Алгоритм обнаружения фальсификаций отдельного цифрового изображения в последовательности снимков на основе выявления аномальных изменений

Н.И. Евдокимова[1], В.В. Мясников[1,2]

[1]Самарский национальный исследовательский университет имени акад. С.П. Королева, Московское шоссе, 34А, Самара, Россия, 443086
[2]Институт систем обработки изображений РАН – филиал ФНИЦ «Кристаллография и фотоника» РАН, Молодогвардейская 151, Самара, Россия, 443001

Аннотация. Настоящая работа посвящена разработке и исследованию алгоритма обнаружения преднамеренных искажений – фальсификаций – отдельного цифрового изображения в серии (временной последовательности) снимков одной сцены. Предлагаемый алгоритм состоит из трех этапов. На первом этапе оценивается ряд ошибок реконструкции фрагментов анализируемого снимка по соответствующим фрагментам «соседних» снимков. Полученные по всему снимку ошибки анализируются для оценки на втором этапе их вероятностного распределения. На заключительном этапе в качестве «подозрительных» фрагментов анализируемого кадра отбираются те, которые представляют собой аномалии – то есть в статистическом смысле маловероятны. Предлагаемый алгоритм, в отличие от ряда существующих, позволяет унифицированным образом проводить обнаружение таких атак, как дублирование фрагментов изображения, врезки фрагментов с «соседних» снимков, врезки фрагментов с изображения, не включенного в рассматриваемую серию и т.п. Представлены результаты исследования предлагаемого алгоритма по эффективности обнаружения некоторых атак.

1. Введение

С расширением количества сфер, использующих цифровые изображения в своей работе, а также доступностью и популяризацией средств их обработки вместе с объемом фальсифицированных изображений возрастает также сложность методов искажений по отношению к их обнаружению. Временные серии изображений показывают динамику сцены и позволяют проводить ее сравнение с течением времени. Так, имея временную последовательность снимков некоторой сцены, с некоторым допущением можно смоделировать снимок, который будет в сцене следующим, или же снимок в серии, который, возможно, был искажен.

Предлагаемый алгоритм позволяет обнаруживать атаки пространственного типа. На настоящий момент времени разработано несколько методов обнаружения атак подобного типа: методы, основанные на уникальности артефактов, оставляемых камерой, методы, основанные на артефактах, сопровождающих кодирование, и методы, использующие временную и пространственную корреляцию [2]. Их основной слабостью является отсутствие устойчивости к различным видам искажений. Разработанный алгоритм использует корреляцию между соответствующими фрагментами разных изображений

в серии изображений и позволяет обнаруживать дублированные в пределах одного изображения фрагменты, врезанные с соседних изображений фрагменты, а также врезанные с изображения, не включенного в рассматриваемую серию, фрагменты.

Для разработки алгоритма обнаружения преднамеренных искажений понятие "искажение"понимается в смысле понятия "аномалия". В общем смысле аномалией называется фрагмент данных, который не соответствует точно заданному понятию нормального поведения [1]. В соответствии с приведенным определением в рамках данной работы аномалией считаются измененные области изображения, при этом предлагаемый алгоритм использует понятие аномалии в смысле наименее вероятных точек.

Работа состоит из двух частей, а именно описания предлагаемого алгоритма и анализа результатов проведенных экспериментов. Описание предлагаемого алгоритма подразделяется на выбор способа описания фрагментов изображения, характеристику метода формирования статистики и определение правила отнесения фрагментов анализируемого изображения к аномалиям.

2. Описание предлагаемого алгоритма

Пусть есть некоторая серия (временная последовательность) изображений одной сцены $I_t\,(n_1, n_2)$, $t = \overline{0, T}$, имеющих размер $N_1 \times N_2$, $n_i \in \overline{0, N_i - 1}$ $(i = \overline{1, 2};\ T \geq 1)$.

Для определенности считается, что проверяется изображение $I_0(n_1, n_2)$, хотя во временной последовательности оно может находиться в любом месте. В скользящем окне с позицией (n_1, n_2) рассматривается некоторая прямоугольная область изображения $D(n_1, n_2) \subseteq \overline{0, N_1 - 1} \times \overline{0, N_2 - 1}$. Конкретной области $D(n_1, n_2)$ соответствуют фрагменты $I_t(m_1, m_2)$, где $(m_1, m_2) \in D(n_1, n_2)$. Для упрощения изложения аргументы области в записи $D(n_1, n_2)$ ниже могут быть опущены. Экспериментальным способом было установлено, что наилучшее в смысле качества обнаружения и времени выполнения окно имеет размер 15×15 пикселей.

