Passive Copy-Move Forgery Detection Using Halftoning-based Block Truncation Coding Feature

Bambang Harjito and Heri Prasetyo
Informatics Department of Mathematics and Natural Sciences Faculty of Sebelas Maret University, Indonesia

E-mail: bambang_harjito@staff.uns.ac.id, and heri.prasetyo@staff.uns.ac.id

ABSTRACT. This paper presents a new method on passive copy-move forgery detection by exploiting the effectiveness and usability of Halftoning-based Block Truncation Coding (HBTC) image feature. Copy-move forgery detection precisely locates the large size or flat tampered regions of an image. On our method, the tampered input image is firstly divided into several overlapping image blocks to construct the image feature descriptors. Each image block is further divided into several non-overlapping image blocks for processing HBTC. Two image feature descriptors, namely Color Feature (CF) and Bit Pattern Feature (BF) are computed from the HBTC compressed data-stream of each image block. Lexicography sorting rearranges the image feature descriptors in ascending manner for whole image. The similarity between some tampered image regions is measured based on their CF and BF under specific shift frequency threshold. As documented in the experimental results, the proposed method yields a promising result for detecting the tampered or copy-move forgery regions. It has proved that the HBTC is not only suitable for image compression, but it can also be used in the copy-move forgery detection.

1. Introduction

HBTC is a simple image compression technique which converts an input image into another image representations, namely two color quantizes and bitmap image. Formerly, HBTC is technique for grayscale image compression. However, some progresses have been made for HBTC to extend its usability in color image compression [1,7], image watermarking, and image retrieval [1-2]. As reported in some literatures, image features derived from HBTC data stream yields a promising result in some applications.

This research extends the usability of the HBTC image feature for passive copy-move forgery detection task. For this task, the tampered or forged regions are determined using similar strategy in [3-6]. The main difference between the proposed method and former schemes [3-8] is on image feature generation. In our proposed method, an image feature is simply derived from HBTC data stream since HBTC offers low computational burden in image feature computation compared to the other schemes.

The rest of this paper is organized as follow: Section II gives a brief introduction about the HBTC image compression for color image. The image feature is subsequently presented in this section.

Section III presents the proposed method for the passive copy-move forgery detection using HBTC image feature. Section IV reports the experimental results on the passive copy-move forgery detection using the proposed method. The conclusion and future directions are given at the end of this paper

2. Feature Extraction

This section gives a brief introduction of the HBTC as well as feature extraction technique using HBTC compressed data stream. The HBTC is formerly proposed for grayscale image compression. Some researches extends the usability of the HBTC for color image compression and image retrieval [1-2,8]. Inspired by its successfulness, this paper exploits the effectiveness of HBTC into the copy-move forgery detection domain. The subsequent subsections explain the basic concept of HBTC and the feature extraction.

2.1. Halftoning-based Block Truncation Coding (HBTC)

The HBTC compresses an image in effective manner by simply dividing the input image into several non-overlapping (or overlapping) image blocks. The HBTC simply encodes each image block into another representation, i.e. two color quantizes and bitmap image. Let f(x,y) be an image block of size m×n. The HBTC performs the image compression of each block f(x,y) as follow :

$$\mathcal{H}\{f(x,y)\} \Rightarrow \{q_{min}, q_{max}, b(x,y)\} \quad (1)$$

for all pixel position x=1,2,…,m and y=1,2,…,n. The q_min and q_max denote the color quantizers, i.e. min and max quantizer, respectively. These two color quantizers can be simply computed by locating the minimum and maximum values over all pixel intensities. If the input image is in color space, the two color quantizers are also in the same color space. The min and max quantizer can be simply determined using the following strategies:

$$q_{min} = \min_{\forall x,y} f(x,y) \quad (2)$$

$$q_{max} = \max_{\forall x,y} f(x,y) \quad (3)$$

The other data stream can also be obtained after HBTC encoding, namely bitmap image b(x,y). In this research, we employ the Ordered Dither Block Truncation Coding to generate bitmap image. Please refer [1-2] to detail explanation of this bitmap image generation. The two color quantizers and the bitmap image can be utilized to generate image feature descriptor in the copy-move forgery detection.

