

Для оценки устойчивости предложенного алгоритма к помехам типа «световое пятно» и аддитивному гауссовому шуму на ЭВМ синтезировались эти помехи. Моделирование работы алгоритма проводилось при изменении отношения сигнал/шум μ от 2 до 15. Эксперименты показали, что для широкого класса объектов вероятность правильного определения их местоположения была не меньше 0,95 при $\mu \geq 5$. Работоспособность алгоритма практически не зависит от наличия помех типа «световое пятно», от изменения параметров геометрических преобразований в пределах 10 %, а также от изменения контрастности объектов относительно фона в пределах 30 %. Эффективность алгоритма зависит от близости геометрической формы искомым и ложных объектов.

В результате экспериментов установлено, что разработанный алгоритм позволяет с высокими показателями достоверности и быстродействия решать практические задачи анализа изображений.

СПИСОК ЛИТЕРАТУРЫ

1. Гороховатский В. А., Кацалан С. Ф., Пуятин Е. П. Анализ изображений в условиях локальных искажений // Автометрия.— 1986.— № 6.
2. Прэтт У. Цифровая обработка изображений.— М.: Мир, 1982.— Кн. 1, 2.
3. Павлидис Т. Алгоритмы машинной графики и обработки изображений — М.: Радио и связь, 1986.
4. Распознавание образов и обработки изображений // ТИИЭР.— 1979.— 67, № 5.
5. Гуткин Л. С. Теория оптимальных методов радиоприема при флуктуационных помехах.— М.: Сов. радио, 1972.
6. Пуятин Е. П., Гороховатский В. А., Ерьско Ю. Н., Стрельченко В. И. Метод экстремально-логической фильтрации в задачах автоматической сегментации изображений // АСУ и приборы автоматики.— Харьков: Вища шк., 1987.— Вып. 83.

Поступила в редакцию 2 июля 1987 г.

УДК 621.397

Н. Н. КРАСИЛЬНИКОВ

(Ленинград)

ОБОБЩЕННАЯ ФУНКЦИОНАЛЬНАЯ МОДЕЛЬ ЗРЕНИЯ И ЕЕ ПРИМЕНЕНИЕ В СИСТЕМАХ ОБРАБОТКИ И ПЕРЕДАЧИ ИЗОБРАЖЕНИЙ

При проектировании и оптимизации параметров систем обработки и передачи изображений серьезные трудности вызывает отсутствие формализованного критерия качества воспроизведения изображений, функционально связанного с электрооптическими характеристиками самой системы.

В настоящее время для оценки качества воспроизведения изображения телевизионной системой используется либо метод экспертных оценок, либо набор так называемых частных критериев, к числу которых относятся: разрешающая способность, отношение сигнал/шум, число воспроизводимых градаций яркости, погрешности воспроизведения цвета и т. д. Их значения определяются параметрами системы и могут быть вычислены. Недостатком первого метода является отсутствие математических выражений, устанавливающих связь между оценкой и электрооптическими характеристиками аппаратуры, недостатком второго — невозможность сравнивать системы, которые различаются между собой по разным частным критериям. Так, например, если система *A* обеспечивает большую разрешающую способность, но меньшее отношение сигнал/шум, чем система *B*, то, не опираясь на опыт, невозможно сказать, какая из них лучше. Из этого следует, что метод экспертных оценок не может быть использован при аналитическом решении оптимизационных задач,

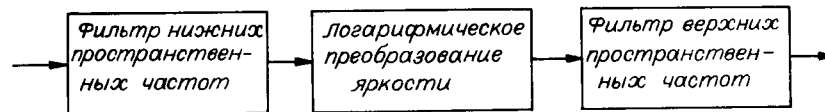


Рис. 1. Функциональная модель зрения, предложенная Ч. Холлом и Э. Холлом

а метод частных критериев — при решении оптимизационных задач, в которых имеют место обменные операции между отдельными частными критериями качества, как, например, в известной задаче апертурной коррекции при наличии шума.

Попытки синтезировать обобщенный критерий качества путем комбинирования частных критериев, а также другие эвристические подходы к решению этой проблемы, как известно, успеха не имели. Объясняется это тем, что в предлагавшихся критериях свойства зрения наблюдателя либо вообще не принимались во внимание, либо принимались, но не в достаточной мере.

