Классификация и сравнительное исследование систем аутентификации JPEG-изображений, основанных на встраивании полухрупких водяных знаков

А.А. Егорова [2], В.А. Федосеев [1, 2]
[1] ИСОИ РАН – филиал ФНИЦ «Кристаллография и фотоника» РАН, 443001, Россия, Самарская область, г. Самара, ул. Молодогвардейская, д. 151,
[2] Самарский национальный исследовательский университет имени академика С.П. Королёва, 443086, Россия, Самарская область, г. Самара, Московское шоссе, д. 34

Аннотация

В статье рассматриваются системы полухрупких цифровых водяных знаков, предназначенные для защиты от несанкционированных изменений изображений, представленных в формате JPEG. Эти системы позволяют обнаруживать и определять местоположение изменений, а некоторые также приближённо восстанавливать исходное содержимое. Приводятся формальные схемы, описывающие процедуры встраивания цифровых водяных знаков и аутентификации. Рассматривается более десятка систем данного типа, предложенных с 2000 года, и предлагается их классификация по различным критериям. Представлены результаты экспериментальных исследований различных систем по оценке уровня искажений, возникающих при встраивании информации, а также погрешности при извлечении. Кроме того, исследована работоспособность систем в смысле обеспечения частичной стойкости к JPEG-сжатию.

Ключевые слова: цифровой водяной знак, ЦВЗ, аутентификация изображений, полухрупкий водяной знак, JPEG, QIM, НЗБ.

Цитирование: **Егорова, А.А.** Классификация и сравнительное исследование систем аутентификации JPEG-изображений, основанных на встраивании полухрупких водяных знаков / А.А. Егорова, В.А. Федосеев // Компьютерная оптика. – 2019. – Т. 43, № 3. – С. 419-433. – DOI: 10.18287/2412-6179-2019-43-3-419-433.

Введение

Стремительное развитие телекоммуникационных систем предоставляет колоссальные возможности по хранению и распространению изображений. В то же время сегодня существует большое число доступных средств обработки изображений, позволяющих с лёгкостью вносить в них изменения. Такие манипуляции могут быть преднамеренно вредоносными или могут непреднамеренно влиять на интерпретацию содержимого изображений, что может, в свою очередь, привести к нарушению авторских прав и нанести тем самым значительные убытки правообладателю. В связи с этим проблема аутентификации изображений весьма актуальна в современном мире. Одним из способов её решения является встраивание цифрового водяного знака (ЦВЗ) в защищаемое изображение (контейнер). ЦВЗ представляет собой малозаметную компоненту, наличие которой может свидетельствовать о подлинности изображения [1]. Совокупность методов и средств, образующих единое решение для встраивания ЦВЗ, мы будем называть системой встраивания ЦВЗ или кратко ЦВЗ-системой. Ключевой характеристикой ЦВЗ-систем является стойкость, под которой понимается возможность извлечения ЦВЗ из искажённого изображения [2, 3].

Различают хрупкие, стойкие (робастные), защищённые (безопасные) и полухрупкие ЦВЗ-системы [1, 4]. Хрупкие ЦВЗ разрушаются при любых модификациях изображений и в большинстве случаев используются для проверки целостности информации. Робастные ЦВЗ обладают стойкостью к непреднамеренным искажениям, характерным для конкретного сценария использования изображений. Класс защищённых ЦВЗ шире робастных. Такие водяные знаки проявляют стойкость ещё и по отношению к преднамеренным атакам со стороны злоумышленника. При этом предполагается, что злоумышленнику известен используемый способ встраивания ЦВЗ. Полухрупкие ЦВЗ используют в ситуации, когда некоторые изменения изображения считаются допустимыми (как правило, это изменения, не оказывающие существенного влияния на содержимое и не нарушающие его структуру). Они применяются в различных задачах криминалистики, медицины, в военной сфере и могут служить для аутентификации данных, получения маски изменений и восстановления искажённой информации. Полухрупкие ЦВЗ, как правило, незаметны глазу и разрушаются под воздействием изменений, не относящихся к «разрешённым», в список которых зачастую относят искажения, возникающие вследствие сжатия с потерями (до определённого уровня качества).

Одним из наиболее распространённых стандартов сжатия изображений с потерями является JPEG, поэтому для него разработано более двух десятков систем встраивания хрупких и полухрупких ЦВЗ. Наибольшее практическое применение среди них нашли системы, производящие внедрение водяного знака в коэффициенты блочного дискретного косинусного преобразования (ДКП) изображения [4 – 18]. Это обусловлено тем, что такие методы тесно связаны с процедурой сжатия и потому способны обеспе-

чить визуальную неразличимость ЦВЗ и одновременно с этим стойкость к JPEG. Особую ценность представляют системы, обеспечивающие стойкость при заданном пользователем диапазоне значений показателя качества JPEG-сжатия (*Quality Factor*, *QF* – целое число от 1 до 100) [6–11].

Несмотря на значительное число существующих полухрупких к JPEG систем встраивания ЦВЗ, в литературе практически отсутствуют работы, посвящённые их сравнению. Одним из немногих исключений является обзорная статья [19], однако она охватывает более широкий класс систем, что делает проведённый анализ менее детальным. Кроме того, в этой работе не проведён расчёт показателей исследуемых систем по единой методике и на едином наборе данных. По этой причине в настоящей работе представлен обзор наиболее значимых ЦВЗ-систем, полухрупких к JPEG, приведена классификация существующих решений по различным критериям и проведён сравнительный анализ их основных показателей.

Статья организована следующим образом. В первом параграфе представлены схемы аутентификации изображений на основе ЦВЗ-систем. Во втором параграфе описываются основные этапы JPEG-сжатия, а также основные структурные компоненты ЦВЗ-систем, предназначенных для JPEG-изображений. В третьем параграфе приводится обзор основных существующих на данный момент полухрупких систем, а также их классификация по различным параметрам. Четвёртый параграф посвящён экспериментальным исследованиям рассмотренных ЦВЗ-систем, проведённым для проверки их работоспособности и сравнения в единых условиях.

1. Аутентификация изображений с использованием ЦВЗ-систем

Схемы, представленные на рис. 1, 2, иллюстрируют процедуру аутентификации изображений с использованием ЦВЗ-систем. На этапе встраивания информации (рис. 1) сначала происходит формирование ЦВЗ, представляющего собой двоичную последовательность (W). Он может как полностью генерироваться на основе секретного ключа, так и формироваться на основе некоторых характеристик контейнера с возможным использованием ключа. В качестве примеров второго подхода можно привести использование хэш-функции от средней яркости изображения в окрестности встраивания или битов знака разностей между значениями пар пикселей, координаты которых определяются ключом. Затем полученный ЦВЗ встраивается в защищаемое изображение I (при этом также может быть использована часть секретного ключа), результатом чего является модифицированное изображение I^W, которое далее передаётся по открытому каналу и может быть изменено. Отметим, что встраивание на практике реализуется таким образом, чтобы сохранить неизменными характеристики, используемые при формировании ЦВЗ. То есть при генерации ЦВЗ по изображению I^W мы должны получить тот же ЦВЗ W.

Отметим, что пунктирными прямоугольниками на рис. 1, 2 обозначены этапы, а пунктирными стрелками – процессы передачи данных, которые могут отсутствовать в отдельных системах. В частности, в некоторых системах отсутствует этап извлечения характеристик контейнера. В этом случае ЦВЗ полностью определяется секретным ключом.

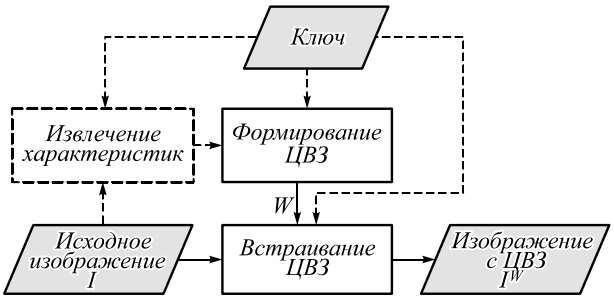


Рис. 1. Аутентификация изображений при помощи ЦВЗ-системы: встраивание информации

Пришедшее на аутентификацию изображение мы будем обозначать $\widetilde{I^W}$ (рис. 2). При отсутствии изменений оно эквивалентно I^W. На принимающей стороне сначала производится формирование ЦВЗ на основе ключа и изображения $\widetilde{I^W}$.

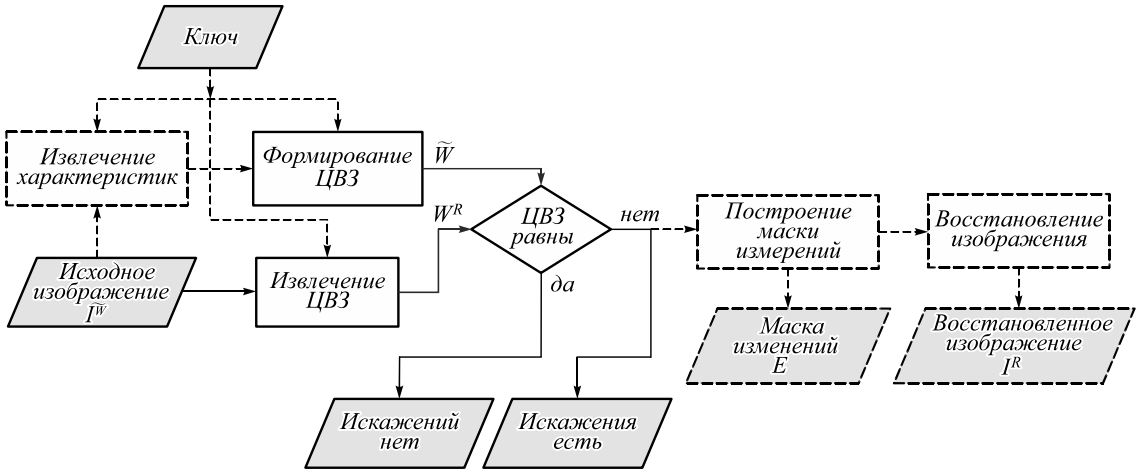


Рис. 2. Аутентификация изображений при помощи ЦВЗ-системы: извлечение информации

Результирующий ЦВЗ (обозначим его $\widetilde{W}$) совпадает с W, если ЦВЗ формируется только лишь на основе ключа или, если изменения в $\widetilde{I^W}$ относительно I^W не затронули характеристики, используемые при формировании ЦВЗ. Одновременно с этим производится извлечение ЦВЗ, встроенного в изображение $\widetilde{I^W}$ (обозначим его W^R). Совпадение $\widetilde{W}$ и W^R означает, что изображение не было подвергнуто «неразрешённым» искажениям. В противном случае изображение $\widetilde{I^W}$ расценивается как искажённое.

