

НЕЙРОМОРФНОЕ ДЕКОДИРОВАНИЕ ВЫБОРОЧНЫХ ПРЕДСТАВЛЕНИЙ ИЗОБРАЖЕНИЙ МЕТОДОМ СОГЛАСОВАННОЙ С ГРАНИЦАМИ ИНТЕРПОЛЯЦИИ

© 2024 г. В. А. Кершнер

Институт радиотехники и электроники им. В.А. Котельникова РАН,
ул. Моховая, 11, стр. 7, Москва, 125009 Российская Федерация

E-mail: vladkershner@mail.ru

Поступила в редакцию 19.12.2023 г.

После доработки 20.05.2024 г.

Принята к публикации 25.05.2024 г.

Рассмотрены методы нейроморфного кодирования и декодирования больших объемов данных на основе моделирования известных нейромеханизмов восприятия информации. Исследованы известные механизмы зрительной системы, такие как агрегация отсчетов рецептивными полями, центрально-латеральное торможение и др. Разработана модель декодирования, реализующая функцию простых клеток первичной зрительной коры, отвечающих за пространственное восприятие контрастов стимулов. Предложена модель декодирования, позволяющая восстанавливать локальные границы объектов на изображении, улучшая при этом визуальное качество изображений в сравнении с качеством восстановления при классической билинейной интерполяции.

Ключевые слова: нейроморфные системы, выборочное представление, нейронное кодирование, система рецептивных полей, адаптивная интерполяция

DOI: 10.31857/S0033849424120064, **EDN:** HNBTUV

ВВЕДЕНИЕ

Последние десятилетия из-за стремительного развития сети Интернет, многочисленных его сервисов и разработок в области искусственного интеллекта объем информации повышается экспоненциально. Вследствие такого роста также расширился арсенал механизмов по использованию этих данных. Однако вместе с этим появилось огромное число специфических проблем, связанных преимущественно с хранением, обработкой и передачей огромных объемов данных как между корпорациями, так и между людьми, что получило название проблема Больших данных.

Наиболее четко проблемы с Большими данными проявляются в отношении аудиовизуального контента, который является наиболее востребованным в сети Интернет – изображения в соцсетях, аудиоматериалы на сайтах электронных книг, видеоролики в видеохостингах и т.п. Согласно статистике, полученной за 2023 г., было сделано около 1.81 трлн цифровых изображений, что составляет около 5 млрд изображений в день. Стоит отметить, что большая часть этих изображений была создана с использованием мобильных устройств. Благодаря развитию CMOS-камер и увеличению памяти смартфонов среднестатистический пользователь

за 2023 г. создал и сохранил на своих устройствах порядка 3000 фотографий и видео. В свою очередь, развитие социальных сетей и мессенджеров привело к увеличению объемов обмена данными до уровня в 6.9 млрд изображений ежедневно [1].

Поскольку вопросы хранения данных в настоящее время достаточно эффективно решаются на основе достижений микроэлектроники, все большее внимание уделяется задаче эффективной их передачи между пользователями (устройствами). В данном вопросе в сравнении с прочими видами данных, видеоданные являются наиболее проблемными, поскольку подразумевают передачу большого объема информации по каналам, имеющим принципиально ограниченную пропускную способность. Данная проблема давно известна под названием эффект «бутылочного горлышка», когда отдельный элемент информационной системы ограничивает ее производительность в целом. Ограничение производительности приводит к увеличению времени передачи информации, что влечет дискомфорт пользователям системы. Отметим, что проблема была актуальной задолго до появления Больших данных и, по сути, стала основным стимулом при создании современной теории информации К. Шенноном и его коллегами. Следует отметить связанную с проблемой «бутылочного горлышка» теорию

скорости–погрешности (rate–distortion theory) [2]. К сожалению, на практике результаты этой теории дают только потенциально возможные характеристики производительности систем обмена информацией вместо реально достижимых.

К настоящему времени большинство способов решения отмеченной проблемы найдено на путях сжатия (возможно с потерями) данных. В частности, разработано огромное множество методов компрессии изображений, среди которых можно упомянуть арифметическое сжатие, метод Хаффмана, широко используемый JPEG (Joint Photographic Experts Group) и др., оперирующие пиксельным представлением изображения. Однако стоит отметить, что существуют и другие оптимальные практические решения проблемы пропускной способности информационных систем. Речь идет о нейросенсорных системах, представленных в первую очередь зрительной системой высших позвоночных животных и человека. Если сравнить эти системы с имеющимися искусственными, то можно выделить, что число рецепторов сетчатки глаза достигает порядка 10^8 , что аналогично количеству "байт" изображения, а число аксонов зрительного нерва, или емкости канала передачи информации, — около 10^6 [3], что указывает на стократную степень сжатия входных данных, полученных с сетчатки глаза, без заметных визуальных дефектов. Детальное обсуждение ряда вопросов по данной теме было представлено в работе [4], где приведены также многочисленные ссылки на материалы по теме.

Цель данной работы — анализ данных, полученных в области адаптации инспирированных механизмами зрительного восприятия процедур нейронного кодирования к проблемам оптимального кодирования и декодирования изображений, и создание нового метода по обработке графических данных.

