

РОССИЙСКАЯ АКАДЕМИЯ НАУК

СИБИРСКОЕ ОТДЕЛЕНИЕ

А В Т О М Е Т Р И Я

№ 2

1995

УДК 621.397

Л. Н. Аксютов

(Санкт-Петербург)

**МЕТОДИКА ОЦЕНКИ КАЧЕСТВА ИЗОБРАЖЕНИЙ В ОПТИЧЕСКИХ
ИНФОРМАЦИОННЫХ СИСТЕМАХ**

Разработана методика, позволяющая оценить общее и информативное качество изображения независимо от физического принципа его формирования. Она базируется на основном психофизическом законе (законе Фехнера) и других закономерностях зрительного восприятия при решении зрительных задач. Параметрами методики являются разрешающая способность изображения, среднеквадратическое отклонение шума и масштаб изображения, угловое разрешение системы формирования изображения, площадь видимой проекции наблюдаемых объектов и расстояние наблюдения. Предложена мера информативности изображений, связанная с гексагональной мозаикой колбочек сетчатки и уровнями информативности, обеспечивающими категориальное распознавание объектов при анализе макро-, мезо- и микроструктуры изображений.

Значение модели зрительной системы (ЗС) в решении проблемы создания эффективных оптических информационных систем, обеспечивающих сбор, преобразование, обработку и анализ видеоданных (изображений), общепризнано [1, 2]. Одним из аспектов этой проблемы является оценка качества изображений различного уровня информативности.

Когда цель наблюдения конкретно не задана, мерой восприятия является общее качество (Q) изображения, основанное на его эстетической оценке. При решении перцептивно-опознавательных задач, постановка которых всегда конкретна, меру восприятия изображения определяет его информативное качество, оцениваемое вероятностью распознавания (P_p) фигуры изображения в контексте анализа сцены.

В настоящее время для оценки качества изображений необходимо знание технических параметров и характеристик системы формирования изображений (СФИ) [3] или самого изображения, основными параметрами которого, влияющими на его качество, являются среднеквадратический шум и разрешающая способность [4]. При оценке качества используются простые зрительные стимулы, а корреляцию результатов оценки с визуальным качеством изображений устанавливают эксперты методом. Полученные таким образом результаты отражают лишь закономерности восприятия [5] и не связаны с информативностью изображений. Цель настоящей статьи состоит в том, чтобы проиллюстрировать разработанную автором методику оценки качества изображений на базе основного психофизического закона (закона Фехнера) и других закономерностей зрительного восприятия.

Согласно современным представлениям, визуальное восприятие происходит в два этапа [6]. На первом, сенсорном этапе ЗС различает простые зрительные стимулы и посредством генетически врожденных механизмов выделяет пространственные неоднородности изображения, которые формируют текстуры и фигуры, регистрируемые в кратковременной (оперативной) памяти (КВП). Эта способность ЗС регламентируется закономерностями восприятия

яркостных пространственных структур светового поля. На втором, перцептивно-опознавательном этапе в соответствии с задачей наблюдения, уровнем информативности изображения и хранящейся в долговременной памяти (ДВП) когнитивной картой образов ЗС различает фигуры изображения и выполняет их описание (представление) с помощью признаков, выработанных в процессе обучения. При этом осуществляется «умственное сопоставление» ощущений (впечатлений), соответствующих релевантному зрительной задаче содержанию КВП и ДВП, и принятие решения об их тождестве. В результате происходит «распознавание» представленного изображения объекта.

Чередование фаз принятия решения на втором этапе восприятия соответствует дедуктивному характеру различия зрительного образа при «послойном» анализе изображения [7], в процессе которого представленный объект последовательно относится к одной из $g = 1, 4$ групп категорий — виду, классу, подклассу и типу [8], образующих иерархическую древовидную структуру компактного описания объектов, детальность которого возрастает от вида ($g = 1$) к типу ($g = 4$).

