

УДОСКОНАЛЕННЯ МЕТОДУ ВИЯВЛЕННЯ ТА ЛОКАЛІЗАЦІЇ ОБЛАСТЕЙ КЛОНУВАННЯ В ЦИФРОВИХ ЗОБРАЖЕННЯХ

P.I. Васалатій, В.І. Матрос, О.Ю. Лебедева, Д.А. Маєвський

Національний університет «Одеська Політехніка»,
просп. Шевченка, 1, Одеса, 65044, Україна; e-mail: vasroma4@gmail.com, matros.s.od@gmail.com

Розглядається удосконалення методу виявлення та локалізації областей клонування в цифрових зображеннях. Значну роль у сучасному світі відіграють цифрові сигнали, зокрема, цифрові зображення, що використовуються у науці, медицині, судових розглядах, пресі тощо. Сьогодні камера є у кожного, якщо подія значуща, то у вашому розпорядженні опиняться тисячі фотографій. Проте змінити фотографію дуже просто. Це можуть зробити навіть любителі. Часто такі фальсифікації практично неможливо виявити неозброєним оком. Однією з найчастіше використовуваних операцій під час фальсифікації цифрових зображень є операція клонування, у ході якої відбувається заміна частини цифрового зображення, частиною того ж цифрового зображення. В роботі розглядається метод виявлення та локалізації областей клонування на основі коефіцієнту кореляції Пірсона. У базовому методі використовуються блоки розміром 8x8 як компроміс між часом обробки зображення та точністю виявлення областей клонування. Збільшення розміру блоку дозволить прискорити роботу базового методу. Іншим направленням удосконалення методу виявлення та локалізації клонованих блоків можна вважати використання маркерів. Маркери – це заздалегідь визначені пікселі в блоку, які використовуються для визначення перспективності блоку для виявлення області клонування. В роботі запропоновано маркери для квадратних блоків та блоків складної форми. Під блоками складної форми в роботі прийняті блоки які складаються з відповідного набору пікселів, які використовуються для виявлення та локалізації областей клонування. Наводяться основні кроки удосконаленого методу виявлення та локалізації областей клонування в цифрових зображеннях. Оцінку ефективності удосконаленого методу виявлення та локалізації областей клонування виконано у вимірюванні часу виявлення та локалізації областей клонування до удосконалення та після.

Ключові слова: зображення, фальсифікація зображень, виявлення клонування областей, блоки складної форми.

Вступ

Суспільство вступило у період свого розвитку, який на загальну думку можна назвати інформаційним. Процес впровадження нових інформаційних технологій у всі сфери життя суспільства немислимий без вирішення питань інформаційної безпеки, зокрема питань, пов'язаних із забезпеченням/виявленням порушень цілісності цифрових контентів.

У сучасному цифрові технології стали невід'ємною частиною життя людини. У сервісі Google Images на сьогоднішній день можна знайти більше 136 мільярдів фотографій. Кожен рік по всьому світу роблять приблизно 1,72 трильйони фотографій. Кожен день користувачі інтернету діляться між собою 3,2 мільярдами зображень. Через це в інтернеті легко зіткнутися з фальсифікованими зображеннями, навіть у довірених новинних виданнях, оскільки у вільному доступі знаходяться такі графічні редактори, як Adobe Photoshop, Figma, GIMP та Photo Pos Pro.

У наш час цифрові зображення (ЦЗ) можуть бути просто змінені за допомогою комп'ютерів та редагування фотографій, тощо. Ці зміни вплинуть на достовірність зображень у законодавстві, політиці, ЗМІ, промисловості та

медицині.

Фальсифікація – це маніпуляції з цифровим зображенням, щоб приховати певну значущу або корисну інформацію зображення. Виявлення фальсифікації цифрових зображень є однією з найбільш критичних аналітичних практик у наш час. Через це зростає необхідність у ефективних методах виявлення фальсифікованого зображення.

Однією з найчастіше використовуваних операцій під час фальсифікації цифрових зображень, реалізованої у всіх графічних редакторах, є операція клонування, у ході якої відбувається заміна частини (частин) цифрового зображення, частиною (частинами) того ж цифрового зображення.

Через загальнодоступність програмного забезпечення, що дозволяє обробляти і редагувати ЦЗ, а також у зв'язку з різноманітністю способів їх фальсифікації, зростає необхідність у вдосконаленні існуючих та розробці нових методів виявлення порушення цілісності ЦЗ, що є обов'язковою складовою частиною будь-якої сучасної комплексної системи захисту інформації. Тому зростає актуальність методів виявлення порушення цілісності цифрових зображень.

