

КОМП'ЮТЕРНІ Й ІНФОРМАЦІЙНІ МЕРЕЖІ І СИСТЕМИ АВТОМАТИЗАЦІЯ ВИРОБНИЦТВА COMPUTER AND INFORMATION NETWORKS AND SYSTEMS MANUFACTURING AUTOMATION

УДК 004.056.5

В.В. Зорило, канд. техн. наук,
Е.Ю. Лебедева, канд. техн. наук,
Одес. нац. политехн. ун-т

КОМПЛЕКСНЫЙ МЕТОД ВЫЯВЛЕНИЯ И ЛОКАЛИЗАЦИИ ОБЛАСТЕЙ КЛОНИРОВАНИЯ В ЦИФРОВЫХ ИЗОБРАЖЕНИЯХ

Введение. Сегодня обнаружение фальсификации цифрового изображения (ЦИ) — одна из важных проблем защиты информации ввиду общедоступности графических редакторов, таких как Adobe Photoshop, Gimp и т.д. Часто при фальсификации ЦИ возникает необходимость дублирования каких-либо объектов (добавить предметы, вставить дополнительную цифру в некий идентификационный номер и т.п.) или сокрытия различных деталей (родинки, шрамы, царапины и т.п.). Удобнее всего с точки зрения сохранения визуального восприятия использовать для этого фрагменты того изображения, которое фальсифицируется, причем замещающие области (ЗО) выбираются, как правило, вблизи обрабатываемого объекта, чтобы минимизировать отличия света/тени, яркости/контрастности, и могут быть очень незначительными по размеру (часто линейные размеры ЗО сравнимы с размерами блока 8×8 стандартного разбиения матрицы ЦИ, а в некоторых случаях и меньше).

При дублировании объекта, как и при его сокрытии, происходит копирование групп пикселей из одной части изображения в другую путем параллельного переноса. Данная операция, так называемое клонирование, реализована в большинстве современных графических редакторов. Обнаружение клонирования сводится к выявлению идентичных групп пикселей ЦИ. Если таковые найдутся, это будет свидетельствовать о наличии фальсификации ЦИ, осуществленной путем клонирования, что укажет на невозможность использования данного ЦИ как достоверного документа. В ряде случаев установления факта наличия клонированных блоков в ЦИ может быть недостаточно: важна точность локализации области клонирования. Однако за высокую точность приходится платить высокой вычислительной сложностью соответствующих алгоритмов, и часто при выборе метода решения задачи предпочтение отдается менее точным, но более быстрым методам. Все это оставляет *актуальным* разработку новых и усовершенствование существующих методов решения рассматриваемой задачи с учетом их временных затрат.

DOI 10.15276/opu.1.45.2015.17

© В.В. Зорило, Е.Ю. Лебедева, 2015

Целью данной работы является повышение эффективности обнаружения нарушений целостности цифровых изображений путем разработки комплексного метода выявления и локализации областей клонирования.

Под эффективностью обнаружения нарушений целостности в работе понимается эффективность выявления области клонирования, которая будет определяться двумя параметрами: эффективностью (ЭО) и точностью (ТО) обнаружения. Эффективность обнаружения будет оцениваться количеством ошибок первого (пропуск имеющего место клонирования), второго рода (определение клонированных областей в оригинальных ЦИ) и быстродействием, а точность – относительной величиной площади выявленной области к реальной площади клона, определяемой в процентах. Реальную площадь области клона определяется как разность между оригинальным и фальсифицированным изображениями. Величина площади измеряется в пикселях.

Для достижения поставленной цели необходимо решить следующие *задачи*:

1. Выбрать эффективный с точки зрения ЭО метод обнаружения клонированных областей, имеющий незначительную вычислительную сложность.
2. Выбрать принципиальное направление повышения точности обнаружения области клонирования.
3. Разработать комплексный метод выявления, локализации областей клонирования.

