

Ю. С. Гулина, В. Я. Колючкин,
Н. Е. Трофимов

РАСПОЗНАВАНИЕ ЧЕЛОВЕКОМ-ОПЕРАТОРОМ СИМВОЛОВ В УСЛОВИЯХ СИЛЬНОГО ЗАШУМЛЕНИЯ

Предложена многоканальная модель зрительной системы человека-оператора, учитывающая информацию как об амплитудах, так и о фазах гармонических составляющих спектра объекта. Обоснована методика расчета вероятности распознавания образов объектов, наблюдаемых человеком. Проведен количественный анализ возможности распознавания символов при наличии помех в стеганографических системах. Приведены результаты экспериментальных исследований, подтверждающие правильность теоретических выводов.

E-mail: julia-sg@yandex.ru; kolambda@gmail.com

Ключевые слова: стеганография, зрительная система, человек-оператор, вероятность распознавания.

Одним из важных показателей качества стеганографических методов является робастность скрытой информации — ее устойчивость к возможным искажениям. При этом в зависимости от области применения этих методов требования к робастности различаются. Например, в некоторых приложениях требуется установить подлинность цифрового объекта, т. е. определить, не был ли объект сфальсифицирован или не подвергался ли он изменениям. В таких случаях целесообразно использовать методы, обеспечивающие наименьшую робастность метки, подтверждающей подлинность. При использовании стеганографических методов для тайной коммуникации или защиты авторских прав, наоборот, важно обеспечить устойчивость скрытой информации к модификациям цифрового объекта, возникающим либо в результате стандартной обработки данных, либо в результате злонамеренных атак [1].

В ряде публикаций [2, 3] для количественной оценки качества стеганографических систем применяется показатель BER (Bit Error Rate — вероятность появления ошибочных битов), который характеризует качество каналов связи. Использование данного показателя в стеганографических системах позволяет оценить возможность извлечения скрытой информации в автоматическом режиме с применением вычислительной техники. При небольших значениях BER для надежного восстановления скрытых данных достаточно обеспечить соответствующую избыточность встраиваемой информации при записи.

Однако на практике возникают ситуации, когда BER имеет достаточно большое значение (больше 0,3). Например, если стегообъект подвергнулся атакам неизвестного типа, существенная часть скрытых данных может быть уничтожена.

Если на систему наложены жесткие требования не только по робастности, но и по незаметности факта наличия встроенной информации, то при априори неизвестном контейнере (а такие случаи представляют наибольший практический интерес) значение BER может сильно зависеть от алгоритма извлечения.

Таким образом, в ряде случаев надежное извлечение скрытых данных автоматическими методами затруднено. В то же время известно, что человек способен различать зрительные образы знакомых объектов в условиях наличия сильных помех [4, 5].

Если к системе стеганографической защиты предъявляются жесткие требования к робастности и незаметности наличия встроенного сообщения, а необходимый объем этого сообщения невелик, предлагается записывать его в виде графического объекта, содержащего, например, символы латинского алфавита или цифры. Надежность восстановления скрытых данных будет определяться способностью зрительной системы человека различать знаки при наличии сильных помех. Известны различные модели зрительной системы человека, на основе которых можно проводить количественную оценку возможности распознавания образов объектов [4—7].

В ряде работ [7, 8] исследованы многоканальные модели зрительной системы человека-оператора при обнаружении и распознавании изображений объектов. Во всех этих моделях в качестве признаков обнаружения или распознавания используются амплитуды спектральных составляющих образов объектов. Эти признаки адекватно отражают процесс обнаружения объектов человеком-оператором. Однако для оценки возможности распознавания их недостаточно, поскольку существенная часть информации об изображении содержится в спектре фаз. Поэтому модель зрительной системы человека-оператора для адекватной оценки процесса распознавания объектов нуждается в доработке.

