

УДК 621.397.2

Ю. С. Гулина, В. Я. Колючкин

МЕТОДИКА РАСЧЕТА ВЕРОЯТНОСТИ РАСПОЗНАВАНИЯ ИЗОБРАЖЕНИЙ ЧЕЛОВЕКОМ-ОПЕРАТОРОМ

Изложена методика расчета вероятности распознавания образов объектов, наблюдаемых человеком-оператором на выходе оптико-электронных приборов наблюдения, основанная на многоканальной модели зрительной системы человека. Приведены результаты экспериментальных исследований, подтверждающих правильность предложенной методики, которая может быть использована при проектировании оптико-электронных приборов наблюдения.

E-mail: julia-sg@yandex.ru

Ключевые слова: оптико-электронные приборы наблюдения, вероятность распознавания, зрительная система, человек-оператор.

Основными показателями качества оптико-электронных приборов (ОЭП) наблюдения являются условные вероятности обнаружения образов наблюдаемых объектов, а также условные вероятности распознавания этих образов человеком-оператором. Предлагаемая методика позволяет учесть не только характеристики излучения объектов и ОЭП наблюдения, но также свойства зрительного восприятия зашумленных изображений человеком-оператором.

Методика основана на многоканальной модели зрительной системы человека-оператора [1]. Согласно этой модели в зрительной системе осуществляется оптимальный прием и последующий анализ финитных пространственно-частотных гармонических составляющих образа объекта, на основе которого в соответствии с критерием максимума апостериорной вероятности (критерием идеального наблюдателя [2]) принимается решение об обнаружении или распознавании образов объектов.

Решение об обнаружении объекта принимается на основе одной или нескольких пространственно-частотных гармоник. Вероятность правильного обнаружения объекта в предположении о независимости каналов, в которых происходит анализ гармоник, описывается выражением [1]

$$P_{\text{обн}} = 1 - \prod_{m,n} (1 - P_{\text{обн}mn}); \quad (1)$$

здесь $P_{\text{обн}mn}$ — вероятность правильного обнаружения mn -й гармоники, которая в свою очередь определяется формулой

$$P_{\text{обн}mn} = \frac{1}{\sqrt{2\pi}} \int_{\zeta_{mn}}^{\infty} \exp\left(-\frac{(\zeta_{mn} - \mu_{mn})^2}{2}\right) d\zeta, \quad (2)$$

где ζ_{mn} — порог принятия решения об обнаружении mn -й гармоники, выраженный в нормированных на среднеквадратическое отклонение (СКО) шума значениях реализаций сигнала на выходе mn -го фильтра и равный

$$\zeta_{mn} = \frac{\ln \Lambda_n}{\mu_{mn}} + \frac{\mu_{mn}}{2}; \quad (3)$$

μ_{mn} — воспринимаемое отношение сигнала к шуму (ОСШ) на mn -й гармонике; Λ_n — пороговое отношение правдоподобия.

Как указано в работе [1], воспринимаемое ОСШ определяется формулами:

$$\mu_{00} = \mu_{\Pi} \left| \tilde{L}'_{\Pi}(0, 0) \right| \sqrt{\frac{2\nu_k \tau_{\text{зр}} A_o}{A_{\text{ш}}}}; \quad (4)$$

$$\mu_{mn} = 2\mu_{\Pi} \sqrt{\frac{2\nu_k \tau_{\text{зр}} A_o}{A_{\text{ш}}}} \cdot \left| \tilde{L}'_{\Pi}\left(\frac{m}{l_x}, \frac{n}{l_y}\right) \right| \cdot \left| \tilde{H}_k\left(\frac{m}{l_x}, \frac{n}{l_y}\right) \right|, \quad (5)$$

где μ_{Π} — пиковое ОСШ; ν_k — частота кадров в ОЭП наблюдения; $\tau_{\text{зр}}$ — постоянная времени, описывающая инерционность зрительного восприятия; A_o — габаритная площадь изображения объекта; $A_{\text{ш}}$ — площадь корреляции аддитивного шума; $\tilde{L}'_{\Pi}\left(\frac{m}{l_x}, \frac{n}{l_y}\right)$ — фурье-образ нормированной функции, описывающей изображение объекта с учетом линейных искажений, которые вносит тракт ОЭП наблюдения; $\tilde{H}_k\left(\frac{m}{l_x}, \frac{n}{l_y}\right)$ — функция контрастной чувствительности [2].

