Об оценке локальной энтропии изображения в скользящем окне

В.В. Сергеев [1,2], А.Ю. Баврина [1,3], И.Д. Зайцев [1], М.Ю. Лазутов [1], Д.А. Шапиро [1]
[1] Самарский национальный исследовательский университет имени академика С.П. Королёва, 443086, Россия, г. Самара, Московское шоссе, д. 34;
[2] Институт систем обработки изображений, НИЦ «Курчатовский институт», 443001, Россия, г. Самара, ул. Молодогвардейская, д. 151;
[3] АО «Самара-Информспутник», 443080, Россия, г. Самара, пр-т Карла Маркса, д. 192, оф. 717

Аннотация

В статье рассмотрен ряд вопросов, связанных с вычислением локальной энтропии в скользящем окне обработки изображения. Предложен новый алгоритм оценки энтропии, обладающий в несколько раз меньшей вычислительной сложностью по сравнению с известными. Алгоритм основан на приближенном представлении гистограммы в виде усеченного разложения в ряд по некоторой системе ортогональных базисных функций, рекурсивного вычисления коэффициентов этого разложения в скользящем окне и последующего пересчета коэффициентов в оценку энтропии. Рассмотрены вопросы выбора ортогонального базиса для представления локальной гистограммы в виде усеченного ряда, показана целесообразность использования базиса Хаара. Описана методика построения иерархического аппроксиматора, осуществляющего быстрый пересчет коэффициентов разложения гистограммы в значение энтропии. Теоретические положения статьи верифицированы экспериментами на изображениях дистанционного зондирования Земли.

Ключевые слова: обработка изображений, формирование локальных признаков, скользящее окно, локальная гистограмма, оценка энтропии, разложение по базису, рекурсивная обработка, иерархическая аппроксимация, погрешность оценки, вычислительная сложность.

Цитирование: **Сергеев, В.В.** Об оценке локальной энтропии изображения в скользящем окне / В.В. Сергеев, А.Ю. Баврина, И.Д. Зайцев, М.Ю. Лазутов, Д.А. Шапиро // Компьютерная оптика. – 2024. – Т. 48, № 5. – С. 714-725. – DOI: 10.18287/2412-6179-CO-1509.

Citation: Sergeyev VV, Bavrina AY, Zaitsev ID, Lazutov MY, Shapiro DA. On estimating the local entropy of an image in a sliding window. Computer Optics 2024; 48(5): 714-725. DOI: 10.18287/2412-6179-CO-1509.

Введение

В практике компьютерной обработки изображений широкое распространение получили операции поэлементной обработки – от простой пороговой обработки, контрастирования или препарирования [1 – 4] до тематической классификации пикселов [5, 6]. В последнем случае, как правило, обрабатываются многокомпонентные (цветные, многоспектральные) изображения, каждый пиксел которых интерпретируется как вектор признаков, поступающий в процедуру классификации и в итоге преобразующийся в номер класса, к которому он относится.

Поскольку отдельные пикселы не могут описывать пространственные связи на изображении, результаты поэлементной классификации, как правило, содержат много ошибок. Для повышения качества классификации принято дополнять изображения синтетическими компонентами – полями различных признаков, учитывающих статистические и/или структурные особенности изображения в окрестности обрабатываемого пиксела [1, 5, 6]. Технология классификации при этом остается поэлементной, однако теперь работает с расширенным вектором характеристик, описывающих пиксел.

Каждое поле признаков имеет те же размеры (в пикселах), что и исходное изображение. Обычно оно формируется в результате локальной обработки одной из компонент изображения по известной схеме «скользящего окна» [1, 3].

Среди разнообразных локальных признаков изображений можно выделить класс гистограммных признаков, основанных на вычислении и использовании эмпирической оценки распределения вероятности (гистограммы) значений пикселов в окне обработки [1, 2, 4, 5]. Подмножество этого класса составляют признаки, вычисляемые по одномерной гистограмме, среди которых важное место занимает локальная энтропия изображения [2, 7, 8] – признак, являющийся предметом рассмотрения настоящей статьи.

Основная проблема, возникающая при формировании гистограммных признаков как синтетических компонент изображения, состоит в высокой вычислительной сложности соответствующих процедур. Действительно, для каждого положения локального окна необходимо построить гистограмму значений яркости в этом окне, а затем по гистограмме вычислить агрегирующую, зачастую нелинейную, характеристику (в нашем случае – энтропию).

Прямое построение гистограммы имеет сложность, пропорциональную площади (числу пикселов) окна, что неприемлемо при работе с окнами достаточно больших размеров. Использование рекурсивных методов пересчета гистограммы при скольжении двумерного прямоугольного окна вдоль одной из пространственных координат изображения [1, 7, 9, 10] позволяет снизить вычислительные затраты до уровня, пропорционального стороне окна, но все равно такого ускорения может оказаться недостаточно для практики. Более перспективным с точки зрения экономии вычислений представляется подход, согласно которому гистограмма заранее представляется в сглаженной форме, в виде суммы нескольких слагаемых – усеченного разложения в ряд по какой-либо системе базисных функций, при этом задача построения гистограммы заменяется на задачу нахождения небольшого числа ненулевых коэффициентов этого разложения [1, 11, 12]. Именно такой подход мы развиваем в настоящей статье.

После того как гистограмма тем или иным способом получена, необходимо выполнить еще одну операцию – преобразовать ее в оценку энтропии. Данное преобразование также вычислительно весьма затратно: даже если заранее протабулировать значения логарифма, входящего в формулу энтропии (см. ниже), неизбежным представляется суммирование большого (равного числу ячеек гистограммы) слагаемых в этой формуле. В работе [8] предпринята попытка заменить вычисление локальной энтропии на ее оценку нейросетью, настраиваемой с помощью машинного обучения. Но, на наш взгляд, достигаемое при этом ускорение оказывается незначительным, а сама процедура – недостаточно универсальной: нейросеть приходится заново настраивать для каждого размера окна и для каждого обучающего набора изображений.

Исходя из сказанного выше, можно заключить, что задача быстрой оценки локальной энтропии изображения в скользящем окне не утратила своей актуальности. В нашей статье мы предлагаем один из вариантов ее эффективного решения.

Потенциал снижения вычислительной сложности алгоритма, решающего поставленную задачу, состоит в переходе от точного вычисления энтропии к ее приближенной оценке. Поэтому разработку нового алгоритма, выбор его параметров нельзя осуществлять в отрыве от анализа погрешности такой оценки, которая, безусловно, должна лежать в некоторых приемлемых пределах. Анализ точности формируемой оценки энтропии, выработка рекомендаций по рациональному распределению допустимой погрешности по этапам осуществляемых преобразований также является предметом рассмотрения в данной статье.

