

DOI: <https://doi.org/10.30898/1684-1719.2021.9.2>

УДК: 004.932.1; 004.932.2; 004.942

КОДИРОВАНИЕ ИЗОБРАЖЕНИЙ ВЫБОРКАМИ ОТСЧЕТОВ, МОТИВИРОВАННОЕ МЕХАНИЗМАМИ РЕГИСТРАЦИИ ИЗЛУЧЕНИЯ В ЗРИТЕЛЬНОЙ СИСТЕМЕ ЧЕЛОВЕКА

В. Е. Анциперов¹, В. А. Кершнер¹, Р.А. Ефимов²

¹ Институт радиотехники и электроники им. В. А. Котельникова РАН,
125009, Москва, ул. Моховая, 11, корп. 7

² Российский университет транспорта (МИИТ),
127994, Москва, ул. Образцова, 9, стр. 9

Статья поступила в редакцию 29 сентября 2021 г.

Аннотация. В данной работе приводятся результаты исследования вопросов наиболее адекватного в смысле моделирования механизмов формирования зрительной системой человека входных видеоданных. Адекватность здесь понимается как максимальное соответствие способов представления данных регистрации падающего излучения материальными детекторами способам кодирования данных в сетчатке зрительной системы. В этой связи в работе обсуждаются общие статистические вопросы формирования (фото) отсчетов и на их основе формализуется концепция идеального устройства формирования (идеальных) изображений. Обсуждаются возникающие на практике проблемы при работе непосредственно с “идеальными” изображениями и предлагается способ их редукции к выборкам отсчетов фиксированного (контролируемого) размера, которые, собственно, и составляют представление данных регистрации. Приведены результаты иллюстративных вычислительных экспериментов по выборочному кодированию обычных цифровых изображений, заданных пиксельными данными. В частности, приведены примеры сгенерированных для заданного цифрового изображения выборок отсчетов разных размеров. На основе приведенных результатов обсуждается зависимость характеристик отсчетных представлений от размеров выборки.

Ключевые слова: кодирование/представление изображения, датчик изображения на основе однофотонного лавинного диода (spad), фотоотсчеты, идеальное устройство формирования изображения, концепция идеального изображения, представление выборки изображения, цифровая (dsp) камера.

Abstract. The paper presents the results of a study of input video data adequate formation/coding in modern imaging systems. Adequacy is understood here as the maximal correspondence between the ways of the radiation registration by material detectors and the ways of data coding in the retina of the human visual system. In this connection, the paper discusses general statistical issues of (photo) counts photoelectric detection and, on this basis, formalizes the concept of an ideal image formation by (ideal) visualization device. The problems arising in practice when working directly with ideal images are discussed and a method of their reduction to count sample of fixed (controllable) size, which, in fact, constitute the representation (coding) of registered data, is proposed. Results of illustrative computational experiments on count coding of the common digital images given by pixel data are presented. Examples of count samples of different sizes generated for the tested digital image are given. Based on the given results, the dependence of characteristics of sampling representations on the parameter of sample size is discussed.

Key words: image coding/representation, spad image sensor, photocounts, ideal imaging device, ideal image concept, image sample representation, dsp camera.

Введение.

Использование механизмов зрительного восприятия в искусственных системах формирования изображений.

В зрительной системе человека регистрация приходящего извне оптического излучения осуществляется с помощью клеток-фоторецепторов, составляющих наружный слой сетчатки, см. рис. 1. Именно фоторецепторы сетчатки (палочки, колбочки) ответственны за сбор первичных данных, сигналов из окружающего нас мира. На основе этих данных в зрительной (нервной)

системе происходит последующая интерпретация, распознавание, наделение семантическим смыслом информации об окружающей среде. Наличие же полезной информации в регистрируемом излучении и возможность ее правильной интерпретации обусловлены тем, что большинство окружающих нас объектов, даже если они сами не являются источниками излучения, способны отражать свет других (искусственных/естественных) источников в соответствии с формой, структурой, химическим составом и т. д. их поверхности.

Принципы преобразования и регистрации оптического излучения в зрительной системе использовались уже в конструкциях первых фотокамер. В процессе эволюции от простейших устройств на основе камеры-обскуры и вплоть до современных цифровых фотоаппаратов, число заимствованных принципов работы с излучением только росло. Отметим, в частности, использование диафрагмы (зрачка), оптической фокусирующей системы (хрусталика) и регистрирующей излучение чувствительной поверхности в виде прозрачной пластинки с нанесённой на неё фотоэмульсией, либо в виде матрицы CCD/CMOS фотодиодов (фоторецепторов). То, что фоточувствительная поверхность глаза (сетчатка) имеет форму внутренней поверхности пустотелого шара, а не плоскости не является в данном случае принципиальным.

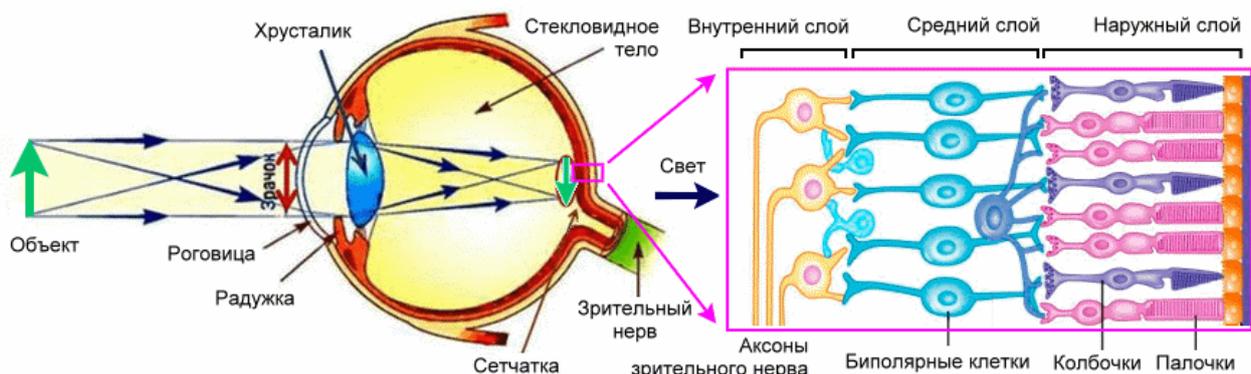


Рис. 1. Схема строения глаза человека и сетчатки.

Эволюция применения зрительных механизмов в искусственных системах формирования изображений наиболее отчетливо проявилась в вопросах регистрации излучения. Действительно, начиная от разработанных Дагером в первой половине 19-го века медных, покрытых серебром пластинок и их

усовершенствования в 20-м веке до листовых и рулонных целлулоидных фотоплёнок с желатиново-серебряной эмульсией (аналоговая фотография), к началу 21 века регистрирующие излучение элементы трансформировались в матрицы фотодиодов - цифро-аналоговые интегральные микросхемы (цифровая фотография). Этот прогресс во многом обязан проявлению в конце 1960-х годов приборов с зарядовой связью (CCD) и последующему изобретению в начале 1990-х годов светочувствительных матриц на комплементарных МОП структурах (CMOS). В практическом плане переход на новые (цифровые) технологии регистрации излучения обеспечил возможность улучшения целого ряда характеристик фотокамер, среди которых важнейшей является повышение пространственного разрешения. Этого удалось достичь за счет последовательного уменьшения размеров фотодиодов (фотодетекторов) матриц вплоть до нескольких микрон. Среди других достижений следует упомянуть увеличение частоты кадров, повышение динамического диапазона, уменьшение потребления энергии и т. д. Отметим, что перечисленные технологические возможности во многом обязаны развитию теоретической физики 20-го века в области взаимодействия излучения с веществом (квантовой электродинамики) [1] и соответствующих исследований в области полупроводниковых структур (в теории твердого тела) [2].