Значительные успехи в направлении учета свойств зрения при создании критериев качества были достигнуты путем разработки и применения для этой цели так называемых функциональных моделей зрения. В качестве примера на рис. 1 показана функциональная модель ахроматического зрения, предложенная Ч. Холлом и Е. Холлом. Первое звено этой модели — фильтр нижних пространственных частот — соответствует линейным преобразованиям изображения в оптической части глаз, второе — поточечное логарифмическое преобразование яркости в величину, пропорциональную ее ощущению, и, наконец, третье звено модели — фильтр верхних пространственных частот — учитывает линейные преобразования при формировании так называемого нейронного изображения [1]. Совместное действие фильтров нижних и верхних пространственных частот приводит к усилению в нейронном изображении спектральных составляющих средних пространственных частот.

Рассмотренная функциональная модель позволяет получать надежные оценки видности флуктуационного шума на изображении путем вычисления среднего квадрата разности между зашумленным и незашумленным компонентами на ее выходе. Так, например, из применения этой модели следует, что шум, энергия которого в основном сосредоточена в области верхних частот, будет меньше заметен на изображении, чем шум с равномерным распределением энергии по спектру, что, как известно, подтверждается экспериментально.

Не останавливаясь на других функциональных моделях зрения, сведения о которых читатель найдет в [2], отметим, что, к сожалению, они позволяют получать надежные оценки только видности флуктуационного шума на изображении. Попытки оценить видность других искажений с помощью этих моделей, например ухудшения резкости при апертурных искажениях, приводят к результатам, которые расходятся с оценками зрителей [2]. Проведенный нами анализ причин несовершенства оценок видности на изображениях искажений, не связанных с флуктуационным шумом (речь идет об оценках, полученных путем использования указанных функциональных моделей зрения), выявил общие для них недостатки, которые обусловлены тем, что ряд важных свойств зрения в этих моделях не учитывается.

На основании нашего анализа предложена обобщенная функциональная модель зрения человека, свободная от перечисленных недостатков, которая показана на рис. 2. В обобщенной функциональной модели, помимо традиционно учитываемых свойств зрения: линейного преобразования изображения $L(x, y, \lambda)$ оптикой глаза ($L(x, y, \lambda)$ — функция, описывающая зависимость яркости изображения от декартовых координат x, y и длины волны λ); цветоделения изображения сетчаткой дневного зрения; логарифмического поточечного преобразования яркости в величину, соответствующую ее ощущению; формирования ахроматиче-

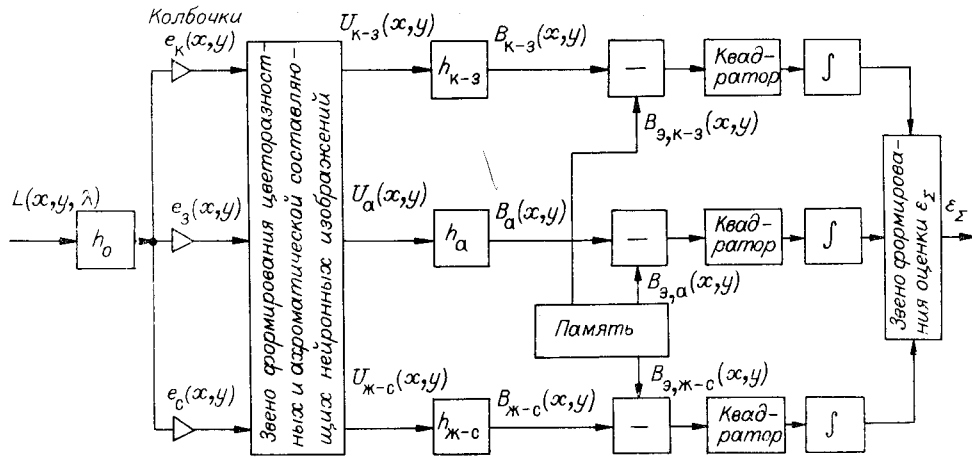


Рис. 2. Обобщенная функциональная модель зрения

ской и цветоразностных составляющих изображения; линейного преобразования при формировании нейронного изображения, определяемого латеральным торможением в ганглиозных клетках, — дополнительно учтены и другие важные для решения рассматриваемой проблемы свойства зрения:

1. Ограниченность поля с высокой остротой зрения (зона ясного видения) углом около 1° , обусловленная тем, что лишь небольшая часть сетчатки глаза — так называемая область фовеа, диаметр которой 0,3 мм, имеет высокую плотность расположения рецепторов, за пределами которой она быстро убывает (это обстоятельство наглядно иллюстрируется рис. 3 [3]).

