

УДК 612.8

Н. Н. Красильников, О. И. Красильникова

(Санкт-Петербург)

**ИССЛЕДОВАНИЕ КОЭФФИЦИЕНТА ЭФФЕКТИВНОСТИ
ЗРИТЕЛЬНОЙ СИСТЕМЫ ЧЕЛОВЕКА
В ПОРОГОВЫХ УСЛОВИЯХ НАБЛЮДЕНИЯ
ДИНАМИЧЕСКИХ ИЗОБРАЖЕНИЙ***

Экспериментально исследованы зависимости коэффициента эффективности при опознавании зашумленных неподвижных и движущихся тестовых объектов от длительности их предъявления. Исследования проведены для различных условий наблюдения. Найдено, что с увеличением длительности наблюдения движущихся изображений тестовых объектов при наличии белого шума коэффициент эффективности вначале убывает. Отмечено, что после достижения минимального значения, которое имеет место приблизительно при 120 мс, наблюдается его возрастание, а затем, после достижения максимального значения, которое имеет место приблизительно при 400 мс, – его монотонное убывание. Выполнено математическое моделирование процессов накопления сигнальной компоненты изображения в зрительной памяти.

Введение. В значительной части информационных систем результаты наблюдения представляются в виде изображений. Эти изображения могут быть как статическими, например, изображения облачного покрова, полученные с метеорологических спутников, изображения, полученные для картографии при аэросъемке, так и динамическими, например, изображения на экранах радаров, представляющие воздушную обстановку в районах аэропорта и используемые для управления воздушным движением. Так как получателем этой информации является человек, то при проектировании таких систем необходимо учитывать свойства его зрения, влияющие на вероятности правильного обнаружения и опознавания наблюдаемых объектов, например авиалайнеров в зоне аэропорта. Другими словами, необходимо иметь математическую модель зрительной системы, посредством которой можно было бы делать предварительные оценки этих вероятностей на стадии проектирования таких систем.

Начиная с 50-х годов 20 века при исследованиях зрительной системы человека стали интенсивно применяться методы теории статистических реше-

* Работа выполнена при поддержке Российского фонда фундаментальных исследований (гранты № 98-06-80001, № 01-06-80297).

ний [1–5], в результате была разработана группа функциональных моделей зрения, известных как «модели согласованной фильтрации» [6–8]. Эти модели позволили получить аналитическим путем ряд известных ранее законов психофизики зрения [9], а также производить оценку вероятностей правильного обнаружения и опознавания объектов на статических зашумленных изображениях (например, фотографиях) при неограниченном времени наблюдения [10]. Эти весьма популярные в настоящее время модели включают в себя линейные и нелинейные звенья, источники внутреннего шума, блок зрительной памяти, а также блок принятия решения в виде оптимального приемника, в котором осуществляется пространственная согласованная фильтрация и принимается решение о принадлежности опознаваемого изображения объекта тому или иному объекту алфавита.

Применительно к случаю наблюдения динамических изображений при ограниченном времени наблюдения такие модели отсутствуют.

Цель данной работы – исследование зрительной системы человека при опознавании неподвижных и движущихся объектов на динамических изображениях и математическое моделирование процессов накопления сигнальной компоненты изображения в зрительной памяти.

Метод исследований. В описываемых исследованиях использован разработанный нами метод сравнительных измерений, который заключается в том, что одни и те же реализации зашумленных пространственным динамическим шумом изображений тестовых объектов предъявляются для опознавания реальным наблюдателям и действующей компьютерной модели идеального наблюдателя, а результаты, полученные в ходе эксперимента, затем сравниваются. Постоянная времени накопления идеального наблюдателя была принята неограниченно большой. Эксперименты выполнялись с динамическими изображениями тестовых объектов, изготовленных в виде цифровых полутоновых черно-белых видеофильмов с частотой смены кадров 25 Гц, при этом длительность предъявления одного кадра составляла 40 мс.

