A Novel Approach for Copy-move Forgery Detection using Bilateral Filtering

N.H. KAPLAN, I.KARABEY AKSAKALLI, U. KILIC, I. ERER

Abstract— Digital image processing methods have a wide area of usage and their complexity is increasing, as well as the tampering methods. A widely used tampering method is copy-move forgery. In this study, a hybrid method combining the Discrete Cosine Transform (DCT) and Bilateral filtering is developed. In this method, first overlapping blocks are obtained from the input image. Then, bilateral filtering and DCT of these blocks are multiplied to obtain the refined block features. The block features are scanned by a zig-zag process followed by a lexicographic sorting. Finally, a similarity detection by a predetermined threshold parameter is applied to detect the forgery. Both visual and quantitative results demonstrated that the proposed method can determine the copy-move forgery regions.

Index Terms— copy-move forgery, bilateral filtering, zigzag scanning, DCT (Discrete Cosine Transform)

I. INTRODUCTION

NOWADAYS, digital images are used in important areas such as medical, law and public. Digital images can be manipulated and regulated easily by malicious people using various image regulation software tools. With the emergence of this software, the reliability of the images and their authentication have become an important problem. Therefore, image fraud detection has become an important research focus. Image fraud detection methods are generally divided into two categories as active and passive approaches. Active approaches are based on additional information embedded into digital images such as digital watermarks or digital signatures. By using these additional information, the originality of the image can be detected. Unfortunately, active approaches require additional information to be embedded in the image by authorized personnel in the process of capturing the false image or at a later stage. If there is no information about the original image, applying an active approach is not useful [1].

On the other hand, passive approaches are used to determine the manipulated image without any additional information. Passive approaches are divided into two groups named tampering detection and source device identification. This approach detects copied image by extracting real features in the image. Tampering detection is also divided into dependent and independent classes. The dependent class of copy-move forgery is the commonly used method in fraud image. The image content is manipulated by copying an object that exists in the image and paste this object to another location within the same image. The transition between copied object and original image is masked using a variety of retouching tools. Since the features such as noise, color and contrast in the source and target regions have a statistical match, the detection of copied zones is a challenge.

Fridrich et al. [1], pioneers of copy-move forgery detection algorithm (CMFD), has handled the various requirements of the detection algorithm. The first requirement is that the detection algorithm should allow the approximate matching of the small image segments. Secondly, while the detection algorithm determines the mismatched fields (false positive), it must have an acceptable execution time. Furthermore, the authors mentioned that a fake segment will likely have a dependent component rather than very small patches or individual pixels.

In this study, a novel method consisting of a combination of Discrete Cosine Transform (DCT) and zigzag scanning is proposed by applying bilateral filtering. DCT is one of the most used watermarking algorithms among many data hiding methods to protect digital multimedia files. It states a limited sequence of data points in terms of a sum of cosine functions in different frequencies. Bilateral filtering has been used in many image processing algorithms [2-4]. Bilateral filters take into consideration both spatial and spectral properties of the image. By this way, the edge information is kept during the filtering process. The method is compared with traditional and state of art methods and the proposed method gives 99.7% accuracy rate in a standard dataset called “CoMoFoD” used for benchmarking the detection of tampering or copied images.
II. RELATED WORK

It is known that there is a correlation between the original image and the pasted object in copy-move forgery [1]. This correlation is used to detect forgery successfully. The methods for approximate matching of copy-moved and real segments using retouching tools or other image processing tools are not fully sufficient to detect forgery, so various approaches for detecting the forgeries are increasing day by day.

In most of the methods proposed for CMFD, the basic procedure is divided into pre-processing, feature extraction, matching, filtering and post-processing stages, respectively [5-8]. In the pre-processing stage, image data is improved to enhance image features or to reduce undesirable distortions within the image. One of the most commonly used methods at this stage is the conversion of RGB color channels of the input image into a single grayscale image [1], [5-9]. Besides, if the image is stored in a compressed format, the files are decompressed in the preprocessing phase [3]. The feature extraction phase is then applied to the selection of related information representing the properties of the image. This phase is carried out in two ways: dividing into blocks and keypoint detection. In block-based approach, the image is divided into blocks in a square or round shape. These blocks can be divided by overlapping or non-overlapping division [8]. Then, properties are extracted from these blocks by using various features (frequency transform, texture and intensity, moments invariant, log-polar transform, dimension reduction etc.) and similarity comparison is performed between the blocks in the image. The third stage, matching, is a process in which the similarities between different segments of the image are detected. This process is carried out for each feature that is extracted and measured to define the manipulated area. In the literature, block-based matching can be performed by several methods such as sorting, hashing, correlation and Euclidean distance [8]. In the key point-based method, the image is subdivided into subfields and the feature points are removed only in certain regions by using methods such as SIFT, SURF, etc. [10]. Then, similar to the block-based method, each of the feature points is paired with each other using methods such as clustering, Euclidean distance, etc. If the focus is on computational complexity, the use of keypoint-based algorithms should be preferred. If the aim is to achieve a higher accuracy rate, block-based methods provide a higher accuracy rate [7]. In the scope of this study, it is aimed to obtain a higher accuracy rate than complexity.

