

Выявление стеганографических вставок типа LSB-замещения в растровых изображениях

Д.Э. Вильховский
vilkhovskiy@gmail.com

С.В. Белим
belimsv@omsu.ru

Омский государственный университет им. Ф.М. Достоевского, Омск, Россия

Аннотация

В статье предложен метод выявления размера и расположения стеганографической вставки типа LSB-замещение. Предложенный метод основан на анализе нулевого слоя битов, к которому применяется алгоритм таксономии. Метод позволяет эффективно выявлять вставки на искусственных изображениях. Для фотографических объектов проводится предварительная обработка, основанная на кластеризации изображения. К каждому кластеру применяется линейное преобразование, коэффициенты которого вычисляются из параметров кластера.

Введение

Наиболее известным методом встраивания стеганографических вставок на сегодняшний день является подмена наименее значащих бит (LSB-замещение) [1]. Данный метод основан на том, что замена от одного до четырех младших бит в байтах цветового представления пикселей изображения остаётся практически незаметной для человеческого глаза. Наибольшей возможностью для скрытия информации обладает синяя компонента, что обусловлено строением сетчатки глаза. На сегодняшний день разработано достаточно много алгоритмов встраивания информации в изображения, аудио- и видеопотоки, однако метод LSB-замещения, являясь исторически первым, имеет широкое распространение. Все методы скрытия информации ориентированы на такое преобразование изображения, которое не заметно визуально. В связи с чем актуальной является задача построения алгоритмов анализа изображений на наличие стеганографических вставок.

На сегодняшний день все существующие алгоритмы ориентированы на определение самого факта присутствия или отсутствия стеганографической вставки в изображении. В работе [2, 3] приведен метод статистического анализа Хи-квадрат, основанный на предположении о случайном распределении младших битов цветовых байтов изображения. Данный метод дает хорошие результаты при равномерном заполнении контейнера и слабо применим при случайном выборе пикселей для замены младших битов. В статье [4] стеганоанализ проводится на основе сравнения младших битов в соседних байтах с использованием формализма цепей Маркова. В статье [5] предложен метод обнаружения стеганографической вставки на основе использования искусственных нейронных сетей. Показано, что при достаточно большом объеме обучающей выборки нейронная сеть способна определять наличие вставки с ошибкой не превышающей 15%. Все известные на сегодняшний день методы стеганоанализа метода LSB-замещения эффективны при заполнении

Copyright © by the paper's authors. Copying permitted for private and academic purposes.

In: Sergey V. Belim, Nadezda F. Bogachenko (eds.): Proceedings of the Workshop on Data, Modeling and Security 2017 (DMS-2017), Omsk, Russia, October 2017, published at <http://ceur-ws.org>

стегакоде не менее чем на 50% [6]. В статье [7] предложен метод обнаружения встроенной информации, основанный на алгоритмах сжатия информации. Основная идея метода состоит в том, что случайные данные сжимаются слабее, чем упорядоченные. Данный подход позволяет с высокой точностью определять наличие стеганографической вставки при заполнении контейнера от 40%. Данный метод получил развитие в работе [8] на основе использования предварительной обработки изображения, что позволило применять его при значительно меньших заполнениях контейнера.

Следует отметить, что на сегодняшний день отсутствуют алгоритмы, определяющие в байты, в которых произошла подмена младшего бита. По постановке данная задача близка к проблеме выявления пикселей, поврежденных импульсным шумом. Для импульсного шума характерно изменение цвета произвольно выбранного байта на случайно выбранную величину. Однако задача поиска встроенного сообщения является более сложной так как величина изменения составляет всего один бит. На сегодняшний день существует несколько методов обнаружения поврежденных пикселей. Прежде всего необходимо выделить метод SD-ROM [9], который, являясь исторически первым, послужил основой для ряда алгоритмов [10, 11]. Также существуют иные подходы, основанные на поиске ассоциативных правил [12, 13], методе кластеризации [14] и методе анализа иерархий [15].

В данной статье предложен алгоритм выявления пикселей изображения, в которых произведена подмена наименее значащего бита при стеганографическом встраивании сообщения, на основе автоматического анализа нулевого слоя.