2.2 Color Feature.

The first feature, namely CF, can be easily computed from HBTC color quantizers. Figure 1 shows the schematic diagram of CF computation. The CF computation requires two color codebooks obtained from Vector Quantization (VQ) over a set of training color images. Let $C = \{c_1, c_2, \dots, c_{N_{min}}\}$ and $D = \{d_1, d_2, \dots, d_{N_{max}}\}$ be the min and max color codebook, respectively, to index the min and max color quantizer. The color indexing processes for min and max quantizers are given as follow:

$$i_{min}(\theta) = \text{argmin}_{k=1,2,\dots,N_{\min}} \|q_{min}, c_k\|_2^2, \quad (4)$$

$$i_{max}(\theta) = \text{argmin}_{k=1,2,\dots,N_{\max}} \|q_{max}, c_k\|_2^2, \quad (5)$$

for all θ, where θ denotes image block position in whole image. Subsequently, the CF for min quantizer can be obtained as:

$$CF_{min}(k) = \Pr\{i_{min}(\theta) = k, \forall\theta\}, \quad (6)$$

for $k = 1,2,\dots,N_{min}$. Herein, the feature dimensionality of CF_{min} is the same as min color codebook size N_{min}. The CF for max quantizer is given as follow:

$$CF_{max}(k) = \Pr\{i_{max}(\theta) = k, \forall\theta\}, \quad (7)$$

for $k = 1,2,\dots,N_{max}$. The feature dimensionality of CF_{max} is the same as max color codebook size N_{max}.

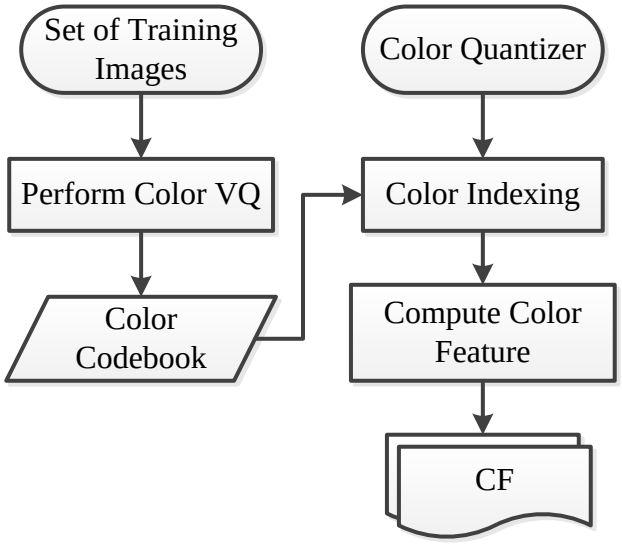

Figure. 1. Schematic Diagram of CF computation.

2.3. Bit Pattern Feature (BF)

Another feature, namely Bit Pattern Feature (BF), can be easily generated from HBTC bitmap image. Figure 2 depicts the schematic diagram of BF computation. Let $B = \{b_1, b_2, \dots, b_{N_b}\}$ be bit pattern codebook of size N_b. This bit pattern codebook can be obtained from binary VQ using soft centroid method under several training images. The bitmap indexing process of bitmap image $b(x, y)$ is formally defined as:

$$i_b(\theta) = \text{argmin}_{k=1,2,\dots,N_b} \|b(x,y), b_k\|_2^2. \quad (8)$$

The BF can be subsequently calculated using the following formula:

$$BF(k) = \Pr\{i_b(\theta) = k, \forall\theta\}, \quad (9)$$

for $k = 1,2,\dots,N_b$. The feature dimensionality of BF is identical to the bit pattern codebook size N_b.