Анализ последних исследований и публикаций. В настоящий момент при решении задач выявления нарушений целостности цифровых сигналов, в частности, цифровых изображений, все большей популярностью пользуется общий подход к анализу состояния и технологии функционирования информационной системы [1], в основе которого лежат следующие положения. Любое ЦИ, формализуемое в виде конечного набора двумерных матриц, однозначно определяется сингулярным спектром (спектрами) и набором (наборами) ортонормированных лексикографически положительных сингулярных векторов соответствующей ему матрицы (матриц), т.е. сингулярные числа (СНЧ) и сингулярные вектора несут в себе всю информацию о состоянии ЦИ и о любых воздействиях на него. Фальсификация, представляющая собой преобразование ЦИ, формально представима в виде возмущения (а значит в виде совокупности возмущений СНЧ и сингулярных векторов) исходной матрицы (множества матриц) ЦИ. Возмущения СНЧ сравнимы с величиной возмущающего воздействия, характеризуют его силу, чего нельзя сказать о сингулярных векторах. Таким образом, в качестве набора формальных параметров, характеризующих ЦИ, анализ которых целесообразно использовать для выявления нарушения его целостности, далее выступает набор СНЧ.

В [2] одним из авторов был разработан метод, основанный на общем подходе к анализу состояния и технологии функционирования информационной системы, способный обнаружить клонированную область, размеры которой сравнимы с размерами блока стандартного разбиения матрицы ЦИ. Метод хорошо себя зарекомендовал с точки зрения эффективности обнаружения клонирования среди существующих современных аналогов [3]. Основная идея метода заключается в следующем: поскольку клонирование представляет собой копирование пикселей, абсолютно естественно для повторяющихся блоков иметь одинаковые СНЧ при условии, что ЦИ после фальсификации сохранено в формате без потерь. Поэтому для поиска ЗО имеет смысл сравнивать СНЧ блоков и выделять, как клонированные, те блоки, СНЧ которых равны [2]. Один из основных шагов данного метода — разбиение матрицы ЦИ на квадратные блоки 8×8 , пересекающиеся так, чтобы любые два соседних блока отличались друг от друга на один столбец (строку). Далее для каждого блока определяется его сингулярный спектр. Всего для блока указанной размерности сингулярных чисел будет 8. Это на порядок меньше, чем количество коэффициентов яркости пикселей или коэффициентов ДКП, на анализе которых основаны аналоги [4], за счет чего поиск клонирования в [2, 3] и происходит быстрее. Однако сингулярное разложение является инструментом для анализа матрицы, что жестко связывает предложенный в [2, 3] метод с прямоугольными блоками разбиения. Прямоугольные блоки, с учетом произвольности формы клонированной области, принципиально не позволяют точно определить границу ЗО в общем случае. Поэтому для повышения ТО необходим более гибкий математический инструмент.

Существует несколько направлений повышения точности обнаружения области клонирования. Подробно они рассмотрены в [5]. Авторы статьи считают, что наиболее перспективным является направление с использованием блоков нестандартной формы, например треугольных, круглых, блоков сложной формы и т.п.

Использование таких блоков в процессе обнаружения областей клонирования возможно, если в качестве инструмента, определяющего клонированные блоки, используется метрика, например коэффициент корреляции [6] между ними. Использование такого подхода к выявлению областей клонирования требует значительных вычислительных затрат. Данный подход был бы более эффективным с точки зрения быстродействия, если бы удалось заранее ограничить область поиска клонирования, то есть проверять лишь подозреваемую часть изображения, а не все изображение в целом.

Таким образом, разрабатываемый комплексный метод выявления, локализации областей клонирования должен состоять из двух основных этапов: первый — предварительный “грубый” поиск областей клонирования “быстрым” методом (в основу положен метод, предложенный в [2, 3]); второй — повышение точности обнаружения с использованием блоков нестандартной формы, возможно уступающий первому по быстродействию.

Комплексный метод выявления, локализации областей клонирования. Рассмотрим основные шаги комплексного метода выявления, локализации областей клонирования:

Шаг 1. Поиск областей клонирования

1.1. Разбить матрицу F ЦИ на пересекающиеся 8×8 -блоки F_{ij} , $i = 1, 2, \dots, (n-7)$, $j = 1, 2, \dots, (m-7)$ так, чтобы каждый блок отличался от соседнего на один столбец (строку).

1.2. Построить матрицу S с элементами $s_{ij} = \sum_{k=1}^4 \sigma_k$, $i = 1, 2, \dots, (n-7)$, $j = 1, 2, \dots, (m-7)$,

где $\sigma_1, \sigma_2, \sigma_3, \sigma_4$ — наибольшие СНЧ соответствующего блока F_{ij} , $i = 1, 2, \dots, (n-7)$, $j = 1, 2, \dots, (m-7)$.