1. Задача вычисления энтропии изображения

Введем основные обозначения и термины, необходимые для дальнейшего изложения, сформулируем принятые ограничения и поставим задачу вычисления поля локальной энтропии изображения.

Пусть $\{x(n_1, n_2)\}$ – матрица пикселов однокомпонентного (полутонового) изображения, n_1, n_2 – их целочисленные индексы. И пусть значения пикселов – целые числа, лежащие в интервале $[0, L-1]$. В практике обработки изображений чаще всего пикселы представляются в байтовом формате ($L=256$).

Считаем пикселы изображения случайными величинами, которые принимают значения в соответствии с некоторым распределением вероятностей $P(x)$, обладающим очевидными свойствами [13]:

$$P(x) \geq 0 \text{ при } 0 \leq x \leq L-1, \sum_{x=0}^{L-1} P(x) = 1 \tag{1}$$

(здесь и ниже индексы можно опустить). Как известно, энтропия (Шеннона) такого распределения вычисляется по формуле [14]:

$$H = -\sum_{x=0}^{L-1} P(x) \log_2 P(x). \tag{2}$$

Экспериментальная оценка распределения вероятности случайной величины осуществляется через гистограмму [13]. В нашем случае случайная величина x – целочисленная, поэтому достаточно просто подсчитать, сколько раз она принимает каждое значение:

$$P(x) \approx \frac{1}{\|D\|} N(x),\ 0 \leq x \leq L-1, \tag{3}$$

где $N(x)$ – число пикселов, равных x, в области D, в котором анализируется изображение, $\|D\|$ – размер этой области в пикселах. Знак приближенного равенства в соотношении (3) обусловлен тем, что область D на практике всегда ограничена, поэтому, строго говоря, называть вероятностью получаемый результат нельзя, это всего лишь относительная частота реализации значений случайной величины в конкретном эксперименте. Иногда на практике удобнее использовать не относительную частоту, а непосредственно массив целочисленных значений $N(x)$, обычно он называется ненормированной гистограммой.

Область D может соответствовать всему анализируемому изображению, тогда оценка (3) будет характеризовать изображение в целом, являться глобальной гистограммой. Или она может соответствовать некоторому относительно небольшому окну обработки, тогда и область, и гистограмму принято называть локальными. Мы рассматриваем именно второй случай.

Итак, «базовый» вариант решения поставленной задачи, с которым мы будем сравнивать наш алгоритм, выглядит следующим образом. По изображению «скользит» окно обработки определенного (назначаемого нами как параметр) размера. Каждое положение окна задает область D, в пределах которой по формуле (3) вычисляется локальная гистограмма, а затем по формуле (2) рассчитывается энтропия, значение которой относится к центру окна. В результате мы получаем искомую синтетическую компоненту изображения – двумерное поле локальной энтропии $H(n_1, n_2)$. Иллюстрацией к сказанному служит рис. 1.

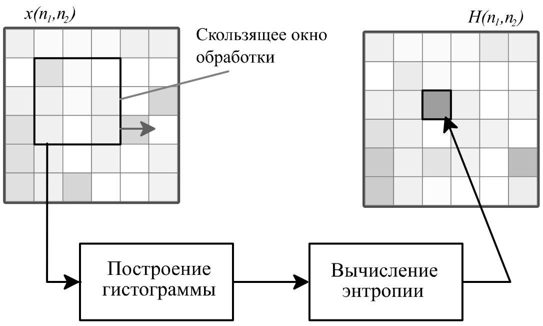


Рис. 1. Базовый алгоритм вычисления поля локальной энтропии

В рамках данной статьи будем считать размеры изображения достаточно большими, чтобы не рассматривать особенности описанной вычислительной процедуры в ситуации, когда скользящее окно, как правило, имеющее относительно малый размер, одной из своих сторон выходит за край изображения. Для обхода указанной коллизии существуют простые стандартные решения [3].

2. Предлагаемый алгоритм вычисления локальной энтропии изображения

Известно, что распределение вероятностей $P(x)$ может быть представлено (и оценено эмпирически) в виде разложения по некоторой системе базисных функций [11]. В нашем случае для такого представления можно использовать одну из известных систем ортогональных базисных функций дискретного аргумента [3, 15–17], заданных на конечном интервале определения функции $P(x)$.

Очевидно, полное описание распределения $P(x)$, соответствующего условиям (1), обеспечивают L базисных функций: $\{f_k(x)\}_{k=0}^{L-1}$. Коэффициенты разложения по ним задаются выражением:

$$F_k = \sum_{x=0}^{L-1} f_k(x) P(x),\ 0 \le k \le L-1. \tag{4}$$

Поскольку базис ортогонален, исходная функция $P(x)$ может быть однозначно восстановлена непосредственно по коэффициентам разложения:

$$P(x) = Q \sum_{k=0}^{L-1} F_k f_k(x), \tag{5}$$

где Q – некоторый коэффициент, зависящий от конкретного базиса. Но мы, преследуя цель экономии вычислений, усечем ряд (5) и будем аппроксимировать $P(x)$ суммой небольшого числа базисных функций:

$$P(x) \approx Q \sum_{k=0}^{K-1} F_k f_k(x), \tag{6}$$

где $K << L$ – число используемых коэффициентов разложения. После представления плотности распределения вероятностей в виде (6) открывается возможность вычисления энтропии не как функции L значений гистограммы (см. формулу (2)), а как функции K коэффициентов ее разложения:

$$H = H(F_0, F_1, ..., F_{K-1}). \tag{7}$$

Такое вычисление даст заведомо приближенный результат оценки энтропии: погрешность возникает из-за усечения ряда при переходе от (5) к (6). Поэтому мы можем допустить и некоторую дополнительную погрешность, не вычисляя сложную многомерную функцию (7) точно, а используя какую-либо ее простую аппроксимацию.

