

РОЗРОБКА АЛГОРИТМУ ВИЯВЛЕННЯ ЗАШУМЛЕННЯ ЯК ФАЛЬСИФІКАЦІЇ ЦИФРОВОГО ЗОБРАЖЕННЯ

В.В. Зоріло, О.Ю. Лебедєва, Н.О. Бензар

Державний університет «Одеська політехніка»,
1, пр. Шевченка, Одеса, 65044, vikazorilo@gmail.com

Маніпуляції із зображеннями знизили нашу довіру до цифрових зображень, завдяки більш витонченим методам підробок, що стають все більшим викликом цілісності зображень та їх справжності. В епоху цифрових технологій наявність потужного та відносно недорогого програмного забезпечення для редагування зображень означає, що створення візуально переконливих фотопідробок поширюється з неймовірною швидкістю. Зростання маніпуляцій із фотографіями має наслідки майже у всіх сферах – від правоохоронних органів та національної безпеки до наукових видань, політики, засобів масової інформації та реклами. Важливість цього питання стає очевидною, якщо врахувати, що частіше за все в сучасному суспільстві ми все ще покладаємось на людей, які судять про достовірність зображення, але сучасні технології дозволяють обманути людське око. Проблема автентичності поширюється майже на всі цифрові зображення, від тих, які використовують як докази в залі суду, до тих, які ми бачимо щодня в газетах та журналах. Необхідна впевненість в тому, що у використуванні цифрові зображення не були внесені будь-які зміни. Оскільки мало уваги у відкритому друці приділяється виявленню обробки зображення різноманітними фільтрами, зокрема, фільтром графічного редактора AdobePhotoshop – шум з рівномірним розподілом, - метою даної роботи є підвищення ефективності виявлення обробки цифрового зображення шляхом розробки алгоритму виявлення шуму з рівномірним розподілом графічного редактора AdobePhotoshop. Розроблено алгоритм виявлення зазначеного фільтра, заснований на аналізі сингулярних чисел блоків матриці цифрового зображення. Експериментально встановлено, що кількість блоків необробленого зображення, найменше сингулярне число яких буде перевищувати значення 16, відносно до загальної кількості блоків не перевищує 0,14%. На основі результатів обчислювального експерименту розроблено алгоритм, ефективність якого оцінено в термінах помилок 1 і 2 роду: помилки першого роду не були виявлені, а помилки другого роду складають 0,89%.

Ключові слова: цифрове зображення, порушення цілісності цифрового зображення, сингулярні числа, матриця зображення, шум з рівномірним розподілом.

Вступ

Є безліч способів змінити зображення: фотомонтаж, клонування, різні фільтри графічних редакторів. І усі способи підробки впливають на математичні параметри зображення по різному. Оскільки немає універсального методу виявлення порушень цілісності цифрового зображення, задача створення нових та вдосконалення існуючих методів пошуку слідів підробки є актуальною.

На сьогоднішньому рівні розвитку цифрові технології можуть імітувати майже все – музику, голос, зображення, рухомі об'єкти. Будь-які звуки генеруються в сучасних аудіоредакторах. Створити графічне зображення, яке буде схоже на звичайну фотографію, можна в графічному редакторі. У програмах 3D-моделювання створюють рухливі моделі, які неможливо відрізнити від об'єктів реальної відеозйомки. Всім цим дуже легко ввести в оману людей.

Цифрові зображення часто використовуються в суді в якості речового доказу або демонстративного матеріалу [1]. Тим не менше, ніколи не було так легко зробити так, щоб фотографії спотворювали правду, ніж зараз. Важливо визначити, чи є інформація, надана

слідству, правдивою та чи не спотворена вона.

Демонстративні докази можуть бути потужним інструментом, який допомагає присяжним зрозуміти важку концепцію, підкріпити їхні переконання, або навіть переконати їх повірити у щось. Правила допуску демонстративних доказів є менш суворими, ніж для речових доказів, що створює ще більший потенціал для допущення фальсифікованих цифрових зображень.

Ще однією галуззю, де фальсифікація зображень є важливою проблемою – це наука. В журналі *ScientificReports* 10 вересня 2018 року було опубліковано суперечливу статтю [2], в якій стверджувалося, що гомеопатичний препарат з рослини отруйного плюща може полегшити біль у щурів. Незважаючи на те, що спочатку шум відбувся через запеклу дискусію щодо ефективності гомеопатії, він посилювався, коли виявилось, що стаття має фальсифіковане зображення [3].