Особая роль суммации возбуждений функциональных структур ЗС в процессе «различение — распознавание» [9] означает, что визуальное качество изображения должно быть эквивалентно суммарному ощущению (σ_s) ЗС, обусловленному световой стимуляцией (σ_{cb}) и стимульной информацией (σ_{np}) о пространственной структуре светового поля. Тогда, согласно закону Фехнера [10], запишем

$$\sigma_s = \sigma_{np} + \sigma_{cb} = b \lg \psi_{np} + d \lg \psi_{cb}, \quad (1)$$

где $\psi = S/S_0$ — фундаментальная величина стимула, равная отношению действующей интенсивности стимула S к ее пороговому S_0 значению; b, d — параметры, характеризующие отклик ЗС на стимул.

Пороговая интенсивность стимула определяется как «абсолютная» (t_a), если она соответствует границе «появление — исчезновение» вызванного стимулом ощущения, или как «операционная» (t_o), когда она обеспечивает уровень ощущения, достаточный для 50 %-ной вероятности различия стимула. Данному уровню соответствует баланс влияющих на восприятие стимула благоприятных и неблагоприятных факторов или максимум распределения мгновенных флуктуаций порога восприятия, которое подчиняется нормальному закону [10]. Это объясняет характерную для психометрической функции визуального различия форму интегральной кривой нормального распределения, центром симметрии которой и является операционный порог различия.

Суммарное ощущение ЗС при анализе изображения должно быть эквивалентно объему информации, необходимому для решения зрительной задачи. В теории поиска [11] этим эквивалентом является поисковый потенциал Φ наблюдателя, обуславливающий возможность снятия пространственной неопределенности объекта визуального поиска (зрительного стимула) с вероятностью $P = 1 - \exp(-\Phi)$. Справедливость тождества $\Phi \equiv \sigma_s$ подтверждается формулой Вейбулла [12], описывающей психометрическую функцию визуального различия контраста:

$$P = 1 - \exp(-K/\alpha)^\beta,$$

где α, β — постоянные, определяющие величину суммарного ощущения ЗС при величине K интенсивности стимула. Для общего случая различия зрительных стимулов можно записать:

$$P_p = 1 - \exp(-\sigma_s), \quad (2)$$

где величина σ_s выражается уравнением (1).

Численное значение ψ_{cb} в уравнении (1) равно отношению яркости L квазиравномерного светового поля изображения к яркости $L_{ta} = 3,16 \times$

$\times 10^{-3}$ кд/м², соответствующей абсолютному порогу остроты различения $\theta_L = 1/\delta_L$, где δ_L — предельный угол пространственного разрешения ЗС при яркости L светового поля. Величина L_{ta} согласуется с яркостью $L = 3 \times 10^{-3}$ кд/м², при которой происходит полный переход от фoveального к периферическому зрению [13]. Увеличение (до определенного предела) яркости L или величины $\lg\psi_{cb}$ вызывает возрастание остроты зрения. При этом, согласно закону константности восприятия света, для всех уровней психометрической функции должно выполняться условие $\sigma_{cb} = \text{const}$, когда в отсутствие светового шума взаимно компенсируется изменение величин $\lg\psi_{cb}$ и θ_L , что выполняется при равенстве $d = \delta_L$ в уравнении (1).

При интенсивности зорительных стимулов, отображающих информацию о пространственной структуре светового поля, равной операциональному порогу ($\psi_{np} = 1$), выполняется равенство $\sigma_s = \sigma_{cb}$, на основании которого и в соответствии с уравнениями (1) и (2) получим

$$\sigma_{cb} = \text{const} = \delta_L \lg(L/L_{ta}) = \ln 2, \quad (3)$$

где величина δ_L выражается в миллирадианах.

Значения δ_L , рассчитанные с помощью соотношения (3) для яркостей $1 < L < 500$ кд/м², в пределах 10 % согласуются с их оценкой по полуэмпирической формуле Луизова [14], определяющей величину δ_L в угловых минутах для максимального контраста изображения.

Составляющая суммарного ощущения σ_{np} в уравнении (1) не обладает константностью и определяется величинами ψ_{np} и b , характеризующими информативность изображения и условия различия наблюдателем структуры светового поля соответственно.