Постановка задачі та мети дослідження

У роботі буде розглядатись фальсифікація цифрових зображень методом клонування, яку можна реалізувати у будь-якому графічному редакторі за допомогою інструментів виділення області, копіювання або за допомогою інших інструментів. Так в графічному редакторі Adobe Photoshop для цих цілей можна використовувати такі інструменти як Штамп (Clone Stamp), Пензель відновлення (Healing Brush) і Латка (Patch). Для більш вдалого клонування зловмисники також можуть використовувати Гумка (Erase) и Розмиття (Blur) для «замітання слідів» своєї фальсифікації.

Метою роботи є підвищення ефективності методу виявлення та локалізації областей клонування в цифрових зображеннях шляхом його модифікації, заснованої на використанні маркерів.

Для досягнення поставленої мети необхідно вирішити такі завдання:

- огляд методів та засобів виявлення та локалізації областей клонування в цифрових зображеннях;
- проаналізувати метод виявлення та локалізації областей клонування в цифрових зображеннях та провести вибір маркерів для блоків, які використовуються для виявлення областей фальсифікації;
- вибір виду складних блоків для виявлення областей клонування;
- удосконалення методу виявлення та локалізації областей клонування в цифрових зображеннях;
- програмна реалізація удосконаленого методу виявлення та локалізації областей клонування в цифрових зображеннях та оцінити його ефективність.

Під ефективністю в цій роботі будемо розуміти зменшення часу роботи методу виявлення та локалізації областей клонування в цифрових зображеннях.

Основна частина

Для захисту та виявлення порушення цілісності ЦІ створюються відповідні методи та алгоритми, які поділяються на дві великі категорії: активні, пасивні. В активних методах деяка інформація попередньо впроваджується у ЦІ. Пасивні методи дозволяють підтвердити цілісність ЦІ або виявити її порушення без впровадження додаткової інформації. У роботі розглядатимуться пасивні методи.

Виходячи з найчастіших операцій, що здійснюються засобами графічних редакторів над цифровими зображеннями під час їх фальсифікації, за способом створення ці фальсифікації можна розбити на такі категорії:

- клонування – коли одна область зображення копіюється і вставляється в іншу область цього зображення;
- фотомонтаж – коли використовуються частини двох чи кількох зображень для створення нового зображення;
- обробка – коли зображення або його частини піддаються таким операціям, як зміна масштабу, поворот, розмиття, зміна яскравості, кольору тощо.

При підробці областей клонування одна частина зображення копіюється та вставляється в інше місце того самого зображення, щоб приховати інформацію або змінити значення зображення [1]. Отже, між оригінальною та клонованою областю існує сильна кореляція, яку можна використовувати як доказ для виявлення підробки копіювання-переміщення.

Існують різні методів виявлення областей клонування, які спираються на різні компоненти, наприклад яскравість пікселів або сингулярні числа, різні ускладнення, наприклад фальсифікація без додаткової обробки або зі збереженням у форматі з втратами. Кожен метод виявлення залежить від поставленої задачі: чи то обробити якомога швидше, чи то обробити якомога точніше.

Одним з перспективних для удосконалення та модифікації методів виявлення клонованих областей є метод виявлення на основі коефіцієнту кореляції Пірсона.

Коефіцієнт кореляції Пірсона (позначають « r ») – в статистиці, показник кореляції (лінійної залежності) між двома змінними X та Y , який набуває значень від -1 до $+1$ включно [2]. Він широко використовується в науці для вимірювання ступеня лінійної залежності між двома змінними.

Коефіцієнт кореляції Пірсона між двома змінними дорівнює сумі добутків відхилень, поділеній на добуток їх стандартних відхилень. Нехай, x – дві вибірки $x^m = (x_1, \dots, x_m)$, $y^m = (y_1, \dots, y_m)$, тоді коефіцієнт кореляції Пірсона розраховують за формулою 1:

$$r_{xy} = \frac{\sum_{i=1}^m (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^m (x_i - \bar{x})^2 \sum_{i=1}^m (y_i - \bar{y})^2} \quad (1)$$

де \bar{x} , \bar{y} – вибіркові середні x^m і y^m .