1. ФОТООТСЧЕТНЫЕ ВЫБОРОЧНЫЕ ПРЕДСТАВЛЕНИЯ

В основе предлагаемого метода лежит специальное представление изображений в виде набора фотоотсчетов, который моделирует результирующий поток случайных дискретных событий, зарегистрированных фоторецепторами сетчатки глаза, в ответ на воздействие непрерывно распределенной интенсивности излучения $I(\vec{x})$, $\vec{x} \in \Omega$. Существующие системы формирования изображений, к примеру CMOS-камеры [5], формируют исходные изображения аналогичным образом, регистрируя отдельные фотоны и генерируя фототок в виде дискретных отсчетов. Представление сигналов в виде потоков событий и их статистическое описание с использованием двумерных точечных пуассоновских процессов можно найти в [6], в свою очередь, схожая модель таких сигналов

может быть получена с использованием разработанного в предыдущих наших работах понятия идеального устройства формирования изображения [4]. В основе представления лежит набор координат отсчетов — зарегистрированных событий $X = \{\vec{x}_i\}$, где \vec{x}_i , $i = 1, \dots, N$ — число случайных точек на светочувствительной поверхности устройства Ω , причем количество зарегистрированных отсчетов N также является величиной случайной, заданной распределением Пуассона со средним значением

$$\bar{N} = \int_{\Omega} \alpha I(\vec{x}) d\vec{x},$$

где коэффициент $\alpha = \eta(h\bar{v})^{-1}$ определяется в том числе квантовой эффективностью η идеального устройства формирования изображения [4].

Собственно набор случайных отсчетов описывается распределением вероятностей случайного числа точек $\{\vec{x}_i\}$ некоторого неоднородного точечного пуассоновского процесса с функцией интенсивности $\alpha I(\vec{x})$ [6]. Поскольку количество отсчетов N является случайной величиной, то указанное описание представляется неудобным для практического использования, в частности при больших значениях N . Поэтому было предложено использовать представление точечных процессов в виде набора случайных точек, как в оригинальном пуассоновском процессе, однако общее число отсчетов фиксировано $N_p \ll \bar{N}$. Фиксированная выборка $X_p = \{\vec{x}_j\}$ случайного размера N_p , полученная из $X = \{\vec{x}_i\}$, может быть задана плотностью распределения [7] следующего вида:

$$\rho(X_p = \{\vec{x}_j\}, | I(\vec{x})) = \prod_{j=1}^{N_p} \rho(\vec{x}_j | I(\vec{x})), \quad (1)$$

$$\rho(\vec{x}_j | I(\vec{x})) = I(\vec{x}_j) / \int_{\Omega} I(\vec{x}) d\vec{x}.$$

Представление изображений с использованием фиксированных выборок $X_p = \{\vec{x}_j\}, j = 1, \dots, N_p$, статистически определенных в (1), было предложено называть выборочным представлением изображений.

На рис. 1 приведен пример сформированного выборочного изображения на основе взятого из стандартной базы данных MPEG7 [8] изображения "butterfly-19". Для формирования выборки $X_p = \{\vec{x}_j\}$, состоящей из $N_p = 10\,000\,000$ отсчетов, исходное изображение размером 429×421 пикселей, глубиной цвета $v = 8$ бит было предварительно конвертировано в формат PNG с той же глубиной цвета, но размером $s \times s = 1\,000 \times 1\,000$ пикселей. При этом в конечном варианте изображения использовались всего два оттенка серого — 100 и 255. Отсчеты генерировались методом отбраковки Монте-Карло с равномерным вспомогательным распределением $u(\vec{x}) = s^{-2}$ и вспомогательной константой $M = 2^v$.

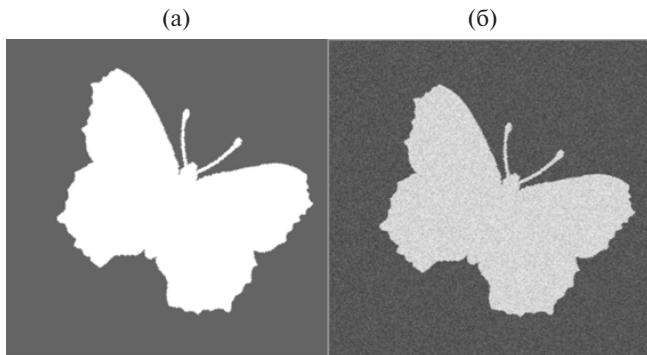


Рис. 1. Представление изображения по выборке отсчетов (выборочное представление): а – оригинальное изображение «butterfly-19» [8], б – выборочное представление размером 10 млн отсчетов.

2. МОДЕЛЬ РЕЦЕПТИВНЫХ ПОЛЕЙ ДЛЯ ОБРАБОТКИ ВЫБОРОЧНЫХ ПРЕДСТАВЛЕНИЙ

Предложенное выборочное представление $X_p = \{\vec{x}_j\}$ наиболее точно описывает данные, полученные с рецепторов внешнего слоя сетчатки – палочек и колбочек. Эти данные генерируются при формировании изображений на основе регистрируемой интенсивности светового сигнала $I(\vec{x})$, $\vec{x} \in \Omega$. Однако, как упоминалось ранее, посылаемые в кору головного мозга импульсы от нейронов внутреннего слоя сетчатки имеют существенное отличие от данных, зарегистрированных на фоторецепторных клетках. Благодаря многочисленным промежуточным интернейронам среднего и внутреннего слоев сетчатки, входной сигнал существенно преобразуется. В результате нейроны на выходе сетчатки передают агрегированные данные от десятков, а порой и тысяч фоторецепторов, расположенных в малых областях входного слоя сетчатки и образующих так называемые рецептивные поля (РП). Изучение систем РП и нейронного преобразования данных от входного слоя фоторецепторов в последовательность нейронных импульсов зрительного нерва обычно связывают с фундаментальными исследованиями Хьюбеля и Визеля [9]; современное представление строения зрительной системы и особенностях функционирования рецептивных полей можно найти, например, в [10].