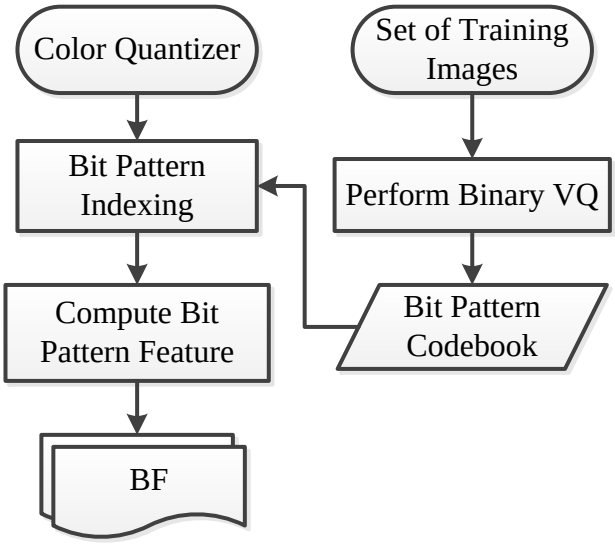

Figure. 2. Schematic Diagram of BF computation.

3. Proposed Method

This section presents the proposed passive copy-move forgery detection in detail. Figure 3 shows the schematic diagram of the proposed method. The proposed method receives the forged image containing some tampered or duplication regions on a single image. The proposed method marks the forged/tampered regions of the input image at the end of detection process. The copy-move forgery detection of the proposed method is then given as follow:

(1) Divide the suspicious image into several overlapping image blocks of fixed size.
Firstly, an input image with some tampered regions is divided into several overlapping blocks to yield some image blocks. The size of image block gives great impact at final performance.

(2) Divide an image block into several non-overlapping image blocks for HBTC processing.
Each overlapping image block from the previous step is further processed in this step. In this step, each image block is divided into several non-overlapping image blocks for HBTC processing. Herein, each image block is encoded and compressed using HBTC to yield two color quantizers (q_{min} and q_{max}) and bitmap image ($b(x, y)$).

(3) Compute an image feature obtained from the HBTC data stream.
In this step, the proposed method extracts the image features from HBTC data stream (two color quantizers and bitmap image). The CF is derived from two color quantizers , whereas the BF is obtained from bitmap image. In our proposed method, the color and bit pattern codebook sizes are selected as $N_{min} = N_{max} = N_b = 8$. To achive faster computation, we simply employ the image features with dimensionality 16.

(4) Perform similarity matching for all image blocks.
After obtaining the image features from previous step, this step performs similarity matching between all image regions based on their image features. The proposed method utilizes the lexicographically sorting to detect tampered region as often used in copy-move forgery detection [3-5]. The shift vector is also employed in the proposed method.

(5) Perform post-processing to remove the isolated regions.

To further improve the detection, an additional step is required to remove small regions or isolated image regions. Some morphological operations can be used in this step. This step produces the final copy-move forgery detected image.

