IJCSE International Journal of Computer Sciences and Engineering Open Access

Review Paper | Volume-4, Special Issue-3, May 2016 | E-ISSN: 2347-2693

Inactive Method of Noncausal 2D Image Splice Recognition Model using Markov Model

Thofa Aysha [1], Manjesh R [2]

[1] *Computer Science and Engineering, Srinivas Institute of Technology, Mangaluru, India*
[2] *Computer Science and Engineering, Srinivas Institute of Technology, Mangaluru, India*

Available online at: www.ijcseonline.org

Abstract— Noncausal Markov model for a 2D signal is one of the inactive methods for spliced image. Image splicing is an image copies or merge a portion of image to same images or different images. The way Noncausal Markov model differs from traditional Markov model is the proposed methodology models a image as a 2-D noncausal signal and captures and analyzes the underlying dependencies between the current node and its neighbors in all directions. These dependencies are obtained through Discrete Cosine Transform and Discrete Wavelet Transform. These parameters give features to differentiate the natural ones with the features of spliced images. The noncausal Markov Model considers the input of block discrete cosine transformation domain, the discrete wavelet transform domain, and the cross-domain features for classification. The Expectation Maximization which is the classifier which classifies based on maximum likelihood of images. The dataset used is UCID dataset where we have uncompressed color images.

Keywords—Noncausal Markov Model, Discrete Cosine Transformation (DCT), Discrete Wavelet Transform(DWT), inactive image splicing recognition, Expectation Maximization(EM).

I. INTRODUCTION

The digital image play technology is important topics in daily life. While the use of digital cameras there were also sophisticated image processing software developed for altering and for manipulation of image. There are many image processing tools that were also developed which manipulate the images and also hiding some useful or important data to make forged images. The goal for image forensics will always be addressed to image integrity and authenticity. The integrity of image loses when they are attacked by image slicing and cloning. This type of forged image is not recognized easily by bare human eye resulting in its authenticity. So there must be method for regaining the trust for integrity and authenticity verification in image processing field. The Freehand and Photoshop are some of the famous tools used to modify and maliciously manipulate digital images. So developing method to verify and recognize affected digital images became very essential when images are used for any law proceedings in judgment in court, for example, for medical purposes and financial document, transportation sector etc. The different types of tampering images are taken care by different digital forgery detection techniques that have been proposed which regains trust for the image authenticity and integrity. One of such methods deals with image splicing recognition.

The digital image forgery detection methodology consists of two type mainly active methods and inactive methods. The active methods checks the image integrity using prior embedded knowledge like watermark or digital signature has marked as forged. Applying this type of method can detect image tampering and find the controlled areas accordingly. Embedding the signatures or watermarks into the digital image in the imaging process causes infeasibility in cameras and because of this prior knowledge given to active method to help the validation of digital images. Conversely with no prior knowledge there developed a method called the passive or inactive methods. They detects image frauds on the Internet which is been forged. This method distinguishes image forgery under the condition that altered image would change the fundamental attributes of the first pictures. The inconsistency characteristic found on the statistical is the proof for discovery of altered image. Among various types of image tampering operations, splicing is the most common and useful operation in creating image forgeries.

The three major steps of Digital Image Forensics are detection, source identification of digital image and image forgery detection. The recognition of computer generated image is very hard for the viewers because of its high quality and high realism realistic that makes them confused. From cameras and scanner the image source is being identified. Every device has its way of footprint of image. This type of forensics has two approaches for image source identification, the first method attempt to resolve the source image class properties and second method identifies the

Corresponding Author: *Mr. Manjesh R, manjuis024@gmail.com*
Dept. of CSE, Srinivas Institute -of Technology, India.

image potential source device through novel hardware characteristics. The forgery detection has algorithms of various kinds and techniques for detection and traces image tampering. The traces are used for proof of tampering.

