Идентификационный параметр для отбора стегоконтейнеров

И. А. Кривошеев, М. А. Линник

Хабаровский Федеральный исследовательский центр Дальневосточного отделения РАН, г. Хабаровск, Россия

Аннотация. В статье предлагаются идентификационные параметры для изображений, используемых при передаче информации методами стеганографии в качестве стегоконтейнеров. Разработанные алгоритмы позволяют ранжировать изображения в поисках самого оптимального, с учетом особенностей зрительной системы человека и индивидуальных структурных особенностей изображения. Проведенные эксперименты показывают работу идентификационных параметров и соответствие заявленным целям.

Ключевые слова: стеганография, стегоконтейнер, стегоанализ, детерминант, LSB, RS-стегоанализ, Хи-квадрат стегоанализ, анализ битовых срезов.

DOI 10.14357/20718632210304

Введение

Цифровая стеганография является одним из направлений защиты информации. Ключевой особенностью стеганографии является то, что встраиваемая информация в объектах (стегоконтейнерах) добавляется путем некоторого изменения их структуры [1-3]. Вносимые искажения должны выполняться с учетом того, чтобы получившийся на выходе объект с информацией не вызывал никаких подозрений у наблюдателя.

Существует понятие пропускной способности стегосистемы. Если ее превысить, то факт наличия встроенной информации будет скомпрометирован с большей вероятностью. Поэтому вопрос оптимального выбора стегоконтейнеров, который также влияет на оценку качества стегосистемы, является нетривиальной задачей.

Данную проблему решают разными путями. Одним из способов решения является вычисление разностных показателей искажения, таких как отношение сигнал-шум (SNR), среднеквадратическая ошибка, предельное отношение сигнал-шум (GNSR), L^p-норма [2, 4, 5]. Они представляют собой формулы, в которых вычисляется разница между оригинальным (пустым) контейнером и стегоконтейнером со встроенной информацией. Таким образом, рассчитанные значения характеризуют величину искажений объекта, неизбежно получаемого в результате изменения своего изначального вида.

Существует ряд корреляционных показателей, таких как качество корреляции (CQ) и других, работа которых основываются на изучении корреляционных зависимостей между изначальным сигналом и искаженным [2].

Другим подходом является построение стегосистемы и выбор носителя информации на основе особенностей зрительной системы человека. Для этого выполняется вычисление таких параметров, как мера контраста, функции контрастной чувствительности (CSF), яркостной

контрастной чувствительности и др. [4]. Некоторые из представленных параметров являются показателями, которые используются для работы технического зрения [4, 6].

Помимо этого в качестве субъективных показателей применяются специальные таблицы рейтингов [2, 7], согласно которым степень искажения оценивается по цифровой шкале. Оценка выполняется специальными экспертами. Так как в данных методах задействован человеческий фактор, то для получения качественной оценки требуется кропотливая работа и высокий уровень квалификации экспертов.

В стеганографии для выбора стегоконтейнеров используются численные оценки стегоконтейнеров, такие как метод структурного подобия (SSIM) [8] и пиковое отношение сигнала к шуму (PNSR) [4]. Кроме того в [9-10] были предложены новые варианты вычисления характеристик контейнеров, позволяющих выбрать наиболее выгодный вариант из рассматриваемой группы.

Используются также и другие алгоритмы оценки качества изображения, такие как, например, норма Минковского [2-3], которые также служат для определения величины искажения изображения.

1. Постановка задачи

В качестве стегоконтейнеров рассматриваются цветные изображения по причине того, что они во-первых позволяют передавать достаточно большое количество данных, а во-вторых работа с данным видом стеганографических контейнеров накладывает меньшее количество ограничений, таких как, например, сглаживание нежелательных негативных факторов и необходимость высоких вычислительных затрат. Описанные недостатки заметно проявляются в ситуациях, когда в качестве носителей стеганографического сообщения используются текст, звук или видео [1-3]. Например, при работе с видео форматами требуется принимать во внимание жесткие ограничения при работе с определенными видами кадров, а также учитывать связь между ними [1]. Таким образом, выбор цифровых изображений в качестве стегоконтейнеров обуславливается высоким балансом в допустимом максимальном объеме переносимой информации, необходимой вычислительной мощности при их обработке и анализе.