Отметим, что преобразования, описываемые соотношениями (4) – (7), должны выполняться для каждого положения окна обработки изображения. Также отметим, что формула (4), корректная математически, непригодна для практического применения. Мы располагаем не самой функцией $P(x)$, а лишь ее эмпирической оценкой (3), получаемой по окну обработки, поэтому на практике следует использовать альтернативную формулу:

$$F_k(n_1, n_2) = \frac{1}{\|D\|} \sum_{(m_1, m_2) \in D} f_k[x(n_1 - m_1, n_2 - m_2)], \\ 0 \le k \le K-1, \tag{8}$$

в которой дополнительно учтен скользящий характер окна и восстановлены индексы, задающие пространственное положение пикселов на входе и выходе осуществляемого преобразования. Здесь m_1, m_2 – локальные индексы, задающие смещение входных пикселов относительно центра окна. В формуле (8), в отличие от (4), отсутствует функция распределения вероятностей, она учитывается здесь неявно, через частоту встречаемости конкретного значения пиксела в окне обработки.

Анализируя структуру формулы (8), легко заметить, что каждое слагаемое в ней – результат простого в реализации поэлементного преобразования изображения:

$$y_k(n_1, n_2) = f_k[x(n_1, n_2)], \tag{9}$$

а сама сумма, с учетом коэффициента перед ней, задает усреднение преобразованных пикселов по окну обработки D.

Таким образом, мы получаем быстрый алгоритм приближенного вычисления локальной энтропии изображения, который можно проиллюстрировать рис. 2.

Важно отметить, что режим скользящего окна позволяет выполнить суммирование пикселов в окне рекурсивно, с минимальными вычислительными затратами [3]. Пусть скользящее окно обработки в формуле (8) имеет прямоугольную форму, симметричную относительно своего центра:

$$D: \left|m_1\right| \le (M_1 - 1)/2, \left|m_2\right| \le (M_2 - 1)/2, \tag{10}$$

где M_1, M_2 – размеры (в пикселах) окна обработки по вертикали и горизонтали. Тогда в каждой ветке алгоритма сумма пикселов $S_k(n_1, n_2)$ вычисляется каскадно (последовательно), например, сначала рекурсивным преобразованием двумерного поля пикселов по вертикали (переменной n_1), а затем – по горизонтали (переменной n_2):

$$\begin{aligned} &w_k(n_1, n_2) = w_k(n_1 - 1, n_2) + y_k(n_1 + \frac{M_1 - 1}{2}, n_2) - \\ &- y_k(n_1 - \frac{M_1 + 1}{2}, n_2), \\ &S_i(n_1, n_2) = S_k(n_1, n_2 - 1) + w_k(n_1, n_2 + \frac{M_2 - 1}{2}) - \\ &- w_k(n_1, n_2 - \frac{M_2 + 1}{2}), \end{aligned} \tag{11}$$

где $w_k(n_1, n_2)$ – промежуточный результат преобразования. Далее переход от суммирования к усреднению пикселов осуществляется очевидным образом:

$$F_k(n_1, n_2) = \frac{1}{M_1 M_2} S_k(n_1, n_2). \tag{12}$$

Построение аппроксиматора – блока быстрого получения приближенных значений функции многих переменных (7), используемого на последнем шаге алгоритма, будет описано в параграфе 6.

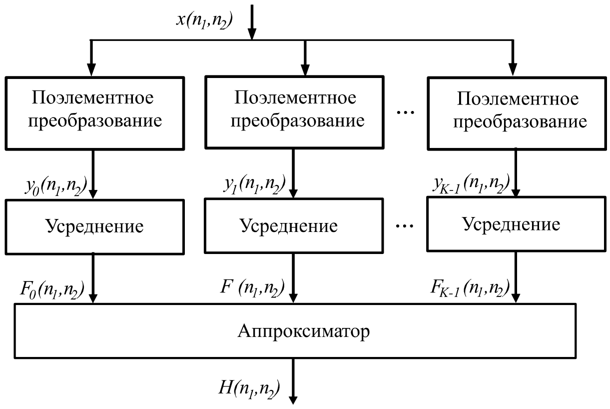
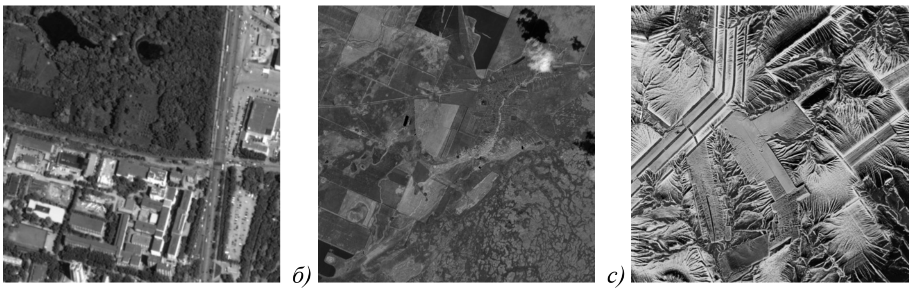


Рис. 2. Алгоритм приближенного вычисления локальной энтропии изображения

3. Тестовые изображения

Для проведения вычислительных экспериментов в качестве тестовых изображений были выбраны три снимка дистанционного зондирования Земли с космических аппаратов QuickBird и Sentinel-2, соответствующие разным типам подстилающей поверхности (см. рис. 3). Каждый снимок представляет собой полутоновое изображение в байтовом формате ($L = 256$) размером 1024×1024 пикселов.



а) б) с)

Рис. 3. Тестовые изображения. (а) QB, (б) S1, (в) S2

Для этих изображений в режиме скользящего окна размером 25×25 пикселов построены локальные гистограммы. Общий объем полученной коллекции гистограмм составил примерно 3 000 000 экземпляров – полагаем, этого достаточно для обоснования качественных выводов проводимого анализа.

4. О точности оценки энтропии в скользящем окне

Прежде чем перейти к экспериментам и выбору параметров предлагаемого алгоритма, нужно ответить на вопрос, насколько точно локальная гистограмма описывает истинное распределение вероятностей пикселов.

Распределение $P(x)$ является для нас гипотетическим, ненаблюдаемым, его экспериментальное измерение (оценка) осуществляется в соответствии с формулой (3), т.е. подсчетом целочисленных значений $N(x)$ ненормированной гистограммы по выборке из $\|D\|$ наблюдаемых случайных величин,

$$\sum_{x=0}^{L-1} N(x) = \left\|D\right\|. \tag{13}$$

Ниже для краткости выкладок временно опустим аргумент x.