Согласно современным представлениям [1,2], по мере прогрессирующего уменьшения размеров фотодетекторов цифровых матриц, характер регистрации излучения с необходимостью приобретает все более выраженный квантовый характер и в пределе переходит в регистрацию отдельных фотонов (фотоэлектронов). Замечательным обстоятельством является то, что этот предельный случай на сегодняшний день уже достигнут рядом технологий. Речь идет о технологиях изготовления так называемых фотоотсчётных датчиков (Photon-Counting Image Sensors) [3] - датчиков работающих в режиме счета фотонов. В качестве примера укажем на электронно-умножительные матрицы с зарядовой связью (EMCCD) [4], однофотонные лавинные диоды (SPAD) [5], лавинные фотодиоды в режиме счётчика Гейгера (GMAPD) [6], см. рис. 2.

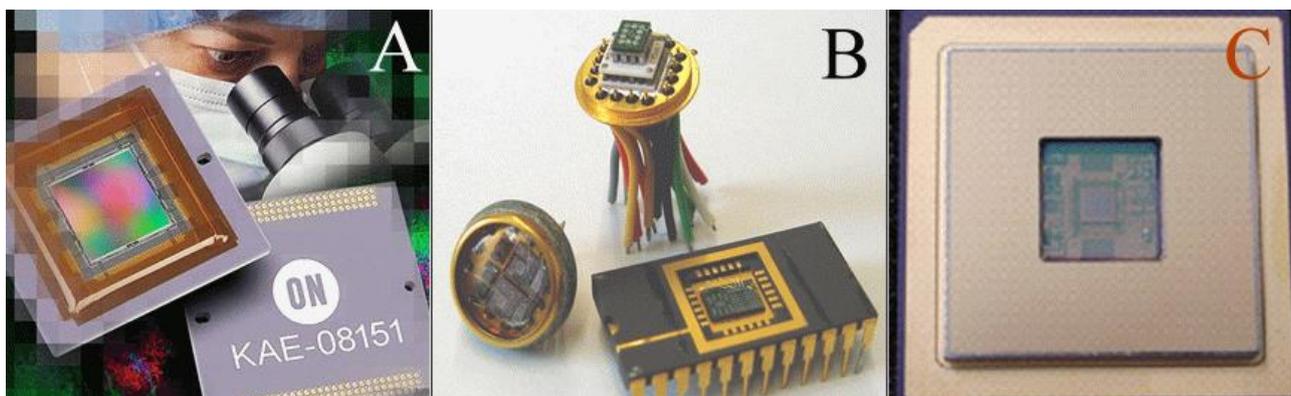


Рис. 2. Современные фотоотсчетные датчики регистрации излучения А) EMCCD, В) SPAD, С) GMAPD.

Поскольку результатом регистрации излучения цифровыми матрицами является электрический ток фотодетекторов, соответствующие интегральные микросхемы легко могут быть включены в различные электронные схемы, содержащие помимо управляющих элементов также микропроцессоры, снабженные, например, приложениями для цифровой обработки сигналов (ЦОС). Учитывая высокую производительность современных микропроцессоров, можно по мере регистрации, в онлайн режиме выполнять не только стандартные задачи предобработки видеосигналов, такие как линейаризация, компенсация темнового тока, компенсация бликов и баланс белого, но и решать гораздо более интеллектуальные проблемы типа классификации изображений, распознавания на них объектов, анализа сцен и т. д. Другими словами, преобразование на выходе светочувствительной матрицы видеосигнала в цифровую область открывает совершенно новые возможности в вопросах формирования и анализа изображений на основе методов ЦОС.



Рис. 3. Приближенная схема методов и средств современной цифровой обработки изображений (ЦОС)

Современная теория обработки цифровых изображений [7], в свою очередь, предоставляет цифровым камерам обширный арсенал методов и средств обработки изображений, начиная от простейших, линейных методов, до сложных алгоритмических процедур машинного зрения. Многообразие методов ЦОС обычно условно подразделяется на три основных уровня - компьютеризованные процессы низкого, среднего и высокого уровня. Процессы низкого уровня обычно включают операции предобработки изображений с целью цветовой коррекции, сглаживания для устранения шумов, или, наоборот, повышения контраста для улучшения качества и т. д. Обработка изображений на среднем уровне охватывает такие задачи, как классификация, сегментация, распознавание объектов и сцен, сжатие в удобную для последующей компьютерной обработки форму и пр., см. рис. 3. Наконец, высокоуровневая обработка включает в себя семантический анализ распознанных объектов, сцен и сюжетов, т. е. в перспективе, осуществляет “осмысление” захваченной изображением действительности, ассоциируемое со зрением.

Отметим, что в методах ЦОС всех трех уровней также широко используются отдельные механизмы зрительной системы человека [8]. Методы низкого уровня при этом эксплуатируют механизмы предварительной обработки изображений, представленные в периферии зрительной системы - в сетчатке глаза, а для среднего и высокого уровней используются модели

нейробиологических механизмов обработки визуальной информации в коре головного мозга [9]. Здесь следует отметить в первую очередь методы машинного обучения [10], такие, как классификация изображений, морфологический анализ, распознавание объектов и т. д.

Безусловно, имитация механизмов зрительного восприятия осуществляется в рамках ЦОС с той или иной степенью приближения. Приближенное моделирование начинается уже на уровне представления данных регистрации. Действительно, большинство традиционных методов ЦОС ориентировано на растровые (bitmap) изображения, представляющие собой дискретные пикселы - оцифрованный результат накопленной за время экспозиции детекторами матрицы энергии регистрируемого излучения. Другими словами, значения пикселов пропорциональны полному числу фотонов, падающих на чувствительные площадки соответствующих детекторов. Наоборот, в сетчатке глаза фоторецепторы реагируют скорее на отдельные фотоны излучения и передают сигнал об их регистрации в ганглиозные клетки внутреннего слоя сразу же, без накопления [9]. Отмеченное отличие в представлении входных данных в искусственных и естественных системах формирования изображений объясняется тем, что до недавнего времени просто отсутствовали технологические возможности реализации методов регистрации изображений в зрительной системе. Однако, как отмечено выше, в последние десятилетия достигнут значительный прогресс в разработке цифровых матриц, работающих в режиме счета фотонов (см. также [11]). В этой связи естественным представляется желание адаптировать процедуры ЦОС к вновь открывшимся возможностям по моделированию видеоданных регистрации ближе имитирующих входные данные зрительной системы, формируемые сетчаткой. Один из возможных подходов к решению этой проблемы -новый метод представления (кодирования) изображений выборками отсчетов, моделирующих поток зарегистрированных фотонов излучения, представлен ниже.