Отметим, что модель зрительной системы человека-оператора [1] хорошо согласуется с экспериментальными данными при решении задачи обнаружения образов объектов. Однако в этом модельном описании проводится обнаружение отдельных гармонических составляющих образов объектов путем сравнения с соответствующими пороговыми значениями. Но для решения задачи распознавания необходимо проводить анализ значений гармоник, рассматриваемых как признаки распознавания, в многомерном признаковом пространстве. Поэтому модель, предложенную в работе [1], нельзя использовать для оценки вероятности распознавания образов объектов.

Предлагаемая методика оценки процесса распознавания изображений человеком-оператором основана на гипотезе о том, что в зрительной системе проводится анализ признаков, которыми являются модули

финитных гармонических составляющих образа объекта, а решение о принадлежности образов объектов к тому или иному классу принимается в соответствии с критерием идеального наблюдателя [2].

Чтобы отнести объект к тому или иному классу, нужно, во-первых, выбрать эффективный словарь признаков распознавания — набор гармоник, достаточных для принятия решения о распознавании, и описать каждый класс объектов, подлежащих распознаванию, в признаковом пространстве. В рамках статистической теории распознавания образов такое описание задается функциями правдоподобия. Во-вторых, требуется определить в пространстве признаков границы между классами, т.е. найти правило, согласно которому будет приниматься решение о принадлежности объекта к тому или иному классу.

С учетом сказанного были внесены коррективы в многоканальную модель зрительной системы человека-оператора [1] для адаптации ее к решению задачи распознавания образов объектов. В частности, блоки пороговых устройств из каждого канала исключены. Вместо них в модифицированную структурную схему модели многоканальной зрительной системы введен блок определения решающих границ (рис. 1). В предпроцессоре отдельные финитные гармонические составляющие образов объектов первично обрабатываются и анализируются, а в логическом блоке строятся решающие границы в признаковом пространстве, образованном совокупностью гармоник, для принятия решения о распознавании.

Рассмотрим задачу распознавания на примере двух классов объектов, изображения которых наблюдаются на экране видеоконтрольного устройства при наличии аддитивного шума с нормальной статистикой. Значения реализаций сигналов на выходе каждого m -го канала представляют собой нормально распределенные случайные величины,

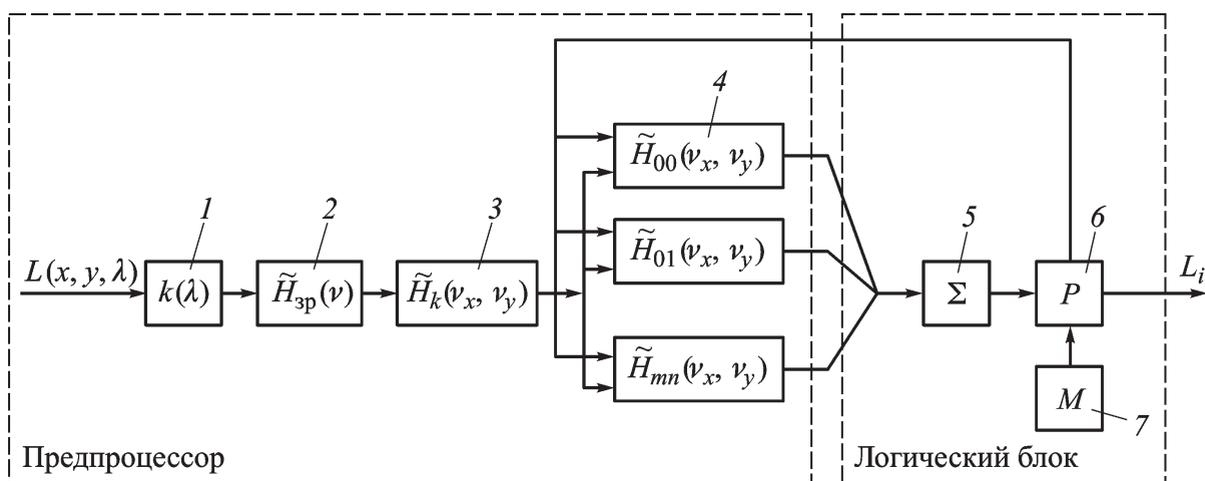


Рис. 1. Структурная схема многоканальной модели зрительной системы человека-оператора при распознавании:

1 — оптический фильтр; 2 — временной фильтр; 3 — пространственный фильтр; 4 — параллельные фильтры; 5 — блок определения решающих границ; 6 — процессор; 7 — блок памяти

которые характеризуются математическим ожиданием и дисперсией. При принятии решения о распознавании двух классов объектов, у которых признаки распознавания являются нормально распределенными случайными величинами, в соответствии с критерием идеального наблюдателя решающее правило можно представить в виде квадратичной формы относительно вектора наблюдения \mathbf{X} следующим образом [3]:

$$\frac{1}{2} (\mathbf{X} - \mathbf{M}_1)^T \Sigma_1^{-1} (\mathbf{X} - \mathbf{M}_1) - \frac{1}{2} (\mathbf{X} - \mathbf{M}_2)^T \Sigma_2^{-1} (\mathbf{X} - \mathbf{M}_2) + \frac{1}{2} \ln \frac{\Sigma_1}{\Sigma_2} \leq \ln \frac{P(\omega_1)}{P(\omega_2)} \rightarrow \mathbf{X} \in \begin{cases} \omega_1 \\ \omega_2 \end{cases}, \quad (6)$$

где $\mathbf{X} = [x_1 \ x_2 \ \dots \ x_n]^T$ — случайный вектор, характеризующий реализации сигнала в n -мерном признаковом пространстве; $\mathbf{M}_j = [m_{1j} \ m_{2j} \ \dots \ m_{nj}]^T$ — вектор математического ожидания случайного вектора \mathbf{X} ; $\Sigma = \begin{bmatrix} \sigma_{11}^2 & \dots & \sigma_{1n}^2 \\ \dots & \dots & \dots \\ \sigma_{n1}^2 & \dots & \sigma_{nn}^2 \end{bmatrix}$ — ковариационная матрица аддитивной помехи, элементами которой являются числа $\sigma_{ij}^2 = \mathbf{M}\{(x_i - m_i) \times (x_j - m_j)\}$; $P(\omega_j)$ — априорные вероятности появления объектов, принадлежащих к классу ω_j .

При искажении изображений объектов белым шумом ковариационная матрица преобразуется к диагональному виду, ее элементы одинаковы и равны значениям дисперсии σ_{mn}^2 помехи на выходе каждого mn -го канала. Компоненты вектора \mathbf{X} при этом не коррелированы, а байесовское решающее правило принимает вид линейной функции относительно \mathbf{X} [3]:

$$(\mathbf{M}_2 - \mathbf{M}_1)^T \mathbf{X} + \frac{1}{2} (\mathbf{M}_1^T \mathbf{M}_1 - \mathbf{M}_2^T \mathbf{M}_2) \leq \ln \frac{P(\omega_1)}{P(\omega_2)} \rightarrow \mathbf{X} \in \begin{cases} \omega_1 \\ \omega_2 \end{cases}. \quad (7)$$

Предположим, что появление объектов различных классов равновероятно. Тогда априорные вероятности, характеризующие принадлежность объектов к тому или иному классу, будут равны, т.е. $P(\omega_1) = P(\omega_2) = 0,5$, и решающее правило имеет вид

$$(\mathbf{M}_2 - \mathbf{M}_1)^T \mathbf{X} + \frac{1}{2} (\mathbf{M}_1^T \mathbf{M}_1 - \mathbf{M}_2^T \mathbf{M}_2) \leq 0 \rightarrow \mathbf{X} \in \begin{cases} \omega_1 \\ \omega_2 \end{cases}. \quad (8)$$

Границы между классами можно определить как решение системы уравнений

$$(\mathbf{M}_2 - \mathbf{M}_1)^T \mathbf{X} + \frac{1}{2} (\mathbf{M}_1^T \mathbf{M}_1 - \mathbf{M}_2^T \mathbf{M}_2) = 0. \quad (9)$$

Решение системы (9) представляет собой многомерную функцию, описывающую гиперплоскость, разделяющую области двух классов объектов в признаковом пространстве.