Величина N – случайное число. Если проигнорировать вытекающий из (13) факт его ограниченности сверху, что вполне допустимо при больших $\|D\|$ и отсутствии аномально больших отдельных вероятностей в распределении (1), оно имеет распределение Пуассона [13], плотность которого в наших обозначениях задается соотношением:

$$W_N(N) = e^{-\|D\|P} \frac{(\left\|D\right\|P)^N}{N!}, \ N \ge 0. \tag{14}$$

Как известно, это случайное число имеет одинаковые математическое ожидание и дисперсию:

$$M_N = D_N = \left\|D\right\|P. \tag{15}$$

Абсолютная среднеквадратическая погрешность оценивания P равна

$$\varepsilon_P = \frac{1}{\|D\|}\sqrt{D_N} = \sqrt{P/\|D\|}, \qquad (16)$$

а относительная среднеквадратическая погрешность P –

$$\sigma_P = \frac{\varepsilon_P}{P} = \sqrt{\frac{1}{\|D\|P}}. \qquad (17)$$

В табл. 1 приведены примеры значений σ_P (в процентах) для некоторых типичных размеров двумерного окна обработки.

Табл. 1. Относительная погрешность оценивания вероятности (в процентах)

Объем выборки *D*	**Оцениваемая вероятность *P***								
	0,001	0,002	0,005	0,01	0,02	0,05	0,1	0,2	0,5
81 (9×9)	***351***	***248***	***157***	***111***	***79***	*50*	*35*	*25*	16
225 (15×15)	***211***	***149***	***94***	***67***	*47*	*30*	*32*	15	9
625 (25×25)	***126***	***89***	***57***	*40*	*28*	18	13	9	6
1225 (35×35)	***90***	***64***	*40*	*29*	20	13	9	6	4

В таблице жирным шрифтом выделены относительные погрешности больше 50 %, курсивом – больше 20 %. Видно, что при небольшом объеме выборки, характерном для скользящего окна, гистограмма описывает истинное распределение вероятностей с очень большой погрешностью, которая может оказаться неприемлема для практики. Однако в рамках данного исследования нас интересует не точность представления гистограммы, а точность оценки энтропии. Определим и ее. Рассмотрим любое слагаемое в формуле (2):

$$h = -P\log_2 P \approx \frac{N}{\|D\|}\log_2\frac{\|D\|}{N} = h(N). \qquad (18)$$

Формула (18) задает нелинейное преобразование случайной величины N в случайную величину h. Её математическое ожидание и дисперсия [13] выражается формулами (19) и (20). А если вернуть аргумент x в обозначения исходного распределения вероятностей и связанных с ним величин, то (19) и (20) перепишется в виде (21) и (22).

$$M_h = \sum_{N=0}^{\infty} h(N)W_N(N) = \frac{e^{-\|D\|P}}{\|D\|}\sum_{N=1}^{\infty}(\log_2\|D\| - \log_2 N)\frac{(\|D\|P)^N}{(N-1)!}, \qquad (19)$$

$$D_h = \sum_{N=0}^{\infty} h^2(N)W_N(N) - M_h^2 = \frac{e^{-\|D\|P}}{\|D\|^2}\sum_{N=1}^{\infty}N(\log_2\|D\| - \log_2 N)^2\frac{(\|D\|P)^N}{(N-1)!} - M_h^2. \qquad (20)$$

$$M_h(x) = \frac{e^{-\|D\|P(x)}}{\|D\|}\sum_{N=1}^{\infty}(\log_2\|D\| - \log_2 N)\frac{(\|D\|P(x))^N}{(N-1)!}, \qquad (21)$$

$$D_h(x) = \frac{e^{-\|D\|P(x)}}{\|D\|^2}\sum_{N=1}^{\infty}N(\log_2\|D\| - \log_2 N)^2\frac{(\|D\|P(x))^N}{(N-1)!} - M_h^2(x). \qquad (22)$$

И далее просуммируем все функционально преобразованные случайные величины в соответствии с формулой (2). Нет оснований считать случайные величины h статистически зависимыми при разных x. Поэтому в силу центральной предельной теоремы теории вероятностей [13] результат суммирования будет иметь распределение, близкое к нормальному, которое с достаточной полнотой описывается своими математическим ожиданием и дисперсией:

$$M_H = \sum_{x=0}^{L-1} M_h(x), \qquad (23)$$

$$D_H = \sum_{x=0}^{L-1} D_h(x). \qquad (24)$$

От дисперсии (24) можно перейти к среднеквадратичной погрешности этой оценки, как абсолютной:

$$\varepsilon_H = \sqrt{D_H}, \qquad (25)$$

так и относительной:

$$\sigma_H = \frac{\varepsilon_H}{H}. \qquad (26)$$

Расчеты по формулам (21) – (24) могут быть выполнены только для конкретного распределения вероятностей $P(x)$. Важно заметить, что энтропия инвариантна к порядку расположения отдельных элементов дискретного распределения вероятности, на нее влияет лишь степень неравномерности значений этих элементов, которые можно расположить в порядке убывания. Поэтому для численных экспериментов было выбрано усеченное экспоненциальное распределение целочисленной случайной величины:

$$P(x) = \frac{1-a}{1-a^L}a^x,\ 0 \le x \le L-1, \qquad (27)$$

где a – единственный параметр распределения, который регулирует его неравномерность ($0 < a < 1$). Обоснованием такого выбора является достаточная близость к распределению (27) усредненных и переупорядоченных гистограмм многих реальных изображений (см. рис. 4).

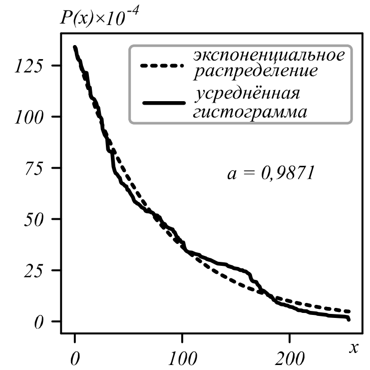


Рис. 4. Сравнение переупорядоченной усредненной гистограммы тестового изображения QB (рис. 3а) с усеченным экспоненциальным распределением (пунктир), L = 256

Неравномерность распределения (27) также можно задать через отношение между максимальным и минимальным значением вероятности, соответственно, на левом и правом краю области определения:

$$\frac{P_{\min}}{P_{\max}} = \frac{P(L-1)}{P(0)} = a^{L-1}. \qquad (28)$$

В табл. 2 приведены показатели точности оценивания энтропии (23) – (26) в зависимости от показателя неравномерности усеченного экспоненциального распределения (28) для различных размеров скользящего окна. Из таблицы видно, что, во-первых, оценка энтропии является смещенной в меньшую сторону, причем смещение особенно велико для малых окон, которые чаще всего используются на практике. Впрочем, систематическая погрешность (смещение) может и не влиять на результат обработки изображения, т.к. ее легко учесть при настройке (обучении) процедуры поэлементной классификации (см. Введение).