Для предотвращения потери важной информации об объекте помимо спектра амплитуд необходимо учитывать его спектр фаз, а значит, строить модель таким образом, чтобы признаки распознавания учитывали распределение как амплитуд, так и фаз в спектре изображения. В модели, предложенной в работе [8], в параллельных каналах зрительной системы осуществляется оптимальная фильтрация изображения. На выходе этих каналов формируются сигналы Z_{mn} , представляющие собой корреляционные интегралы, в которых отсутствует информация о фазах φ_{mn} гармонических составляющих спек-

тра объекта. Но за счет того, что в оптимальных фильтрах осуществляется фазовая подстройка для максимизации значения корреляционного интеграла на выходе, информация о фазах объекта может сниматься с фильтров и использоваться для распознавания объектов. Таким образом, модель многоканальной зрительной системы человека-оператора при распознавании образов объектов будет иметь вид, представленный на рис. 1.

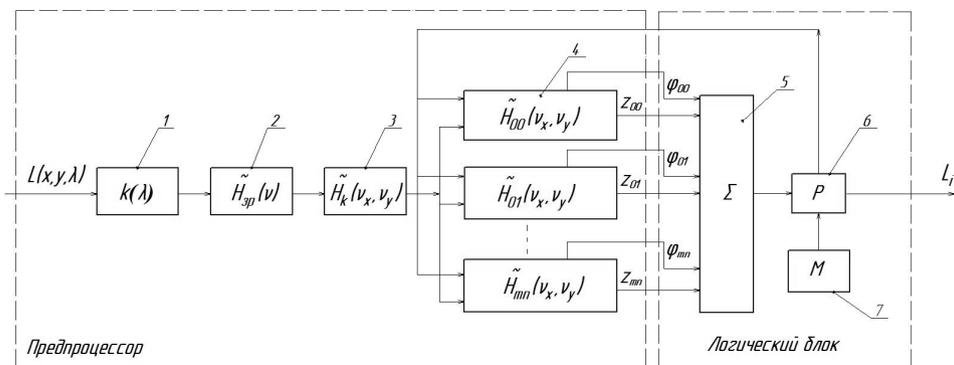


Рис. 1. Структурная схема многоканальной модели зрительной системы человека-оператора при распознавании:

1—4 — оптический, временной, пространственный и параллельные фильтры; 5 — блок определения решающих границ; 6 — процессор; 7 — память

Рассмотрим процесс распознавания изображений человеком-оператором в соответствии с доработанной многоканальной моделью зрительной системы человека-оператора.

Известно [7], что спектр объекта, ограниченного в пределах прямоугольника со сторонами $l_x \times l_y$, можно представить в соответствии с теоремой Котельникова соотношением

$$\tilde{L}(v_x, v_y) = L_M \sum_{m,n=-\infty}^{\infty} \tilde{L}_H \left(\frac{m}{l_x}, \frac{n}{l_y} \right) \text{sinc} \left[\pi l_x \left(v_x - \frac{m}{l_x} \right) \right] \text{sinc} \left[\pi l_y \left(v_y - \frac{n}{l_y} \right) \right],$$

где ΔL_M — максимальное значения яркости наблюдаемого объекта;

$\tilde{L}_H \left(\frac{m}{l_x}, \frac{n}{l_y} \right)$ — значение mn -гармоники нормированной функции яркости.

В зрительной системе человека этот спектр фильтруется пространственно-частотными и временным фильтрами, в результате на выходе mn -х каналов формируются сигналы Z_{mn} , представляющие

собой максимальные значения корреляционных интегралов, усредненных по времени [7]:

$$Z_{mn} = \max_{x,y} \left\{ \overline{U_{mn}(x, y, t)} \right\} = S_{mn} + N_{mn}.$$

Сигнальная составляющая на выходе mn -го канала для нулевой и mn -й гармоник описывается соответственно следующими выражениями:

$$S_{00} = \frac{k_B \Delta L_M}{A_0} \tilde{L}_H(0, 0); \quad (1)$$

$$S_{mn} = \frac{k \Delta L_M}{A_0} \left| \tilde{L}_H \left(\frac{m}{l'_x}, \frac{n}{l'_y} \right) \right|, \quad (2)$$

где k — коэффициент, учитывающий энергетические соотношения при преобразовании оптического сигнала; $A_0 = l'_x l'_y$ — габаритная площадь изображения объекта, имеющего размеры l'_x и l'_y .