Если значения реализаций сигналов на выходе каждого mn -го канала пронормировать на СКО σ_{mn} аддитивного шума, то математическими ожиданиями для случайных векторов, описывающих образы объектов, будут являться значения воспринимаемых ОСШ на mn -х гармониках спектров изображений объектов. С учетом этого замечания можно ввести следующие обозначения нормированного вектора реализаций и нормированного вектора математического ожидания:

$$\mathbf{Z} = \frac{1}{\sigma_{mn}} [x_1, x_2, \dots, x_n]^T = [\zeta_1, \zeta_2, \dots, \zeta_n]^T,$$

$$\mathbf{Mu}_j = \frac{1}{\sigma_{mn}} \mathbf{X}_j = [\mu_{00}^j, \mu_{10}^j, \dots, \mu_{mn}^j]^T.$$

Систему уравнений для нахождения границ между классами в признаковом пространстве нормированных гармонических составляющих можно представить в виде

$$(\mathbf{Mu}_2 - \mathbf{Mu}_1)^T \mathbf{Z} + \frac{1}{2} (\mathbf{Mu}_1^T \mathbf{Mu}_1 - \mathbf{Mu}_2^T \mathbf{Mu}_2) = 0. \quad (10)$$

Вероятность правильного распознавания j -го класса

$$P_{\text{пр}j} = \int_{\Omega_j} \frac{1}{(2\pi)^{\frac{n}{2}}} \exp \left(-\frac{(\zeta_{00} - \mu_{00}^j)^2 + \dots + (\zeta_{10} - \mu_{10}^j)^2}{2} \right) d\zeta_{00} \dots d\zeta_{mn}, \quad (11)$$

где Ω_j — область j -го класса в пространстве признаков.

Безусловная вероятность правильного распознавания любого из двух классов объектов

$$P_{\text{пр}} = P(\omega_1) P_{\text{пр}1} + P(\omega_2) P_{\text{пр}2}. \quad (12)$$

Для проверки адекватности предложенной модели были проведены экспериментальные исследования. В экспериментах принимали участие 11 специально обученных операторов. В процессе обучения на мониторе видеоконтрольного устройства операторам предъявлялись полутонные изображения объектов различного масштаба на равномерном фоне. Типовые изображения объектов приведены на рис. 2.

При проведении экспериментов операторам предъявлялись зашумленные изображения объектов (рис. 3) и ставилась задача определить принадлежность к одному из двух классов. Значение СКО шума постепенно увеличивалось до такой степени, когда ни один из операторов

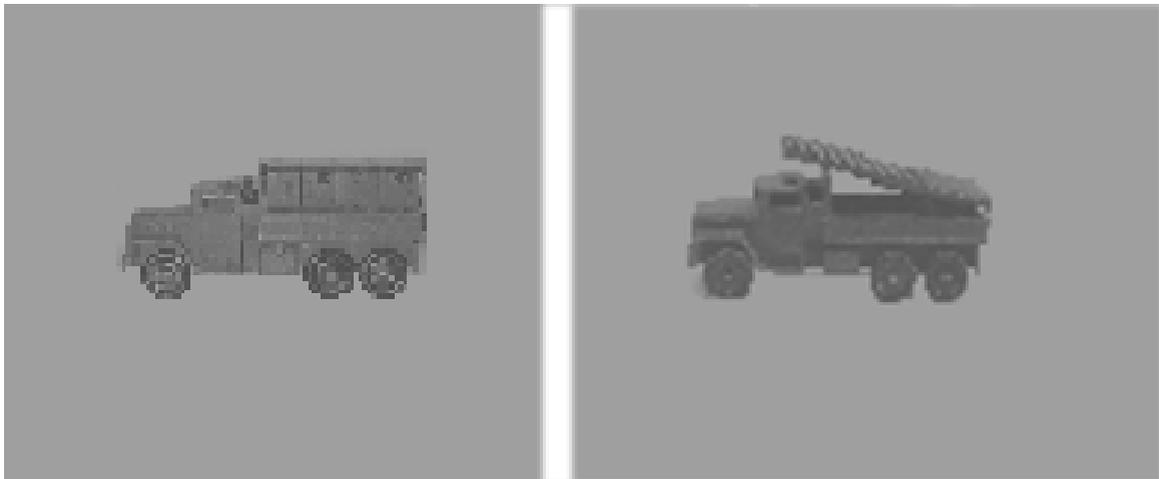


Рис. 2. Типовые изображения объектов

уже не мог распознать объект. Условия наблюдения, а именно среднюю яркость экрана и расстояние наблюдения, каждый оператор мог выбирать по своему усмотрению.