For the detection of splice image, the image has to be classified and then detected. For block based classification, an image is divided into certain number of blocks. And with the extraction of statistical characteristics feature vector from each of the block, a class for image is formed. The important thing here is, it should consider the context information also. The forged images can contain hidden information in it. This also should be detected in detection algorithm.

II. Related Work

The image splicing detection always classifies the features and extracts the statistical characteristics for identifying the differences from the base images.

A. *Hidden Markov Model Image Classification*

In traditional block-based classification, by arranging the statistics extracted from the block, a feature vector is formed for each block and decides which class it belongs and ignoring context data. To improve the classification by context two dimensional (2D) hidden Markov models (HMM) was presented. By considering feature vectors, the HMM is dependent through an underlying state process which has transition probabilities on the states of neighboring blocks from both horizontal and vertical directions. As the dependency in two dimensions is obtained and the HMM parameters are estimated by the EM algorithm. To classify an image, the classes with maximum likelihood are searched every one block [2].

The image is then classified according to the feature vectors. The 2-D HMM assumes that the feature vectors are generated by a Markov model that may change state once every block. In classification, blocks are not classified one by one in such an order. The classification algorithm attempts to find the optimal combination of classes jointly for many blocks at once. The above assumption can be summarized by two points. First, the state is a sufficient statistics for estimating transition probabilities, i.e. they are conditionally memory less. Second, the state transition is first-order Markovian in a 2-D sense [2]. The second assumption is that for every state, the feature vectors follow a Gaussian mixture distribution. Once the state of a block is known, the feature vector is conditionally independent of the other blocks. The task of the classifier is to estimate the 2-D HMM from training data and to classify images by finding the combination of states with the maximum a posteriori probability given the observed feature vectors.

For the feature to be extracted for the assumed HMM, estimate the transition probabilities and covariance matrices of the Gaussian distributions parameters. The EM algorithm provides an iterative computation of maximum likelihood estimation when the observed data are incomplete. The estimation algorithm iteratively improves the model estimation by the following two steps.

1) Given the current model estimation, the observed feature vectors and classes, the mean vectors and covariance matrices are updated.
2) The transition probabilities are updated.

The iterative algorithm starts by setting an initial state for each feature vector. For every class, feature vectors labeled as this class are sequenced in a raster order; and the states corresponding to this class are assigned in a round-robin way to those vectors [2].

B. *Steganalysis for splicing detection*

Stego data is also a important channel in digital images. To check whether the given image has hidden data or not, Steganalysis for splicing detection was developed. Steganalysis methods are divided as specific steganalysis and universal steganalysis [3]. The Encrypted information bits are embedded using steganography into the cover image. Splicing is to replace or merged with one or more portion on a base image from the same host image or other source images. So the statistical characters with steganography slightly differ from splicing. For a human eye spliced images looks like unspliced even though they altered with some part. By embedding the data to cover image, steganalytic attacks is avoided whereas splice images only changes parts in host images. So stegnography is considered global and splicing is considered local. Hence touching these stego images and spliced images cause disturbance on the smoothness, regularity, continuity, consistency and periodicity of the images. Therefore, they do cause some statistical artifacts and are thus detectable using certain statistical attack approaches. With well designed natural image model, which can separate stego images and spliced images from natural images, both steganography and splicing are detectable by machine learning schemes [3].

C. *New Developments in Color Image Tampering Detection*

An efficient method for passive-blind color image forgery detection is presented here. Statistical features are obtained by applying DCT to a given test image and a set of 2-D arrays. Because we have observed sensitivity to color image, the image features are extracted from Cr channel, a chroma channel in YCbCr color space. To effectively evaluate over a color image dataset boosting feature selection is applied to having feature dimensionality reduced so as to make detection accuracy general and computational complexity decreased.

