International Journal on Recent and Innovation Trends in Computing and Communication ISSN: 2321-8169
Volume: 4 Issue: 9 42 - 46

Improved Tampering Localization in Digital image Forensics: Comparative Study Based on Maximal Entropy Random Walk and Multi-Scale Fusion

Madhavi Kunte
Department of Information Technology,
Pimpri Chinchwad College of Engineering,Pune,India
Email: madhavikunte44@gmail.com

Shweta Gangawane
Department of Information Technology,
Pimpri Chinchwad College of engineering,Pune,India
Email:shwetagangawane3195@gmail.com

Ketaki Kate
Department of Information Technology,
Pimpri Chinchwad College of Engineering,Pune,India
Email: ketukate@gmail.com

Rohini. G. Pise
Asst. Prof. Department of Information Technology,
Pimpri Chinchwad College of Engineering,Pune,India
Email:rohinipise @gmail.com

Abstract— Nowadays the increasing ease of editing digital photographs has spawned an urgent need for reliable authentication mechanism capable of precise localization of potential malicious forgeries. In this paper we compare two different Techniques to analyze which technique can be used more efficiently in localization of Tampered Region In Digital Image .First Technique is Maximal Entropy Random Walk in which Strong localization property of this random walk will highlight important regions and to diminish the background- even for noisy response maps.

Our evaluation will show that the proposed method can significantly perform both the commonly used threshold-based decision, and the recently proposed optimization approach with a Markovian prior.

The second Technique which is based on Multi-Scale Fusion will investigate a multi-scale analysis approach which merge multiple candidate tampering maps, obtained from the analysis with different windows, to obtain a single, more efficient tampering map with better localization resolution. We propose three different techniques for multi- scale fusion, and verify their feasibility .In this slant we consider popular tampering scenario to distinguish between singly and doubly compressed regions

Keywords:c *Maximum Entropy Random Walk, Multi-Scale Fusion, Tampering Localization, Markov random walk, JPEG splicing ,Digital Image Forensics.*

*****

I.INTRODUCTION

Precise localization of tampered image regions is one of the most challenging problems in digital image forensics .While many forensic features are known to differ between pristine and tampered images, their application for blind localization poses additional challenges .Though proactive image protection schemes can deliver precise identification of the tampered regions and even restore their original appearance with very high-quality they can only be exploited in a strictly controlled environment ,since they use a carefully designed digital watermark that needs to be available as side information. We propose to use the maximal entropy random walk (MERW) on a graph to post-process the response map. Random walks, including MERW, have been successfully adopted for many problems including image segmentation, link prediction and saliency estimation . Our approach is inspired by recent successful use of MERW for detection of salient objects in digital photographs In conventional random walk, only local knowledge is available to the walker when making the transitions. In MERW the walker is aware of the full graph structure, which is beneficial for many applications. MERW is characterized by a strong localization property which will highlight important regions of the map, and attenuate the background.

The Three novel fusion methods that exploit the dependencies between successive scales of analysis .Firstly, we consider an energy-minimization approach, which uses a MRF to model the prior knowledge about the tampering maps. Secondly, we consider two dual heuristic strategies , referred to as bottom-up and top-down fusion, which exploit the expected dependencies between the tampered regions in small and large-scale analysis. Our analysis is performed on the example case of a popular tampering scenario involving splicing of JPEG images that produces regions with different compression history. The forensic features of choice are the mode-based first digit features (MBFDF) . While recent studies have already demonstrated that MBFDF can be used successfully for sliding window- based localization , we conduct a much more detailed analysis with densely sampled compression levels and emphasis on multi-scale localization.

II.MOTIVATION

The major challenges in tampering is the precise localization of tampered region in the image ,selection of proper edges and vertices in the tampered image by this two technique Maximal Entropy Random Walk and Multi-scale Fusion .based on the edges that would be classified, and also after a huge number of calculations and selection of parameters the accuracy of both the technique is calculated, hence we are proceeding with the supervised learning techniques like sliding window analysis, binarization, post-processing that could help to locate the tampered region more accurately.

