

---

## Психофизиологически обоснованный метод оценки количества информации в изображении

*Д.В. Явна, В.В. Бабенко*

*Южный федеральный университет, Ростов-на-Дону*

**Аннотация:** В работе предлагается способ количественного оценивания информации в изображении, базирующийся на подсчёте числа моделируемых простых нейронов проекционной коры, которые будут реагировать на содержащиеся в изображении элементарные зрительные признаки – светлые полосы в тёмном окружении и яркостные границы. Предполагается, что чем больше таких признаков содержится в изображении, тем больше информации поступит в мозг. Способ основывается на модельных представлениях о механизмах локального анализа в зрительной системе человека.

**Ключевые слова:** сенсорный признак, растровое изображение, информационная энтропия, информационная избыточность, стриарный нейрон, функция Габора, компьютерное моделирование.

### Введение

В этой работе мы предлагаем способ оценки количества информации в изображении, которое поступает в проекционные зоны коры головного мозга наблюдателя. Для простоты изложения речь ведётся о растре, представленном градациями серого, однако предлагаемый подход может быть расширен и на цветное изображение. Важно отметить, что количество реагирующих нейронов, определяемое в ходе расчёта, является количеством оп-нейронов в модели, а не в живом мозге, однако модель может быть сделана более реалистичной, так что в принципе имеется возможность примерно подсчитать и число реагирующих простых нейронов проекционной коры человека. С другой стороны, это может сильно усложнить расчёты и не дать существенного улучшения оценки с точки зрения практического применения.

Целью данной работы является создание психофизиологически обоснованной оценки количества *зрительной* информации в растре как изображении, наблюдаемом с определённого расстояния, а не информации вообще в сообщении, это изображение кодирующем.

Когда оценивают количество информации в растровом изображении (или связанную с этой величиной его *информационную избыточность*), обычно используют подход, основанный на вычислении информационной энтропии, восходящий к К.Шеннону. Например, в простейшем случае информационная энтропия изображения  $H$  может быть оценена как *средняя энтропия сообщения*

$$H = - \sum_{k=0}^K P(k) \log_2(P(k)) \quad (1)$$

где  $K$  – количество уровней яркости,  $P(k)$  – вероятность появления  $k$ -го уровня яркости, а информационная избыточность  $r$  – как

$$r = b - H \quad (2)$$

где  $b$  – наименьшее число бит, которыми может быть представлен дискретизированный сигнал [1]. Если вероятность появления  $k$ -го уровня яркости неизвестна, она может быть оценена по частоте появления пикселя данной яркости в конкретном изображении. Такой подход является довольно распространённым, в частности, оценка энтропии изображения по гистограмме частот реализована в виде функции библиотеки расширений для обработки изображений в среде Matlab (Image processing toolbox) [2].

Такая избыточность называется *статистической* [3]. Очевидно, что оценки (1) и (2) связаны с содержанием изображения очень слабо – они отражают в первую очередь распределение яркости и количество уровней её описания. Например, если взять растровое изображение объекта и перемешать пиксели случайным образом, значения средней информационной энтропии  $H$  и избыточности  $r$  не изменятся, а доступная субъекту информация об изображённом объекте с большой вероятностью исчезнет. Учитывая то, что в реальных изображениях значения смежных пикселей коррелированы [3], информационная энтропия изображения  $H_c$  из  $m$  точек

может быть оценена с учётом условной вероятности появления определённого значения  $j$ -го пикселя  $p_i(j)$ , если предыдущим был  $i$ -ый:

$$H_c = - \sum_i^m p(i) \sum_j^m p_i(j) \log_2 p_i(j) \quad (3)$$

В случае, когда статистическая связь между пикселями отсутствует, формулы (1) и (3) дадут одинаковую оценку энтропии[3]. Очевидно также, что в выражении (3) – формуле *условной энтропии* 1-го порядка, – используется далеко не единственный способ учёта статистической связи между пикселями: аналогичным образом могут учитываться, например, связи между парами пикселей или их группами, имеющими притом разную пространственную конфигурацию.

Снижению статистической избыточности (повышению энтропии) и уменьшению информационного объёма передающих изображения сообщений служат методы *энтропийного кодирования*, применяющиеся, в частности, при сжатии растровой информации без потери качества [см. напр. 3, 4]. Эффективность применения этих методов зависит во многом от степени декоррелированности изображения и способа выполнения декорреляции, снижающей как  $H$ , так и  $H_c$ , но не изменяющей информационный объём растра. Следует отметить, что декорреляция рассматривается и как биологический механизм оптимизации нейронального кодирования в проводящих путях от сетчатки к мозгу [5, 6].