Стоит отметить, что выходные нейроны сетчатки, представленные набором ганглиозных клеток, весьма разнообразны как по форме, так и по функциям – в сетчатке выделяются более 20 типов этих клеток. Соответственно, функции и строение различных РП значительно дифференцированы, поскольку определяются типами связанных с этими полями ганглиозных клеток, нейронов внутреннего слоя и соответствующих им фоторецепторов. В работе рассматривается только семейство карликовых клеток, основная функция которых – кодирование

пространственного распределения интенсивности $I(\vec{x})$ по всей поверхности сетчатки Ω . Отметим, что большинство РП обладает центро-антагонистической структурой, которая определяет реакцию на характер освещения либо затемнения. Так, ON-клетки РП возбуждаются при световой стимуляции центра поля и тормозятся при стимуляции его периферии, и наоборот, OFF-клетки возбуждаются при стимуляции РП-периферии и тормозятся при стимуляции центра [9]. Наличие двух типов клеток обусловлено особенностями нервного кодирования положительных/отрицательных перепадов в стимулах (ON-клетка активируется, когда стимуляция центра превышает среднюю стимуляцию по РП, OFF-клетка – наоборот).

Что касается структуры РП, то известно, что соседние ON- и OFF-клетки имеют значительно перекрывающиеся поля, а РП-клетки одинакового типа практически не перекрываются. При этом неперекрывающиеся РП каждого типа плотно прилегают друг к другу, образуя сетчатую структуру, покрывающую всю сетчатку глаза [11]. Если допустить пересечение ON-полей только с соседними OFF- полями и отсутствие пересечений с прочими ON- полями, при этом допускавшая соприкосновение их границ, можно окончательно формализовать пространственную геометрию РП. Аналогичное справедливо при взаимной замене ON-клеток на OFF-клетки, благодаря их симметрии. Формальное представление области Ω квадратными РП с круглыми центрами, которое используется в дальнейшем для алгоритмических целей, представлено на рис. 2. (Подробнее о структуре РП см. в [12].)

Ввиду ранее упомянутого симметричного относительно друг друга расположения ON- и OFF- полей, можно рассмотреть систему РП гибридного типа, способную кодировать как положительные (по образу ON-полей), так и отрицательные (по образу OFF-полей) перепады интенсивности светового сигнала. На основе приведенных выше

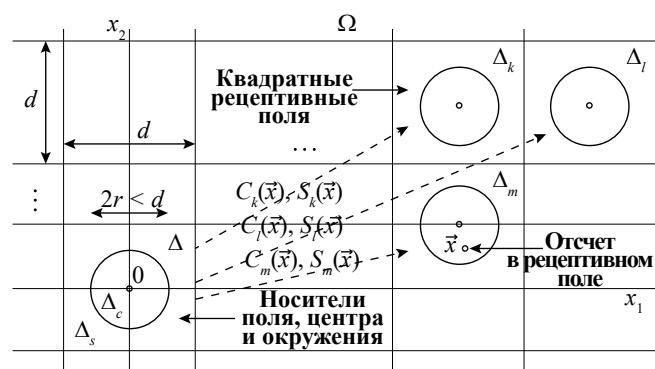


Рис. 2. Разбиение поверхности изображения Ω системой рецептивных ON-полей $\{C_k(\vec{x}), S_k(\vec{x})\}$ с квадратными носителями Δ_k , Δ_s , расположенными в узлах регулярной квадратной решетки.

данных мы можем формализовать модель плотности вероятности отсчета \vec{x}_j на поверхности изображения $\Omega - \rho(\vec{x}_j | I(\vec{x}))$ (1), как параметрическое семейство плотностей вероятностей $\mathbb{P} = \{\rho(\vec{x}; \vec{\theta}) | \vec{\theta} \in \Theta\}$, представляющих собой смеси K пар компонентов $\{C_k(\vec{x}), S_k(\vec{x})\}, k = 1, \dots, K$:

$$\rho(\vec{x}; \vec{\theta}) = \sum_{k=1}^K w_k C_k(\vec{x}) + v_k S_k(\vec{x}), \quad (2)$$