Информативность сетчатого изображения обусловлена зоной рецептивных полей (РП) ЗС, перекрываемых стимулом: чем больше эта зона, тем больше пространственных свойств стимула выявляет суммация его локальных воздействий на нейронные уровни ЗС. Пространственная организация РП вблизи фовеа сетчатки повторяет гексагональную мозаику ее колбочек [15], следовательно, число РП, перекрывающих видимую проекцию целевого объекта (ВПО), равно числу колбочек, перекрываемых фигурай изображения ВПО на фовеа. Поэтому в качестве меры информативности изображения выбрано число E_s элементов разрешения (пикселов) СФИ (в рассматриваемом случае — ЗС), приходящихся на площадь ВПО или ее фигуры изображения. Тогда фундаментальной величине стимула ψ_{np} в уравнении (1) эквивалентен линейный инвариант информативности ψ , определяемый для данных условий наблюдения как отношение E_s к операциональному порогу $E_{s,to}$:

$$\psi = (E_s/E_{s,to})^{1/2} = \alpha/\alpha_{to} = a_{to}/a = A_{to}/A = D_{to}/D, \quad (4)$$

где α — число пикселов СФИ, согласованной по пространственному разрешению с ЗС, которые приходятся на воспринимаемый ЗС линейный размер стимула для заданного уровня психометрической функции $P_p = f(\psi)$ его различия; $A = aD/F = \delta \times D$, a — линейный размер пикселя СФИ или колбочки фовеа ЗС; δ , F — угловое разрешение и фокусное расстояние СФИ (ЗС) соответственно; A — линейное разрешение СФИ (ЗС) в плоскости объекта при расстоянии D до него; уровень величины (t_o) соответствует операциональному порогу различия.

Фокусное расстояние ЗС, аккомодированной для наблюдения стимула на расстоянии D , м, в настоящей методике определяется по формуле

$$F_{ZC} = \exp \left[10^{1/2} - 0,008/(D - 0,1) \right] \text{ мм},$$

аппроксимирующей с погрешностью < 1 % табулированные в [14] значения F_{ZC} для расстояний наблюдения от 0,25 м до бесконечности. Легко вычислить, что при среднем размере колбочки фовеа $a = 3,16$ мкм угловое разрешение ЗС

при наблюдении контрастного изображения с расстояния, например, $D = 1,5$ м равно 0,134 мрад.

Зрительные иллюзии убеждают, что визуальное восприятие стимулов обусловлено их окружением. Окружением обусловлена также альтернативность (v) различия фигур изображения при распознавании наблюдаемых объектов, которая совместно с локальной неоднородностью фона определяет фоновую помеху решению зрительной задачи, уменьшающую вследствие возникающей при этом неопределенности уверенность различения стимулов и, следовательно, вероятность распознавания объектов. Аналогичное воздействие на характеристики различия фигуры изображения оказывает световой шум изображения.

С учетом тождеств (3) и (4) уравнение (1) при замене $b = \eta_v$, можно записать в виде

$$\sigma_s = \eta_v \ln \psi + q \ln 2, \quad (5)$$

где q — показатель влияния светового шума на восприятие изображения (для незашумленного изображения $q = 1$); η_v — показатель уверенности различия, соответствующий правильному решению зрительной задачи в условиях фоновой помехи, которая эквивалентна вероятности ложных тревог $P_{\text{ЛТ}}$: $\eta_v = |\ln P_{\text{ЛТ}}|$.