Помимо статистической избыточности, в литературе широко [3, 7, 8] обсуждается *психофизическая избыточность* изображений, связанная с особенностями работы зрительной системы. В общем случае можно утверждать, что психофизически избыточное изображение содержит информацию, которая по ряду причин не воспринимается наблюдателем. В частности, это могут быть малые объекты низкого контраста или низкоамплитудные высокочастотные шумы. Методы снижения

психофизической избыточности (повышения информативности кодирующих изображение сообщений) принято называть методами сжатия с *потерей качества*, хотя на практике речь может идти как раз о повышении субъективной оценки *качества* воспринимаемого изображения.

В настоящее время методы снижения психофизической избыточности растров хорошо разработаны и востребованы. Например, в быту мы постоянно сталкиваемся с форматом сжатия jpeg, имеющим в качестве одной из важнейших основ упрощение пространственно-частотного спектра изображения. Эти методы в основном направлены на удаление из изображения информации, восприятие которой ограничено параметрами контрастной чувствительности зрения, с целью снижения энтропии раstra и достижения лучшего сжатия методами энтропийного кодирования.

Предлагаемая нами оценка количества информации в изображении входит в число множества оценок, подобных оценкам условной энтропии изображения, учитывающим взаимное расположение пикселей. Правда, в отличие от оценок такого типа, предлагаемая является психофизиологически обоснованной. С другой стороны, она позволяет узнать количество информации в изображении, избегая учёта психофизически избыточной информации. Предлагаемый подход к оценке количества информации в изображении может принести пользу при переходе от пиксельного описания изображения к векторному и объектному. Кроме того, он может использоваться при определении количества информации, получаемой испытуемым в психофизическом эксперименте.

### **Способ вычисления предлагаемой оценки**

Как известно, основным путём, по которому информация передаётся от сетчатки в кору больших полушарий мозга человека, является ретино-геникуло-стриарный путь. По классической теории Хьюбела и Визела [9],

---

ретиальные проекции оканчиваются на *простых* нейронах стриарной коры. Эти клетки выполняют функцию полосовых пространственно-частотных фильтров, характеризующихся также ориентационной и фазовой избирательностью. Структура их рецептивного поля, имеющего вытянутую форму и оппонентную организацию, может быть описана двумерной функцией Габора [10, 11]. Значения функции, используемой как ядро цифрового фильтра, удобно рассчитывать по формуле, предложенной в [12]:

$$G_{\lambda\theta\phi\sigma\gamma}(x,y) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right) \cos\left(2\pi\frac{x'}{\lambda} + \phi\right), \quad (4)$$

где

$$x' = x \cos \theta + y \sin \theta,$$

$$y' = -x \sin \theta + y \cos \theta.$$

Стандартное отклонение гауссовой составляющей  $\sigma$  связано с  $b$  – параметром, определяющим ширину полосы пропускания фильтра по пространственной частоте, – следующим соотношением:

$$b = \log_2 \frac{\frac{\sigma}{\lambda} \pi + \sqrt{\frac{\lambda^2}{2\sigma^2}}}{\frac{\sigma}{\lambda} \pi - \sqrt{\frac{\lambda^2}{2\sigma^2}}},$$

откуда

$$\sigma = \lambda \frac{1}{\pi} \sqrt{\frac{\lambda^2}{2} \left[ \frac{2^b + 1}{2^b - 1} \right]}. \quad (5)$$

Параметр  $\lambda$  выражения (5) задаёт длину волны косинусоидальной составляющей ядра функции, тем самым определяя центр полосы пропускания фильтра по пространственной частоте;  $\theta$  – пространственная ориентация функции, определяющая соответственно ориентационную избирательность фильтра;  $\phi$  – фазовый сдвиг косинусоидальной составляющей, задающий фазовую избирательность фильтра;  $\gamma$  –

«эллиптичность» функции, связанная с полосой пропускания фильтра по ориентации.