In the literature, many block-based methods have been proposed for copy-move forgery detection and among these methods, DCT (Discrete Cosine Transform) based methods generally present higher performance. Since DCT is used in the scope of this study, literature studies involving this method are given priority. Parveen et al. [7] have proposed a pixel-based method for copy-move forgery detection by converting the color image into a gray-scale image. Then the image is split into 8x8 block sizes and applied feature extraction using DCT (Discrete Cosine Transform) on different datasets. After the feature extraction, k-means algorithm is used for block clustering and the radix sort algorithm is used for feature matching. Wandji et al. [11] convert each block into DCT coefficients and apply feature extraction obtained from DCT coefficients with Red, Green and Blue (RGB) channels. Afterward, similar block pairs are searched by ordering the properties alphabetically. In the last stage, duplicated regions, if any, are detected by using Euclidean distance as the similarity criterion. With this method, it has been proved that there is no decrease in the performance of the image with a slight rotation, JPEG compression, scrolling, scaling, blur and noise addition. A similar study is performed by Huang et al. [12] and DCT coefficients are used as a feature. It is observed that PCA (Principle Component Analysis) analysis method [13] is better for feature extraction in copy-move forgery detection when compared to detection forgery based on lexicographic sorting of DCT block coefficients. Wandji et al. [11] converted each block into DCT coefficients, then apply feature extraction obtained from these coefficients with red, green and blue (RGB) channels. Afterwards, similar block pairs are searched by ordering the properties alphabetically. In the last stage, duplicated regions, if any, are detected by using Euclidean distance as the similarity criterion.

Using the block-based method, not only the detection of the copied object, but also the image tampering is detected. Manu and Mehtre [14] proposed two methods for detecting forgery by identifying the boundaries and location of the tampered image without any prior knowledge. The first method is the classification of tamper patterns with standard deviations of block discrete cosine transformations (BDCT) of textures. In the second method, the entropy of the histograms and the image quality artifacts due to image forgery are combined with the texture patterns of the first method. In these methods, a SVM based classifier is used to determine whether the image is tampered or not. It is emphasized that the method is implemented on various datasets and it is stated that the highest accuracy belongs to the second method with 98.89%. In another study [15], the RGB image is converted to grayscale and the resulting image is divided into blocks of m x m pixels. After the 2D DCT coefficients of the blocks have been calculated, zigzag scanning is performed on each block to reposition the coefficients to a feature vector. Extracted feature vectors are sorted by lexicographic sorting. Lastly, duplicated blocks are determined with Euclidean distance. The proposed method is used for extracting false positive and false negative values in the CoMoFoD dataset using different overlapping block sizes. As a result of this method, it is stated that the accuracy performance of different overlapping block sizes is affected by the size of the forged area, and the 4x4 overlapping block size causes a high false-positive compared to 8x8 overlapping block size, decreasing the accuracy of tampering detection in terms of precision.

III. PROPOSED METHOD

On the basis of all copy-move forgery detection methods in the literature, it is seen that there should be determined at least one copy of the same object. The areas within the image, which are exactly the same, are compared in size and shape. For this purpose, the image is divided into overlapping blocks by using block-based detection method. Then, feature extraction is performed from each block. Thus, each block has some features and the possibility of replicating regions having similar features is considered.
A. DCT (Discrete Cosine Transform)

During feature extraction, Discrete Cosine Transform (DCT) is first applied to the blocks in the proposed method. The intensity of the image at pixel \((x, y)\) is \(I(x, y)\) and the block size is \(b\). The overlapping blocks of the image are represented by corresponding DCT coefficients as \(D(u, v)\) as per equation (1).

\[
D(u, v) = \frac{2}{b} C(u) C(v) \sum_{x=0}^{b-1} \sum_{y=0}^{b-1} I(x, y) \cos \frac{\pi u(2x + 1)}{2b} \cos \frac{\pi v(2y + 1)}{2b} \tag{1}
\]

where

\[
C(u) = \begin{cases} 
\frac{1}{\sqrt{2}} & \text{if } u = 0 \\
1 & \text{otherwise}
\end{cases}
\tag{2}
\]

B. Zigzag Scanning

Zigzag scanning is used to group Direct Current (DC) low frequency and Alternative Current (AC) high frequency coefficients obtained from the DCT process. This scanning method maps 8x8 blocks to 1x64 block. It increases the rate of image and video compression. A sample of zigzag scanning process is shown in Fig. 1.