Для представлення цифрового зображення розміром $n \times m$ пікселів в роботі використовується: $n \times m$ – матриці R, G, B (колірна схема RGB); $n \times m$ – матриця яскравості Y (колірна схема YUV).

У роботі [3] пропонується метод виявлення областей клонування шляхом використання коефіцієнта Пірсона який має наступні кроки:

1. Перевести зображення з режиму RGB до YUV.
2. Ініціалізувати область, що треба виявити $Res = \emptyset$ та розбити матрицю яскравості цифрового зображення на множину квадратних блоків, що перетинаються, $C = \{c_1, c_2, \dots, c_s\}$, $\bigcup_{i=1}^s c_i = Y$, розміром $p \times p$ (тут кожний наступний блок відрізняється від попереднього зсувом на 1 піксель вправо, вліво, вниз та угору).
3. Кожний блок c_i , $i = 1, \dots, s$ розглянути в парі з усіма c_j , $j = i+1, \dots, s$, відповідно. Для кожної пари розраховується коефіцієнт кореляції Пірсона.
 - 3.1 Якщо кореляція дорівнює 1, то блоки c_i та c_j – це оригінальний та клонований, після чого до результату додаються обидва блоки: $Res = Res \vee c_i \vee c_j$.
 - 3.2 Якщо кореляція не дорівнює одиниці, то перейти до наступної пари блоків.
4. Вивести знайдену область Res .

Ілюстративний приклад роботи описаного базового методу виявлення та локалізації областей клонування представлений на рисунку 1.

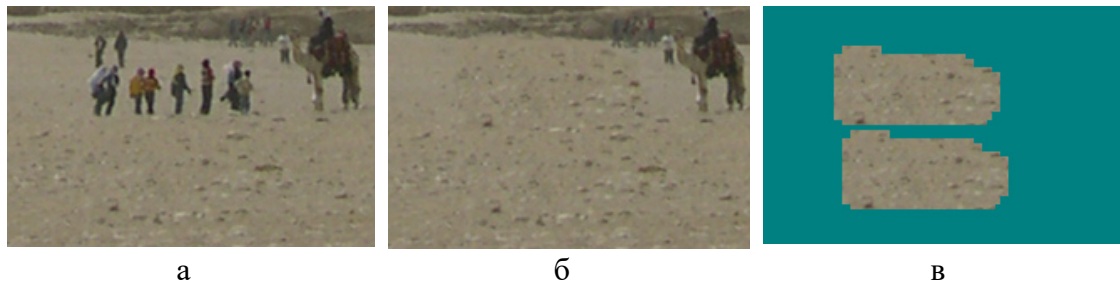


Рис. 1. Результат роботи методу виявлення та локалізації областей клонування: оригінальне зображення (а); фальсифіковане (б); результат виявлення стандартними блоками (в)

У базовому методі використовуються блоки розміром 8×8 через компроміс між часом обробки зображення та точністю виявлення областей клонування. Хоча зменшення розміру блоку дозволяє уточнювати область фальсифікації, чим більше зменшується розмір блоку, тим більше буде кількість помилок другого роду. Збільшення розміру блоку прискорює роботу базового методу.

Іншим напрямленням удосконалення методу виявлення та локалізації клонованих блоків можна вважати використання маркерів. Маркери – це заздалегідь визначені пікселі обох блоків, які будуть зрівнюватись. Для квадратних блоків різних розмірів в роботі було запропоновано використовувати кутові пікселі у якості маркерів (рис. 2) **Ошибка! Источник ссылки не найден..**

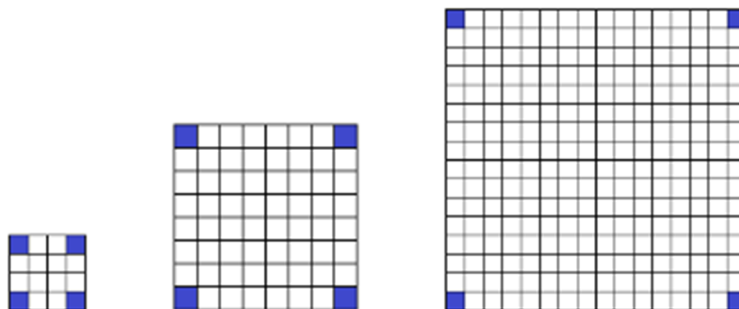


Рис. 2. Маркери блоку квадратної форми розміру 4×4 , 8×8 , 16×16

Для блоків, розміром 4×4 в якості маркеру для цілого блоку беруться лише 4 значення яскравості, що відповідають пікселям, які знаходяться в кутах блоку. Це значно підвищить ефективність роботи алгоритму виявлення областей клонування, але якщо брати блоки розміром 16×16 , виявлена область клонування буде меншою, ніж є насправді. Оптимальним є використання маркерів для квадратних блоків розміром 8×8 .