1. Кодирование (представление) изображений выборками отсчетоведение.

Для обоснования предлагаемого ниже способа кодирования (представления) видеоданных предварительно обсудим используемые модельные представления о механизмах регистрации излучения сетчаткой глаза человека. Эти представления мотивированы известными биофизическими фактами о механизмах взаимодействия светового излучения с рецепторами сетчатки человеческого глаза [9] и обоснованы положениями квантовой электродинамики [1] в ее полуклассическом приближении [12]. Далее в работе эти представления формализуются в виде модели идеального изображения [13], регистрируемого идеальным устройством формирования изображений, содержащим большой массив точечных детекторов [14].

Прототипом модели идеального устройства формирования изображений в зрительной системе человека является сетчатка глаза. В качестве мотивации представленной ниже модели перечислим несколько касающихся ее основных фактов. Сетчатка человека содержит около 10 миллионов колбочек и 100 миллионов палочек, способных регистрировать отдельные фотоны видимого излучения. Их плотность достигает величин 100000 - 160000 рецепторов на мм² (у хищных птиц, которым приходится выискивать малые объекты на достаточно большом расстоянии, на квадратный миллиметр сетчатки приходится около миллиона рецепторов). Попутно заметим, что сигналы, приходящие в мозг по зрительному нерву (нервные импульсы), строго говоря не являются непосредственно данными регистрации излучения рецепторами, а формируются на их основе сложной системой клеток среднего и внутреннего слоев сетчатки, см. рис. 1. Итого, видеоинформация, поступающая в мозг, передается по аксонам ганглиозных клеток (зрительным нервам), число которых на ~ два порядка меньше (около миллиона) от числа рецепторов.

В качестве прототипа идеального устройства формирования изображений искусственного происхождения, который можно рассматривать и как техническую реализацию последнего, можно использовать матрицу однофотонных лавинных диодов (SPAD) [11], либо какой-либо из ее аналогов

[3]. В настоящее время достигнуты серьезные результаты в производстве таких матриц: на участке $11 \times 11 \text{ мм}^2$ с шагом в $9,4 \text{ мкм}$ размещается $1024 \times 1000 = 1\,024\,000$ микро-детекторов светового излучения (пикселей). Каждый элемент имеет динамическую бинарную память для хранения события регистрации фотоотсчета, обновление памяти осуществляется с частотой $24\,000$ кадра в секунду [11]. Соответственно, поток информации от SPAD-матрицы может достигать 25 Гбит/сек , что существенно превосходит информационный поток, полученный с помощью зрительной системы человека, составляющей около 50 Мбит/сек [15].

При всех отличиях, приведенных выше прототипов, они имеют ряд общих черт. Оба прототипа содержат имеющую конечные размеры фоточувствительную область, в которой находится очень большое число детекторов/фоторецепторов. Каждый фотодетектор способен регистрировать отдельные фотоны падающего на него излучения. Все события регистрации фотонов детекторами, хранятся прототипами в течение некоторого времени. Перечисленные свойства можно положить в основу достаточно общей концепции идеального устройства формирования изображений, являющегося формальным обобщением, как представленных прототипов, так и ряда других систем формирования изображений (фотопластинок, фотоплёнок с желатиново-серебряной эмульсией и пр.).

В этой связи формулируем следующее определение. Под идеальным устройством формирования изображений будем понимать некоторую двумерную область Ω с координатами $\vec{x} = (x_1, x_2)$ площади S , на которой вплотную друг к другу расположены точечные детекторы [14] регистрируемого излучения с исчезающе малой площадью чувствительной поверхности ds . Соответственно, общее число детекторов в области равно $N = S/ds$. В предположении $ds \rightarrow 0$ число N предполагается сколь угодно большим, в пределе $N \rightarrow \infty$. Таким образом, формально, идеальное устройство формирования изображений Ω представляет из себя практически

“непрерывную” чувствительную поверхность, ее точки с координатами \vec{x} задают положения идеальных “точечных” детекторов.

Когда на чувствительную поверхность Ω падает (стационарное в течение времени экспозиции) излучение интенсивности $I(\vec{x})$, $\vec{x} \in \Omega$, некоторые точечные детекторы идеального устройства регистрируют отдельные фотоны. Событие регистрации в течение времени кадра T некоторым детектором падающего на него фотона будем называть (фото) отсчетом и приписывать этому отсчету координаты \vec{x} детектора. В рамках полуклассической теории взаимодействия излучения с веществом отсчеты является случайными событиями и в пределе $ds \rightarrow 0$ задаются вероятностями $P(\vec{x}) = \alpha TI(\vec{x})ds$, где $\alpha = \eta(h\nu)^{-1}$, $h\nu$ средняя энергия фотона (h - постоянная Планка, ν - характерная частота излучения), безразмерный коэффициент $\eta < 1$ является квантовой эффективностью материала детектора и имеет смысл вероятности отсчета на один фотон (при $w = TI(\vec{x})ds = h\nu$). Тем самым, при регистрации излучения интенсивности $I(\vec{x})$ с каждым точечным детектором $\vec{x} \in \Omega$ можно связать бинарную случайную величину $\sigma \in \{0, 1\}$, принимающую значения $\sigma = 1$ и $\sigma = 0$ в зависимости от того, зарегистрирован или нет этим детектором отсчет. Условное (при заданной интенсивности $I(\vec{x})$) распределение вероятностей величины σ имеет вид распределения Бернулли:

$$\begin{aligned} P(\sigma = 1|\vec{x}) &= P(\vec{x}) = \alpha TI(\vec{x})ds, \\ P(\sigma = 0|\vec{x}) &= 1 - \alpha TI(\vec{x})ds. \end{aligned} \quad (1)$$

Заметим, что, согласно распределению (1), среднее число отсчетов $\bar{\sigma}$ в точке \vec{x} равно $\alpha TI(\vec{x})ds$. Соответственно, интеграл $\bar{n} = \alpha T \iint_{\Omega} I(\vec{x})ds$ задает среднее число всех отсчетов, зарегистрированных на Ω за время T .

Выбирая случайным образом из массива всех N точечных детекторов идеального устройства какой-то один, имеющий координаты \vec{x} (с равномерной вероятностью $Q(\vec{x}) = N^{-1}$, можно, используя (1), найти совместное распределение случайных величин \vec{x} и σ :

$$\begin{aligned} P(\sigma = 1, \vec{x}) &= P(\sigma = 1|\vec{x})Q(\vec{x}) = \frac{\alpha T I(\vec{x}) ds}{N}, \\ P(\sigma = 0, \vec{x}) &= P(\sigma = 0|\vec{x})Q(\vec{x}) = \frac{1 - \alpha T I(\vec{x}) ds}{N}. \end{aligned} \quad (2)$$

Суммируя распределение (2) по всем N точечным детекторам, получим маргинальное безусловное распределение σ , задающее вероятности (частоты) появления/отсутствия отсчетов:

$$\begin{aligned} P(\sigma = 1) &= \frac{\alpha T}{N} \iint_{\Omega} I(\vec{x}) ds = \frac{\bar{n}}{N}, \\ P(\sigma = 0) &= 1 - \frac{\alpha T}{N} \iint_{\Omega} I(\vec{x}) ds. \end{aligned} \quad (3)$$

Окончательно, поделив вероятность $P(\sigma = 1, \vec{x})$ (2) на соответствующую вероятность $P(\sigma = 1)$ (3), получим условную вероятность иметь отдельному отсчету координаты \vec{x} в области Ω :

$$P(\vec{x}|\sigma = 1) = \frac{P(\sigma=1, \vec{x})}{P(\sigma=1)} = \frac{I(\vec{x}) ds}{\iint_{\Omega} I(\vec{x}) ds} = \frac{\bar{\sigma}(\vec{x})}{\bar{n}}, \quad (4)$$

что не следует путать с вероятностью отсчета $P(\sigma = 1|\vec{x})$ (1) в точке \vec{x} .

