УДК 123.456.789

Построение стегосистемы для растровых изображений на основе стохастической модуляции с учётом статистики младших бит

Е. Ю. Елтышева

В данной статье предлагается метод скрытия информации в младших битах растровых изображений, который был дополнен и улучшен с помощью применения известной схемы ±1 встраивания. В отличие от ±1 метода, наша схема кодирует внедряемое сообщение в соответствии с оценками вероятностей появления значений 0 и 1 в младших битах с целью минимального искажения статистики, свойственной «пустому» файлу-контейнеру. Способ встраивания бит в матрицу изображения аналогичен способу ±1. Проводится стегоанализ разработанного алгоритма и демонстрируется его преимущество в сравнении с первоначальным ±1 методом внедрения.

Ключевые слова: стеганография, стегоанализ, LSB-методы, статистическая модель, арифметическое кодирование, идеальные стегосистемы, метод ±1.

1. Введение

Главной задачей стеганографии является организация передачи секретных данных таким образом, чтобы были скрыты и содержание сообщения, и сам факт его передачи. Для этого используют некоторые файлы, называемые контейнерами, в которые встраивается сообщение так, чтобы никто, кроме отправителя, не имел доступа одновременно к пустому и заполненному контейнеру. В качестве контейнеров могут использоваться различные файлы, не вызывающие подозрения при передаче. Например, аудиофайлы, видеофайлы, различные исполняемые файлы и цифровые фотографии. В данной работе в качестве контейнера мы используем 24-битовое растровое изображение в системе цветности RGB. Изображение представляется в виде матрицы, каждый пиксель которой соответствует видимой точке и задаётся значениями яркости трёх составляющих: красного (R), зелёного (G) и синего (B) цвета. Яркость каждой составляющей представлена 8-битным числом и может изменяться в диапазоне от 0 до 255. Скрываемая информация записывается в наименее значимые (младшие) биты цветовых составляющих, поэтому предлагаемый метод относится к широко известному классу LSB-методов. Таким образом, эмпирическая ёмкость таких методов составляет 1/8 от размера изображения.

В совершенной стегосистеме [1] пустые и заполненные контейнеры статистически неразличимы, что делает задачу выявления скрытого сообщения достаточно трудной. В работе [2] нами была предложена схема скрытия информации в младших битах растрового изображения, использующая кодирование внедряемого сообщения в соответствии с оценками вероятностей появления значений 0 и 1 в младших битах при минимальном искажении статистики, свойственной пустому контейнеру. В данной работе предлагается соединить такую схему с известным методом ±1 встраивания [3]. Способ внедрения сообщения по типу ±1 позволит усовершенствовать схему в плане стойкости к стегоанализу, в то же время использование статистической модели приблизит схему к совершенной.

2. Описание алгоритма

Изложим основные этапы построения стегосистемы, предложенной в работе [2]. На первом этапе изображение необходимо разделить на две части: одну часть для оценки статистики младших бит и вторую для внедрения информации с учётом этой статистики. Второй этап заключается в построении самой статистической модели. Вследствие специфики выбранного вида контейнеров, была построена адаптивная модель, перестраивающая оценки вероятностей при переходе от одного фрагмента изображения к другому. Оценка вероятностей производилась в окне размером 16×16 пикселей со сдвигом на 8 пикселей, по наиболее существенным элементам контекста, в наибольшей степени статистически связанных с младшими битами. На третьем этапе происходит последовательное формирование младших бит путём арифметического декодирования внедряемого сообщения с распределением вероятностей, определяемым соответствующим контекстом. Зашифрованное сообщение в данном случае рассматривается как код, построенный ранее соответствующим арифметическим кодером. Для извлечения зашифрованного сообщения мы применяем арифметическое кодирование, которому указываются те же самые распределения вероятностей, что и декодеру.

В нашей модели оценка статистики младших бит производилась по формуле Кричевского – Трофимова. То есть, если в некотором контексте X значение 0 встретилось $n_{X,0}$ раз, а 1 – $n_{X,1}$ раз, то можно получить оценку условных вероятностей по формуле:

$$P(0 \mid X) = \frac{2n_{X,0} + 1}{2(n_{X,0} + n_{X,1}) + 2}, \quad P(1 \mid X) = 1 - P(0 \mid X).$$

Из этого следует, что в ситуации, когда в данном контексте встречалось только одно значение, вероятность появления второго будет очень маленькой, но не нулевой. Таким образом, при внедрении сообщения, в область младших бит, состоящую только из нулей, может попасть единица. Наглядно данный пример можно увидеть на рис. 1.

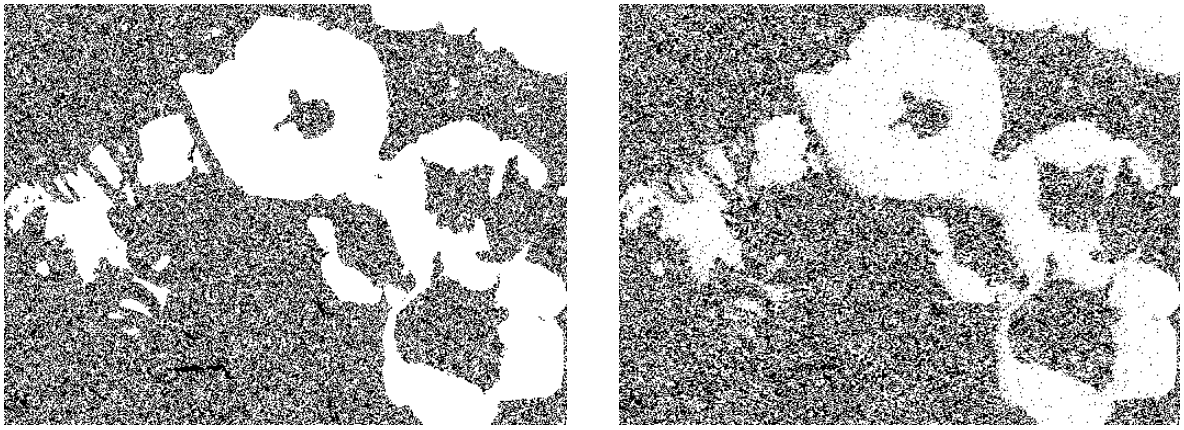

Рис.1. Распределение младших бит до (слева) и после (справа) внедрения информации по описанной модели

Чтобы избежать подобных искажений, необходимо пропускать пиксели, для которых возникает такая ситуация. В результате данной модификации, наша предыдущая модель несколько улучшилась, что можно видеть на рис. 2.

