УДК 004.056.5

А.А. Кобозева[1]

[1]д.т.н., проф., зав.каф.информатики и управления защитой информационных систем Одесского национального политехнического университета, Одесса

МАТЕМАТИЧЕСКАЯ ОСНОВА МЕТОДА ВЫЯВЛЕНИЯ КЛОНИРОВАННЫХ УЧАСТКОВ ИЗОБРАЖЕНИЯ, ПОДВЕРГНУТЫХ КОРРЕКЦИИ ЦВЕТА

На основе общего подхода к анализу состояния и технологии функционирования информационных систем, разработанного автором ранее, построен математический базис метода выявления клонированных участков изображения, подвергнутых коррекции цвета.

Ключевые слова: клонирование, цифровое изображение, матрица, сингулярные числа, сингулярные векторы, коррекция цвета, *n*-оптимальный вектор

Введение

Различного рода фальсификации цифровых контентов, в частности изображений, средствами графических редакторов, таких как Adobe Photoshop, Adobe Illustrator, Adobe Flash, CorelDRAW и другие, являются в настоящий момент чрезвычайно распространенными, приводя к актуальности разработки методов пассивной защиты информации [1], которые часто оказываются в роли «догоняющих» для быстро и успешно развивающихся методов неавторизованных изменений информационных контентов [2].

Одно из первых мест среди инструментов графических редакторов при ранжировании по частоте использования при несанкционированных изменениях цифровых изображений (ЦИ) занимает клонирование, выявлению которого в современной печати уделяется много внимания [3,4], однако полностью рассматриваемая задача до сих пор не решена. Так до настоящего момента не существует возможности даже при удачном определении и локализации клонированной области и ее прообраза определить, какая из выделенных областей является прообразом; непреодоленные трудности вызывает выявление клонирования в случае использования в качестве постобработки клонированных участков коррекции цвета. Актуальность этих двух задач в свете роста компьютерной преступности в настоящее время, легкости проведения фальсификаций цифровых изображений, необходимости доказательства их целостности и установления аутентичности трудно переоценить.

Цель статьи и постановка исследования

Целью статьи является разработка математического базиса для метода выявления клонированных участков изображения, подвергнутых коррекции цвета, на основе матричного анализа и общего подхода к анализу состояния и технологии функционирования информационных систем (ОПАИС) [5].

Для достижения поставленной цели в работе необходимо решить следующие *задачи*:

1. Получить формальное представление результата коррекции цвета в виде возмущений параметров, входящих в полный набор, определяющих состояние/изменение состояния цифрового изображения.

2. Определить характерные особенности возмущений формальных параметров, определяющих ЦИ, в результате коррекции цвета.

Основная часть

В соответствии с ОПАИС состояние любой информационной системы, в частности, ЦИ, формально описывается совокупностью однозначно определяющих ее параметров – сингулярных чисел (СНЧ) и сингулярных векторов (СНВ) соответствующей матрицы (матриц).

Для простоты изложения рассматриваются монохромные ЦИ с $n \times n$-матрицей F. В случае цветного ЦИ (например, в формате RGB) все рассмотренные ниже выкладки могут быть применены для каждой из матриц цвета.

Для простоты изложения предположим, что все ЦИ подвергается коррекции цвета. Матричное представление такой операции в соответствии с ее реализацией в графическом редакторе Adobe Photoshop имеет следующий вид:

$$\overline{F} = F \pm K,$$

где $\overline{F}$ - матрица результирующего ЦИ, K - $n \times n$-матрица коррекции:

$$K = \begin{pmatrix} a & a & \dots & a \\ a & a & \dots & a \\ \dots & \dots & \dots & \dots \\ a & a & \dots & a \end{pmatrix}, \tag{1}$$

где a - натуральное число, определяющее количество градаций цвета, на которое происходит коррекция.

Построим для матрицы K нормальное сингулярное разложение [5]:

$$K = U_K \Sigma_K V_K^T, \tag{2}$$

где U_K, V_K - ортогональные матрицы левых и правых СНВ соответственно, левые СНВ являются лексикографически положительными [6], $\Sigma_K = diag\left(\sigma_1^K, ..., \sigma_n^K\right)$ - матрица СНЧ K: $\sigma_1^K \geq \sigma_2^K \geq ... \geq \sigma_n^K \geq 0$.