III. MAXIMAL ENTROPY RANDOM WALK

A random walk is mathematical formalization of a path

that consists of a succession of random steps. For example, path traced by a molecule as it is in a liquid or a gas state, the search path of a foraging animal, due to super string behavior, the price of a fluctuating stock and the financial status of a gambler can all be modeled as random walks, even if they may not be truly random in reality. The term random walk was first introduced by Karl Pearson in 1905.
Uses of random walks in many fields are: ecology, economics, psychology, computer science, physics, chemistry, and biology. Observed behaviors are explained by random walks of many processes in these fields, and hence serve as a fundamental model for the stochastic activity which was recorded earlier. Various different types of random walks are

- Random Walk on Markov Chain
- Random Walk on graphs
- Random Walk on lines
- Random Walk on Plane

Random walks also vary with the time parameter. The walk is in discrete time, and indexed by the natural numbers, However, some walks take their steps at random times, and in that case the position is defined for the continuum of times.

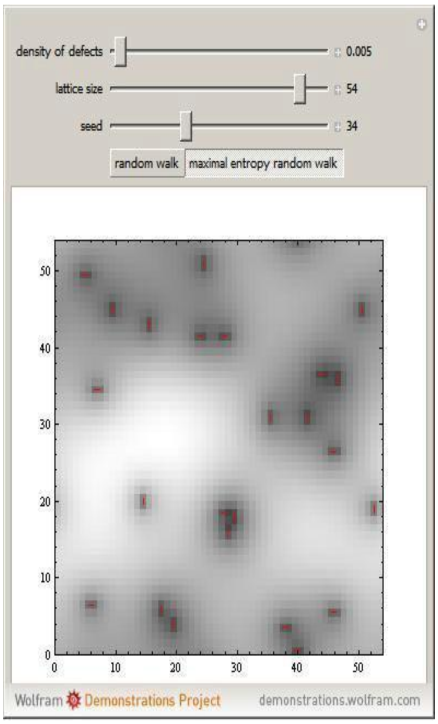


FIG 1: Random Walk

The plots compares the probability of finding a particle performing a random walk on a 2D square lattice with randomly istributed defects having Generic Random Walk (GRW) and Maximal Entropy Random Walk (MERW). The darker a region, the lower is the stationary probability of finding a particle there. In GRW the particle chooses one of the neighboring nodes with equal probability, while for MERW the particle moves in such a way that all trajectories of given length between two given points are improbable. By changing the density of defects, that is, of randomly erased links (marked as red in the plot), you can see that the probability spans the whole lattice for GREW while MERW is placed in the largest region free of errors.
Applications of random walks in statistical physics, economics, biophysics, and many other disciplines. We compare two types of random walk on the graphs: GRW (Generic Random Walk) and MERW (Maximal Entropy Random Walk). The particle which is performing GRW jumps in a single step to one of the neighboring nodes of the graph with equal probability. One can show that whenever the graph (lattice) is regular, different random walk trajectories of the same nodes that are having start and end and have the same length are improbable. One can, therefore, equivalently define GRW by requiring that all trajectories with the same length and endpoints are improbable.
In GRW the particle diffuses on the whole lattice while the diffusion area is constrained to the largest lattice region which is free of defects for MERW. It is very similar to the phenomenon of localization which is known from quantum physics. Once the particle is trapped in this region, it will have no chance to leave it again because of the "entropy barrier". The localization occurs for any density of defects if the lattice is sufficiently large. If the defects are distributed at random and uniformly, the localization volume grows with the logarithm of the system size.