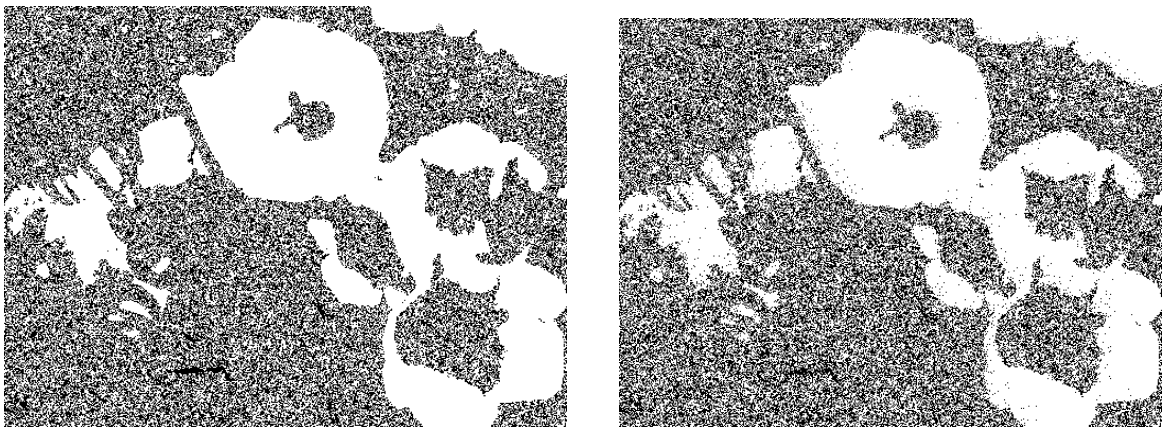

Рис. 2. Распределение младших бит до (слева) и после (справа) внедрения информации с учётом улучшения схемы

Учитывая последнюю модификацию, очевидно, что фактическая средняя ёмкость контейнера уменьшится. По исследованиям, проведённым на наборе из 800 bmp-файлов, средняя фактическая ёмкость контейнеров составила 37 % и 24 % соответственно для первоначальной и модифицированной версии стегосистемы.

Таким образом, мы имеем стеганографический метод, позволяющий в некоторой степени сохранить исходное распределение символов контейнера. Результаты стегоанализа, приведённые в работе [2], показывают, что метод нуждается в дальнейших улучшениях. С этой целью в данной работе предлагается объединить описанную стегосистему с известным методом ±1 встраивания, который является частным случаем стохастической модуляции [3]. Такой метод позволяет встроить сообщение в матрицу растровых изображений путём добавления лёгкого шума с определённым случайным распределением вероятностей. Данный механизм встраивания маскирует искажения контейнера под некоторый шум, полученный вследствие специфики фотокамеры. Недостатком данной схемы является то, что распределение вероятностей младших бит практически полностью переходит в случайное. Это можно легко заметить, визуально сравнив распределение взятых изолированно младших бит до и после внедрения (рис. 3):

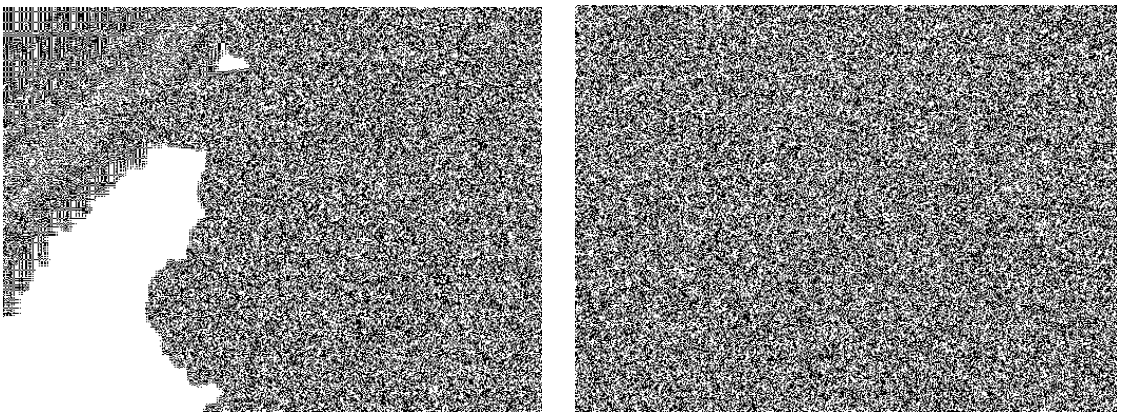

Рис. 3. Распределение младших бит до (слева) и после (справа) внедрения сообщения ±1 методом на 100 %

Достаточно сравнить некоторое множество файлов, заполненных методом ±1, чтобы заподозрить тенденцию «затирания» рисунка младших бит. В контрастных фотографиях с яркими областями особенно явно должно просматриваться распределение, как на рис. 1 слева. Однако для того, чтобы сравнить результаты применения стегосистемы ±1 и нашей модели, необходимо для обоих методов

установить равную фактическую ёмкость контейнеров. На рис. 4 приведены изображения младших бит для пустого контейнера (слева) и контейнера, заполненного на 24 % от общей ёмкости для методов LSB с помощью схемы ±1.