По результатам, полученным в ходе экспериментов, были определены значения вероятности распознавания изображений объектов различных размеров и при разных значениях пикового ОСШ. На рис. 4 приведены результаты экспериментов, полученные при распознавании изображений объектов, имеющих габаритные размеры 38×68 пикселей и наблюдаемые при различных значениях пикового ОСШ, а также расчетные графики зависимости вероятности распознавания таких изображений от пикового ОСШ. Как следует из приведенных на рис. 4 результатов, экспериментальная и расчетные зависимости различаются, причем при увеличении числа признаков в области значений пикового ОСШ, меньших 0,15, различие возрастает до 11 %, при больших ОСШ — уменьшается до 2 %. При четырех признаках расчетная вероятность распознавания и экспериментальная различаются не более чем на 5 % во всем диапазоне значений пикового ОСШ.

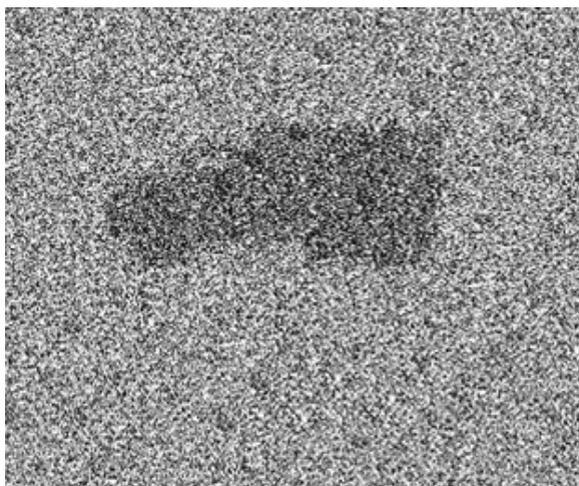


Рис. 3. Изображение объекта, искаженное белым шумом

Экспериментальные и расчетные результаты исследований влияния размеров изображений на вероятность распознавания приведены на рис. 5. При расчетах в качестве признаков использовались амплитуды гармоник, в том числе амплитуда \tilde{L}_{00} нулевой гармоники и амплитуды \tilde{L}_{01} , \tilde{L}_{10} , \tilde{L}_{11} первых гармонических составляющих образов объектов. Анализируя полученные результаты, выявили, что различие