В дальнейшем удобно вместо условной вероятности координат отсчета (4) использовать соответствующую ей плотность $\rho(\vec{x}|I(\vec{x})) = P(\vec{x}|\sigma = 1)/ds$, в которой явно указано условие распределения отсчета при заданной интенсивности $I(\vec{x})$. Переписывая (4) для плотности $\rho(\vec{x}|I(\vec{x}))$, получим следующий важный результат:

$$\rho(\vec{x}|I(\vec{x})) = \frac{I(\vec{x})}{\iint_{\Omega} I(\vec{x}) ds}, \quad (5)$$

который декларирует, что плотность распределения вероятностей отсчета $\vec{x} \in \Omega$ при регистрации излучения идеальным устройством формирования изображений совпадает с нормированной интенсивностью $I(\vec{x})$ этого излучения на Ω . Отметим в этой связи универсальный характер (5): условная плотность распределения вероятностей не зависит ни от квантовой эффективности материала детекторов η , ни от спектра (в том числе характерной частоты $\bar{\nu}$) излучения, ни от времени фрейма T . Более того, она не зависит и от нормировки $I(\vec{x})$ на полную мощность излучения $W = \iint_{\Omega} I(\vec{x}) ds$, поскольку определяется только формой интенсивности - нормированной версией $I(\vec{x})/W$. От

перечисленных параметров безусловно зависят статистические характеристики совокупности отсчетов, к примеру, их среднее число $\bar{n} = \alpha TW$, а, следовательно, и распределение P_n их общего числа n . Однако, подчеркнем это еще раз, для плотности распределения вероятностей отсчета (5) достаточной статистикой является только и только нормированная интенсивность $I(\vec{x})/W$.

На основе сформулированной концепции идеального устройства формирования изображений теперь может быть дано определение и собственно модели идеального изображения. Под идеальным изображением понимается (упорядоченный) набор $X = (\vec{x}_1, \dots, \vec{x}_n)$, $\vec{x}_i \in \Omega$ всех n случайных отсчетов, регистрируемых идеальным устройством в течение времени фрейма T . Идеальное изображение является, таким образом, существенно случайным объектом, который следует различать с отдельными его реализациями. Отметим, что случайный характер идеального изображения определяется не только случайными координатами \vec{x}_i с распределениями (5), но и их случайным числом n в наборе X .

Полное статистическое описание идеальных изображений в виде всех конечномерных плотностей распределений $\{\rho(\vec{x}_1, \dots, \vec{x}_n, n|I(\vec{x}))\}$, $\vec{x}_i \in \Omega, n = 0, 1, \dots$ может быть получено в предположении условной независимости отсчетов \vec{x}_i (при условии $I(\vec{x})$ и n). Стандартный вывод этого описания, опирающийся на распределения координат отсчетов (5) и пуассоновскую аппроксимацию (в пределе $ds \rightarrow 0, N \rightarrow \infty, Nds = S = const$) для P_n - распределения вероятностей их общего числа n , можно найти, например, в [16]:

$$\begin{aligned} \rho(\vec{x}_1, \dots, \vec{x}_n, n|I(\vec{x})) &= \rho(\vec{x}_1, \dots, \vec{x}_n, |n, I(\vec{x})) \times P_n(I(\vec{x})) = \\ &= \prod_{i=1}^n \rho(\vec{x}_i|I(\vec{x})) \times \frac{\bar{n}^n}{n!} \exp(-\bar{n}), \quad \bar{n} = \alpha T \iint_{\Omega} I(\vec{x}) ds. \end{aligned} \quad (6)$$

Отметим, что статистическое описание (6) соответствует некоторому двумерному точечному пуассоновскому негомогенному процессу с интенсивностью точек $\lambda(\vec{x}) = \alpha T I(\vec{x})$, а приведенный результат отражает тот известный факт, что при сделанных предположениях совокупность бернулиевых испытаний $\{(\vec{x}, \sigma)\}$ хорошо аппроксимируются пуассоновским процессом [17].

Заметим, что в отличие от распределения вероятностей одного отсчета (5), плотности $\rho(\vec{x}_1, \dots, \vec{x}_n, n | I(\vec{x}))$ (6) через P_n и $\bar{n} = \alpha TW$ зависят от параметров α, T, W , т.е. отчасти они утрачивают свойство универсальности. Тем не менее известно, что если зафиксировать n , то условные плотности $\rho(\vec{x}_1, \dots, \vec{x}_n | n, I(\vec{x}))$ от этих параметров уже не зависят [17]. Последнее обстоятельство часто позволяет упростить анализ разделив связанные с n “энергетические” оценки от “геометрических”, структурных оценок, связанных только с конфигурацией набора координат отсчетов $X_n = (\vec{x}_1, \dots, \vec{x}_n)$ [18].

Модель идеального изображения и его статистическое описание (6) полезны в теоретических исследованиях, например, в вопросах поиска оптимальных процедур анализа изображений, в частности распознавания или идентификации объектов на изображениях [18]. Кроме того, при низких интенсивностях $I(\vec{x})$ регистрируемых излучений (не больших $\bar{n} = \alpha TW$), статистическое описание (6) вполне успешно используется в областях флуоресцентной микроскопии, позитронно-эмиссионной томографии (ПЭТ), однофотонной эмиссионной компьютерной томографии (ОФЭКТ), оптической и инфракрасной астрономии и т. д. [19].

Однако, при обычных интенсивностях излучений, соответствующих, например, дневному свету, практическое использование идеальной модели изображения оказывается проблематичным. Дело в том, что потоки фотонов, падающих на площадку $S \sim 1 \text{ м}^2$ за 1 сек, например, от солнца в обычных условиях огромны - на земной поверхности в ясный день они составляют $\sim 10^{15} - 10^{16}$ фотонов [9]. Для идеальных устройств, допускающих режим счета фотонов, даже если они формируют один отсчет на ~ 10 фотонов (при квантовой эффективности $\eta = 0.1$), число отсчетов за секунду составит величину $\bar{n} \sim 10^{15}$ (1000000 Гбит/сек = 1 Пбит/сек). Очевидно, что работа с такими потоками данных потребует в практических задачах слишком больших ресурсов. Поэтому желательно разработать иные подходы к вопросам кодирования/представления изображений.