Перед тим як вирахувати коефіцієнт кореляції при порівнянні двох блоків, будемо порівнювати між собою маркери блоків. Якщо маркери у двох блоків співпадають, то припускаємо, що маємо справу зі схожими (клонуваними та оригінальними) блоками та тільки тоді будемо вираховувати коефіцієнт кореляції для остаточного підтвердження цього припущення.

Тоді удосконалений метод для виявлення та локалізації областей клонування буде мати наступні шаги:

1. Ініціалізувати область, що треба виявити $res = \emptyset$ та розбити матрицю яскравості Y цифрового зображення на множину блоків, що перетинаються,

розміром $p \times p$ пікселей $C = \{c_1, c_2, \dots, c_s\}$ таких, що: $\bigcup_{i=1}^s c_i = Y$, (тут кожний наступний блок c_i відрізняється від попереднього c_{i-1} зсувом на 1 піксель вправо, вліво, вниз та угору).

2. Кожний блок c_i , $i = 1, \dots, s$, розглянути в парі з усіма c_j , $j = i + 1, \dots, s$, відповідно. Для кожної пари:

2.1 Для кожної пари блоків c_i та c_j :

2.2 Отримати набори маркерів mc_i та mc_j . Якщо маркери mc_i не дорівнюють mc_j , то розглядати наступну пару блоків. Інакше виконати наступні шаги:

2.2.1 Розрахувати коефіцієнт кореляції: $cor = correlation(c_i, c_j)$.

2.2.2 Якщо $cor = 1$, то блоки c_i и c_j являються оригінальним та клонованим,

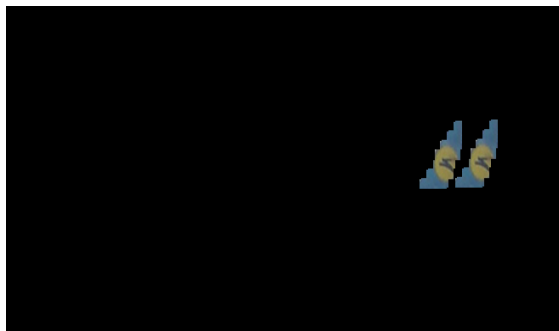
$$res = res \cup c_i \cup c_j.$$

3. Вивести знайдену область res .

Розглянемо приклад результату роботи удосконаленого методу. Знайдені клоновані області при запусках без використання маркерів та з використанням маркерів будуть однакові (рисунок 3).



а



б



в

Рис. 3. Результат роботи алгоритму: зображення, що обробляється (а); пошук по блокам без використання маркерів (б); пошук по блокам з використання маркерів (в)

Оцінку ефективності удосконаленого методу виявлення та локалізації областей клонування виконано у вимірюванні часу виявлення та локалізації областей клонування до удосконалення та після. Нижче в таблиці 1 наведені дані експерименту для деяких зображень при використанні блоків 8×8 .

Таблиця 1

Дані експериментів при використанні блоків 8x8

| № зображення | Розмір зображення | Час роботи базового методу (годин:мінут:секунд) | Час роботи удосконаленого методу (годин:мінут:секунд) |
|--------------|-------------------|---|---|
| 1 | 240 x 480 | 00:09:44 | 00:00:54 |
| 2 | 272 x 400 | 00:21:59 | 00:00:49 |
| 3 | 320 x 880 | 00:47:52 | 00:03:51 |
| 4 | 320 x 832 | 00:39:43 | 00:03:27 |
| 5 | 272 x 704 | 00:21:00 | 00:01:49 |
| 6 | 320 x 480 | 00:14:30 | 00:01:11 |
| 7 | 880 x 512 | 02:03:02 | 00:10:07 |
| 8 | 240 x 512 | 00:08:28 | 00:00:44 |
| 9 | 240 x 352 | 00:05:41 | 00:00:23 |
| 10 | 176 x 352 | 00:02:17 | 00:00:11 |

Маркери можна використовувати не тільки для квадратних блоків, а і для блоків різної форми. В роботі також розглядалися блоки складної форми. Під блоками складної форми будемо розуміти блоки які складаються з відповідного набору пікселів, які використовуються для виявлення та локалізації областей клонування.

