Телевизионные системы, передача и обработка изображений

УДК 519.688

КОМБИНИРОВАННЫЙ АЛГОРИТМ ВСТРАИВАНИЯ ЦИФРОВЫХ ВОДЯНЫХ ЗНАКОВ В ИЗОБРАЖЕНИЯ ГРАФИЧЕСКИХ ФОРМАТОВ ДЛЯ СКРЫТОЙ ПЕРЕДАЧИ ИНФОРМАЦИИ

Шелухин Олег Иванович
доктор технических наук, профессор, заведующий кафедрой «Информационная безопасность и автоматизация», Московский технический университет связи и информатики.
E-mail: heluhin@mail.ru.
Адрес: г. Москва, ул. Авиамоторная, д. 8а.

Смычёк Михаил Александрович
кандидат технических наук, главный специалист отдела проектирования сетей связи, АО «Гипрогазцентр».
E-mail: m-smychek@mail.ru.
Адрес: г. Нижний Новгород, ул. Алексеевская, 26.

Аннотация: Для встраивания цифровых водяных знаков (ЦВЗ) в неподвижные изображения формата JPEG предложен комбинированный алгоритм, объединяющий возможности встраивания больших объёмов данных с помощью пространственного метода в формат BMP и сжатия данных путём JPEG кодирования. Реализация представленного стеганографического алгоритма в виде программного обеспечения показала, что количество битов, пригодных для встраивания, зависит от исходного JPEG-файла (характер зависимости определяется JPEG-преобразованием) и линейно зависит от числа LSB, используемых для встраивания. Сохранение сообщения в игнорируемых сегментах файла JPEG позволяет скрыть большое количество байтов сообщения, однако факт такого сокрытия легко обнаружить. В результате оценки качества внедряемой информации сделан вывод о возможности встраивания большого объёма стеганографических данных в изображения без изменения визуальной составляющей изображения – то есть не заметной человеческому глазу. Показана низкая устойчивость разработанного алгоритма противостоять атакам вида «поворот изображения» и аддитивному шуму. Показано, что разработанный алгоритм может быть использован для скрытия и обмена данными в условиях воздействия внешних атак при условии использования дополнительного помехоустойчивого кодирования.

Ключевые слова: встраивание информации, пространственная область изображения, цифровой водяной знак, видеоконтейнер, атака, оценка качества.

Введение

Исследования показывают [1-4], что использование только одной области изображения (пространственной или частотной) не позволяет осуществлять сокрытие в них больших объёмов данных, и может быть реализовано только для встраивания цифровых водяных знаков (ЦВЗ) в виде коротких последовательностей байтов. Вместе с тем, анализ структуры формата файлов JPEG показывает, что в нём присутствуют маркеры определяющие сегменты, которые не участвуют в JPEG преобразовании и не влияют на визуализацию изображения, а потому – игнорируются программами просмотра. Это позволяет сделать вывод о возможности разработки алгоритма встраивания больших объёмов данных в изображения графических форматов JPEG и BMP с использованием форматного метода. В качестве базового формата для исследований и разработок был выбран формат JPEG как наиболее распространённый в сценариях повседневного использования цифровой графики, в частности – цифровых фотографий.

Целью статьи является оценка возможностей комбинированного алгоритма встраивания ЦВЗ, объединяющего возможности пространственного и форматного методов скрытия больших объёмов данных в неподвижных изображениях, его практическая реализация и оценка эффективности алгоритма при воздействии на изображение, содержащее ЦВЗ, типовых графических атак в виде шума, мозаики, поворота, кристаллизации, jpg-сжатия.

Возможности использования формата JPEG для встраивания метаданных

Анализ структуры формата файлов JPEG показывает, что в нём находятся маркеры, определяющие сегменты, которые не влияют на декодирование изображение и поэтому игнорируются программами просмотра. Маркеры, не участвующие в JPEG декодировании, представлены в таблице 1.

Таблица 1. Маркеры JPEG, игнорируемые программами просмотра

Маркер	**Байты**	**Длина**	**Назначение**
COM	0xFFFE	переменный размер	Комментарий
SOF9	0xFFC9	переменный размер	Начало фрейма (арифметическое кодирование, обычно не поддерживается)
SOF10	0xFFCA	переменный размер	Обычно не поддерживается
DNL	0xFFDC	переменный размер	Может быть проигнорировано
DAC	0xFFCC	переменный размер	Определение арифметической таблицы, обычно не поддерживается
APP15	0xFFEF	переменный размер	Обычно игнорируется

Каждый маркер имеет два байта для записи размера маркера (включая эти два байта), то есть максимальный размер маркера может быть 65 535 байт. На основании этого можно сделать вывод о том, что максимальное количество дополнительной информации, которое может быть встроено в неиспользуемые маркеры составляет (65535 · 6) = 393 210 байт.

