Белобокова Ю.А., МГТУ имени Н.Э. Баумана,
Клышинский Э.С., МИЭМ НИУ ВШЭ

Защита информационного содержания цифровых фотографий

методом многократной маркировки цифровыми водяными знаками

В статье излагается метод защиты цифровых изображений от модификации. Изображение разделяется на прямоугольные блоки, каждый из которых подписывается двумя цифровыми водяными знаками. Первый водяной знак служит для поиска блоков и контролирует их взаимное расположение. Второй является цифровой подписью и контролирует неизменность содержимого блока изображения. Метод позволяет определять не только изменения всего изображения в целом, но и какая именно часть изображения подвергалась модификации (искажению, вставке постороннего изображения, реплицированию)

В условиях активного развития информационных технологий, в частности, усовершенствования функций графических редакторов и лёгкости доступа к мультимедийным данным, размещённым в компьютерных сетях, проблема защиты информационного содержания цифровых фотографий от подделок и незаконного использования приобретает всё большую актуальность.

При заимствовании фотография может подвергнуться кадрированию, обработке фильтрами графических редакторов и инструментами коррекции изображений, также могут быть удалены или добавлены какие-либо информационные фрагменты. Возможно также изменение формата файла фотографии, её масштабирование и повороты. Следовательно, разрабатываемый инструмент защиты цифровых фотоизображений должен быть устойчив к последствиям воздействий на них. Кроме того, необходимо разработать способ определения факта изменения информационного содержания фотографий.

Предназначенные для использования в электронных средствах массовой информации не анимированные полутоновые и полноцветные фотографии обычно сохраняют в цифровых форматах PNG и JPEG. К достоинствам этих форматов можно отнести кроссплатформенность, возможность обработки практически во всех графических редакторах, хорошие показатели качества изображений. Изображения в формате JPEG за счёт возможности их сжатия с потерями имеют меньший размер по сравнению с аналогичными, сохранёнными в формате PNG. При этом качество цифровых фотографий с большим коэффициентом сжатия может сильно проигрывать фотографиям в формате PNG, в котором используется алгоритм сжатия без потерь. Следовательно, разрабатываемый инструмент защиты должен учитывать особенности форматов защищаемых изображений.

В качестве основы для инструмента защиты были выбраны внедряемые в цифровые фотоизображения невидимые метки, также называемые цифровыми водяными знаками (ЦВЗ). Обычно цифровая фотография маркируется ЦВЗ в соответствии с определённым ключом [1], при этом внедряемый знак должен быть устойчивым к воздействиям на фотографию и незаметным, то есть не должен вносить в неё видимых искажений. Для проверки целостности информационного содержания промаркированного изображения производят поиск встроенной информации и в случае успешного нахождения и извлечения проверку на соответствие исходной.

Научная новизна разрабатываемой методики защиты информационного содержания цифровых изображений заключается в возможности разделения оригинальных и измененных фрагментов изображений. Это достигается благодаря маркировке двумя видами ЦВЗ: bitmap-логотипом, общим для всей фотографии, который позволяет подтвердить подлинность фрагментов, и строками бит, в дальнейшем именуемыми электронными цифровыми подписями (ЭЦП), внедряемыми в изображение в определённом порядке. Несовпадение извлечённого из фрагмента изображения и рассчитанного значения ЭЦП при наличии маркировки bitmap-логотипом указывает на факт изменения этого фрагмента или всего изображения в целом.

Исходная цифровая фотография, сохранённая в цветовой модели RGB, представляется в виде двумерной матрицы $Pict = |pix_{i,j}|$, где $i \in [1;h]$, $j \in [1;w]$, w и h – количество пикселей по горизонтали и вертикали. Каждый из пикселей фотоизображения представляется в виде кортежа $pix = <r, g, b>$, где r, g, b – значения интенсивностей цвета пикселя по цветовым каналам ($r, g, b \in [0;255]$). Матрица Pict разбивается на подматрицы P, которые, в свою очередь, разделяются на три цветовые плоскости.

Принцип встраивания ЦВЗ был выработан с учётом системы зрения человека (СЗЧ). Доказано [1], что человеческий глаз воспринимает 7 бит из 8 в красном канале цифрового изображения, 8 из 8 в зелёном канале и 4 из 8 в синем канале, поэтому внедрение водяных знаков производится в красную и синюю цветовые плоскости.