Известно, что в зрительной системе человека имеются 6 пространственно-частотных каналов, имеющих полосы пропускания шириной примерно 1,25-2,5 октавы по пространственной частоте с пиками на 0,5; 1; 2; 4; 8 и 16 периодах на угловой градус [13]. Обозначим их число как  $n_\lambda = 6$ . В модели мы задавали полосу пропускания  $b$  в 2 октавы для всех каналов. Полоса пропускания каждого канала по ориентации составляет порядка 30 градусов [14]. Число ориентационных предпочтений в зрительной коре, вероятно, очень велико, однако в модели мы, исходя из достаточности такого количества при выбранной ширине полосы пропускания, ограничились  $n_\theta = 6$  пиковыми ориентациями. Значения ядер фильтров для разных ориентационных каналов, настроенных на одну пространственную частоту, показаны на Рис. 1. Следует отметить, что в проведённых нами расчётах учитывались только ответы клеток оп-типа ( $\varphi=0$ ).

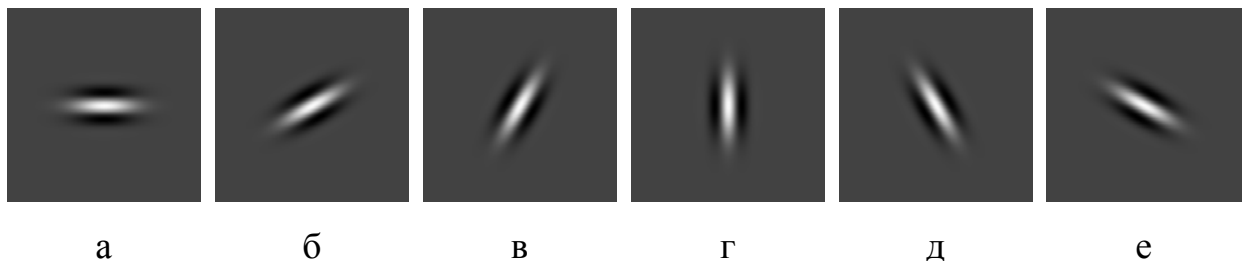


Рис. 1. – Значения функции Габора для  $\theta=0$  (а), 30 (б), 60 (в), 90 (г), 120 (д) и 150 (е) град.

Таким образом, для оценки количества активированных «нейронов» мы выполняли 36 операций линейной фильтрации с разными пиковыми значениями ориентационной и пространственно-частотной избирательности. На Рис. 2 даны примеры полосовой фильтрации изображения фильтрами Габора, настроенными на разную ориентацию. Значение оценки рассчитывалось отнесением суммы числа «активированных» элементов во



Основной проблемой используемого нами подхода, решение которой к тому же представляет отдельный интерес для исследователя зрительной системы, является отсутствие критерия важности информации, поступающей по разным пространственно-частотным каналам. Очевидно, что информация, приходящая в рецептивное поле простого нейрона, относящегося к низкочастотному каналу, описывает относительно большую область пространства, но сравнительно обобщённо. Известно также, что число простых стриарных нейронов в мозге положительно связано с частотой, на которую они настроены [9].

Так, в недавно проведённом нами психофизическом исследовании с идентификацией объектов на цифровом изображении, обработанном методом, описанным в [16], испытуемые в 94% случаев (доверительный интервал по [17] 89% до 97%,  $p < 0,05$ ) правильно идентифицировали пол незнакомого человека, основываясь на информации о менее чем 50% площади лица, описанного на частоте 2 периода на угловой градус. В то же время лицо, значительно большая площади которого описана на частоте 1 периода на угловой градус, правильно идентифицировалось лишь в 74% (от 68% до 82%,  $p < 0,05$ ) случаев.

Менее значимым, но требующим упоминания ограничением, заложенным при расчёте оценки, является фактическая бинаризация ответов «нейронов», хотя реальный стриарный нейрон кодирует величину контраста в своём рецептивном поле значительным количеством уровней активации.

### **Обсуждение**

Как известно, в растровой графике, в отличие от векторной, нет специальных средств кодирования образной (перцептивного уровня) и даже символической информации. Главным и единственным примитивом растровой графики выступает пиксель – светящаяся или светоотражающая