В роботі проводився експеримент з використанням блоків складної форми розміром 16x16. В експерименті використовувались складні блоки різних форм. За результатом експерименту було вирішено залишити для виявлення областей клонування 3 блоки (рисунок 4), де сірим позначені пікселі, що використовуються.

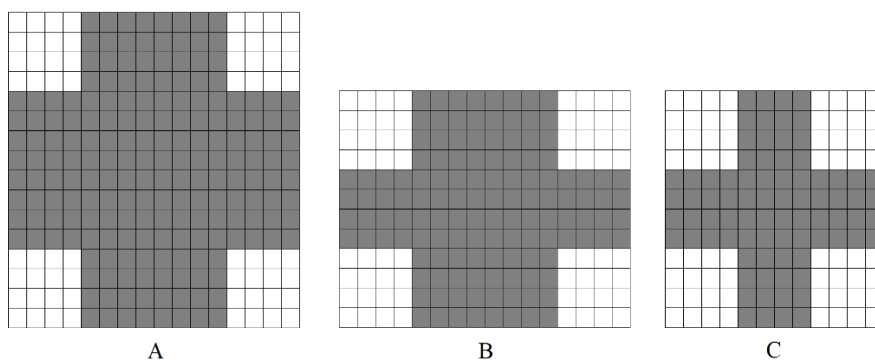


Рис. 4. Типи блоків складної форми

Блок А залишився 16x16. Розмір блоку В змінився на 16x12, а блоку С – на 12x12. Кутові пікселі в блоках прибрані, щоб краще відстежувати нерівні контури області клонування.

Проводився експеримент на цифрових зображеннях з використанням обраних типів блоків. Результати роботи використання блоків складної форми для виявлення та локалізації областей клонування продемонстровані на рисунку 5.

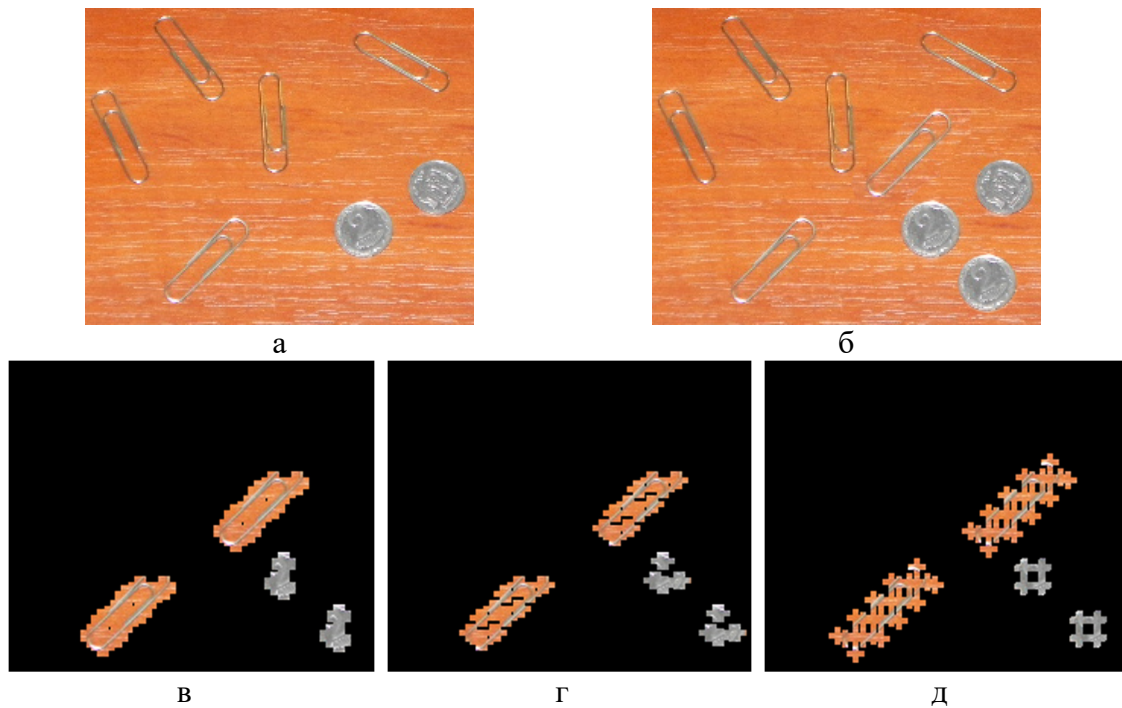


Рис. 5. Результат виявлення областей клонування з використанням складних блоків: оригінальне зображення (а); фальсифіковане (б); складними блоками типу А (в); складними блоками типу В (г); складними блоками типу С (д);