Если предположить, что спектральная плотность мощности шума равномерна, т. е. шум белый, то, как показано в работе [7], дисперсия составляющей (шумовой составляющей) N_{mn} шума корреляционного интеграла

$$\sigma_{mn}^2 = \frac{k^2 \sigma_{ш}^2 A_{ш}}{2 A_0 \tau_{зр} \nu_k}, \quad (3)$$

где $\sigma_{ш}^2$ — дисперсия аддитивного шума; $A_{ш}$ — площадь корреляции аддитивного шума; $\tau_{зр}$ — постоянная времени, описывающая инерционность зрительного восприятия; ν_k — частота кадров.

При восстановлении скрытых данных из стего-объекта характер шумовой составляющей может быть различным, однако в ряде практических случаев приближение белого шума не дает существенных ошибок.

Из выражений (1) и (2) следует, что корреляционные интегралы несут в себе информацию об амплитудах гармонических составляющих спектров объектов, а информация о фазах гармоник может быть получена от соответствующих оптимальных фильтров mn -х каналов. Таким образом, в логический блок зрительной системы для распознавания передается информация как об амплитудах, так и о фазах гармонических составляющих спектров объектов.

Однако из амплитуд и фаз невозможно сформировать однородное признаковое пространство. Поэтому в качестве признаков распознавания целесообразно использовать действительные и мнимые части гармонических составляющих.

Переходя к действительным и мнимым частям сигналов, поступающих в логический блок с выходов параллельных каналов, получаем

$$Z_{Re_{mn}} = (S_{mn} + N_{mn}) \cos(\varphi_{mn}), \quad (4)$$

$$Z_{Im_{mn}} = (S_{mn} + N_{mn}) \sin(\varphi_{mn}). \quad (5)$$

Полученные сигналы являются реализациями случайных процессов. Предлагается признаками распознавания образов наблюдаемых объектов считать действительные и мнимые части детерминированных составляющих этих реализаций.

В таком случае векторы признаков, описывающих различные классы объектов, можно записать в виде

$$\mathbf{X}_j = \left[x_{Re_{10}}^j, \dots, x_{Re_{mn}}^j, \dots, x_{Im_{10}}^j, \dots, x_{Im_{mn}}^j \right]^T, \quad (6)$$

где $x_{Re_{mn}}^j$ — действительная часть mn -й гармонической составляющей, а $x_{Im_{mn}}^j$ — ее мнимая часть, причем

$$x_{Re_{mn}}^j = S_{mn} \cos(\varphi_{mn});$$

$$x_{Im_{mn}}^j = S_{mn} \sin(\varphi_{mn}).$$

Нулевую гармонику использовать в качестве признака распознавания нецелесообразно, поскольку она несет в себе информацию только о средней яркости объекта.

Для расчета вероятности распознавания по выбранным признакам необходимо определить дисперсию действительной и мнимой частей гармонических составляющих. При подстройке, осуществляемой во время согласованной фильтрации, фаза как сигнальной, так и шумовой составляющих изменяется на величину $-\varphi_{mn}$. Поскольку фаза белого шума случайна и равномерно распределена в диапазоне значений $0 \dots 2\pi$, добавление к ней любой постоянной величины изменит только ее начало отсчета и не повлияет на дисперсию самого шума. В этом случае дисперсию как действительной, так и мнимой частей гармонической составляющей можно описать выражением (3).

Тогда вероятность распознавания j -го класса можно рассчитать по формуле

$$P_{pj} = \frac{1}{(2\pi\sigma_{mn})^{\frac{N}{2}}} \int_{\Omega_j} \exp \left(-\frac{\left(z_1 - x_{Re_{10}}^j \right)^2 + \dots + \left(z_{N/2} - x_{Re_{mn}}^j \right)^2}{2\sigma_{mn}^2} \right) \times \\ \times \exp \left(-\frac{\left(z_{N/2+1} - x_{Im_{10}}^j \right)^2 + \dots + \left(z_N - x_{Im_{mn}}^j \right)^2}{2\sigma_{mn}^2} \right) dz_1 \dots dz_N, \quad (7)$$

где N — число признаков распознавания; Ω_j — область j -го класса в пространстве признаков, определяемая по критерию идеального наблюдателя при равновероятном появлении объектов обоих классов.