This maximal entropy random walk is equivalent to generic random walk if it takes place on regular lattice, but it is not if the underlying lattice is irregular. Particularly, we consider a lattice with weak dilution. We show that how the stationary probability of finding a particle which is performing a maximal entropy random walk localizes in one of the largest nearly spherical region of the lattice which is free of defects. This localization phenomenon is purely classical in nature, is explained in terms of the Lifshitz states of a certain random operator.

The transition probabilities can be formally defined as:

$$p_{ij} = \frac{w_{ij} u_j}{\lambda u_i}$$

where is the largest eigenvalue of the weight matrix, and denote the i -th and the j-th component of the corresponding eigenvector . Since the weight matrix is non-negative, both and are non-negative according to the Frobenius-Perron theorem.

MERW possesses a strong localization property, which will lead to suppression of low-degree ( ) nodes and stronger highlighting of high-degree nodes. Hence, by properly choosing the weights of the graph, it is possible to exploit this phenomenon for exposition of important image regions.

IV MERW FOR TAMPERING LOCALIZATION

The localization procedure is summarized in Fig. 2. The input to the model is a real-valued response map of a forensic detector with detection scores for individual authentication units denoted as . The extremes of the response range represent the degree to which a certain forensic feature is present / absent (or alternatively the confidence of the classifier regarding the region's authenticity).

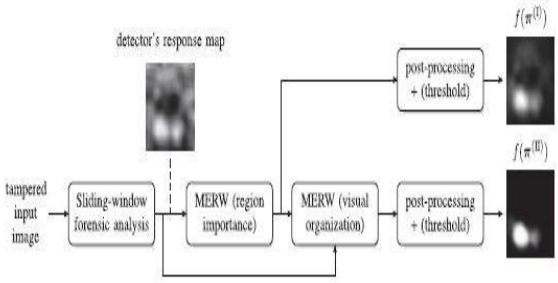


Fig 2. Proposed localization framework

Successive maximal entropy random walks highlight important regions and improve visual organization of the decision map typical post- processing includes map normalization, brightness correction and optimally binarization

We span a fully-connected graph on non-overlapping image blocks of size pixels (px)1. Similarly to , we consider a two-layer approach where two successive MERWs aim to identify important regions of the input map, and improve visual organization of the output map, respectively. We will separately assess the improvement of each of these layers. The stationary probability densities and , after normalization to [0,1]& prospective post-processing or binarization, will constitute the final decision map for the considered scenarios.

V. MULTI-SCALE ANALYSIS IN DIGITAL IMAGE FORENSICS

We consider a popular forgery, in which, as a result ofcontent replacement, some fragments of the JPEG image have different compression history and exhibit either single or double compression artifacts. In this section, we provide a detailed analysis of the multi- scale localization problem based on MBFDFs. We make our discussion as general as possible to facilitate easier generalization to other forensic features.

The editing digital photographs has spawned an urgent need for reliable authentication mechanisms, capable of precise localization of potential malicious forgeries. Though proactive image protection schemes can deliver precise identification of the tampered regions and even restore their original appearance with very high-quality , they can only be exploited in a strictly controlled environment, since they use a carefully designed digital watermark that needs to be available as side information.

The rapidly developing field of digital image forensics aims to deliver passive authentication mechanisms that analyze intrinsic fingerprints introduced on various stages of the image acquisition pipeline . As a result, such techniques can be applied to existing, non-watermarked images. However, since they are built on the foundations of machine learning and statistical signal analysis, reliable detection of forensic features renders precise tampering localization a challenging problem Multi-scale analysis is a well-established approach in various areas of image processing and computer vision. It has been used for object detection image blending, depth of field blending , denoising , speeding up convergence of optimization algorithms finding similar patches (e.g., in super-resolution , and many more. The most relevant works for the application at hand include object detection, image texture segmentation and depth map estimation which involve decision fusion. Object detection involves identification of only the bounding box of the object. Hence, simple voting mechanisms typically suffice. In the texture segmentation problem, the fusion was performed by averaging the candidate likelihood functions, weighted by discriminative capabilities of different scales. For patch-based depth map estimation, the authors used a MRF to model the multi-scale fusion problem.

