ISSN:0975-9646

An Active Approach based on Independent Component Analysis for Digital Image Forensics

Anil Dada Warbhe[#], R. V. Dharaskar[*]

[#]*Department of Electronics Engineering,*
Manoharbhai Patel Institute of Engineering and Technology, Gondia, India

[*]*Disha Technical Campus, DMAT*
Raipur, India

***Abstract*— Advancements in digital imaging technologies has elevated many new issues and challenges concerning the authenticity and integrity of digital images. Digital images can now be easily captured and edited for creating forgery without leaving any obvious clues of such operations. These capabilities undermine the credibility of digital images in all aspects. Digital image forensics is has gained tremendous importance in last one decade among the research community. Digital Image Forensics (DIF) aims at determining the origin and potential authenticity of a digital images. One of the fundamental problems digital image forensics techniques attempt to solve is the identification of the source of a digital image. In case if the image is challenged in the court of the law it is very important to prove and identify the original image from the tampered and forged one. In this paper we present an active approach which is based on Independent Component Analysis (ICA). The experiments carried out proves that the ICA can be used as an effective and robust tool for identifying the original image from the forged one.**

***Keywords*— digital image forensics, image forgery, image security, Image Authentication, Active Forgery detection.**

I. INTRODUCTION

In recent years, due to the wide spread availability of smart gadgets like mobile phones, hand held computing devices which are often equipped with digital cameras; capturing images and uploading it over the World Wide Web became relatively straightforward. The accelerated advancement of social media, popular messaging applications like Whatapp and online storage websites make these personal data, such as digital images, more accessible than ever before. Digital images are prone to manipulations and if there images are accessed by a person with malign intent, he or she can make use of them to create forgery and spread it in the network. In situations, such as for evidence in a court of law or insurance digital imaging, even a small amount of ambiguity can change the judgment. Therefore, protecting digital image integrity has become important. In such scenarios the authenticity of the digital image can be assured by utilizing tamper detection algorithms.

The tamper detection algorithms pertaining to the digital image forensics are classified as active tamper detection approaches and passive detection approaches. Passive tamper detection approach do not require the knowledge of any prior information about the content. The core assumption for this class of techniques is the assumption that original non-forged content owns some inherent statistical pattern introduced by the generative processing. Such patterns are always consistent in the un-forged content, but they are very likely to be altered after some tampering processes. Although visually imperceptible, such changes can be detecting by statistical analysis of the content itself, without the need of any a-priori information. Thus they are said to be passive and blind. [1].

On contrary, in active approach makes use of image watermarking schemes and digital signatures. So any tampering operation can be detected by extracting the watermark or the digital signatures and the images can be authenticated. But researchers are lacking confidence in it due to its critical limitations. The major drawback is that a watermark must be inserted at the time of recording, which would limit this approach to specially equipped digital cameras [2]. Also it’s a costlier affair and do need the original source image handy. Some of the active approach based tamper detection methods from the literature are discussed below.

Lin et al. [3] proposed a digital image authentication algorithm which divides the original image into blocks of 16×16 pixels. Each block of an image is processed using the Discrete Cosine Transform (DCT). Then, the coefficients of the converted pixels are embedded with the generated watermark information. The proposed algorithm achieved a 90% tamper detection rate for digital images with slight compressions.

Ho et al. [4] proposed a semi-fragile tamper detection algorithm utilized the Pinned Sine Transform (PST) domain. In their proposed algorithm, the digital image is divided into blocks of 8×8 pixels, and the pinned and boundary field of all the divided blocks is generated using PST. The embedding procedure for their proposed algorithm locates the high frequency coefficients of each divided block in the pinned field.

Zhao et al. [5] proposed two tamper detection algorithms using active and passive techniques. The proposed active watermarking performed the embedding operation using Slant Transform (SLT). Their proposed algorithm divided the original image into blocks of 8×8 pixels and extracted the seven most significant bits of the pixels to generate the watermark information. The algorithm is a semi-fragile watermarking algorithm and achieved 98% tamper detection rate for copy-move tampering attacks. These semi-fragile algorithms do not have self-recovery capability.