Матрица K имеет единичный ранг, что говорит о единственности отличного от нуля ее сингулярного числа [7]:

$$\sigma_1^K > 0, \quad \sigma_2^K = ... = \sigma_n^K = 0. \tag{3}$$

Одним из наиболее важных числовых параметров, характеризующих состояние ЦИ, является энергия E сигнала. Учитывая различные способы вычисления энергии [5] и соотношение (3), для K имеем:

$$E = \sum_{i,j=1}^{n} a^2 = a^2 n^2, \; E = \sum_{i=1}^{n} \left(\sigma_i^K\right)^2 = \left(\sigma_1^K\right)^2,$$

откуда, с учетом положительности σ_1^K :

$$\sigma_1^K = a \cdot n. \tag{4}$$

Тогда в (2)

$$\Sigma_K = diag(a \cdot n, 0, ..., 0).$$

Сингулярное разложение (2) может быть записано в форме внешних произведений [8]:

$$K = \sum_{i=1}^{n} \sigma_i^K u_i^K \left(v_i^K\right)^T, \tag{5}$$

где $u_i^K, v_i^K, \; i = \overline{1,n}$, - i-ые столбцы матриц U_K, V_K (левые и правые СНВ матрицы K) соответственно.

Из (5) с учетом (3) вытекает, что

$$K = \sigma_1^K u_1^K \left(v_1^K\right)^T. \tag{6}$$

Таким образом, значения элементов матрицы K вида (1) никак не зависят от набора СНВ, отвечающих СНЧ $\sigma_2^K, ..., \sigma_n^K$ (эти СНВ могут быть выбраны произвольно), а определяются лишь максимальным СНЧ и соответствующими ему левым и правым СНВ $u_1^K = \begin{pmatrix} u_{11}^K & u_{21}^K & \dots & u_{n1}^K \end{pmatrix}^T$, $v_1^K = \begin{pmatrix} v_{11}^K & v_{21}^K & \dots & v_{n1}^K \end{pmatrix}^T$, для которых в соответствии с [9]:

$$u_{j1}^K > 0, \quad v_{j1}^K > 0, \quad j = \overline{1,n}. \tag{7}$$

Из (6) следует, что матрица K может быть представлена в виде:

$$K = \sigma_1^K \begin{pmatrix} u_{11}^K \\ u_{21}^K \\ \dots \\ u_{n1}^K \end{pmatrix} \begin{pmatrix} v_{11}^K & v_{21}^K & \dots & v_{n1}^K \end{pmatrix} = \sigma_1^K \begin{pmatrix} u_{11}^K v_{11}^K & u_{11}^K v_{21}^K & \dots & u_{11}^K v_{n1}^K \\ u_{21}^K v_{11}^K & u_{21}^K v_{21}^K & \dots & u_{21}^K v_{n1}^K \\ \dots & \dots & \dots & \dots \\ u_{n1}^K v_{11}^K & u_{n1}^K v_{21}^K & \dots & u_{n1}^K v_{n1}^K \end{pmatrix}, \quad (8)$$

откуда с учетом (1) для элементов первой строки матрицы K имеем:

$$\begin{cases} \sigma_1^K u_{11}^K v_{11}^K = a, \\ \sigma_1^K u_{11}^K v_{21}^K = a, \\ \dots\dots\dots\dots \\ \sigma_1^K u_{11}^K v_{n1}^K = a \end{cases} \quad (9)$$

Из (9) с учетом (4) вытекает, что

$$\begin{cases} u_{11}^K v_{11}^K = \dfrac{1}{n}, \\ u_{11}^K v_{21}^K = \dfrac{1}{n}, \\ \dots\dots\dots \\ u_{11}^K v_{n1}^K = \dfrac{1}{n} \end{cases} \Leftrightarrow \quad v_{11}^K = v_{21}^K = \dots = v_{n1}^K = \frac{1}{n \cdot u_{11}^K} \quad (10)$$

Поскольку вектор v_1^K нормированный, из (10) вытекает:

$$\left\| v_1^K \right\|_2 = 1 = \sqrt{\left(\frac{1}{n \cdot u_{11}^K} \right)^2 \cdot n},$$

откуда в силу (7) $u_{11}^K = \dfrac{1}{\sqrt{n}}$. Тогда из (10) вытекает, что правый СНВ v_1^K, отвечающий максимальному СНЧ K, определяется как

$$v_1^K = \begin{pmatrix} \dfrac{1}{\sqrt{n}} & \dfrac{1}{\sqrt{n}} & \dots & \dfrac{1}{\sqrt{n}} \end{pmatrix}^T = n^O \in R^n,$$

где n^O - n-оптимальный вектор пространства R^n [9].