При обнаружении факта искажений далее может выполняться построение маски изменений E – бинарного изображения, значения «1» в котором указывают на то, что соответствующий пиксель изображения $\widetilde{I^W}$, вероятно, был изменён, в то время как нулевые значения говорят об обратном. Заключительным опциональным этапом, присутствующим лишь в некоторых системах, является этап восстановления искажённых фрагментов изображения $\widetilde{I^W}$ на основе извлечённого ЦВЗ W^R.

2. Построение ЦВЗ-систем, адаптированных к стандарту сжатия JPEG

2.1. Процедура JPEG-сжатия

Сжатие с потерями в формате JPEG включает следующие ключевые шаги [20]:

1. Исходное изображение I размера $N_1 \times N_2$ делится на непересекающиеся блоки I_i размера 8×8, где $i = 1, \ldots, N$ – номер блока, а $N = N_1 N_2 / 64$ – общее количество непересекающихся блоков.

2. Для каждого блока I_i, состоящего из 64 отсчётов $I_i(n_1, n_2)$, $0 \le n_1, n_2 \le 7$, вычисляется прямое ДКП. Полученные значения мы будем обозначать $B_i(m_1, m_2)$, $0 \le m_1, m_2 \le 7$. Коэффициенты, расположенные вблизи левого верхнего угла, характеризуют низкочастотную составляющую. Угловой элемент $B_i(0,0)$, значение которого равно средней яркости пикселей блока I_i, мы будем называть DC-отсчётом, все прочие – AC-отсчётами.

3. Производится квантование коэффициентов каждого блока $B_i(m_1, m_2)$ с использованием матрицы квантования Q_{QF} размера 8×8, соответствующей заданному пользователем значению параметра качества сжатия QF (от 1 до 100), по формуле:

$$D_i(m_1, m_2) = \text{round}\left(\frac{B(m_1, m_2)}{Q_{QF}(m_1, m_2)} \right). \quad (1)$$

Матрицы квантования для различных значений QF рассчитываются по следующей формуле:

$$Q_{QF} = \begin{cases} Q_{50} \cdot \text{round}\left(50 / QF\right), & QF < 50, \\ Q_{50} \cdot \left(2 - 0{,}02 QF\right), & 50 \le QF < 100. \end{cases} \quad (2)$$

Для $QF = 100$ матрица полностью состоит из единиц. Опорная матрица Q_{50} имеет вид:

$$Q_{50} = \begin{bmatrix} 16 & 11 & 10 & 16 & 24 & 40 & 51 & 61 \\ 12 & 12 & 14 & 19 & 26 & 58 & 60 & 55 \\ 14 & 13 & 16 & 24 & 40 & 57 & 69 & 56 \\ 14 & 17 & 22 & 29 & 51 & 87 & 80 & 62 \\ 18 & 22 & 37 & 56 & 68 & 109 & 103 & 77 \\ 24 & 35 & 55 & 64 & 81 & 104 & 113 & 92 \\ 49 & 64 & 78 & 87 & 103 & 121 & 120 & 101 \\ 72 & 92 & 95 & 98 & 112 & 100 & 103 & 99 \end{bmatrix}. \quad (3)$$

Таким образом, согласно (2), уменьшение QF приводит к росту значений в матрице Q_{QF}, что, в свою очередь, приводит к возрастанию числа нулей среди $D_i(m_1, m_2)$, а следовательно, и к уменьшению размера архива.

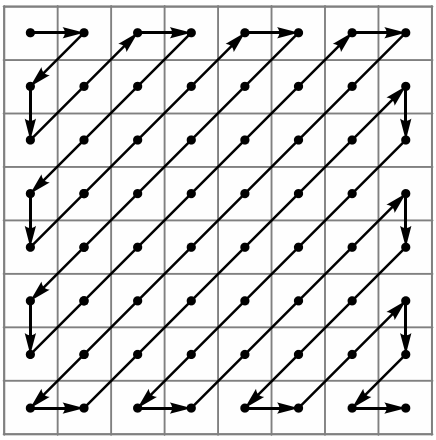


Рис. 3. Зигзагообразная развёртка спектра ДКП

4. Обход значений $D_i(m_1, m_2)$ в зигзагообразном порядке, как показано на рис. 3, и последующее статистическое кодирование этих значений. Далее в статье мы будем для удобства обозначать ДКП-коэффициенты $D_i(j)$, где $j = 1 \ldots 64$ – номер коэффициента при зигзагообразном обходе.

Схема рассмотренной процедуры сжатия JPEG показана на рис. 4.

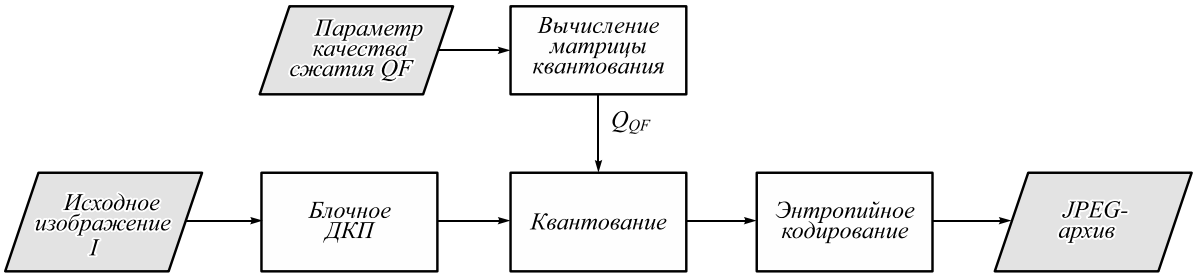


Рис. 4. Схема алгоритма сжатия JPEG

2.2. Основные особенности ЦВЗ-систем, полухрупких к JPEG-сжатию

Большинство существующих ЦВЗ-систем, полухрупких к JPEG, базируется на двух свойствах:

– отношение между коэффициентами, находящимися в одинаковых позициях в двух разных блоках, до и после сжатия не изменяется;

– квантованное с некоторым (первоначальным) шагом значение можно восстановить после последующего квантования с меньшим шагом, если к искажённому числу повторно применить процедуру квантования с первоначальным шагом:

$$\mathrm{quan}(x,\Delta_1)=\mathrm{quan}(\mathrm{quan}(\mathrm{quan}(x,\Delta_1),\Delta_2),\Delta_1), \qquad (4)$$

где

$\Delta_1 \geq \Delta_2$ и $\mathrm{quan}(x,\Delta)=\Delta\cdot\mathrm{round}(x/\Delta)$.

Первое свойство обычно используется при генерации зависящего от характеристик контейнера ЦВЗ [6, 13, 17] и в системах, способных восстанавливать искажённую информацию [5, 10]. Второе лежит в основе систем, которые позволяют задавать минимально допустимый уровень сжатия *QF* [5, 6, 7, 9, 17].

Во всех системах рассматриваемого класса работа с изображениями и непосредственно встраивание производится на уровне блоков коэффициентов блочного ДКП. При этом в общем случае исходное изображение и изображение с ЦВЗ представляются в несжатом виде, как показано на рис. 5. Поэтому встраиванию предшествует этап расчёта ДКП, а вслед за встраиванием выполняется обратное ДКП. Таким образом, результатом применения данной схемы является изображение I^W, подготовленное для последующего JPEG-сжатия в том смысле, что встроенный ЦВЗ обладает частичной стойкостью к JPEG в заданном диапазоне значений *QF*.

Этап формирования ЦВЗ на рис. 5 опущен, поскольку в нём нет никакой дополнительной конкретизации по сравнению с рис. 1. Наличие двух альтернативных вариантов на рис. 5 показывает, что встраивание может происходить как в уже квантованные ДКП-коэффициенты, так и непосредственно во время самой процедуры квантования (например, когда модификация информации производится при помощи алгоритмов семейства QIM, речь о которых пойдёт в параграфе 3).

Состав операций, выполняемых при извлечении информации, очевиден из рис. 2 и 5, поэтому позволим себе опустить схему данного этапа.

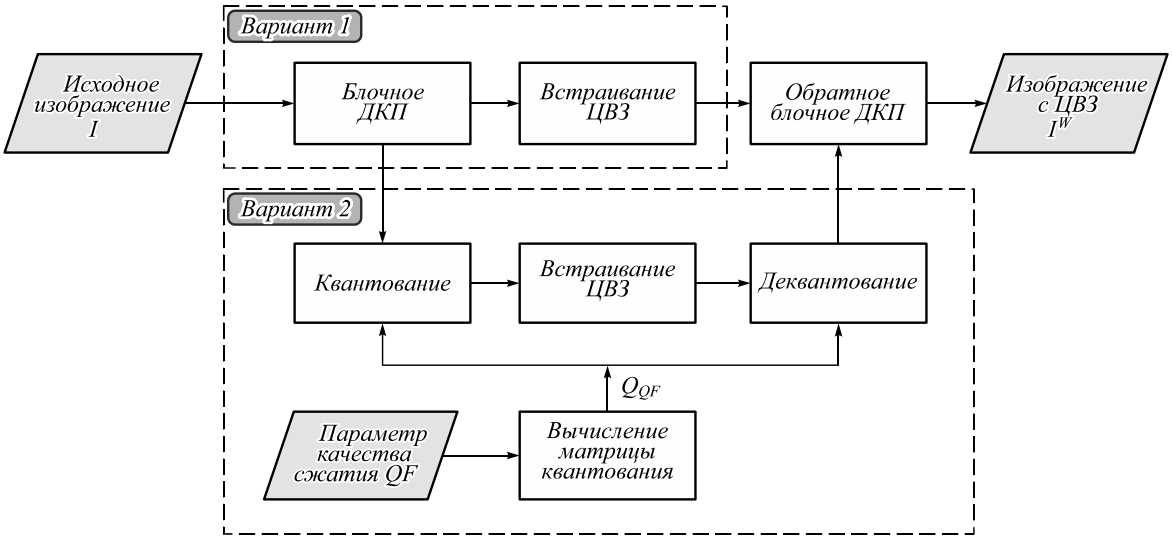


Рис. 5. Схема встраивания ЦВЗ в системах, полухрупких к JPEG-сжатию

3. Классификация систем встраивания ЦВЗ, полухрупких к JPEG-сжатию

3.1. По используемому методу встраивания

Одной из ключевых характеристик полухрупких к JPEG ЦВЗ-систем является метод встраивания ЦВЗ, то есть способ модификации спектральных компонент.