Lee and Lin [6] proposed a dual watermarking scheme that embeds two copies of the watermark payload into two different positions. Their proposed method offers a second chance for self-recovery in case the primary information is destroyed.

Huang et al. [7] proposed a fragile image authentication for color images that utilizes the concept of block truncation coding (BTC) for the embedding procedure; the 2-D transformation is used for determining the embedding bit location. Their proposed authentication methods create a block dependency in such a way that their algorithm is vulnerable to security attacks, such as the collage attack and four-scanning attack.

Patra et al. [8] proposed a fragile self-recovery watermarking for digital image authentication. In their proposed scheme, Chinese remainder theorem (CRT) is used and provides a computational advantage with additional security measures. However, because the block size for dividing the original image is 8×8 pixels, their scheme fails to accurately localize the tampered regions and the quality of the recovered image for attacks with a small tamper ratio is 36.77 dB.

Change et al. [9] proposed fragile watermarking for tamper detection and recovery. Their proposed scheme utilized local binary patterns (LBP) to construct tamper detection watermark bits. In Chang's proposed scheme, a discrete Torus is used for block mapping and two levels of tamper detection will be conducted to locate the tampered regions.

Tong et al. [10] proposed a chaos-based fragile watermarking for self-recovery. In their scheme, a new chaos-based algorithm for encrypting the watermark and generating block mapping is applied. Their proposed algorithm scrambles the original image by utilizing chaotic maps and divides the image into blocks of 2×2 pixels to obtain better tamper localization.

The existing tamper detection methods are divided in two groups: the first which contains the algorithms that can only authenticate the digital image and locate the tampered regions. And, the second group which contains the algorithms that have the ability to recover the destroyed sections of the digital images along with locating the tampered sections. Tamper detection algorithms with self-recovery capabilities are desirable for protection of the content and the integrity of the digital images. Localizing the tampered sections along with self-recovery capability can make a difference in several profound circumstances, such as accident imaging for insurance companies. These algorithms utilize certain block information as a watermarking payload to be embedded in the same or different blocks.

In this paper we introduce a non-intrusive way of image authentication and forgery detection using independent component analysis. Which do not need any watermark or digital signature to be embedded into the image at the time of its capture. What is needed is the only the original image and the tampered image. The method can be made more robust by adding the digital signature into the image. The paper is organized as follows. In section 2, we present the methodology used for tamper detection. Section 3, illustrates the results and finally conclusion in section 4.

II. METHODOLOGY

Recently, there has been an increasing interest in statistical models for learning data representations. A very popular method for this task is independent component analysis (ICA), the concept of which was initially proposed by Comon [11]. The ICA algorithm was initially proposed to solve the blind source separation (BSS) problem i.e. given only mixtures of a set of underlying sources, the task is to separate the mixed signals and retrieve the original sources [12]. Neither the mixing process nor the distribution of sources is known in the process. A simple mathematical representation of the ICA model is as follows.

Consider a simple linear model which consists of N sources of T samples i.e. $S_i = [S_i(1), ..., S_i(t), ..., S_i(T)]$. The symbol there represents time, but it may represent some other parameter like space. M weighted mixtures of the sources are observed as X, where $X_i = [X_i(1), ..., X_i(t), ..., X_i(T)]$. This can be represented as –

$$X = A\,S + n; \qquad (1)$$

Where

$X = (X_1, X_2, X_3 \ldots\ldots\ldots, X_M)$; $S = (S_1, S_2, S_3 \ldots\ldots\ldots, S_N)$ and $n = (n_1, n_2, n_3 \ldots\ldots\ldots, n_k)$.

S and n represent the additive white Gaussian noise (AWGN).