Fusion techniques (not necessarily multi-scale) are also used in saliency detection. The input maps may represent various saliency estimation techniques , various types of saliency (static vs. dynamic) , or even co-saliency among many images A recent study compares the performance of many fusion methods [, most of which are just variations of well known averaging, maximization, or multiplication approaches with custom weights for individual components. The best performance was obtained by the simplest unweighted averaging. Decision fusion in saliency can also involve combining bottom-up and top-down clues . Bottom-up saliency refers to low-level image features, e.g., high-contrast regions that attract attention. Top-down saliency refers to high-level, task-oriented recognition tasks, e.g., face, object or text recognition.

Successful fusion of multi-scale tampering maps constitutes a challenging problem and requires new techniques to properly exploit inter-scale dependencies between the candidate maps. Existing methods commonly assume independence of the scores both in the spatial and in the scale dimension] A brief argument about this simplification for single- scale PRNU-based localization can be found in . Therefore, in multi-scale fusion of tampering maps, the aspect of inter-scale correlations calls for a separate, more detailed analysis.

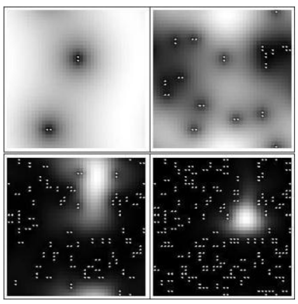


Fig. 3. Idealized fusion of 3 candidate maps c(s), corresponding to a small, medium, and large analysis windows, into a single accurate tampering map.

B. Localization Using First Digit Features

Our feature space contains 180 features - the MBFDFs of all 9 digits from first 20 AC coefficients. Classification is performed by a SVM classifier with a radial basis function (RBF) kernel and trained to yield probability estimates of the decisions

VI. Comparative Study Based On Maximal Entropy Random Walk And

Multi-Scale Fusion Maximal Entropy Random Walk This maximal entropy random walk is equivalent to generic random walk if it takes place on a regular lattice, but it is not if the underlying lattice is irregular. MERW, have been successful adopted for many problems including image segmentation, link prediction and saliency estimation .

Maps Used By The Technique A Real Valued Response Map:The real valued response map works an input to the model of a forensic detector with detection scores for individual authentication units denoted as ciε [0,1]. The extremes of the response range represent the degree to which a certain forensic feature is present/absent (or alternatively the confidence of the classifier regarding the region's authenticity.

B Decision Map

We consider a two-layer approach where two successive MERWs aim to identify important regions of the input map, and improve visual organization

of the output map, respectively. We will separately assess the improvement of each of these layers. The stationary probability densities and, after normalization to [0,1]& prospective post-processing or binarization, will constitute the final decision map for the considered scenarios Complexity of the considered graph makes it impossible to compute the optimal weights analytically. Hence, we consider a few representative heuristic selection strategies motivated by the properties of MERW. We have also evaluated more sophisticad strategies. However, they provided marginal improvement and hence will not be presented in this paper.

Since the walker will concentrate on highly connected regions (localization property), greater weights should be assigned between tampered nodes. Edges between pristine regions should have low weights. Hence, one possible strategy is to make the weight proportional to the average response of the source and target node . A more aggressive strategy would ignore the source node, and encourage transitions to target nodes with confident response .

PARAMETERS USED IN MERW

A. Selection of Graph's Weights

Complexity of the considered graph makes it impossible to compute the optimal weights analytically. Hence, we consider a few representative heuristic selection strategies motivated by the properties of MERW. We have also evaluated more sophisticated strategies. However, they provided marginal improvement and hence will not be presented in this paper.

Since the walker will concentrate on highly connected regions (localization property), greater weights should be assigned between tampered nodes. Edges between pristine regions should have low weights. Hence, one possible strategy is to make the weight proportional to the average response of the source and target node.