В существующих на сегодняшний день решениях наибольшее распространение получили метод изменения наименее значимых бит (НЗБ-встраивание) [4] и методы на основе управляемого переквантования (QIM, Quantization Index Modulation) [21]. Однако помимо них используются и другие подходы, в частности, базирующиеся на изменении позиции последнего ненулевого элемента (LNZ, Last Non Zero) [12], на использовании табличных преобразований (Mapping Table) [13] и на методе расширения спектра (Spread Spectrum Watermarking) [22].

НЗБ-встраивание в полухрупких к JPEG методах производится после выполнения процедуры квантования путём замены одного или нескольких квантованных ДКП коэффициентов изображения битами водяного знака. Алгоритм прост в реализации, не требует больших вычислительных ресурсов и позволяет скрыть в изображении достаточно большой объём информации, что обусловливает его активное использование [6, 7]. Однако в то же время НЗБ-встраивание в спектральной области может сопровождаться значительными искажениями контейнера, а факт использования этого метода может быть обнаружен.

В отличие от НЗБ-встраивания, QIM (подробно рассмотренный в [3, 23, 24]) осуществляет одновре-

менно квантование данных и встраивание водяного знака, производя модуляцию значений коэффициентов ДКП в соответствии со значениями битов встраиваемого водяного знака. QIM более устойчив к атакам по сравнению с НЗБ-встраиванием и потенциально может приводить к меньшим погрешностям. Особенностью QIM является тот факт, что под этим названием подразумевается не один метод, а целое семейство методов, отличающихся используемыми функциями квантования. В известных системах полухрупких ЦВЗ также используются различные версии QIM [9 – 11].

Альтернативным способом внедрения информации в JPEG-изображения является смена чётности позиции последнего ненулевого элемента (LNZ) блока квантованных ДКП коэффициентов. Такая операция позволяет встроить один бит водяного знака в каждый блок. В полухрупкой системе Fallahpour & Megias, 2016 [12], стойкой к JPEG-сжатию и аддитивному шуму, смена чётности LNZ-компоненты производится в тех блоках, где её позиция превышает заданное пороговое значение, остальные блоки пропускаются. Такой подход прост в реализации, не требует больших затрат по времени, что позволяет применять его в системах реального времени.

Другой альтернативный способ встраивания водяных знаков основан на использовании табличного отображения. В полухрупкой системе Mursi et al., 2009 [13] такая таблица представляет собой соответствие номеров коэффициентов, расположенных в случайном порядке, последовательности из нулей и единиц. Встраивание происходит следующим образом: допустим, в некоторый коэффициент нужно встроить бит со значением «1». Тогда если в таблице ему соответствует «1», то его значение не изменяется, в противном случае его значение заменяется на значение ближайшего по соседству коэффициента, которому в таблице соответствует «1».

Наконец, метод встраивания информации с расширением спектра состоит в распределении небольшого числа бит ЦВЗ среди большого числа компонент изображения-контейнера с использованием функции, зависящей от секретного ключа [22]. В системах, использующих такой подход [14, 15], встраивание производится путём внесения небольшой аддитивной или мультипликативной шумоподобной последовательности в большое число коэффициентов. Для аутентификации учитывается поблочная корреляция модифицированных коэффициентов с заданной шумоподобной последовательностью. Высокое значение корреляции свидетельствует об аутентичности рассматриваемого блока. Системы этой группы показывают хорошие результаты на однородных областях изображений, кроме того, их важным преимуществом является повышенная стойкость ЦВЗ к искажениям за счёт распределения встраиваемой информации среди множества коэффициентов ДКП. Однако эффективность таких систем снижается на текстурированных фрагментах. Кроме того, они не позволяют регулировать минимально допустимый уровень сжатия *QF*, при котором сохраняется их стойкость к JPEG.

В табл. 1 представлены основные существующие ЦВЗ-системы, для каждой из которых во втором столбце приводится метод изменения ДКП-коэффициентов.

Табл. 1. Основные свойства полухрупких ЦВЗ-систем: метод встраивания ЦВЗ, частотная область и число изменяемых коэффициентов

ЦВЗ-система	**Метод изменения коэффициентов**	**Область частот, в которую производится встраивание ЦВЗ**	**Число изменяемых коэффициентов блока**
Lin & Chang, 2000 [6]	НЗБ	НЧ: $j = 7, ..., 9$ для обнаружения изменений и $j = 10, ..., 15$ для восстановления	Разное. Основной вариант: 3 для обнаружения изменений и 6 для восстановления
Lin & Chang, версия Cox et al. [1]	НЗБ	ВЧ: $j = 37, ..., 64$	4
Ho & Li, 2004 [7]	НЗБ	Разная (наибольшие ненулевые коэффициенты блока)	4
Huang, 2013 [8]	НЗБ	НЧ: $j = \{5;8;9;13\}$	4
Ye et al., 2003 [9]	QIM	НЧ, $j = \{5;8;9;13\}$СЧ: $j = 2, ..., 20$	4
Preda & Vizireanu, 2015 [10]	QIM	НЧ: $j = 2, ..., 2N_w$	Разное
Wang et al., 2011 [11]	QIM	НЧ (для обнаружения изменений), НЧ, СЧ, ВЧ (для восстановления)	6 для обнаружения изменений 4 для восстановления
Fallahpour & Megias, 2016 [12]	Смена чётности позиции LNZ	Разная (зависит от установленного значения порога)	4
Mursi et al., 2009 [13]	Табличное отображение	НЧ: $j = 2, ..., 6$	5
Lin et al., 2000 [14]	Расширение спектра	НЧ, СЧ: $j = 2, ..., 36$	35, при этом встраивается 1 бит на блок
Al-Mualla, 2007 [15]	Расширение спектра	НЧ, СЧ: $j = 2, ..., 36$	35, при этом встраивается 1 бит на блок

3.2. По местоположению и количеству изменяемых коэффициентов

Большинство полухрупких к JPEG систем встраивают водяной знак в низкочастотные (НЧ) коэффициенты (за исключением DC-отсчёта) или среднечастотные (СЧ) [6, 8 – 11]. Потенциальный недостаток такого подхода заключается в том, что низкие частоты значительно информативнее высокочастотных,

поэтому их изменение может привести к заметному ухудшению качества защищаемого изображения. С другой стороны, для НЧ-коэффициентов ниже погрешность квантования, как видно из матрицы (2). Известны работы [1, 7, 11], в которых внедрение водяного знака производится в несколько высокочастотных коэффициентов (ВЧ) и при этом, согласно утверждениям авторов, обеспечивается высокое качество извлечения ЦВЗ. Также встраивание в высокочастотные коэффициенты нередко происходит и в системе [12] на основе LNZ-изменения, так как последний ненулевой коэффициент блока может располагаться в области высоких частот.

В зависимости от назначения ЦВЗ-системы или её особенностей использования объём встраиваемого ЦВЗ может меняться. В подавляющем большинстве систем в каждый блок размера 8×8 встраивается одинаковое число бит ЦВЗ. Однако в некоторых системах, в частности, в [12], в зависимости от распределения значений коэффициентов ДКП, часть блоков может быть пропущена при встраивании.

Также следует отметить, что в большинстве систем число изменяемых коэффициентов блока соответствует числу бит, которое приходится на этот блок (будем обозначать его N_W). Исключение составляют системы, основанные на методе расширения спектра [14, 15], которые изменяют большое число компонент спектра ДКП для того, чтобы встроить 1 бит информации.

Если система используется только для обнаружения факта изменений или для построения маски изменений, число бит, приходящихся на один блок, колеблется в диапазоне от 1 до 6. Если в системе реализуется процедура частичного восстановления изображения после обнаружения несанкционированных изменений, то, как правило, дополнительно встраиваются ещё не менее 4 бит информации [6, 11]. Таким образом, суммарное количество может превысить 10 бит на блок.

В столбцах 3 и 4 табл. 1 представлены данные по основным системам в части местоположения и количества изменяемых коэффициентов ДКП.

3.3. По способу формирования ЦВЗ

Ещё одним важным свойством, по которому могут быть классифицированы рассматриваемые ЦВЗ-системы, является способ формирования водяного знака (рис. 1). В простейшем случае ЦВЗ представляет собой псевдослучайную последовательность и генерируется на основе секретного ключа [10, 12, 14, 15]. Однако в ряде систем для повышения стойкости ЦВЗ может формироваться на основе каких-либо характеристик (инвариант) защищаемого изображения. Кроме того, нередки случаи, когда ЦВЗ представляет собой конкатенацию псевдослучайной последовательности и вычисляемых характеристик защищаемого изображения. Рассмотрим некоторые примеры.

В системах [6, 13] при генерации ЦВЗ принимается во внимание одно из отмеченных в подпараграфе 2.1 свойств JPEG-сжатия, согласно которому отношение между квантованными коэффициентами, находящимися в одинаковых позициях в двух разных блоках, до и после сжатия не изменяется. Это обусловлено тем, что во время квантования по формуле (1) каждый блок ДКП коэффициентов поэлементно делится на одну и ту же матрицу квантования Q_{QF}, которая соответствует заданному значению параметра качества QF. В результате этого коэффициенты разных блоков, находящиеся в одинаковых позициях, делятся на один и тот же элемент матрицы Q_{QF}, таким образом, отношение между ними сохраняется, а его изменение свидетельствует о том, что изображение подвергалось модификациям.

В Ho & Li, 2004 [7] встраиваемая информация зависит от значений характеристик, вычисленных по значениям коэффициентов, взятых из соседних блоков, следующим образом: на основе случайной последовательности, генерируемой с использованием секретного ключа, определяется группа коэффициентов в рассматриваемом блоке и в восьми соседних с ним. Затем для каждой из четырёх групп вычисляются две характеристики: одна зависит от знака коэффициентов группы, вторая – от их величины.