2. Наличие саккадических (скачкообразных) движений глаз во время рассматривания изображения, при которых ось зрения фиксируется вблизи его контуров (рис. 4, а — в) [4].

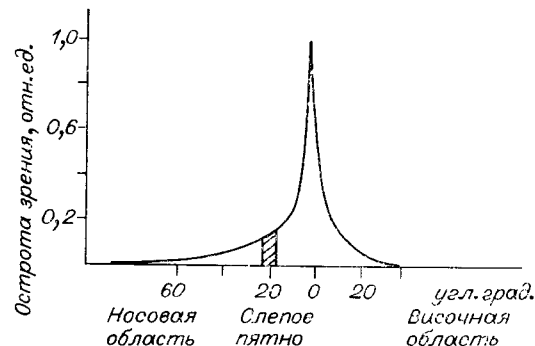
3. Подавление средней составляющей после поточечного логарифмического преобразования изображения в рецепторах, благодаря чему достигается инвариантность к яркости нейронного изображения [5] (в психофизических опытах следствие этого преобразования проявляется в виде явления одновременного контраста).

4. Уменьшение контраста изображения, обусловленное несовершенством оптики глазных сред.

5. Найденная и изученная нами закономерность различения объектов (фрагментов) на ахроматических изображениях, искаженных помехой [6], которая впоследствии была подтверждена Бернардом [7]. Сущность этой закономерности состоит в том, что обнаружение и различение объектов (фрагментов изображений) на искаженных изображениях выполняются в соответствии со среднеквадратичной мерой близости искаженного и неискаженного изображений, что обусловлено пространственной фильтрацией изображений в высших отделах зрительной системы. Из этой закономерности, в частности, следует, что в зрительной системе человека имеется механизм,

Рис. 3. Зависимость остроты зрения от места раздражения на сетчатке:

по оси абсцисс отложено угловое расстояние проекции изображения раздражителя от середины фовеа сетчатки (град); по оси ординат — острота зрения (отн. ед.)



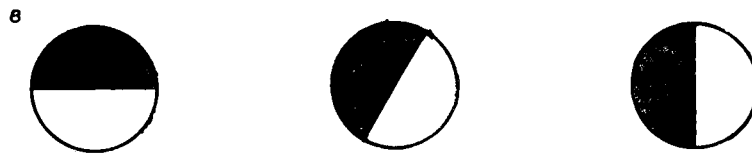
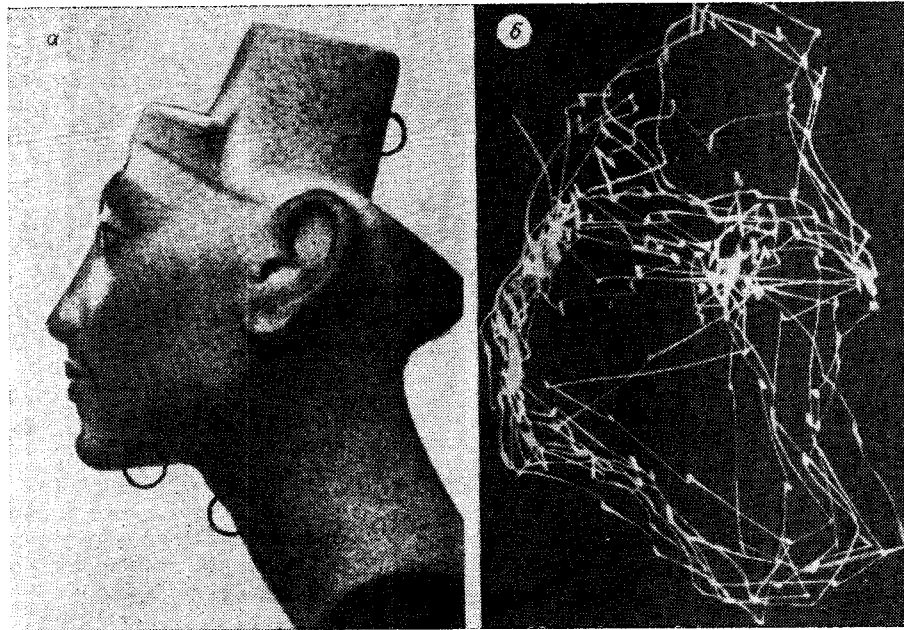


