An Algorithm for Image Time Series Forgery Detection Based on the Anomalies Detection

Nadezhda Evdokimova and Vladislav Myasnikov

Samara National Research University, Samara 443086, Russia,
nadezh.evdokimova@gmail.com

Abstract. The present work is devoted to the development and investigation of the algorithm for detection intentional distortions (forgeries) of a single digital image in an image time series (time sequences) of one scene. The proposed algorithm consists of three stages. At the first stage, a set of errors that were calculated during reconstruction the fragments of the analyzed image by the 'neighboring' ones are estimated. After errors were calculated throughout image, we analyze their distribution on the second stage. At the final stage, fragments of the analyzed image that are anomalies are selected as 'suspicious'. The proposed algorithm, unlike existing algorithms, will allow unified detection of such attacks as intra-image copy-move and inter-image copy-move. Also, copy-move fragments may fall under geometric transformation, linear enhancement and other distortions. The investigation results of intra-image copy-move and inter-image copy-move detection using the proposed solution are presented.

Keywords: time series, image time series, image forgery, detection, anomaly

1 Introduction

Nowadays the image forgery methods complexity increases in relation to their detection complexity. This is due to the number spheres increasing using digital images in their work, as well as their processing tools availability and popularization. Image time series show the scene dynamics and allow it to be compared over time. So, having an image time series of some scene, with some admission, you can model an image that will be next in the scene, or an image that may have been distorted.

This algorithm will allow detection spatial tampering attacks. At the moment, several detection techniques were development. They are technique based on camera's fingerprint, technique based on coding artifacts, and techniques that use temporal and spatial correlation [2]. Their main weakness is absence of robustness to different distortions. Proposed algorithm uses correlation between corresponding fragments of different images in the image series and allows detection intra-image copy-move and inter-image copy-move.

We will mean forgery in the sense of an anomaly to develop an algorithm for image forgery detection. In the general sense, an anomaly is a data fragment that does not correspond to the precisely defined concept of normal behavior [1]. In accordance with the definition given above, within the framework of this paper, we will consider forgery image regions as anomalies. The proposed algorithm uses the concept of anomaly in the sense of the least probable points [3][1].

The work consists of two parts, namely, description of the proposed algorithm and analysis of experiments results. The description of the proposed algorithm is subdivided into the three sections. The first section contains a characteristic of the image fragments description method. The second section defines the statistic construction method. The third section explains the rule for assigning fragments of the analyzed image to anomalies.

2 Description of the proposed algorithm

Let $I_t\left(n_1, n_2\right)$, $t = \overline{0, T}$ is an image series (time sequence) of one scene. Every image in the series has the same size $N_1 \times N_2$, $n_i \in \overline{0, N_i - 1}$ ($i = \overline{1, 2}$; $T \geq 1$).

For definiteness, we assume that the image $I_0(n_1, n_2)$ is checked, although it can be located in the sequence anywhere. We analyze a certain square region of the image $D(n_1, n_2) \subseteq \overline{0, N_1 - 1} \times \overline{0, N_2 - 1}$ in a sliding window with a position (n_1, n_2). For certain region $D(n_1, n_2)$, we have image fragments $I_t(m_1, m_2)$, where $(m_1, m_2) \in D(n_1, n_2)$. To simplify the exposition, the arguments of the region in the record $D(n_1, n_2)$ can be omitted. Experimentally, we have spotted the best in terms of detection quality and runtime window has 15×15 size.

2.1 Image fragments description

For each possible position of the window D in the image plane, the corresponding fragments $I_t(m_1, m_2)$ are successively divided into k fragments, $k = 2^p$, by clustering by brightness (a k-means clustering algorithm is used). An example of such splitting under $k = 2^2$ is shown in Figure 1:

$$\begin{bmatrix} 3 & 5 & 12 \\ 1 & 14 & 2 \\ 4 & 9 & 17 \end{bmatrix} = \begin{bmatrix} 3 & 5 & 0 \\ 1 & 0 & 2 \\ 4 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 9 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 12 \\ 0 & 14 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 17 \end{bmatrix}$$

Fig. 1. Splitting of the fragment $I_t(m_1, m_2)$ into $k = 2^2$ fragments

The new fragments of the image obtained in this way on step k can be denoted by $I_t^j(n_1, n_2)$, $j = \overline{0, k - 1}$. Next, we solve the problem of image I_0 fragment representation for this region by means of corresponding by the window D position fragments $I_1^0, I_1^1, ..., I_1^{k-1}; ...; I_T^0, I_T^1, ..., I_T^{k-1}$ linear combination, that is:

$$I_0 \approx \sum_{t=1}^{T} \sum_{j=0}^{k-1} \alpha_t^j I_t^j \tag{1}$$

[1] A comprehensive coverage of the field of outlier analysis from a computer science point of view can be found in [4]

using mean squared deviation ε_k^2 minimization:

$$\varepsilon_k^2 \cong \frac{1}{|D|} \sum_{(m_1,m_2)\in D} \left(I_0(m_1,m_2) - \sum_{\substack{1\le t\le T\\ 0\le j\le k-1}} \alpha_t^j I_t^j(m_1,m_2) \right)^2 \tag{2}$$
$$\rightarrow \min_{\alpha_1^0,\ldots,\alpha_1^{k-1},\ldots,\alpha_T^0,\ldots,\alpha_T^{k-1}} .$$