Эти видеофильмы различались между собой: среднеквадратичным значением аддитивного квазибелого гауссова пространственного шума σ , тестовыми объектами, скоростью и направлением движения тестовых объектов, а также местом и моментом времени появления тестовых объектов на изображениях. В описываемых экспериментах было использовано четыре тестовых объекта, показанных на рис. 1. При выполнении экспериментов определялись вероятности правильного опознавания тестовых объектов при наличии пространственного динамического шума реальными наблюдателями P_h и действующей компьютерной моделью идеального наблюдателя P_{id} , на основании которых производилось их сравнение.

В качестве меры сравнения мы использовали так называемый коэффициент эффективности, введенный Г. Барлоу в 1962 г. и широко применяемый в настоящее время в психофизике зрения [11].

Согласно определению коэффициент эффективности k представляет собой отношение пороговой энергии изображения тестового объекта при опознавании (обнаружении) его идеальным наблюдателем E_{id} к пороговой энергии изображения тестового объекта при опознавании (обнаружении) человеком E_h для заданной величины вероятности



Рис. 1

его правильного опознавания (обнаружения) P :

$$k = E_{id} / E_h. \quad (1)$$

Обратим внимание на то, что если в случае реальных наблюдателей каких-либо ограничений относительно наличия или отсутствия априорной информации о параметрах наблюдаемых объектов не делается, то в случае идеального наблюдателя согласно определению все параметры опознаваемых (обнаруживаемых) объектов полагаются априорно известными. Следствием того, что реальный и идеальный наблюдатели в разных опытах могут располагать различной априорной информацией о наблюдаемых объектах, является влияние на величину коэффициента эффективности как собственно совершенства (эффективности) зрительной системы, так и имеющейся у реального наблюдателя априорной информации. Более подробно эта тема изложена в работе [12].

Значение коэффициента эффективности в данной работе определялось, как уже было сказано, исходя из найденных значений вероятностей правильного опознавания тестовых объектов P . На основании этих значений рассчитывались величины отношений сигнала к шуму ψ , которые получаются в результате фильтрации зашумленных изображений зрительной системой и действующей моделью. Расчеты выполнялись по формуле, которая применительно к случаю, когда алфавит состоит из четырех изображений, имеет вид [10]:

$$\psi = \frac{\Phi^{(-1)}(2P - 1) + 0,68}{0,57}, \quad (2)$$

здесь $\Phi^{(-1)}(y)$ – функция, обратная интегралу вероятности

$$\Phi(y) = \frac{2}{\sqrt{2\pi}} \int_0^y e^{-x^2/2} dx.$$

Затем по найденным значениям отношений сигнала к шуму вычислялся коэффициент эффективности

$$k = \left[\frac{\psi_h}{\psi_{id}} \right]^2, \quad (3)$$

где ψ_h и ψ_{id} – отношения сигнала к шуму, вычисленные на основании измеренных вероятностей правильного опознавания тестовых объектов в опытах с наблюдателями и действующей моделью идеального наблюдателя соответственно. Приведенная методика расчета k подробно описана в [12].

В экспериментах приняли участие два тренированных наблюдателя и было использовано по 12 реализаций зашумленных изображений, различающихся между собой реализациями шума. Измерения проводились при различных уровнях динамического шума и различных длительностях предъявления тестового объекта. Размер тестовых объектов составлял $5,9 \times 5,9$ мм, расстояние наблюдения 500 мм. Исследовались как неподвижные, так и движущиеся тестовые объекты.