В Mursi et al., 2009 [13] в состав ЦВЗ входит секретная последовательность и разница между коэффициентами соседних блоков, что делает ЦВЗ устойчивым ещё и к искажению типа «копирование-вставка» [25, 26].

Другой характеристикой контейнера, которая может использоваться при генерации ЦВЗ, является значение самого информативного коэффициента блока – DC. Такой подход применён в системе Huang, 2013 [8], где биты ЦВЗ представляют собой зашифрованную комбинацию бит DC-компонент блоков контейнера. В Ye et al., 2003 [9] в процессе формирования ЦВЗ всегда используется DC-коэффициент текущего блока, а также три коэффициента, определяемые случайной последовательностью. Все они являются параметрами хэш-функции, которая подписывается секретным ключом правообладателя.

В системе Preda & Vizireanu, 2015 [10] ЦВЗ представляет собой значение хэш-функции, параметрами которой являются координаты блока и секретная последовательность. Это также делает ЦВЗ устойчивым к искажению типа «копирование-вставка».

В Wang et al., 2011 [11] для генерации битов ЦВЗ, отвечающих за восстановление изображения, встраиваются дополнительные 4 бита информации, значения которых определены как нормализованные средние значения подблоков контейнера размером 4×4.

Сводные сведения по способам формирования ЦВЗ для различных систем представлены в табл. 2.

3.4. По назначению

Наконец, ещё одной важнейшей характеристикой ЦВЗ-систем, используемых для аутентификации изображений, является их назначение, то есть конкретная решаемая задача аутентификации. Возможны следующие варианты (см. рис. 2) в порядке от простого к сложному:

1) выявление факта искажений без детализации их местоположения;

2) обнаружение и локализация искажений (то есть построение маски изменений);

3) построение маски изменений и приближённое восстановление искажённых областей.

Большинство ЦВЗ-систем рассматриваемого класса решают вторую задачу, что отмечено в табл. 2, при этом в подавляющем большинстве случаев обнаружение искажений происходит с точностью до блока размером 8×8. Использование встраиваемой последовательности большой длины снижает вероятность пропуска искажённых блоков, в то время как использование специальных способов формирования ЦВЗ на основе характеристик контейнера, рассмотренных в подпараграфе 3.3, повышает точность построения маски изменений за счёт обеспечения стойкости к отдельным типам искажений.

Табл. 2. Основные свойства полухрупких ЦВЗ-систем: принцип формирования ЦВЗ и назначение системы

ЦВЗ-система	**Используются ли характеристики изображения при формировании ЦВЗ**	**Назначение**
Lin & Chang, 2000 [6]	Да, кодируются знаки разниц между первыми шестью AC коэффициентами текущего блока с некоторым другим	Обнаружение изменений, восстановление
Lin & Chang, версия Cox et al. [1]	Нет	Обнаружение изменений
Ho & Li, 2000 [7]	Да (зависимость от 9 соседних блоков)	Обнаружение изменений
Huang, 2013 [8]	Да, в генерации ЦВЗ участвуют 5 бит DC коэффициента блока	Обнаружение изменений
Ye et al., 2003 [9]	Да, в генерации ЦВЗ участвует DC и три AC коэффициента из первых двадцати	Обнаружение изменений
Preda & Vizireanu, 2015 [10]	ЦВЗ зависит от координат блока контейнера и случайной последовательности, генерируемой при помощи секретного ключа	Обнаружение изменений
Wang et al., 2011 [11]	ЦВЗ для обнаружения изменений – псевдослучайная последовательность. ЦВЗ для восстановления – нормализованные средние значения подблоков	Обнаружение изменений, восстановление
Fallahpour & Megias, 2016 [12]	Нет	Обнаружение изменений
Mursi et al., 2009 [13]	Да, зависит от разницы между коэффициентами двух соседних блоков, но включает также псевдослучайную последовательность	Обнаружение изменений
Lin et al., 2000 [14]	Нет	Обнаружение изменений
Al-Mualla, 2007 [15]	Нет	Обнаружение изменений

Решение третьей задачи предусмотрено в системах Lin & Chang, 2000 [6] и Wang et al., 2011 [11]. В первой из них биты ЦВЗ, используемые для восстановления, содержат JPEG-архив исходного изображения, уменьшенного вдвое по каждому из измерений, полученный с использованием низкого значения *QF* (в качестве примера приводится значение 25). Таким образом, 4 блока 8×8 защищённого изображения в совокупности задают приближённую копию одного блока 16×16 исходного изображения, при этом, согласно данным табл. 1, для этого используется 24 бита. Распределение фрагментов архива среди блоков определяется ключом.

В системе Wang et al., 2011 [11] в каждый блок 8×8 встраиваются по 4 бита, используемых для приближённого восстановления четырёх коэффициентов (DC и трёх первых AC-коэффициентов) другого блока. Соответствие между блоками устанавливается при помощи ключа. Восстановление ДКП-коэффициентов осуществляется при помощи линейной регрессии, коэффициенты которой определяются заранее на этапе настройки системы.

4. Экспериментальные исследования

4.1. Основные показатели рассмотренных систем

К числу основных объективных характеристик полухрупких ЦВЗ-систем для JPEG-изображений, которые приводят большинство авторов оригинальных статей, относятся минимально допустимое значение показателя *QF*, к которому можно обеспечить стойкость, а также качество защищённого изображения, как правило, измеряемое при помощи показателя *PSNR* (Peak Signal-To-Noise Measure) [1, 6]. В табл. 3 отражены данные по этим показателям, взятые из оригинальных работ. Как видно из таблицы, они недостаточно полны, чтобы сделать вывод о преимуществе тех или иных систем.

В отношении минимального значения *QF*, это объясняется тем, что ввиду упомянутого в подпараграфе 2.2 свойства квантования, использующие его системы теоретически вовсе не должны иметь ограничений на значения *QF*. Однако погрешности при квантовании защищённого изображения в пространственной области и рамки допустимого качества изображений, определяемые практическими аспектами, приводят авторов к введению ограничений. В большинстве случаев эти ограничения вводятся умозрительно и не продиктованы результатами конкретных исследований.

Значение показателя *PSNR* между исходным и защищённым изображениями главным образом зависит от следующих факторов:

1) числа встраиваемых бит ЦВЗ;
2) местоположения изменяемых коэффициентов;
3) используемого метода встраивания;
4) значения *QF*;
5) используемого набора данных.

Табл. 3. Сравнительные показатели полухрупких к JPEG ЦВЗ-систем (по материалам оригинальных статей)

ЦВЗ-система	**Минимальные значения *QF*, к которым можно обеспечить стойкость системы**	***PSNR*, дБ**
Lin & Chang, 2000 [6]	Не указано (исследованы $QF \geq 50$)	43,02 (для $QF = 50$) для обнаружения изменений 32,75 (для $QF = 50$) для восстановления
Lin & Chang, версия Cox et al. [1]	Не указано	Нет данных
Ho & Li, 2000 [7]	50	32,7 (для $QF = 50$)
Huang, 2013 [8]	50	41,2 (для $QF = 50$)
Ye et al., 2003 [9]	Не указано (исследованы $QF \geq 75$)	Около 40 (для $QF = 75$)
Preda & Vizireanu, 2015 [10]	Не указано (исследованы $QF \geq 50$)	44,63 (для $QF = 50$)
Wang et al., 2011 [11]	65 (для восстановления)	37,61 (для $QF = 75$)
Fallahpour & Megias, 2016 [12]	Без ограничений	41
Mursi et al., 2009 [13]	80	Нет данных
Lin et al., 2000 [14]	Нет связи с QF	36,67
Al-Mualla, 2007 [15]	Нет связи с QF	35,53

Факторы $1-3$ являются неотъемлемыми свойствами системы, в то время как 4 и 5 характеризуют условия проведения вычислительного эксперимента. Практический интерес представляет исследование влияния факторов $1-3$ на значение *PSNR*, но единственный способ отделить их от $4-5$ – это независимые эксперименты в унифицированных условиях.

В последующих подпараграфах приводятся результаты проведённых нами сравнительных экспериментальных исследований некоторых систем встраивания ЦВЗ из числа рассмотренных в параграфе 3. При проведении исследований решались следующие задачи:

- оценка уровня искажений, возникающих при встраивании защитного ЦВЗ;
- проверка работоспособности в смысле обеспечения частичной стойкости ЦВЗ к JPEG-сжатию.

4.2. Системы, отобранные для исследования

При отборе систем, участвовавших в экспериментальных исследованиях, мы руководствовались главным образом необходимостью обеспечить полный охват в части методов модификации коэффициентов ДКП, рассмотренных в подпараграфе 3.1. Среди систем одной группы приоритет отдавался системам, получившим большее распространение, а также более простым в реализации. Полный перечень отобранных систем можно увидеть в табл. 4. Рассмотрим кратко их основные отличительные особенности, дополняющие их свойства, отражённые в параграфе 3 и табл. 1, 2.

В системе Lin & Chang, 2000 [6] используется изменённая матрица квантования $Q'_{QF} = Q_{QF} + 1$ (то есть все значения увеличиваются на единицу). По-видимому, это сделано для повышения стойкости. Встраивание информации в квантованные коэффициенты производится по стандартной для НЗБ-методов формуле:

$$D_i^W\left(j_k\right) = 2\left\lfloor D_i\left(j_k\right)/2 \right\rfloor + W_{i,k}, \tag{5}$$

где j_k, $k = 0, ..., N_W - 1$ – позиции изменяемых коэффициентов ДКП в зигзагообразной развёртке, а $W_{j,k}$ – это k-й бит ЦВЗ, встраиваемый в i-й блок.

Кроме системы Lin & Chang, 2000, мы анализировали также её версию, приведённую в книге [1] (обозначена Lin & Chang, версия Cox et al., 2000). Уже из табл. 1, 2 видно, что она имеет значительные отличия от оригинала. В дополнение к этому изменены и процедуры встраивания и извлечения информации. 28 высокочастотных коэффициентов ($j = 37, ..., 64$) при встраивании разделяются на $N_W = 4$ группы по 8 коэффициентов в соответствии с секретным ключом. Сумма по модулю 2 наименее значимых бит восьми коэффициентов в каждой группе должна равняться встраиваемому биту информации. Если это условие соблюдается, то изменений не происходит. В противном случае инвертируется младший бит одного из коэффициентов в группе.