Некоторое время назад нами было предложено следующее решение приведенной выше проблемы уменьшения размерности представления идеальных изображений [14]. Зафиксируем с самого начала некоторое приемлемое значение размеров представления $k \ll \bar{n}$ и, рассматривая представление идеального изображения $X = \{\vec{x}_i\}$ как некоторую генеральную совокупность случайных отсчетов, осуществим из нее случайную выборку в k отсчетов $X_k = \{\vec{x}_j\}$. Очевидно, такое “выборочное” представление при размерах $k \ll \bar{n}$, в полном соответствии с подходом классической статистической теории, будет по-прежнему представлять идеальное изображение X . Назовем X_k представлением изображения выборкой случайных отсчетов или, короче, выборочным представлением. Статистическое описание выборочных представлений легко следует из (6) посредством интегрирования $\rho(\vec{x}_1, \dots, \vec{x}_n, n | I(\vec{x}))$ по не выбранным в X_k отсчетам и суммирования результата по числу $l = n - k = 0, 1, \dots$ этих не учитываемых отсчетов:

$$\begin{aligned} \rho(X_k | I(\vec{x})) &= \rho(\vec{x}_1, \dots, \vec{x}_k | n, I(\vec{x})) \times \sum_{l=0}^{\infty} P_{k+l}(I(\vec{x})) = \\ &= \prod_{j=1}^k \rho(\vec{x}_j | I(\vec{x})) \times P_{n \geq k}(I(\vec{x})). \end{aligned} \quad (7)$$

где $P_{n \geq k}(I(\vec{x}))$ обозначает вероятность того, что в идеальном изображении содержится более чем k отсчетов. Учитывая пуассоновский характер вероятностей P_n и их асимптотическое стремление при $\bar{n} \rightarrow \infty$ к гауссовому распределению со средним \bar{n} , несложно показать (см. рис. 4), что в случае $\bar{n} \gg 1$ вероятность $P_{n < k}(I(\vec{x}))$ будет меньше ε , как только $k \leq 2\varepsilon\bar{n}$, соответственно, при этом $P_{n > k}(I(\vec{x}))$ будет отличаться от единицы менее чем на величину ε .

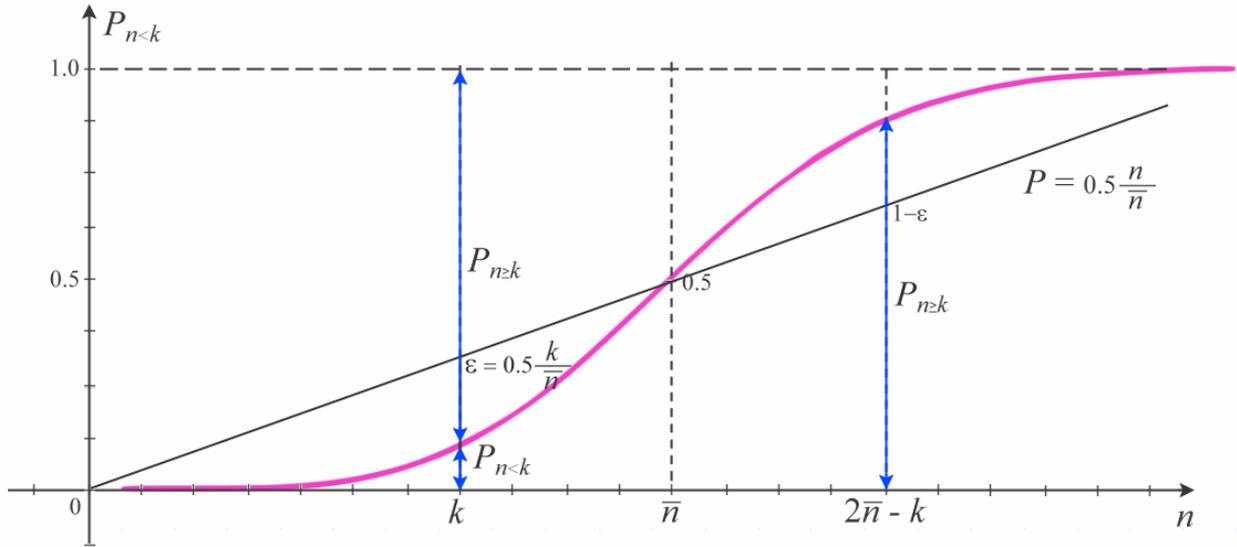


Рис. 4. К выводу оценок $P_{n < k} < \varepsilon$ и $P_{n \geq k} > 1 - \varepsilon$ при $k = 2\varepsilon\bar{n}$.

Считая далее, что для представлений изображений выборками случайных отсчетов X_k их размеры удовлетворяют $k < 2\varepsilon\bar{n}$, из (7), с точностью до произвольно малого ε , полагая $P_{n > k}(I(\vec{x})) \cong 1$, получим:

$$\rho(X_k | I(\vec{x})) = \prod_{j=1}^k \rho(\vec{x}_j | I(\vec{x})). \quad (8)$$

Так, если, к примеру, $\varepsilon = 0.0011$ и $\bar{n} \sim 10^{15}$, получим, что (8) имеет место вплоть до $k \sim 2 \times 10^{12}$, впрочем, и эта оценка является, скорее всего, сильно заниженной.

Заметим, что статистическое описание выборочных представлений (8) может быть получено сразу же из (5) при допущении, что вероятности координат \vec{x}_j любого из отсчетов X_k не зависят от того, какие и сколько отсчетов кроме него содержатся в выборке. Как показано выше, это предположение о независимости отсчетов равносильно предположению, что идеальное изображение практически, наверное, содержит (гораздо) больше отсчетов, чем размер выборочного представления k .

Статистическое описание (8) выборочных представлений $X_k = \{\vec{x}_1, \dots, \vec{x}_k\}$ оказывается значительно удобнее полного статистического описания идеальных изображений (6) не только в силу фиксированного размера представления $k \ll \bar{n}$, но и в силу ряда других обстоятельств. Оно, во-первых, фиксирует условную

независимость и одинаковое условное распределение (iid свойство) всех k отсчетов \vec{x}_j . При этом, во-вторых, плотности распределений отдельных отсчетов $\rho(\vec{x}_j | I(\vec{x}))$ задаются нормированной интенсивностью $I(\vec{x})$ в области Ω . Отсюда, в-третьих, также, как и для условных плотностей $\rho(\vec{x}_1, \dots, \vec{x}_n | n, I(\vec{x}))$ отсчетов, для описания (8) выполняется свойство универсальности - оно не зависит ни от квантовой эффективности материала детекторов η , ни от спектра излучения, ни от времени фрейма T . Перечисленные свойства выборочных представлений (8) обеспечивают подходящий вид задания входных данных для многих хорошо разработанных статистических подходов и методов машинного обучения, включая наивный байесов подход [10].