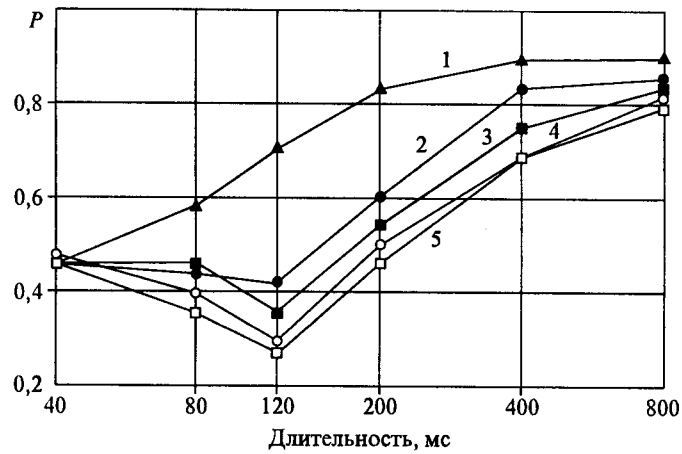


Рис. 2

Зависимость коэффициента эффективности от длительности предъявления неподвижных тестовых объектов в присутствии динамического шума. В этой серии экспериментов все параметры опознаваемых тестовых объектов были априорно известны, зашумленные изображения тестовых объектов предъявлялись наблюдателям и компьютерной модели внутри квадратного окна, размеры которого совпадали с размерами тестовых объектов. Наблюдатель подавал команду компьютеру, после чего на экран дисплея для опознавания выводилось зашумленное динамическим шумом изображение тестового объекта в течение N кадров, длительность предъявления одного кадра составляла 40 мс. На рис. 2 (кривая 1) приведены экспериментальные точки, представляющие полученную в этой серии экспериментов зависимость вероятности правильного опознавания тестовых объектов наблюдателями от длительности их предъявления. На рис. 3 в виде точек 1 приведены

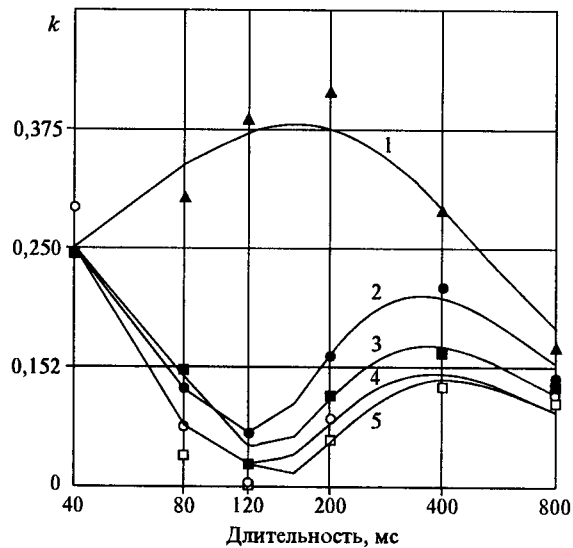


Рис. 3

значения коэффициента эффективности для различных длительностей предъявления тестовых объектов, рассчитанные на основании вероятностей их правильного опознавания, полученных в этой серии экспериментов. Из графиков видно, что найденная зависимость $k = f(t)$ имеет максимум, который может быть объяснен отличием процессов накопления в зрительной системе человека от аналогичных процессов для идеального наблюдателя, а также наличием внутреннего шума зрительной системы. Действительно, при наблюдении изображений неподвижных тестовых объектов на изображениях, зашумленных динамическим шумом, и зрительная система, и идеальный наблюдатель осуществляют усреднение шума вследствие накопления во времени. Однако в отличие от идеального наблюдателя, для которого время накопления принято неограниченным, в зрительной системе оно ограничено, так как в ней одновременно протекают два противоположных процесса: процесс накопления и процесс утечки. Вследствие этого при длительностях предъявления тестовых объектов, превышающих 200 мс, когда начинает проявляться влияние утечки, наступает уменьшение коэффициента эффективности. Малое значение коэффициентов эффективности при длительностях предъявления, меньших 200 мс, можно объяснить влиянием внутреннего шума зрительной системы, что подтверждается приводимыми далее расчетами.