Форматная составляющая комбинированного метода встраивания ЦВЗ предполагает работу как раз с такими сегментами, а её использование при создании комбинированного алгоритма позволит решить задачу сокрытия значительных объёмов.

Разработка алгоритма внедрения ЦВЗ

Алгоритм предполагает преобразование файла из формата JPEG в формат BMP, последующую запись в BMP файл данных по методу наименее значимых бит (НЗБ – LSB – LeastSignificantBits), а затем преобразование файла обратно в JPEG [5, 6]. Для работы алгоритма необходимо решить ряд проблем, основной из которых является то, что JPEG - это формат сжатия с потерями. В результате, если BMP файл преобразовать в JPEG, а затем обратно, то на выходе получится искажённое исходное сообщение, что иллюстрирует невозможность достоверного извлечения встроенной информации. Для устранения этого недостатка был разработан алгоритм компенсации потерь при межформатных преобразованиях вида JPEG-RGB_BMP-JPEG. Последовательность процесса внедрения ЦВЗ представлена на рис. 1.

На первом этапе, при формировании стегопосылки, происходит преобразование исходного изображения формата JPEG в BMP, в котором каждый пиксель изображения кодируется с помощью трёх байт. Каждый из байтов отвечает за один из трёх потоков основных цветов– R (RED – красного), G (GREEN – зеленого) и B (BLUE – синего), формирующих RGB структуру. В результате, благодаря тому, что изображение формата BMP имеет больший размер по сравнению с JPEG, максимальный объём данных для скрытия тоже увеличивается.

На втором этапе происходит встраивание данных в синюю составляющую цвета BMP изображения по методу НЗБ, предполагающему замену младших бит каждого байта потока ис-

ходного изображения на биты из потока скрываемых данных, как это показано на рис.2.

На третьем этапе осуществляется подсчёт разницы межформатных потерь при заданном качестве преобразования BMP→JPEG (quality).

Вычисление минимально необходимого числа бит для кодирования каждого элемента массива осуществляется по формуле:

$$K = \text{ceil}(\log_2(\max(mas) - \min(mas))) + 1, \quad (1)$$

где ceil – функция округления для ближайшего большего целого, mas – массив разностей.

На четвёртом этапе происходит преобразование BMP изображения в JPEG.

На пятом этапе осуществляется запись вспомогательной информации в игнорируемые маркеры JPEG.

Оценка эффективности алгоритма внедрения ЦВЗ

Для подсчета количества данных, которое можно скрыть в одном изображении предлагается использовать соотношение

$$N_{data} = (w \cdot h - const) \cdot \frac{n_{бит}}{i}, \quad (2)$$

где N_{data} - количество бит информации, которую можно скрыть в одном изображении; w, h- ширина и высота исходного изображения в