где параметры $\vec{\theta} = \{w_k, v_k\}$ модели \mathbb{P} представляют собой положительные веса смеси, а компоненты смеси $C_k(\vec{x})$ и $S_k(\vec{x})$ могут быть интерпретированы как наборы компонент центра и антагонистического окружения k -го РП соответственно. Причем компоненты центра и окружения задаются положительными, нормированными плотностями распределения вероятностей с компактными носителями $\Delta_k^c = \{\vec{x} | C_k(\vec{x}) > 0\}$ и $\Delta_k^s = \{\vec{x} | S_k(\vec{x}) > 0\}$, которые составляют в объединении носители k -х РП $\Delta_k = \Delta_k^c \cup \Delta_k^s$ (см. рис. 2):

$$\int_{\Delta_k^c} C_k(\vec{x}) d\vec{x} = \int_{\Delta_k^s} S_k(\vec{x}) d\vec{x} = 1. \quad (3)$$

Учитывая тождественность плотности $\rho(\vec{x}_j | I(\vec{x}))$ нормированной версии интенсивности сигнала $I(\vec{x})$, зафиксированную в (1), становится очевидным смысл введенных параметров $\vec{\theta} \in \Theta$. Уравнение (2), по существу, задает разложение интенсивности $I(\vec{x})$ по системе локальных распределений с точностью до нормирующего множителя, как это принято в любом мультиразрешающем анализе [13]. При этом соотношениями (3) определена нормировка заданных функций.

Далее, если предположить, что носители центра Δ_k^c и антагонистического окружения Δ_k^s k -го РП не имеют общих точек $\Delta_k^c \cap \Delta_k^s = \emptyset$, то можно добавить соотношения ортогональности в дополнение к приведенным соотношениям нормировки (3):

$$\int_{\Delta_k^c} S_k(\vec{x}) d\vec{x} = \int_{\Delta_k^s} C_k(\vec{x}) d\vec{x} = 0. \quad (4)$$

Учитывая, что множество носителей РП $\{\Delta_k\}$ составляет точное мозаичное разбиение поверхности сетчатки, т.е. все они плотно покрывают все пространство Ω , но при этом носители центра Δ_k^c и окружения Δ_k^s попарно не пересекаются, можно, так же как в мультиразрешающем анализе, выразить параметры $\vec{\theta} = \{w_k, v_k\}$ модели через соответствующие интегралы плотности $\rho(\vec{x}; \vec{\theta})$ (2) по соответствующим носителям РП:

$$\begin{aligned} w_k &= \int_{\Delta_k^c} \rho(\vec{x}; \vec{\theta}) d\vec{x}, \\ v_k &= \int_{\Delta_k^s} \rho(\vec{x}; \vec{\theta}) d\vec{x}, \end{aligned} \quad (5)$$

где параметры w_k, v_k могут быть интерпретированы как вероятности попадания отсчета \vec{x}_j в центр Δ_k^c либо в окружение Δ_k^s k -го РП. Также можно представить параметры w_k, v_k (5) как средние значения

характеристических функций, для которых справедливы следующие утверждения: $\Pi_k^c(\vec{x}) = 1$, если $\vec{x} \in \Delta_k^c$, иначе $\Pi_k^c(\vec{x}) = 0$, и $\Pi_k^s(\vec{x}) = 1$, если $\vec{x} \in \Delta_k^s$, иначе $\Pi_k^s(\vec{x}) = 0$ по всей поверхности Ω .

Поскольку плотность $\rho(\vec{x}; \vec{\theta})$ не известна, а в отношении ее известна только выборка $X_p = \{\vec{x}_j\}$, то связи (5) не могут быть использованы для нахождения w_k, v_k . Однако, используя стандартный прием из [16], с учетом асимптотики закона больших чисел можно заменить средние $\Pi_k^c(\vec{x})$ и $\Pi_k^s(\vec{x})$ их выборочными (эмпирическими) средними, что приводит к следующему представлению параметров:

$$\begin{aligned} w_k &= \frac{1}{N_s} \sum_{j=1}^{N_s} \Pi_k^c(\vec{x}_j) = \frac{n_k^c}{N_p}, \\ v_k &= \frac{1}{N_s} \sum_{j=1}^{N_s} \Pi_k^s(\vec{x}_j) = \frac{n_k^s}{N_p}, \end{aligned} \quad (6)$$

где n_k^c и n_k^s – числа отсчетов соответственно в центре и в окружении k -го РП.

Стоит отметить, что приближенные значения параметров (6) не зависят от вида компонент $C_k(\vec{x})$ и $S_k(\vec{x})$, а только от вида их носителей Δ_k^c и Δ_k^s . Отсюда следует, что для приближенной оценки плотности вероятности $\rho(\vec{x}; \vec{\theta})$ (3) достаточны лишь числа n_k^c и n_k^s отсчетов в центрах и окружениях рецептивных полей. Другими словами, выборочное представление $X_p = \{\vec{x}_j\}$ изображения может быть сокращено (сжато) в рассматриваемом случае до представления “числами заполнения” $Y_p = \{n_k^c, n_k^s\}$, которые в данном контексте являются достаточными статистиками выборочного представления.

3. КОДИРОВАНИЕ ВЫБОРОЧНЫХ ДАННЫХ, СЖАТЫХ С ИСПОЛЬЗОВАНИЕМ СИСТЕМЫ РЕЦЕПТИВНЫХ ПОЛЕЙ

Используя наборы чисел заполнения $Y_{N_p} = \{n_k^c, n_k^s\}$, оценим, каким образом эти данные можно закодировать для дальнейшей передачи на последующие этапы обработки. Для удобства можно считать систему рецептивных полей однородной на всем пространстве Ω , то есть все поля идентичны друг другу по выполняемым функциям и по своему строению, как это показано на рис. 2. Типичное поле имеет площадь $\sigma = |\Delta|$, состоящую из компактного центра площади $\sigma_c = |\Delta^c|$ и антагонистического окружения с площадью $\sigma_s = |\Delta^s|$, причем эти площади не пересекаются, откуда следует, что $\sigma = \sigma_c + \sigma_s$.