Последнее замечание касается следующего обстоятельства. Поскольку $\rho(\vec{x}_j | I(\vec{x}))$ (5) не зависит от абсолютных значений интенсивности, а определяется только ее нормированной версией - формой интенсивности $I(\vec{x}) / \iint_{\Omega} I(\vec{x}) ds$, статистическое описание выборочных представлений (кодированных изображений) (8) также не зависит от единиц измерения $I(\vec{x})$. В частности, если интенсивность зарегистрированного излучения задана пикселями $\{n_i\}$ некоторого растрового изображения, полученного в результате оцифровки $I(\vec{x})$ с параметром разрешения цифрового квантования $Q = \Delta I$, то от Q форма описания (8) будет зависеть не непосредственно через его значение, а только через параметр битовой глубины пикселя $v = \log_2(I_{max}/\Delta I)$ (разрядность пикселя) - стандартной характеристики цифровых изображений. Несколько примеров выборочного кодирования цифровых изображений приводятся в следующем разделе.

2. Эксперименты по кодированию цифровых изображений выборками отсчетов.

В связи с последним замечанием отметим, что процедуру выборочного кодирования (представления) цифровых изображений можно по существу свести к нормировке $\pi_i = n_i / \sum n_i$ значений пикселей $n_i \sim I_i$ изображения и

последующему семплированию k случайных отсчетов из полученного распределения вероятностей $\rho(\vec{x}_j | I(\vec{x})) \approx \pi_j$. Отметим при этом, что в области машинного обучения существует большой арсенал методов организации процедур семплирования, объединенных общим названием методов Монте-Карло [20]. Они включают в себя такие известные методы, как семплирование по значимости, семплирование с отклонением, алгоритмы семплирования Метрополиса–Гастингса, Гиббса и т. д. Имея подобный арсенал методов можно с разных позиций оптимизировать процедуру семплирования – с позиций вычислительной эффективности, репрезентативности, специфики задачи и т. д. При этом следует отметить, что для некоторых методов семплирования вообще не требуется даже предварительная нормировка – достаточно того условия, что все пиксели ограничены величиной 2^v , где v – параметр битовой глубины пикселя изображения.

Для примера на рис. 5 приведены представления выборками случайных отсчетов стандартного тестового изображения House, часто используемого в публикациях по обработке изображений [21]. Изображение House изначально задано в формате TIFF, имеет размеры 512×512 пикселей, цветное, с глубиной цвета $v = 24$ бита. Для уменьшения объема вычислений изображение было конвертировано в GIF формат того же размера $s \times s$, $s = 512$, но в серую палитру с глубиной цвета $v = 8$ бит (см. рис. 5А)). Семплирование выборочных представлений размерами $k = 100.000, 500.000, 1.000.000, 2.000.000$ и $5.000.000$ отсчетов (рис. 5 В–F)) осуществлялось одним из самых простых методов – семплированием с отклонением (rejection sampling [20]) при равномерном вспомогательном распределении $g(\vec{x}) = (s \times s)^{-1} = 512^{-2}$ и граничной константе оценки верхней границы правдоподобия $c = 2^v = 256$. Выбор $g(\vec{x})$ и c обусловлен ограничениями $n_j < 2^v$, которое в случае среднего $\bar{m} = \sum n_j / s^2 > 1$ приводят к следующей мажорации плотности распределения отсчетов:

$$\rho(\vec{x}_j | I(\vec{x})) \approx \pi_j = \frac{n_j}{\sum n_j} < \bar{m} \frac{n_j}{\sum n_j} = \frac{n_j}{s \times s} < \frac{2^v}{s \times s} = c g(\vec{x}_j). \quad (9)$$

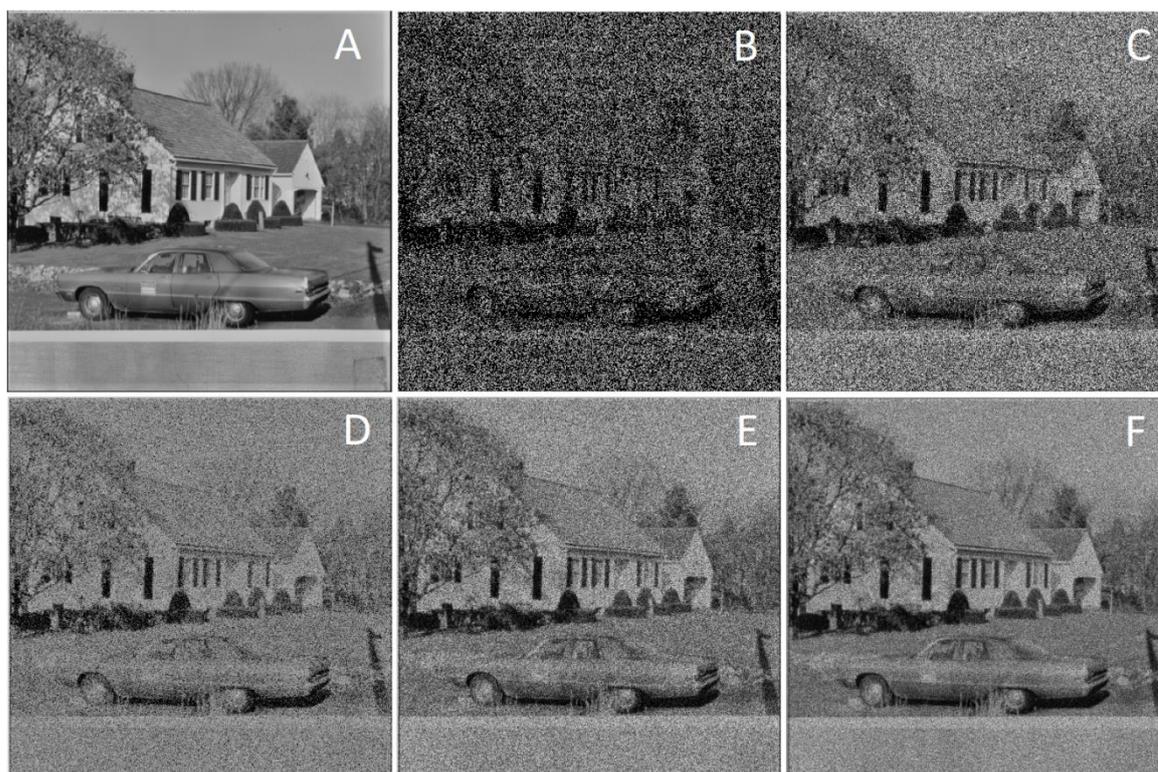


Рис. 5. Представление изображения House [21] выборками случайных отсчетов:
 А) - исходное изображение в формате TIFF, В-Ф) - представления выборки
 размеров, соответственно $k = 100.000$, 500.000 , $1.000.000$, $2.000.000$ и $5.000.000$
 отсчетов.

Отметим, что алгоритмическая реализация этой процедуры семплирования сводится к случайному выбору равномерно распределенных в области Ω изображения (площади $s \times s$, с координатами - числами с плавающей точкой) случайных векторов \vec{x}_j и включения их в выборку отсчетов X_k при выполнении равносильного (9) теста $u_j < n_j$, где j – индекс содержащего \vec{x}_j пикселя, а u_j – реализация равномерно распределенной на $(0, \bar{m}s)$ случайной величины (по этому поводу см. [20]). В данной реализации нормировки пикселей n_j в явном виде не требуется, необходима только предварительная оценка величины среднего пикселя $\bar{m} = \sum n_j / s^2$.