The effective study of image chroma brought to color image splicing detection in made us obtain knowledge into our feature set. Our features that we considered are extracted from one image chroma channel. We consider natural image model derived from the following two combinations first one is features derived from the image pixel 2-D array and those derived from the MBDCT coefficient 2-D arrays and second is that of moments of characteristic functions based features and Markov-based features. It is expected that with various block sizes, the coefficients of MBDCT 2-D arrays efficiently capture such diverse and complicated changes compared to those of a single-block-size BDCT array. To capture tampered character, the moment-based features combine the information from first-order statistic and second-order statistic. A prediction error 2-D array reduces the influence caused by diversity of image content; as a result, high frequency contents, where tampering artifacts reside in are enhanced [4].

D. Detection Based On Stationary Distribution of Markov Chain

An efficient method for passive-blind color image forgery detection is presented here. Statistical features are obtained by applying DCT to a given test image and a set of 2-D arrays. Because we have observed sensitivity to color image, the image features are extracted from Cr channel, a chroma channel in YCbCr color space. To effectively evaluate over a color image dataset boosting feature selection is applied to having feature dimensionality reduced so as to make detection accuracy general and computational complexity decreased [5].

The effective study of image chroma brought to color image splicing detection in made us obtain knowledge into our feature set. Our features that we considered are extracted from one image chroma channel. We consider natural image model derived from the following two combinations first one is features derived from the image pixel 2-D array and those derived from the MBDCT coefficient 2-D arrays and second is that of moments of characteristic functions based features and Markov-based features [5]. It is expected that with various block sizes, the coefficients of MBDCT 2-D arrays efficiently capture such diverse and complicated changes compared to those of a single-block-size BDCT array. To capture tampered character, the moment-based features combine the information from first-order statistic and second-order statistic. A prediction error 2-D array reduces the influence caused by diversity of image content; as a result, high frequency contents, where tampering artifacts reside in are enhanced.

Here we model edge information for inactive image tampering detection method. Edge image of image chroma component as a finite-state Markov chain is modeled and extract low dimensional feature vector from its stationary distribution for tampering detection. The support vector machine (SVM) is used as classifier to evaluate the effectiveness of the proposed method and avoid overfitting problem.

Capture its interpixel dependencies through modeling the edge image of finite-state Markov chain (MC). Markov chain is a utilized to model adjacent pixels statistical dependency for steganalysis and splicing detection. The transition probability matrix (TPM) can be used to characterize it. We find the chroma channel of a color image is more suitable for image splicing detection than illuminance channel. A RGB color image is transformed into YCbCr color space and chroma component (Cb or Cr channel) is used. Then, we determine the stationary distribution of the thresholded edge image of chroma to server as a feature vector [5].

E. Markovian Rake Transform for Digital Image Tampering Detection

Here also we present one more effective framework for passive-blind color image forgery detection is presented. By applying Markovian rake transform to image luminance component, we generate image statistical features. Markovian rake transform is the applications of Markov process to difference arrays which are derived from the quantized block discrete cosine transform 2-D arrays with multiple block sizes. With large-scale image dataset designed for tampering detection, the efficiency of generated features has been confirmed where relevant issues were found and corresponding adjustment measures have been taken [6].

With the natural image model based on Markovian rake transform (MRT) on image luminance, MRT applies Markovian process to difference 2-D arrays independently derived from block discrete cosine transform with multiple block sizes [6]. The proposed MRT features of size 486, abbreviated as MRT-486, consist of Markov process based features generated with three different block sizes: 4×4, 8×8 and 16×16. Image luminance is chosen for feature extraction because it does not suffer from any uneven treatment applied to color information. Although the dimensionality of MRT-486 seems high, they actually require relatively low computational time because only rather simple operations are involved in feature extraction. Even with low complexity, MRT-486 perform fairly well. Moreover, the proposed features outperform any combination of MP features, provided that the number of block sizes utilized in MRT is less than or equal to three. The tests on images outside the dataset used, although the success rate is still far from satisfactory, unveil the practicality of our proposed natural image model to a certain extent. This indicates that image tampering detection is still encountering enormous challenges, especially of the tests on real-life tampered images [6].