Более серьезную проблему представляет случайная погрешность оценки энтропии, которая также велика для окон малого размера. Интересно заметить, что дисперсия (и среднеквадратичная погрешность) оценки слабо зависит от степени неравномерности распределения вероятностей значений пикселов, а относительная среднеквадратичная погрешность (26) вообще остается практически постоянной.

Для байтовых пикселов, для которых максимально возможное значение энтропии равно 8, абсолютная среднеквадратичная погрешность оценки энтропии для окна 15×15 лежит в диапазоне 0,35 – 0,45, а для окна 25×25 – в диапазоне 0,2 ÷ 0,25.

5. Выбор базиса для приближённого представления гистограммы

Поиск системы базисных функций для разложения (5) и его усеченного варианта (6), оптимальной для решения задачи оценки энтропии, является предметом отдельного исследования и лежит за рамками настоящей статьи. Для демонстрации работоспособности и эффективности предлагаемого алгоритма мы ограничимся рассмотрением трех известных систем функций дискретного аргумента, ортогональных на интервале $[0, L-1]$:

- дискретного преобразования Фурье [1, 3 – 5, 15, 16] в вещественной форме (синусная и косинусная компоненты каждой гармоники считаются двумя функциями разложения);
- дискретного косинусного преобразования (ДКП-2) [15];
- дискретного преобразования Хаара [1, 4, 16].

Табл. 2. Показатели точности оценивания энтропии в скользящем окне

Объём выборки $\|D\|$		P_{min}/P_{max}	1	0,1	0,01	$1\cdot10^{-3}$	$1\cdot10^{-4}$
		a	1,0	0,9910	0,9821	0,9733	0,9645
		H	8,0	7,716	7,153	6,637	6,233
	81 (9×9)	M_H	6,043	5,939	5,719	5,492	5,288
		D_H	0,4158	0,3917	0,3448	0,3018	0,2676
		ε_H	0,6448	0,6258	0,5872	0,5494	0,5173
		σ_H	0,0806	0,0811	0,0821	0,0828	0,0830
	225 (15×15)	M_H	7,071	6,866	6,468	6,101	5,798
		D_H	0,1877	0,1717	0,1439	0,1216	0,1055
		ε_H	0,4332	0,4144	0,3793	0,3487	0,3248
		σ_H	0,0541	0,0537	0,0530	0,0525	0,0521
	625 (25×25)	M_H	7,670	7,389	6,864	6,404	6,044
		D_H	0,0703	0,0648	0,0542	0,0455	0,0392
		ε_H	0,2651	0,2545	0,2328	0,2132	0,1981
		σ_H	0,0331	0,0330	0,0325	0,0321	0,0318
	1225 (35×35)	M_H	7,842	7,554	6,997	6,506	6,126
		D_H	0,0354	0,0330	0,0279	0,0234	0,0202
		ε_H	0,1882	0,1817	0,1671	0,1531	0,1422
		σ_H	0,0235	0,0235	0,0234	0,0231	0,0228

Для выбора наилучшего базиса из перечисленных был проведен следующий вычислительный эксперимент. Для каждой локальной гистограммы, взятой из заранее сформированной коллекции (см. параграф 3) и рассматриваемой в качестве оценки распределения вероятностей $P(x)$, вычислялся ее спектр, т.е. все 256 коэффициентов ее разложения в заданной системе базисных функций (см. формулу (4)). Далее спектр усекался, т.е. в нем оставлялось только K младших коэффициентов разложения, остальные обнулялись. После чего по формуле (6) выполнялся обратный переход к теперь уже приближенной гистограмме.

На рис. 5 на примере конкретной локальной гистограммы показано, как выглядят ее приближения, полученные с помощью усеченного спектра в трех рассматриваемых системах базисных функций.

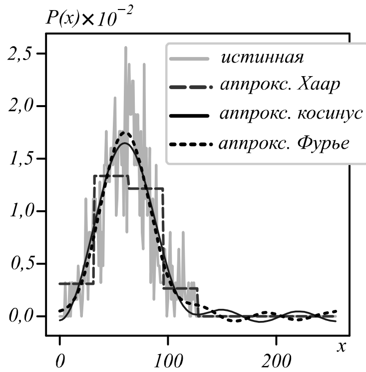


Рис. 5. Приближённое представление гистограммы с помощью усечённого спектра различных базисов: Фурье, косинус, Хаар (K = 8)

Функция, полученная в результате описанной процедуры, аппроксимирует гистограмму, но, строго говоря, таковой не является. Чтобы ее считать гистограммой, необходимо осуществить нормировку, обеспечивающую выполнение условия (1). Как видно из рис. 5, эта функция иногда может принимать отрицательные значения, которые перед нормировкой следует отсечь (приравнять нулю). Далее по формуле (2) находились значения энтропии для точной локальной гистограммы и ее аппроксимации, соответственно, H и H_y, вычислялась их разность – погрешность оценивания энтропии за счет усечения спектра гистограммы:

$$\Delta_y = H_y - H. \tag{29}$$

Затем по множеству локальных гистограмм, полученных для конкретного тестового изображения, находились статистические характеристики этой погрешности: среднее значение M_y и среднеквадратичное отклонение ε_y. На рис. 6 и рис. 7 представлены их значения в зависимости от K – числа используемых коэффициентов разложения гистограммы.

По этим рисункам видно, что погрешность оценивания энтропии по аппроксимациям гистограммы быстро убывает с увеличением используемых коэффициентов разложения. Среднеквадратичное отклонение этой погрешности оказалось минимальным для базиса Хаара. Его, очевидно, и следует использовать в предлагаемом алгоритме.

Необходимо отметить специфику базиса Хаара: иерархический характер его локальных базисных функций. Поскольку неизвестно, в какой части диапазона своих аргументов будет располагаться конкретная гистограмма, при ее аппроксимации целесообразно полностью исчерпывать определенный иерархический уровень базисных функций, т.е. выбирать число используемых коэффициентов разложения равным целой степени двойки: $K = 4, 8, 16...$

Как видно из рис. 7, при $K = 8$ среднеквадратичная погрешности оценки энтропии, возникающая за счет усечения спектра в базисе Хаара, лежит в диапазоне 0,25 – 0,45, то есть становится сопоставимой со среднеквадратичной погрешностью оценки энтропии из-за ограниченного числа пикселов в окне (см. параграф 4). Такое значение K мы и будем использовать далее.