Описанные выше параметры и методы, представленные в литературе [4-6, 11-13] не являются универсальными, и часто не учитывают структурные особенности изображения, т.е. стегоконтейнера. Существуют различные критерии для оценки подобных параметров.

Известно, что человеческое зрение имеет наибольшую чувствительность к колебанию в самых высоких диапазонах цвета [1-3, 14]. Поэтому следует избегать использования в качестве контейнеров изображения, в которых преобладают интенсивности цвета такого рода.

Стоит отметить, что искажения будут наиболее характерны в однотонных изображениях, чем в изображениях с большим количеством объектов или неоднородностью структуры. В противном случае встроенная информация будет легко детектироваться не только с помощью специальных методов оценки [15], но и даже при зрительном наблюдении.

Таким образом, при разработке алгоритма оценки следует обратить внимание на то, чтобы учитывать специфичные особенности зрительной системы человека, которые проявляются в обнаружении искажений в ярких цветах, и структурный состав контейнера.

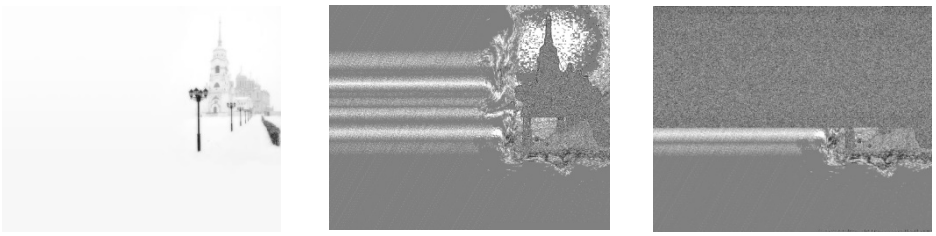
Используя в качестве ключевого параметра восприимчивость человеческого глаза к ярким цветам, было принято решение в качестве базовой формулы использовать выражение для вычисления интенсивности цвета [16]:

$$C = 0.299 * R + 0.587 * G + 0.114 * B, \quad (1)$$

где $R \in 0\ldots255$, $G \in 0\ldots255$, $B \in 0\ldots255$ – цветовые компоненты пикселя.

Значения формулы лежат в узком диапазоне от 0 до 255, поэтому ими удобно оперировать в вычислениях.

Для обнаружения стеганографических вставок используются ряд атак. В качестве атак выступают методы стегоанализа. Помимо статистических методов, которые вычисляют нарушения в корреляционных зависимостях между элементами изображения, существует ряд методов визуального стегоанализа, суть которых заключается в изучении внешнего вида контейнера. В большинстве случаев [1-3] данные методы используются в ка-



а) б) в)

Рис. 1. Иллюстрация стеганографической атаки анализом битовых срезов

а) исходное изображение; изображение, построенное с помощью младших бит
б) без встроенной информации
в) со встроенной информацией

честве предварительной оценки. Однако, метод анализа битовых срезов [15] позволяет сформировать изображения, построенные с помощью младших бит. Появление инородной области, резко выделяющейся на фоне остальной части изображения, позволяет легко определить наличие встроенной информации. Особенно, если встраивание производилось в каждый пиксель подряд.

На Рис. 1 представлен пример такой атаки. На таком изображении можно четко различить встроенную информацию.

Одной из причин заметности данных областей является перезапись границ объектов на изображении. Так как аналитик знает, как выглядит оригинал, несоответствие в изображении, построенном на младших битах, вызовет ряд подозрений.

Существует ряд алгоритмов для выделения границ на изображении [14]. С их помощью можно сформировать изображения, на которых границы объектов будут оформлены в виде четных отдельных линий, что позволит выделить в них отдельные группы пикселей (Рис. 2).

С учетом описанных выше факторов при разработке алгоритма следует опираться на такие особенности, как распределение и концентрацию объектов на рассматриваемом изображении, а также варьирование цветового диапазона. Выдвигается гипотеза, что ситуации, при которых на анализируемом контейнере расположено малое число объектов (однородность структуры изображения) или же в нем преобладают высокие области интенсивности цвета, являются нежелательными.