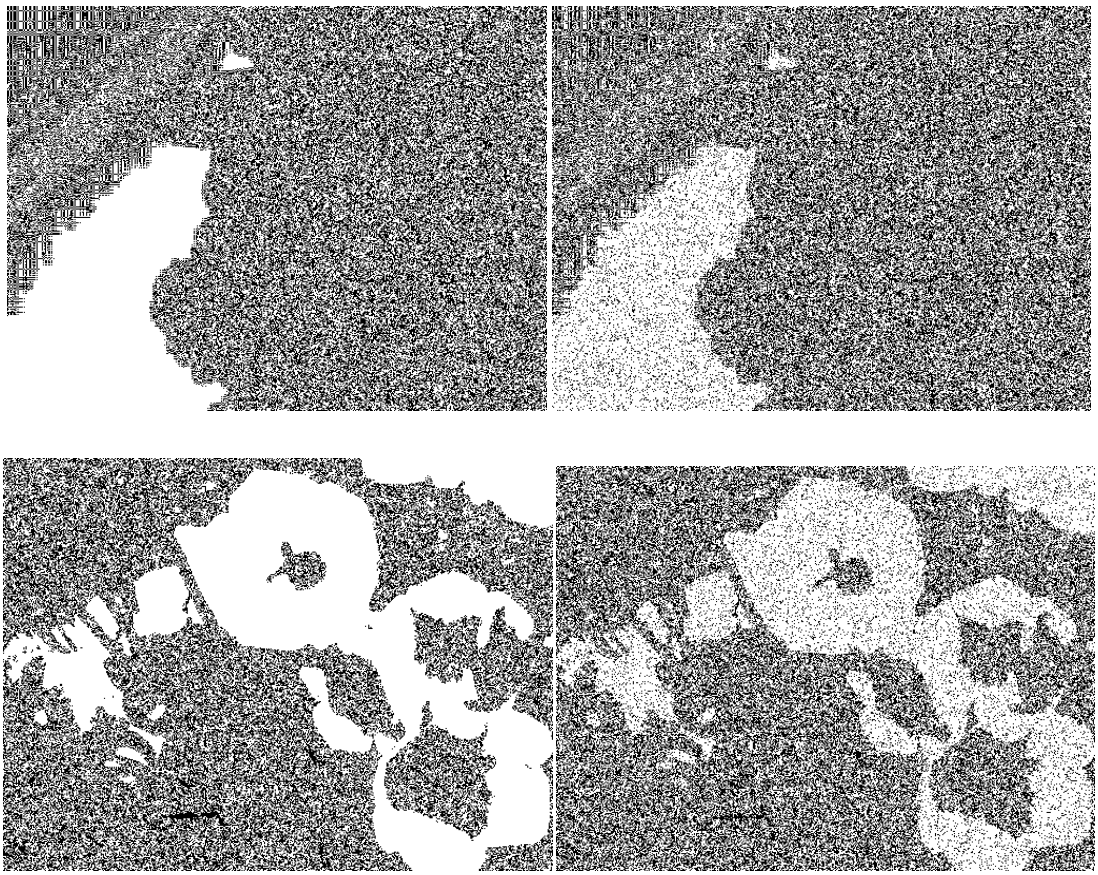

Рис. 4. Распределение младших бит до (слева) и после (справа) внедрения сообщения ±1 методом на 24 %

Более конкретно можно судить о внедрении сообщения с «затиранием» распределения, оценив сравнительную энтропию в младших битах, что будет показано далее.

Таким образом, необходимо применить предложенную нами ранее статистическую модель для вычисления распределения вероятностей при внедрении по методу ±1. Назовём объединённый метод СТ-1 и изобразим схему работы метода на рис. 5.

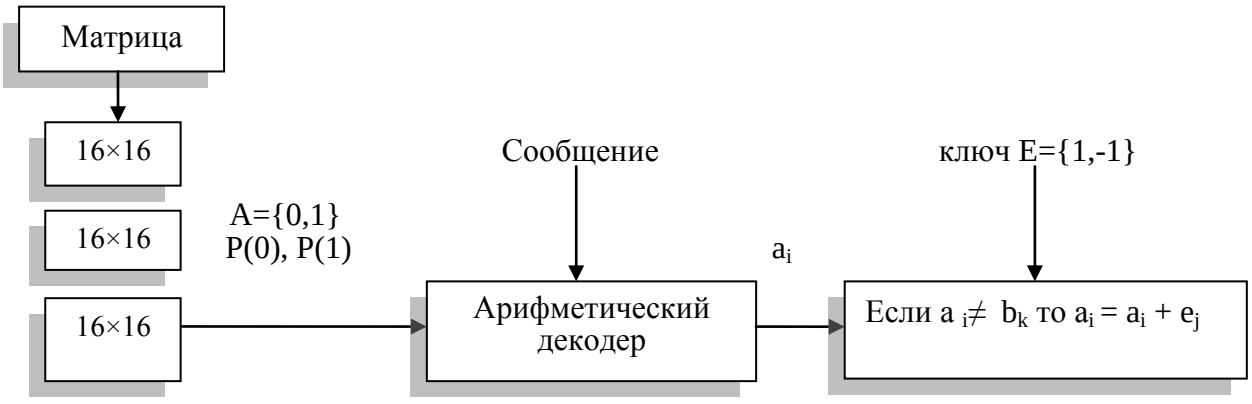

Рис. 5. Схема работы метода СТ-1

Следовательно, для того, чтобы записать бит, полученный на выходе арифметического декодера, мы применяем схему обычного метода ±1. При этом используется секретный ключ в виде псевдослучайной последовательности, которая определяет нужную операцию – увеличение значения текущего байта на единицу или его уменьшение. Для краевых значений 0 и 255 предусмотрены следующие действия: если $b_k = 0$ то $b_k = b_k + 1$, если = 255 то $b_k = b_k - 1$. Таким образом, в результате работы данного алгоритма встраивания, в младших битах будет записана информация, процесс извлечения которой не отличается от изложенного в нашей работе [2].