Экспериментальные исследования, аналогичные описанным выше, были выполнены также для случаев, когда наблюдателям априорно были неизвестны координаты тестовых объектов на изображениях, а также время их появления и исчезновения. В результате этих исследований установлено, что отсутствие априорной информации о времени появления и исчезновения тестовых объектов на зашумленных изображениях лишь незначительно уменьшает коэффициент эффективности, в то время как отсутствие априорной информации о координатах тестовых объектов заметно уменьшает его величину [13].

Зависимость коэффициента эффективности от длительности предъявления движущихся тестовых объектов в присутствии динамического шума. Эта серия экспериментов отличалась от предыдущей тем, что опознаваемые реальными наблюдателями тестовые объекты во время наблюдения двигались на экране. Величина угловой скорости V и направление движения объекта были наблюдателю либо априорно известны, либо один или оба эти параметра были априорно неизвестны. В экспериментах, когда скорость движения тестового объекта реальному наблюдателю была априорно неизвестна, сообщалось, что величина этой скорости постоянна в предъявляемой реализации и с равной вероятностью может принимать любое значение в пределах от 0 до V_{\max} . Кроме того, перед проведением опыта наблюдателю предварительно демонстрировалось движение тестового объекта на экране со скоростью V_{\max} . В этой серии экспериментов среднее квадратичное значение динамического шума составляло $\sigma = 62,5$, изображения предъявлялись наблюдателям в центре квадратного окна, размеры которого выбраны таким образом, чтобы в течение всего времени предъявления изображения движущийся тестовый объект оставался в его пределах. Аналогичные результаты были получены и при других уровнях шума.

На основании измеренных вероятностей правильного опознавания тестовых объектов реальными наблюдателями и компьютерной моделью идеального наблюдателя, приведенных на рис. 2, по формулам (2) и (3) были рас-

считаны значения коэффициента эффективности k для различных длительностей предъявления движущихся тестовых объектов при различных условиях наблюдения.

На рис. 2 и 3 показаны зависимости вероятностей правильного опознавания P и коэффициента эффективности k от длительности предъявления движущихся тестовых объектов в виде экспериментальных точек для следующих условий наблюдения: кривая 2 – направление и скорость движения тестовых объектов априорно известны, при этом $V = 0,85$ град/с; 3 – направление движения тестовых объектов априорно известно, скорость случайна, а ее максимальная величина составляет $V_{\max} = 4,33$ град/с; 4 – направление движения тестовых объектов и скорость случайны, а ее максимальная величина составляет $V_{\max} = 4,33$ град/с; 5 – направление движения тестовых объектов и скорость априорно известны, при этом $V = 4,33$ град/с.

Эти зависимости отличаются от зависимостей, полученных в опытах с неподвижными тестовыми объектами (экспериментальные точки 1), тем, что при увеличении длительности наблюдения движущегося тестового объекта вначале имеет место уменьшение вероятности его правильного опознавания реальным наблюдателем, а следовательно, и уменьшение коэффициента эффективности k , которые в наших экспериментах достигали минимума при 120 мс. Затем при дальнейшем увеличении длительности наблюдения начался рост и вероятности правильного опознавания, и коэффициента эффективности, который достигал своего максимального значения при 400 мс. При дальнейшем увеличении длительности наблюдения вероятность правильного опознавания тестового объекта продолжала медленно увеличиваться, а коэффициент эффективности начинал монотонно уменьшаться.