В системе Huang, 2013 [8] при встраивании информации вместо (5) используется формула:

$$D_i^W\left(j_k\right) = 2\left\lfloor D_i\left(j_k\right)/2 \right\rfloor + \alpha W_{i,k}, \tag{6}$$

где $0{,}5 < \alpha < 1$ – параметр, характеризующий интенсивность ЦВЗ. По умолчанию в экспериментах мы использовали значение $\alpha = 0{,}6$, предложенное в оригинальной работе.

В системе Preda & Vizireanu, 2015 [10] позиции j_k выбираются согласно ключу из промежутка $[2; 2N_W]$. В системе используются следующие формулы встраивания и извлечения информации:

$$B_i^W\left(j_k\right) = \operatorname{round}\left(\frac{B_i\left(j_k\right)}{2Q_{QF}\left(j_k\right)} - W_{i,k}\right)2Q_{QF}\left(j_k\right) + W_{i,k}Q_{QF}\left(j_k\right), \tag{7}$$

$$W_{i,k}{}^R = \operatorname{round}\left(B_i^W\left(j_k\right)/Q_{QF}\left(j_k\right)\right) (\operatorname{mod} 2). \tag{8}$$

Для сопоставления различных версий QIM в рамках работы мы исследовали также простую модификацию системы Preda & Vizireanu, 2015, которую обозначили Sign-QIM. Она отличается главным образом тем, что знак компоненты водяного знака зависит от того, в какую сторону происходит округление модифицируемого ДКП-коэффициента функцией *round*. Благодаря этому ошибка в коэффициенте j_k, вызванная встраиванием информации, гарантированно не превышает $Q_{QF}(j_k)$:

$$B_i^W\left(j_k\right) = Br_i\left(j_k\right) + S_i\left(j_k\right) \cdot W_{i,k} \cdot Q_{QF}\left(j_k\right), \tag{9}$$

где

$$Br_i\left(j_k\right)=\text{round}\left(\frac{B_i\left(j_k\right)}{2Q_{QF}\left(j_k\right)}\right)2Q_{QF}\left(j_k\right), \quad (10)$$

$$S_i\left(j_k\right)=\begin{cases}1, & B_i\left(j_k\right)\geq Br_i\left(j_k\right),\\ -1, & \textit{иначе}.\end{cases} \quad (11)$$

Ещё одна система на основе метода QIM – Wang et al., 2011 [11] – отличается от рассмотренных тем, что не изменяет значения коэффициентов ДКП, отстоящих не более чем на βQ_{QF}, $0<\beta<0{,}5$, от квантованного значения. Эта мера позволяет снизить искажения при встраивании информации. Точную формулу встраивания мы не будем приводить по причине её громоздкости. В экспериментах использовалось значение $\beta=0{,}25$.

Три другие системы: Fallahpour & Megias, 2016 [12], Mursi et al., 2009 [13] и Lin et al., 2000 [14] – уже были подробно рассмотрены в подпараграфе 3.1.

4.3. Порядок проведения экспериментов

При проведении экспериментов в качестве меры точности извлечения использовался показатель *BER* (Bit Error Rate), равный отношению числа неверно извлечённых бит к длине ЦВЗ. В качестве меры качества изображения со встроенным ЦВЗ использовался показатель *PSNR*.

Эксперименты проводились на 10 изображениях размерами 512×512 из набора данных USC-SIPI Image Database [27] (“Lenna”, “Peppers”, “Baboon”, “Bridge” и др.); результаты расчёта показателей *BER*, *PSNR* усреднялись.

При встраивании информации использовались два варианта значений минимального *QF*, к которому требовалось обеспечить стойкость ЦВЗ: 50 и 70. При исследовании всех систем водяной знак генерировался на основе ключа и не зависел от защищаемого изображения. Такое упрощение является оправданным для двух задач, решавшихся в ходе экспериментального исследования (сформулированных в параграфе 4.1). Если бы мы дополнительно исследовали, например, точность построения маски изменений, то это потребовало бы реализации адаптивных процедур формирования ЦВЗ, описанных в подпараграфе 3.3.

4.4. Оценка искажений, вызванных встраиванием защитного ЦВЗ

Для оценки уровня искажений, возникающих в результате встраивания защитного ЦВЗ, мы исследовали системы в трёх различных режимах:

1) число встраиваемых бит на блок N_W и позиции изменяемых коэффициентов j_k, $k=0, ..., N_W-1$ заданы согласно табл. 1 (авторские значения);
2) N_W одинаковое для всех систем (рассматривались значения 1 и 4), а j_k заданы согласно табл. 1;
3) N_W и j_k одинаковые для всех систем.

Таким образом, в первом режиме мы анализировали оригинальные системы, во втором – сочетание метода встраивания ЦВЗ и области встраивания, в третьем – исключительно метод встраивания. Если оригинальное значение N_W было меньше 4, то лишние позиции j_k дополнялись наиболее логичным образом. В третьем режиме в качестве j_k использовались 4 коэффициента с наименьшими значениями $Q_{QF}(j_k)$ согласно (2). Для систем Lin & Chang (версия Cox et al.) и Lin et al., 2000 такая установка противоречила принципу встраивания информации, поэтому они в этой части исследования не участвовали.

Результаты экспериментов отражены в табл. 4 – 6. Для системы Fallahpour & Megias, 2016 во всех таблицах в скобках указана доля ЦВЗ, которую удалось встроить (в этой системе предусмотрена возможность пропуска блоков). Помимо *PSNR,* также указаны значения *BER* при извлечении информации из неискажённого изображения.

С точки зрения искажений наилучший результат в первом режиме показала система Sign-QIM – модификация Preda & Vizireanu, 2015. Однако после выравнивания N_W лучшие результаты показала система Wang et al., 2011. Следует отметить, что и по ошибке извлечения также лидируют эти две системы.

Если сравнивать различные методы встраивания, можно сделать вывод о безусловном преимуществе систем на базе QIM. Далее с небольшим отставанием следует система, использующая табличное отображение, и системы на базе НЗБ-встраивания. Встраивание на основе изменения позиции LNZ-элемента приводит к существенным искажениям даже при сниженном объёме встраиваемой информации.

Система Lin et al., 2000 на основе расширения спектра сильнее всего искажает изображение. Помимо этого, она обладает высокой ошибкой извлечения и не связана со значением *QF*, поэтому применение данной системы как полухрупкой по отношению к JPEG-сжатию представляется нецелесообразным.

Для большинства остальных систем ошибка извлечения довольно низкая, за исключением системы Huang, 2013, в которой ошибка может быть существенной в некоторых случаях. По-видимому, это обусловлено наличием параметра α в формуле (5).

На рис. 6, 7 показаны увеличенные фрагменты изображений с ЦВЗ, встроенным при помощи различных систем (с использованием оригинальных значений N_W). Эти рисунки иллюстрируют характер искажений, возникающих в результате встраивания ЦВЗ.

4.5. Исследование стойкости ЦВЗ при JPEG-сжатии

В рамках второго эксперимента мы исследовали влияние показателя качества сжатия *QF** защищённого изображения в формате JPEG на точность последующего восстановления ЦВЗ. Если встраивание ЦВЗ произведено с показателем *QF*, то в идеальном случае полухрупкая система должна себя вести следующим образом: при $QF^*\geq QF$ значение ошибки извлечения *BER* должно быть близким к нулю, а при $QF^*<QF$ оно должно быть близким к 0,5, что соответствует вероятности случайного угадывания каждого бита ЦВЗ в отдельности. На практике могут встречаться немногочисленные ненулевые значения *BER* при $QF^*\geq QF$, объясняемые округлением значений пикселей после

обратного ДКП, и, как следствие, искажением спектральных компонент при извлечении информации. Если *BER* достигает уровня 0,01 и выше, как правило, это свидетельствует уже о несовершенстве ЦВЗ-системы. В области $QF^* < QF$ неизбежно присутствует переходная фаза, когда график *BER* постепенно снижается с 0,5 до 0. Этот промежуток должен быть как можно короче.

Табл. 4. Исследование качества изображений с ЦВЗ и точности извлечения при отсутствии искажений при встраивании с параметром QF = 50

ЦВЗ-система	**Оригинальные значения N_W**			$N_W = 1$		$N_W = 4$	
	N_W	***PSNR, дБ***	***BER***	***PSNR, дБ***	***BER***	***PSNR, дБ***	***BER***
Lin & Chang, 2000 [6]	3	37,75	1,9E-04	41,59	1,5E-04	36,57	2,1E-04
Huang, 2013 [8]	4	39,06	1,2E-03	47,34	1,8E-01	39,06	1,2E-03
Lin & Chang, версия Cox et al. [1]	4	27,19	2,8E-04	34,81	2,7E-04	27,19	2,8E-04
Preda & Vizireanu, 2015 [10]	3	39,62	4,9E-05	45,26	7,3E-05	38,40	9,2E-05
Sign-QIM	3	**42,99**	**0**	48,93	**0**	41,77	**0**
Wang et al., 2011 [11]	6	38,83	4,1E-06	**48,94**	**0**	**42,06**	**0**
Fallahpour & Megias, 2016 [12]	1	35,39 (96%)	3,2E-03	35,39 (96%)	3,2E-03	35,39 (24%)	3,2E-03
Mursi et al., 2009 [13]	5	37,13	2,1E-04	45,82	1,2E-04	36,90	3,4E-04
Lin et al., 2000 [14]	1	33,66	5,3E-02	33,66	5,3E-02	–	–

Табл. 5. Исследование качества изображений с ЦВЗ и точности извлечения при отсутствии искажений при встраивании с параметром QF = 70