---





точка изображения, для которой известны относительные местоположение, яркость и/или цвет. Наборы точечных примитивов *объективно* представляют сцены визуального мира (реального, представляемого или воображаемого) уже потому, что атрибуты пикселей интерпретируются как вполне объективные физические величины. Растровые изображения создаются, хранятся и обрабатываются в основном для того, чтобы содержащаяся в них визуальная информация могла вновь стать доступной зрению человека, животного или, с существенными оговорками, технической системы. С другой стороны, субъективная информация, извлекаемая из изображения, вполне реальна и, как и физическая, может быть закодирована и количественно измерена; правда, по причине её субъективности количество возможных способов кодирования очень велико. Так, мы можем предложить нескольким людям нарисовать совершенно одинаковую сцену или дать её словесное описание – вероятность того, что интерпретации хотя бы двух человек совпадут, очень мала. Важно отметить, что по степени обобщённости субъективная информация является крайне неоднородной. Например, наблюдатель, осматривающий свой рабочий стол, может обнаружить, что он жёлтого цвета, что на нём расположены предметы – линейки, карандаши, ноутбук, что каждый из предметов обладает и характерными для своего класса, и специфическими признаками. Кроме того, наблюдатель может заметить беспорядок на своём рабочем месте, тем самым извлекая обобщённую информацию о свойствах сцены. Оценить субъективную информацию зрительной модальности количественно очень сложно. К сожалению, используемый авторами подход является преимущественно объективным и позволяет лишь косвенно оценить количество информации, поступающей в сознание субъекта. Создание такой оценки на данный момент можно отнести лишь к области очень отдалённых целей когнитивных нейронаук; возможно, путь к её разработке проложен

---

созданием иерархических моделей зрительной обработки [см. обзор 18] и представлениями о «нейронах сознания» и пирамиде гештальта [19].

**Работа выполнена при финансовой поддержке Минобрнауки России по теме № 213.01-11/2014-4 в рамках задания №2014/174 на выполнение государственных работ в сфере научной деятельности (базовая часть государственного задания).**

### Литература

1. Грузман И.С., Киричук В.С., Косых В.П., Перетягин Г.И., Спектор А.А. Цифровая обработка изображений в информационных системах: Учебное пособие. Новосибирск: Изд-во НГТУ, 2002. 352 с.
  2. Gonzalez R.C., Woods R.E., Eddins S.L. Digital image processing using MATLAB. New Jersey: Prentice Hall, 2003. 609 p.
  3. Красильников Н.Н. Цифровая обработка 2D- и 3D-изображений. Санкт-Петербург: БХВ-Петербург, 2011. 608 с.
  4. Востриков Д.А., Скакунов В.Н. Исследование способов передачи видеоданных по гидроакустическому каналу связи // Инженерный вестник Дона. 2012. №2. URL: [ivdon.ru/ru/magazine/archive/n2y2012/862](http://ivdon.ru/ru/magazine/archive/n2y2012/862).
  5. Тарасов В.В., Якушенков Ю.Г. Многоспектральные оптико-электронные системы // Специальная техника. 2002. № 4. С. 56-62.
  6. Pitkow X., Meister M. Decorrelation and efficient coding by retinal ganglion cells // Nature Neuroscience. 2012. Vol. 15, no. 4. P. 628-635.
  7. Немировский В.Б., Стоянов А.К. Предобработка изображений одномерными точечными отображениями // Известия Томского политехнического университета. 2011. Т. 319, №5. С. 107-111.
  8. Титов А.М. Методы сжатия спектра телевизионных сигналов изображения // Телекоммуникации и транспорт. 2012. №9. С. 136-139.
  9. Хьюбел Д. Глаз, мозг, зрение. Москва: Мир. 1990. 239 с.
-



10. Daugman J.G. Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. *J. Opt. Soc. Am. A*. 1985. Vol. 2, no. 7. P. 1160–1169.
  11. Jones J.P., Palmer L.A. An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex // *J. Neurophysiol.* 1987. Vol. 58, no. 6. P. 1233-1258.
  12. Grigorescu C., Petkov N., Westenberg M.A. Contour detection based on non-classical receptive field inhibition // *IEEE Trans. Image Processing.* 2003. Vol. 12, no. 7. P. 729–739.
  13. Wilson H.R., McFarlane D.K., Phillips G.C. Spatial frequency tuning of orientation selective units estimated by oblique masking // *Vision Research.* 1983. Vol. 23, no. 9. P. 873–882.
  14. DeValois R. L., DeValois K. K. *Spatial vision.* New York: Oxford University Press. 1988. 382 p.
  15. Bruce N.D.B., Tsotsos J.K. Saliency based on information maximization // *Advances in Neural Information Processing Systems* 18. Cambridge MA, London UK: MIT Press, 2005. P. 155-162.
  16. Явна Д.В., Бабенко В.В. Метод извлечения из цифровой фотографии лица информации, достаточной для его идентификации // *Инженерный вестник Дона.* 2014. № 3. URL: [ivdon.ru/ru/magazine/archive/n3y2014/2494](http://ivdon.ru/ru/magazine/archive/n3y2014/2494).
  17. Agresti A., Coull B.A. Approximate is better than “exact” for interval estimation of binomial proportions // *The American Statistician.* 1998. Vol. 52, no. 2. P. 119-126.
  18. Poggio T., Serre T. Models of visual cortex // *Scholarpedia.* 2013. Vol. 8, no. 4. URL: [dx.doi.org/10.4249/scholarpedia.3516](http://dx.doi.org/10.4249/scholarpedia.3516).
  19. Соколов Е.Н. Нейроны сознания // *Психология.* 2004. Т. 1, № 2. С. 3–15.
-