2.1. Описание фрагментов изображения

Для каждого возможного положения окна D в плоскости изображения соответствующие фрагменты $I_t(m_1, m_2)$ последовательно разбиваются на k фрагментов, $k = 2^p$, путем кластеризации по яркости (используется алгоритм кластеризации внутренних средних). Пример такого разбиения для $k = 2^2$ представлен на рисунке 1:

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

Рисунок 1. Разбиение фрагмента $I_t(m_1, m_2)$ (слева) на $k = 2^2$ фрагмента

Полученные таким образом для k-ого шага новые фрагменты изображения обозначаются как $I_t^j(n_1, n_2)$, $j = \overline{0, k-1}$. Далее для каждого возможного положения области D в плоскости изображения решается задача представления фрагмента изображения I_0 для этой области с помощью линейной комбинации соответствующих (по положению области D) фрагментов $I_1^0, I_1^1, ..., I_1^{k-1}; ...; I_T^0, I_T^1, ..., I_T^{k-1}$, т.е.:

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

с помощью минимизации СКО ε_k^2:

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

Затем находятся два вида ошибок: собственно СКО ε_k^2 (2) и нормированное СКО $\tilde{\varepsilon}_k^2$, характеризующее величину сопряженности и задаваемое формулой (3):

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

Данная процедура выполняется последовательно для $k = 4, 8, 16$. Результатом работы первого этапа для каждого положения (n_1, n_2) области анализа $D(n_1, n_2)$ являются наборы величин нормированного СКО представления фрагментов анализируемого кадра, которые, для удобства дальнейшего использования, будут обозначаться в виде вектора:

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

2.2. Метод формирования статистики

На следующем этапе работы алгоритма множество полученных векторов $\vec{x}(n_1, n_2)$ представляется в системе координат $\tilde{\varepsilon}_4^2\tilde{\varepsilon}_8^2\tilde{\varepsilon}_{16}^2$. Данное множество векторов $\vec{x}(n_1, n_2)$ располагается в трехмерном кубе со стороной, равной 1, как показано на рисунке 2.

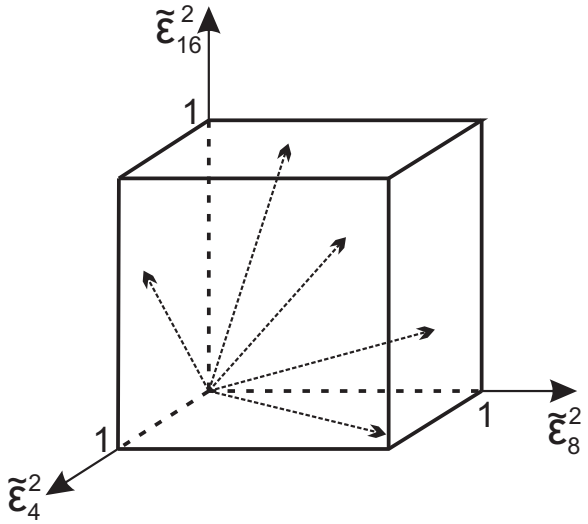


Рисунок 2. Множество векторов $\vec{x}(n_1, n_2)$ в системе координат $\tilde{\varepsilon}_4^2\tilde{\varepsilon}_8^2\tilde{\varepsilon}_{16}^2$

2.3. Нахождение аномалий

На изображениях, полученных в реальных условиях, нет абсолютно статичных (не изменяющихся с течением времени) объектов. Это связано как с реальными свойствами

камер, обладающих собственными шумами, так и с особенностями тракта передачи информации от камеры к системе обработки. В процессе передачи информации изображение перед отправкой подвергается компрессии, а затем декодируется. Данные манипуляции часто приводят к дополнительным искажающим особенностям системы. Кроме того, на сцене, которую захватывает камера, часто присутствуют объекты, обладающие определенными динамическими характеристиками, несмотря на то, что они статичны в общем смысле - деревья, колеблющиеся на ветру, отдаленные объекты, скрывающиеся за туманом и т.п.