MULTI-SCALE ANALYSIS IN DIGITAL IMAGE FORENSICS

Multi-scale fusion involves computation of the final tampering map, given a set of candidate maps obtained from various scales of analysis..

1 Detailed Analysis of Candidate Maps

Forensic classifiers may not obey asymmetry of the decision confidence when distinguishing between tampered and pristine regions. This phenomenon will also be visible in this study, as the decisions in favor of double compression tend to have higher confidence leading to subtle different behavior depending on the considered tampering scenario. In order to facilitate more generalization to other forensic features, we do it by considering two tampering scenarios corresponding to the presence of double JPEG compression either inside or outside of the tampered regions, in short referred to as the double- inside or single-inside scenarios, respectively[11]. Regardless of the scenario, we expect the candidate maps to indicate tampered regions with scores ≈ 1 and pristine regions with scores ≈ 0. Hence, for the single inside scenario the candidate scores will correspond to 1 - classification scores.

2 Determining Candidate Map Reliability

Small scale analysis yields less reliable candidate maps, which may unnecessarily introduce noise to the fusion process. Additionally, large scale analysis may produce empty candidate maps if the tampered region is smaller than the analysis window. Since such candidate maps do not contribute information useful for tampering localization, ideally they should be ignored by the fusion procedure. In this section, we describe a simple algorithm that allows to quickly estimate whether a candidate map is useful or not[11].

CONSIDERED FUSION STRATEGIES

Different Algorithm

A. Fusion by Energy Minimization

A Bayesian approach to tampering localization would involve finding the optimal tampering map tˆ that maximizes the posterior probability given a set of candidate maps:

tˆ = argmax P(t|c(s) : s = 1,2,..., S). (1)

t∈{0,1} N

Then, ignoring the irrelevant constant term, the problem can be rewritten as:

tˆ = argmax P(c(s) : s = 1,2,..., S|t)P(t).

(2) t∈{0,1}N

Due to analytical tractability issues, full independence of the observations for individual authentication units is commonly assumed, leading to a simpler formulation:

N

$\hat{\mathbf{t}}$ = argmax*N* *P* $(c_i^{(s)} : s = 1, 2, \ldots,$
S|*ti*)*P*(**t**). (3)
t∈{0,1} *i*=1

Analogously, we find it more practical to assume interscale independence of the candidate scales at this point, and introduce a simpler heuristic mechanism for exploiting these dependencies at a later stage. Then, the problem becomes:

N *S*

$\hat{\mathbf{t}}$ = argmax*N* $P(c_i^{(s)} \; t_i) P(\mathbf{t})$. (4)
t∈{0,1} *i*=1 *s*=1

The prior of the tampering map *P*(**t**) can be conveniently modeled with a MRF. Then, the decision for each authentication unit will depend only on its direct neighborhood. Assuming a 1st order neighborhood, the decision regarding t_i will depend only on up to 4 of its neighbors t_j : $j \in \Xi_i =$ $\{i-1, i+1, i-N_h, i+N_h\}$, corresponding to the top, bottom, left, right neighbors, respectively. Obviously, at the image borders the set of neighbors i needs to be pruned accordingly.

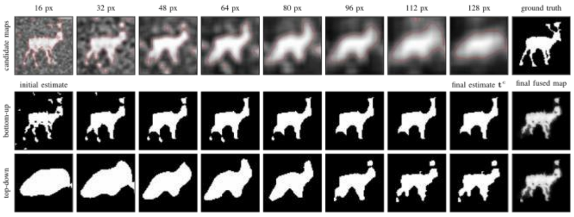
In practice it is often more convenient to represent the MRF in terms of Gibbs potentials, and reformulate the problem into energy minimization [44], [45]. Gibbs potentials use probabilities in the form:

$$P(\mathbf{t}) = Z^{-1} e^{-U(\mathbf{t})} = Z^{-1} e^{-\sum_{c \in C} V_c(\mathbf{t})}, \qquad (5)$$

where Z is a normalizing constant, and U is an energy function defined as a sum of potentials V_c on individual *cliques* - small groups of neighboring nodes in a graphical model of the MRF (authentication units in the problem at hand). Similarly to Chierchia et al. [7], we resort to the popular Ising model [45] which considers single- element and two-element

