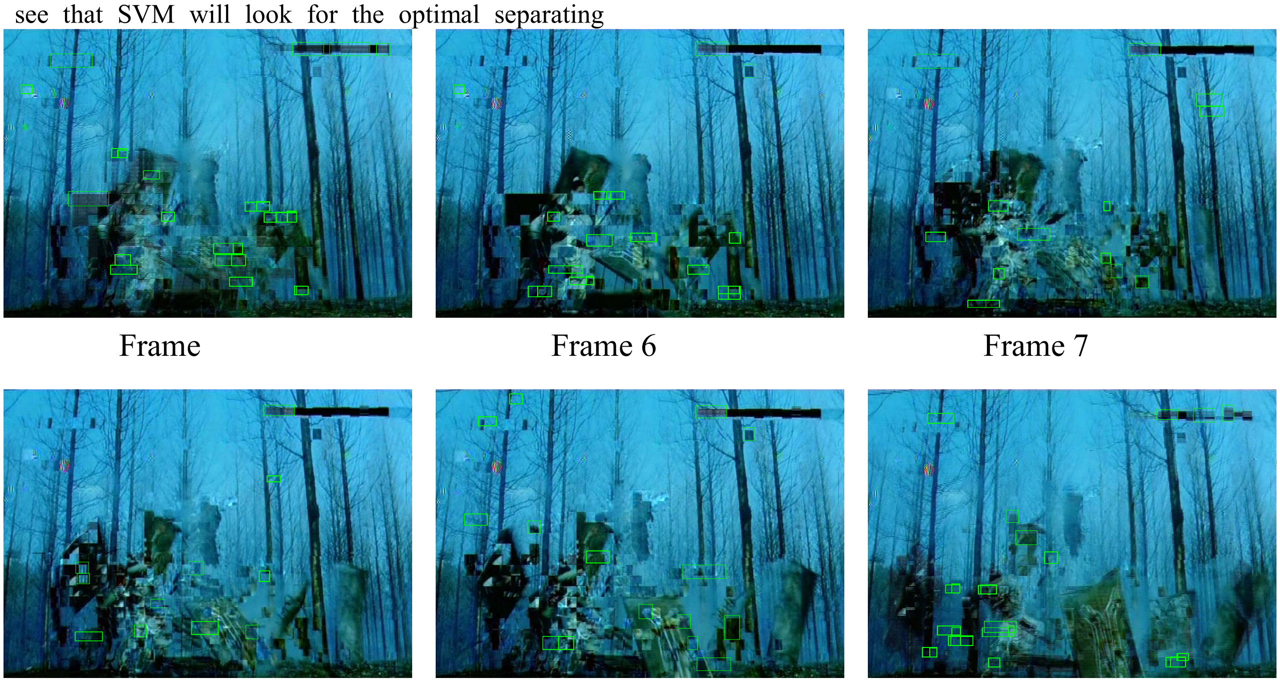
5.2. Unbalanced training data set

The described feature extraction method for selecting examples produces an unbalanced distribution of the examples in the training set. In our experiment, the number of negative examples is ten times the number of positive examples.

To the unbalanced data set, using C-SVM to classify the data can not get satisfied result. In our experiment, we use the same training set with 4800 data records. In unbalanced data set SVM can get 41 nBSV (number of Bounded Support Vectors) and 66 nSV (number of Support Vectors) and in balanced data set SVM can get 116 BSV and 156 NSV. This means it will cause higher error classification rate in the unbalanced data set than in the balanced data set.

We interpret the regularization parameter C has the cost associated to a misclassified example, then it is easy to see that SVM will look for the optimal separating hyperplane in regions of the feature space with few examples. As a result, a simple approach for facing with unbalanced training data set is to assign different misclassification costs to each class. We let C^+ and C^- are the costs we pay when misclassify a positive and negative example, respectively. Then, $C = k_1C^+ + k_2C^-$, where k_1 and k_2 are weight of C^+ and C^-. In [11], l^+ and l^- are the number of positive and negative examples in the training set, respectively, and $k_1 = \frac{l^+}{l^+ + l^-}$, $k_2 = \frac{l^-}{l^+ + l^-}$. We present a new method to compute the k_1 and k_2, then $k_1 = \frac{1}{2} + \beta$, $k_2 = \frac{1}{2} - \beta$ where $\beta = \frac{l^+}{l^-}$.