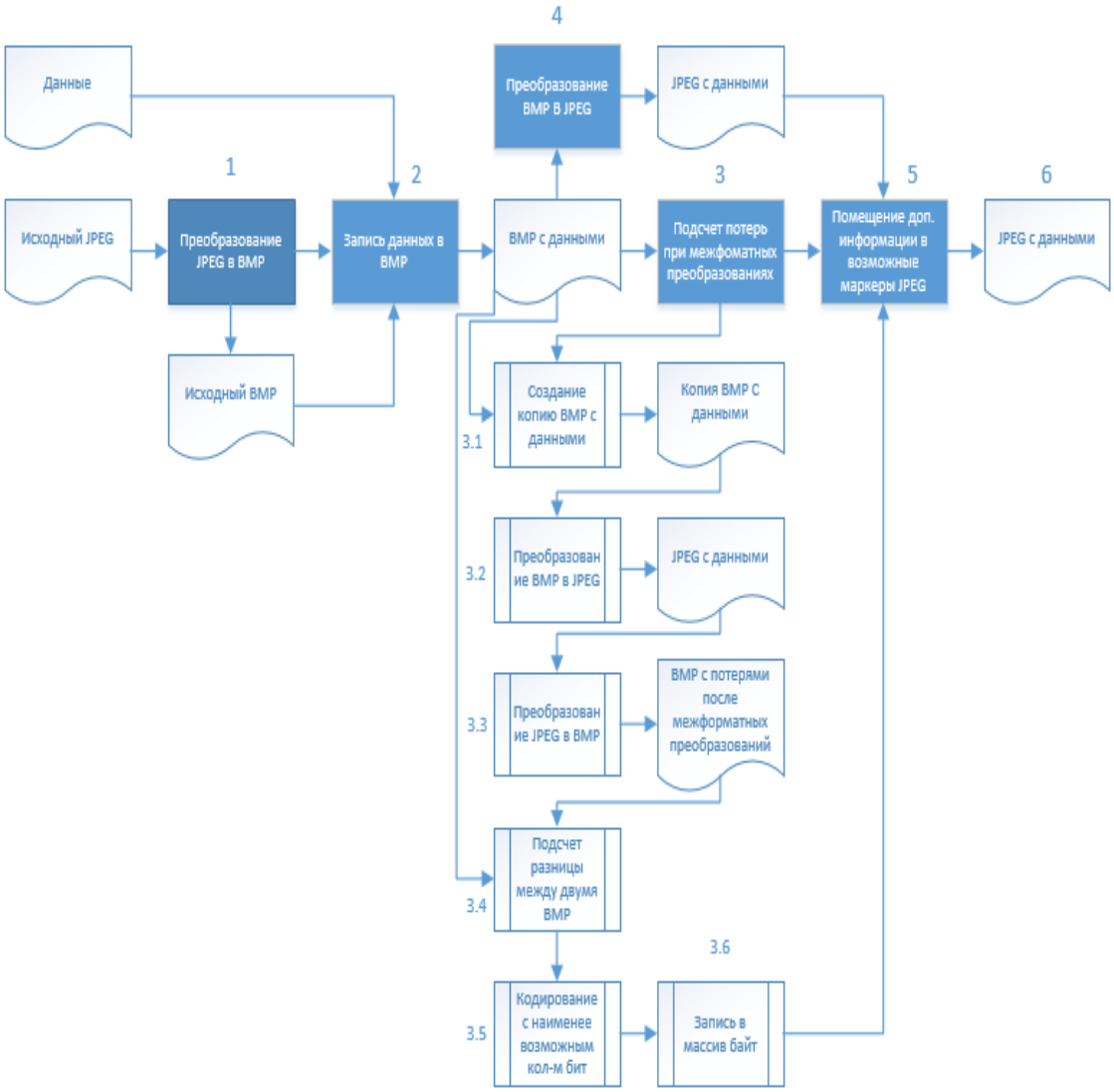

Рис. 1. Алгоритм скрытия данных поэтапно

пикселях; i – интервал скрытия (интервал, через которые берутся пиксели для записи); $n_{бит}$– глубина скрытия (количество используемых бит в потоке для записи информации). Константа const = 14 – характеризует количество пикселей для записи служебной информации.

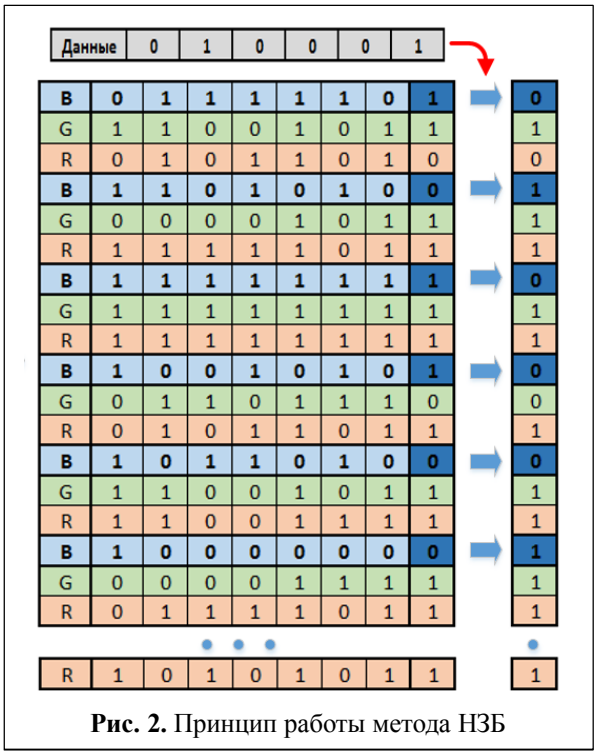

Рис. 2. Принцип работы метода НЗБ

Количественная оценка качества скрытия

На рис.3 приведены количественные оценки качества скрытия стегоданных в изображении для разных параметров качества (quality) преобразования BMP → JPEG. Оценка проводилась с помощью разработанного универсального комплекса для исследования алгоритмов скрытия данных.