Поэтому на основе выше сказанного существует потребность в проектировании алгоритма, который позволил бы определять наиболее оптимальный вариант из некоторой группы изображений, который при использовании в качестве стегоконтейнера обеспечивал бы большую защищенность, чем другие.

а) б)

Рис. 2. Пример обработки изображения с помощью оператора выделения границ

а) оригинальное изображение
б) изображение с выделенными границами (черным цветом)

2. Алгоритм идентификационного параметра

В данной работе на первом этапе предлагается алгоритм, который предоставляет возможность оценивать предлагаемый контейнер или выбирать из нескольких предложенных один, используя визуальный критерий.

Основой, как обычно, является разбиение изображения на определенные области. Для дальнейшего описания следует ввести некоторые понятия и определения.

Область – это совокупность пикселей, которые имеют значения интенсивности цвета в заданном диапазоне. Пиксель принадлежит области, если среди соседних элементов как минимум один из них входит в тот же диапазон, что и рассматриваемый.

Если пиксель принадлежит к области, и при этом хотя бы один из соседних пикселей не принадлежит к области, то он является граничным пикселем своей области.

Дополнительно в качестве отдельной группы выделяются изолированные пиксели, которые со всех сторон окружены элементами, не принадлежащими к диапазону рассматриваемого пикселя. Эта группа выступает в качестве одного из ограничивающих факторов.

Под площадью области понимается количество всех элементов принадлежащих к ней.

На Рис. 3 показана схема, иллюстрирующая введенные параметры и алгоритм выявления областей.

Зрительная система человека наиболее четко различает колебания в ярких цветах, поэтому факт встраивания в крайние диапазоны цвета будет заметен с высокой вероятностью. По этой причине следует обратить внимание на самые высокие значения интенсивности. Поэтому для рассмотрения были выбраны области, которые попадают в диапазон значений пикселей от 170 до 255. Ранжирование контейнеров будет производиться с учетом анализа в данном промежутке.

Таким образом, алгоритм анализа изображения состоит из следующих шагов:

1) Введем массив

$$A_k = \begin{bmatrix} a_{0,0}, \ldots, a_{0,l-1} \\ \cdots, \quad a_{i,j}, \quad \cdots \\ a_{l-1,0}, \ldots, a_{l-1,l-1} \end{bmatrix}, \qquad (2)$$

где $k \in 0 \ldots n$, n – количество блоков, $a_{i,j}$ – флаг, указывающий на принадлежность пикселя к рассматриваемой области , $i \in 0, \ldots, l - 1$, $j \in 0, \ldots, l - 1$ – координаты пикселя в блоке, l – длина или ширина блока (в пикселях).

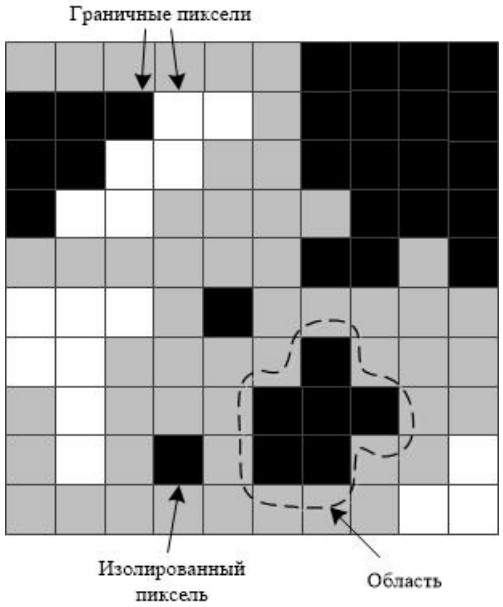


Рис. 3. Схема выбора областей на изображении

2) Выполняется обход изображения слева направо с анализом значений пикселей. При достижении границы ширины изображения анализ следующего пикселя выполняется с начала следующей строки, после чего обход продолжается до тех пор, пока не будет достигнута граница ширины изображения, а затем действия повторяются.

3) Значения пикселей вычисляются по формуле (1).

4) Если значение пикселя не относится к выбранному диапазону, то элементу $a_{i,,\,j}$ ставится в соответствие значение 0.

5) В ином случае анализируются соседние пиксели. Если хотя бы один из граничных пикселей принадлежит к выбранной области, то значением $a_{i,j}$ является значение 1. Если же ни один из элементов не входит в выбранный диапазон, то данный пиксель является изолированным и $a_{i,j}$ ставится значение 0.