It is assumed that there are at least as many observations as sources i.e. M = N. The M × N matrix A is represented as –

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1N} \\ \vdots & \ddots & \vdots \\ a_{M1} & \cdots & a_{MN} \end{bmatrix}; \qquad (2)$$

A relates *X* and *S*. *A* is called the mixing matrix. The estimation of the matrix *S* with knowledge of *X* is the linear source separation problem. The source separation problem cannot be solved if there is no knowledge of either *A* or *S*, apart from the observed mixed data *X*. If the mixing matrix *A* is known and the additive noise n is negligible, then the original sources can be estimated by evaluating the pseudo inverse of the matrix *A*, which is known as the un-mixing matrix *B*, such that

$$BX = BAS = S \qquad (3)$$

For cases where the number of observations M equals the number of sources N (i.e. $M = N$), the mixing matrix A is a square matrix with full rank and $B = A^{-1}$.

The necessary and sufficient condition for the pseudo-inverse of *A* to exist is that it should be of full rank. When there are more observations than the sources (i.e. $M > N$), there exist many matrices *B* which satisfy the condition $BA = I$. Here the choice *B* depends on the components of S that we are interested in. When the number of observations is less than the number of sources (i.e. $M < N$), a solution does not exist, unless further assumptions are made. On the other side of the problem, if there is no prior knowledge of the mixing matrix *A*, then the estimation of both *A* and *S* is known as a blind source separation (BSS) problem. A very popular technique for solution of a BSS problem is independent component analysis [13]. Estimation of the underlying independent sources is the primary objective of the BSS problem. The problem defined in (3), under the assumption of negligible Gaussian noise n, is solvable with the following restrictions:

- The sources (i.e. the components of *S*) are statistically independent.
- At most, one of the sources is Gaussian distributed.
- The mixing matrix is of full rank.

Basically almost all ICA algorithms are good for separation of the instantaneous mixture of the non-Gaussian sources. We here assume that the mixture is instantaneous. Firstly we feed two images, i.e., forged and original image, we mix both the images instantaneously. Then this mixture is fed to an improved version of the FastICA [13-15]. We have used EFICA [16], algorithm which is asymptotically efficient, i.e., its accuracy given by the residual error variance attains the Cramér-Rao lower bound. The error is thus as small as possible. Finally, we get the estimated independent components as separate image output, making us able to identify the forged regions.

III. Experimental Results

Experiment is performed using several images using the original and forged images. The output of the method cleary detects the forged regions.