С целью проверки адекватности предложенной модели зрительной системы и выведенных математических соотношений были проведены экспериментальные исследования. На первом этапе исследований оценивали вероятность различения двух символов, наблюдаемых на равномерном фоне при наличии аддитивного белого гауссова шума. Измерения выполняли для двух пар символов (рис. 2), причем каждая пара была представлена в двух различных масштабах. Эксперименты проводили при динамическом шуме, реализации которого обновлялись с частотой кадров $\nu_k = 25$ Гц. Формирование таких изображений на экране монитора ЭВМ осуществлялось с помощью специально написанной компьютерной программы.



Рис. 2. Пары символов, предъявляемые для распознавания

В экспериментах принимали участие четырнадцать предварительно обученных операторов. Каждому оператору предъявлялся зашумленный символ (рис. 3) и ставилась задача определить, какой из предложенной пары символов был ему предъявлен. Когда оператор затруднялся сделать выбор, он мог отметить этот факт. При обработке экспериментальных данных такому исходу испытания ставилось в соответствие значение вероятности распознавания, равное 0,5. При выборе правильного ответа значение среднего квадратического от-

клонения (СКО) шума на следующем предъявляемом изображении увеличивалось. В случае выбора неправильного ответа или непринятия оператором решения, значение СКО шума на следующем предъявляемом изображении уменьшалось. При последующих правильных ответах значение, на которое увеличивалось СКО, уменьшалось. Этот прием при проведении испытаний позволил оценить вероятность распознавания в наиболее интересном для последующего анализа диапазоне отношений сигнал/шум.

После того как оператор совершал несколько ошибок, ему предъявлялись символы из следующей пары. Условия наблюдения — среднюю яркость экрана, расстояние наблюдения и время предъявления символа — каждый оператор мог выбирать по своему усмотрению.

По результатам, полученным при проведении экспериментов, были определены значения вероятности распознавания человеком оператором изображений символов различных размеров и при различных значениях пикового отношения сигнал/шум. На рис. 4, а, б представлены экспериментальные значения вероятности распознавания пары символов L и Q размером 16×24 и 8×12 пикселей, а также расчетные зависимости вероятности распознавания изображений этих символов от пикового отношения сигнал/шум.

Как показано в работе [9], для распознавания хорошо известных изображений объектов достаточно использовать от трех до пяти гармоник по каждому из сечений пространственного спектра объектов. Поэтому при расчете вероятности распознавания символов L и Q использовались первые три гармоники по каждому из сечений пространственного спектра символов, т. е. восемь нижних гармоник. На рис. 4, в, г приведены аналогичные данные, полученные для пары символов I и I . Однако при расчете вероятности их распознавания использовалось большее число гармонических составляющих, чем для предыдущей пары символов, а именно 23 гармоники, т. е. вся значимая часть пространственно-частотного спектра изображения. Это обусловлено тем, что символы I и I очень похожи, и информация об их различии содержится в более высоких гармониках.

Из анализа полученных результатов можно сделать следующие выводы:

- экспериментальные данные отличаются от расчетных не более чем на 15 %;
- наблюдаемое к концу эксперимента увеличение частоты принятия правильных решений по сравнению с расчетной можно объяснить обучением операторов. Это наглядно иллюстрирует зависимость, приведенная на рис. 4, з.