Каждое РП может определять общее число принадлежащих ему отсчетов n , состоящее из числа отсчетов в центре n_c и окружении n_s , причем количество отсчетов в окружении может быть определено в виде линейной комбинации $an + bn_c$. Коэффициенты a и b могут быть произвольными как

по величине, так и по знаку, что интерпретируется в данном случае, как возбуждающее либо тормозящее действие соответствующих областей поля. В частном случае эти коэффициенты могут принимать единичные значения, при которых справедлива зависимость $n_s = n - n_c$. Поскольку регистрация отсчетов имеет случайный характер, числа n , n_c и n_s также являются случайными величинами, однако одновременно с этим являются взаимозависимыми. Для выборочных представлений $X_p = \{\vec{x}_j\}$ данные числа являются пуассоновскими случайными величинами [4]. Учитывая статистическую независимость n_c и n_s ввиду того условия, что области центра и окружения отдельного РП являются непересекающимися, можно представить распределения вероятностей в следующем виде:

$$\begin{aligned} P_c(n_c|\lambda) &= \frac{(\sigma_c \lambda)^{n_c}}{n_c!} \exp\{-\sigma_c \lambda\}, \\ P_s(n_s|\mu) &= \frac{(\sigma_s \mu)^{n_s}}{n_s!} \exp\{-\sigma_s \mu\}, \end{aligned} \quad (7)$$

где λ и μ по сути представляют интенсивность отсчетов в центре и в окружении РП, которые можно представить следующим образом:

$$\lambda = \frac{\alpha}{\sigma_c} \int_{\Delta^c} I(x) dx, \quad \mu = \frac{\alpha}{\sigma_s} \int_{\Delta^s} I(x) dx. \quad (8)$$

Можно заметить, что средние значения распределений (7) связаны с (8) следующим образом: $\bar{n}_c = \sigma_c \lambda$ и $\bar{n}_s = \sigma_s \mu$. Поэтому значения отсчетов в центре и окружении РП n_c и n_s соответственно, являясь несмещенными оценками своих средних \bar{n}_c и \bar{n}_s , предоставляют также соответствующие им несмещенные оценки \bar{n}_c / σ_c и \bar{n}_s / σ_s значений интенсивностей отсчетов λ и μ .

Путем обычного перемножения распределений (7) можно получить совместные распределения для отсчетов n_c и n_s . Если же от этих данных перейти сначала к значениям n_c в центре и n на всем РП и далее к $\delta = n_c - (\sigma_c/\sigma)n$ и n , то после ряда преобразований с использованием аппроксимации биноминального распределения гауссовским и дальнейшим упрощением, можно перейти к следующей статистической модели данных:

$$\begin{aligned} P(\delta, n|\lambda, \mu) &= P(\delta|n, \lambda, \mu)P(n|v), \\ P(\delta|n, \lambda, \mu) &= \frac{1}{\sqrt{2\pi\gamma^2}} \exp\left(-\frac{(\delta - \varepsilon)^2}{2\gamma^2}\right), \\ \varepsilon &= \frac{\sigma_c \sigma_s}{\sigma^2} \left(\frac{\lambda - \mu}{v}\right)n, \quad \gamma^2 = \frac{\sigma_c \sigma_s}{\sigma_i^2} \frac{\lambda \mu}{v^2} n, \\ P(n|v) &= \frac{(\sigma v)^n}{n!} \exp\{-\sigma v\}, \quad v = \frac{\sigma_c}{\sigma} \lambda + \frac{\sigma_s}{\sigma} \mu. \end{aligned} \quad (9)$$

Однако для полного статистического описания модели рецептивного поля следует использовать априорное совместное распределение для значений интенсивностей отсчетов центра и окружения λ и μ , которое можно выбрать, например, в виде:

$$\rho(\lambda|\mu) = \omega \delta(\lambda - \mu) + (1 - \omega) \varphi(\lambda), \quad (10)$$

где ω и $(1 - \omega)$ представляют веса, которые можно трактовать как вероятность 0-гипотезы H_0 о том, что интенсивности λ и μ зависят, либо наоборот, как вероятность независимости значений λ и μ , т.е. альтернативы этой гипотезы. Следовательно, $\varphi(\lambda)$ является априорным безусловным распределением вероятностей каждой из интенсивностей λ и μ .