6) После выполнения описанных операций по анализу изображения массив A становится картой, показывающей расположение и принадлежность элементов к выбранной области.

7) Следующим шагом является подсчет пикселей, являющихся границами изображения. Для этого сформируем изображение на основе рассматриваемого, в котором были выделены границы с помощью оператора выделения границ [14]. После обработки данным оператором формируется изображение в оттенках серого. Далее изображение переводится в черно-белый формат. Введем массив

$$B_k = \begin{bmatrix} b_{0,0}, \ldots, b_{0,l-1} \\ \cdots, \quad b_{i,j}, \quad \cdots \\ b_{l-1,0}, \ldots, b_{l-1,l-1} \end{bmatrix} \qquad (3)$$

Он представляет собой массив пикселей исходного изображения, где ненулевые значения получают элементы $b_{i,\,j}$ белого цвета, иначе говоря границы изображения.

На основе приведенного выше алгоритма были разработаны два метода вычисления идентификационных параметров.

Первый определяет отношение суммы пикселей из выбранного диапазона к их общему числу:

$$Sum = \frac{\sum C - \sum E}{\sum_{i=0}^{H}\sum_{j=0}^{W} P - \sum U}, \qquad (4)$$

где C – значение пикселя в области рассматриваемого диапазона, вычисляемого согласно выражению (1), $\sum U$ – сумма значений изолированных пикселей, P – значения пикселей по всей области изображения, W, H – ширина и высота изображения, $\sum E$ – сумма значений пикселей, которая была вычислена с помощью оператора выделения границ.

Второй параметр вычисляет значение, которое является отношением площади областей пикселей к их общему числу:

$$Sq = \frac{\sum p - \sum e}{W * H - \sum u}, \qquad (5)$$

где $\sum p$ – количество пикселей в области рассматриваемого диапазона, вычисляемого согласно выражению (1), $\sum u$ – количество изолированных пикселей, W, H – ширина и высота изображения, $\sum e$ – количество пикселей, которое было вычислено с помощью операторов выделения границ.

Существующие алгоритмы [4-5, 10] оценивают изображения без учета его особенностей, таких как количество и размер объектов, распределения яркости и прочих признаков. Разработанный алгоритм по сравнению с существующими аналогами позволяет выполнять оценку стегоконтейнеров на основе их структурного содержания. Таким образом, отбор выполняется более гибким способом, исходя из специфических атрибутов рассматриваемых изображений в группе.

3. Вычислительный эксперимент

Для проведения экспериментов были выбраны 10 изображений размера 1024x1280, для каждого из которых были произведены вычисления параметров по формулам (4) и (5).

Изображения для экспериментов были подобраны на основе их структурного и яркостного содержания. Во внимание принималась неоднородность структуры и преобладание ярких или темных тонов. Поэтому для проверки выдвинутой гипотезы в выбранном тестовом