Распределение младших бит до и после заполнения контейнера по предложенному нами методу внедрения изображено на рис. 6.

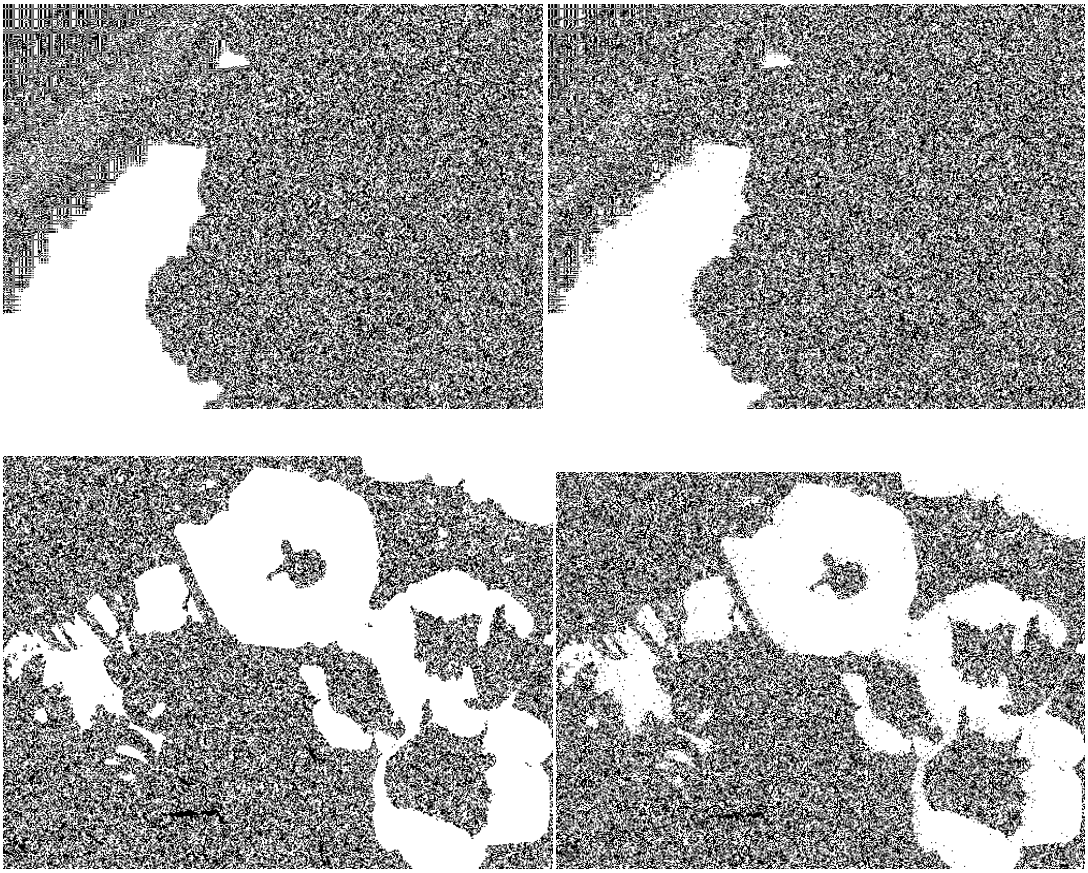

Рис. 6. Распределение младших бит до (слева) и после (справа) заполнения контейнера по предложенной схеме СТ-1. Ёмкость составляет 24 %

3. Исследование преимуществ объединенного метода СТ-1

Для того чтобы оценить отличие предложенного алгоритма от ±1 метода внедрения, был разработан специальный тест, который состоит в исследовании распределения вероятностей младших бит в определённых областях изображения-контейнера. Например, светлые области отлично выделяются на рис. 7 (слева). Такие области, как белые, так и чёрные, мы можем использовать для сравнительной оценки энтропии. При этом, чтобы определить нужные фрагменты матрицы пикселей для вычисления энтропии в заполненном контейнере, необходимо использовать более старшие биты, которые остаются без изменения при внедрении информации. Очевидно, что распределение восьмых бит более схоже с распределением седьмых бит, поэтому их мы можем использовать для определения нужных областей, которые далее будем называть *полезными*. Рассмотрим распределение взятых изолированно восьмых (самых младших) бит и седьмых бит (рис. 7).

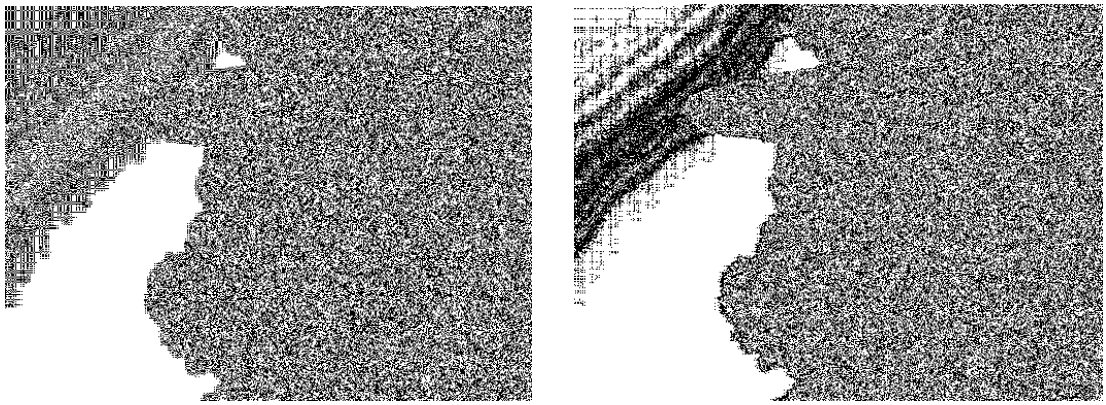

Рис. 7. Распределение восьмых (слева) и седьмых (справа) бит пустого изображения-контейнера