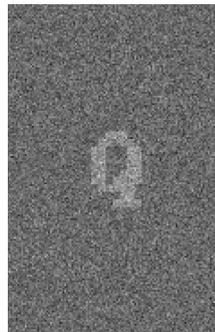


Рис. 3. Зашумленный символ

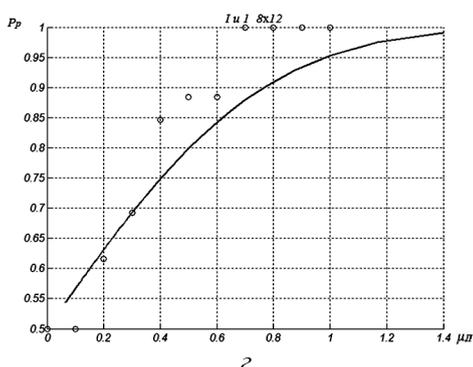
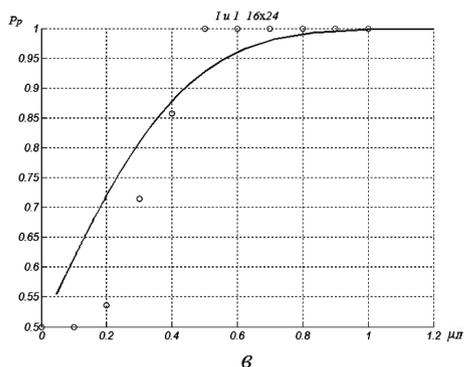
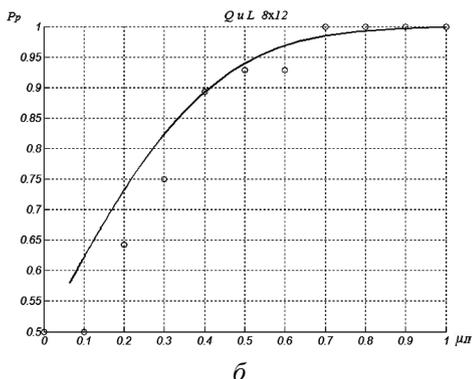
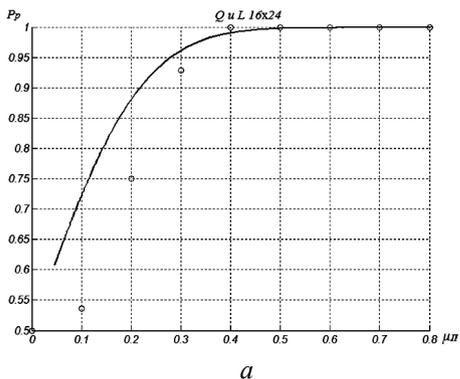


Рис. 4. Зависимость вероятности распознавания от пикового отношения сигнал/шум для различных пар символов при вариациях размеров этих символов: *а* — пара символов *L* и *Q* с размером 16×24 пикселей; *б* — пара символов *L* и *Q* с размером 8×12 пикселей; *в* — пара символов *I* и *I* с размером 16×24 пикселей; *г* — пара символов *I* и *I* с размером 8×12 пикселей; черточки — теоретически рассчитанная зависимость; кружки — экспериментально полученные значения

Следует отметить, что проведенные исследования позволили провести количественную оценку возможности извлечения из стегообъектов скрытых данных, наблюдаемых человеком-оператором на фоне аддитивных помех. Для количественной оценки надежности извлечения скрытых символов целесообразно использовать вероятность распознавания.

СПИСОК ЛИТЕРАТУРЫ

1. Грибунин В. Г., Оков И. Н., Туринцев И. В. Цифровая стеганография. М.: Солон-Пресс. 2002. 272 с.
2. Marvel L. M., Retter C. T. The Use of Side Information in Image Steganography // International Symposium on Information Theory and Its Applications, Honolulu, Hawaii, USA., November 5–8, 2000.
3. Tulasi Y. V. N., Krishna Prasad T. S. R., Narayana S. Steganography – Security through Images // 2nd International conference in RF Signal Processing Systems. P. 452–455.

4. Красильников Н. Н. Теория передачи и воспроизведения изображений. Теория передачи изображений и ее приложения. М.: Радио и связь, 1986. 248 с.
5. Сивяков И. Н. Исследование информационных возможностей оптико-электронных систем наблюдения: автореф. дис. ...канд. техн. наук. СПб., 2007. 20 с.
6. Пространственное зрение / В.М. Бондаренко, М.В. Данилова, Н.Н. Красильников и др. СПб.: Наука. 1999. 218 с.
7. Бенуни А. А., Колючкин В. Я. Модель зрительной системы человека-оператора // Вестник МГТУ им. Н.Э. Баумана. Сер. Приборостроение. 2002. № 4. С. 43–52.
8. Гулина Ю. С., Колючкин В. Я. Методика расчета вероятности распознавания изображений человеком-оператором // Вестник МГТУ им. Н.Э. Баумана. Сер. Приборостроение. 2012. № 1. С. 100–107.
9. Глезер В. Д. Зрение и мышление. Л.: Наука, 1985. 246 с.

Статья поступила в редакцию 28.09.2012.