Рис. 4. К пояснению процесса оценки видности искажений:
a — изображение; окружностями отмечены зоны ясного видения, рассчитанные для случая, когда расстояние рассматривания в 4 раза превышает высоту изображения; *б* — точки фиксации и траектории движения оси зрения при рассматривании изображения; *в* — фрагменты контуров в зоне ясного видения

посредством которого формируется («вычисляется») некоторая величина, являющаяся монотонной функцией среднего квадрата разности сравниваемых нейронных изображений. Логично предположить, что эта величина будет мерой видности искажений на изображениях. Результаты экспериментальных исследований подтверждают данное предположение.

Согласно предложенной нами обобщенной функциональной модели зрения, процесс наблюдения изображения протекает следующим образом. Ось зрения, скачкообразно изменяя свое положение вследствие саккадических движений глаз, фиксируется вблизи контуров (см. рис. 4, *a, б*) так, что в зоне ясного видения располагаются фрагменты контуров, подобные тем, что показаны на рис. 4, *в*. Поскольку одновременно все изображение в деталях зритель рассмотреть не может вследствие ограниченности зоны ясного видения, полное изображение наблюдаемой сцены формируется в его памяти последовательно во времени (на так называемом инвариантном экране детекторов в терминологии психологов). Так как искажения на наблюдаемом изображении, обусловленные помехами, потерей резкости и другими причинами, различимы только в зоне ясного видения, т. е. на фрагментах границ, попадающих в зону ясного видения, то и оценка их зрительной системой может производиться только на этих фрагментах.

Рассмотрим подробно, как осуществляется оценка искажений изображения, исходя из этой концепции. Несовершенство оптики глазных сред при проецировании изображения на сетчатки глаз приводит к уменьше-

нию его контраста, а также к ослаблению в нем верхних пространственных частот, которое в первом приближении может быть учтено посредством функции рассеяния точки $h_0(x, y)$. Сетчатки осуществляют цветоделение изображения с образованием трех цветоделенных составляющих $e_k(x, y)$, $e_z(x, y)$, $e_c(x, y)$ (красной, зеленой и синей), из которых затем путем комбинирования, поточечного нелинейного преобразования и подавления средней слагающей формируются две цветоразностные $U_{к-з}(x, y)$ и $U_{ж-с}(x, y)$ (красный минус зеленый и желтый минус синий) и одна ахроматическая $U_a(x, y)$ составляющие изображения:

$$U_{к-з}(x, y) = \log \frac{e_k(x, y)}{e_z(x, y)}; \quad U_{ж-с}(x, y) = \log \frac{e_k(x, y) + e_z(x, y)}{e_c(x, y)};$$

$$U_a(x, y) = \log [e_k(x, y) + e_z(x, y) + e_c(x, y)] -$$

$$- C \log [e_k(x, y) + e_z(x, y) + e_c(x, y)],$$

где C — коэффициент, близкий единице, который учитывает неполное подавление средней слагающей ($C \cong 0,925$), черта над слагаемым — символ усреднения по координатам x, y . Далее в сетчатках эти составляющие подвергаются пространственной фильтрации, результатом которой является ограничение их спектра со стороны верхних пространственных частот, особенно в каналах цветоразностных составляющих изображений, и ослабление нижних пространственных частот за счет латерального торможения в канале ахроматической составляющей изображения. Свойства указанных фильтров могут быть описаны введением соответствующих функций рассеяния точки (профилей активности ганглиозных клеток в терминологии [5]) $h_{к-з}(x, y)$, $h_{ж-с}(x, y)$, $h_a(x, y)$. Получившиеся в результате перечисленных преобразований три нейронных изображения наблюдаемого в данный момент фрагмента контура искаженного изображения