Ще один приклад стався, коли дослідники, пов'язані з Міжнародним агентством Всесвітньої організації охорони здоров'я з дослідження раку (IARC), робили маніпулювання зображеннями в багатьох наукових роботах, які вони публікували протягом багатьох років. У період між 2005 і 2014 роками Массімо Томмазіно, керівник групи інфекцій та біології раку в IARC у Ліоні, Франція, та його колишній колега Узма Хасан опублікували безліч статей із підробленими зображеннями [4].

На нараді з плагіату в Лондоні в 2009 році, Вірджинія Барбур, головний редактор журналу *PLoS Medicine*, повідомила, що проблема маніпуляції зображеннями «підкралася» до редакторів журналів з моменту появи редагуючого програмного забезпечення, такого як, наприклад, *Adobe Photoshop* [5].

Часто потрібно більше, ніж добре навчене око, щоб виявити такі маніпуляції. Тут дуже доречні інструменти цифрової криміналістики. Програмне забезпечення використовується для виявлення навмисних змін зображень у наукових роботах. Навчання людей користуватися такими ресурсами також може зайняти час, але, як показують ці тривожні висновки, це є важливим для гарантування достовірності та наукової обґрунтованості.

Щоб впоратися з проблемою фальсифікації зображень, виникла сфера криміналістики цифрових зображень як основна галузь знань, зосереджена на перевірці справжності та цілісності цифрових зображень, яка забезпечила певну довіру до них. Важливість і актуальність криміналістики цифрових зображень залучила різних дослідників до створення методів виявлення фальсифікації зображень.

Методи виявлення цифрової підробки зображень поділяються на два підходи: активний та пасивний [6]. При активному підході цифрове зображення вимагає попередньої обробки зображення, наприклад, вбудовування водяних знаків або формування електронного підпису. Однак для цього типу технологій потрібно спеціальне програмне або апаратне забезпечення, щоб вставити або витягти інформацію.

На відміну від методів, що базуються на водяних знаках та підписах, пасивні методи не потребують цифрового підпису чи вбудовування будь-якого водяного знаку. Основний підхід, який використовується в переважній більшості алгоритмів, полягає в аналізі ковзного вікна [7]. Під ковзним вікном розуміється деяка локальна область зображення, яка послідовно проходить всі пікселі зображення. З порівняння характеристик центрального пікселя вікна з іншими пікселями вікна робиться висновок про його пошкодження. Відмінність методів детектування пошкоджених пікселів один від одного полягає у виборі розмірів ковзного вікна і алгоритму прийняття рішення про пошкодження пікселя [8-10]. Однак ефективність пошуку пошкоджених пікселів істотно залежить від налаштувань. Наприклад, алгоритм *SD-ROM*, що лежить в основі великої кількості інших алгоритмів, містить чотири параметри, які вибираються користувачем. Варіювання даних параметрів може змінювати ефективність пошуку пошкоджених пікселів від 62% до 96%. При цьому залишається високим відсоток помилкових спрацьовувань.

Поліпшення результатів вдається домогтися з використанням алгоритмів, які аналізують не тільки невелику локальну область, а й все зображення в цілому. В якості таких можна відзначити алгоритми на основі асоціативних правил, методах підтримки прийняття рішень, а також ті, що засновано на методах сегментації зображень. У всіх перерахованих алгоритмах один або два налаштовуючих параметра і досить висока ефективність виявлення пошкоджених пікселів. При цьому вдається істотно знизити відсоток помилкових спрацьовувань.

Один з фільтрів, що їх використовують при обробці цифрового зображення, фільтр графічного редактора AdobePhotoshop–шум з рівномірним розподілом. У відкритому друці не виявлено згадувань про методи виявлення цього шуму.

Мета роботи

Метою роботи є підвищення ефективності виявлення обробки цифрового зображення шляхом розробки алгоритму виявлення шуму з рівномірним розподілом графічного редактора AdobePhotoshop.

Для досягнення поставленої мети необхідно вирішити наступні задачі:

- визначити параметри цифрового зображення для аналізу;
- провести обчислювальний експеримент з використанням зображень, оброблених фільтром фоторедактору Adobe Photoshop «Шум з рівномірним розподілом»;
- розробити алгоритм методу виявлення обробки зображення фільтром «Шум з рівномірним розподілом» і провести аналіз його ефективності.

Зазначений фільтр графічного редактора AdobePhotoshop проявляється на зображенні у вигляді випадковим чином розташованих точок, які відрізняються від зображення більш світлим або темним відтінком кольору. Усюди далі під зашумленням ЦЗ будемо розуміти зашумлення з рівномірним розподілом графічного редактора AdobePhotoshop.