ЦВЗ-система	**Оригинальные значения N_W**			$N_W = 1$		$N_W = 4$	
	N_W	***PSNR, дБ***	***BER***	***PSNR, дБ***	***BER***	***PSNR, дБ***	***BER***
Lin & Chang, 2000 [6]	3	41,38	1,8E-04	44,94	2,0E-04	40,28	1,8E-04
Huang, 2013 [8]	4	43,00	2,7E-03	51,23	2,1E-01	43,00	2,7E-03
Lin & Chang, версия Cox et al. [1]	4	31,72	1,2E-03	39,48	1,9E-03	31,72	1,2E-03
Preda & Vizireanu, 2015 [10]	3	43,77	7,3E-05	49,71	2,7E-04	42,66	8,5E-05
Sign-QIM	3	**47,33**	**0**	52,77	**0**	46,26	6,1E-06
Wang et al., 2011 [11]	6	43,27	4,1E-06	**52,80**	**0**	**46,40**	**0**
Fallahpour & Megias, 2016 [12]	1	37,60 (99 %)	5,1E-03	37,60 (99%)	5,1E-03	37,60 (25 %)	5,1E-03
Mursi et al., 2009 [13]	5	42,10	1,1E-04	49,50	1,2E-04	41,36	3,6E-04
Lin et al., 2000 [14]	1	33,66	5,3E-02	33,66	5,3E-02	–	–

Табл. 6. Исследование качества изображений с ЦВЗ и точности извлечения при отсутствии искажений при стандартизации числа бит и положения модифицируемых коэффициентов

ЦВЗ-система	***QF=50***				***QF=70***			
	$N_W = 1$		$N_W = 4$		$N_W = 1$		$N_W = 4$	
	PSNR, дБ	*BER*	*PSNR, дБ*	*BER*	*PSNR, дБ*	*BER*	*PSNR, дБ*	*BER*
Lin & Chang, 2000 [6]	41,59	1,5E-04	36,57	2,1E-04	44,94	2,0E-04	40,28	1,8E-04
Huang, 2013 [8]	44,06	6,1E-02	39,07	1,2E-03	47,48	1,4E-01	43,01	2,9E-03
Preda & Vizireanu, 2015 [10]	45,26	7,3E-05	38,40	9,2E-05	49,71	2,7E-04	42,66	8,5E-05
Sign-QIM	48,93	**0**	41,77	**0**	52,77	**0**	46,26	6,1E-06
Wang et al., 2011 [11]	**48,94**	**0**	**42,06**	**0**	**52,80**	**0**	**46,40**	**0**
Fallahpour & Megias, 2016 [12]	35,39 (96%)	3,2E-03	35,39 (24%)	3,2E-03	37,60 (99%)	5,1E-03	37,60 (25%)	5,1E-03
Mursi et al., 2009 [13]	45,82	1,2E-04	36,90	3,4E-04	49,50	1,2E-04	41,36	3,6E-04

Результаты этого эксперимента для $QF = 50$ представлены на рис. 8. Заметно, что для всех систем имеются отклонения от идеальной картины. Для оценки степени этих отклонений для каждой кривой были рассчитаны следующие показатели:

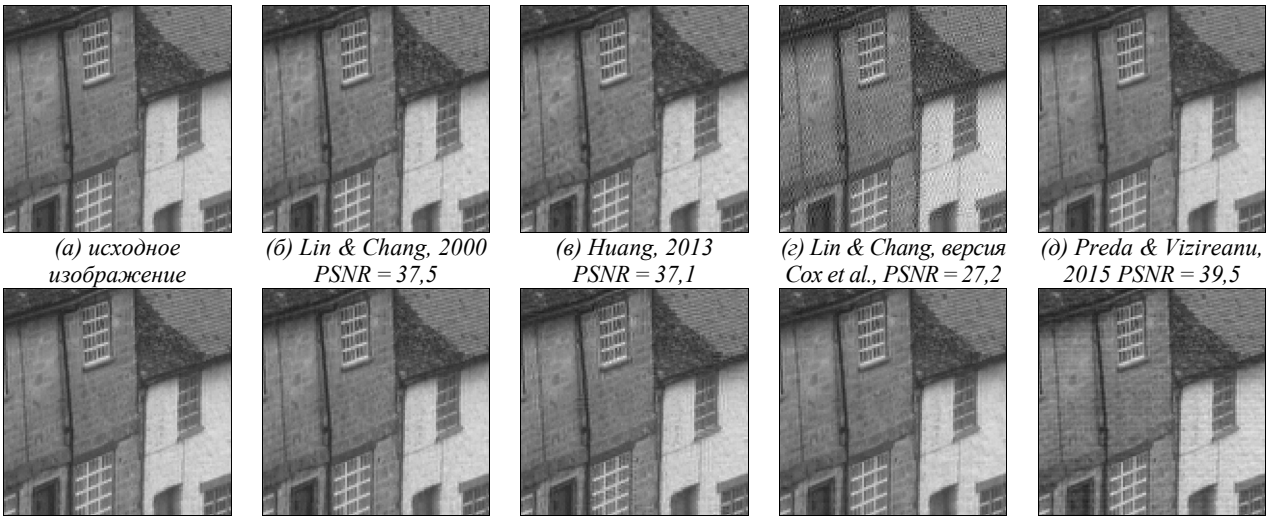
$$err_{FN} = \sum_{QF^*=QF-25}^{QF-1} \left(0.5 - BER\left(QF^*\right)\right), \quad (12)$$

$$err_{FP} = \sum_{QF^*=QF}^{QF+24} BER\left(QF^*\right). \quad (13)$$

Их значения приведены в табл. 7. Более значимым из них является err_{FP}.

Табл. 7. Показатели отклонения значений BER от их теоретической оценки в серии экспериментов с JPEG-сжатием

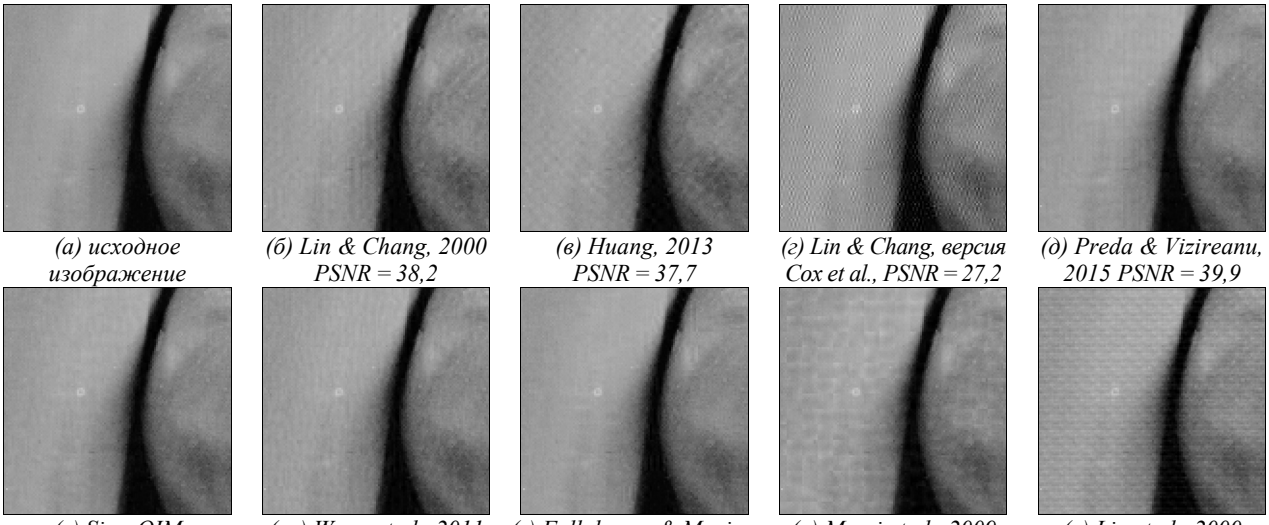
ЦВЗ-система	err_{FN}	err_{FP}
Lin & Chang, 2000 [6]	8,247	0,041
Huang, 2013 [8]	6,176	1,574
Lin & Chang, версия Cox et al. [1]	8,461	0,629
Preda & Vizireanu, 2015 [10]	7,120	0,033
Sign-QIM	7,085	**0,024**
Wang et al., 2011 [11]	**3,796**	2,001
Fallahpour & Megias, 2016 [12]	8,278	11,558
Mursi et al., 2009 [13]	8,642	0,029
Lin et al., 2000 [14]	9,318	2,440
Huang, 2013 [8], α=1	7,178	0,026
Wang et al., 2011 [11], β=0	**2,691**	1,112



(а) исходное изображение | *(б) Lin & Chang, 2000 PSNR = 37,5* | *(в) Huang, 2013 PSNR = 37,1* | *(г) Lin & Chang, версия Cox et al., PSNR = 27,2* | *(д) Preda & Vizireanu, 2015 PSNR = 39,5*

(е) Sign-QIM PSNR = 43,0 | *(ж) Wang et al., 2011 PSNR = 38,8* | *(з) Fallahpour & Megias, 2016, SNR = 36,7* | *(и) Mursi et al., 2009 PSNR = 37,5* | *(к) Lin et al., 2000 PSNR = 33,7*

Рис. 6. Встраивание ЦВЗ в изображение "Goldhill" с использованием оригинальных значений N_W (QF = 50, увеличение 6×)



(а) исходное изображение | *(б) Lin & Chang, 2000 PSNR = 38,2* | *(в) Huang, 2013 PSNR = 37,7* | *(г) Lin & Chang, версия Cox et al., PSNR = 27,2* | *(д) Preda & Vizireanu, 2015 PSNR = 39,9*

(е) Sign-QIM PSNR = 43,0 | *(ж) Wang et al., 2011 PSNR = 38,8* | *(з) Fallahpour & Megias, 2016, SNR = 38,7* | *(и) Mursi et al., 2009 PSNR = 36,6* | *(к) Lin et al., 2000 PSNR = 33,7*

Рис. 7. Встраивание ЦВЗ в изображение "Peppers" с использованием оригинальных значений N_W (QF = 50, увеличение 6×)