![Fig.1. Sample of zigzag scanning process on 8x8 picture](image)

C. Bilateral Filtering

Let Bilateral Filter0(I) (BF0(I)) be equal to the original image \(I\) at level 0, BF1(I) be the approximation image at level 1, \(p\) be the vector defining the location of the pixel to be filtered, \(S\) be the set of the neighbor pixels of \(p\) and \(q\) be the vector defining the location of a pixel in the set \(S\). In the classical bilateral filtering, the \(l\)th approximation layer of a current pixel location \(p\) is estimated by:

\[
BF(I)_p = \frac{1}{W_p} \sum_{q \in S} G_{\sigma_s}(\| p - q \|) G_{\sigma_r}(\| I_p - I_q \|) I_q \tag{3}
\]

with normalization parameter:

\[
W_p = \sum_{q \in S} G_{\sigma_s}(\| p - q \|) G_{\sigma_r}(\| I_p - I_q \|) \tag{4}
\]

and Gaussian kernel as:

\[
G_{\sigma}(x) = \frac{1}{2\sigma^2} e^{-\frac{x^2}{2\sigma^2}} \tag{5}
\]

The first detail plane is obtained by extracting the approximation image from the original image as

\[
D^1(I) = BF^0(I) - BF^1(I) \tag{6}
\]

with \(BF0(I) = I\) denoting the original image.

To obtain a multiscale decomposition, the spatial parameter \(\sigma_s\) is doubled and the range parameter \(\sigma_r\) is halved at each level.

The detail layer at level \(j\) can be obtained by the difference between the two consecutive approximation layers as

\[
D^j(I) = BF^{j-1}(I) - BF^j(I) \tag{7}
\]

The original image can be reconstructed by simply adding the detail layers and the final level approximation layer as,

\[
I = \sum_{i=1}^{L} D^i(I) + BF^{L+1}(I) \tag{8}
\]

D. Algorithm Framework

This section describes the procedures performed for the proposed method, respectively. The framework of the method is shown in Fig. 2.

Step-1: The input image has been converted into an mxn sized gray-scale image using Luminance values. A colored image is converted into grayscale using a standard formula of YUV conversion: \(Y=0.299R+0.587G+0.114B\). In the formula; \(R, G, B\) represent three different color components of the RGB color model.

Step-2: The grayscale obtained in Step 1 is divided into 8x8 blocks. The purpose is to identify overlapping (m-b+1) (n-b+1) blocks by shifting one pixel consisting of fixed-size bxb square windows from the top left to the bottom right corner of the image.

Step-3: After blocking, Bilateral filtering described in (3) is applied to each block to obtain the first level approximation layer for each block. In order to carry on decomposition Bilateral filtering is applied to the filtered blocks by doubling the spatial parameter and halving the spectral parameter to obtain the second level approximation layers for each block. The difference between these two approximation layers construct the second detail layer. Since each block is 8x8, the resulting detail layer will also be 8x8 and is named as Bilateral blocks.

Step-4: A normal image is a time-based image. Although the human eye can easily detect transitions at low frequency, it
cannot detect transitions in the high-frequency region. Therefore, copy-move forgery is usually performed in high-frequency regions. In this step, DCT is applied to 8x8 unfiltered blocks to convert from time domain to frequency domain. In DCT, the DC value of each block represents the brightness in that block.

**Step-5:** DCT applied blocks and Bilateral filtered blocks are multiplied and new 8x8 DCT-Bilateral blocks are created.

**Step-6:** Zigzag scanning has been applied to DCT-Bilateral filtered blocks to obtain a 1x64 sized vector and first 16 values of the vector are considered. \( P \times B^2 \) part of 1x64 sized vector is taken to reduce dimension in DCT-based image fraud. \( P \in [0,1] \) is generally selected as 0.25 [12]. Thus, vector size drops to 1x16. In JPEG compression, fake data is searched in the low frequency region (first 16 values) due to the data loss occurs in the high frequency region.