1 Постановка задачи

Будем анализировать изображения, в которых может быть встроена информация в виде стеганографической вставки в младший бит синей компоненты. Анализ именно синей компоненты обусловлен тем, что встраивание в неё наименее заметно визуально, в связи с чем рекомендуется ее использование для скрытия сообщений. Аналогичным образом могут быть проанализированы и другие компоненты без потери общности предлагаемого метода. Будем исходить из двух предположений. Во-первых, достоверно неизвестно есть ли стеганографическая вставка или нет. Во-вторых, будем считать, что стеганографическая вставка заполняет некоторую прямоугольную область, размеры и положение которых неизвестны. Задачей ставится не только определение наличия стеганографической вставки, но и области, в которую произведено встраивание. Второе предположение существенно осложняет задачу, так как возможна ситуация, при которой заменены все младшие пиксели синей компоненты. Будем решать задачу, исходя из предположения о неполной подмене нулевого слоя.

Нулевой слой представляет собой матрицу из нулей и единиц. Закономерности распределения нулевых и единичных значений нулевого слоя без встраивания обусловлены структурой изображения. Встраивание сообщения вносит изменения в нулевой слой, изменяя плотность распределения единичных значений. На рисунке 1 представлено изображение, его нулевой слой без встраивания и нулевой слой с встроенным сообщением.

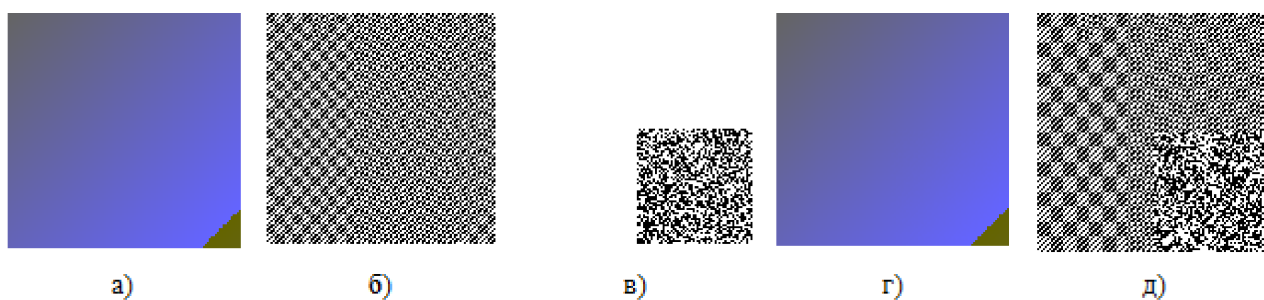


Рис. 1: Сравнение нулевых слоев: а) исходное изображение, б) нулевой слой исходного изображения, в) карта встраиваемых пикселей, г) изображение с встроенным сообщением, д) нулевой слой изображения с встроенным сообщением

Как видно из рисунка 1 наличие встроенного сообщения может быть обнаружено визуально из анализа нулевого слоя. Поставим себе целью автоматическое определение области встраивания.

2 Алгоритм выделения области встраивания

Для выделения области встраивания используем алгоритм, построенный на основе алгоритма таксономии FOREL [16]. В своем классическом виде FOREL объединяет точки в таксоны, лежащие внутри окружности. В нашем случае мы будем строить таксоны прямоугольной формы.

Введем показатель плотности единичных значений p . Если в некоторой области изображения содержится N пикселей и N_1 из них имеет единичное значение, показатель плотности единичных значений $p = N_1/N$.

Будем искать прямоугольные области, имеющие плотность единичных значений заданной величины p_0 . Также в качестве входного параметра алгоритма зададим параметр R_0 , определяющий начальный размер таксона. Алгоритм состоит из следующих шагов:

Шаг 1. Выбираем начальное значение размера таксона $R = R_0$.

Шаг 2. Выбираем случайным образом точку с координатами (x_1, y_1) , которая играет центра таксона. Строим квадрат, левый верхний угол которого имеет координаты $(x_1 - R, y_1 - R)$, а правый нижний угол координаты $(x_1 + R, y_1 + R)$.

Шаг 3. Ищем координаты центра масс точек, лежащих внутри построенного квадрата (x_2, y_2) .

Шаг 4. Если точки (x_1, y_1) и (x_2, y_2) совпадают, то переходим к Шагу 5, иначе $x_1 = x_2$, $y_1 = y_2$ и переходим к Шагу 2.

Шаг 5. Вычисляем показатель плотности единичных значений p .

Шаг 6. Если $p > p_0$, то $R := 1.1R$ и переходим к Шагу 3.

Шаг 7. Если $p < p_0$, то $R := 0.9R$ и переходим к Шагу 3.

Шаг 8. Если $p = p_0$, то переходим к Шагу 2.