В соответствии с этим фактом, невозможно получить вектор с координатами $(0;0;0)$ после представления фрагментов аутентичного изображения с помощью линейной комбинации соответствующих фрагментов соседних изображений. Сделанное утверждение позволяет определить правило отнесения фрагментов, которым соответствуют векторы с координатами $(0;0;0)$, к аномалиям. Данный тип аномалий соответствует фрагментам, врезанным с "соседних"снимков.

С другой стороны, ошибки $\tilde{\varepsilon}_4^2$, $\tilde{\varepsilon}_8^2$, $\tilde{\varepsilon}_{16}^2$ представления аутентичного фрагмента изображения должны иметь значения, не превышающие некоторых заданных пороговых значений. Соответственно, векторы ошибок, значения которых по хотя бы одной из координат превышают пороговое значение, являются аномалиями. Данный тип аномалий соответствует фрагментам, дублированным в пределах анализируемого изображения или взятым с изображения, не включенного в рассматриваемую серию. Очевидно, значения ошибок представления одного и того же фрагмента должны уменьшаться с увеличением количества кластеров, на которые разбивается фрагмент. Принимая во внимание данный факт, обоснованно использовать разные пороговые значения для $\tilde{\varepsilon}_4^2$, $\tilde{\varepsilon}_8^2$ и $\tilde{\varepsilon}_{16}^2$. Таким образом, должно выполняться соотношение (5):

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

После этапа формирования статистики анализируются гистограммы значений ошибок представления фрагментов и выбираются пороговые значения в соответствии с соотношением (5) как показано на рисунке 3. В качестве порогового значения выбирается первый локальный минимум. Так, для гистограмм распределения ошибок, представленных на рисунке 3, были выбраны следующие пороговые значения:

- для $\tilde{\varepsilon}_4^2$ - $T_{\tilde{\varepsilon}_4^2} = 1.5 \times 10^{-8}$;
- для $\tilde{\varepsilon}_8^2$ - $T_{\tilde{\varepsilon}_8^2} = 1.5 \times 10^{-8}$;
- для $\tilde{\varepsilon}_{16}^2$ - $T_{\tilde{\varepsilon}_{16}^2} = 0.9 \times 10^{-8}$.

Затем куб, содержащий множество всех возможных векторов ошибок $\vec{x}(n_1, n_2)$, разбивается на три области:

1) Центр системы координат;
2) Параллелепипед, примыкающий к центру системы координат;
3) Оставшаяся часть куба.

Согласно определенным выше типам аномалий, векторы, лежащие в первой области (векторы с координатами $(0;0;0)$), соответствуют врезанным с соседних изображений серии фрагментам. Векторы, лежащие во второй области, указывают на аутентичные фрагменты анализируемого изображения. Векторы, лежащие в третьей области, соответствуют фрагментам, дублированным в пределах одного изображения, или фрагментам, врезанным с изображения, не включенного в рассматриваемую серию. Данное разбиение представлено на рисунке 4.

После определения множества векторов, расположенных в релевантной области, они помечаются как подозрительные. На основе множества подозрительных векторов