$$B_{к-з}(x, y) = h_{к-з}(x, y) * U_{к-з}(x, y);$$

$$B_{ж-с}(x, y) = h_{ж-с}(x, y) * U_{ж-с}(x, y);$$

$$B_a(x, y) = h_a(x, y) * U_a(x, y),$$

где $*$ — символ свертки, сравниваются с тремя соответствующими нейронными изображениями неискаженного изображения фрагмента (этадона) $B_{з,к-з}(x, y)$, $B_{з,ж-с}(x, y)$, $B_{з,a}(x, y)$, хранящимися в памяти наблюдателя, и вычисляются среднеквадратичные значения их разностей, которые затем комбинируются в результирующую оценку ε_z . Эта оценка инвариантна к яркости наблюдаемого изображения, но зависит от относительного расстояния рассматривания l_0 , представляющего собой отношение расстояния, с которого ведется наблюдение, к высоте изображения.

Значение ε_z в зависимости от решаемой задачи используется следующим образом.

1. При оценке степени ухудшения визуального качества изображения, обусловленного как помехами, так и различного вида искажениями, величина ε_z используется как мера ухудшения качества (мера видности на изображении помех и искажений). При этом пороговое изменение оценки ε_z , соответствующее едва заметному ухудшению качества восприятия изображения $\Delta \varepsilon_z$, как показывает опыт, не зависит от величины ε_z . Вместе с тем величина $\Delta \varepsilon_z$ зависит от средней яркости изображения, поскольку в конечном счете она определяется флюктуационными процессами, имеющими место в зрительной системе.

2. При различении (опознавании) зашумленных изображений объектов вычисляется набор значений ε_{zi} , получающихся от сравнения зашумленного изображения, подлежащего анализу, с каждым из незашумленных изображений объектов $N_{пз}$, составляющих исходный алфавит и хранящихся в памяти наблюдателя, а затем принимается решение путем сравнения каждого из значений ε_{zi} с соответствующими порогами [6].

В частном случае, если априорные вероятности предъявления каждого из изображений объектов одинаковы и равны $1/N_{из}$, а функции потерь C_i выбраны так, что при правильном решении $C_i = 0$, а при ошибочном — $C_i = 1$ — решение принимается в пользу того образца i , для которого ε_{zi} минимальна.

Таким образом, предлагаемая нами обобщенная функциональная модель зрения дает возможным образом ввести критерий качества воспроизведения изображений в виде оценки ε_z .

Найдем для примера значение ε_z применительно к случаю, когда наблюдается ахроматическое изображение, содержащее как помехи, так и искажения другого рода. Запишем выражение для распределения интенсивности возбуждения в нейронном изображении

$$B_a(x, y) = h_a(x, y) * \left\{ \log \left[h_0(x, y) * L(x, y) + \frac{\overline{L(x, y)}}{K_{30}} \right] - \right. \\ \left. - C \log \left[h_0(x, y) * L(x, y) + \frac{\overline{L(x, y)}}{K_{30}} \right] \right\}; \\ B_{к-з}(x, y) = 0; \quad B_{к-с}(x, y) = 0,$$

где K_{30} — величина, посредством которой учитывается уменьшение контраста, обусловленное несовершенством оптики глаз. По нашим оценкам, $K_{30} = 5$. В большинстве практических задач, как показывают расчеты, можно, не допуская заметной ошибки, упростить эту формулу, заменив свертки в квадратных скобках на $L(x, y)$ и аппроксимировав при этом $h_a(x, y)$ выражением

$$h_a(x, y) = \frac{1}{(1 - C_1) r^2 \pi^2} \left(e^{-\frac{x^2 + y^2}{r^2}} - \frac{C_1}{a^2} e^{-\frac{x^2 + y^2}{r^2 a^2}} \right),$$

где C_1 — постоянный коэффициент; $C_1 = 0,6$; r — условный радиус возбуждения рецептивного поля сетчатки, $r = 1,04 \cdot 10^{-2}$ мм; a — отношение радиуса зоны торможения к радиусу зоны возбуждения рецептивного поля сетчатки, $a = 4$ [5]. При этом

$$\varepsilon_z = \sqrt{\int \int_{s_\phi} [B_a(x, y) - B_{э,а}(x, y)]^2 dx dy},$$

где s_ϕ — площадь фовеа.