Разработанная методика защиты основана на алгоритме, разработанном Кохом и Жао [2]. Известно, что ЦВЗ, встроенные в частотную область изображения, обладают большей устойчивостью [3], при этом для повышения устойчивости водяных знаков желательно выбирать алгоритм внедрения, аналогичный алгоритму сжатия [4]. Выбранный алгоритм основан на дискретном косинусном преобразовании (ДКП), которое также используется в алгоритме сжатия JPEG. Поскольку необходимо промаркировать каждую из матриц изображения, bitmap-логотип должен быть небольшого размера. Алгоритм Коха и Жао позволяет многократно встраивать небольшие ЦВЗ, не требуя для внедрения негладких и многоконтурных блоков. Кроме того, извлечение скрытой информации при использовании этого алгоритма происходит по так называемой слепой схеме, то есть без исходного изображения.

Суть алгоритма Коха и Жао заключается в следующем. В красной и синей цветовых плоскостях подматриц защищаемого изображения с использованием определённого ключа выбирают области размером 8x8 пикселей по числу встраиваемых бит водяного знака. К выбранным областям применяют двумерное ДКП:

$$DCT(u,v) = \sqrt{\frac{2}{m}}a(v)\left(\sum_{j=0}^{m-1}\left(\sqrt{\frac{2}{n}}a(u)\sum_{i=0}^{n-1}Fi,j\cos\left(\frac{\pi(2i+1)u}{2n}\right)\right)\cos\left(\frac{\pi(2j+1)v}{2m}\right)\right) \quad (1)$$

где:

- DCT (u,v) – значения элементов матрицы частотных коэффициентов;
- Fi,j – значения элементов целочисленной матрицы изображений;
- n – количество столбцов матрицы;
- m – количество строк;
- i, j – позиция текущего элемента матрицы изображения;
- u, v – позиция формируемого элемента частотной матрицы;
- $u \in [0,i-1]$, $v \in [0,j-1]$.

При $u = 0 \rightarrow a(u) = \frac{1}{\sqrt{2}}$, при $u > 0 \rightarrow a(u) = 1$

При $v = 0 \rightarrow a(u) = \frac{1}{\sqrt{2}}$, при $v > 0 \rightarrow a(v) = 1$

Таким образом, целочисленные матрицы пикселей преобразуются в матрицы частотных коэффициентов. Обозначим K_1 и K_2 два коэффициента из области средних частот. Для встраивания бита водяного знака, имеющего значение 0, коэффициенты корректируют так, чтобы выполнялось условие:

$$|K_1| - |K_2| > p \quad (2)$$

где p – целочисленный параметр, влияющий на силу встраивания.

Для встраивания бита водяного знака, имеющего значение 1, добиваются выполнения условия:

$$|K_1| - |K_2| < -p \quad (3)$$

После встраивания в матрицы частотных коэффициентов бит водяных знаков к изменённым матрицам применяют обратное ДКП. При этом сначала выполняется одномерное ДКП по строкам, затем по столбцам:

$$F(u,j) = \sqrt{\frac{2}{m}}\sum_{v=1}^{m-1}a(v)DCT(u,v)\cos\left[\frac{\pi(2j+1)v}{2m}\right] \quad (4)$$

$$F(i,j) = \sqrt{\frac{2}{n}}\sum_{u=1}^{n-1}a(uv)F(u,j)\cos\left[\frac{\pi(2i+1)u}{2n}\right] \quad (5)$$

Для встраивания бит логотипа используются подматрицы красной и синей цветовых плоскостей, для встраивания бит ЭЦП – только центральная область подматриц синей цветовой плоскости.

Для проверки целостности информационного содержания промаркированной цифровой фотографии Pict` необходимо последовательно решить две задачи: определить разбиение фотоизображения на подматрицы и проверить корректность водяных знаков каждой найденной подматрицы.

Для решения первой задачи используется bitmap-логотип. В связи с этим для проверки целостности информационного содержания фотографии должен быть известен ключ, с помощью которого биты логотипа внедрялись в подматрицы; также желательно наличие самого bitmap-логотипа.

Если промаркированная фотография не модифицировалась, расположение подматриц будет неизменным, но в случае кадрирования или добавления в неё инородных фрагментов координаты начал подматриц будут смещены. Алгоритм поиска бит логотипа состоит в следующем.

На первом этапе смещения по горизонтали и по вертикали равны нулю. Для текущего смещения выделяют подматрицу заданного размера и проводят в ней поиск бит логотипа по формулам (1)-(5). Если все биты найдены в соответствии с ключом и полученный логотип совпадает с исходным, считается, что найден базовый блок. В противном случае текущая позиция по горизонтали смещается на один пиксель на интервале [1;wP]. Если достигнут предел интервала, текущее смещение по горизонтали присваивается равным 0, и производят смещение по вертикали на интервале [1;hP]. Если алгоритм перебрал все точки в заданном диапазоне, считают, что базовый блок не найден, после чего случайным образом генерируют смещение по вертикали и горизонтали и повторяют алгоритм еще раз (или при необходимости несколько раз).