Из (8) для элементов первого столбца матрицы K получаем систему уравнений, аналогичную (9):

$$\begin{cases} \sigma_1^K u_{11}^K v_{11}^K = a, \\ \sigma_1^K u_{21}^K v_{11}^K = a, \\ \dots\dots\dots\dots\dots\dots \\ \sigma_1^K u_{n1}^K v_{11}^K = a \end{cases},$$

откуда следует, что левый СНВ $u_1^K = \left(u_{11}^K \quad u_{21}^K \quad \dots \quad u_{n1}^K\right)^T$, отвечающий максимальному СНЧ, также совпадает с n^O пространства R^n.

Пусть

$$F = U \Sigma V^T$$

- нормальное сингулярное разложение матрицы F ЦИ, где матрицы $U = \left(u_1,...,u_n\right)$, $\Sigma = diag\left(\sigma_1,...,\sigma_n\right)$, $V = \left(v_1,...,v_n\right)$ (u_i, v_i, $i = \overline{1,n}$, - столбцы матриц U,V - левые и правые СНВ F) определяются аналогично матрицам U_K, Σ_K, V_K в (2). В соответствии с [9], левый и правый СНВ матрицы F , отвечающие максимальному СНЧ σ_1, близки к n — оптимальному вектору пространства R^n. Эти векторы отвечают первым столбцам матриц U,V соответственно. Тогда, допуская незначительную погрешность, можно считать, что:

$$\overline{F} = F \pm K = U \Sigma V^T \pm U_K \Sigma_K V_K^{\ T} \approx$$

$$\approx \left(n^O \quad u_2 \ \dots \ u_n\right) \begin{pmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_n \end{pmatrix} \left(n^O \quad v_2 \ \dots \ v_n\right)^T \pm$$

$$\mp \left(n^O \quad u_2^{\ K} \ \dots \ u_n^{\ K}\right) \begin{pmatrix} \sigma_1^{\ K} & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} \left(n^O \quad v_2^{\ K} \ \dots \ v_n^{\ K}\right)^T =$$

$$
=\left(n^O \quad u_2 \; \dots \; u_n\right)\begin{pmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_n \end{pmatrix}\left(n^O \quad v_2 \; \dots \; v_n\right)^T \pm
$$

$$
\mp\left(n^O \quad u_2 \; \dots \; u_n\right)\begin{pmatrix} \sigma_1{}^K & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}\left(n^O \quad v_2 \; \dots \; v_n\right)^T =
$$

$$
=\left(n^O \quad u_2 \; \dots \; u_n\right)\begin{pmatrix} \sigma_1 \pm an & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_n \end{pmatrix}\left(n^O \quad v_2 \; \dots \; v_n\right)^T \qquad (11)
$$

Формула (11) представляет из себя нормальное сингулярное разложение матрицы $\overline{F}$ - результата цветокоррекции изображения с матрицей F. Из (11) вытекает, что результатом цветокоррекции изображения является увеличение/уменьшение в случае осветления/затемнения изображения максимального СНЧ матрицы F на величину, равную произведению an (в случае, когда элементы матрицы $\overline{F}$ не выходят за границы $[0{,}255]$). Аналогичное заключение будет иметь место в случае цветокоррекции для отдельных частей (блоков) ЦИ, являющихся результатом клонирования.

Выводы

В работе разработан математический базис метода выявления клонированных участков ЦИ в случае их постобработки при помощи цветокоррекции. Установлено, что в случае отсутствия ошибок округления в процессе формирования элементов матрицы $\overline{F}$ (все элементы $\overline{F}$ после цветокоррекции попадают в промежуток $[0{,}255]$):

1.Для клонированных участков (блоков) ЦИ после цветокоррекции СНЧ соответствующих им матриц будут одинаковыми, кроме максимальных, которые отличаются друг от друга на величину an.