Оскільки ці блоки краще себе проявляють на периметрі області клонування, то, щоб уникнути пустих дир, що будуть утворені через відсутність кутових пікселів блоку, то у подальшому до методу необхідно додати додаткову перевірку. Якщо коефіцієнт кореляції між двома складними блоками дорівнюватиме одиниці, то буде проведена додаткова перевірка на кореляцію, використовуючи вже квадратні блоки того ж розміру.

Для обраних блоків складної форми було вирішено використовувати маркери, які продемонстровано на рисунку 6.

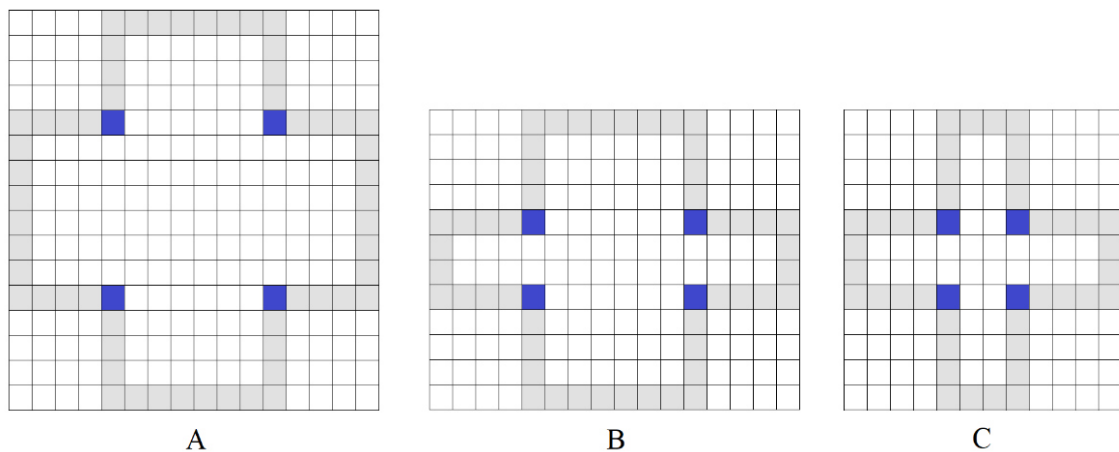


Рис. 6. Маркери блоків складної форми

Результати експериментів з використання блоків складної форми наведені нижче у таблиці 2 та 3.

Таблиця 2

Час роботи методу без маркерів при використанні блоків складної форми у форматі (ГГ:ХХ:СС)

| Номер зображення | Розмір | Час роботи | | | | |
|------------------|---------|---------------------|-----------------------|----------------------------|----------------------------|----------------------------|
| | | квадратні блоки 8x8 | квадратні блоки 16x16 | блок складної форми типу А | блок складної форми типу В | блок складної форми типу С |
| 1 | 301x235 | 0:03:08 | 0:01:36 | 0:01:44 | 0:01:49 | 0:02:26 |
| 2 | 352x176 | 0:02:27 | 0:01:17 | 0:01:13 | 0:01:29 | 0:01:56 |
| 3 | 256x208 | 0:01:56 | 0:00:57 | 0:01:07 | 0:01:07 | 0:01:27 |
| 4 | 216x216 | 0:01:30 | 0:00:40 | 0:00:47 | 0:00:47 | 0:01:09 |
| 5 | 704x272 | 0:25:06 | 0:12:25 | 0:17:40 | 0:17:10 | 0:18:05 |
| 6 | 880x320 | 1:02:02 | 0:26:20 | 0:44:54 | 0:37:54 | 0:43:31 |
| 7 | 832x320 | 0:49:50 | 0:23:19 | 0:28:38 | 0:28:03 | 0:41:06 |
| 8 | 512x240 | 0:11:34 | 0:04:57 | 0:06:39 | 0:06:01 | 0:09:58 |

Як можна побачити, навідмінно від квадратних блоків, при використанні складних блоків обробка проходить значно швидше. При тому чим більше розмір зображення, тим більша буде різниця часу обробки.