F. Copy-Move Forgery Detection using DWT

A copy-move image forgery is to hide some portion of image or to add some portion that is being duplicated. In both of the case image reliability and authenticity is compromised. So an improved algorithm based on Discrete Wavelet Transform (DWT) is used to recognize such duplicated forgery. In this technique at first DWT (Discrete Wavelet Transform) is applied to the input image for a reduced dimensional representation. Then the compressed image is divided into overlapping blocks. After that Lexicographic sorting is performed and duplicated blocks are identified. Due to DWT usage, detection is first carried out on lowest level image representation. This approach increases accuracy of detection process and reduces the time needed for the detection process [7].

G. Problems in related works

There were conflicts in Conventional block-based classification algorithms that classify block class using feature vector and not caring about context information. So to improve context based block classification two dimensional (2D) hidden Markov models were introduced. By using steganalysis method, it aims at detecting secret and hidden information inside a image. The BDCT domain and spatial domain features were considered in causal Markov model to detect image tampering in chroma channels for 1D image. A while after, causal Markov model using both Discrete Cosine Transformation (DCT) and Discrete Wavelet Transform (DWT) domains and splice image was recognized by the cross-domain Markov features.

The problem here is all the method explained here considered image splicing recognition considers one dimensional image and considered the dependency difference between the current node and its previous node taking certain direction such as either horizontal or vertical. So here the 2D nature of images was vomited which is considered in proposed image splicing model and it takes dependency in all directions. The simplified Markov model used before does not consider the image characteristics sufficiently.

III. PROPOSED METHOD FOR IMAGE SPLICING DETECTION

The already developed 1-D Markov model only considers the state conditions in the 1-D sense along certain direction. In 2D images every state is dependent on its neighbor. In spite of the fact that the noncausal model could consider how much ever basic data as could be expected, it could become computationally infeasible. So here we consider four neighbors to rearrange the noncausal model to decrease the computational complexity mentioned above which finds a 2-D noncausal Markov model. There is no scientific answer for the 2-D noncausal model, we decompose the noncausal model into a few 2-D causal models and examine these models consecutively to estimate the 2-D noncausal model. The 2-D causal submodels can be considered by the state transition probability matrix, probability density function of every state, and the prior probability of every state. The statistical dependency characteristic is considered by transition probability matrix and measures reliance qualities among neighboring nodes in the model [1].

The 2-D hidden noncausal Markov model is defined as consists of three parts.

1. Considering state set and the observation set with spatial coordinates, and the number of rows and columns in the image. Each observation is considered as the state. Observations can be characterized by useful issues, for instance, BDCT coefficient contrast array, DWT coefficient cluster or pixel intensity distinction cluster etc.
2. Consider the state dependence in the image, where set of neighboring states that will affect the current state.
3. Consider total number of states. The distribution of observation set relating to state is formulated as function of both the values. If the state set is obscure, the model would advance to be a 2-D noncausal shrouded Markov model which is computationally costly. To fix this problem, we isolate the state set into different levels according to the distribution of observations.

The proposed detection method consists of two parts: feature extraction and classification. The 2D noncausal Markov model-based features are extracted from the source arrays, and then the extracted features are fed into Expectation Maximization to determine whether the target image has been spliced or not.

A. Feature Extraction

Image splicing would introduce specific statistical artifacts in DCT domain and DWT domain. By modeling adjacent coefficient difference array from DCT and DWT domains is observations input for the 2-D noncausal Markov model.

1) DCT Domain Features: In image processing, most image content-related information is usually contained in a few low-frequency Neglecting the low-frequency coefficients in the image could put emphasis on some content-irrelevant traces on images causing image tampering. With the good information packing in DCT, it considers all such low frequencies.. The region containing tampering traces is usually much lesser than the whole image in most of the spliced images. Hence DCT usually focuses on the local information, so for this features DCT is adopted in image-splicing detection. The dependences among intrablock BDCT coefficients and interblock BDCT coefficients are considered in this method of image splicing detection as a feature. For a r × r BDCT transform, the entire image is divided into a number of nonoverlapping r × r blocks. Difference arrays (both horizontal and vertical) are considered as the observations and the relationship between the state and the observations. Example for 8x8 blocks sized DCT is shown in Fig 1.