Анализ этих результатов показал, что уменьшение k в начальный период наблюдения обусловлено невозможностью зрительной системы при малых длительностях наблюдения отслеживать смещения проекции движущегося тестового объекта на сетчатках глаз от кадра к кадру. По этой причине накопленный сигнал будет обусловлен рядом наложенных друг на друга взаимно смещенных проекций тестового объекта, которые при больших смещениях могут даже маскировать друг друга. Накопление же динамического шума, присутствующего на изображениях, не зависит от движения тестового объекта и происходит по квадратичному закону вследствие его некоррелированности. В результате при увеличении длительности наблюдения движущегося тестового объекта в начальный период будет иметь место уменьшение отношения сигнала к шуму в зрительной системе и, как следствие, уменьшение коэффициента эффективности. После того, как зрительная система определит направление и скорость движения тестового объекта и начнет его сопровождение, а это, как показали эксперименты, происходит через 120–160 мс после начала демонстрации, ситуация изменяется. С этого времени начинается «синфазное» накопление сигнальной компоненты, приводящее к увеличению отношения сигнала к шуму в зрительной системе и соответственно к увеличению коэффициента эффективности. Так продолжается до тех пор, пока не проявляется влияние утечки в динамической памяти зрительной системы. После этого вероятность правильного опознавания тестового объекта стремится к некоторому пределу, как это видно из рис. 2, а значит, стремится к некоторому пределу и отношение сигнала к шуму в зрительной системе. Из этого следует, что, поскольку в случае идеального наблюдателя согласно принятому определению отношение сигнала к шуму непрерывно возрастает, коэффициент эффективности начинает монотонно убывать.

Анализ этих и аналогичных им экспериментальных данных позволил сделать следующие выводы.

1. Чем выше угловая скорость движения тестового объекта при априорно известном его начальном положении, тем меньше коэффициент эффективности.

2. При прочих равных условиях отсутствие априорной информации о направлении движения тестового объекта приводит к уменьшению коэффициента эффективности.

3. При известном значении максимальной угловой скорости V_{\max} движения тестовых объектов отсутствие априорной информации о ее действительной величине V приводит к увеличению коэффициента эффективности. Этот, казалось бы, весьма странный результат объясняется тем, что в данном случае среднее значение скорости $\langle V \rangle$ в 2 раза меньше ее максимального значения V_{\max} , а следовательно, при отсутствии априорной информации о ее действительной величине наблюдателю предъясняются реализации, где объект движется с меньшей скоростью, чем в случае, когда он располагает этой информацией и скорость движения равна V_{\max} .

4. Наличие априорной неопределенности относительно начального положения движущихся тестовых объектов на изображениях при малом времени их предъяснения приводит к уменьшению коэффициента эффективности, однако, если время наблюдения велико, то движение в ряде случаев приводит к увеличению коэффициента эффективности. Другими словами, в этих случаях сам факт движения позволяет легче определять координаты опознаваемых объектов.

Анализ процесса накопления сигнала в зрительной памяти. Согласно современным представлениям при наблюдении изображений распределения освещенности на сетчатках глаз преобразуются в распределения возбуждений по сетчаткам, которые пропорциональны распределению контрастов K [14]. Благодаря инерционности зрительной системы в ней имеет место накопление этих возбуждений, распределение которых называют нейронным изображением. Затем к нейронному изображению добавляется внутренний шум зрительной системы [8]. Инерционность зрительной системы может быть описана с помощью импульсной функции $h(t)$, которую мы аппроксимируем экспонентой $h(t) = \alpha e^{-\alpha t}$, где $1/\alpha$ – постоянная времени, t – время.

При предъяснении реальному наблюдателю видеопленки, представляющего собой последовательность изображений движущегося объекта, на которые наложен динамический шум, в динамической памяти имеет место накопление сигнальной и шумовой компонент. С учетом изложенного выше выражение для сигнальной компоненты может быть записано в виде

$$s = a_0 K \sum_{n=1}^N F(nT, V) e^{-\alpha T(N-n)}, \quad (4)$$

где T – длительность кадра; TN – время наблюдения; n – номера кадров; a_0 – коэффициент, согласующий размерности; $F(nT, V)$ – функция, посредством которой учитывается «несинфазность» накопления сигнальной компоненты вследствие взаимных смещений проекций на сетчатки глаз движущегося объекта при отсутствии его сопровождения зрительной системой в начале наблюдения. Величина $F(nT, V)$ зависит как от времени, прошедшего от

начала наблюдения движущегося объекта, так и от угловой скорости его движения V .