Таблиця 3

Час роботи методу з маркерами при використанні блоків складної форми у форматі (ГГ:ХХ:СС)

| Номер зображення | Розмір | Час роботи | | | | |
|------------------|---------|---------------------|-----------------------|----------------------------|----------------------------|----------------------------|
| | | квадратні блоки 8x8 | квадратні блоки 16x16 | блок складної форми типу А | блок складної форми типу В | блок складної форми типу С |
| 1 | 301x235 | 0:00:14 | 0:00:03 | 0:00:03 | 0:00:04 | 0:00:06 |
| 2 | 352x176 | 0:00:16 | 0:00:02 | 0:00:04 | 0:00:05 | 0:00:07 |
| 3 | 256x208 | 0:00:11 | 0:00:02 | 0:00:03 | 0:00:03 | 0:00:05 |
| 4 | 216x216 | 0:00:09 | 0:00:01 | 0:00:02 | 0:00:02 | 0:00:04 |
| 5 | 704x272 | 0:01:53 | 0:00:26 | 0:00:27 | 0:00:37 | 0:00:51 |
| 6 | 880x320 | 0:04:07 | 0:00:56 | 0:01:06 | 0:01:24 | 0:01:54 |
| 7 | 832x320 | 0:03:45 | 0:00:50 | 0:00:56 | 0:01:12 | 0:01:41 |
| 8 | 512x240 | 0:00:45 | 0:00:10 | 0:00:11 | 0:00:14 | 0:00:21 |

Висновки

За результатами аналізу методу виявлення та локалізації областей клонування в цифрових зображеннях був проведений вибір маркерів для квадратних блоків та блоків складної форми, які можуть використовуватися для виявлення областей фальсифікації.

Удосконалено метод виявлення та локалізації областей клонування шляхом використання маркерів в процесі виявлення областей клонування, що дозволило прискорити час роботи методу.

Список літератури

1. Rani P., Rani J. Copy-move forgery attack detection in digital images. *International Journal of Engineering Research and Technology IJERT*. 2015. V.4. Is.6. P. 118-132.
2. Коефіцієнт_кореляції_Пірсона. Матеріал з вікіпедії – вільної енциклопедії. URL: https://uk.wikipedia.org/wiki/Коефіцієнт_кореляції_Пірсона
3. Лебедева, Е.Ю., Лебедев Ю.Ф. Исследование метрик используемых при обнаружении клонированных участков изображений в задачах выявления фальсификации. *Вісник національного технічного університету ХПІ*. 2011. №35. С.25 – 31.

IMPROVEMENT OF THE METHOD OF DETECTION AND LOCALIZATION OF CLONING AREAS IN DIGITAL IMAGES

R. Vasalatiy, V. Matros, O. Lebedieva, D. Majevsky

National Odesa Polytechnic University, ave. Shevchenko, 1, Odesa, 65044, Ukraine
e-mail: vasroma4@gmail.com, matros.s.od@gmail.com

The paper considers the improvement of the method of detection and localization of cloning areas in digital images. A significant role in the modern world is played by digital signals, in particular, digital images used in science, medicine, court proceedings, the press, etc. Today, everyone has a camera, if the event is significant, you will have thousands of photos at your disposal. However, changing the photo is very easy. Even amateurs can do it. Often, such falsifications are almost impossible to detect with the naked eye. One of the most commonly used operations in the forgery of digital images is the cloning operation, in which a part of a digital image is replaced by a part of the same digital image. The paper considers the method of detection and localization of cloning regions based on the Pearson correlation coefficient. The basic method uses 8x8 blocks as a compromise between image processing time and the accuracy of cloning region detection. Increasing the block size will speed up the base method. Another way to improve the method of detecting and locating cloned blocks is to use markers. Markers are predefined pixels in a block that are used to define the perspective of the block to identify the cloning area. Markers for square blocks and blocks of complex shape are proposed in the work. Blocks of a complex shape in the work are considered to be blocks consisting of a suitable set of pixels, which are used to detect and localize areas of cloning. The main steps of the improved method of detection and localization of cloning regions in digital images are presented. Evaluation of the effectiveness of the improved method of detection and localization of cloning areas was performed by measuring the time of detection and localization of cloning areas before and after improvement.

Keywords: image, image falsification, area cloning detection, blocks of complex shape.