1.3. Построить матрицу клонирования (МК) C с элементами c_{ij} , $i = 1, 2, \dots, (n-7)$, $j = 1, 2, \dots, (m-7)$; c_{ij} соответствует блоку F_{ij} ЦИ. Для определения c_{ij} сравнить s_{ij} попарно со всеми элементами матрицы S :

Если для s_{ij} найдется элемент s_{kl} , $k \neq i, l \neq j$, матрицы S , что $s_{ij} = s_{kl}$

то $c_{ij} = 1$,

иначе $c_{ij} = 0$.

1.4. Элементы $c_{ij} = 1$ матрицы C соответствуют клонированным блокам ЦИ.

1.5. Формирование матрицы нового изображения $F' = \bigcup c_{ij}$, содержащего найденные клонированные блоки.

Шаг 2. Уточнение области клонирования

2.1. Получить координаты (x_1, y_1) и (x_2, y_2) прямоугольной области с найденными областями клонирования из матрицы изображения F' .

2.2. Сформировать матрицу яркости Y' полученной прямоугольной области по координатами (x_1, y_1) и (x_2, y_2) .

2.3. Разбить матрицу яркости Y' на пересекающиеся блоки $D = \{d_1, d_2, \dots, d_t\}$,

$\bigcup_{j=1}^t d_j = Y'$ размерами $p \times p$.

2.4. Для рассматриваемой пары блоков d_i, d_j , $i=1, \dots, t$, $j=i+1, \dots, t$ получить нестандартные разбиения d_i^r и d_j^r соответственно, где r – номер разбиения. Для каждого разбиения:

2.4.1. Вычислить меру близости $\delta = \text{Correlation}(d_i^r, d_j^r)$.

2.4.2. Найти пары блоков d_i^r и d_j^r , подозрительные на принадлежность областям клона и его прообразу, путем анализа значения меры близости δ .

Результаты. Иллюстративный пример работы разработанного комплексного метода выявления, локализации областей клонирования представлен на рис. 1. В качестве подозрительной области в этом примере была выбрана область, содержащая блоки белого цвета на соответствующей матрице (рис. 1, *з*). На втором шаге комплексного метода использовались круглые блоки. Из результатов видно, что при использовании на втором шаге круглых блоков относительная величина выявленной области относительно реальной области клона увеличилась в 1,9 раз в сравнении с использованием стандартного разбиения.