Как видно из рис.3, основные параметры оценки ухудшаются в большей степени из-за снижения качества полученного изображения, и лишь в малой степени зависят от встроенных данных.

Атаки на разработанный стеганографический алгоритм

Для оценки устойчивости алгоритма к внешним воздействиям осуществлялось преобразование на промежуточное BMP изображение стегоконтейнера в соответствии с наиболее распространёнными видами атак [8 - 10] с помощью разработанного универсального программного комплекса [11] для исследования стеганографических алгоритмов проводились.

Для оценки устойчивости водяного знака изображение, к воздействию внешних атак содержащее ЦВЗ, подвергалось следующим графическим атакам: обрезка изображения на заданное количество пикселей справа и снизу; шум (аддитивный); шум «Соль и Перец» (salt&pepper) с разными значениями плотности шума; поворот; мозаика; кристаллизация; jpg-сжатие, - все они, кроме последней, осуществлялись в AdobePhotoshop.

Для мозаики и кристаллизации использовались специальные фильтры, параметром которых является размер ячейки: чем больше её размер, тем сильнее искажается изображение (с ЦВЗ, и сам извлечённый ЦВЗ). Выбраны следующие размеры ячеек: для атаки кристаллизация: 3,4,5,6,7; для атаки мозаика: 2,4,6,8,10.

Шум добавлялся при помощи специального фильтра, параметром которого являлось значение в %: 1,5,10,15,20.

Атака «поворот» осуществлялась при помощи AdobePhotoshop, выбирались значения поворота: 1,5,10,15,20 град.

Для оценки качества изображения до и после встраивания ЦВЗ использованы метрики среднеквадратической ошибки (MSE) и отношения сигнал/шум (SNR), характеризующие изменение качества контейнера. Формула для вычисления метрики MSE имеет вид:

$$MSE = \frac{1}{M \cdot N} \cdot \sum_{m=0}^{N-1} \sum_{n=0}^{M-1} (C(m,n) - S(m,n))^2 , \quad (3)$$

где C – пустой контейнер; S – контейнер, содержащий ЦВЗ; m, n - координаты пикселя в

плоскости изображения; M; N – размеры изображения.

(справа) в формате QR кода, размером 276 х 276 с глубиной цвета 1 (монохромное).

Качество	100%			95%			85%			75%		
Глубина скрытия	2											
Плотность скрытия	1											
	ЦВЗ есть	ЦВЗ нет	Разность	ЦВЗ есть	ЦВЗ нет	Разность	ЦВЗ есть	ЦВЗ нет	Разность	ЦВЗ есть	ЦВЗ нет	Разность
Количество бит для кодирования межформатных преобразований	4 бит	-	-	7 бит	-	-	8 бит	-	-	8 бит	-	-
Размер информации в маркерах	19,54 Кбайт	-	-	34,19 Кбайт	-	-	39,07 Кбайт	-	-	39.07 Кбайт	-	-
Размер jpeg файла с иформацией в маркерах	468,91 Кбайт	-	-	195,89 Кбайт	-	-	109,31 Кбайт	-	-	87.52 Кбайт	-	-
Среднеквадратичная ошибка	0,32515	0,11934	0,20581	10,39707	10,28198	0,11509	19,86811	19,82519	0,04292	25,16902	25,13227	0,03675
Норм. среднеквадратичная ошибка	0,00002	0,00001	0,00001	0,00052	0,00052	0	0,001	0,001	0	0,00126	0,00126	0
Отношение сигнал/шум	47,87276	52,22586	-4,35310	32,82443	32,87277	-0,04834	30,01197	30,02137	-0,0094	28,98488	28,99122	-0,00634
Максимальное отношение с/ш	53,01002	57,36312	-4,35310	37,9617	38,01004	-0,04834	35,14924	35,15863	-0,00939	34,12214	34,12849	-0,00635