Then we calculate both types of errors the mean squared deviation and the normalized mean squared deviation that characterizes the conjugacy value and given by:

$$\tilde{\varepsilon}_k^2 = \frac{\varepsilon_k^2}{\sum\limits_{i=0}^{N_1-1} \sum\limits_{j=0}^{N_2-1} I_0(i,j)^2}. \tag{3}$$

We perform this procedure for $k = 4, 8, 16$. The first stage result for each position of the analysis region is a normalized mean squared deviations set of the analyzed image fragment representation. For convenience of further use, we designate them as a vector:

$$\bar{x}(n_1,n_2) \equiv \left(\tilde{\varepsilon}_4^2(n_1,n_2), \tilde{\varepsilon}_8^2(n_1,n_2), \tilde{\varepsilon}_{16}^2(n_1,n_2)\right)^T \tag{4}$$

2.2 Statistic construction method

We represent the obtained vectors $\bar{x}(n_1,n_2)$ set in the coordinate system $\tilde{\varepsilon}_4^2\tilde{\varepsilon}_8^2\tilde{\varepsilon}_{16}^2$. This set is located in the three-dimensional cube with sides equal to 1 as shown in Figure 2:

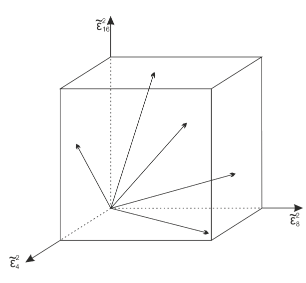


Fig. 2. The set of vectors $\bar{x}(n_1,n_2)$ in the coordinate system $\tilde{\varepsilon}_4^2\tilde{\varepsilon}_8^2\tilde{\varepsilon}_{16}^2$

2.3 Anomalies finding

There are no absolute unchanging objects on images obtained in real conditions. This is due both to real cameras properties that have their own noises and the information transfer path properties of from the camera to the processing system. In this path, the image is compressed before shipment and then decoded, which often leads to additional system distortions. Moreover, there are often objects on the scene that have certain dynamic characteristics although they are static in our understanding. For example, it may be trees swaying in the wind.

In accordance with this fact, it is impossible to obtain an errors vector with coordinates $(0; 0; 0)$ after authentic image fragment representation by means of neighboring images fragments linear combination. So we can conclude the errors vector with coordinates $(0; 0; 0)$ corresponds to the image region, which is a duplicate inserted from one of neighboring images of the image time series.

On the other hand, the errors $\tilde{\varepsilon}_4^2$, $\tilde{\varepsilon}_8^2$, $\tilde{\varepsilon}_{16}^2$ of an authentic fragment representation must have values that do not exceed a certain threshold. It is obvious the error value of the same fragment representation decreases with the clusters number increasing. Therefore, it is justified to use different thresholds for $\tilde{\varepsilon}_4^2$, $\tilde{\varepsilon}_8^2$ and $\tilde{\varepsilon}_{16}^2$. So the following relation should be observed:

$$T_{\tilde{\varepsilon}_4^2} \geq T_{\tilde{\varepsilon}_8^2} \geq T_{\tilde{\varepsilon}_{16}^2} \tag{5}$$

After the statistic construction stage, we analyze the distribution histograms and select the thresholds according to the relation (5) as shown in Figure (3). We choose first local minimum and consider its value a threshold.

Then the cube with the errors vectors $\bar{x}(n_1, n_2)$ set is splitted into three areas:

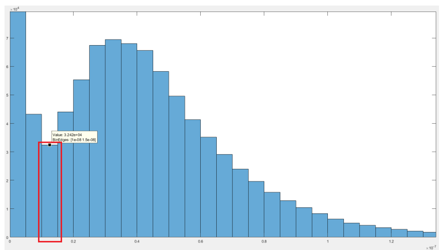
1) Origin of the coordinate system;
2) A parallelepiped that is adjacent to the origin;
3) Rest area of the cube.

In accordance with the above, vectors from the first area correspond image region which is a duplicate inserted from one of neighboring images of the image time series (inter-image copy-move). Vectors from second area refer to authentic image regions and vectors from third area correspond to fragments which is the duplicate within one image (intra-image copy-move). This splitting is shown in Figure 4.