Выражение же для дисперсии шумовой компоненты после добавления внутреннего шума имеет вид

$$\sigma_{\Sigma}^2 = a_0^2 \sigma^2 \sum_{n=1}^N e^{-2\alpha T(N-n)} + \sigma_{\text{in}}^2,$$

где σ_{in}^2 – дисперсия внутреннего шума зрительной системы. При этом отношение сигнала к шуму будет определено как

$$\Psi_h = s/\sigma_{\Sigma} = \frac{K \sum_{n=1}^N F(nT, V) e^{-\alpha T(N-n)}}{\sqrt{\sigma^2 \left[\sum_{n=1}^N e^{-2\alpha T(N-n)} + (\sigma_{\text{in}}/a_0\sigma)^2 \right]}}. \quad (5)$$

Учитывая, что в случае идеального наблюдателя имеет место идеальное накопление, а внутренний шум отсутствует, запишем соответствующее выражение для отношения сигнала к шуму:

$$\Psi_{\text{id}} = \frac{K}{\sigma} \sqrt{N}. \quad (6)$$

Подставляя Ψ_h и Ψ_{id} из формул (5) и (6) в формулу (3), получим выражение для коэффициента эффективности, «предсказываемое» функциональной моделью

$$k = \left(\frac{\Psi_h}{\Psi_{\text{id}}} \right)^2 = \frac{\left[\sum_{n=1}^N F(nT, V) e^{-\alpha T(N-n)} \right]^2}{N \left[\sum_{n=1}^N e^{-2\alpha T(N-n)} + (\sigma_{\text{in}}/a_0\sigma)^2 \right]}. \quad (7)$$

На рис. 3 приведены зависимости коэффициента эффективности от длительности предъявления неподвижных и движущихся тестовых объектов в присутствии аддитивного динамического квазибелого гауссова шума, рассчитанные по этой формуле. При расчетах значения α , $\sigma_{\text{in}}/a_0\sigma$ и $F(nT, V)$ определялись из условия наименьшего расхождения в смысле среднего квадрата расчетных и экспериментальных данных. Было найдено, что $1/\alpha = 150$ мс, $\sigma_{\text{in}}/a_0\sigma = 1,733$. Значения функции $F(nT, V)$ приведены на рис. 4 для различных условий наблюдения, при этом точками обозначены ее найденные значения. Эти значения соответствуют следующим условиям наблюдения: кривая 1 – тестовые объекты неподвижны; 2 – направление и скорость движения тестовых объектов априорно известны, при этом $V = 0,85$ град/с; 3 – направление движения тестовых объектов априорно известно, скорость случайна, а ее максимальная величина составляет $V_{\text{max}} = 4,33$ град/с; 4 – направление движения тестовых объектов и скорость случайны, а ее максимальная величина составляет $V_{\text{max}} = 4,33$ град/с; 5 – направление движения тестовых объектов и скорость априорно известны, при этом $V = 4,33$ град/с.

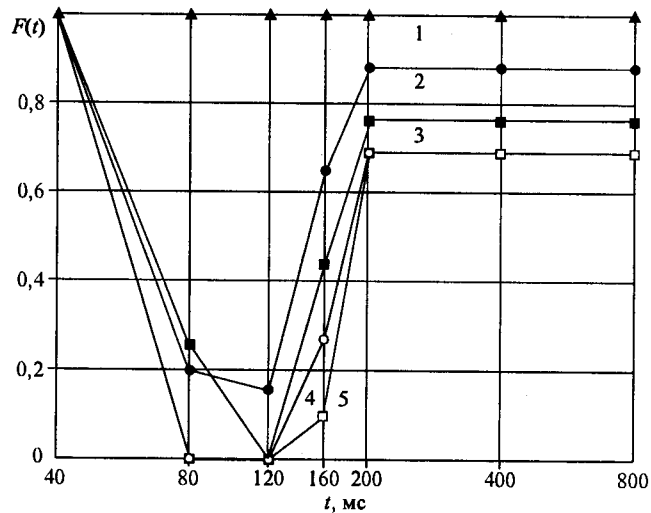


Рис. 4

Сравнительно небольшие расхождения между данными, полученными на основании модели, и данными, полученными в результате эксперимента, подтверждают адекватность математической модели (см. рис. 3).