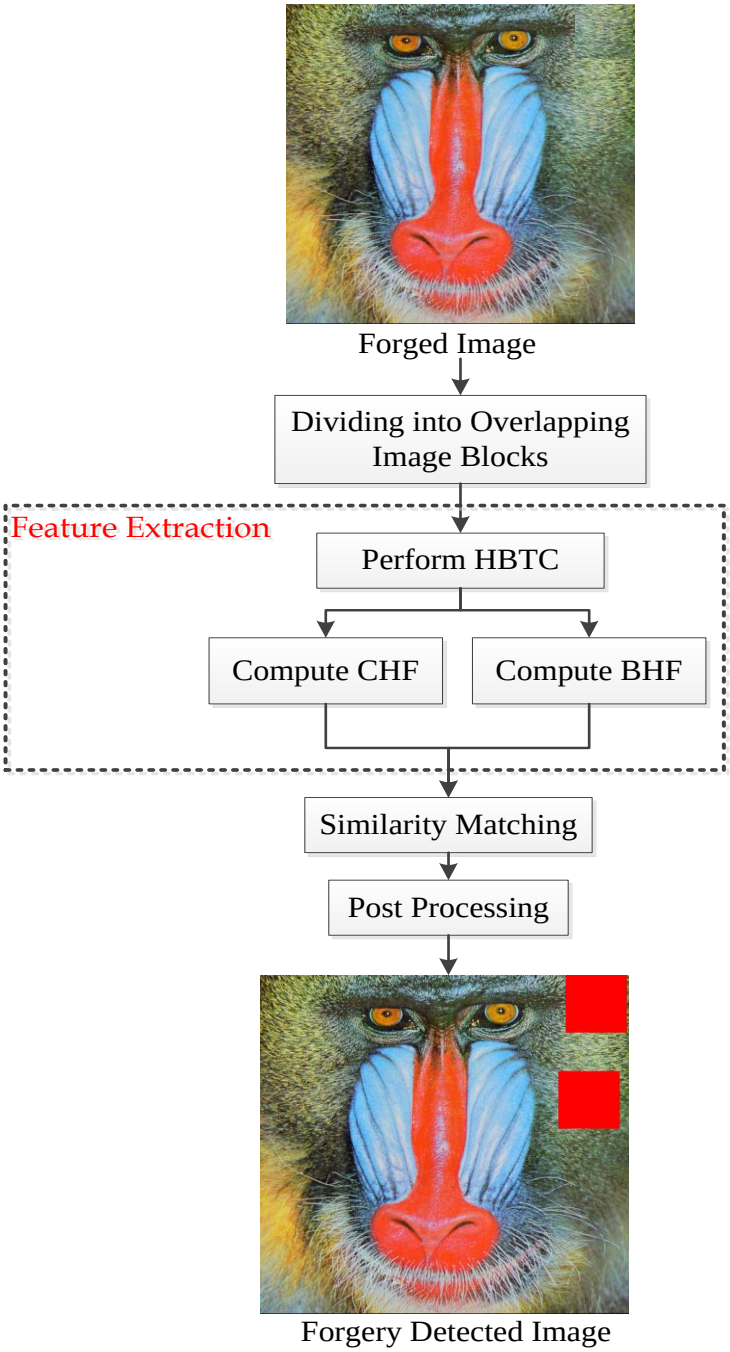

Figure. 3. Schematic Diagram of the Proposed Method

4. Experimental Results

Some experiments were carried out to investigate the usability and effectiveness of the proposed method. The performance of the proposed method is examined under three images from SIPI-USC dataset denoted as Baboon, Peppers, and Lake. The tampered regions for each image are manually determined and copied into the other region to create a forged image. The correctness of the proposed method is visually judge when the tampered region can be correctly detected by the algorithm and

confirmed by human vision. In addition, the proposed method is objectively measured using two quantitative measurements

4.1. Evaluation Metrics

To objectively investigate the performance of proposed method, two quantitative metrics can be used namely accuracy ratio (p) and false negative ratio (f). The value of accuracy ratio decreases when the proposed method detects non-forged region as forged region. The accuracy ratio (p) can be formally defined as:

$$p = \frac{|D_1 \cap R_1| + |D_2 \cap R_2|}{|D_1| + |D_2|}, \quad (10)$$

The other metric, namely false negative ratio, closes to zero while the algorithm detects almost all real forged regions as forged image. The false negative ratio (f) is formally defined as

$$f = \frac{|D_1 \cap R_1| + |D_2 \cap R_2|}{|D_1| + |D_2|} - p, \quad (11)$$

Where D_1 and D_2 denotes copied and pasted/tampered regions, respectively, in forged/fake image. The symbols R_1 and R_2 represent copied and pasted/tampered regions, respectively, detected by the copy-move forgery detection algorithm.