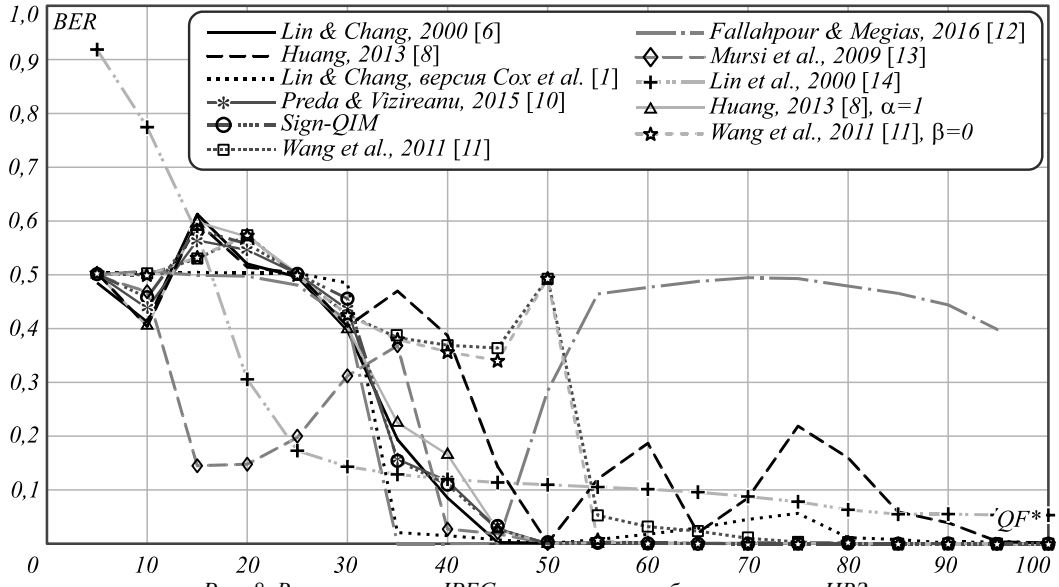


Рис. 8. Влияние уровня JPEG-сжатия на ошибку извлечения ЦВЗ

Как показывают результаты, в области $QF^* \geq QF$ большинство систем демонстрируют низкие значения *BER*. Среди исключений для систем Huang, 2013 и Wang et al., 2011 большое число ошибок объясняется наличием параметров α и β. Если минимизировать влияние этих параметров, приравняв их к 1 и 0 соответственно (что делает эти системы близкими к другим системам на основе НЗБ-встраивания и QIM соответственно), то уровень ошибок падает. Впрочем, это повлечёт снижение *PSNR* для изображений с ЦВЗ. Для системы Lin & Chang, версия Cox et al. высокий уровень ошибок обусловлен зависимостью между различными коэффициентами в блоке. Ошибка наступает в случае изменения любого из 8 коэффициентов в группе, в то время как для других систем всё определяется одним коэффициентом. Наибольший уровень ошибок наблюдается у системы Fallahpour & Megias, 2016. Это объясняется переменным количеством бит, встраиваемым в каждый блок. Таким образом, для данной системы ошибка при извлечении ЦВЗ может проявляться не только в инверсии бита, но и в его пропуске. Последнее приводит к рассинхронизации встроенного и извлечённого ЦВЗ и, следовательно, резкому росту *BER*. Наконец, система Lin et al., 2000, не использующая при встраивании параметр *QF*, не позволяет добиться формы ступенчатой функции на графике. Таким образом, данная система не позволяет обеспечить полухрупкость ЦВЗ.

Среди прочих систем наилучшее значение показателя err_{FP} достигается для системы Sign-QIM, немногим уступают ей Huang, 2013 (при $\alpha = 1$), Mursi et al., 2009, Preda & Vizireanu, 2015 и Lin & Chang, 2000.

В области $QF^* < QF$ почти все системы имеют переходную фазу, однако следует отметить, что большинство из них уже при небольшом отличии QF^* от QF превышают порог в 0,05. Этот факт можно использовать в качестве индикатора наличия недопустимых искажений изображения со встроенным ЦВЗ. Как показывает табл. 7, наименьшие отклонения от теоретических значений наблюдаются для системы Wang et al., 2011.

Заключение

В настоящей работе рассмотрен класс полухрупких ЦВЗ-систем, предназначенных для аутентификации JPEG-изображений. Приведено формальное схематическое описание сценария использования систем данного класса. Подробно рассмотрено свыше десятка таких систем, и предложена их классификация по различным критериям: по используемому методу встраивания, по местоположению и количеству изменяемых коэффициентов, а также по способу формирования ЦВЗ. Так, в рамках первой классификации выделены следующие методы: НЗБ-встраивание, QIM, изменение позиции LNZ компоненты, табличное отображение и расширение спектра.

В рамках исследовательской части работы проведена оценка уровня искажений, возникающих при встраивании информации, а также погрешности при извлечении ЦВЗ. Данное исследование показало, что основное влияние на качество результирующего изображения оказывает расположение модифицируемых коэффициентов ДКП, а именно соответствующие им значения матрицы Q_{QF}. Так, система Lin&Chang, версия Cox et al., несмотря на механизм минимизации искажений при квантовании, существенно проигрывает прочим системам за счёт использования высокочастотных коэффициентов.

Сравнивая различные методы встраивания ЦВЗ, можно сказать, что наилучшим образом проявили себя системы на основе QIM, и главным образом система Wang et al., 2011. Помимо высокого качества защищённых изображений, эта система обеспечивает безошибочное извлечение ЦВЗ из неискажённого контейнера. Помимо этой системы, хорошо показала себя система Sign-QIM, являющаяся модификацией системы Preda & Vizireanu, 2015, а также Mursi et al., 2009 на основе табличного отображения.

Кроме того, исследована работоспособность систем в смысле обеспечения частичной стойкости к JPEG-сжатию. Исследование показало, что системы Lin & Chang, 2000, Preda & Vizireanu, 2015, Sign-QIM, Mursi et al., 2009, а также Huang, 2013 и Wang et al., 2011 (при должном подборе множителей α и β) позволяют выполнять полухрупкое встраивание. Оставшиеся три системы (Lin&Chang, версия Cox et al., Fallahpour&Megias, 2016, Lin et al., 2000) по разным причинам не позволяют этого добиться.

Благодарности

Работа выполнена за счет гранта Российского научного фонда (проект № 18-71-00052).

Литература

1. **Cox, I.** Digital Watermarking and Steganography: Morgan Kaufmann / I. Cox, J. Kilian, F.T. Leighton, T. Shamoon. – Elsevier, 2008. – 624 p.
2. **Федосеев, В.А.** Цифровые водяные знаки и стеганография: учеб. пособие / В.А. Федосеев. – Самара: Издательство Самарского университета, 2019. – 144 с.
3. **Федосеев, В.А.** Унифицированная модель систем встраивания информации в цифровые сигналы / В.А. Федосеев // Компьютерная оптика. – 2016. – Т. 40, № 1. – С. 87-98.
4. **Barni, M.** Watermarking systems engineering: Enabling digital assets security and other applications / M. Barni, F. Bartolini. – CRC Press, 2004. – 500 p.
5. **Lin, C.-Y.** Issues and solutions for authenticating MPEG video / C.-Y. Lin, S.-F. Chang // Proceedings of SPIE. – 1999. – P. 54-65.
6. **Lin, C.-Y.** Semifragile watermarking for authenticating JPEG visual content / C.-Y. Lin, S.-F. Chang // Electronic Imaging. – 2000. – P. 140-151.
7. **Ho, C.K.** Semi-fragile watermarking scheme for authentication of JPEG images / C.K. Ho, C.-T. Li // International Conference on Information Technology: Coding and Computing, 2004. Proceedings. – 2004. – Vol. 1. – P. 7-11.
8. **Huang, L.-Y.** Authentication watermarking algorithm resisting JPEG compression based on preliminary quantization / L.-Y. Huang // Information Technology Journal. – 2013. – Vol. 12(16). – P. 3723-3728.

9. **Ye, S.** A quantization-based image authentication system / S. Ye, Z. Zhou, Q. Sun, E. Chang, Q. Tian // Proceedings of the 2003 Joint Fourth International Conference on Information, Communications and Signal Processing, 2003 and the Fourth Pacific Rim Conference on Multimedia. Proceedings of the 2003 Joint. – 2003. – Vol. 2. – P. 955-959.
10. **Preda, R.O.** Watermarking-based image authentication robust to JPEG compression / R.O. Preda, D.N. Vizireanu // Electronics Letters. – 2015. – Vol. 51(23). – P. 1873-1875.
11. **Wang, H.** A novel fast self-restoration semi-fragile watermarking algorithm for image content authentication resistant to JPEG compression / H. Wang, A. Ho, X. Zhao // Digital Forensics and Watermarking. – 2011. – Vol. 7128. – P. 72-85.
12. **Fallahpour, M.** Flexible image watermarking in JPEG domain / M. Fallahpour, D. Megias // 2016 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT). – 2016. – P. 311-316.
13. **Mursi, M.** A DCT-based secure JPEG image authentication scheme / M. Mursi, G.M.R. Assassa, H. Aboalsamh, K. Alghathbar // World Academy of Science, Engineering and Technology. – 2009. – Vol. 53. – P. 681-687.
14. **Lin, E.T.** Detection of image alterations using semifragile watermarks / E.T. Lin, C.I. Podilchuk, E.J. Delp // Security and Watermarking of Multimedia Contents II. – 2000. – Vol. 3971. – P. 152-164.
15. **Al-Mualla, M.E.** Content-adaptive semi-fragile watermarking for image authentication / M.E. Al-Mualla // 2007 14th IEEE International Conference on Electronics, Circuits and Systems. – 2007. – P. 1256-1259.
16. **Wong, P.H.W.** Data hiding technique in JPEG compressed domain / P.H.W. Wong, O.C.L. Au, J.W.C. Wong // Proceedings of SPIE – The International Society for Optical Engineering, 2001. – P. 309-320.
17. **Xiao, J.** A semi-fragile watermarking distinguishing JPEG compression and gray-scale-transformation from malicious manipulation / J. Xiao, Z. Ma, B. Lin, J. Su, Y. Wang // 2010 IEEE Youth Conference on Information, Computing and Telecommunications (YC-ICT). – 2010. – P. 202-205.
18. **Евсютин, О.О.** Улучшенный алгоритм встраивания информации в сжатые цифровые изображения на основе метода PM1 / О.О. Евсютин, А.С. Кокурина, А.А. Шелупанов, И.И. Шепелев // Компьютерная оптика. – 2015. – Т. 39, № 4. – С. 572-581. – DOI: 10.18287/0134-2452-2015-39-4-572-581.
19. **Yu, X.** Review on semi-fragile watermarking algorithms for content authentication of digital images / X. Yu, C. Wang, X. Zhou // Future Internet. – 2017. – Vol. 9, Issue 4. – 56.
20. **Wallace, G.K.** The JPEG still picture compression standard / G.K. Wallace // IEEE Transactions on Consumer Electronics. – 1992. – Vol. 38, Issue 1. – P. xviii-xxxiv.
21. **Chen, B.** Quantization index modulation: A class of provably good methods for digital watermarking and information embedding / B. Chen, G. Wornell // IEEE Transaction on Information Theory. – 2001. – Vol. 47, Issue 4. – P. 1423-1443.
22. **Cox, I.J.** Secure spread spectrum watermarking for multimedia / I.J. Cox, J. Kilian, F.T. Leighton, T. Shamoon // IEEE Transactions on Image Processing. – 1997. – Vol. 6, Issue 12. – P. 1673-1687.
23. **Mitekin, V.** A new QIM-based watermarking algorithm robust against multi-image histogram attack / V. Mitekin, V. Fedoseev // Procedia Engineering. – 2017. – Vol. 201. – P. 453-462. – DOI: 10.1016/j.proeng.2017.09.687.
24. **Митекин, В.А.** Алгоритмы встраивания информации на основе QIM, стойкие к статистической атаке / В.А. Митекин, В.А. Федосеев // Компьютерная оптика. – 2018. – Т. 42, № 1. – С. 118-127. – DOI: 10.18287/2412-6179-2018-42-1-118-127.
25. **Глумов, Н.И.** Обнаружение дубликатов на изображениях / Н.И. Глумов, А.В. Кузнецов // Компьютерная оптика. – 2011. – Т. 35, № 4. – С. 508-512.
26. **Глумов, Н.И.** Обнаружение на изображениях искусственных изменений локального происхождения / Н.И. Глумов, А.В. Кузнецов // Автометрия. – 2011. – Т. 47(3). – С. 4-12.
27. The USC-SIPI Image Database [Электронный ресурс]. – Режим доступа: http://sipi.usc.edu/database/ (дата обращения 07.04.2019).