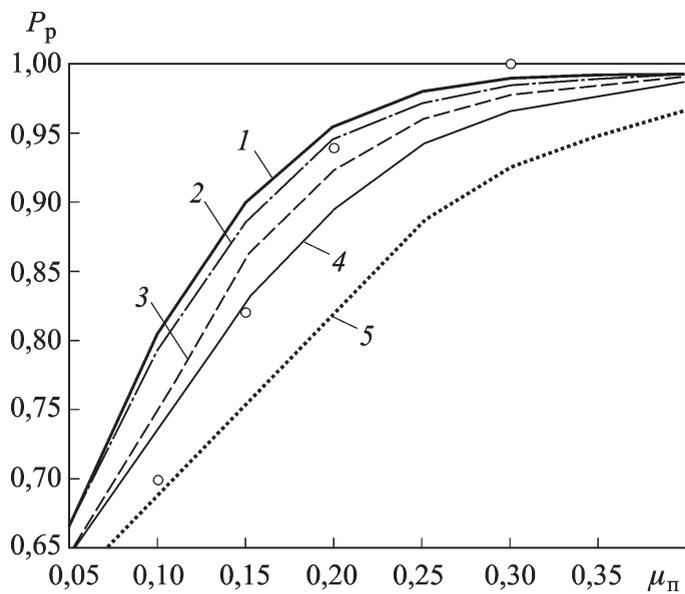


Рис. 4. Зависимость вероятности распознавания объекта размером 38×68 пикселей от пикового ОСШ для различного числа признаков: 1 — расчетная зависимость при 6 признаках (амплитуды гармоник $\tilde{L}_{00}, \tilde{L}_{01}, \tilde{L}_{10}, \tilde{L}_{11}, \tilde{L}_{20}, \tilde{L}_{21}$); 2 — то же при 5 признаках (амплитуды гармоник $\tilde{L}_{00}, \tilde{L}_{01}, \tilde{L}_{10}, \tilde{L}_{11}, \tilde{L}_{20}$); 3 — то же при 4 признаках (амплитуды гармоник $\tilde{L}_{00}, \tilde{L}_{01}, \tilde{L}_{10}, \tilde{L}_{11}$); 4 — то же при 3 признаках (амплитуды гармоник $\tilde{L}_{00}, \tilde{L}_{01}, \tilde{L}_{10}$); 5 — то же при 2 признаках (амплитуды гармоник $\tilde{L}_{00}, \tilde{L}_{10}$); \circ — экспериментально полученные значения

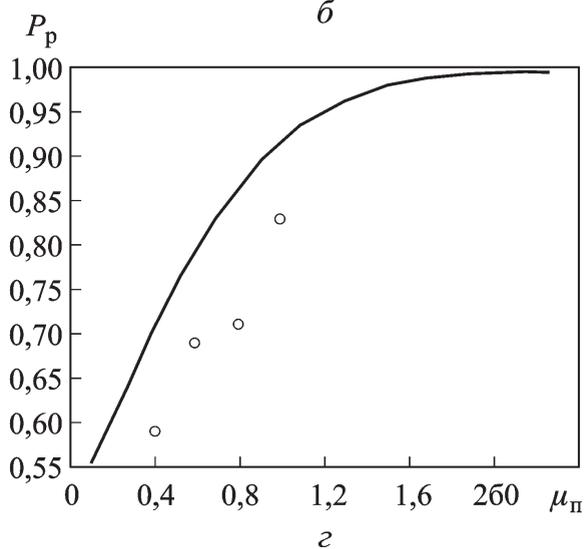
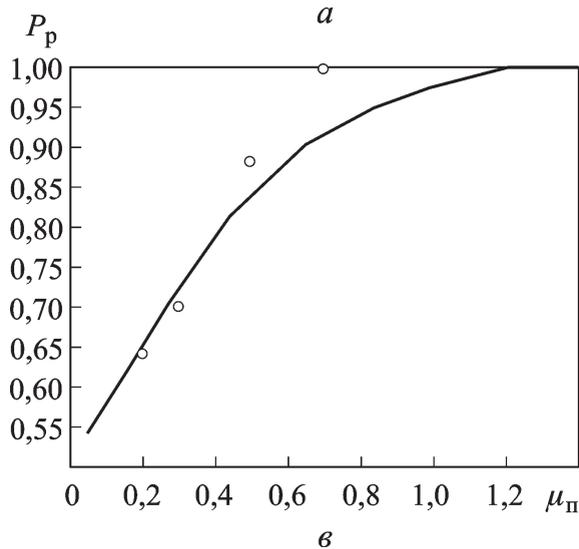
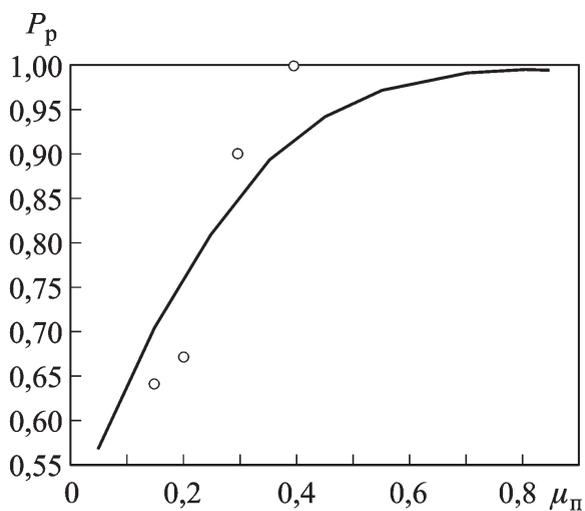
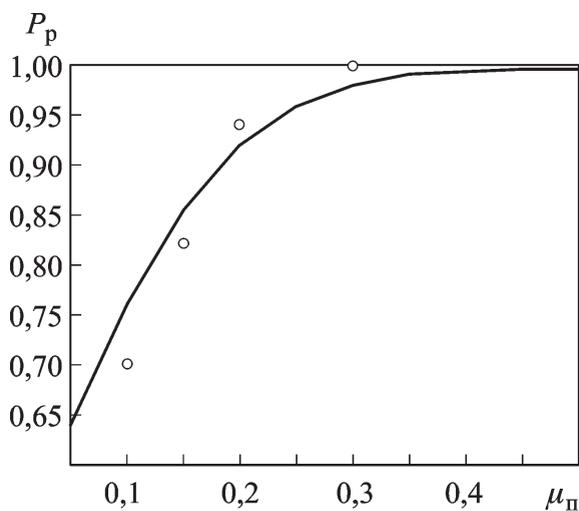