Подставив выражение (5) в уравнение (2), получим формулу для психометрической функции визуального различия незашумленного изображения, когда интенсивность зрительных стимулов превышает операциональный порог ($\psi = 1$, для которого $P_p = 0,5$):

$$P_p = 1 - 0,5\psi^{-\eta_v}, \quad \psi > 1. \quad (6)$$

В силу центральной симметрии психометрической функции относительно операционального порога для нижепороговой интенсивности стимулов

$$P_p = 0,5\psi^{\eta_v}, \quad \psi < 1. \quad (7)$$

Для реализации выбора из исходного множества X некоторого подмножества, разнообразие которого определяется величиной W , необходимо количество информации, равное в соответствии с мерой Хартли [16] энтропии $H(X)$ исходного множества:

$$H(X) = C \ln W, \quad (8)$$

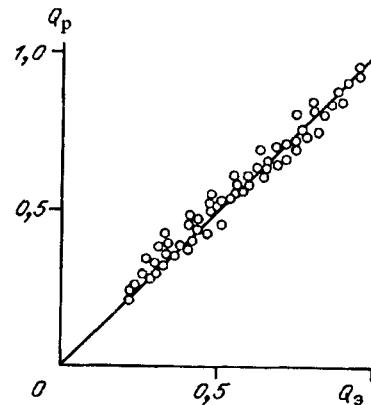
где $C > 0$ — константа, не зависящая от W и равная производной от функции $f(W)$ в точке $W = 1$.

Значение η_v в уравнениях (5) — (7) численно равно удвоенной производной психометрической функции $P_p(\psi)$ в пороговой точке $\psi = 1$, что свидетельствует об аналогии уравнения (8) и первого слагаемого в уравнении (5). Второе слагаемое в этом уравнении при $q = 1$ выражает оцениваемое той же мерой количество информации, необходимое для выбора из двух возможностей: вызывают ощущение или нет яркостные стимулы светового поля. Таким образом, суммарное ощущение ЗС при восприятии изображений имеет смысл количества информации, необходимого для решения зрительной задачи.

Уравнения (4) — (7) позволяют осуществить прогнозирование вероятности «различения — распознавания» (метод ПВР) объектов при визуальном анализе сцены, на основе которого разработана предлагаемая методика оценки качества изображений. Параметры в методике определены по результатам анализа представительных данных экспериментов по различению объектов и распознаванию их категориальной принадлежности при визуальном наблюдении натурных и искусственных сцен или дешифрировании (ДШ) их фотографических и оптико-электронных изображений.

Результаты и их обсуждение. В работе [4] методом экспертных оценок определялось общее качество (Q_s) фото- и телевизионных сюжетных изобра-

Рис. 1. Сравнение расчетной Q_p и экспертной Q_s оценок общего качества изображений:
объем выборки $n = 63$ [4], коэффициент выборочной корреляции
 $\hat{r} = 0,985$, выборочная дисперсия $\hat{\sigma}^2 = 0,0015$



жений в функции их разрешающей способности $R = 6,5—77 \text{ мм}^{-1}$ и среднеквадратического шума $\sigma_N = 0,036—0,086$, измеряемых инструментально.

На основании принципа «обмена» между параметрами, влияющими на визуальное восприятие, световой шум при разрешающей способности R изображения можно заменить ее эквивалентным операциональным порогом $R_{io} = k\sigma_N \text{ мм}^{-1}$. При этом допущении анализ данных эксперимента из [4] показал, что воспринимаемое экспертами общее качество зашумленных изображений соответствует выполнению равенства

$$Q_p = 0,5\sigma_s, \quad (9)$$

в котором σ_s определяется уравнением (5) при следующих значениях входящих в него параметров: $\psi = R/R_{io} = R/k\sigma_N$ при $k = 1 \text{ мм}^{-1}$, $\eta_v = 0,59$, $q = -3,15$ для значений $R < 11 \text{ мм}^{-1}$ и $\eta_v = 0,33$, $q = -1,0$ для $R > 11 \text{ мм}^{-1}$. Корреляция значений Q_p и Q_s показана на рис. 1. Приведенные значения параметров η_v и q отражают известные качественные особенности визуального восприятия зашумленных изображений, а значение $R = 11 \text{ мм}^{-1}$ согласуется с эмпирически установленным разрешением $R = 10 \text{ мм}^{-1}$, разделяющим изображения, пригодные для категориального ($R > 10 \text{ мм}^{-1}$) или только для топологического ($R < 10 \text{ мм}^{-1}$) ДШ [7].