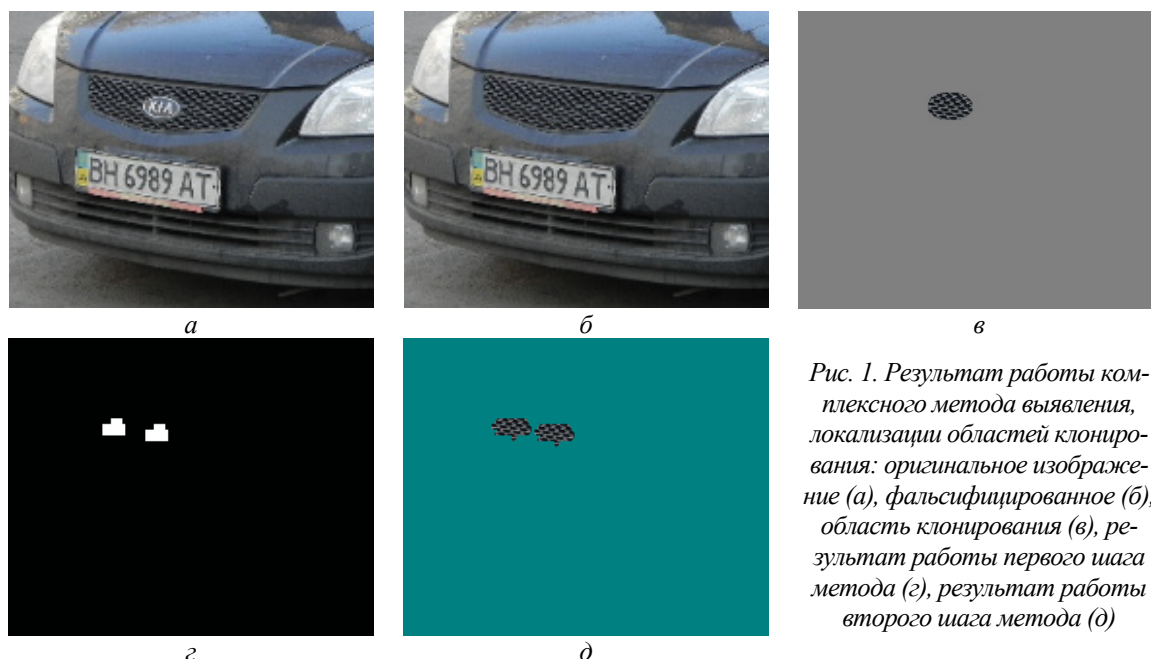


Рис. 1. Результат работы комплексного метода выявления, локализации областей клонирования: оригинальное изображение (а), фальсифицированное (б), область клонирования (в), результат работы первого шага метода (з), результат работы второго шага метода (д)

Выводы. Разработан комплексный метод выявления, локализации областей клонирования как фальсификации цифрового изображения, объединяющий в себе преимущества двух наиболее эффективных современных методов. На первом этапе комплексного метода осуществляется предварительный “грубый” поиск клонированных областей изображения. Затем на ограниченной после шага 1 комплексного метода области ЦИ выполняется более точное выделение границ замещающей области с использованием блоков нестандартной формы. Таким образом, достигается более высокая точность локализации областей клонирования по сравнению с методом, основанным на общем подходе к анализу состояния и технологии функционирования информационной системы (идентифицируемая клонированная область увеличилась в среднем на 47 %), и уменьшается суммарное время, затрачиваемое на обработку ЦИ по сравнению с методом, основанным на анализе коэффициентов корреляции между двумя блоками нестандартной формы (уменьшение времени работы комплексного метода на втором этапе зависит от того, насколько уменьшилась область обработки ЦИ после предварительного поиска клонирования, а также от размера самой клонированной области – на тестируемых изображениях понадобилось в среднем в 3 раза меньше времени, чем при проверке всего ЦИ). В результате повышается эффективность выявления и локализации областей клонирования в цифровом изображении. С учетом

того, что видео последовательность может рассматриваться как совокупность кадров, предложенный метод может быть адаптирован для цифрового видео.

Литература

1. Кобозева, А.А. Общий подход к анализу состояния информационных объектов, основанный на теории возмущений / А.А. Кобозева // Вісн. СНУ ім. В. Даля. — 2008. — № 8(126), Ч. 1. — С. 72 — 81.
2. Зорило, В.В. Выявление клонирования как фальсификации цифрового изображения / В.В. Зорило // Вісн. НТУ “ХПІ”. Серія: Системний аналіз, управління та інформаційні технології. — 2011. — № 35. — С. 31 — 38.
3. Зорило, В.В. Методы повышения эффективности выявления нарушения целостности цифрового изображения / В.В. Зорило // Інформаційна безпека. — 2013. — № 2(10). — С. 34 — 41.
4. Fridrich, J. Detection of copy-move forgery in digital images [Электронный ресурс] / J. Fridrich, D. Soukal, J. Lukáš // Proceedings of the Digital Forensic Research Workshop. — Cleveland, OH, 2003. — Режим доступа: <http://www.ws.binghamton.edu/fridrich/research/copymove.pdf> (Дата обращения: 10.03.2015).
5. Lebedeva, Ye.Yu. Improvement of the accuracy of detection of cloned regions in digital images / Ye.Yu. Lebedeva // Інформатика та математичні методи в моделюванні. — 2014. — Т. 4, № 3. — С. 278 — 284.
6. Лебедева, Е.Ю. Исследование метрик используемых при обнаружении клонированных участков изображений в задачах выявления фальсификации / Е.Ю. Лебедева, Ю.Ф. Лебедев // Вісн. НТУ “ХПІ”. Серія: Системний аналіз, управління та інформаційні технології. — 2011. — № 35. — С. 25 — 31.