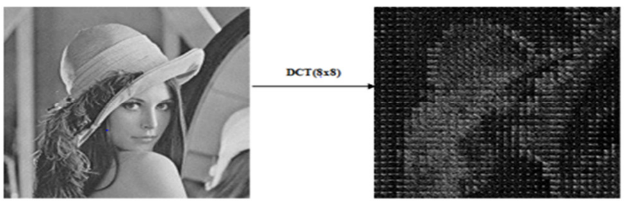

Fig 1: DCT with block size 8x8

2) *DWT Domain Features:* Three-level discrete wavelet decomposition was modeled for considering dependencies of coefficients across positions, levels, and orientations for classification. Image splicing has being reflected by the high-frequency wavelet subbands (detail coefficients array). We apply one-level DMWT on the source images and the detail coefficient arrays at three orientations, that is, horizontal CH1, vertical CV1, and diagonal CD1 are treated as the observation arrays for the proposed noncausal Markov model as shown in Fig 2. Multilevel wavelet decomposition introduces excessive redundant information which could probably confuse the classifier and degrade the detection performance.

B. Classification

In image-splicing detection the extracted feature vector is fed into a classifier to check whether it is spliced or not. Hence EM is adopted to evaluate the effectiveness of the features for its robustness to overfitting problem by estimating and maximization. Because EM is a machine learning method, we split the feature set extracted into two sets: one for training and the other for testing. The training set is considered to find the optimal hyperplane for classification, and the testing set is used to check the effectiveness of the proposed features.

The Image splicing is detected using following steps as shown in Fig 3:

1) For a given image, apply DCT and DWT to transform the image respectively.

2) Get the observation arrays and state arrays. Construct 2-D non-causal Markov models for these arrays.

3) Split the non-causal model into four causal submodels.

4) Extract the parameter set.

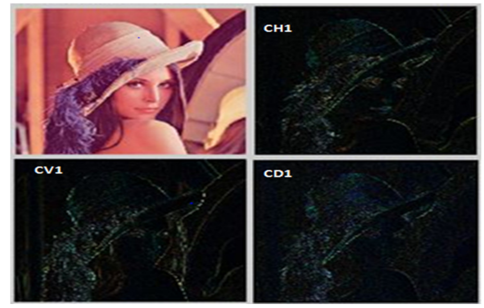

Fig 2: One-level DWT

5) Repeat the procedure 1–4 to get all the feature vectors in the image data set and group all of them into a feature set.

6) Randomly split the feature set into training and testing set proportionally.

7) Conduct EM parameter selection on the distributed 2D grid.

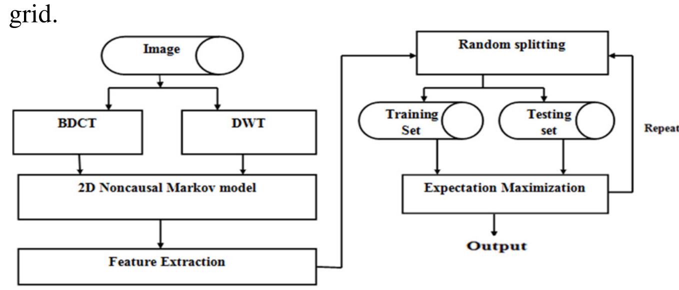

Fig 3: Noncausal Image Splicing Detection Steps.

8) Input the parameter set into EM and get training samples by the training algorithm.

9) Fed feature vectors in the testing set into the trained EM in step 8 and thus the detection accuracy can be obtained.

10) Repeat the steps 6 and 9 several times and employ the average detection accuracy to evaluate the effectiveness of the extracted feature to reduce the detection performance caused by the various selections of the training set.