2.СНВ матриц, отвечающих клонированным областям после цветокоррекции, остаются без изменения.

3.Число градаций цветокоррекции определяется по разности первых СНЧ и размеру блока.

Для полученных результатов существенным является предположение о том, что все элементы $\overline{F}$ после цветокоррекции попадают в промежуток $[0{,}255]$. При невыполнении этого условия при непосредственной разработке метода для распознавания искомых блоков ЦИ потребуется определение порогового значения как для отличий СНЧ, начиная со второго, так и для возможной погрешности числа градаций a.

Литература

4. Нариманова Е.В. Проверка целостности цифрового сигнала [Текст] : монография / Е.В. Нариманова. – Донецк: Изд. Цифровая типография, 2011. – 180 с.
5. Бобок, И.И. Адаптация стеганоаналитического метода, основанного на теории возмущений, для задачи выявления нарушения целостности цифрового изображения / И.И. Бобок, Е.В. Малахов // Інформатика та математичні методи в моделюванні. — 2012. — Том 2, №4. — С. 297–303.
6. Зорило, В.В. Метод выявления симметричного клонирования при фальсификации цифрового изображения / В.В.Зорило, А.А.Кобозева, Е.Ю.Лебедева // Інформатика та математичні методи в моделюванні. — 2013. — Том 3, №1. — С. 5-12.
7. Зорило, В.В. Выявление клонирования как фальсификации цифрового изображения / В.В. Зорило // Вісник Національного технічного університету «ХПІ». Збірник наукових праць. Тематичний випуск «Системний аналіз, управління та інформаційні технології». — Х.: НТУ «ХПІ», 2011. — № 35 — С.31–38.
8. Кобозева, А.А. Анализ информационной безопасности: монография / А.А. Кобозева, В.А. Хорошко. — К.: ГУИКТ, 2009. — 251 с.
9. C.Bergman, J.Davidson. Unitary Embedding for Data Hiding with the SVD. Security, Steganography, and Watermarking of Multimedia Contents VII, SPIE Vol. 5681,San Jose, CA, Jan. 2005.
10. Бахвалов, Н.С. Численные методы [Текст] : учебное пособие для студ. физико-математических спец. вузов; Рекомендовано МО РФ / Н.С. Бахвалов, Н.П. Жидков, Г.М. Кобельков. — 6-е изд. — М. : БИНОМ. Лаборатория знаний, 2008. — 636 с.
11. Деммель, Д. Вычислительная линейная алгебра [Текст] : теория и приложения / Д. Деммель; Пер. с англ. Х.Д. Икрамова. — М. : Мир, 2001. — 430 с.
12. Кобозева, А.А. Анализ чувствительности сингулярных векторов матрицы изображения как основа стеганоалгоритма, устойчивого к сжатию / А.А.Кобозева, М.А.Мельник // Захист інформації. – 2013. - №2. – С.49-58.

Надійшла до редколегії 20.01.2013 р.

Рецензент: д.т.н., проф. Петров А.С.

Кобозева А.А.

МАТЕМАТИЧНА ОСНОВА МЕТОДУ ВИЯВЛЕННЯ КЛОНОВАНИХ ДІЛЯНОК ЗОБРАЖЕННЯ, ПІДДАНИХ КОРЕКЦІЇ КОЛЬОРУ

На основі загального підходу до аналізу стану й технології функціонування інформаційних систем, розробленого автором раніше, побудований математичний базис методу виявлення клонованих ділянок зображення, підданих корекції кольору.

Ключові слова: клонування, цифрове зображення, матриця, сингулярні числа, сингулярні вектори, корекція кольору, *n*-оптимальний вектор.

Kobozeva A.A.

MATHEMATICAL BASIS OF METHOD FOR DETECTING THE CLONED IMAGE AREAS AFTER COLOR CORRECTION

Mathematical basis of method for detecting the cloned image areas after color correction was developed on the basis of the general mathematical approach to analysis of state and functioning of information system that was previously developed by the author.

Keywords: clonning, digital image, matrix, singular value, singular vector, color correction, n-optimal vector.