For intermediate values of C, the solution of the function (1) is a tradeoff between the maximum margin and the minimum number of misclassified points. So we use the $\beta = \frac{l^+}{l^-}$ can assign the C to an intermediate values.



Frame | Frame 6 | Frame 7

Frame 8 | Frame 9 | Frame 11

Fig. 4. The Detected Mosaic Frame

5.3. Experiment results and application

We apply the degraded video detection algorithm to the Media Assets Management, which manages the storage, distribution and content analysis of a great deal of digital visual media. We use the detection algorithm to detect the defects of the digital visual media. We select 500 images from video frames. The frame size is 720×576 pixels. The test sets consist of a variety of cases, including frame with MB and frame without MB. The algorithm performs robustly on the majority of the frames as shown in Fig. 4.

Recall and false alarm rate are used to evaluate algorithm. Recall rate is the percentage of real frame with MB detected as mosaic block. False alarm rate is the percentage of frame without MB detected as frame with MB. The 96.1% and 5.6% recall/false alarm rates show the good performance of this algorithm. Although the false alarm rate will increase in some complex background, especially when the mosaic block is mixed with background, the frame with mosaic block can be achieved through adjusting the Numblock value.

6. Conclusions

In this paper, we propose an automatic degraded video detection approach to find the blocking artifacts in video frames. Template matching is utilized as finding the candidate blocking artifacts coordinates, SVM-based classification is employed to distinguish mosaic blocking artifacts from the candidate ones. As a consequence, a unified method for automatic mosaic blocking artifacts detection based on support vector machine and template matching is constructed. This method can be widely used to solve other problems for degraded video detection, such as digital drop out, betacam dropoput [1].

Because some Chinese characters will have appeared blocking feature in edge image, our future work will focus on dealing with distinguishing Chinese character from real mosaic blocking artifacts.

Acknowledgment

This work reported in this paper is supported by the National Natural Science Foundation of China under Grant No.90612013, the 111 Project under Grant No.B08004, the Cosponsored Project of Beijing Committee of Education under Grant No. SYS100130422, Microsoft Research Asia.

References

[1] Anil C. Kokaram. “On Missing Data Treatment for Degraded Video and Film Archives: A Survey and a New Bayesian Approach”. IEEE Trans. Image Processing, Vol.13, No.3, Mar 2004.

[2] R. Storey. “Electronic detection and concealment of film dirt,” SMPTE J., pp. 642-647, June 1985.

[3] A.C. Kokaram, “Motion Picture Restoration: Digital Algorithms for Artefact Suppression in Degraded Motion Picture Film and Video”. Berlin, Germany: Springer-Verlag, 1998.

[4] M. J. Nadenau and S. K. Mitra, “Blotch and scratch detection in image sequences based on rank ordered differences,” in Proc. 5th Int. Workshop on Time-Varying Image Processing and Moving Object Recognition, Sept. 1996.

[5] A. Kokaram, R. Bornard, A. Rares, D. Sidorov, J.-H. Chenot, L. Laborelli, and J. Biemond, “Robust and automatic digital restoration systems: Coping with reality,” in Proc. Int. Broadcasting Convention (IBC), 2002.

[6] F. Paisan and A. Crise. “Restoration of signals degraded by impulsive noise by means of a low distortion, nonlinear filter”. Signal Process., vol. 6, pp. 67–76, 1984.

[7] L. Joyeux, S. Boukir, B. Besserer, and O. Buisson, “Reconstruction of degraded image sequences. Application to film restoration.” Image Vis. Comput., no.19, pp.503-516, 2001.

[8] L. Tenze, G. Ramponi, and S. Carrato, “Robust detection and correction of blotches in old films using spatio-temporal information,” Proc. SPIE, Jan. 2002.

[9] BURGES CJ C. “A tutorial on support vector machines for pattern recognition”. Data Mining and Knowledge Discovery, 1998, 2(2):121-167

[10] V.N. Vapnik, The Nature of Statistical Learning Theory, Springer-Verlag, New York, ISBN 0-387-94559-8, 1995.

[11] N. Ancona, G. Cicirelli, E. Stella and A. Distante. “Ball detection in static images with support vector machines for classification”. Image Vision Computing. 2003:21(8),675-692.