Рис.3. Сводные данные оценки качества скрытия стегоданных

Для оценки метрики *SNR* удобнее пользоваться логарифмической шкалой:

$$SNR = 10 \cdot \log_{10}\left(\frac{\sum_{m=0}^{N-1}\sum_{n=0}^{M-1} C(m,n)^2}{\sum_{m=0}^{N-1}\sum_{n=0}^{M-1}\left(C(m,n)-S(m,n)\right)^2}\right) =$$

$$= 20 \cdot \log_{10}\left(\frac{\sum_{m=0}^{N-1}\sum_{n=0}^{M-1} C(m,n)}{\sum_{m=0}^{N-1}\sum_{n=0}^{M-1}\left(C(m,n)-S(m,n)\right)}\right) \quad (4)$$

Исходные данные, характеризующие перечисленные атаки, представлены на рис. 4: bmp изображение (слева), размером 512 х 512 с глубиной цвета 24 бит, содержащее в себе ЦВЗ

Данные внедрены с глубиной скрытия 2 и плотностью скрытия 1.

На основании проведённых расчётов, а также на основе визуального анализа можно сделать вывод о том, что атака «обрезка изображения» практически полностью разрушает внедрённый ЦВЗ в том случае, если изображения обрезается с боков. Если изображение обрезается только снизу, это может не повлиять на внедренные данные, поскольку данные начинают внедряться последовательно сверху–вниз, справа – налево.

Рис. 4. Видеоконтейнер (слева) и извлеченный ЦВЗ (справа) до проведения атак

На рис.5 и рис.6 представлены результаты некоторых атак.

Атака «зашумление гауссовым белым шумом» полностью разрушает внедренный ЦВЗ. Результаты извлечения ЦВЗ из картинки, подвергшейся зашумлению с помощью шума «соль и перец» с плотностью 0,05; 0,1; 0,15; 0,2; 0,25 показали, что подобные атаки практически не разрушают внедрённый ЦВЗ, о чём свидетель-

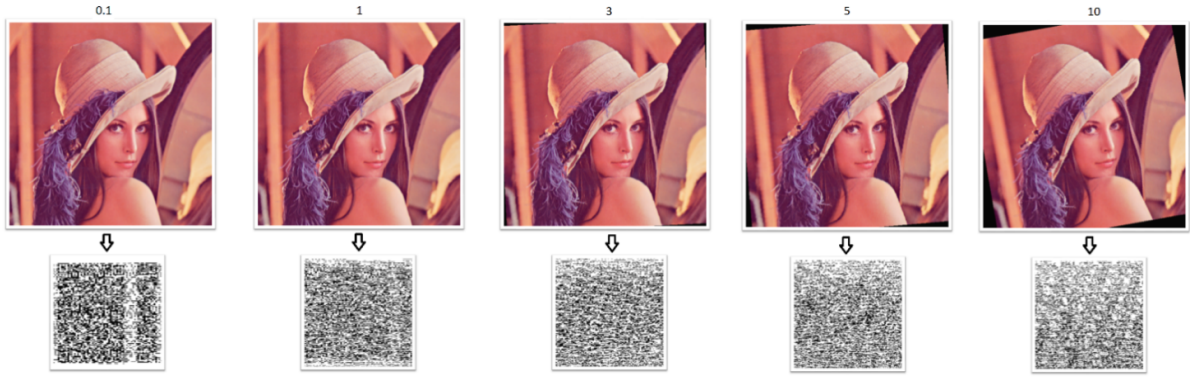

Рис. 5. Результаты извлечения ЦВЗ из неподвижного изображения, подвергшегося повороту на 0,1˚; 1˚; 3˚; 5˚; 10˚

ствуют гораздо меньшие ошибки при оценке извлечённого ЦВЗ. Тем не менее восстановить закодированное в QR коде сообщение с помощью программ для считывания не удалось.

Таким образом, в условиях воздействия внешних атак сохранить высокое качество извлечения ЦВЗ можно только при условии использования дополнительного помехоустойчивого кодирования.

Выводы

Разработан стеганографический алгоритм скрытия большого объёма данных в неподвижных изображениях, эффективность которого сравнима с встраиванием данных по алгоритму LSB в изображения формата BMP.

Реализация представленного стеганоалгоритма в виде программного обеспечения показала, что количество битов, пригодных для встраивания, зависит от исходного JPEG-файла (характер зависимости определяется JPEG-преобразованием) и линейно зависит от числа LSB, используемых для встраивания. Сохранение сообщения в игнорируемых сегментах файла JPEG позволяет скрыть большое количество байтов сообщения, однако факт такого сокрытия легко обнаружить.