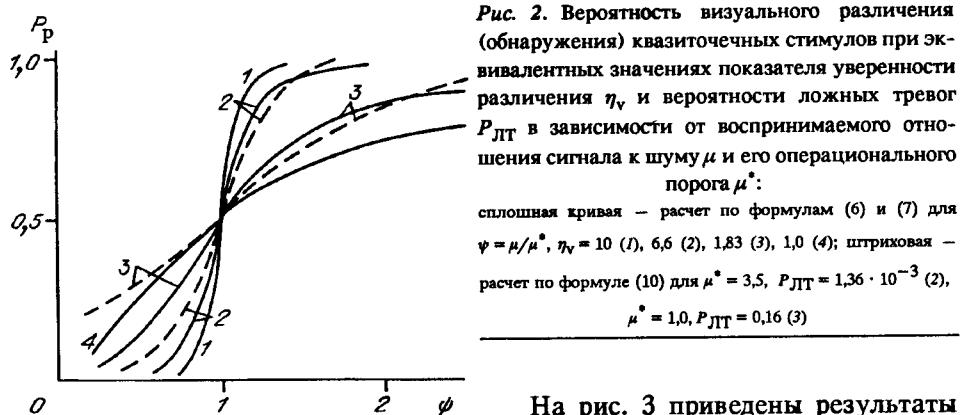
Вероятность визуального различия точечных и квазиточечных стимулов определяется воспринимаемым ЗС отношением сигнал/шум (μ) и его операциональным порогом (μ_{io}) и может быть вычислена для заданной вероятности ложных тревог с помощью известной из статистической теории обнаружения сигналов формулы:

$$P_p = (2\pi)^{-1/2} \int_{-\infty}^{\frac{\mu - \mu_{io}}{\sigma}} \exp(-z^2/2) dz, \quad (10)$$

где z — аргумент интеграла вероятности.

В рамках предлагаемой методики эту вероятность можно вычислить в предположении справедливости равенства $\psi = \mu/\mu_{io}$, что дает возможность сравнивать результаты расчетов по формулам (10) и (6) для эквивалентных значений η_v и P_{lt} . Как видно из рис. 2, эти результаты для $\psi > 1$ практически согласуются друг с другом.

При визуальном различении точечных стимулов величина α_{io} из тождества (4) соответствует информативности изображения, когда число перекрывающихся колбочек может быть не более двух, что соответствует линейному «критерию Джонсона» [17], равному $1,4 \pm 0,35$. Следовательно, начальный «нулевой» уровень информативности ($n = 0$) изображения, при котором ЗС реализует первую стадию восприятия, характеризуется для различия точечных стимулов операциональным порогом $E_{s,n} = 2$. Для решения более сложных зрительных задач второй стадии восприятия уровень информативности и операциональный порог различия должны быть выше.



На рис. 3 приведены результаты выполненного с помощью метода ПВР анализа экспериментальных данных по распознаванию объектов на фото- и оптико-электронных изображениях. Пересечение оси α_{lo} прямыми, вдоль которых кластеризуются точки, представляющие данные экспериментов, определяет значения α_n , связанные с α_{lo} соотношением

$$\alpha_{lo} = \alpha_n \times \eta_v^{-\kappa}.$$

Полученные результаты подтверждают установленный в [7] факт существования «серийного» (стандартного) и «эвристического» (исследовательского) режимов ДШ изображений, которым в приведенном соотношении отвечают соотношения $\kappa = \kappa_c = 1,25$ и $\kappa = \kappa_u = 0,25$ соответственно. Кроме того, они позволили установить, что различие по фигуре изображения принадлежности объекта к одной из $g = 1, 4$ групп категорий обеспечивается при соответствующем уровне $n = 1, 4$ информативности изображения, который характеризуется максимальным операциональным порогом числа $E_{s,n} = 2 \times 10^{n/2}$ колбочек фовеа, перекрываемых изображением ВПО (фигуры изображения объекта), причем $\alpha_n = (E_{s,n})^{1/2}$.