В том случае, когда оценка ε_z используется в качестве критерия верности воспроизведения изображения в телевизионном вещании, она, подобно критерию, введенному Аллнаттом и Льюисом [8], имеет смысл оценки видности искажений и помех на изображении. Однако в отличие от критерия Аллнатта и Льюиса, измеряющего ухудшение изображения в пикселях, оценка ε_z обладает следующими достоинствами. Во-первых, ε_z аналитически выражается через параметры телевизионной системы. Во-вторых, являясь интегральной оценкой, ε_z учитывает все виды искажений воспроизводимого изображения, благодаря чему не возникает проблемы суммирования их отдельных составляющих, которая до сих пор, как известно, не решена в отношении критерия Аллнатта и Льюиса [8]. Кроме того, ε_z может быть использована для вычисления порогов восприятия различных видов искажений и помех, а также для решения других задач. Проиллюстрируем изложенное, для чего вычислим зависимости порогов восприятия контраста δ_k , четкости δ_s и зашумленности δ_ϕ от определяющих их значений контраста K , числа разрешаемых на изображении линий m и отношения размаха сигнала к среднеквадратичному значению шума ψ_{max} и сравним результаты с экспериментальными данными, взятыми из [9]. Вывод формулы, устанавливающей связь между порогом восприятия и определяющим его параметром системы, проведем на примере вычисления δ_k . Так как ε_z является функцией параметров системы передачи изображений, и в частности K , m , ψ_{max} , т. е.

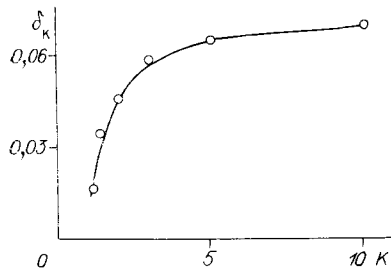


Рис. 5. Зависимость δ_K от K при $l_0 = 4$

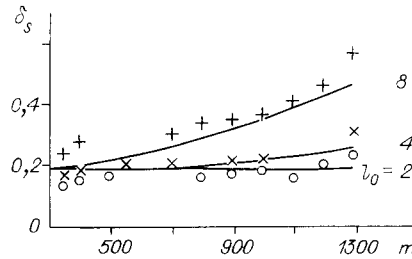


Рис. 6. Зависимость δ_s от m

$\varepsilon_\Sigma = \varphi(K, m, \psi_{\max}, \dots)$, то, обозначая $d\varepsilon_\Sigma/dK = \varphi'_K$, запишем

$$\frac{dK}{K} = \frac{d\varepsilon_\Sigma}{\varphi'_K K}$$

Переходя от дифференциалов к конечным приращениям, будем иметь

$$\delta_K = \frac{\Delta K}{K} = \frac{\Delta\varepsilon_\Sigma}{\varphi'_K K}$$

где $\Delta\varepsilon_\Sigma$ — пороговое изменение оценки ε_Σ , при котором зритель едва замечает изменение ошибки воспроизведения изображения. Значение φ'_K вычисляется исходя из известных характеристик системы передачи изображений и обобщенной функциональной модели зрения. Аналогичным образом находятся формулы для δ_s и δ_ψ , причем по определению [9]

$$\delta_s = 1 - \left(1 - \frac{\Delta m}{m}\right)^2; \quad \delta_\psi = \frac{\Delta\psi_{\max}}{\psi_{\max}}$$

где Δm — пороговое приращение числа разрешаемых линий; $\Delta\psi_{\max}$ — пороговое приращение отношения сигнал/шум. Не приводя выводов формул для расчета функции φ'_K и аналогичных ей по смыслу функций $\varphi'_s = d\varepsilon_\Sigma/dm$ и $\varphi'_\psi = d\varepsilon_\Sigma/d\psi_{\max}$ ввиду их чрезвычайной громоздкости, продемонстрируем результаты расчетов на ЭВМ зависимостей δ_K , δ_s , δ_ψ от соответствующих параметров, которые показаны на рис. 5—7. На этих же рисунках приведены результаты экспериментальных измерений порогов восприятия δ_K , δ_s , δ_ψ , взятые из [9]. Поскольку в этой работе указаны не все параметры экспериментальной установки и тестовых изображений, мы вынуждены были в расчетах использовать наиболее типичные значения недостающих параметров. Так, сквозная световая характеристика экспериментальной установки аппроксимировалась степенной функцией с показателем степени 2,8, а световая характеристика кинескопа — степенной функцией с показателем степени 2. Апертурно-частотная характеристика кинескопа аппроксимировалась гауссовой функцией, имеющей спад на граничной частоте до 0,64, коэффициент усреднения шума, обусловленный инерционностью зрения, принят равным 1,75 [10]. Параметры изображений были взяты следующими: $\Delta L/\bar{L} = 0,93$, $k_{\text{пф}} = 5,6$, где ΔL — среднее значение скачка яркости на контуре; \bar{L} — среднее значение яркости в наблюдаемом фрагменте; $k_{\text{пф}}$ — пик-фактор. Как следует из рисунков, расхождение результатов расчета и эксперимента не выходит за пределы ошибок измерений.

