

В. В. ЗОРИЛО, аспирант ОНПУ, г. Одесса

ВЫЯВЛЕНИЕ КЛОНИРОВАНИЯ КАК ФАЛЬСИФИКАЦИИ ЦИФРОВОГО ИЗОБРАЖЕНИЯ

Розроблено та програмно реалізовано метод, який дозволяє виявити та локалізувати фальсифікацію цифрового зображення, проведено за допомогою інструменту «Штамп», реалізованого у більшості графічних редакторів. Розроблений метод значно перевершує аналогів з швидкості.

Разработаны программно реализован метод, позволяющий выявить и локализовать фальсификацию цифрового изображения, проведенную с помощью инструмента «Штамп», реализованного в большинстве графических редакторов. Разработанный метод значительно превосходит аналоги по быстродействию.

The detection digital image falsification method is developed and implemented. The falsification made by using Adobe Photoshop instrument "Clone". These methods calculate complicity is much less than this one parameter in other similar methods.

Введение. В век глобальной компьютеризации и интенсивного развития информационных технологий особенно остро ставится вопрос защиты информационных систем. Вместе с информационными технологиями развивается и компьютерная преступность, используя для своих противозаконных действий все более изощренные методы. В частности, такой вид компьютерной преступности, как подделка информации, может преследовать различные цели. Итогом подделки является то, что потребителю информации будут предоставлены недостоверные данные. Примером могут служить подтасовка результатов выборов или же хищение различного вида товаров путем ввода в программу фальшивых данных; подделка, изготовление или сбыт поддельных документов, штампов, печатей и бланков; изготовление или сбыт поддельных кредитных либо расчетных карт и иных платежных документов.

Ежегодно на борьбу с киберпреступностью развитые страны мира выделяют миллионы долларов. Новые компьютерные атаки требуют постоянного развития методов защиты информации. Практически во всех областях нашей жизни (медицина, охранные системы, системы безопасности, судебные разбирательства) используются цифровые аудио-, видео-, изображения. Их подделка доступна всем компьютерным пользователям. Чаще всего фальсификации подвергаются цифровые изображения (ЦИ). Ввиду этого особенно *актуальным* на сегодняшний день является умение отличить подделку от подлинного сигнала. В [1] была разработана общая методология анализа свойств, состояния и технологии функционирования произвольной информационной системы, которая успешно адаптируется для решения вопросов, связанных с идентификацией фальсификаций цифровых сигналов [2, 3].

В данной работе продолжается исследование возможностей обнаружения несанкционированного изменения цифрового сигнала, в частности, фотомонтажа ЦИ, и разработка методов решения этой актуальной проблемы.

Постановка проблемы и цель исследования. Из практики известно, что наиболее часто при подделке фотографий необходимо убрать какой-то предмет, либо, наоборот, продублировать его. В Adobe Photoshop для решения этой задачи чаще всего, если не всегда, применяют инструмент «Штамп». Данный инструмент используется для переноса клона объекта из одной части изображения в другую путем параллельного переноса, как правило, в пределах одной и той же фотографии. Как нам известно из открытой печати, наиболее успешным методом обнаружения такой фальсификации является метод, основанный на корреляции коэффициентов дискретного косинусного преобразования матрицы изображения [4], однако данный метод в реальных условиях требует значительных вычислительных затрат. Как показано в [2], любое вмешательство в ЦИ отразится на сингулярных числах (СНЧ) его матрицы и приведет к некоторым их особенностям. Выявление таких особенностей даст возможность отличить подлинное изображение от подвергнутого обработке штампом, что, в свою очередь, в случае наличия фальсификации позволит не использовать исследуемое ЦИ, например, в качестве доказательства чего-либо в суде. *Целью* данной работы является разработка метода выявления фальсификации цифрового изображения, выполненной посредством инструмента «Штамп» в Adobe Photoshop. Для достижения поставленной цели необходимо решить следующие задачи:

- выделить и обосновать целесообразность математических параметров, определяющих ЦИ, являющихся объектом сравнения при поиске клонированных участков;
- выявить характерные особенности математических параметров, определяющих изображение, при различных форматах хранения: с потерями и без потерь;
- разработать практический метод отделения ЦИ, подвергнувшегося обработке штампом, от изображения, которое не подвергалось данной операции; реализовать его программно;
- апробировать разработанный метод в реальных условиях.