Это делается в связи с тем, что по краям фотографии могли быть добавлены инородные области. Количество итераций алгоритма выбирается таким образом, чтобы гарантировать вероятность того, что изображение выше заданного уровня полностью не промаркировано. В случае если базовый блок был найден, начиная с этого блока оставшаяся фотография делится на подматрицы.

Проверка целостности информационного содержания также проводится по формулам (1)-(5). Расположение логотипов однозначно указывает на местонахождение центральной области, содержащей ЭЦП. С помощью выбранного метода проводят извлечение цифровой информации из этой области по формулам (1)-(5). Далее по алгоритму размещения строк бит проводят расчёт значения ЭЦП. Если рассчитанная и найденная контрольные суммы совпадают, подматрица считается корректной и не подвергавшейся модификации. В противном случае считают, что подматрица была подвергнута изменениям.

Возможны три исхода проверки подматриц. В первом случае, когда логотип не был найден, считается, что подматрица не содержалась в исходной фотографии, либо изображение в подматрице было подвергнуто изменениям. Во втором случае

весь логотип корректен, но контрольная сумма не совпадает с рассчитанной.

В этом случае можно предполагать, что фотография подвергалась небольшим изменениям, не попавшим на область внедрения бит логотипа. В третьем случае сохранились как логотип, так и ЭЦП. Корректность цифровой подписи в этом случае будет указывать на то, что подматрица не была модифицирована.

Таким образом анализируется корректность всех подматриц фотографии. Модифицированные подматрицы могут смыкаться в единые области, и за счет этого можно сделать вывод о границах, в которых проводилась модификация промаркированной цифровой фотографии.

Для проведения экспериментов с использованием данной методики была разработана программа, выполненная на языке программирования C#.

Защищаемая фотография маркируется созданным пользователем логотипом с выбираемым параметром p (см. рис. 1).

Значение p влияет на стойкость вложений при сжатии изображения [4], но при достаточно высоких значениях маркировка изображения становится заметной (см. рис. 2). Опытным путём было установлено, что оптимальным является значение p, равное 40.

После маркировки логотипом и ЭЦП полученные фотографии сохранялись в формате PNG.

К промаркированным цифровым фотографиям применялись кадрирование, клонирование фрагментов, добавление фрагментов непромаркированных изображений, осуществлялись смена формата изображения, воздействие инструментами коррекций изображений Photoshop и сжатие с потерями в формате JPEG. Искаженные таким образом цифровые фотографии проверялись на стойкость вложений.

Проведённая проверка изменённых цифровых изображений на наличие вложений подтвердила первоначальную гипотезу.

По изменениям в расположении внедренных меток можно судить о характере изменения изображения. Если маркировка bitmap-логотипами нерегулярна, это говорит о том, что производились действия по удалению, добавлению или сдвигу фрагментов цифровой фотографии. На рис. 3 показан результат обработки фотографии с клонированным фрагментом (показан контрастным цветом, поскольку блоки изображения были смещены). Чёрные диагональные метки показывают, что контрольные суммы ЭЦП не соответствуют рассчитанным.

Если порядок расположения ЭЦП не соответствует заданному, можно определить, осуществлялось ли кадрирование цифровой фотографии. На рис. 4 показан результат проверки на фотографии с обрезанным правым краем. Логотип распознан, но чёрные диагональные метки в блоках, содержащих ЭЦП, в которой зашифрован размер изображения по ширине, свидетельствуют об изменении горизонтального размера.

При смене формата цифровой фотографии с PNG на JPEG было выявлено, что стойкость внедренных меток зависит от степени сжатия изображения. Результаты тестирования на различных фотографиях показали, что встроенный логотип обнаруживается в случае сохранения изображения с коэффициентом качества, равным 10-12. При сохранении изображения с коэффициентом качества 9 и ниже логотип не детектируется. Перевод изображения из формата JPEG в формат PNG не ухудшает стойкость логотипа.