Ввиду простоты формирования выборок случайных отсчетов $X_k = \{\vec{x}_1, \dots, \vec{x}_k\}$, и универсальности их статистического описания (8), связанные с ними представления полезны во многих задачах анализа, классификации, идентификации изображений, см. [14]. Однако, в ряде приложений, связанных, например, с визуальным восприятием, выборочные представления встречаются

с рядом проблем. Действительно, как это видно на фрагментах рис. 5, выборочные представления имеют менее гладкую, более зернистую текстуру, чем, например, обычные растровые представления.

При этом не следует забывать, что традиционные цифровые изображения, представленные в ряде популярных форматов, в том числе взятое для примера растровое TIFF изображение House на рис. 5А, являются результатом достаточно сложной обработки исходных изображений, полученных с помощью цифровых камер или отсканированных фотопластинок/фотопленок, которые обычно не обладают требуемым качеством. В этой связи, по аналогии с задачами подготовки растровых представлений, может быть также поставлена задача повышения качества визуального восприятия и выборочных представлений. Эта весьма важная тема, однако, находится за пределами данной работы. Оставляя ее решение для последующих исследований ограничимся в заключении лишь результатами простейшего подхода к сглаживанию зашумленных изображений на основе метода окна Парзена-Розенблатта.

Сглаживание методом окна Парзена-Розенблатта тесно связано с непараметрическим восстановлением по выборке данных плотности их распределения вероятностей на основе ядерного оценивания последней [22]. Поскольку в нашем случае плотность распределения независимых отсчетов $\rho(\vec{x}|I(\vec{x}))$ в выборочном представлении $X_k = \{\vec{x}_1, \dots, \vec{x}_k\}$ кратна зарегистрированной на растровом изображении интенсивности $I(\vec{x})$ (5), ядерная оценка

$$\hat{\rho}(\vec{x} | X_k) = \frac{1}{kh} \sum_{j=1}^k K\left(\frac{\vec{x}-\vec{x}_j}{kh}\right), \quad (10)$$

восстанавливает как плотность распределения отсчетов, так и с точностью до нормировки интенсивность $I(\vec{x})$ (форму интенсивности).

В (10) ядро $K(\vec{x})$ предполагается неотрицательной, нормированной, симметричной функцией с единичным вторым моментом на двумерной плоскости с координатами $\vec{x} = (x_1, x_2)$. Другими словами, ядро является простейшим сглаживающим окном единичной ширины. Параметр $h > 0$

является параметром сглаживания, он часто называется также шириной окна [22]. В наших экспериментах в качестве ядра использовалось двумерное гауссово распределение $K(\vec{x}) = N(\vec{x} | \vec{0}, E)$, параметр ширины окна не использовался – по умолчанию предполагалось $h = 1$. Размеры изображения выбирались теми же $s \times s$, $s = 512$, что и у исходного изображения, т. е. предполагалось равенство параметра h размеру пикселя, выбранного в качестве масштаба.

Для примера на рис. 6. представлены изображения, полученные из выборок фотоотсчетов (рис. 5), соответствующих выборочным представлениям размерами $k = 100.000, 500.000, 1.000.000, 2.000.000$ и $5.000.000$ отсчетов (рис. 2 В–F). Стоит отметить, что изображение, полученное уже из относительно небольшой выборки отсчетов ($k = 1.000.000$) в целом визуально воспринимается как изображение приемлемого качества, несмотря на наличие шумов и смазанность элементов. При увеличении числа отсчетов сглаживание позволяет получать больше полезной информации, однако также увеличивается и время, затрачиваемое на формирование отсчетов и восстановление изображения.

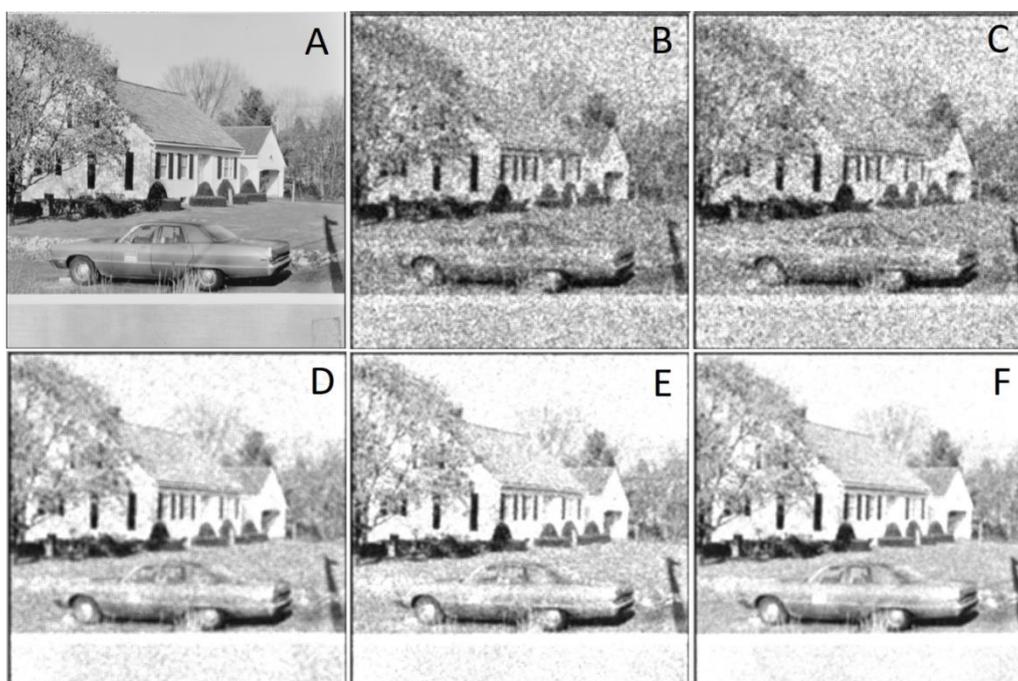


Рис. 6. Восстановление изображения House [21] с помощью выборок случайных отсчетов: А) - исходное изображение в формате TIFF, В–F) - восстановленные изображения по выборкам размеров, соответственно $k = 100.000, 500.000, 1.000.000, 2.000.000$ и $5.000.000$ отсчетов.

Заключение.

В работе было описано представление изображений выборками случайных отсчетов фиксированного размера. При формировании такого «выборочного» представления изображения стоит отметить не только очевидное уменьшение объема исходного изображения, но и другие характеристики, полученные во время работы с данными. В частности стоит отметить время, затраченное на получение выборок разного объема, которое зависит от количества случайных отсчетов, выбранных для формирования изображения. При малых количествах отсчетов время формирования выборок изображения House оказалось незначительным, для $n = 100.000$ оно составило в среднем $t = 0.56$ секунд, при увеличении количества отсчетов в 10 раз (до $n = 1.000.000$) время составило около $t = 3.76$ секунд. Стоит отметить, что представление с количеством отсчетов равным $n = 1.000.000$ уже имеет достаточный набор свойств, позволяющий достаточно точно сопоставить представленные на нем объекты с аналогичными объектами на исходном изображении. А значит, время, необходимое на формирование максимально правдоподобного «выборочного» представления для изображений имеющих схожие исходные характеристики (размер 512×512 пикселей, цветное, с глубиной цвета 24 бит), не будет превышать 5 секунд. При дальнейшем формировании выборок с $n = 2.000.000$ и $n = 5.000.000$ время составило около $t = 7.3$ секунд и $t = 13.9$ секунд соответственно. Представления, полученные из больших выборок имеют незначительные отличия от того, что было получено при $n \sim 10^6$, в основном это проявляется в постепенном устранении шумов и более четком формировании мелких объектов.