Используя стандартные статистические методы, на основе статистической модели (9), (10) с помощью данных δ и n могут быть получены апостериорные распределения интенсивностей λ и μ , а также их оптимальные оценки, тесты проверки 0-гипотезы H_0 и т.д. Выбрав для значений λ и μ апостериорные распределения $\rho(\lambda, \mu|n_c, n)$, можно выразить первые моменты, которые могут представлять МАР (оценки апостериорного максимума) оценок $\bar{\lambda}(\delta, n)$ и $\bar{\mu}(\delta, n)$:

$$\begin{aligned} \bar{\lambda}(\delta, n) &= \begin{cases} n / \sigma, & |\delta| \leq D\sqrt{n}, \\ n_c / \sigma_c, & |\delta| > D\sqrt{n} \end{cases}, \\ \bar{\mu}(\delta, n) &= \begin{cases} n / \sigma, & |\delta| \leq D\sqrt{n} \\ n_s / \sigma_s, & |\delta| > D\sqrt{n} \end{cases}, \end{aligned} \quad (11)$$

где $D^2 = 2\sigma_c \sigma_s \ln\{\bar{\Lambda}_0\} / \sigma^2$ – пороговый коэффициент, представленный исключительно через параметры задачи, включая параметр априорного правдоподобия $\bar{\Lambda}_0$, который, при разумных предположениях, может быть задан следующим приближенным выражением

$$\bar{\Lambda}_0 = \frac{4\omega}{\pi(1-\omega)} \sqrt{\frac{\sigma_c \sigma_s \bar{n}}{\sigma^2}}, \quad (12)$$

где $\bar{n} = \sigma \hat{v}$, \hat{v} – характерный масштаб априорного распределения вероятностей $\varphi(\lambda)$. На рис. 3 представлено кодирование данных δ и n в пороговые оценки $\bar{\lambda}(\delta, n)$ и $\bar{\mu}(\delta, n)$ на решетке 50×50 рецептивных полей (исходное изображение было представлено на рис. 1).

Кодирование (11) можно интерпретировать как регрессию данных при их сжатии [14], если допустить, что n/σ является общей оценкой для значений λ и μ на основе предиктора n/σ (при условии, что имеет место 0-гипотеза H_0), а $\delta/\sigma_c = n_c/\sigma_c - n/\sigma$, $\delta/\sigma_s = n_s/\sigma_s - n/\sigma$ являются остатками для этой оценки. Если сравнить данный подход с методом регрессии данных LASSO (Least Absolute Shrinkage and Selection Operator) [14], можно отметить зависимость общего числа отсчетов n на области σ отдельного РП

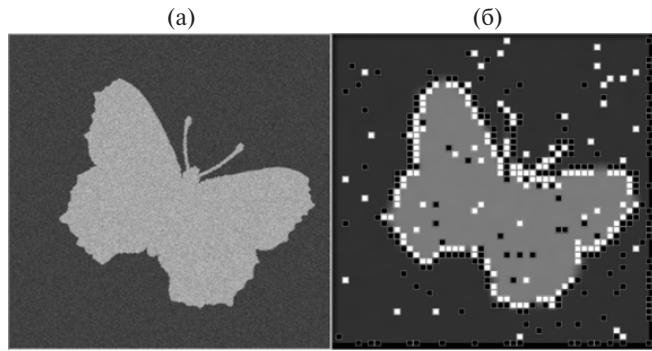


Рис. 3. Иллюстрация процедуры кодирования (11) на решетке рецептивных полей размером 50×50 выборочного представления изображения «butterfly-19» [8] (см. рис. 1): а – выборочное представление, б – РП с ненулевыми значениями δ , белым цветом выделены ON-ответы ($\delta > 0$), черным – OFF-ответы ($\delta < 0$).

и выражения порогового значения $D\sqrt{n}$. Причем количество отсчетов связано с зависимостью интенсивности $I(\vec{x})$ сигнала изображения и пуассоновским шумом. Таким образом, (11) представляет собой нелинейный центро-латеральный пороговый фильтр.

4. ВОССТАНОВЛЕНИЕ ИЗОБРАЖЕНИЙ С ИСПОЛЬЗОВАНИЕМ СОГЛАСОВАННОЙ С ГРАНИЦАМИ ИНТЕРПОЛЯЦИИ

Используя известные методы интерполяции изображений с низким разрешением, можно провести оценку эффективности предложенной процедуры кодирования (11). Среди таких методов можно выделить ряд широко используемых интерполяционных алгоритмов (билинейная интерполяция, бикубическая интерполяция и метод ближайшего соседа), алгоритмы регуляризации, спектральное представление. В данной работе был выбран простейший метод билинейной интерполяции, что, однако, не исключает возможность использования более эффективных в смысле точности методов, в частности связанных с нейронными сетями, используемыми для решения задачи формирования интерполяционных изображений.

Поскольку часть кода с числами отсчетов $\{n_k\}$ по сути является выходом сглаживающего фильтра со скользящим окном площади σ , можно однозначно говорить о размытой форме искажений исходного изображения (рис. 4а). В свою очередь, декодирование, связанное с частью кода, представляющей так называемые “дефекты” $\{\delta_k\}$, реализуется ступенчатым фильтром, принимающим постоянные и положительные значения в области центра РП, и также постоянные, но уже отрицательные – в его окружении, причем полный интеграл отклика равен

нулю. Данный тип фильтров широко распространен в цифровой обработке изображений, и в частном случае полученный фильтр совпадает с фильтром COSO (center-ON-surround-OFF), предложенным в [15]. Последний был предложен в качестве имитации отклика LoG-фильтра (Laplacian of Gaussian) и использован Марром и Хилдертром в теории обнаружения границ на изображениях [16].