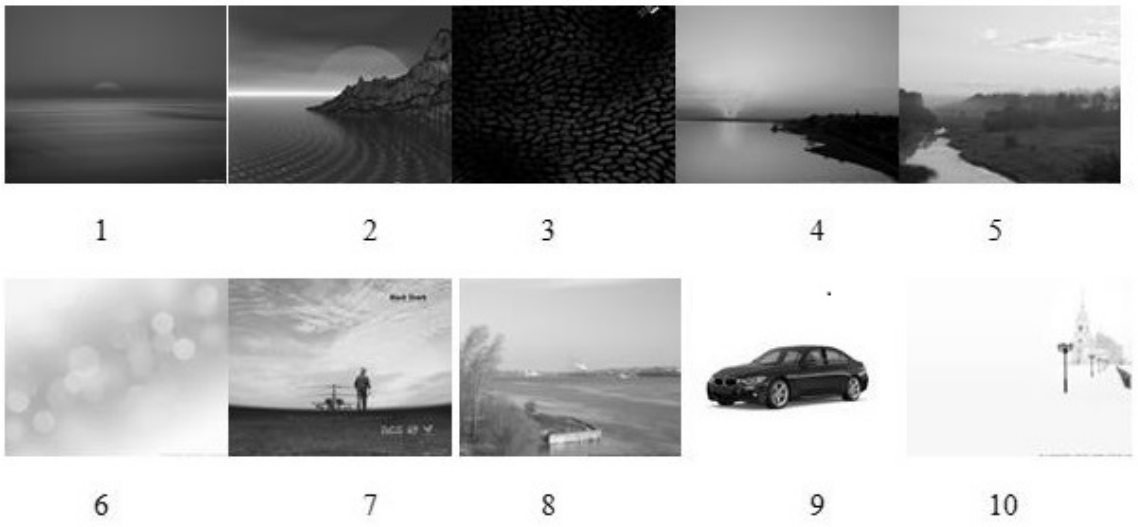


Рис. 4. Группа отсортированных изображений

Табл. 1. Ранжирование группы изображений

№ изображения	1	2	3	4	5	6	7	8	9	10
Sq (10^{-5})	1	2	10	26	76	104	175	442	477	528

наборе находятся изображения, в которых одновременно имеются экземпляры с высокими и низкими значениями яркости цвета, а также – с большим и малым количеством объектов.

Рассмотрим результаты наиболее предпочтительного варианта ранжирования, второго метода (5): отношение площади областей пикселей к их общему числу, поскольку результаты вычислений показали наибольшее соответствие выдвинутой гипотезе. Результаты выполнения ранжирования представлены на Рис. 4, где изображения отсортированы по возрастанию идентификационного параметра. В Табл. 1 приведены вычисленные значения.

Как можно заметить, изображения расположены согласно возрастанию яркости цвета и по уменьшению неоднородности структуры рисунка. Т.е. по увеличению риска обнаружения встроенных данных. Численным моделированием, при встраивании одинаковой информации в эти изображения, были получены результаты, которые при стегоатаке подтвердили правильность выбора идентификационного параметра.

Эффективность результата работы алгоритма была оценена с помощью стеганографической атаки методом RS-стегоанализа [15] при встраивании 30% информации от общего доступного объема. Результатом этого метода получается численное значение, которое показывает приблизительный размер посторонней встроенной информации. На Рис. 5 показаны результаты RS-стегоанализа для рассматриваемой группы изображений, отсортированных по введенному идентификационному параметру.

Можно заметить, что изображения, согласно результатам стегоанализа изображения также расположены по возрастанию. Таким образом, работа введенного алгоритма удовлетворяет выдвинутым предположениям. Поэтому можно говорить о том, что разработанный алгоритм позволяет выбрать наиболее эффективные изображения, которые будут использоваться в качестве контейнеров

Заключение

В результате изучения особенностей восприятия цвета зрительной системой человека и при учете структурных особенностей изображений был разработан алгоритм, позволяющий разбить элементы (пиксели) изображения на группы. На основе данного метода был предложен ряд параметров для ранжирования изображений.

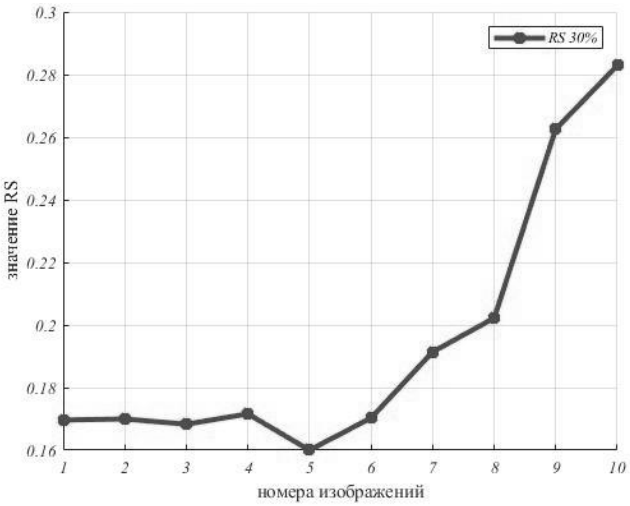


Рис. 5. Результаты RS-стегоанализа

Разработанные алгоритмы являются тонкими и гибкими в применении, потому что они учитывают структуру изображения и его цветовое содержание.

Введенные идентификационные параметры предлагается использовать в стеганографии в качестве критерия отбора наиболее подходящих изображений в качестве контейнеров для передачи информации.