Основна частина

Основним параметром зашумлення у редакторі Photoshop є відсоток шуму. Чим більший відсоток, тим сильніший і помітніший шум. Зашумлення з великим відсотком у більшості випадків викликає підозри до ЦЗ, що є небажаним, якщо метою цього зашумлення було приховання результатів стеганографічної атаки чи фальсифікації.

У зв'язку з цим, для проведення дослідження було обрано значення шуму в 8% та 10%. Для цифрових зображень ефект зашумлення таких значень не є помітним при перегляді зображення в цілому. Вплив на ЦЗ шуму з різним відсотком показано на рисунках 1-4.



Рис. 1. Оригінальне зображення.



Рис. 2. Шум 10%.

Як бачимо, для цифрових зображень ефект зашумлення у 10% не є помітним.



Рис. 3. Шум у 50%.

Шум у 50% вже є сильно помітним. Бачимо, що чим більший відсоток зашумлення, тим помітніший шум.



Рис. 4. Шум у 200%.

Вибір параметрів для аналізу обґрунтовано роботами [11-12]. В даних роботах доведено, що сингулярне розкладання однозначно визначає матрицю цифрового

зображення, тобто будь-які зміни в цифровому зображенні позначаються на сингулярних числах і векторах його матриці. Максимальне сингулярне число відповідає головним чином низьким частотам. Зі зменшенням сингулярного числа вклад низьких частот стає меншим, а високих – більшим. Сингулярні числа на відміну від сингулярних векторів є добре зумовленими, тобто зміни сингулярних чисел є порівняними з величиною збурюючої дії. Таким чином, для забезпечення принципової можливості отримання оцінки сили впливу зашумлення як набір формальних параметрів, які характеризують ЦЗ, з урахуванням мети роботи доцільно використовувати сингулярні числа (СНЧ) матриці ЦЗ [11-12].

Для визначення особливості поведінки СНЧ при наявності шуму було проведено аналіз сингулярних чисел блоків 8×8 матриць незашумлених ЦЗ та зображень, що піддавалися процедурі зашумлення. Аналіз проводився на множині V із 300 цифрових зображень, 100 з яких було знято на камеру смартфона Redmi Note 8T та збережено у форматі JPG (множина $V1$), 100 – на непрофесійну фотокамеру (множина $V2$), які також зберігаються у форматі JPG високої якості, 100 – професійні фотознімки з американської бази фотографій NSCR [13], які було взято у форматі без втрат TIFF (множина $V3$). Всі цифрові зображення мали різну глибину різкості зображуваного простору. Шум з рівномірним розподілом в 30% (із 400 можливих) накладався в графічному редакторі AdobePhotoshop за допомогою меню «Додати шум» з типом розподілення «Рівномірний», що рівномірно розподіляє шум, на відміну від розподілу «За Гаусом», що створює різнорідну структуру шумів [14]. Такий тип розподілення був обраний тому, що рівнорозподілений шум менш помітний на обробленому зображенні, тому для цілей фальсифікації є більш бажаним.

Усі ЦЗ представлено у колірній моделі RGB. Для роботи із зображенням у контексті поставленої мети не принципово, яку колірну компоненту аналізувати, адже для матриць R, G, B будуть приблизно однакові результати, а перевірка за всіма компонентами зображення значно подовжить тривалість роботи алгоритму. Тому обираємо будь-яку колірну компоненту і аналізуємо блоки цієї матриці.

В результаті аналізу було встановлено, що сингулярні числа блоків матриці зображення зростають до значень, які в нормальних умовах майже не зустрічаються. Тобто накладання шуму збільшує високочастотну складову до аномальних значень, що можна використовувати для ідентифікації шуму на зображенні. Зокрема, знайдено граничне значення найменшого сингулярного числа блоку матриці, рівне 16, що дозволяє відокремити переважну більшість зашумлених ЦЗ від незашумлених.

В результаті проведеного експерименту було визначено, що основна кількість блоків матриці, що мають значення останніх сингулярних чисел більше 16, в незашумлених цифрових зображеннях знаходиться в межах від 0 до 0,14% від загальної кількості блоків матриці ЦЗ (таблиця 1).

Таблиця 1.