Так как в нашем методе скрытия оценивание статистики младших бит происходило последовательно в отдельно взятых квадратах 16×16, то и в данном случае целесообразно производить оценку по матрице в квадратах того же размера. Итак, считая энтропию распределения вероятностей седьмых бит в каждом квадрате матрицы изображения, мы можем определить полезные квадраты. Оценка энтропии производится по следующей формуле:

$$H = -P_0 \cdot \log_2 P_0 - P_1 \cdot \log_2 P_1, \quad (1)$$

где P_0 и P_1 – это вероятности появления нулевого и единичного бита соответственно. Далее, если $H < 0.98$ для текущего квадрата, то он считается полезным, и мы можем взять соответствующий ему квадрат из восьмых бит, который будет составлять область для оценки энтропии младших бит в заполненном контейнере. Таким образом, по всей матрице изображения набирается некоторое множество квадратов, которые составляют полезную область. На рис. 8 справа чёрным цветом обозначены полезные квадраты для соответствующего распределения младших (восьмых) бит на том же рисунке слева.

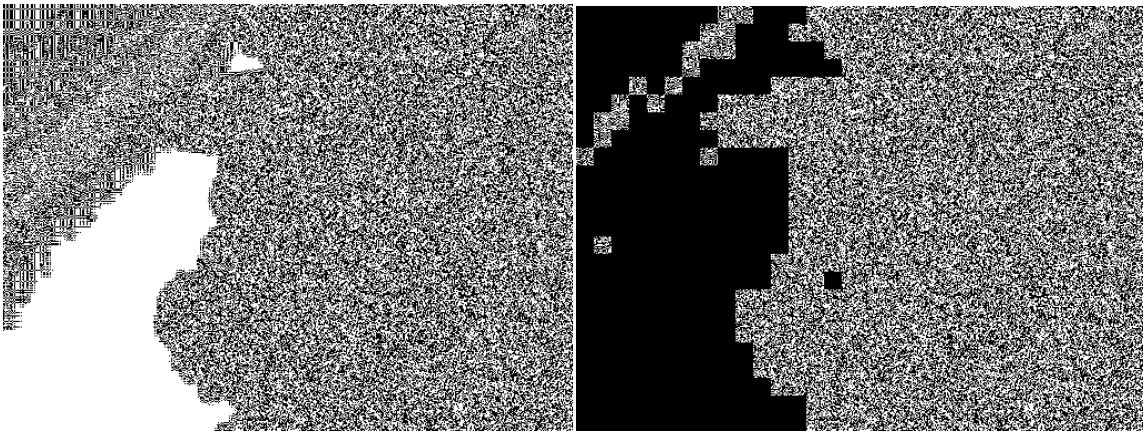

Рис. 8. Распределение младших бит (слева) и соответствующая полезная область (справа) на примере исходного изображения

Полученная область младших бит, состоящая из квадратов размером 16×16 пикселей, используется для того, чтобы найти минимальное значение энтропии отдельно взятых квадратов полезной области. На заданном наборе пустых контейнеров сравним количество файлов, для которых полученный минимум энтропии попадает в определённый диапазон. Результаты

данного теста на наборе из 400 контейнеров размерами 512×384 и 1024×768 приведены в таблице 1.

Таблица 1. Результаты теста на наборе из 400 пустых контейнеров

Значение минимума энтропии	Доля файлов из набора пустых контейнеров размером 512×384, %	Доля файлов из набора пустых контейнеров размером 1024×768, %
$0.9 < H \leq 1$	58	44
$0.8 < H \leq 0.9$	7	13
$0.7 < H \leq 0.8$	2	4
$0.6 < H \leq 0.7$	2	2
$0.5 < H \leq 0.6$	2	2
$0.4 < H \leq 0.5$	2	0
$0.3 < H \leq 0.4$	2	1
$0.2 < H \leq 0.3$	2	2
$0.1 < H \leq 0.2$	1	2
$0 < H \leq 0.1$	4	2
$H = 0$	18	28

Мы видим, что большинство файлов имеет минимальную по областям энтропию, близкую к единице, а также существенен процент файлов с наличием чисто чёрных или чисто белых областей, о чём свидетельствует последняя строка табл. 1. Очевидно, минимальное значение энтропии квадратов зависит от качества и размера полезной области изображения. Например, фотография с распределением младших бит, показанных на рис. 9, является типичным изображением, для которого полезная область оказалась слишком маленькой и некачественной.

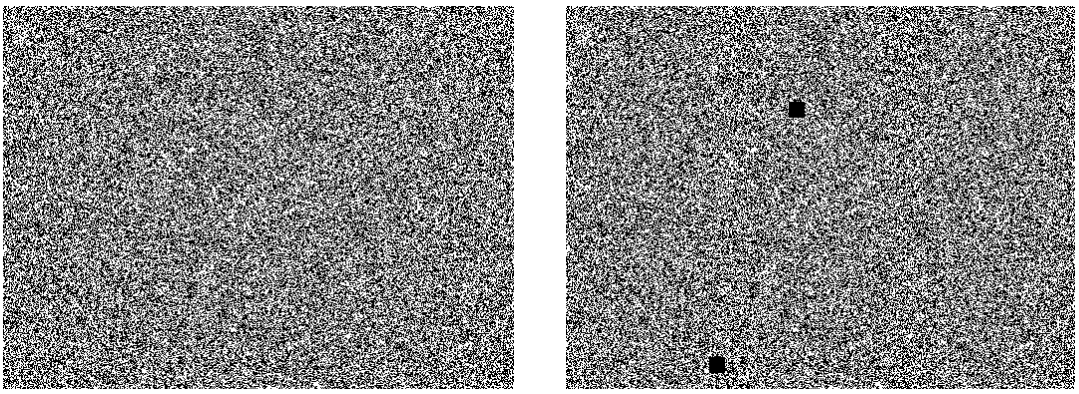

Рис. 9. Распределение младших бит (слева) и полезная область (справа) для пустого «шумного» изображения-контейнера

Для таких файлов, которые составили более половины от общего набора фотографий, заполнение по методу ±1 не внесёт видимых искажений в распределение младших бит, так как оно изначально является близким к случайному распределению и не подходит для нашего теста.