**Step-7:** The quantization process is performed by dividing each vector composed with zigzag scanning into 16.

**Step-8:** Matrix A with \( (N-7 \times M-7, 16) \) elements is ordered lexicographically. In this step, vectors obtained from each block are placed in matrix A. Each vector is compared with all vectors placed after itself. Euclidean distance is used for similarity detection.

*Euclidean distance:* This method measures the difference between vectors and if the result is less than a certain threshold value, it is determined that these vectors are similar. The choice of the threshold value is very important, since it is used to determine how similar vectors are. In this step, lexicographically sorting has been performed to reduce comparison complexity. This reduction may result in disappearance of the information about which block the vector belongs. To avoid this, block values are stored by adding 2-bit location information to the end of the vector.

**Step-9:** After the lexicographic sorting, each element in the matrix A is compared with the element up to threshold value after itself using Euclidean distance. The initial coordinates of the vectors that are found to be similar are stored in a separate sequence. This sequence is called the shift-vector. If an area is selected as similar, the density of shift vectors is checked to ignore the false detected areas. The dense area is determined as the place of copying.

**Step-10:** The number of shift-vector is considered and if this number is up to threshold value, these areas are marked as a forgery.

**Step-11:** In the final step, the accuracy ratio and false negative values are calculated. The higher is the block size, the higher is the possibility of false negatives. The false-negative value increases when shift vector is used but this error is detected if there is multiple forgeries.

---

**IV. EXPERIMENTAL RESULTS**

**A. Dataset**

According to our investigations, although many methods have been proposed to detect copy-move forgery, there was no common database created to test the proposed methods until 2013. In 2013, a common database called “CoMoFoD” is created by Video Communication Laboratory (VCL) [16] to detect copy-move forgery. The CoMoFoD dataset [17] contains two separate image sets containing 512 images of 512 x 512 and 60 images of 3000 x 2000. Each group has unreal images with five different manipulations such as translation, rotation, scaling, merging, and distortion with various composite operations. Also, each image set contains 40 different images containing the original image, color masking, dual masking, and fake images as shown in Fig. 3.
In addition, there are image forms in distorted original and fake using various last operation techniques for each image such as JPEG compression, blurring, noise addition, brightness change, contrast adjustment and color reduction [16-18].

B. Parameter Adjustment and Visual Comparisons

The performance of the bilateral filtering is highly correlated to three parameters, namely window size, spatial parameter and range parameter. In order to determine the optimum parameters, the proposed method is applied to several images and the accuracy rate of the resulting image is recorded. According to this comparisons, the optimum window size is determined as 3x3, spatial parameter as 1 and range parameter as 0.1.

After determining the optimum parameters, the proposed method has been tested on 40 different images on the CoMoFoD dataset. Among the 40 tested images, a copied region is pasted to more than one region in 12 images and more than one different regions are reproduced in 7 images with copy-move forgery. Fig. 4 shows the resulting images obtained by the proposed method we applied to the CoMoFoD dataset. According to the figure, the copy move forgeries are detected in a good accuracy, and a very little amount of false positives.

In order to make visual comparisons, the proposed method, DCT, Local Binary Pattern (LBP) and DCT-LBP methods have been applied to several images in CoMoFoD dataset and sample results are shown in Fig. 5. Fig.5.a and Fig.5.b show the input image and the mask determining the copy move forgery area, respectively. Fig. 5.c, Fig. 5.e, Fig. 5.g and Fig. 5.i demonstrates the results of the DCT, LBP, DCT-LBP and the proposed methods, respectively. Fig. 5.d, Fig. 5.f, Fig. 5.h and Fig. 5.j show the copy move forgery areas obtained by the DCT, LBP, DCT-LBP and the proposed methods, respectively. According to Fig. 5., the closest result is achieved by the proposed method, with lower false positives. In order to make an objective comparison, quantitative assessments are made as well.

C. Quantitative Comparisons

In the copy-move forgery detection technique, two different assessments are made according to whether the copied area is regular or irregular.

In areas that are properly copied in square or rectangular form, the TP (True Positive) ratios of block-based copy-move forgery detection algorithms may be high and the FP (False Positive) ratios low. Considering this situation, the success of the proposed method on the experimental images is high.