Алгоритм выполняется пока все точки нулевого слоя не будут объединены в некоторые таксоны.

В качестве областей, в которые может быть встроено сообщение, выбираем таксоны, размер которых не меньше, чем 10% размеров исходного изображения. Результаты работы данного алгоритма для изображения с равномерной заливкой представлены на рисунке 2.

Как видно из рисунка 2 в результате работы алгоритма на искусственном изображении с равномерной заливкой область встраивания определяется достаточно точно. Однако уже для искусственных изображений с градиентной заливкой возникают сложности, так как непрерывное изменение цвета изображения в целом проявляется как полосы одинаковых значений на нулевом слое (рисунок 1,б). Данные полосы могут быть устранены с помощью предварительной обработки изображения.

3 Алгоритм предварительной обработки изображений

Как было показано выше нулевой слой изображения с градиентной заливкой представляет собой полосы из нулей и единиц. Используем линейное преобразование:

$$d(x, y) = ax + by - e,$$

где $a = c(x + 1, y) - c(x, y)$, $b = c(x, y + 1) - c(x, y)$, $c(x, y)$ – значения цвета пикселя, расположенного в точке с координатами (x, y) . Определим e , как минимальное значение $c(x, y)$ на множестве всех пикселей изображения. В случае, когда цвет изображения представляет собой заливку с постоянным градиентом, функция $d(x, y)$ будет иметь постоянное значение ($d(x, y) = const$). Алгоритм принятия решений об изменении пикселя может быть применен к функции $d(x, y)$, а не к функции $c(x, y)$.

Для применения данного линейного преобразования к фотографическим изображениям необходимо определить области градиентной заливки. Будем вычислять вторые производные от функции $c(x, y)$ и выявлять области, в которых они имеют нулевое значение. В связи с тем, что области идеальной градиентной заливки на фотографических изображениях встречаются крайне редко будем требовать выполнение трех более «мягких» условий:

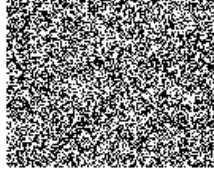
$$\left| \frac{\partial^2 c(x, y)}{\partial x^2} \right| \leq 2, \quad \left| \frac{\partial^2 c(x, y)}{\partial y^2} \right| \leq 2, \quad \left| \frac{\partial^2 c(x, y)}{\partial x \partial y} \right| \leq 2.$$

Не строгое неравенство вместо равенства нулю введено для того, чтобы учесть небольшие отклонения от градиентной заливки и не потерять встроены биты.

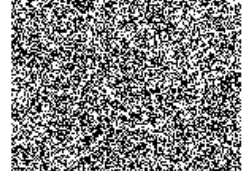
После выявления связанных областей, удовлетворяющих условиям на вторые производные, необходимо определить коэффициенты функции $d(x, y)$. Для их нахождения был использован метод наименьших квадратов. После чего вычислялось значение функции $d(x, y)$, к которому применялся алгоритм поиска областей встраивания сообщений.



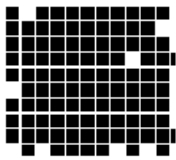
а)



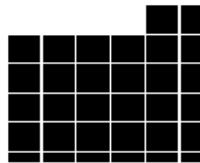
б)



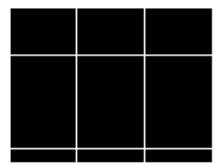
в)



г)



д)



е)



ж)

Рис. 2: Результаты работы алгоритма по автоматическому выделению области встраивания для изображения с равномерной заливкой: а) изображение со встроенной вставкой, б) карта встроенных пикселей, в) нулевой слой, г) автоматически выделенная область встраивания при $R_0 = 15$, д) автоматически выделенная область встраивания при $R_0 = 30$, е) автоматически выделенная область встраивания при $R_0 = 50$, ж) автоматически выделенная область встраивания при $R_0 = 100$