4.2. Effectiveness of the Proposed Method

This subsection reports the result of proposed copy-move forgery detection over three image sets. For each image, the copied regions are manually determined by user and placed into the other position to yield a forged image. The proposed method tries to detect the copy-move regions for these three images. The successfulness of proposed method is visually examined and quantitatively measured. Fig. 4 shows the detection results of the proposed method. As shown in this figure, the proposed method gives the good result on detecting the forged/tampered regions which can be easily confirmed by human vision. For Baboon image, the accuracy and false negative ratios are $p = 0.98$ and $f = 0.03$. The Pappers image produces the accuracy and false negative ratios at $p = 0.98$ and $f = 0.01$. The proposed method yields the best forgery detection result for Lake image (compared to other two images) with accuracy and false negative ratios $p = 0.99$ and $f = 0.002$. It can be drawn from this experiment that the proposed method offers a promising result in the passive copy-move forgery detection task.

4.3. Comparison with Former Existing Schemes

The additional experiments were also conducted to further investigate the performance of the proposed method. Herein, the proposed method is objectively compared to the former scheme [5] over all three images. To make a fair comparison, both methods try to detect the forged regions under an identical feature dimensionality, i.e. 16. The accuracy and false negative ratios are calculated for each image. The average value is subsequently computed over all image tests for both methods. Table I summarizes the performance comparison between the proposed method and former scheme [5]. The proposed method outperforms the former scheme in the passive copy-move forgery detection problem. It proves that the HBTC feature is also suitable for forgery detection, not only for image retrieval and compression.