Анализ имеющихся экспериментальных и теоретических данных показывает, что при решении перцептивно-опознавательных задач наблюдатель последовательно анализирует макро-, мезо- и микроструктуру изображения, стараясь реализовать достижимый уровень его информативности, позволяющий различить категориальную принадлежность: объектов в сцене ($n = n_o = 1, 4$), деталей объекта ($n = n_{do} = 5, 8$) и элементов детали ($n = n_{ed} = 9, 12$). Возможность распознавания целевого объекта, включенного в одну из этих структур изображения, определяется реализованным уровнем его информативности, который связан с индексом группы категорий $g = 1, 4$ равенством

$$g = n = n_o = n_{do} - 4 = n_{ed} - 8.$$

Так, при $n = n_o = 2$ изображение объекта воспринимается квазито-

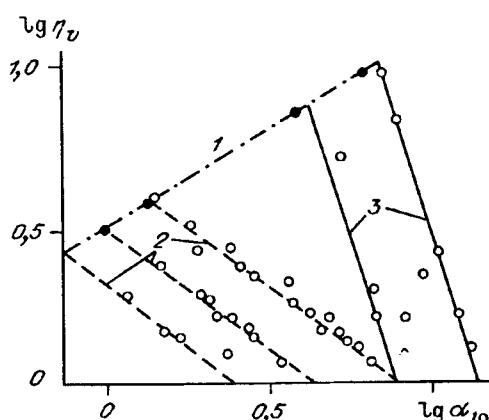


Рис. 3. Результаты анализа экспериментальных данных по распознаванию объектов сцены (различию макроструктуры изображения) при стандартном (2) и исследовательском (3) дешифрировании изображений:
 1 – безальтернативное различие фигуры изображения; • – критерий Джонсона [17]

Рис. 4. Корреляция расчетных P_p и экспериментальных P_e данных категориального распознавания объектов сцены по изображениям, полученным оптико-электронной системой наблюдения:

$$w = 29, \hat{r} = 0,968, \sigma^2 = 0,0069$$

чечным пятном («блобом»), по которому можно определить лишь ориентацию объекта, а при $n = n_{\text{до}} = 8$ на изображении ВПО распознается тип деталей объекта.

В общем случае распознавание объектов по изображению связано с фоновой помехой, основными компонентами которой являются «похожие» фигуры и локальные неоднородности фона. Альтернативность распознавания, эквивалентная максимальной фоновой помехе, вследствие ограниченности объема оперативной памяти ЗС [18] в методе ПВР определяется выбором из трех–пяти возможностей при распознавании вида и типа целевых объектов соответственно. Это при решении зрительных задач обуславливает размер вероятности ложных тревог $P_{\text{ЛТ}} = 1/e \cong 0,37$, для которого $\eta_v = 1$. В случае безальтернативного различия, для которого разработан критерий Джонсона, максимальным является размер η_v , определяемый для стандартного (с) и исследовательского (и) ДШ с помощью эмпирических выражений:

$$\eta_{v, c} = M_0^{(g+4)/3} \quad \text{и} \quad \eta_{v, i} = M_0^{(g+4)/2},$$

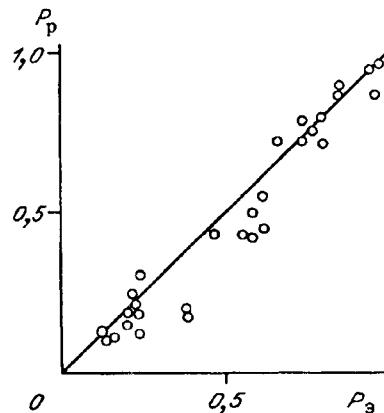
где $g = \sqrt[4]{4}$, $M_0 = 10^{1/4}$ — постоянная, имеющая смысл коэффициента изменения линейного масштаба сетчатого изображения при переходе от одного уровня его информативности к последующему.