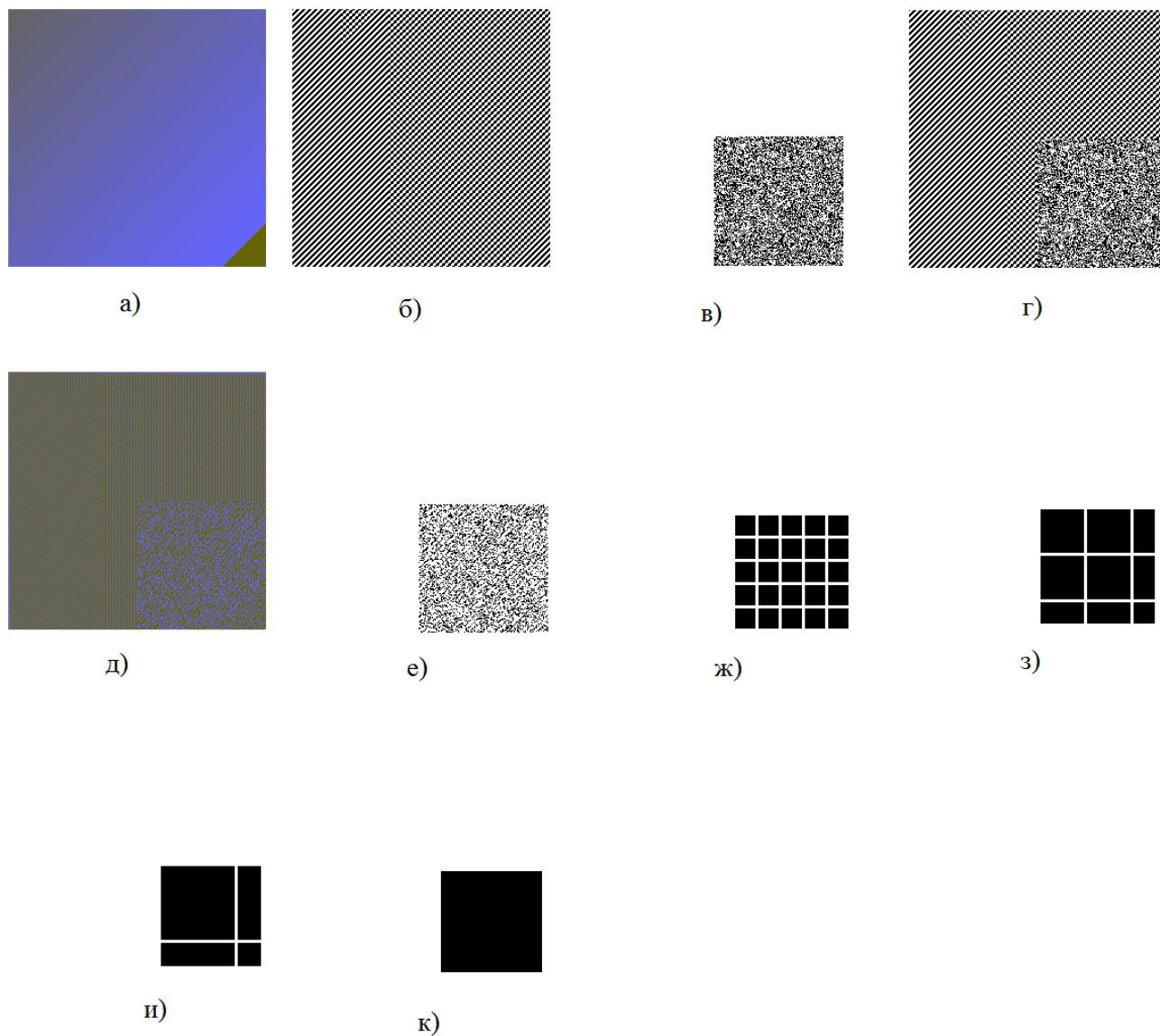


Рис. 3: Результаты работы алгоритма по автоматическому выделению области встраивания для изображения с градиентной заливкой: а) исходное изображение, б) нулевой слой исходного изображения, в) карта встроенных пикселей, г) нулевой слой со встроенным сообщением, д) изображение после предварительной обработки, е) нулевой слой после предварительной обработки изображения, ж) автоматически выделенная область встраивания при $R_0 = 15$, з) автоматически выделенная область встраивания при $R_0 = 30$, и) автоматически выделенная область встраивания при $R_0 = 50$, к) автоматически выделенная область встраивания при $R_0 = 100$

4 Компьютерный эксперимент и результаты

Компьютерный эксперимент проводился для различных цветных изображений, как искусственных, так и фотографических. Встраиваемое сообщение представляло собой текстовую строку, представленную в виде последовательности битов. Встраивание производилось в синюю компоненту, как наименее различимую человеческим глазом. Встраиваемые биты заполняли прямоугольную область. В качестве задачи стегоанализа ставилось определение области, в которой расположены подмененные биты.

Первоначально алгоритм тестировался на прямоугольном искусственном изображении с градиентной заливкой. Результаты представлены на рисунке 3.