Используя оценку ε_Σ в качестве критерия видности помех и искажений на изображении, можно определить предельные значения параметров системы передачи изображений, при которых вносимые ею искажения зрителю не заметны. Для этого необходимо лишь ε_Σ приравнять $\Delta\varepsilon_\Sigma$ и вычислить значение соответствующего параметра системы. Из этих расчетов, в частности, следует, что при косом шуме (гауссовом шуме с треугольным спектром) и $l_0 = 4$ отношение сигнал/шум, обеспечивающее

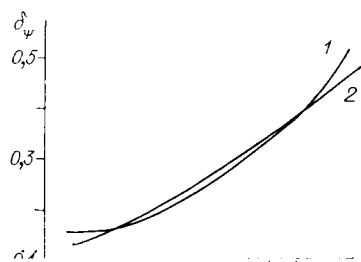


Рис. 7. Зависимость δ_ψ от ψ_{\max} при косом шуме и $l_0 = 4$:
1 — экспериментальная, 2 — теоретическая кривая

незаметность шума на изображении, равно 41,5 дБ, что не расходится с экспериментальными данными, опубликованными в [8].

В заключение отметим, что введение применения обобщенной функциональной модели зрения, является возможность автоматизации выбора оптимальных электрооптических характеристик системы передачи изображения при ее моделировании на ЭВМ. Кратко ее суть заключается в следующем. В память ЭВМ вводится реальное изображение, которое целиком (или его типичный фрагмент) обрабатывается программой, моделирующей преобразования как в системе передачи, так и в зрительной системе. Затем в результате сравнения исходного и обработанного изображений (фрагментов) вычисляется оценка ϵ_z . Эта оценка используется программой в процессе работы для автоматического подбора методом итераций электрооптических характеристик моделируемой системы, обеспечивающих минимум ϵ_z .

Экспериментальная проверка показала высокую эффективность этого метода как с точки зрения существенной экономии времени и ресурсов, затрачиваемых на моделирование систем передачи и обработки изображений, так и с точки зрения качества достигаемых при этом результатов.

СПИСОК ЛИТЕРАТУРЫ

1. Hall C. F., Hall E. L. A nonlinear model for the spatial characteristics of the human visual system // IEEE Trans. Syst., Man, Cybern.—1977.—SMC-7, N 3.—P. 161.
2. Гранрат Д. Дж. Роль моделей зрения человека в обработке изображений // ТИИЭР.—1981.—69, № 5.
3. Физиология сенсорных систем. Ч. 1: Физиология зрения/Под ред. Г. В. Гершуни.—Л.: Наука, 1971.
4. Ярбус А. Л. Роль движений глаз в процессе зрения.—М.: Наука, 1965.
5. Глезер В. Д. Зрение и мышление.—Л.: Наука, 1985.
6. Красильников Н. Н. Теория передачи и восприятия изображений.—М.: Радио и связь, 1986.
7. Прэтт У. Цифровая обработка изображений.—М.: Мир, 1982.—Т. 1.
8. Певзнер Б. М. Качество цветных телевизионных изображений.—М.: Радио и связь, 1988.
9. Антипин М. В. Интегральная оценка качества телевизионного изображения.—Л.: Наука, 1970.
10. Красильников Н. Н. Помехоустойчивость телевизионных устройств.—М.; Л.: Госэнергоиздат, 1961.

Поступила в редакцию 3 мая 1989 г.