В результате оценки качества внедряемой информации сделан вывод о возможности встраивания большого объёма стегоданных в изображения без изменения визуальной составляющей изображения – то есть не заметной человеческому глазу. Показана низкая устойчивость разработанного алгоритма противостоять атакам вида «поворот изображения» и аддитивному шуму.

Показано, что разработанный алгоритм может быть использован для скрытия и обмена данными в условиях воздействия внешних атак при условии использования дополнительного

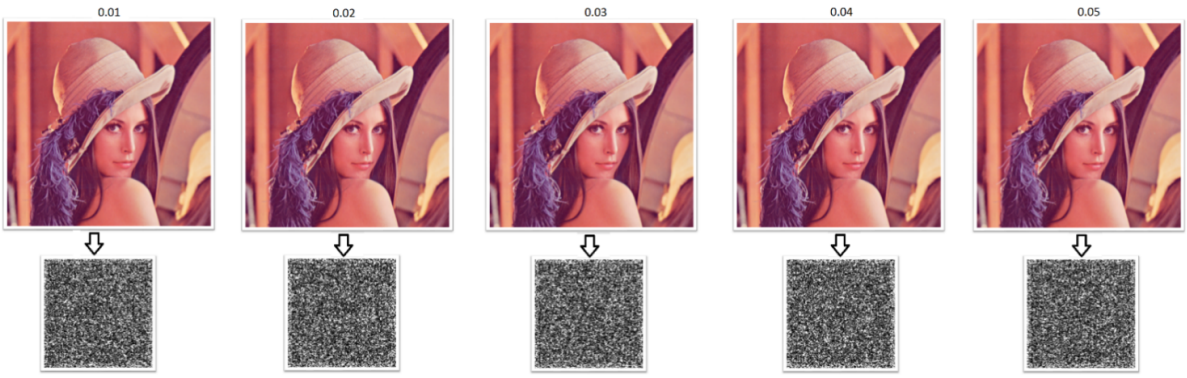

Рис. 6. Результаты извлечения ЦВЗ из неподвижного изображения, подвергшегося зашумлению с помощью гауссова белого шума с дисперсией 0,01; 0,02; 0,03; 0,04; 0,05

помехоустойчивого кодирования

Литература

1. Шелухин О.И., Канаев С.Д. Стеганография. Алгоритмы и программная реализация. Под.редакцией проф. Шелухина О.И.– М.: Горячая линия – Телеком, 2016. – 616 с.

2. Трегулов Т.С. Скрытая передача данных в цифровых изображениях формата JPEG // Научные труды SWorld. 2012. Т. 4. № 2. С. 88-91.

3. Чекасин А.И., Стрельцов Е.В., Поташникова А.В., Учет особенностей формата JPEG при стеганографическом кодировании // Известия высших учебных заведений. Электроника. 2009. № 4 (78). С. 61-65.

4. Cox I.J. Digital watermarking and steganography / M. Miller, J. Bloom, J. Fridrich. – San Francisco: Morgan Kaufmann Publishing, 2008. – 624 p.

5. Кувшинов С.С. Методы и алгоритмы сокрытия больших объемов данных на основе стеганографии: автореферат диссертации на соискание ученой степени кандидата технических наук // Санкт-Петербургский государственный университет информационных технологий, механики и оптики. Санкт-Петербург, 2010

6. Коробейников А.Г., Кувшинов С.С., Блинов С.Ю., Лейман А.В., Нестеров С.И., Разработка стеганоалгоритма на базе форматных и пространственных принципов сокрытия данных // Научно-технический вестник информационных технологий, механики и оптики. 2012. № 1 (77). С. 116-119.

7. Грибунин В.Г., Оков И.Н., Туринцев И.В. Цифровая стеганография. – М.: Солон-Пресс, 2009.– 272 с.

8. Kim, B.S. Robust digital image watermarking method against geometrical attacks / B.S. Kim, J.G. Choi, C.H. Park, J.U. Won, D.M. Kwak, S.K. Oh, C.R. Koh, K.H. Park //Real-Time Imaging. – 2003. – Vol. 9.– P. 139-149.

9. Johnson N.F., Duric Z., Jajodia S. Information Hiding: Steganography and Watermarking. – Attacks and Countermeasures. – Kluwer Academic Publishers. 2001. 160 p.