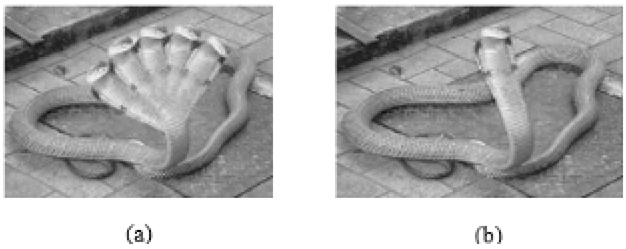

Fig 1: The five head king cobra hoax, (a) a forged image (b) original image

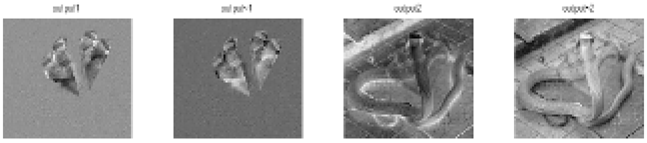

Fig 2: extracted forged section and original image

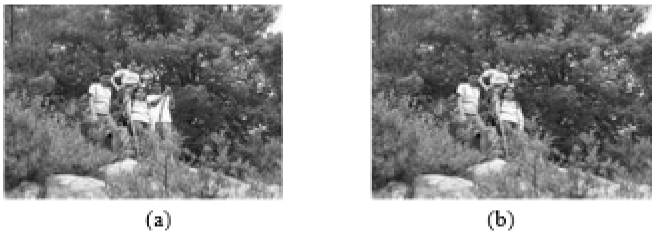

Fig 3: four student image, (a) a forged image (b) original image

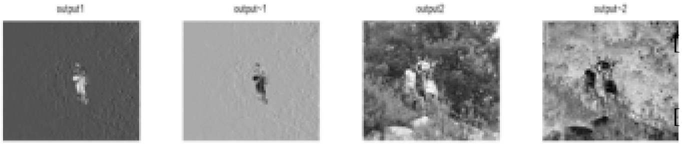

Fig 4: extracted forged section from the forged image

IV. Conclusions

In this paper we have presented an active approach to identify and authenticate the original digital image from the forged or tampered image. Forgery detection based on blind source separation using independent component analysis. The experiments included show how this new method successes in extracting and detecting the image forgery if any in the image. Though this method is good at detecting the forgery in the images the main limitation of this method is that, it needs the forged image as well as the original image which is been forged. Also to increase its robustness the original image can be embedded as a watermark in the image itself. This limitation can be overcome by using and applying a single channel independent component analysis on a single forged image to extract the forgery.

References

[1] Conotter, Valentina, et al. "Active and passive multimedia forensics." (2011).

[2] Farid, Hany. "Image forgery detection." Signal Processing Magazine, IEEE 26.2 (2009): 16-25.

[3] E.T. Lin, C.I. Podilchuk, E.J. Delp III, Detection of image alterations using semifragile watermarks, in: Electronic Imaging, International Society for Optics and Photonics, Bellingham, Washington, USA, 2000, pp. 152–163.

[4] A.T. Ho, X. Zhu, W. Woon, A semi-fragile pinned sine transform watermarking system for content authentication of satellite images, in: Geoscience and Remote Sensing Symposium, 2005. IGARSS'05. Proceedings. 2005 IEEE International, Seoul, Korea, vol. 2, IEEE, 2005, p. 4.

[5] X. Zhao, P. Bateman, A.T. Ho, Image authentication using active watermarking and passive forensics techniques, in: Multimedia Analysis, Processing and Communications, Springer-Verlag, Berlin, Heidelberg, 2011, pp. 139–183.

[6] T.-Y. Lee, S.D. Lin, Dual watermark for image tamper detection and recovery, Pattern Recognit. 41 (11) (2008) 3497–3506.

[7] S.-C. Huang, C.-F. Jiang, et al., A color image authentication and recovery method using block truncation code embedding [j], J. Mar. Sci. Technol. 20 (1) (2012) 49–55.

[8] B. Patra, J.C. Patra, Crt-based fragile self-recovery watermarking scheme for image authentication and recovery, in: IEEE International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS 2012), pp. 430–435.

[9] J.-D. Chang, B.-H. Chen, C.-S. Tsai, Lbp-based fragile watermarking scheme for image tamper detection and recovery, in: 2013 IEEE International Symposium on Next-Generation Electronics (ISNE), IEEE, Tamsui, New Taipei City, Taiwan, 2013, pp. 173–176.

[10] X. Tong, Y. Liu, M. Zhang, Y. Chen, A novel chaos-based fragile watermarking for image tampering detection and self-recovery, Signal Process.: Image Commun. 28 (3) (2013) 301–308.

[11] P. Comon, “Independent Component Analysis-A new concept?” Signal Processing, vol. 36, pp. 287-314, 1994.

[12] J.F.Cardoso, “Blind Signal Separation: Statistical Principles”, Proc. of IEEE, vol. 9, no. 10, pp. 2009-2025, 1998.

[13] Aapo Hyvärinen et al., “Independent Component Analysis: Algorithms and Applications”, Neural Networks, 13(4-5):411-430, 2000.

[14] Aapo Hyvärinen et al., “Independent Component Analysis: Algorithms and Applications”, Neural Networks, 13(4-5):411-430, 2000.

[15] A. Hyvarinen, “Fast and robust fixed-point algorithms for independent component analysis”. IEEE Trans. Neural Netw.,vol.10,no.3,pp.624-634,May 1999.

[16] Koldovský, Z., Tichavský, P., and Oja, E.: Efficient Variant of Algorithm FastICA for Independent Component Analysis Attaining the Cram´er-Rao Lower Bound, IEEE Tr. Neural Networks, 17 (2006) 1265–1277.