Результати аналізу блоків СНЧ незашумлених ЦЗ

Відсоток блоків, що мають значення останнього СНЧ більше 16, відносно загальної кількості блоків	Кількість зображень
0-0,14%	297
$\geq 0,14\%$	3

У зображеннях формату JPG кількість таких блоків у 100% випадків менше 0,14% від загальної кількості блоків. У зображеннях формату TIFF кількість цих блоків в деяких випадках перевищує цей поріг (таблиця 2).

Таблиця 2.

Результати аналізу блоків СНЧ незашумлених ЦЗ різних форматів

Відсоток блоків, що мають значення останнього СНЧ більше 16, відносно загальної кількості блоків	Кількість зображень		
	JPG (V1)	JPG (V2)	TIFF (V3)
0-0,14%	100	100	97
≥0,14%	0	0	3

Дослідження проводилося на тій самій вибірці V із 300 зображень різного формату та різного розміру.

У зображеннях з накладеним шумом у 10% переважна кількість блоків, значення найменших СНЧ чисел яких більше за 16, перевищує 0,14% від загальної кількості блоків матриці зображення. У ЦЗ з шумом у 8% відсоток цих блоків такий, що однозначно відділити зашумлене цифрове зображення від норми неможливо (таблиці 3-4). Тому встановлюємо мінімальний відсоток зашумлення, яке можна виявити – 10%. Поріг кількості блоків цифрового зображення, значення останніх СНЧ яких більше 16, встановлюємо у 0,14% для оптимальної ефективності розпізнавання зашумлення.

Таблиця 3.

Результати аналізу блоків СНЧ ЦЗ з різним відсотком зашумлення

Відсоток блоків, що мають значення останнього СНЧ більше 16, відносно загальної кількості блоків	Кількість зображень	
	10%	8%
0-0,14%	0	251
≥0,14%	300	49

Таблиця 4.

Результати аналізу блоків СНЧ ЦЗ різних форматів з зашумленням 8%

Відсоток блоків, що мають значення останнього СНЧ більше 16, відносно загальної кількості блоків	Кількість зображень		
	JPG (V1)	JPG (V2)	TIFF (V3)
0-0,14%	100	87	64
≥0,14%	0	13	36

На підставі проведених досліджень розроблено алгоритм виявлення зашумлення ЦЗ, який складається з наступних етапів.

1. Отримати матрицю цифрового зображення R .
2. Розбити матрицю R на 8×8 -блоки b_{ij} , що не перетинаються (i, j – індекси блоку, що розглядається; $1 \leq i \leq \lfloor \frac{n}{8} \rfloor$, $1 \leq j \leq \lfloor \frac{m}{8} \rfloor$).
3. Обчислити сингулярні числа для кожного блоку b_{ij} .
4. Обчислити кількість блоків k , значення найменших сингулярних чисел яких більше 16.
5. Визначити відношення до загальної кількості блоків матриці ЦЗ (d).
6. Якщо $d \geq 0,14\%$, то зображення піддавалося процедурі зашумлення за допомогою накладання фільтру «Шум з рівномірним розподілом»; інакше – зображення не піддавалося обробці фільтром.

Аналіз ефективності розробленого методу показав, що помилки 2 роду склали 0,89%, а помилки 1 роду не були виявлені.

Висновки

У результаті виконання роботи було проведено аналіз існуючих методів перевірки цілісності цифрового зображення, який показав, що розробка та реалізація нових алгоритмів для виявлення фальсифікації цифрового зображення є надзвичайно актуальною задачею захисту інформації.

Було досліджено вплив фільтру графічного редактора AdobePhotoshop «Шум з рівномірним розподілом» на сингулярні числа блоків матриці цифрового зображення. В результаті встановлено, що аналіз найменших сингулярних чисел блоків дозволяє відділити оброблене зображення від необробленого.

В якості показника, що вказує на наявність обробки зображення, використано відношення блоків, найменше сингулярне число яких більше 16, до загальної кількості блоків. За результатами обчислювального експерименту встановлено, що коли значення даного відношення перевищує 0,14%, зображення є обробленим, інакше воно не є обробленим зазначеним фільтром.

Розроблено алгоритм виявлення фальсифікації цифрового зображення, кількість помилок 2 роду якого складає 0,89%, помилки першого роду не виявлені.