Заметим, что все три рассмотренные системы ортогональных функций, включая систему Хаара, в качестве младшей функции $f_0(x)$ используют константу. Значение соответствующего ей коэффициента разложения F_0 предопределено условием нормировки (1). Поэтому в схеме алгоритма (рис. 2) можно исключить крайнюю левую ветвь и тем самым, во-первых, несколько уменьшить объем вычислений, во-вторых, упростить аппроксиматор, снизив на единицу число его входов. В этом случае при $K = 8$ аппроксиматор будет принимать на вход коэффициенты $F_1 \div F_7$, т.е. окажется семиканальным.

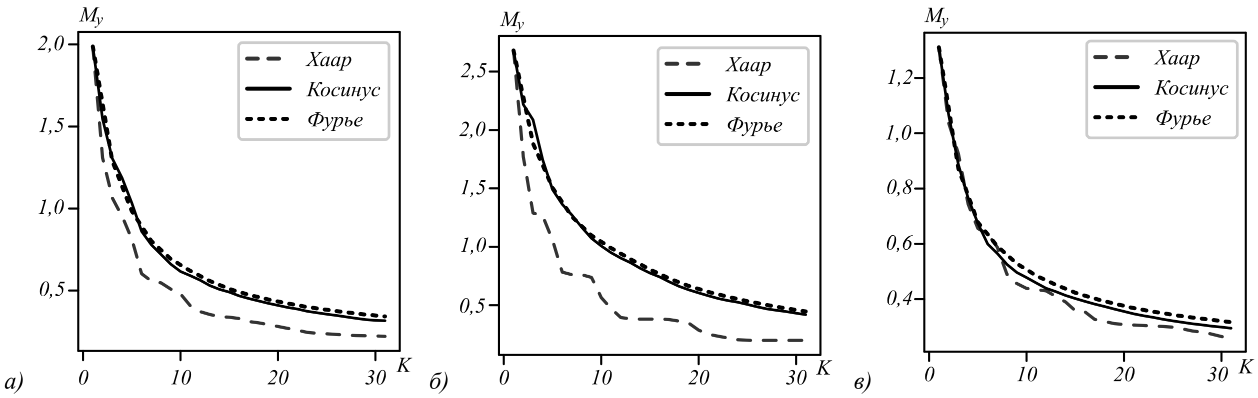


Рис. 6. Зависимости среднего значения погрешности оценивания энтропии от числа используемых коэффициентов разложения гистограммы в различных базисах (Хаар, косинус, Фурье) для трёх тестовых изображений. (а) QB, (б) S1, (в) S2

6. Построение аппроксиматора

Известно много вариантов построения аппроксиматора – блока получения приближенных значений функции многих переменных (7), см., например, обзор [18]. Однако большинство известных аппроксиматоров, включая широко используемые в последнее время нейросетевые [8, 12, 18], ориентированы на достижение требуемой точности аппроксимации, но не на минимизацию ее вычислительной сложности. Поэтому мы предлагаем использовать в нашем алгоритме простой метод многомерной иерархической аппроксимации [3], возможно, не самый точный, но обеспечивающий наиболее быстрое получение значений аппроксимируемой функции.

Кратко напомним принцип построения иерархического аппроксиматора применительно к нашему алгоритму. В общем случае, если не учитывать, что младшая базисная функция является известной константой (см. предыдущий параграф), строится K-мерный гиперкуб, в котором по осям откладываются значения коэффициентов разложения, нормированные к удобному диапазону возможного изменения, например, к стандартному интервалу байтовых целых чисел [0, 255]. В процессе обучения (настройки) этот гиперкуб последовательно разбивается по осям и порождает в памяти компьютера древовидную структуру. В каждой из областей, полученных в результате разбиения, осуществляется аппроксимация функции константой. Области с малой погрешностью аппроксимации принимаются за терминальные вершины дерева. Области, в которых погрешность велика, подвергаются дальнейшему разбиению. Условная «двумерная» (для $K=2$) иллюстрация к изложенному приведена на рис. 8*а*.

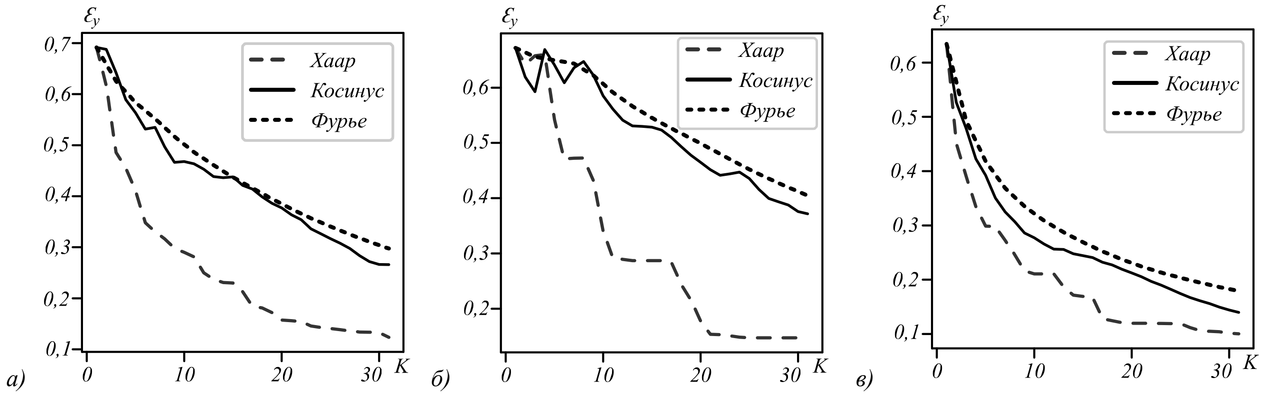


Рис. 7. Зависимости среднеквадратичного отклонения погрешности оценивания энтропии от числа используемых коэффициентов разложения гистограммы в различных базисах для трех тестовых изображений. (а) QB, (б) S1, (в) S2