К интересным выводам приводит анализ рис. 4. Если при опознавании неподвижных объектов все значения функции $F(t, V) = 1$, то в случае опознавания движущихся объектов $F(120, V) < F(80, V) < F(40, V)$. Как уже отмечалось, это обусловлено смещением положения проекции тестового объекта на сетчатках от кадра к кадру; накапливаемые в динамической памяти изображения будут накладываться друг на друга со сдвигом, что влечет за собой уменьшение накопленной компоненты сигнала.

Заключение. Полученные результаты в значительной степени проливают свет на алгоритм работы зрительной системы наблюдателя при опознавании неподвижных и движущихся объектов на динамических изображениях, давая ключ к решению ряда практических задач, связанных с проектированием информационных систем.

СПИСОК ЛИТЕРАТУРЫ

1. Svets I. A., Tanner W. P., Birdsall T. G. Decision processes in perception // Phys. Rev. 1961. 68, N 5. P. 197.
2. Байдаков М. П., Красильников Н. Н., Пастухов О. В., Потоцкий В. К. К вопросу о различении наблюдателем зашумленных изображений // Автометрия. 1970. № 3. С. 20.
3. Байдаков М. П., Красильников Н. Н., Пастухов О. В. О различении человеком изображений в гауссовом шуме // Автометрия. 1973. № 1. С. 7.
4. Barnard T. W. An Image Evaluation Method. Norwalk, Connecticut: Perkin-Elmer Corporation, 1971.
5. Красильников Н. Н. Влияние шумов на контрастную чувствительность и разрешающую способность приемной телевизионной трубки // Техника телевидения. 1958. Вып. 25. С. 26.

6. **Burgess A.** Visual signal detection. On Bayesian use of prior knowledge and cross correlation // JOSA. A. 1985. **2**, N 9. P. 1498.
7. **Rovamo J., Mustonen J., Nasanen R.** Modeling contrast sensitivity as a function of retinal illuminance and grating area // Vision Res. 1994. **34**. P. 1301.
8. **Красильников Н. Н., Шелепин Ю. Е., Красильникова О. И.** Фильтрация в зрительной системе человека в условиях порогового наблюдения // Оптический журнал. 1999. **66**, № 1. С. 5.
9. **Красильников Н. Н.** Функциональная модель зрения // Оптический журнал. 1997. **64**, № 2. С. 72.
10. **Красильников Н. Н.** Теория передачи и восприятия изображений. М.: Радио и связь, 1986.
11. **Barlow H.** Measurements of the quantum efficiency of discrimination in human scotopic vision // Journ. Physiol. 1962. **160**. P. 169.
12. **Shelepin Y. E., Krasilnikov N. N., Krasilnikova O. I.** What visual perception model is optimal in terms of signal-to-noise ratio? // Proc. of SPIE. 2000. **3981**. P. 116.
13. **Krasilnikova O. I., Krasilnikov N. N., Shelepin Y. E.** Measurements of the efficiency of the human visual system in observing dynamic images // Twenty-fourth European Conf. on Visual Perception. Kusadasi, Turkey. 2001. Vol. 30 (suppl.). P. 82.
14. **Бондарко В. М., Данилова М. В., Красильников Н. Н. и др.** Пространственное зрение. С.-Пб.: Наука. 1999.

*Санкт-Петербургский государственный университет
аэрокосмического приборостроения,
E-mail: nnk_k23@aanet.ru*

*Поступила в редакцию
24 января 2003 г.*