Дискретность операционного порога информативности связана, по-видимому, с гексагональной мозаикой колбочек сетчатки, поскольку с высокой точностью выполняется равенство $M_0 = 0,866^{-4}$, где 0,866 определяет отношение высоты к стороне равностороннего треугольника, являющегося элементом гексагона. Косвенным подтверждением этого предположения можно считать факт, связанный с введением при переходе к системам ТВ высокой точности формата кадра $16 : 9 = M_0$ вместо $4 : 3$ [19]. Предпочтительность нового формата была установлена экспериментально по критерию его лучшего визуального восприятия.

В качестве подтверждения достоверности предлагаемой методики на рис. 4 приведен пример корреляции прогноза вероятности распознавания (P_p), выполненного с помощью метода ПВР, и результатов независимого представительного дешифрирования оптико-электронных изображений (P_e), полученных группой подготовленных наблюдателей. Параметры корреляции в подрисуночной подписи типичны для метода ПВР.

Подход, основанный на оценке уровней информативности цифровых изображений по методу ПВР, позволил разработать эффективный алгоритм вычисления по изображению площади ВПО сложной конфигурации [20].

Разработана и апробирована методика оценки общего и информативного качества изображений, интегрирующая основные закономерности зрительно-го восприятия и распознавания объектов. Методика оценки реализуется по-средством измерения общезвестных параметров изображения или системы его формирования.



СПИСОК ЛИТЕРАТУРЫ

1. Гранрат Д. Дж. Роль моделей зрения человека в обработке изображений // ТИИЭР. 1981. 69, № 5.
2. Берт П. Дж. Интеллектуальное восприятие в пирамидальной зрительной машине // ТИИЭР. 1988. 76, № 8.
3. Barton P. G. I. Evaluation of subjective image quality with square-root integral method // JOSA. 1990. 7A, N 10. P. 2024.
4. Антипов М. В., Гласман К. Ф., Кириллова И. Н. и др. Оценка качества фотографических изображений // Техника кино и телевидения. 1984. № 11.
5. Красильников Н. Н. Новое в развитии обобщенной функциональной модели зрения для информационных систем // Автометрия. 1992. № 2.
6. Глезер В. Д. Зрение и мышление. Л.: Наука, 1985.
7. Рубахин В. Ф. Психологические основы обработки первичной информации. Л.: Наука, 1974.
8. Завалишина Д. Н., Ломов Б. Ф., Рубахин В. Ф. Уровни и этапы принятия решения // Проблемы принятия решения. М.: Наука, 1976.
9. Власова М. М. Механизмы опознавания неизвестного // Проблемы космической биологии. М.: Наука, 1989. Т. 62.
10. Бардин К. В., Забродин Ю. М. Проблемы психологического отражения свойств объективного мира на сенсорно-перцептивном уровне // Психологические исследования восприятия и памяти. М.: Наука, 1981.
11. Абчук В. А., Суздаль В. Г. Поиск объектов. М.: Сов. радио, 1977.
12. Vos J. J., van Meeteren A. PHIND: an analytic model to predict target acquisition distance with image intensifiers // Appl. Opt. 1991. 30, N 8. P. 958.
13. Pelli D. G. Uncertainty explains many aspects of visual contrast detection and discrimination // JOSA. 1985. 2A, N 9. P. 1508.
14. Луизов А. В. Глаз и свет. Л.: Энергоатомиздат, 1983.
15. Хьюбел Д. Глаз, мозг, зрение. М.: Мир, 1990.
16. Шилейко А. В., Кочнев В. Ф., Химушин Ф. Ф. Введение в информационную теорию систем. М.: Радио и связь, 1985.
17. Уэзерелл У. Оценка качества изображений // Проектирование оптических систем. М.: Мир, 1983.
18. Шехтер М. С. Зрительное опознание. Закономерности и механизмы. М.: Педагогика, 1981.
19. «Монтрэ-91». Секция «ТВ-вещание» // Техника кино и телевидения. 1991. № 11.
20. Аксютов Л. Н., Костицина Т. Г. Определение площади объектов по цифровым космическим снимкам // Исследование Земли из космоса. 1994. № 6.

Поступила в редакцию 26 апреля 1993 г.