A. Bottom-Up and Top-Down Fusion

The proposed bottom-up (BU) and top-down (TD) fusion are heuristic approaches exploiting the observation that large scale analysis is expected to be more reliable, yet with a worse resolution of analysis.



As a reference for performance comparison we consider 4 fusion strategies representing various approaches to the problem.

VII. CONCLUSION

Our study shows the comparision between Maximal Entropy Random Walk and MultiScale Fusion In conclusion, the major contributions of our work include:

1. This paper shows fusion of candidate maps obtained on multiple scales can improve the tampering localization of sliding window based detectors.
2. Maximal Entropy Random Walk can be successfully adopted for tampering localization in digital image forensics.
3. In this paper we conclude that Maximal Random Walk Technique gives more accuracy than that of the Multi-scale Fusion Technique.
4. In Multi-scale Fusion it involves various approaches like top-down, bottom-up, Energy Minimization. Whereas Maximal Random Walk Technique does not include all these approach of multi-scale fusion for other forensic features (e.g., PRNU or splicing detectors based on rich feature sets). Positive results would indicate feasibility of a combined multi-scale and multi-modal approach

VIII. REFERENCES

[1]. M. Chen, J. Fridrich, M. Goljan, and J. Lukas, "Determining image origin and integrity using sensor noise," IEEE Trans. Inf. Forensics Secur., vol. 3, no. 1, pp. 74–90, 2008.

[2]. G. Chierchia, G. Poggi, C. Sansone, and L. Verdoliva, "A Bayesian-MRF approach for PRNUbased image forgery detection," IEEE Trans. Inf. Forensics Secur., vol. 9, no. 4, pp. 554–567, 2014.

[3]. I. Amerini, R. Becarelli, R. Caldelli, and A. Del Mastio, "Splicing forgeries localization through the use of first digit features," in Proc. IEEEInt. Workshop on Information Forensics and Security, 2014.

[4]. D. Cozzolino, D. Gragnaniello, and L. Verdoliva, "Image forgery detectionbased on the fusion of machine learning and block matching methods," in , 2013 [Online]. Available: http://arxiv.org/abs/1311.6934

[5]. L. Verdoliva, D. Cozzolino, and G. Poggi, "A feature-based approach for image tampering detection and localization," in Proc. IEEE Int. Workshop on Information Forensics and Security, 2014.

[6]. G. Chierchia, D. Cozzolino, G. Poggi, C. Sansone, and L. Verdoliva, "Guided filtering for PRNU-based localization of small-size image forgeries," in IEEE Int. Conf. Acoustics, Speech and Signal Processing,

[7]. P. Korus and A. Dziech, "Efficient method for content reconstruction with self-embedding," IEEE Trans. on Image Processing, vol. 22, no. 3, pp. 1134–1147, March 2013.

[8]. P. Korus, J. Bialas, and A. Dziech, "Towards practical self-embedding for JPEG-compressed digital images," IEEE Trans. on Multimedia, vol. 17, no. 2, pp. 157–170, Feb 2015.

[9]. S. Sarreshtedari and M.A. Akhaee, "A sourcechannel coding approach to digital image protection and self-recovery," IEEE Trans. on Image Processing, vol. 24, no. 7, pp. 2266–2277, July 2015.

[10]. H. Farid, "Image forgery detection," IEEE Signal Processing Magazine, vol. 26, no. 2, pp. 16–25, March 2009.

[11]. P.Korus "Multiscale Fusion For Improved Localization Of Malicious Tampering Of Digital Images,"