Литература

1. Шелухин О.И. Канаев И.Д. Стеганография. Алгоритмы и программная реализация. М.: Горячая линия. 2017. 592 с.
2. Коханович Г.Ф., Пузыренко А.Ю. Компьютерная стеганография. Теория и практика. К.: МК-Пресс. 2006. 288 с.
3. Грибунин В.Г., Оков И.Н. Цифровая стеганография. М.: Солон-пресс. 2002. 272 с.
4. Thung K.H., Paramesran R. A Survey of Image Quality Measures // TECHPOS, International Conference. 2009. pp. 1-4.
5. Noroozi E., Daud S. B. M., Sabouhi A. Critical Evaluation on Steganography Metrics // Proceedings of 2011 International Conference on Electrical Engineering and Applications. 2013. pp. 927-931.
6. Беззубик В. В., Белашенков Н. Р. Определение функции контрастной чувствительности для систем технического зрения. // Известия высших учебных заведений. Приборостроение - 2013. Т. 56. № 9. С. 73-79.
7. Shnayderman A., Gusev A., Eskicioglu A.M. An SVD-Based Gray-Scale Image Quality Measure for Local and Global Assessment // IEEE Transactionson image processing. 2006. vol. 15. № 2. pp. 422-429.
8. Wang Z., Bovik A.C., Sheikh H.R. Image quality assessment: From error visibility to structural similarity // IEEE transaction on Image Processing. 2004. Vol. 13. № 4. pp. 600-612.
9. Setiadi D. PSNR vs SSIM: imperceptibility quality assessment for image steganography // Multimedia Tools and Applications. 2021. Vol. 80. №6. pp. 1-22.
10. Roy R., Changder S. Quality Evaluation of Image Steganography Techniques: A Heuristics based Approach // International Journal of Security and its Applications. 2016. Vol. 10. № 4. pp. 179-196.
11. Кривошеев И.А., Линник М.А. Статический способ стеганографического встраивания информации на основе LSB // Системы и средства информатики. 2020. Т. 30, № 3. С. 56-66.
12. Кривошеев И.А., Линник М.А., Кожевникова Т.В. Способ встраивания информации в цветное изображение // Патент РФ на изобретение №2738250 от 26.03.2020. Бюл. № 35.
13. Кривошеев И.А., Линник М. А. К вопросу об оценке устойчивости стеганографической системы // Ученые заметки ТОГУ. 2017. Т. 8, № 2. С. 433-437.
14. Анисимов Б.В. Распознавание и цифровая обработка изображений. М.: Высш. школа. 1983. 295 с.
15. Pfitzmann A., Westfeld A. Attacks on Steganographic Systems. Breaking the Steganographic Utilities EzStego, Jsteg, Steganos, and S-Tools and Some Lessons Learned. IH 1999. LNCS, 1768: pp. 61-76.
16. Гонсалес Р., Вудс Р. Цифровая обработка изображений. М. Техносфера. 2005. 1072 с.

Кривошеев Игорь Александрович. Хабаровский Федеральный исследовательский центр Дальневосточного отделения РАН г. Хабаровск, Россия. Главный научный сотрудник, доктор технических наук. Количество печатных работ: 149. Область научных интересов: информационная безопасность и защита информации, численное моделирование, обработка изображений. E-mail: igork@as.khb.ru

Линник Максим Анатольевич. Хабаровский Федеральный исследовательский центр Дальневосточного отделения РАН г. Хабаровск, Россия. Младший научный сотрудник. Количество печатных работ: 9. Область научных интересов – информационная безопасность, численное моделирование, обработка изображений. E-mail: linnik.max1995@mail.ru

Identification Parameter for the Selection of Stego-Carriers

I. A. Krivosheev, M. A. Linnik

Khabarovsk Federal Research Center of the Far Eastern Branch of the Russian Academy of Sciences, Khabarovsk, Russia

Abstract. The article proposes a number of identification parameters for images used in the transfer of information by steganography methods as stego-carriers. The developed algorithms make it possible to rank images in search of the most optimal, taking into account the peculiarities of the human visual system and individual structural features of the image. The experiments carried out show the work of the identification parameters and their compliance with the stated goals.