References

1. Kobozeva, A.A. (2008). The general approach to the analysis of the state of the information objects based on perturbation theory. *Bulletin of Volodymyr Dahl East Ukrainian National University*, 8(1), 72-81.
2. Zorilo, V.V. (2011). Detection of cloning as digital image falsification technique. *Herald of the National Technical University “Kharkiv Polytechnic Institute”: System Analysis, Management and Information Technologies*, 35, 31-38.
3. Zorilo, V.V. (2013). Methods of improving efficiency of digital image integrity violation detection. *Informative Safety*, 2, 34-41.
4. Fridrich, J., Soukal, D. and Lukáš, J. (2003). Detection of copy-move forgery in digital images. In *Proceedings of the 2003 Digital Forensics Research Workshop (DFRWS)* (pp. 55-61). Academic Press.
5. Lebedeva, Ye.Yu. (2014). Improvement of the accuracy of detection of cloned regions in digital images. *Informatics and Mathematical Methods in Simulation*, 4(3), 278-284.
6. Lebedeva, Ye.Yu. and Lebedev, Yu.F. (2011). Study of metrics used in the detection of cloned areas during identify of images falsification. *Herald of the National Technical University “Kharkiv Polytechnic Institute”: System Analysis, Management and Information Technologies*, 35, 25-31.

АНОТАЦІЯ / АННОТАЦИЯ / ABSTRACT

В.В. Зорило, О.Ю. Лебедева. **Комплексний метод виявлення і локалізації областей клонування у цифрових зображеннях.** Важливою проблемою комплексної системи захисту інформації є розробка методів виявлення порушень цілісності цифрових зображень. Одним із найчастіших способів порушення цілісності при несанкціонованих змінах є клонування, яке реалізоване в більшості графічних редакторах, що робить актуальним розв’язання задачі його виявлення. Крім виявлення самого факту клонування також важливим є точне визначення контурів клонованої області в зображенні. У роботі запропоновано поліноміальний комплексний метод виявлення областей клонування як фальсифікації цифрового зображення, що дозволяє локалізувати клоновані ділянки з точністю, яка перевершує сучасні аналоги. Комплексний метод базується на двох ефективних сучасних методах виявлення клонування, заснованих на загальному підході до аналізу стану і технології функціонування інформаційної системи, і включає в себе два етапи: “грубе швидке” виявлення клону; “повільне” уточнення контуру клону з використанням обмеженої виділеної на першому етапі підобласті зображення, що містить клон. Запропонований метод може бути адаптовано для цифрового відео.

Ключові слова: фальсифікація зображення, клонування, цифрове зображення.

В.В. Зорило, Е.Ю. Лебедева. Комплексный метод выявления и локализации областей клонирования в цифровых изображениях. Важной проблемой комплексной системы защиты информации является разработка методов выявления нарушений целостности цифровых изображений. Одним из наиболее частых способов нарушения целостности при несанкционированных изменениях является клонирование, реализованное в большинстве графических редакторах, что делает актуальным решение задачи его обнаружения. Помимо обнаружения самого факта клонирования также важным является точное определение контуров клонированной области в изображении. В работе предложен полиномиальный комплексный метод выявления областей клонирования как фальсификации цифрового изображения, позволяющий локализовать клонированные участки с точностью, превосходящей современные аналоги. Комплексный метод базируется на двух эффективных современных методах обнаружения клонирования, основанных на общем подходе к анализу состояния и технологии функционирования информационной системы, и включает в себя два этапа: “грубое быстрое” выявление клона; “медленное” уточнение контура клона с использованием ограниченной выделенной на первом этапе подобласти изображения, содержащей клон. Предложенный метод может быть адаптирован для цифрового видео.

Ключевые слова: фальсификация изображения, клонирование, цифровое изображение.

V.V. Zorilo, Ye.Yu. Lebedeva. Complex method of detection and localization of cloning areas in digital images. Developing methods aiming at identification of digital images integrity violation is an important problem of complex system of information security. Cloning method, realized in the most of graphic editors, is often used to compromise the integrity of the image related to unauthorized changes. Therefore, the task of cloning identification is a relevant one. In addition to the cloning identification it is also important to clarify the contours of the cloned area. In this paper we propose a polynomial comprehensive method to identify areas of cloning as a falsification of the digital image. It allows to locate the cloned areas with a higher precision compared to modern analogues. Complex method is based on two modern clone detection techniques, using a common approach to the analysis of the state and technology of the information system functioning. It includes two stages: “fast rough” cloning area identification and a “slow” identification of the contour of the cloning area using a limited image subarea containing a clone. The method proposed can be adapted for the digital video.

Keywords: image falsification, cloning, digital image.

Поступила в редакцию 11 марта 2015 г.