Исследуем заполненные контейнеры с помощью описанного теста и сравним с результатом тестирования соответствующих пустых контейнеров (таблицы 2 и 3).

Таблица 2. Результаты теста на файлах размером 512×384 точек

Значение минимума энтропии	Доля файлов из набора пустых контейнеров, %	Доля файлов из набора заполненных контейнеров по методу ±1, %	Доля файлов из набора заполненных контейнеров по методу CT-1, %
$0.9 < H \leq 1$	58	62	59
$0.8 < H \leq 0.9$	7	6	7
$0.7 < H \leq 0.8$	2	5	2
$0.6 < H \leq 0.7$	2	5	2
$0.5 < H \leq 0.6$	2	6	3
$0.4 < H \leq 0.5$	2	7	1
$0.3 < H \leq 0.4$	2	8	2
$0.2 < H \leq 0.3$	2	1	2
$0.1 < H \leq 0.2$	1	0	1
$0 < H \leq 0.1$	4	0	6
$H = 0$	18	0	15

Таблица 3. Результаты теста на файлах размером 1024×768 точек

Значение минимума энтропии	Доля файлов из набора пустых контейнеров, %	Доля файлов из набора заполненных контейнеров по методу ±1, %	Доля файлов из набора заполненных контейнеров по методу CT-1, %
$0.9 < H \leq 1$	44	52	47
$0.8 < H \leq 0.9$	13	9	12
$0.7 < H \leq 0.8$	4	4	4
$0.6 < H \leq 0.7$	2	2	1
$0.5 < H \leq 0.6$	2	6	1
$0.4 < H \leq 0.5$	0	7	1
$0.3 < H \leq 0.4$	1	20	1
$0.2 < H \leq 0.3$	2	0	3
$0.1 < H \leq 0.2$	2	0	1
$0 < H \leq 0.1$	2	0	2
$H = 0$	28	0	27

Из табл. 2 и 3 мы видим, что применение данного теста позволило выявить существенную разницу в методах ±1 и CT-1. При заполнении контейнеров по алгоритму ±1 области с изначально нулевой неопределённостью нарушились за счёт добавления случайного шума, что можно наблюдать в последних строках табл. 2 и 3; также возрос процент файлов с энтропией, близкой к единице. На рис. 10 и 11 наглядно показаны результаты сравнительных таблиц 2 и 3 для наборов файлов размерами 512×384 и 1024×768.

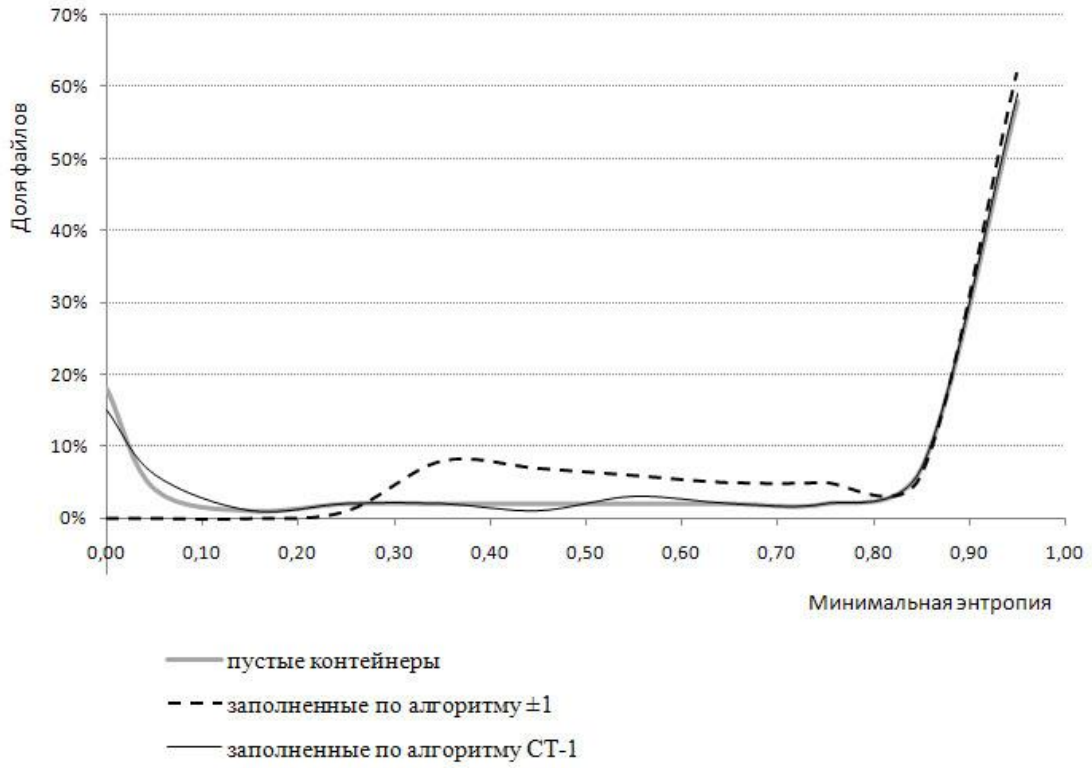

Рис. 10. Результаты теста на минимум энтропии для файлов размером 512×384 – пустых и заполненных по алгоритмам ±1 и СТ-1