Уменьшение объема исходного изображения методом формирования фотоотсчетов, позволило нам эффективно выполнять последующие этапы по его обработке. Полученное фотоотсчетное изображение было с некоторой степенью точности восстановлено до исходного методами сглаживания. Степень гладкости изображения, а также время его формирования из выборки непосредственно зависит от объема используемой выборки. Данный метод моделирования изображения позволил более точно передавать образы на

изображениях, полученных при частотной модуляции с малым числом фотоотсчетов. Отдельно стоит отметить увеличение среднего уровня яркости изображения, возникающее благодаря устранению большого числа пробелов (черных точек) между фотоотсчетами, что также позволяет выделять малозаметные детали на изображении. Также было зафиксировано время, затраченное на сглаживание выборок разных размеров. При малых количествах фотоотсчетов $n = 100.000$, 500.000 и $1.000.000$ время составило около $t = 0.53$, 1.51 и 2.88 секунд соответственно. Для больших выборок $n = 2.000.000$ и $5.000.000$ потребовалось в среднем $t = 5.45$ и 13.24 секунд. Также, как и в случае с формированием фотоотсчетных изображений, наилучшие параметры изображения достигаются при использовании выборок с количеством фотоотсчетов $n \sim 106$, однако стоит также отметить и малые выборки, при обработке которых можно получить понятное для наблюдателя изображение.

Таким образом формирование фотоотсчетных изображений позволяет не только уменьшить объем исходного изображения, но и ускорить все дальнейшие процессы, направленные на его обработку, а качество получаемой на выходе информации будет существенно зависит от количества сформированных фотоотсчетов.

Вышеуказанные процессы, направленные на формирование идеального изображения, дают новые возможности в разработке и совершенствовании разного рода приемников изображения, в частности таких как однофотонные лавинные диоды (SPAD), работающих на прием единичных фотонов.

Финансирование.

Работа выполнена за счет бюджетного финансирования в рамках государственного задания в ИРЭ им. В.А. Котельникова РАН.

Литература

1. Fox M. *Quantum Optics: An Introduction*. New York, Oxford U. Press. 2006. 400 p. <https://doi.org/10.1063/1.2784691>

2. Holst G.C. *CMOS/CCD sensors and camera systems*. Bellingham, SPIE Press. 2011. 408 p. <https://doi.org/10.1117/3.2524677>
3. Fossum E.R., Teranishi N., et al. *Photon-Counting Image Sensors*. Basel, MDPI. 2017. 366 p. <https://doi.org/10.3390/books978-3-03842-375-1>
4. Robbins M. Electron-Multiplying Charge Coupled Devices-EMCCDs. *Single-Photon Imaging*. 2011. P.103-121. https://doi.org/10.1007/978-3-642-18443-7_6
5. Dutton N.A.W., Gyongy I., Parmesan L., et al. A SPAD-based QVGA image sensor for single-photon counting and quanta imaging. *IEEE Transactions on Electron Devices*. V.63. №1. 2016. P.189-196.
<https://doi.org/10.1109/TED.2015.2464682>
6. Aull B.F., Schuette D.R., Young D.J., et al. A study of crosstalk in a 256x256 photon counting imager based on silicon Geiger-mode avalanche photodiodes. *IEEE Sensors Journal*. V.15. №4. 2015. P.2123-2132.
<https://doi.org/10.1109/JSEN.2014.2368456>
7. Gonzalez R.C., Woods R.E. *Digital Image Processing (3rd Edition)*. Prentice Hall Inc. 2007. 976 p.
8. Gabriel C.G., Perrinet L., et al. *Biologically Inspired Computer Vision: Fundamentals and Applications*. Weinheim, Wiley-VCH. 2015. 480 p.
<http://doi.org/10.1002/9783527680863>
9. Rodieck, R.W. *The First Steps in Seeing*. Massachusetts, Sinauer Associates. 1998. 568 p.
10. Barber D. *Bayesian Reasoning and Machine Learning*. Cambridge, Cambridge Univ. Press. 2012.
11. Morimoto K., Ardelean A., et al. Megapixel time-gated SPAD image sensor for 2D and 3D imaging applications. *Optica*. 2020. V.7. №4. P.346-354.
<https://doi.org/10.1364/optica.386574>
12. Goodman J.W. *Statistical Optics (2nd. Edition)*. New York, Wiley. 2015. 544 p.

13. Pal N.R., Pal S.K. Image model, poisson distribution and object extraction. *International Journal of Pattern Recognition and Artificial Intelligence*. 1991. V.5. №3. P.459-483. <https://doi.org/10.1142/S0218001491000260>
14. Antsiperov V. Maximum Similarity Method for Image Mining. *Proceedings of the Pattern Recognition ICPR International Workshops and Challenges*. 2021. V.12665. P.301-313. https://doi.org/10.1007/978-3-030-68821-9_28
15. Koch K., McLean J., et al. How much the eye tells the brain? *Current biology*. 2006. V.16. №14. P.1428-1434. <https://doi.org/10.1016/j.cub.2006.05.056>
16. Streit R.L. *Poisson Point Processes: Imaging, Tracking and Sensing*. New York, Springer. 2010. 273 p. <https://doi.org/10.1007/978-1-4419-6923-1>
17. Gallager R. *Stochastic Processes: Theory for Applications*. Cambridge, Cambridge University Press. 2013. 553 p.
<https://doi.org/10.1017/CBO9781139626514>
18. Antsiperov V. *Machine Learning Approach to the Synthesis of Identification Procedures for Modern Photon-Counting Sensors. Proceedings of the 8th International Conference on Pattern Recognition Applications and Methods ICPRAM*. 2019. P.814-821. <https://doi.org/10.5220/0007579208140821>
19. Bertero M., Boccacci P., Desidera G., Vicidomini G. Image deblurring with Poisson data: from cells to galaxies. *Inverse Problems*. 2009. V.25. №12. P.123006. <https://doi.org/10.1088/0266-5611/25/12/123006>
20. Robert C.P., Casella G. *Monte Carlo Statistical Methods (2-nd edition)*. New York, Springer-Verlag, 2004. 649 p. <https://doi.org/10.1007/978-1-4757-4145-2>
21. The USC-SIPI Image Database. *University of Southern California*. 2021/08/21. <http://sipi.usc.edu/database/>
22. Silverman B.W. *Density Estimation for Statistics and Data Analysis*. London, Chapman & Hall/CRC. 1998. <https://doi.org/10.1201/9781315140919>

Для цитирования:

Анциперов В.Е., Кершнер В.А., Ефимов Р.А. Кодирование изображений выборками отсчетов, мотивированное механизмами регистрации излучения в зрительной системе человека. *Журнал радиоэлектроники* [электронный журнал]. 2021. №9. <https://doi.org/10.30898/1684-1719.2021.9.2>