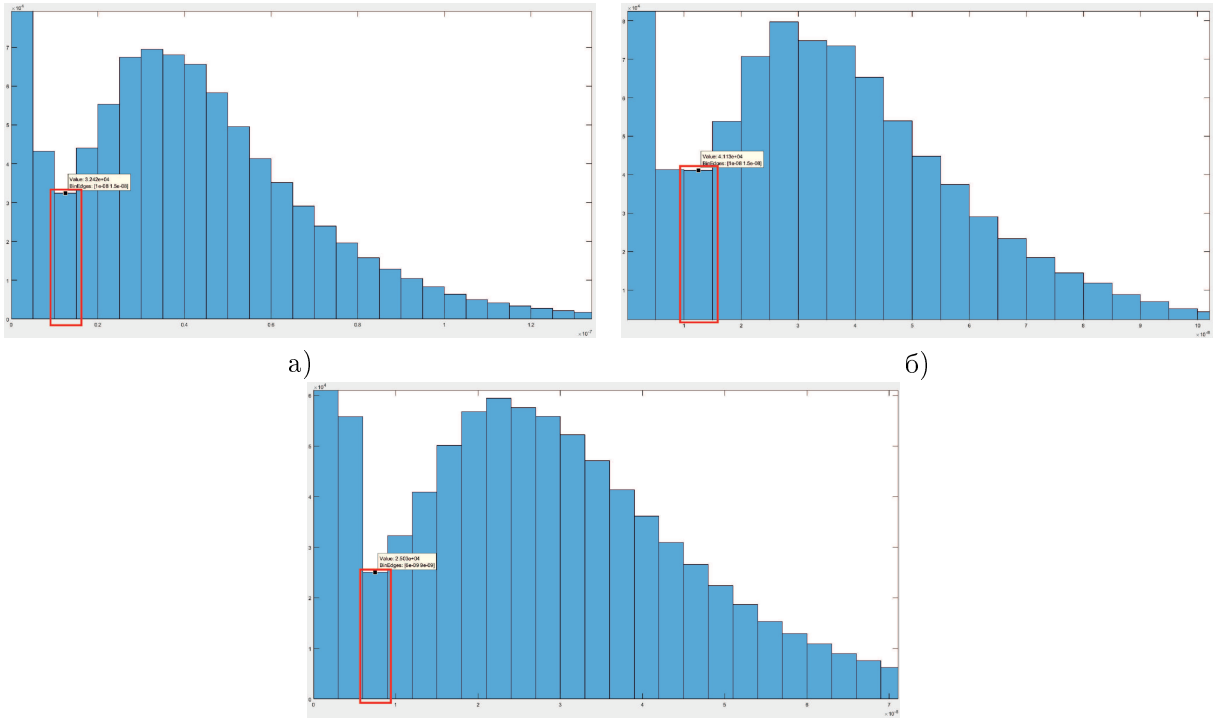


Рисунок 3. Выбор пороговых значений с помощью гистограммы распределения ошибок: а) для $\tilde{\varepsilon}^2_4$, б) для $\tilde{\varepsilon}^2_8$, в) для $\tilde{\varepsilon}^2_{16}$

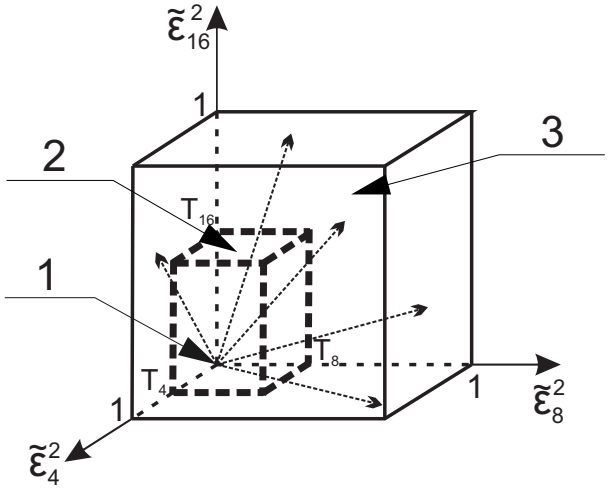


Рисунок 4. Разбиение множества возможных векторов $\vec{x}(n_1, n_2)$ на три области

создается соответствующая ему бинарная маска. На следующем шаге полученная маска обрабатывается шумоподавляющим фильтром, удаляющим регионы площадью менее некоторого заданного значения. Так, после процедуры шумоподавления из множества подозрительных векторов ошибок остаются только векторы ошибок, соответствующие

искаженным областям.

3. Анализ результатов экспериментов

Для проведения экспериментальных исследований используется стандартный ПК (Intel Core i5-4460 3.2 ГГц, 16 Гб ОЗУ) с использованием пакета прикладных программ MATLAB R2016b.

В качестве объектов исследований были выбраны пять серий изображений, по шесть изображений в каждой, полученных с использованием одной и той же камеры, делающей снимки каждые 10 секунд. Полученные изображения имели размер 920×1380. Для применения предложенного алгоритма все изображения серий были преобразованы в полутоновые.