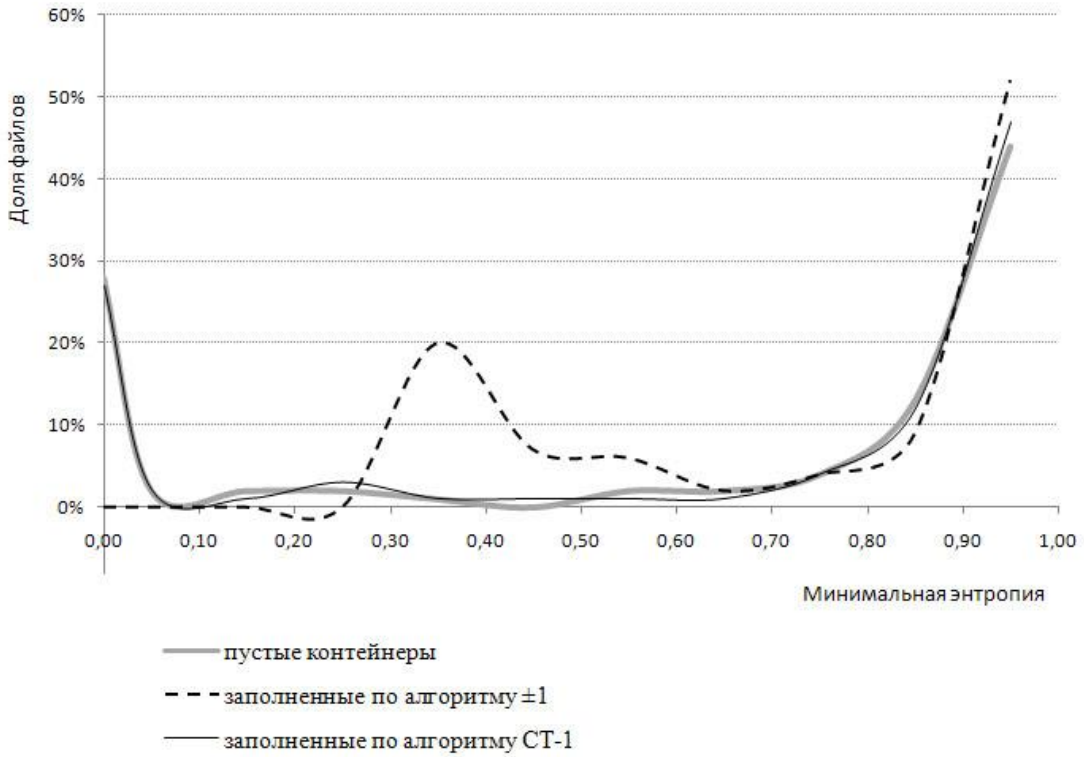

Рис. 11. Результаты теста на минимум энтропии для файлов размером 1024×768 – пустых и заполненных по алгоритмам ±1 и СТ-1

При заполнении контейнера с помощью ±1 энтропия младших бит увеличивается, что свидетельствует о нарушении первоначального распределения вероятностей, в то время как

при заполнении контейнеров совмещённым методом СТ-1 тенденция значений энтропии сохраняется.

4. Стегоанализ предложенного метода

Для оценки алгоритма мы будем использовать метод RS [3]. Анализирующая программа RS-анализа выдаёт количество встроенной информации (L) в процентах от эмпирической ёмкости контейнера, которая высчитывается как при LSB-встраивании.

$$C_{\text{LSB}} = 3wh \text{ бит.}$$

По значению L можно судить о том, заполнен был контейнер или пуст: при $L \geq 5\%$ RS-анализ классифицирует контейнер как заполненный. Также следует отметить, что в цифровой стеганографии существует два рода ошибок. Предположим, имеются две гипотезы: H_s, означающая, что контейнер содержит стегосообщение, и противоположная ей гипотеза H_c, означающая, что контейнер не содержит встроенной информации. Правило решения заключается в том, что каждому контейнеру сопоставляется одна из двух гипотез. В этой задаче возможны два типа ошибок: ошибка первого рода, которая заключается в установлении гипотезы H_s, когда контейнер пуст, и ошибка второго рода, когда принято решение H_c при наличии встроенной информации в контейнере.

Наша первоначальная схема, использующая только статистическую модель, абсолютно точно выявлялась методом RS. Теперь, когда схема улучшена, мы продемонстрируем, насколько повысилась стойкость метода к известной RS-атаке. Для начала оценим процент возникновения ошибок первого рода для тестового набора из 800 файлов. В табл. 4 показано, что 16 % файлов размером 512×384 и 9 % файлов размером 1024×768 составляют ошибку первого рода.

Таблица 4. RS анализ на наборе пустых контейнеров

L		0%	1 – 4%	5% и более
Доля файлов	512×384	5%	79%	16%
	1024×768	33%	58%	9%

Заполним контейнеры случайными данными (имитация зашифрованного сообщения) с помощью совместного алгоритма. Напомним, что эффективная ёмкость контейнеров для данного метода составляет 24 %.

Таблица 5. RS-анализ на наборе контейнеров, заполненных по алгоритму СТ-1

L		0%	1-4%	5% и более
Доля файлов	512×384	5%	80%	15%
	1024×768	30%	59%	9%

Результаты стегоанализа, приведённые в таблицах 4 и 5, свидетельствуют о том, что процент обнаружения встроенной информации по предложенному методу приблизительно равен проценту файлов, в которых была ошибка первого рода. Таким образом, мы показали, что метод устойчив к RS-атаке. Далее сравним с результатами стегоанализа для метода ±1 (табл. 6).