Keywords: steganography, stego-carrier, stegoanalysis, determinant, LSB, RS-steganalysis, Chi-square stegoanalysis, bit-slice analysis.

DOI 10.14357/20718632210304

References

1. Shelukhin, O. I., Kanaev, I.D. 2017. Steganografija. Algoritmy I programmnaja realizacija [Steganography. Algorithms and software implementation] Moskow: Hot line. 592 p.
2. Kokhanovich, GF, Puzyrenko, A.Y. 2006. Komp'juternaja steganografija. Teorija i praktika [Computer steganography. Theory and practice.]. K.: MK-Press,. 288 p.
3. Gribunin, V.G., Okov, I.N. 2002. Cifrovaja steganografija [Digital steganography] Moskow: Solon-press. 272 p.
4. Thung, K.H., Paramesran, R. A Survey of Image Quality Measures // TECHPOS, International Conference. 2009. pp. 1-4.
5. Noroozi, E., Daud, S. B. M., Sabouhi, A. Critical Evaluation on Steganography Metrics // Proceedings of 2011 International Conference on Electrical Engineering and Applications. 2013. pp. 927-931.
6. Bezzubik, V. V., Belashenkov, N.R. 2013. Opredelenie funkcii kontrastnoj chuvstvitel'nosti dlja sistem tehnicheskogo zrenija. [Determination of the contrast sensitivity function for vision systems.] // Izvestija vysshih uchebnyh zavedenij. Priborostroenie [Proceedings of higher educational institutions. Instrumentation]. 56(9): 73-79.
7. Shnayderman A., Gusev A., Eskicioglu A.M. An SVD-Based Gray-Scale Image Quality Measure for Local and Global Assessment // IEEE Transactionson image processing. 2006. vol. 15. № 2. pp. 422-429.
8. Wang Z., Bovik A.C., Sheikh H.R. Image quality assessment: From error visibility to structural similarity // IEEE transaction on Image Processing. 2004. Vol. 13. № 4. pp. 600-612.
9. Setiadi D. PSNR vs SSIM: imperceptibility quality assessment for image steganography // Multimedia Tools and Applications. 2021. Vol. 80. №6. pp. 1-22.
10. Roy R., Changder S. Quality Evaluation of Image Steganography Techniques: A Heuristics based Approach // International Journal of Security and its Applications. 2016. Vol. 10. № 4. pp. 179-196.
11. Krivosheev, I.A., Linnik, M.A. 2020. Staticheskij sposob steganograficheskogo vstraivanija informacii na osnove LSB [Static way of steganographic information embedding based on LSB]. Sistemy i sredstva informatiki [Systems and Means of Informatics]. 30(3): 56-66.
12. Krivosheev, I.A., Linnik, M.A.., Kozhevnikova T.V. 2020. Sposob vstraivanija informacii v cvetnoe izobrazhenie [Method of embedding information into a color image] Patent RF No. 2738250.
13. Krivosheev, I.A., Linnik, M.A. 2017. K voprosu ob otsenke ustoychivosti steganograficheskoy sistemy [On the issue of assessing the stability of a steganographic system]. Uchenye zametki TOGU [TOGU Science Notes]. 8(2): 433-437.
14. Anisimov, B.V. 1983. Raspoznavanie i cifrovaja obrabotka izobrazhenij [Recognition and digital processing of images]. Moskow.: High School. 295 p.
15. Pfitzmann, A., Westfeld, A. Attacks on Steganographic Systems. Breaking the Steganographic Utilities EzStego, Jsteg, Steganos, and S-Tools and Some Lessons Learned. IH 1999. LNCS, 1768: pp. 61-76.
16. Gonzalez R., Woods R. Tsifrovaya obrabotka izobrazheniy [Digital image processing]. Moskow: Tekhnosfera. 1072 p.

Krivosheev I. A. Doctor of Science in technology, leading scientist, Khabarovsk Federal Research Center of the Far Eastern Branch of the Russian Academy of Sciences, Turgeneva street, 51, Khabarovsk, 680000 Russia, e-mail: igork@as.khb.ru

Linnik M. A. Junior scientist, Khabarovsk Federal Research Center of the Far Eastern Branch of the Russian Academy of Sciences, Turgeneva street, 51, Khabarovsk, 680000 Russia, e-mail: linnik.max1995@mail.ru