Коды $\{\delta_k\}$ оказываются тесно связанными с оператором Лапласа, нули их линейной интерполяции вдоль ребер решетки РП представляют собой точки пересечения с ними границ изображений, поэтому для реконструкции изображений, сохраняющих содержательные детали, можно использовать многочисленные методы вдоль-реберной интерполяции (edge-directed interpolation) [17]. Стоит еще раз отметить, что подобные интерполяции

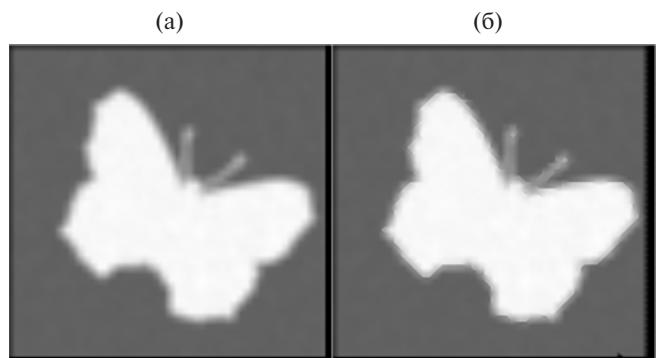


Рис. 4. Восстановленное (декодированное) изображение «butterfly-19» [8] (см. рис. 1), заданное выборочным представлением $\{n_k, \delta_k\}$ на решетке 50×50 : а – сглаженное изображение, декодированное только с использованием “плавной” части $\{n_k\}$ кода, б – интерполяция вдоль ребер, заданных деталями $\{\delta_k\}$.

предоставляют на сегодняшний день наилучшее перспективное качество реконструированных изображений.

Для иллюстрации реконструкции (декодирования) изображений была использована простейшая билинейная, согласованная с границами интерполяция для кодированного выборочного представления «butterfly-19» (см. рис. 1б). Область представления Ω была покрыта набором из 2500 квадратных рецептивных полей, расположенных в узлах квадратной решетки 50×50 (см. рис. 2 и детали в [12]). В каждом k -м узле вычислялись значения $\{\delta_k, n_k\}$, после чего δ_k подвергалось цензуре с порогом $D\sqrt{n_k}$. Классическая билинейная интерполяция, показанная на рис. 4а, демонстрирует реконструкцию на основе только части данных $\{n_k\}$, интерполируя их сначала линейно вдоль вертикальных ребер решетки, а затем линейно по всем строкам всех

ячеек на основе уже интерполированных значений вертикальных ребер. Билинейная интерполяция, направленная вдоль детектированных границ изображения, показана на рис. 4б. В отличие от классической билинейной интерполяции, она дополнительно оперирует значениями дефектов $\{\delta_k\}$, которые представляют некоторое подобие оператора Лапласа. Если рассматривать величину дефекта δ_k подробно, то он представляет собой более общее значение, чем лапласиан, что позволяет использовать его подобно оператору Марра в задачах обнаружения границ. В таком случае билинейная интерполяция также использует двухпроходную реконструкцию. При первом проходе значения $\{n_k\}$ также интерполируются по всем вертикальным и горизонтальным ребрам решетки. Но в случае, когда в узлах данного ребра решетки значения δ_i и δ_j отличны от нуля и разных знаков, это ребро решетки рассматривается как пересекающее границу изображения, причем за точку пересечения принимается середина ребра решетки. В результате, если ребро решетки пересекается границей изображения, значения n_i и n_j в его узлах кусочно-постоянно продолжаются вдоль ребра решетки до встречи ровно посередине. Если же ребро не пересекает границу изображения, для него производится линейная интерполяция $n_i \leftrightarrow n_j$, как и в классическом случае. На втором проходе значения в ячейках решетки линейно интерполируются по значениям на их ребрах. Если ни одно ребро ячейки не пересекает границы изображения, или таких ребер одно или более двух, то интерполяция производится по строкам ячеек, как при классической интерполяции. Если же в ячейке имеется ровно два ребра решетки пересекающихся с границами изображения, то, соединяя точки этих пересечений прямой линией, мы восстанавливаем фрагмент границы изображения в ячейке и проводим линейную интерполяцию по прямым, параллельным этой линии. Результат (см. рис. 4б) представляется обладающим лучшим визуальным качеством даже по сравнению с классической билинейной интерполяцией (см. рис. 4а).

ЗАКЛЮЧЕНИЕ

В работе были представлены методы кодирования и декодирования входных данных на основе нейроморфной модели, основанной на известных нейромеханизмах восприятия зрительной системы. В качестве данных для обработки было использовано формализованное ранее представление входных данных в виде пуассоновского набора случайных точек, представленного в виде выборочного представления. Исходное изображение было взято из стандартной базы данных MPEG7 [4] и предварительно обработано. Модель кодирования данных была формализована в контексте концепции рецептивных полей. Разработанная процедура порогового кодирования

информации была определена как сжатие данных, схожее с принципом нелинейного центрально-латерального порогового фильтра.

Предложенный метод восстановления данных был формализован в виде модели связанного ON-OFF-декодирования, как реконструкция ранее сжатого изображения на основе участков изменения пространственных контрастов. Оценка качества восстановленного изображения была произведена путем визуального сравнения с изображением, восстановленным с использованием алгоритма классической интерполяции. Преимуществом используемого метода оказалось более четкое выделение краевых областей на изображении.