Таблица 6. RS-анализ на наборе контейнеров, заполненных с помощью метода ±1

L		0%	1-4%	5% и более
Доля файлов	512×384	4%	75%	22%
	1024×768	31%	52%	18%

По результатам, приведённым в табл. 6, видно, что процент файлов, в которых обнаружена информация, превышает процент ошибки первого рода на 7 % и 9 % соответственно для файлов размером 512×384 и 1024×768. Таким образом, алгоритм СТ-1 является более стойким к RS-анализу. Для более наглядного результата приведём рис. 12, на котором чёрная шкала демонстрирует долю файлов, классифицированных методом RS, как заполненные.

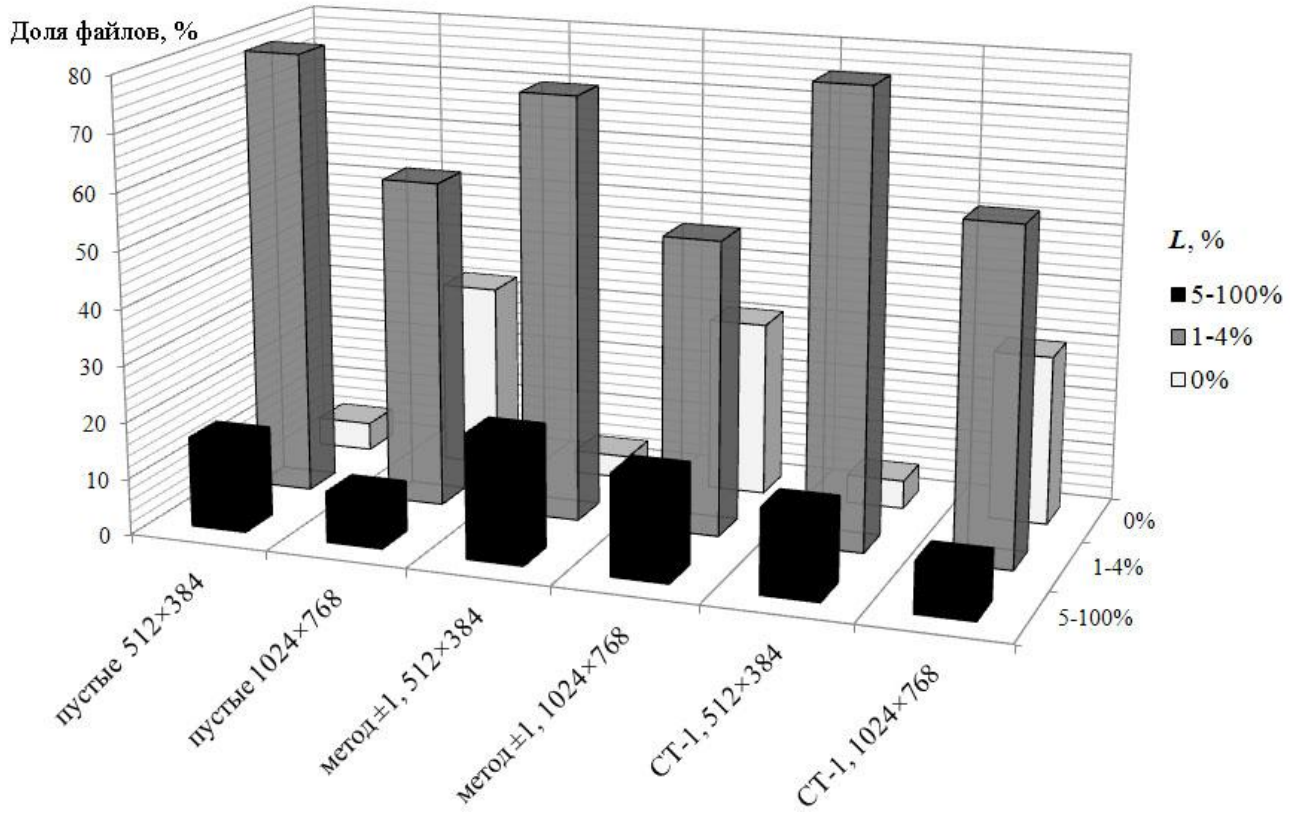

Рис. 12. Результаты RS-анализа для наборов контейнеров – пустых, заполненных методами ±1 и СТ-1 на 24 %

Литература

1. Cachin C. An information-theoretic model of steganography // Lecture Notes in Computer Science (Proc. 2nd Information Hiding Workshop). Springer Verlag, 1998. V. 1525. P. 306-318.
2. Елтышева Е. Ю., Фионов А. Н. Построение стегосистемы на базе растровых изображений с учётом статистики младших бит // Вестник СибГУТИ. 2009. № 1. С. 67-84.
3. Digital Image Steganography Using Stochastic Modulation, with M. Goljan, *Proc. SPIE, Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents V*, vol. 5020, Santa Clara, California, P. 191-202, 2003.

Статья поступила в редакцию 19.03.2011;
переработанный вариант — 03.06.2011.

Елтышева Екатерина Юрьевна
Ассистент кафедры прикладной математики и кибернетики СибГУТИ, (630102, Новосибирск, ул. Кирова, 86) тел. (383) 2-698-272, e-mail: Katya@1nsk.ru.

Constructing Stegosystem for Raster Images on the Base of Stochastic Modulation Adjusted for Statistics of Least Significant Bits

K. Eltysheva

In the present paper we suggest the method of hiding message in the least significant bits of raster images which was complemented and improved by using well-known scheme ±1. In contrast to ±1 method, our scheme codes embedded message under rates of appearing values 0 and 1 in the least significant bits for the purpose of statistic minimum distortion peculiar to empty container. The algorithm of embedding bits in the image matrix is similar to method ±1. Steganalysis of developed algorithm is running and its advantage in comparison to primary one is demonstrated.
Keywords: steganography, stegoanalysis, LSB-methods, statistical model, arithmetic coding, perfect stegosystems, method ±1.