10. Oostveen J., Kalker T., Linnart J.-P.z. Optimal detection of multiplicative watermarks // Proceedings of the 10th European Signal Processing Conference, EUSIPCO'00, Tampere, Finland, September 2000.

11. Шелухин О.И., Кажемский М.А. Разработка программного комплекса для исследования стеганографических алгоритмов скрытия информации/ Труды Международный форум информатизации (МФИ-2015), 24 ноября 2015 г. С. 161.

Поступила 23 ноября 2016 г.

English

The combined algorithm of digital water marks embedding to graphic format images for hidden information transmission

Oleg Ivanovich Shelukhin – Professor, Doc.Tech.Sci. Head of the Department "Information Security", Moscow Technical University of Communication and Information Science.

E-mail: heluhin@mail.ru.

Address: Moscow, Aviamotornaya St., 8a.

Mikhail Aleksandrovich Smychyok – Cand.Tech.Sci., Principal Discipline Engineer Communication Networks Design Department JSC Giprogaztsentr.

E-mail: m-smychek@mail.ru.

Address: Nizhny Novgorod, Alekseevskaya St., 26.

Abstract: The JPEG file structure analysis shows that there are markers in its defining segments which are not involved in JPEG conversion and do not affect image visualization, and therefore - are ignored by viewing programs. Thus, we can draw a conclusion about possibility to develop algorithm for embedding large amount of data to JPEG and BMP graphic images with using a format method. JPEG format was taken as a basic format for research and development as the most widespread standard for digital graphics daily use, especially digital photos. JPEG data stream is converted into BMP data stream at the first stage. Thus, the stream size increases due to the change of data coding principle of color properties in image sections. To minimize the amount of change only the low bit of such byte is used by default in algorithm that reduces probability of change detection to a minimum. Further the modes are implemented that enable to use replacement methods of the least-significant bits (LSB). The developed mechanism of loss compensation in cross-format conversions is intended for successful data extraction in the receiving end. Implementation of the presented steganographic algorithm in the form of the software demonstrated that the quantity of the bits suitable for embedding depends on the initial JPEG file (dependence type is defined by JPEG conversion), and linearly depends on LSB quantity used for embedding. Saving the message in the ignored segments of the JPEG file enables to hide a large number of the message bytes, however it is easy to find the fact of such concealment. As a result of quality evaluation of the

embedded information, the conclusion is drawn on possibility of embedding large amount of steganographic data to images without changing the image visual component - that is not seen by human eye. Low stability of the developed algorithm is shown to stand against actions as image rotation and additive noise. It is pointed out that the developed algorithm can be used for concealment and data exchange in the environment of external attacks if only additional noise resistant coding is applied.

Key words: embedding information, image spatial domain, digital water mark, video container, attack, quality evaluation.

References

1. Shelukhin O.I., Kanayev S.D. Steganography. Algorithms and software implementation. Ed. by Prof. Shelukhin O. I. - M.: Goryachaya liniya – Telekom, 2016. - 616 p.

2. Tregulov T.S., Hidden data transmission in JPEG digital images. - SWorld Scientific works. 2012. V. 4. No. 2. P. 88-91.

3. Chekasin A.I., Streltsov E.V., Potashnikov A.V. Sagittariuses., JPEG feature-oriented steganographic coding.- Izvestiya vysshikh uchebnykh zavedeny. Elektronika. 2009. No. 4 (78). P. 61-65.

4. Kuvshinov S.S. Steganography-based methods and algorithms of hiding large amounts of data: the abstract of the thesis on competition for the degree of Candidate of Technical Sciences. - St. Petersburg State University of Information Technologies, Mechanics and Optics. St. Petersburg, 2010.

5. Korobeinikov A.G., Kuvshinov S.S., Blinov S.Yu., Leyman A.V., Nesterov S. I., Steganography algorithm development based on format and space data hiding principles. - Nauchno-tekhnichesky vestnik informatsionnykh tekhnology, mekhaniki i optiki. 2012. No. 1 (77). P. 116-119.

6. Gribunin V.G., Okov I.N., Turintsev I.V. Digital steganography. - M.: Solon-Press, 2009. - 272 p.

7. Shelukhin O.I., Kazhemskiy M.A. Software development of steganography algorithm research for hiding information. Works. International Forum of Infromation System Development (MFI-2015), November 24, 2015, p. 161.