Сведения об авторах

Егорова Анна Александровна, 1995 года рождения. В 2017 году окончила Самарский национальный исследовательский университет имени академика С.П. Королёва (Самарский университет) с отличием по специальности «Информационная безопасность автоматизированных систем». В настоящее время является аспирантом Самарского университета. Основные сферы научных интересов: обработка изображений, обнаружение искажений на цифровых изображениях, защита информации. E-mail: *2358anna@gmail.com* .

Федосеев Виктор Андреевич, 1986 года рождения, в 2009 году окончил Самарский государственный аэрокосмический университет имени академика С.П. Королёва (ныне – Самарский университет) по специальности «Прикладная математика и информатика», кандидат физико-математических наук (2012). В настоящее время работает доцентом кафедры геоинформатики и информационной безопасности Самарского университета и научным сотрудником в ИСОИ РАН – филиале ФНИЦ «Кристаллография и фотоника» РАН. Области научных интересов: анализ изображений, цифровые водяные знаки, стеганография. E-mail: *vicanfed@gmail.com* .

ГРНТИ: 28.23.15

Поступила в редакцию 8 апреля 2019 г. Окончательный вариант – 22 апреля 2019 г.

A classification of semi-fragile watermarking systems for JPEG images

A.A. Egorova [2], V.A. Fedoseev [1,2]

[1]IPSI RAS - Branch of the FSRC "Crystallography and Photonics" RAS, Molodogvardeyskaya 151, 443001, Samara, Russia,

[2]Samara National Research University, 443086, Russia, Samara, Moskovskoye Shosse 34

Abstract

The article discusses semi-fragile digital watermarking systems designed to protect JPEG images against unauthorized changes. These systems allow detecting and locating changes, and some of them are able to restore the original content. Formal schemes describing the procedures for watermark embedding and authentication are given. We consider more than a dozen systems of this type and classify them according to various criteria. We also consider the results of experimental studies of these systems estimating their level of distortions arising from watermark embedding. In addition, we test the systems whether they are able to operate in the semi-fragile way.

Keywords: digital watermarking, image authentication, semi-fragile watermarking, JPEG, QIM, LSB.

Citation: Egorova AA, Fedoseev VA. A classification of semi-fragile watermarking systems for JPEG images. Computer Optics 2019; 43(3): 419-433. DOI: 10.18287/2412-6179-2019-43-3-419-433.

Acknowledgements: The work was funded by the Russian Science Foundation under grant #18-71-00052.

References

[1] Cox I, Kilian J, Leighton FT, Shamoon T. Digital Watermarking and Steganography: Morgan Kaufmann. Elsevier; 2008.

[2] Fedoseev VA. Digital Watermarking and Steganography: A Tutorial. 2nd ed. Samara: Samara University; 2019.

[3] Fedoseev VA. A unified model for information hiding systems. Computer Optics 2016; 40: 87-98.

[4] Barni M, Bartolini F. Watermarking systems engineering: Enabling digital assets security and other applications. CRC Press; 2004.

[5] Lin C-Y, Chang S-F. Issues and solutions for authenticating MPEG video. Proceedings of SPIE 1999; 54-65.

[6] Lin C-Y, Chang S-F. Semifragile watermarking for authenticating JPEG visual content. Electronic Imaging 2000; 140-151.

[7] Ho CK, Li CT. Semi-fragile watermarking scheme for authentication of JPEG images. International Conference on Information Technology: Coding and Computing, 2004. Proceedings 2004; 1: 7-11.

[8] Huang L-Y. Authentication watermarking algorithm resisting JPEG compression based on preliminary quantization. Information Technology Journal 2013; 12(16): 3723-3728.

[9] Ye S, Zhou Z, Sun Q, Chang E, Tian Q. A quantization-based image authentication system. Proceedings of the 2003 Joint Fourth International Conference on Information, Communications and Signal Processing, 2003 and the Fourth Pacific Rim Conference on Multimedia. Proceedings of the 2003 Joint 2003; 2: 955-959.

[10] Preda RO, Vizireanu DN. Watermarking-based image authentication robust to JPEG compression. Electronics Letters 2015; 51(23): 1873-1875.

[11] Wang H, Ho A, Zhao X. A novel fast self-restoration semi-fragile watermarking algorithm for image content authentication resistant to JPEG compression. Digital Forensics and Watermarking. – 2011. – Vol. 7128. – P. 72-85.

[12] Fallahpour M, Megias D. Flexible image watermarking in JPEG domain. 2016 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT) 2016; 311-316.

[13] Mursi M, Assassa GMR, Aboalsamh H, Alghathbar K. A DCT-based secure JPEG image authentication scheme. World Academy of Science, Engineering and Technology 2009; 53: 681-687.

[14] Lin ET, Podilchuk CI, Delp EJ. Detection of image alterations using semifragile watermarks. Security and Watermarking of Multimedia Contents II 2000; 3971: 152-164.

[15] Al-Mualla ME. Content-adaptive semi-fragile watermarking for image authentication. 2007 14th IEEE International Conference on Electronics, Circuits and Systems 2007; 1256-1259.

[16] Wong PHW, Au OCL, Wong JWC. Data hiding technique in JPEG compressed domain. Proceedings of SPIE – The International Society for Optical Engineering 2001; 309-320.

[17] Xiao J, Ma Z, Lin B, Su J, Wang Y. A semi-fragile watermarking distinguishing JPEG compression and gray-scale-transformation from malicious manipulation. 2010 IEEE Youth Conference on Information, Computing and Telecommunications (YC-ICT) 2010; 202-205.

[18] Evsutin OO, Kokurina AS, Shelupanov AA, Shepelev II. An improved algorithm for data hiding in compressed digital images based on PM1 method. Computer Optics 2015; 39(4): 572-581. DOI: 10.18287/0134-2452-2015-39-4-572-581.

[19] Yu X, Wang C, Zhou X. Review on Semi-Fragile Watermarking Algorithms for Content Authentication of Digital Images. Future Internet 2017; 9(4): 56.

[20] Wallace GK. The JPEG still picture compression standard. IEEE Transactions on Consumer Electronics 1992; 38(1): xviii-xxxiv.

[21] Chen B, Wornell G. Quantization index modulation: A class of provably good methods for digital watermarking and information embedding. IEEE Transaction on Information Theory 2001; 47(4): 1423-1443.

[22] Cox IJ, Kilian J, Leighton FT, Shamoon T. Secure spread spectrum watermarking for multimedia. IEEE Transactions on Image Processing 1997; 6(12): 1673-1687.

[23] Mitekin V, Fedoseev V. A new QIM-based watermarking algorithm robust against multi-image histogram attack. Procedia Engineering 2017; 201: 453-462. DOI: 10.1016/j.proeng.2017.09.687.

[24] Mitekin VA, Fedoseev VA. New secure QIM-based information hiding algorithms. Computer Optics 2018; 42(1): 118-127. DOI: 10.18287/2412-6179-2018-42-1-118-127.

[25] Glumov NI, Kuznetsov AV. Copy-move image forensics detection. Computer Optics 2011; 35(4): 508-512.

[26] Glumov NI, Kuznetsov AV. Detection of local artificial changes in images. Optoelectronics, Instrumentation and Data Processing 2011; 47(3): 207-214. DOI: 10.3103/S8756699011030010.

[27] The USC-SIPI Image Database. Source: ⟨http://sipi.usc.edu/database/⟩.

Author's information

Anna Aleksandrovna Egorova (b. 1995) graduated with honors from Samara National Research University (Samara University) majoring in Information Security of Computer-Aided Systems in 2017. Nowadays she is a postgraduate at Samara National Research University. Main research interests: image processing, digital image forgery detection and information security. E-mail: *2358anna@mail.ru*.

Victor Andreevich Fedoseev (b. 1986) graduated (2009) from Samara State Aerospace University (presently, Samara National Research University), majoring in Applied Mathematics and Computer Science. Candidate degree in Computer Science (2012). Currently he is an associate professor at the Geoinformatics and Information Security вepartment at Samara University and a research scientist at Image Processing Systems Institute of RAS – Branch of the FSRC "Crystallography and Photonics" RAS. His scientific interests include image processing and analysis, digital watermarking and steganalysis. E-mail: *vicanfed@gmail.com*.

Received April 8, 2019. The final version – April 22, 2019.