Рис. 5. Зависимость вероятности распознавания от пикового ОСШ при вариациях размеров изображений объектов:

— — теоретически рассчитанная зависимость; \circ — экспериментально полученные значения 38×68 пикселей (а); 19×34 пикселей (б); 10×17 пикселей (в); 5×9 пикселей (г)

экспериментальных и расчетных значений вероятности не превышает 8% при размерах изображений объектов от 38×68 пикселей до 10×17 пикселей, но при размерах изображений 5×9 пикселей отличие экспериментальных от расчетных значений достигает 15%. Это свидетельствует о том, что при расчетах использовалось недостаточное число признаков распознавания. Кроме этого не учтено влияние помех пространственной дискретизации, возникающих при формировании изображений.

В заключение отметим, что полученные в ходе исследований результаты свидетельствуют о правильности предложенного подхода к описанию свойств зрительной системы, а изложенная методика может быть использована для приближенной оценки вероятности распознавания зашумленных изображений, наблюдаемых человеком-оператором на экране видеоконтрольного устройства ОЭП наблюдения.

СПИСОК ЛИТЕРАТУРЫ

1. Б е н у н и А. А., К о л ю ч к и н В. Я. Модель зрительной системы человека-оператора // Вестник МГТУ им. Н.Э. Баумана. Сер. Приборостроение. – 2002. – № 4. – С. 43–52.
2. К р а с и л ь н и к о в Н. Н. Теория передачи и воспроизведения изображений. Теория передачи изображений и ее приложения. – М.: Радио и связь, 1986. – 248 с.
3. Ф у к у н а г а К. Введение в статистическую теорию распознавания образов: Пер. с англ. – М.: Наука, 1979. – 368 с.

Статья поступила в редакцию 18.05.2011

Юлия Сергеевна Гулина родилась в 1986 г., окончила МГТУ им. Н.Э. Баумана в 2009 г. Аспирантка кафедры “Лазерные и оптико-электронные системы” МГТУ им. Н.Э. Баумана. Специализируется в области проектирования оптико-электронных приборов наблюдения.

Yu.S. Gulina (b. 1986) graduated from the Bauman Moscow State Technical University in 2009, Post-graduate of “Laser and Optical-Electronic Systems” department of the Bauman Moscow State Technical University. Specializes in the field of design of optical-electronic observation instruments.

Василий Яковлевич Колочкин родился в 1948 г., окончил в 1972 г. МВТУ им. Н.Э. Баумана и в 1986 г. МИФИ. Д-р техн. наук, профессор кафедры “Лазерные и оптико-электронные системы” МГТУ им. Н.Э. Баумана. Автор более 100 научных работ в области системного проектирования оптико-электронных приборов.

V.Ya. Kolyuchkin (b. 1948) graduated from the Bauman Moscow Higher Technical School in 1972. D. Sc. (Eng.), professor of “Laser and Optical-Electronic Systems” department of the Bauman Moscow State Technical University. Author of 125 publications in the field of system design of optical-electronic instruments.