## References

1. Gruzman I.S., Kirichuk V.S., Kosykh V.P., Peretyagin G.I., Spektor A.A. Tsifrovaya obrabotka izobrazheniy v informatsionnykh sistemakh: Uchebnoe posobie [Digital image processing in information systems: Textbook (Rus)]. Novosibirsk: NSTU, 2002. 352 p.
  2. Gonzalez R.C., Woods R.E., Eddins S.L. Digital image processing using MATLAB. New Jersey: Prentice Hall, 2003. 609 p.
  3. Krasil'nikov N.N. Tsifrovaya obrabotka 2D- i 3D-izobrazheniy [Digital processing of 2D- and 3D-images (Rus)]. Sankt-Peterburg: BHV-Peterburg, 2011. 608 p.
  4. Vostrikov D.A., Skakunov V.N. Inženernyj vestnik Dona (Rus). 2012. No. 2. URL: [ivdon.ru/ru/magazine/archive/n2y2012/862](http://ivdon.ru/ru/magazine/archive/n2y2012/862).
  5. Tarasov V.V., Yakushenkov Yu.G. Spetsial'naya tekhnika (Rus). 2002. No. 4. P. 56-62.
  6. Pitkow X., Meister M. Decorrelation and efficient coding by retinal ganglion cells. Nature Neuroscience. 2012. Vol. 15, no. 4. P. 628-635.
  7. Nemirovskiy V.B., Stoyanov A.K. Izvestiya Tomskogo politekhnicheskogo universiteta (Rus). 2011. Vol. 319, no. 5. P. 107-111.
  8. Titov A.M. T-Comm - Telecommunications and Transport (Rus). 2012. No. 9. P. 136-139.
  9. Kh'yubel D. (Hubel D.) Glaz, mozg, zrenie [Eye, Brain, and Vision (Rus translation)]. Moskva: Mir. 1990. 239 p.
  10. Daugman J.G. Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. J. Opt. Soc. Am. A. 1985. Vol. 2, no. 7. P. 1160–1169.
  11. Jones J.P., Palmer L.A. An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex. J. Neurophysiol. 1987. Vol. 58, no. 6. P. 1233-1258.
-



12. Grigorescu C., Petkov N., Westenberg M.A. Contour detection based on non-classical receptive field inhibition. *IEEE Trans. Image Processing*. 2003. Vol. 12, no. 7. P. 729–739.

13. Wilson H.R., McFarlane D.K., Phillips G.C. Spatial frequency tuning of orientation selective units estimated by oblique masking. *Vision Research*. 1983. Vol. 23, no. 9. P. 873–882.

14. DeValois R. L., DeValois K. K. *Spatial vision*. New York: Oxford University Press. 1988. 382 p.

15. Bruce N.D.B., Tsotsos J.K. Saliency based on information maximization. *Advances in Neural Information Processing Systems 18*. Cambridge MA, London UK: MIT Press, 2005. P. 155-162.

16. Yavna D.V., Babenko V.V. *Inženernyj vestnik Dona (Rus)*. 2014. № 3. URL: [ivdon.ru/ru/magazine/archive/n3y2014/2494](http://ivdon.ru/ru/magazine/archive/n3y2014/2494).

17. Agresti A., Coull B.A. Approximate is better than “exact” for interval estimation of binomial proportions. *The American Statistician*. 1998. Vol. 52, no. 2. P. 119-126.

18. Poggio T., Serre T. Models of visual cortex. *Scholarpedia*. 2013. Vol. 8, no. 4. URL: [dx.doi.org/10.4249/scholarpedia.3516](http://dx.doi.org/10.4249/scholarpedia.3516).

19. Sokolov E.N. *Psikhologiya (Rus)*. 2004. Vol. 1, no 2. P. 3–15.