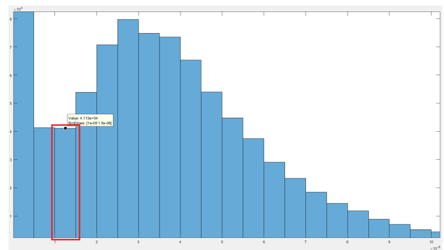
After extraction of errors vectors from the relevant area and labeling them as suspicious, we create corresponding binary mask. After this, we process it with a noise filter that removes regions from the binary mask that have a square less than some value. So after this, we keep only errors vectors which corresponding forgery regions in the set of suspicious errors vectors.

3 Analysis of experiments results

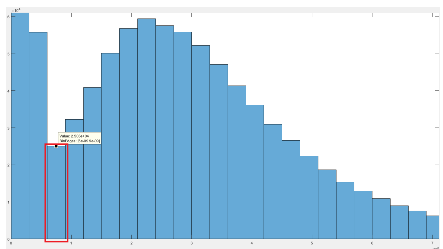
The experiments were carried out on a desktop PC with Intel Core i5-4460 processor and 16 GB RAM using MATLAB R2016b software.



a



b



c

Fig. 3. Selecting the thresholds according to the distribution histogram: a) for $\tilde{\varepsilon}_4^2$, b) for $\tilde{\varepsilon}_8^2$, c) for $\tilde{\varepsilon}_{16}^2$. So, $T_{\tilde{\varepsilon}_4^2} = 1.5 \times 10^{-8}$, $T_{\tilde{\varepsilon}_8^2} = 1.5 \times 10^{-8}$, $T_{\tilde{\varepsilon}_{16}^2} = 0.9 \times 10^{-8}$.

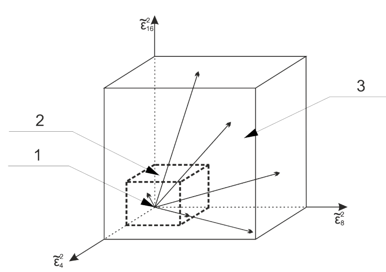


Fig. 4. Splitting the cube with set of vectors $\bar{x}(n_1, n_2)$ into three areas: 1 - the area of vectors corresponding to duplicates taken from another image of the time series; 2 - the area of vectors corresponding to authentic image fragments; 3 - the area of vectors corresponding to duplicates taken from the same image.

Five image time series were obtained using the same camera. The camera was still all the time. It has captured the scene and token image every 10 sec. As result of this procedure, we have got five image time series with six images in every series. Next, we transform all images to gray-scale. Obtained images have 920×1380 dimension. These time series were chosen as the objects of experiments. We developed a copy-move generation procedure which enables to add distortions and control their parameters.

The experiment results that were carried out on all image series with duplicate taken from another image of the image series are showed in the Table 1. Example of this detection is shown in Figure 5.

Table 1. Experiments result of inter-image copy-move detection

Image Series	Precision	Recall	F1 Value
1	1	0.84	0.91
2	1	0.95	0.97
3	0.58	0.74	0.65
4	0.8	0.82	0.8
5	0.87	0.84	0.85

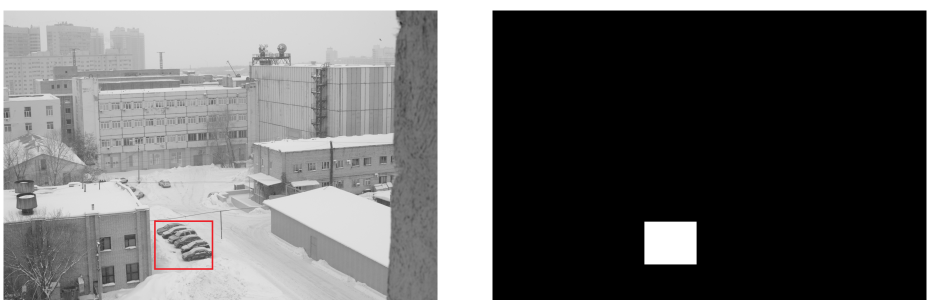


Fig. 5. Inter-image copy-move detection result

The experiment results with duplicate taken from the same image are shown in the Table 2 and in the Figure 6

Table 2. Experiments result of intra-image copy-move detection

Image Series	Precision	Recall	F1 Value
1	0.55	0.84	0.66
2	0.53	0.74	0.62
3	0.62	0.85	0.72
4	0.55	0.68	0.61
5	0.63	0.78	0.70

Fig. 6. Intra-image copy-move detection result

References

1. Chandola, V., Banerjee, A., Kumar, V.: Anomaly detection: A survey. ACM Computing Surveys (CSUR) 41(3), 15.1–15.58 (2009)
2. Christian, A., Sheth, R.: Digital Video Forgery Detection and Authentication Technique - A Review. International Journal of Scientific Research in Science and Technology(IJSRST), 2 (6), 138–143 (2016)
3. Denisova, A.Yu., Myasnikov, V.V.: Anomaly detection for hyperspectral imaginary. Computer Optics 38(2), 287–296 (2014)
4. Aggarwal, C.C.: Outlier Analysis. 2nd ed., Springer (2017)