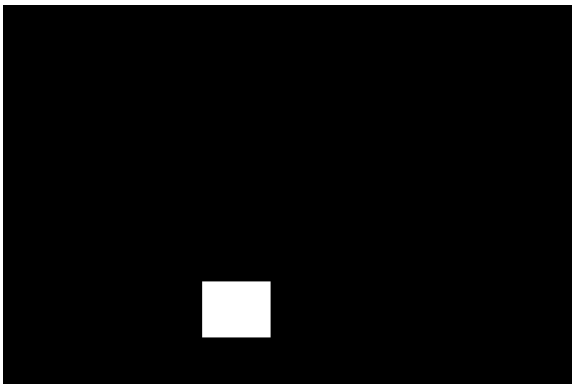
Для проведения экспериментов была разработана процедура генерации изображений, содержащих дублированные в пределах одного изображения фрагменты и содержащих фрагменты, врезанные с других изображений серии.

Результаты экспериментов, проведенных на сериях изображений, содержащих в себе одно изображение с врезанным из другого изображения серии фрагментом, приведены в таблице 1. Пример обнаружения данного вида искажений представлен на рисунке 5.

Таблица 1. Результаты обнаружения врезанного с соседнего изображения серии фрагмента

Номер серии	Precision	Recall	F1
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

а)



б)

Рисунок 5. Пример обнаружения врезанного с соседнего изображения серии фрагмента: а) исходное искаженное изображение, б) бинарная маска, полученная в результате применения алгоритма обнаружения

Результаты экспериментов, проведенных на сериях изображений, содержащих в себе одно изображение с дублированными фрагментами, приведены в таблице 2. Пример обнаружения данного вида искажений представлен на рисунке 6.

Таблица 2. Результаты обнаружения дублированного фрагмента

Номер серии	Precision	Recall	F1
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

а)

б)

Рисунок 6. Пример обнаружения дублированного фрагмента: а) исходное искаженное изображение, б) бинарная маска, полученная в результате применения алгоритма обнаружения.

4. Заключение

В работе предложен алгоритм обнаружения преднамеренных искажений отдельного цифрового изображения во временной серии снимков одной сцены. Проведенные исследования показали, что предложенный алгоритм позволяет с достаточно высокой в смысле метрики F1 точностью определять наличие фрагментов, врезанных с других изображений серии, и их расположение. Однако применение алгоритма в отношении обнаружения дублированных в границах одного изображения областей дает не высокие значения метрики F1 из-за достаточно большого количества ложных срабатываний (значение метрики Precision). В связи с этим предложенный алгоритм нуждается в некоторых улучшениях, которые будут представлены в последующих работах.

5. Благодарности

Работа выполнена при поддержке Федерального агентства научных организаций (соглашение № 007-ГЗ/Ч3363/26).

6. Литература

[1] Chandola, V. Anomaly detection: A survey / V. Chandola, A. Banerjee, V. Kumar // ACM Computing Surveys (CSUR). — 2009. — Vol. 41(3). — P. 15.1–15.58.

[2] Christian, A. Digital Video Forgery Detection and Authentication Technique - A Review / A. Christian, R. Sheth // International Journal of Scientific Research in Science and Technology(IJSRST). — 2016. — Vol. 2(6). — P. 138–143.

Detecting forgery of image time series based on the anomalies detection

N.I. Evdokimova[1], **V.V. Myasnikov**[1,2]

[1]Samara National Research University, Moskovskoye shosse, 34, Samara, Russia, 443086
[2]Image Processing Systems Institute of RAS - Branch of the FSRC "Crystallography and Photonics" RAS, Molodogvardejskaya street 151, Samara, Russia, 443001

Abstract. Increasing complexity of image forgery methods is an actual problem nowadays. This problem rises due to the expansion of fields that use digital images in their work. Image time series show the dynamics of the scene and allow it to be compared over time. This paper proposes a new algorithm for detecting forgeries of single digital image in an image time series described a scene. First part of paper provides description of proposed algorithm consisted of three stages. The aim of first stage is getting a set of errors that were computed during reconstruction of analyzed image using other images of series. Errors distribution histograms is constructed and estimated on the second stage. On the third stage, anomaly types are determined and decision rule for each anomaly type is set up. Finally, fragments of the analyzed image that are anomalies are selected as 'suspicious'. Second part of paper contains investigation results of intra-image copy-move and inter-image copy-move detection using the proposed algorithm.

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