Очень часто при фотомонтаже появляется необходимость дублирования каких-либо объектов (добавить людей в толпе, вставить дополнительное окно в стене дома, изменить номерной знак и т.п.) либо сокрытия различных деталей (убрать человека, дерево, самолет и т.п.). Удобнее всего «перекрывать» убираемый объект небольшими фрагментами, которые целесообразно брать из того же ЦИ вблизи обрабатываемого объекта, чтобы минимизировать отличия света/тени, яркости/контрастности. Дублирование объекта из тех же соображений целесообразно проводить, используя одну и ту же фотографию. В обоих случаях происходит копирование (параллельный перенос) групп пикселей из одной части ЦИ в другую. Задача заключается в том, чтобы об-

наружить группы повторяющихся пикселей ЦИ. Если таковые найдутся, это будет свидетельствовать о применении «Штампа» в данной области.

Выделение формальных параметров, определяющих ЦИ. В [1] было показано, что в качестве формального набора параметров, которые однозначно определяют свойства, состояние, изменение состояния любой информационной системы, в частности, ЦИ, можно использовать множество СНЧ и СНВ соответствующих матриц. Поскольку СНВ при даже очень малом возмущающем воздействии часто ведут себя принципиально непредсказуемо, для решения рассматриваемых в работе задач в качестве искомого набора формальных параметров будем использовать только СНЧ. Такой выбор, кроме всего прочего, даст возможность, как будет показано ниже, значительно уменьшить вычислительную сложность разработанного алгоритма выявления клонированных областей ЦИ по сравнению с существующими аналогами. Исследуем свойства СНЧ блоков матриц ЦИ, подвергнутых обработке штампом. Для этого создадим ситуацию, когда в изображении 16×16 пикселей встречаются идентичные блоки: передвинем верхний левый блок 8×8 пикселей вправо. Для каждого блока найдем множество СНЧ. Результат приведен в таблице. Проведенный эксперимент показал, что повторяющиеся блоки имеют абсолютно одинаковые СНЧ. Это вполне естественно, ведь штамп – не что иное, как клонирование пикселей. Как нам известно, основную зрительную информацию несут не контуры ЦИ, а фоновые части фотографии, или низкие и, возможно, средние составляющие частоты сигнала изображения. В [1] было установлено, что этим частотам соответствуют наибольшие и средние СНЧ. Таким образом, в целях уменьшения вычислительных затрат достаточно сравнивать между собой по 4 наибольших СНЧ каждого блока.

Мы получим желаемый результат, и это потребует на порядок меньше вычислительных затрат, чем сравнение блоков матрицы яркости пикселей или сравнение коэффициентов дискретного косинусного преобразования блоков матрицы ЦИ в [4]. Более того, если СНЧ блока сложить между собой, то в одинаковых блоках эти суммы будут равны. Вероятность равенства сумм СНЧ отличных друг от друга блоков очень мала, а значит, поиск штампа можно свести к сравнению сумм четырех СНЧ блоков матрицы изображения, что позволит еще сократить вычислительные затраты. В данном эксперименте блок был продублирован строго в соседний блок, то есть наше разбиение на блоки 8×8 не нарушило линейной зависимости между пикселями при нахождении СНЧ. Однако в реальных условиях такая ситуация маловероятна. Следовательно, необходимо выбрать такое разбиение на блоки, которое дало бы возможность адекватно провести исследование на наличие штампа.

СНЧ соответствующих блоков 8×8 матрицы исследуемого изображения

		Координаты блока			
		(1,1)	(1,2)	(2,1)	(2,2)
№ СНЧ	1	830,330	830,330	821,814	826,171
	2	15,692	15,692	23,018	99,608
	3	8,171	8,171	9,394	54,277
	4	3,932	3,932	6,143	16,375
	5	1,841	1,841	4,612	11,160
	6	1,499	1,499	3,382	7,317
	7	0,657	0,657	0,126	2,480
	8	0,373	0,373	0,028	0,167

Пусть имеется изображение, подвергшееся обработке инструментом «Штамп» в Adobe Photoshop, хранящееся в формате без потерь, для определенности – формат TIF. После фальсификации изображение также сохранено в формате без потерь. Будем анализировать одну из трех матриц изображения. Для поиска идентичных областей матрицы разобьем ее на пересекающиеся блоки 8×8 следующим образом: каждый блок отличается от предыдущего на один столбец (двигаемся вправо со сдвигом на 1 пиксель) до $(n-7)$ -го столбца матрицы изображения; дойдя до конца строки сдвигаемся на одну строку вниз и т.д. до $(m-7)$ -й строки. Тогда, имея изображение $m \times n$ необходимо сравнить между собой $(m-7) \times (n-7)$ блоков. Матрице изображения поставим в соответствие так называемую матрицу клонирования (МК), каждый элемент которой соответствует блоку 8×8 и принимает значения с учетом суммы четырех наибольших СНЧ: 0, если данный блок не имеет дубликатов (клонов), а 1 – иначе.