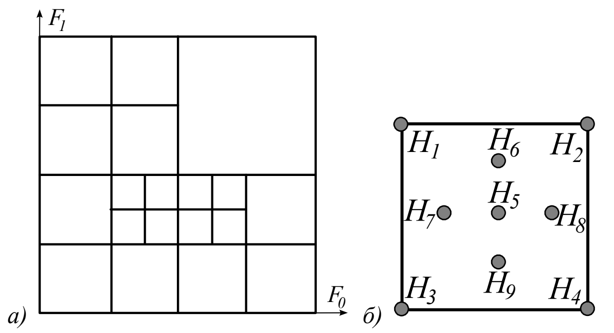


Рис. 8. Построение иерархического аппроксиматора. (а) Результирующая схема разбиения гиперкуба (остались терминальные вершины), (б) выборочный анализ точности аппроксимации в отдельной области

Если считать аргументы аппроксимируемой функции (7) приведенными к целочисленному байтовому формату, то K-мерный гиперкуб потенциально вмещает в себя 2^{8K} ее значений – столь большое число точек делает затруднительной или даже невозможной полную проверку точности аппроксимации. Поэтому мы предлагаем выполнять выборочную проверку, опираясь на предположение о достаточно гладком поведении аппроксимируемой функции. На рис. 8*б* условно показана отдельная область гиперкуба, для которой нужно принять решение либо об отнесении ее к терминальной вершине дерева, либо о разбиении. Рассматриваются 2^K угловых точек области (очевидно, она тоже является гиперкубом, но меньшего размера, нежели исходный). В них вычисляются значения функции H_i (H_1, H_2, H_3 и H_4 на рис. 2*б*) и проверяется выполнение условия:

$$\max_i\{H_i\} - \min_i\{H_i\} \le 2\Delta_A\,, \tag{30}$$

где i – условный номер вычисляемого значения функции, Δ_A – допустимая максимальная погрешность её аппроксимации. При нарушении неравенства (30) область подлежит дальнейшему разбиению. Если неравенство выполняется, то множество контролируемых значений расширяется, например, добавлением центральной точки области (на рис.8*б* – H_5) и других точек (H_6, H_7, H_8, H_9 и т.д.), число которых определяется лишь лимитом времени и вычислительных ресурсов для настройки аппроксиматора. Если для расширенного множества контролируемых точек также выполняется условие (30), область считается терминальной вершиной дерева и ей приписывается аппроксимирующее значение функции:

$$\bar{H} = \frac{1}{2}(\max_i\{H_i\} - \min_i\{H_i\}). \tag{31}$$

Остановимся на выборе значения Δ_A в формуле (30). Случайное отклонение значения энтропии, возвращаемого аппроксиматором, от значения, точно вычисленного по K коэффициентам спектрального разложения гистограммы, очевидно, лежит в интервале $[-\Delta_A, \Delta_A]$. Если предположить, что это отклонение равномерно распределено на указанном интервале, то его среднеквадратичное значение равно

$$\varepsilon_A = \Delta_A \big/ \sqrt{3}. \qquad (32)$$

Эта величина характеризует погрешность оценки энтропии за счет аппроксимации функции (7), добавляемую к погрешности ε_y, возникающей за счет усечения спектра гистограммы. Суммарная среднеквадратичная погрешность оценивания задается формулой

$$\varepsilon_\Sigma = \sqrt{\varepsilon_y^2 + \varepsilon_A^2}. \qquad (33)$$

Из (33) следует, что добавочная среднеквадратичная погрешность (32) несущественно (для определенности – на 10 %) увеличит погрешность из-за усечения спектра, т.е. $\sigma\Sigma = 1{,}1\sigma_y$, если

$$\varepsilon_A = 0{,}46\varepsilon_y. \qquad (34)$$

Из (32) и (34) получаем:

$$\Delta_A = 0{,}80\varepsilon_y. \qquad (35)$$

В параграфе 5 было установлено, что при описании гистограммы восемью младшими компонентами спектра в базисе Хаара минимальное значение $\varepsilon_y \approx 0{,}25$, ориентируясь на него, из (35) находим:

$$\Delta_A = 0{,}2.$$

Выше также было упомянуто о преобразовании аргументов аппроксимируемой функции (7) в целочисленный байтовой формат. Такое преобразование должно опираться на исходный диапазон изменения указанных величин. Если коэффициент F_k разложения гистограммы принимает значения в диапазоне $[\min F_k, \max F_k]$, то его линейное преобразование к диапазону [0, 255], очевидно, задается формулой

$$\tilde{F}_k = 255\frac{F_k - \min F_k}{\max F_k - \min F_k}, \qquad (36)$$

где $\tilde{F}_k$ – нормированный коэффициент – входной параметр иерархического аппроксиматора. Экспериментально установлено, что для используемых тестовых изображений средние значения коэффициентов разложения гистограмм по базису Хаара близки к нулю, но имеют сильный разброс среднеквадратичных отклонений ε_F (см. рис. 9).

Рисунок показывает, что среднеквадратичные отклонения коэффициентов немонотонно зависят от номера коэффициента (в интересующем нас диапазоне), но нигде не превышают максимального значения $\varepsilon_F = 0{,}045$. Поэтому, ориентируясь на известное правило «трех сигм», можно принять с запасом:

$$-\min F_k = \max F_k = 0{,}135,\ 1 \le k \le 7.$$

7. Вычислительная сложность оценки энтропии в скользящем окне

Сравним вычислительную сложность предлагаемого алгоритма с базовым алгоритмом оценки энтропии по локальной гистограмме в скользящем окне, кратко описанным в параграфе 1. Теоретическую (алгоритмическую) сложность оценим числом операций сложения/вычитания (U_+), умножения ($U_\times$) и обращения к таблице (U_T), выполняемых в каждом положении скользящего окна, а также суммарным показателем:

$$U_\Sigma = U_+ + U_\times + U_T. \qquad (37)$$

Как уже было оговорено в параграфе 1, мы не рассматриваем особенности работы алгоритмов на краях изображения.

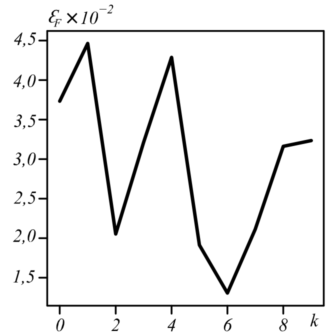


Рис. 9. Среднеквадратичные отклонения коэффициентов разложения гистограмм по базису Хаара

Базовый алгоритм. Сначала формируется гистограмма. Значение каждого пиксела в окне рассматривается как адрес в предварительно заполненной нулями таблице длиной в L ячеек, по которому в таблицу добавляется единица (или, если сразу формируется нормированная гистограмма, то величина $\|D\|^{-1}$), см. формулу (3). Каждое такое действие – одно обращение к таблице и одно сложение. При заданном положении окна его нужно выполнить для каждого пиксела, т.е. $\|D\|$ раз. Если окно прямоугольное с размерами $M_1 \times M_2$ пикселов, то $\|D\| = M_1 M_2$. Но мы будем использовать известный рекурсивный алгоритм формирования гистограммы [1, 7, 9, 10], более эффективный в вычислительном плане, в нем удается снизить повторение описанного выше действия до $2(M_1+1)$ раз (предполагается, что окно движется по горизонтали, а M_1 – его вертикальный размер).