Рисунок 1. Фрагмент окна программы. Исходная фотография, bitmap-логотип и параметр p

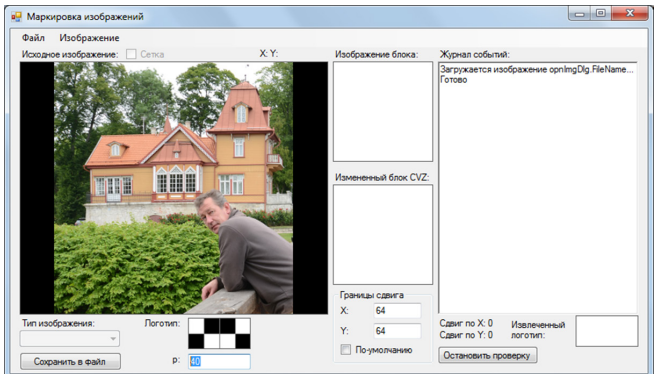


Рисунок 2. Слева – промаркированное изображение со значением p=40. Справа – промаркированное изображение со значением p=400, видны следы маркировки

Рисунок 3. Фрагмент окна программы. Фотография с клонированным фрагментом

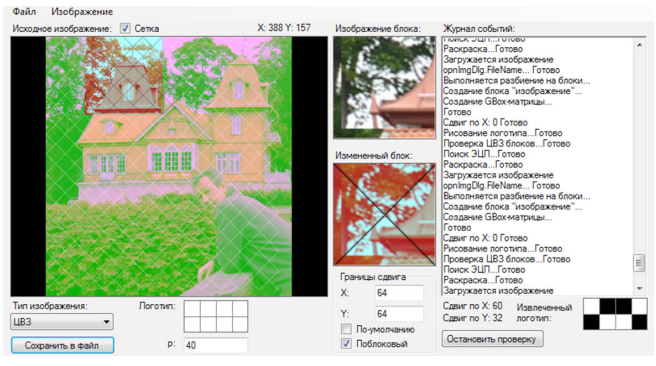
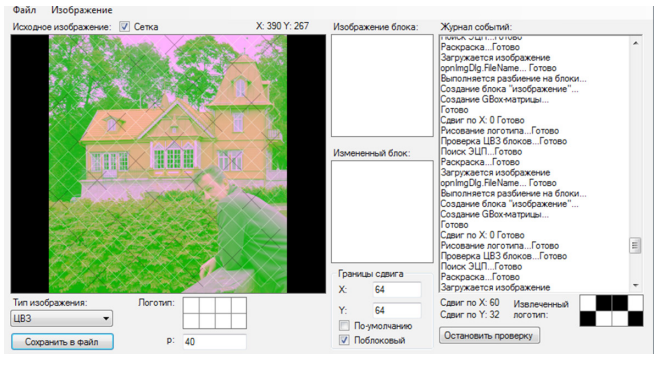


Рисунок 4. Фрагмент окна программы. Изображение, обрезанное справа



Изменение ЭЦП подматриц в зависимости от степени сжатия изображения в формате JPEG показано на рис. 5.

Разработанный метод защиты информационного содержания цифровых фотографий позволяет выявить факты клонирования или удаления фрагментов, добавления инородных объектов, кадрирования и воздействия инструментов коррекции изображений.

Используемый метод в определённых пределах устойчив к смене формата изображения, границы его устойчивости определяются алгоритмом Коха и Жао. При этом метод не является жёстко привязанным к этому алгоритму, в целях повышения устойчивости вложения водяных знаков возможна смена алгоритма внедрения.

Программная реализация показала практическую значимость и корректность разработанного метода. EOF

Авторы выражают благодарность Янусу Кюльмхаллику за помощь в проведении экспериментов.

1. В.Г. Грибунин, И.Н. Оков, И.В. Туринцев. Цифровая стеганография. Наука и учеба. 2002. – 288 с.
2. Koch E., Zhao J. Towards Robust and Hidden Image Copyright Labeling // IEEE Workshop on Nonlinear Signal and Image Processing. 1995. – P. 123-132.
3. Г.Ф. Конахович, А.Ю. Пузыренко. Компьютерная стеганография. Теория и практика. – Киев: «МК-Пресс», 2006.
4. А.В. Аграновский, А.В. Балакин, В.Г. Грибунин, С.А. Сапожников. Стеганография, цифровые водяные знаки и стеганоанализ. – М.: «Вузовская книга», 2009.

Ключевые слова: защита цифровых изображений, цифровые водяные знаки.

Рисунок 5. Фотографии после смены формата. Слева изображение в формате JPEG, сохранённое с коэффициентом качества 10, справа – изображение, сохранённое с коэффициентом качества 12

Securing of Digital Photos' Content Using Multiple Marking by Digital Watermark

Annotation. The article introduce a new method of securing of digital images. The image on hand is divided into rectangular blocks. Every block is signed by two digital watermarks. The aim of the first one is helping to detect every block on the image. The second one stores a digital signature of the block and thus preserves its content. Therefore the introduced method allow not only detecting of the whole image corruption

Keywords: Securing of digital images, digital watermarks.