Вычислительный эксперимент. Рассмотрим ЦИ (рис. 1, а). С помощью штампа уберем ветку в правой части ЦИ.

Сохраним фотомонтаж без потерь (рис. 1, б). Построим для фальсифицированного изображения МК. Блокам, не имеющим клонов, соответствуют нули матрицы. Единицы возникают в той области, где ранее находилась ветка, а также там, откуда были взяты фрагменты для ее «перекрытия». Визуально небо на фотографии не вызывает никаких подозрений, но на матрице четко выделены области клонирования.

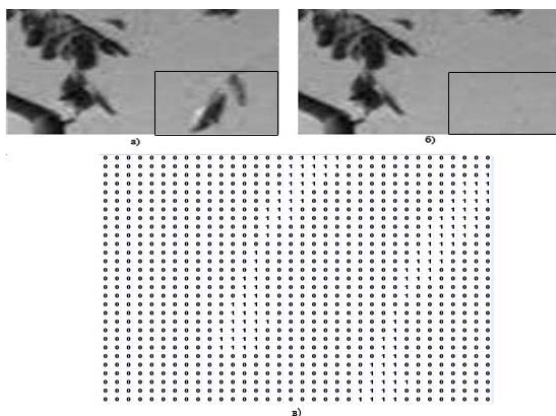


Рис. 1 – Отражение «штампа» на матрице клонирования: *a* – исходное ЦИ; *б* – фальсифицированное ЦИ; *в* – МК фальсифицированной части ЦИ

Эксперимент проведен удачно. Однако могут возникнуть неудобства в случае больших размеров областей, обработанных штампом, МК трудно сопоставить с реальными размерами фотографии. В этом случае представим МК в виде бинарного изображения, где черный цвет соответствует нулю МК, белый – единице. Для иллюстрации рассмотрим изображение, обработанное штампом, хранившееся без потерь до и после проведения фальсификации (рис. 2). Уберем птицу с помощью штампа. Визуально изображение на рис. 2, *б* не вызывает подозрений. Однако анализ показывает, что в области изображения, где ранее находилась птица, явно выделены белые участки изображения, соответствующие штампу (рис. 2, *г*).

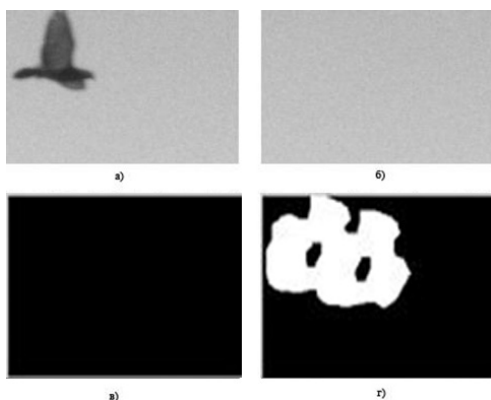


Рис. 2 – Визуализация фальсификации изображения большого размера: *a* – исходное ЦИ; *б* – фальсифицированное ЦИ; *в* – бинарное представление МК исходного ЦИ; *г* – бинарное представление МК фальсифицированного ЦИ

Большинство современных ЦИ хранится в формате jpeg (с потерями). Проверим работу метода для изображения, которое до фальсификации хранилось в формате jpeg, а после фальсификации было сохранено в формате без потерь (рис. 3). На наличие фальсификации подозревается номер автомобиля. В большинстве случаев области, подозрительные на фальсификацию, известны. Тогда нет необходимости исследовать изображение целиком. Достаточно вырезать только номерной знак и исследовать его. Это поможет сократить вычислительные затраты. Вычислительный эксперимент показал, что если такую фотографию обработать штампом и сохранить без потерь, то данный метод позволит выявить и локализовать область клонирования (рис. 3, *е*). В результате анализа видно, что белые области бинарного изображения соответствуют первому и третьему нулю номера авто. Фальсификация выделена успешно. Таким образом, в нашей ситуации совершенно не имеет значения формат ЦИ до обработки штампом.

На базе проведенного анализа разработан метод выявления фальсификации ЦИ, проведенной посредством инструмента «Штамп» Adobe Photoshop.