КОНФЛИКТ ИНТЕРЕСОВ

Авторы заявляют об отсутствии конфликта интересов.

ФИНАНСИРОВАНИЕ РАБОТЫ

Работа выполнена за счет бюджетного финансирования в рамках государственного задания в Институте радиотехники и электроники им. В. А. Котельникова РАН (AAAA-A 19-119041590070-1).

СПИСОК ЛИТЕРАТУРЫ

1. *Lu Z., Huang D., Bai L. et al.* // arXiv preprint arXiv:2304.13023. 2023. <https://doi.org/10.48550/arXiv.2304.13023>
2. *Pinkston J. T.* // IEEE Trans. 1969. V. IT-15. № 1 P. 66. <https://doi.org/10.1109/TIT.1969.1054274>
3. *Milner D., Goodale M.* The Visual Brain in Action. Oxford: Univ. Press, 2006. <https://doi.org/10.1093/acprof:oso/9780198524724.001.0001>
4. *Antsiperov V., Kershner V.* // Pattern Recognition Applications and Methods, ICPRAM 2021–2022. Lecture Notes in Computer Sci. P. 13822. Cham: Springer, 2023. https://doi.org/10.1007/978-3-031-24538-1_3
5. *Yang M., Sun X., Jia F. et al.* // Polymers. 2022. V. 14. № 10. P. 2019. <https://doi.org/10.3390/polym14102019>
6. *Keeler H. P.* Notes on the Poisson Point Process. Technical Report. Berlin: Weierstrass Inst. 2016. 36 p. <https://h paulkeeler.com/wp-content/uploads/2018/08/PoissonPointProcess.pdf>
7. *Antsiperov V.* // Proc. 11th Int. Conf. on Pattern Recognition Applications and Methods – ICPRAM. 3–5 Feb. 2022. Setúbal: SciTePress – Science and Technology Publ., 2022. P. 354. <https://doi.org/10.5220/0010836800003122>
8. *Latecki L. J., Lakamper R., Eckhardt T.* // Proc. IEEE Conf. Computer Vision and Pattern Recognition, CVPR-2000. Hilton Head Island. 15 Jun. N.Y.: IEEE, 2000. P. 424.

- https://doi.org/10.1109/CVPR.2000.855850
9. Hubel D. H., Wiesel T. N. *Brain and Visual Perception: The Story of a 25-year Collaboration*. Oxford: Univ. Press, 2004.
https://doi.org/10.1016/0001-6918(64)90136-2
10. Keller A. J., Roth M. M., Scanziani M. // *Nature*. 2020. V. 582. № 7813. P. 545.
https://doi.org/10.1038/s41586-020-2319-4
11. Hoon M., Okawa H., Santina L. D., Wong R. O. // *Progress in Retinal and Eye Research*. 2014. V. 42. P. 44.
https://doi.org/10.1016/j.preteyes.2014.06.003
12. Antsiperov V. // *Proc. 12th Int. Conf. on Pattern Recognition Applications and Methods (ICPRAM 2023)*. Lisbon. 22–24 Feb. 2023. Setúbal: SciTePress – Science and Technology Publ., 2023. P. 517.
https://doi.org/10.5220/0011792800003411
13. Fish J., Wagner G. J., Keten S. // *Nature Mater.* 2021. V. 20. № 6. P. 774.
https://doi.org/10.1038/s41563-020-00913-0
14. Ranstam J., Cook J. A. // *J. British Surgery*. 2018. V. 105. № 10. P. 1348.
https://doi.org/10.1002/bjs.10895
15. Tam W. S., Kok C. W., Siu W. C. // *J. Electron. Imaging*. 2010. V. 19. № 1. P. 013011.
https://doi.org/10.1117/1.3358372
16. Marr D., Hildreth E. // *Proc. Royal Society of London. Ser. B. Biol Sci.* 1980. V. 207. № 1167. P. 187.
https://doi.org/10.1098/rspb.1980.0020. PMID6102765.
17. Yu S., Zhang R., Wu Sh. et al. // *Biomedical Engineering Online*. 2013. V. 12. P. 1.
https://doi.org/10.1186/1475-925X-12-102

NEUROMORPHIC DECODING OF SAMPLE IMAGE REPRESENTATIONS BY THE BOUNDARY-CONSISTENT INTERPOLATION METHOD

V. A. Kershner

*Kotel'nikov Institute of Radio Engineering and Electronics, Russian Academy of Sciences,
Mokhovaya Str., 11, Build. 7, Moscow, 125009 Russian Federation
E-mail: vladkershner@mail.ru*

Received December 19, 2023, revised May 20, 2024, accepted May 25, 2024

The paper discusses methods for encoding and decoding large amounts of data using a neuromorphic model based on known neuromechanisms for the perception of visual information. Known mechanisms of the visual system, such as aggregation of counts by receptive fields, central-lateral inhibition, etc., have been studied. A decoding model has been developed that implements the function of simple cells of the primary visual cortex responsible for spatial perception of stimulus contrasts. The proposed decoding model makes it possible to restore local boundaries of objects in an image, while improving the visual quality of images in comparison with the quality of restoration with classical bilinear interpolation.

Keywords: neuromorphic systems, sample representation, neural coding, receptive field system, adaptive interpolation