The results of the proposed method in Table 1 are given in comparison with the other methods used previously using TP (True Positive) and FP (False Positive) criteria. The results are calculated based on the pixels that the algorithms correctly identified and marked as false positive. The calculation of TP and FP is given in equation (9) and (10), respectively. True positive (TP) represents the ratio of both result of the method and points labeled by the reference data.

\[
TP = \frac{(C_1 \cap F_1) + (C_2 \cap F_2)}{(C_1 + C_2)}
\]  

The higher is the TP value, the better is the method. False-positive (FP) represents the proportion of points that are labeled by the method but not labeled in the reference image. C₁ and C₂ are copy-move areas, F₁ and F₂ are the fields determined by the method.

\[
FP = \frac{(C_1 \cup F_1) + (C_2 \cup F_2)}{(C_1 + C_2)} - TP
\]

The lower is the FP value, the better is the method.

**TABLE I**

| Method                        | Block Size | 8x8 | 12x12 |
|-------------------------------|------------|-----|-------|
| DCT [18]                      | 0.75       | 0.12| 0.65  | 0.40  |
| LBP [18]                      | 0.82       | 0.25| 0.67  | 0.60  |
| DCT-LBP [18]                  | 0.89       | 0.08| 0.85  | 0.12  |
| DCT-Bilateral (proposed method)| **0.96**   | **0.042**| **0.95**| **0.051** |

In order to make visual comparisons, the proposed method, DCT, Local Binary Pattern (LBP) and DCT-LBP methods have been applied to several images in CoMoFoD dataset and
In Table 1, average accuracy rate (TP) and false positive values (FP) are determined using some copy-move forgery detection methods in literature and the proposed DCT-Bilateral method by applying on the same images in CoMoFod dataset. When the results in Table 1 are compared, the best TP and FP scores are achieved by proposed DCT-Bilateral method. The worst TP values are obtained by DCT method, whereas the worst FP values are achieved by LBP method. Therefore, it is observed that the copy-move forgery detection in DCT-Bilateral applied images yields at least 7% more accurate results. It seems that DCT and LBP alone give less accuracy rates than combined methods.

Fig. 4. (a), (e) and (i) represent orijinal images in ComoFod dataset; (b),(f) and (j) represent mask of these images; (c), (g), (k) DCT – Bilateral applied images; (d), (h), (l) represent mask of (c), (g), (k)

The images and their masks obtained from the proposed DCT-Bilateral method and the other methods in literature are listed in Fig. 4. According to the obtained masks, only LBP applied method finds more false positive rates than the only DCT applied method. Similarly, when DCT-LBP and DCT-Bilateral images are compared, it is seen that DCT-Bilateral gives higher accuracy and lower false positive rates in both 8x8 and 12x12 images. These findings are also overlapped with given values in Table 1.

V. CONCLUSION

In this study, a block-based method is proposed for the first time using DCT and Bilateral filtering to detect copy-move forgery. Two different Gaussian kernels, namely spatial and range kernels of the bilateral filtering keep the edge information in the process of filtering. After dividing the image into overlapping bxb blocks, we applied bilateral filtering and DCT to each block separately. Then we multiply the obtained new blocks with each other to obtain the feature vectors. When we compare the proposed method with other copy-move forgery methods, it is seen that the true positive rate of the proposed method is higher. The proposed method is the first method in copy-move forgery detection technique using DCT and Bilateral combination. This method also works on both gray-level images and color images. Although the proposed method in this study gives good results after blurring and noise addition on the image, it performs poorly in forgery by performing scaling or rotating operations on the copied image while performing copy-move forgery.

Fig. 5. (a) and (b) orijinal images; (c) and (d) only applied DCT image; (e) and (f) only applied LBP image; (g) and (h) DCT-LBP image; (i) and (j) DCT-Bilateral image

REFERENCES

[1] A. J. Fridrich, B. D. Soukal, and A. J. Lukáš, “Detection of copy-move forgery in digital images,” in Proceedings of Digital Forensic Research Workshop, 2003: Citeseer.

[2] N. H. Kaplan and I. Erer, "Bilateral pyramid based pansharpening of multispectral satellite images," in
UR KILIC graduated from Harran University in Turkey and he received his MSc in Computer Engineering from Ataturk University, Turkey. After the graduation from the Ataturk University, he started PhD in 2016 at Karadeniz Technical University. He is currently work as a research assistant at Erzurum Technical University. His primary research interests are IOT security, image processing and digital signal processing.

ISIN ERER is an Associate Professor in the Department of Electronics and Communication Engineering, Istanbul Technical University, Istanbul, Turkey. Her research interests include: Statistical signal processing, image processing for remote sensing, high-resolution radar imaging and ground penetrating radar.