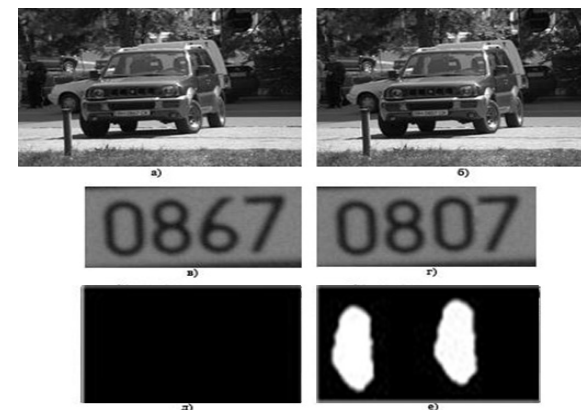


Рис. 3 – Поиск фальсификации: *a* – исходное ЦИ; *б* – фальсифицированное ЦИ; *в* – исследуемый номер исходного ЦИ; *г* – исследуемый номер фальсифицированного ЦИ; *д* – бинарное представление МК исходного ЦИ; *е* – бинарное представление МК фальсифицированного ЦИ

Основные шаги данного метода следует выполнять в таком порядке:

- Разбить матрицу изображения на пересекающиеся блоки 8×8 предложенным выше способом.
- Для каждого блока найти множество сингулярных чисел.
- Сложить наибольшие 4 СНЧ между собой в каждом блоке.
- Взять первое значение суммы $S(1 \ 1)$.

- Сравнить со всеми остальными значениями. Если $S(1\ 1) = S(ij)$, тогда значениям МК, соответствующим номерам блоков, содержащих $S(1\ 1)$ и $S(ij)$, присвоить значение «1». Если совпадений не найдено, тогда значению МК, соответствующему номеру блока, содержащего $S(1\ 1)$ присвоить «0».
- Взять значение суммы $S(1\ 2)$. Повторить для него предыдущий пункт. Затем следующее значение суммы, до тех пор, пока все значения сумм не будут проверены.
- Построить МК.
- Представить МК в виде бинарного изображения.
- Белые области бинарного изображения соответствуют клонированным частям исследуемого ЦИ.

Апробация разработанного метода. В целях апробации данного метода в реальных условиях проверим фотографию обложки известной книги (рис. 4, а), подозрительную на фальсификацию, проведенную вышеописанным способом. Проверка программой, реализующей наш метод, показывает, что часть тестируемого ЦИ под фамилией автора на рисунке соответствует белым областям бинарного представления МК (рис. 4, б).

Это говорит о том, что это ЦИ вероятнее всего могло быть подвергнуто обработке посредством инструмента «Штамп». Более того, как мы знаем и можем в этом убедиться, имея данную книгу – в действительности у книги 2 автора, что не оставляет сомнений в проведении фальсификации. Апробация разрабо-

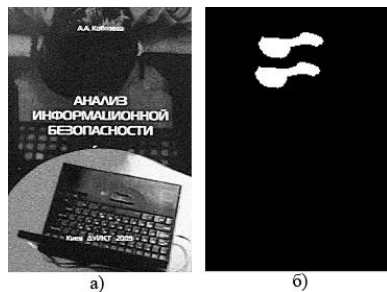


Рис. 4 – Поиск фальсификации: а – тестируемое ЦИ; б – бинарное представление МК тестируемого ЦИ

танного метода показала, что данный метод точно определяет и локализует как область, обработанную штампом, так и область, из которой были взяты фрагменты для перекрытия фальсифицируемого объекта.

Выводы. Итогом данной работы явился практический метод выявления фальсификации ЦИ, проведенной наиболее распространенным способом – с помощью инструмента «Штамп» программы Adobe Photoshop. Данный метод позволяет точно определить и локализовать области, которые обработаны штампом, а также области, из которых были взяты фрагменты для перекрытия фальсифицируемого объекта. Если изображение, подвергнутое обработке штампом, сохранено без потерь, то сравнение сингулярных чисел блоков матрицы изображения дает точный результат и требует меньше вычислитель-

ных затрат, чем предложенные в открытой печати методы решения данной проблемы. Все поставленные задачи выполнены. Цель достигнута.

Список литературы: 1. Кобозева А. А. Анализ информационной безопасности / А. А. Кобозева, В. А. Хорошко. – К. : Изд. ГУИКТ, 2009. – 251 с. 2. Кобозева А. А. Основы общего подхода к решению проблемы обнаружения фальсификации цифрового сигнала / А. А. Кобозева // Електромашинобудування та електрообладнання. – 2009. – Вип. 72. – С. 35–41. 3. Кобозева А. А. Метод виявлення фальсифікації цифрового зображення в умовах збурних дій / А. А. Кобозева, В. В. Зоріло // Збірник наукових праць Військового інституту Київського національного університету ім. Т. Шевченка. – 2009. – Вип. № 20. – С. 147–154. 4. J. Fridrich, D. Soukal, J. Lukas / Detection of copy-move forgery in digital images // Digital Forensic Research Workshop. – 2003. – № 3. – PP. 90–105.

Надійшла до редколегії 07.06.2011