Как видно из рисунка 3 алгоритм предварительной обработки изображений позволяет обрабатывать изображения с градиентной заливкой также эффективно, как и одноцветные изображения. Как равномерной так и градиентной заливки на результатах работы алгоритма сказывается выбор начального размера таксона R_0 .

Список литературы

- [1] E. Adelson. *Digital Signal Encoding and Decoding Apparatus*. U.S. Patent. No. 4,939,515, 1990.
- [2] N. Provos, P. Honeyman. Detecting steganographic content on the internet. *Technical Report CITI 01-1a*, University of Michigan, 2001.
- [3] A. Westfeld, A. Pfitzmann. Attacks on Steganographic Systems: Breaking the Steganographic Utilities EzStego, Jsteg, Steganos and STools- and Some Lessons Learned. *3rd International Workshop on Information Hiding*, 2000.
- [4] V.A. Golub, M.A. Dryuchenko. Steganographic information detection in JPEG files with the help of complex usage of several stego-attacks. *Infocommunication Technologies*, 7(1):44–50, 2009.
- [5] A.Zh. Abdenov, L.S. Leonov. Ispolzovaniye neyronnykh setey v slepykh metodakh obnaruzheniya vstroynnoy steganograficheskoy informatsii v tsifrovyykh izobrazheniyakh. *Polzunovskiy vestnik*, 2:221–225, 2010.
- [6] A. Westfeld, A. Pfitzmann. Attacks on Steganographic Systems. Breaking the Steganographic Utilities EzStego, Jsteg, Steganos and S-Tools - and Some Lessons Learned. *Lecture Notes in Computer Science*, 1768:61–75, 2000.
- [7] M.Yu. Zhilkin. Stegoanaliz graficheskikh dannykh v razlichnykh formatakh. *Doklady TUSURa*, 2(18):63–64, 2008.
- [8] V. A. Monarev. Sdvigovyy metod obnaruzheniya skrytoy informatsii. *Vestnik SibGUTI*, 4:62–68, 2012.
- [9] E. Abreu, M. Lightstone, S.K. Mitra, S.K. Arakawa. A new efficient approach for the removal of impulse noise from highly corrupted images. *IEEE Transactions on Image Processing, IEEE Transactions on*, 5:1012–1025, 1996.
- [10] R. Garnett, T. Huegerich, C. Chui, W. He. A Universal Noise Removal Algorithm with an Impulse Detector. *IEEE Trans Image Process*, 14(11):1747–1754, 2012.
- [11] S.V. Sorokin, M.A. Shcherbakov. Realizatsiya SD-ROM filtra na osnove kontseptsii nechekoy logiki. *Izvestiya vysshikh uchebnykh zavedeniy. Povolzhskiy region*, 3:56–65, 2007.
- [12] S.V. Belim, A.O. Mayorov-Zilbernegel. Algorithm for Searching the Broken Pixels and Eliminating Impulse Noise in Images Using a Method of Association Rules. *Science and Education of the Bauman MSTU*, 12:716–737, 2014. URL: <http://technomag.bmstu.ru/doc/744983.html>.
- [13] S.V. Belim, A.O. Mayorov-Zilbernegel. Image Restoration With Static Gaps On The Basis Of Association Rules. *Herald of computer and information technologies*, 12:18–23, 2014.
- [14] S.V. Belim, P.E. Kutlunin. Impulse noise detection in image using a clustering algorithm. *Herald of computer and information technologies*, 3:3–10, 2016.

- [15] S.V. Belim, S.A. Seliverstov. The Analytic Hierarchy Method-Based Algorithm for Restoring Broken Pixels on the Noisy Images. *Science and Education of the Bauman MSTU*, 11:521–534, 2014. URL: <http://technomag.bmstu.ru/doc/742145.html>.
- [16] N.G. Zagoruyko. *Prikladnyye metody analiza dannykh i znaniy*. Novosibirsk, IM SO RAN, 1999.

Detection the Stego-Insertions Like LSB-Substitution in Bitmap Images

Danil E. Vilkhovskiy, Sergey V. Belim

In article the detection method of the size and position of a stego-insertion type LSB-substitution is suggested. The suggested method is based on the analysis of a zero bits layer to which the taxonomy algorithm is applied. The method allows to reveal effectively insertions in artificial images. For photographic objects the preliminary processing based on image clusterings is carried out. The linear transformation which coefficients are calculated from cluster parameters is applied to each cluster.