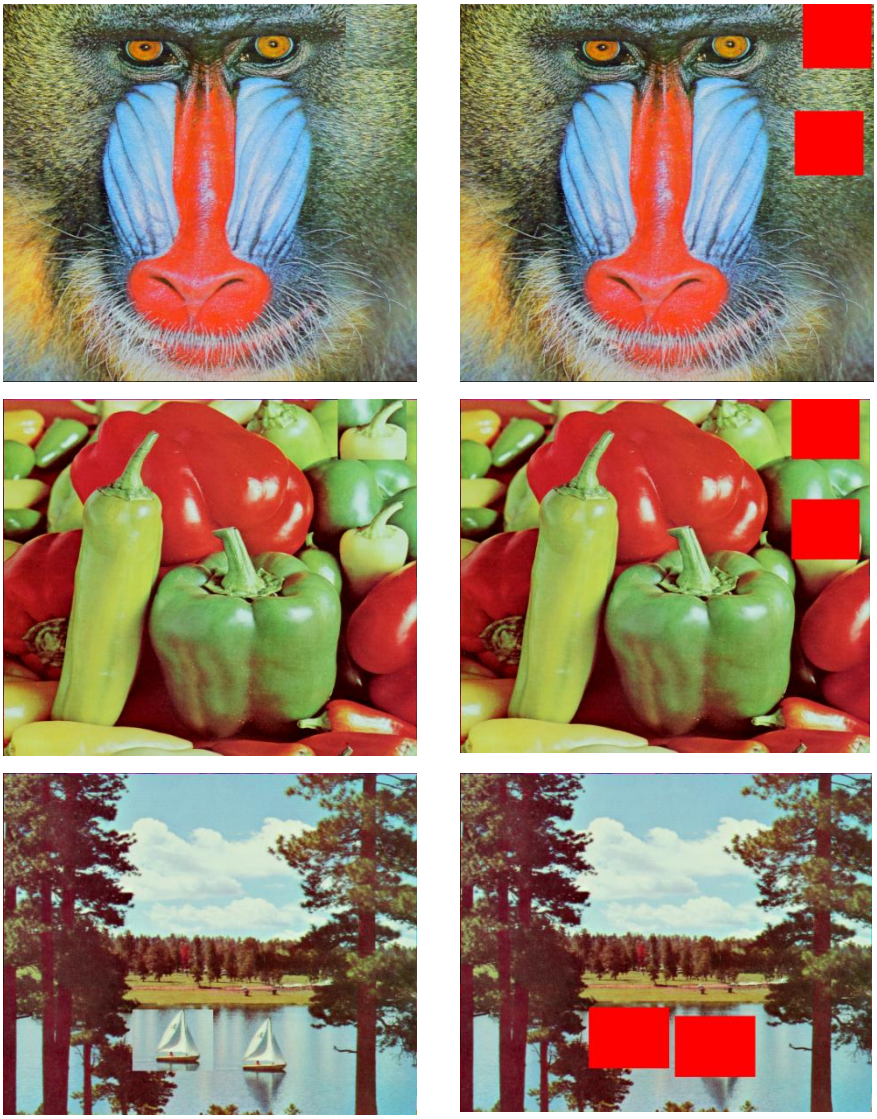

Figure. 4. The results of the copy-move forgery detection using the proposed method for three images. First column shows the forged image, and second column gives the detected forged region using the proposed scheme

Table I. Performance comparison between the proposed method and former scheme [5].

Test Image	DCT-SVD-Based Scheme		Proposed Method	
	p	*f*	*p*	*f*
Baboon	0.97	0.03	0.98	0.03
Peppers	0.96	0.04	0.98	0.01
Lake	0.98	0.03	0.99	0.02
Average	0.97	0.03	**0.98**	**0.02**

5. Conclusions

A simple approach on passive copy-move forgery detection has been presented in this paper. The proposed method exploits the usability of HBTC feature on measuring similarity between the two image regions, i.e. copied and forged regions. The proposed method yields better performance compared to the former scheme under the same feature dimensionality. Another feature generated from HBTC data stream can be investigated to further improve the forgery detection performance. The rotation invariants as well as translation problem and multi resolution condition can also be considered for future works and developments.

Acknowledgment

The authors would like to thank the Institute for Research and Community Services of Sebelas Maret University for funding to this research in the academic year of 2016.

References

[1] J. M. Guo, and H. Prasetyo, "Content-based image retrieval using features extracted from halftoning-based block truncation coding," IEEE Trans. Image Process., vol. 24, no. 3, 2015.

[2] J. M. Guo, H. Prasetyo, and C. C. Yao, "Image retrieval using indexed histogram of void-and-cluster block truncation coding," Signal Process., vol. 123, pp. 143-156, 2016.

[3] O. M. Al-Qershi, and B. E. Khoo, "Passive detection of copy-move forgery in digital images: State-of-the-art," Forensic Sci. Int., vol 231, pp. 284-295, 2013)

[4] N. B. A. Warif, A. W. A. Wahab, M. Y. I. Idris, R. Ramli, R. Salleh, S. Shamshirband, and K. K. R. Choo, "Copy-move forgery detection: Survey, Challenges, and Future Directions," J. Network Comp. Apps., 2016, http://dx.doi.org/10.1016/j.jnca.2016.09.008

[5] J. Zhao, and J. Guo, "Passive forensics for copy-move image forgery using a method based on DCT and SVD," Forensic Sci. Int., vol. 233, pp. 158-166, 2013.

[6] E. J. Delp and O. R. Mitchell, "Image coding using block truncation coding," IEEE Trans. Commun., vol. 27, pp. 1335–1342, Sept. 1979.

[7] C. S. Huang and Y. Lin, "Hybrid block truncation coding," IEEE Signal Process. Lett., vol. 4, no. 12, pp. 328–330, Dec. 1997.

[8] J. M. Guo, and M. F. Wu, "Improved Block Truncation Coding Based on the Void-and-Cluster Dithering Approach," IEEE Trans. Image Processing, vol. 18, no. 1, pp. 211-213, Jan. 2009