IV. EXPERIMENTAL RESULTS

The proposed method is expected to find the underlying dependencies of current node and neighbor nodes in all direction. The spliced 2D image should be detected after undergoing the calculation of DCT and DWT values as the input for Noncausal model with EM classifier. The image that is given should be uncompressed image from dataset UCID.

V. CONCLUSION

The Markov model was the famous and effective tools for image forgery detection. All among the image splice detection approaches based on the Markov models treats the image as a 1-D causal signal. Therefore only state dependencies between adjacent states along certain directions i.e. horizontal, vertical were considered in traditional model.

To overcome this consideration, a 2-D noncausal image Markov model is proposed in this paper that considers 2-D image. Each state in the proposed model depends on its neighbor states simultaneously. We decompose the noncausal model into four 2-D causal submodels and determine the value these submodels sequentially. These causal submodels are elaborated by three parameters i.e. the prior probabilities of each state, the parameters of probability density function of each state and the state transition probability matrix. All these model parameters are the input features for classification. We apply two domain i.e. the BDCT domain and DMWT domain and make the model effective. To classification Noncausal Markov Model is used with EM algorithm to detect the spliced images.

REFERENCES

[1] Xudong Zhao, Shilin Wang, Shenghong Li and Jianhua Li, "Passive Image-Splicing Detection by a 2-D Noncausal Markov Model", IEEE Transactions On Circuits And Systems For Video Technology, Vol. 25, No. 2, February **2015**.

[2] J. Li, A. Najmi and R. M. Gray, "Image classification by a twodimensional hidden Markov model," IEEE Trans. Signal Process, Vol. 48, No. 2, pp. **517–533**, Feb. **2000**.

[3] Y. Q. Shi, C. Chen, G. Xuan, and W. Su, "Steganalysis versus splicing detection," in Proc. 6th Int. Workshop Digit. Watermarking (IWDW), Guangzhou, China, pp. **158–172**, **2008**.

[4] P. Sutthiwan, Y.-Q. Shi, J. Dong, T. Tan, and T.-T. Ng, "New developments in color image tampering detection," in Proc. IEEE Int. Symp. Circuits Syst. (ISCAS), Paris, France, pp. **3064–3067,** May/Jun. **2010**.

[5] W. Wang, J. Dong, and T. Tan, "Image tampering detection based on stationary distribution of Markov chain," in Proc. 17th IEEE Int. Conf. Image Process. (ICIP), Hong Kong, pp. **2101**–**2104**, Sep. **2010**.

[6] S. Patchara, Y. Q. Shi, H. Zhao, T.-T. Ng, and W. Su, "Markovian rake transform for digital image tampering detection," Trans. Data Hiding Multimedia Secur. VI, Vol. 6730, pp. **1**–**17**, Jul. **2011**.

[7] Preeti Yadav, Yogesh Rathore, "Detection of Copy-Move forgery of images using Discrete Wavelet Transform", International Journal on Computer Science and Engineering (IJCSE), Vol. 4 , pp **565**-**570**, April **2012.**

AUTHORS PROFILE

Ms. Thofa Aysha is currently perceiving her Master of Technology degree in Computer Science and Engineering in Srinivas Institute of Technology Mangaluru of Visvesvaraya Technological University (VTU), Belagavi. She received her B.E degree in Information Science and Engineering in Vivekananda College of Engineering and Technology Puttur of Visvesvaraya Technological University (VTU) university, Belagavi. Her interest is in the area of Digital Image Processing, Advanced Algorithm.

Mr. Manjesh R is working as Assistant Professor in Srinivas Institute of Technology, Mangaluru, from past five years. He has received his Master of Technology in Computer Cognition Technology from University of Mysore, Mysuru. He has guided many UG and PG students projects and published many papers in national and international conferences and journal. His interested area are Data Compression, Data Mining and Image Processing.