Список літератури

1. Parry Z. Digital manipulation and photographic evidence: defrauding the courts one thousand words at a time. *Journal of Law, Technology & Policy*. 2009. №2009. P. 175-202.
2. Magar S. Ultra-diluted Toxicodendron pubescens attenuates pro-inflammatory cytokines and ROS- mediated neuropathic pain in rats. *Scientific Reports*. 2018. № 8. P.1-11.
3. Guglielmi G. Peer-reviewed homeopathy study sparks uproar in Italy. *Nature*. 2018. № 562. P. 173-174.
4. Schneider L. Who cures cancer in Photoshop: URL: <https://forbetterscience.com/2018/10/11/who-cures-cancer-in-photoshop>.
5. Gilbert N. Science journals crack down on image manipulation. URL: <https://www.nature.com/articles/news.2009.991/box/1>.
6. Kumar B.S., Karthi S., Karthika K., Cristin R. A Systematic Study of Image Forgery Detection. *Journal of computational and theoretical Nanoscience*. 2018. № 15 (8). P. 2560–2564.
7. Garnett R., Huegerich T., Chui C., He W. A Universal Noise Removal Algorithm with an Impulse Detector. *IEEE Trans Image Process*. 2005. № 14 (11). P. 1747-1754.
8. Сорокин С.В., Щербаков М.А. Реализация SD-ROM фильтра на основе концепции нечеткой логики. *Известия высших учебных заведений. Поволжский регион*. 2017. № 3. С. 56-65.
9. Красовский Г.Я., Усс М.Л. Фильтрация изображений, искаженных импульсными помехами точечного и строчного типа, на основе систем интегрированных функций. *Радиоелектронні і комп'ютерні системи*. 2003. № 2. С. 47-55.
10. Chan R., Ho C., Nikolova M. Salt-and- pepper noise removal by median-type noise detectors and detail-preserving regularization. *IEEE Trans. Image Proc*. 2005. № 14 (10). P. 1479-1485.

11. Кобозева А.А. Использование теории возмущений для обнаружения фальсификации цифрового изображения. *Проблеми інформатиз. та управл. Матеріали міжнар. наук.-техн. конф. «Комп'ютерні системи та мережні технології»*. 2008. № 1. С. 16-22.
12. Кобозева А.А., Рыбальский О.В., Трифонова Е.А. Матричный анализ – основа общего подхода к обнаружению фальсификации цифрового сигнала. *Вісник Східноукраїнського національного університету ім. В. Даля*. 2008. № 8 (126), Ч.1. С.62–72.
13. NRCS Photo Gallery: [Электронный ресурс] // United States Department of Agriculture. Washington, USA. Режим доступа: <http://photogallery.nrcs.usda.gov> (Дата звернення: 30.05.2021).
14. Фильтры подменю «Шум» в Photoshop Elements. URL:<https://helpx.adobe.com/ru/photoshop-elements/using/noise-filters.html> (Дата звернення: 30.05.2021).

DEVELOPMENT OF AN ALGORITHM FOR DETECTION OF NOISE AS FALSIFICATION OF DIGITAL IMAGE

Zorilo V., Lebedeva O., Benzar N.

Odessa National Polytechnic University,
1, Shevchenko Ave, Odessa, 65044, Ukraine, whiteswanhelena@gmail.com

Image manipulation has decreased our confidence in digital images, thanks to more refined editing techniques which have become more and more of a factor in the integrity of images and their validity. In the era of digital technology, the availability of powerful and relatively inexpensive image editing software means that the creation of visually reproducible picture frames is expanding at an incredible speed. The growth of manipulations with photographs has consequences in almost all spheres - from law enforcement agencies and national security to scientific publications, politics, mass media and advertising. The importance of this issue becomes obvious, considering that in the modern society we still rely on people to judge the authenticity of the image, but modern technology allows to deceive the human eye. The problem of authenticity extends to all digital images, from those used as evidence in courtrooms to those we see every day in newspapers and magazines. It is necessary to make sure that the figures used have not been altered in any way. Since little attention is paid to the detection of image processing with various filters, in particular, the filter of the graphic editor AdobePhotoshop - noise with even distribution, - The aim of this work is to improve the efficiency of digital image processing by developing an algorithm for detecting noise with a uniform distribution of the graphical editor Adobe Photoshop. Developed an algorithm for detecting the specified filter based on the analysis of singular numbers of the digital image matrix blocks. It was experimentally established that the number of non-cropped image blocks, the smallest singular number of which will exceed the value of 16, in relation to the total number of blocks does not exceed 0.14%. Based on the results of the computational experiment the algorithm was developed, the efficiency of which was evaluated in terms of type 1 and type 2 errors: type 1 errors were not detected, while type 2 errors amounted to 0.89%.

Keywords: digitalimage, violation of digital image integrity, singular numbers, image matrix, noise with uniform distribution.