После того, как гистограмма получена, по формуле (2) производится вычисление энтропии. Вычисление непосредственно по этой формуле требует многократного выполнения операций логарифмирования и умножения. Но можно заранее протабулировать значения слагаемых, входящих в (2), в зависимости от вероятности, т.е. свести вычисления к обращению к таблице. После L-кратного извлечения из таблицы значений слагаемых остается $(L-1)$ раз выполнить их сложение для вычисления энтропии.

Таким образом, получаем следующие показатели алгоритмической сложности базового алгоритма:

$$\begin{aligned} &U_+ = 2(M_1+1) + L - 1, \\ &U_\times = 0,\ U_T = 2(M_1+1) + L. \end{aligned} \qquad (38)$$

Предлагаемый алгоритм. Здесь вычислительный процесс разделяется на K параллельных ветвей (см. рис. 2). В каждой ветви сначала выполняется поэлементное преобразование изображения в соответствии с формулой (9). Следуя общепринятой технологии LUT «Look Up Table» [4], это преобразование сводится к обращению к таблице, в которую заранее занесены значения преобразованных пикселов. Поскольку число положений скользящего окна примерно равно числу пикселов в изображении, можно считать, что для одного положения окна в одной ветви алгоритма выполняется одно обращение к таблице.

Затем в соответствии с разностными уравнениями (11) выполняется рекурсивное суммирование пикселов преобразованного изображения в пределах окна. На его реализацию требуется четыре операции сложения /вычитания независимо от размеров окна. Кроме того, чтобы перейти от суммы к среднему значению, нужно выполнить одно умножение, см. формулу (12).

В результате получаем значение коэффициента разложения гистограммы по выбранному базису. Это значение затем масштабируется к диапазону, нужному для аппроксиматора. Согласно формуле (36), это линейное преобразование выполняется за одно сложение и одно умножение.

Далее все ветви алгоритма объединяются в иерархическом аппроксиматоре. Поскольку для конкретного набора входных переменных неизвестно, на каком уровне построенного дерева его вершина окажется терминальной, невозможно дать точную оценку алгоритмической сложности аппроксиматора. Поэтому мы ограничимся ее оценкой сверху, считая, что все ветви дерева доходят до нижнего, восьмого уровня. Для каждой входной переменной на каждом уровне необходимо выполнить операцию сравнения с порогом (по сути – вычитания) и обращения к иерархической таблице. В целом для всех переменных эта процедура должна повториться ровно $8K$ раз.

Итого, показатели алгоритмической сложности предложенного алгоритма равны:

$$U_+ = 13K,\ U_\times = 2K,\ U_T = 9K. \qquad (39)$$

В табл. 3 для конкретных значений параметров алгоритмов ($L=256$, $M_1=25$, $K=7$) приведены значения величин (38), (39), суммарного показателя вычислительной сложности (37), а также относительный выигрыш предлагаемого алгоритма по сравнению с базовым.

Табл. 3. Алгоритмическая сложность алгоритма оценки энтропии ($L = 256$, $M_1 = 25$, $K = 7$)

	U_+	$U_\times$	U_T	U_Σ
Базовый алгоритм	308	0	307	615
Предлагаемый алгоритм	91	14	63	168
Относительный выигрыш	3,38	-	4,87	3,66

Из таблицы видно, что теоретически предлагаемый алгоритм в 3-4 раза обыгрывает базовый алгоритм по числу выполняемых операций.

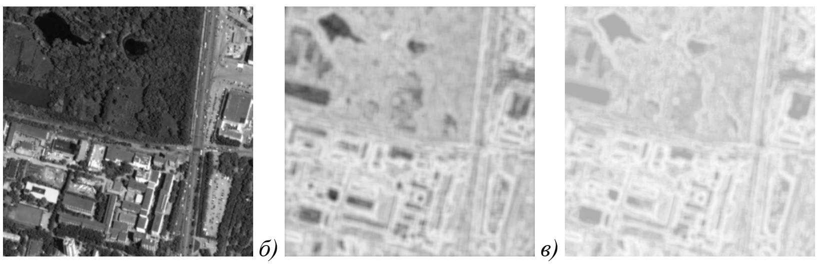
8. Результаты экспериментов

Результаты представленного выше анализа не являются точными, поскольку опираются на ряд допущений. Особенно это касается оценки вычислительной сложности алгоритмов, которая должна непременно учитывать особенности современных компьютеров: число ядер процессора, скрытые функции операционной системы, особенности транслятора и т.д.

Чтобы подтвердить и отчасти верифицировать теоретические результаты статьи, были разработаны программы, реализующие и базовый, и новый алгоритм оценки энтропии. Был использован компьютер со следующими характеристиками: Intel Pentium Gold G6405 4.10GHz, 2 ядра, ОЗУ Kingston Hyper X Fury Black 8 GB, 2400 МГц.

Время обработки на этом компьютере одного тестового изображения (полутоновое, 1024×1024 байтовых пиксела) составило: 0,325 с для базового алгоритма, 0,074 с для нового. То есть экспериментально подтвержденный выигрыш нового алгоритма в скорости составляет 4,39 раз.

На рис. 10 представлено тестовое изображение QB (см. рис. 3*а*) и поля его локальной энтропии в окне 25×25, построенные программами, реализующими базовый и предложенный алгоритм. Чтобы лучше визуально представить поля энтропии, значения их пикселов увеличены в 32 раза.



а) б) в)

Рис. 10. Иллюстрация построения поля энтропии. (а) тестовое изображение, (б) поле энтропии по базовому алгоритму, (в) поле энтропии по предложенному быстрому алгоритму

В дополнение к рис. 10 на рис. 11 приведены одномерные графики локальной энтропии, построенные для средней строки изображения QB.

Для всех трех тестовых изображений (рис. 3) среднеквадратичные погрешности оценки энтропии предлагаемым алгоритмом составили соответственно

0,469, 0,417 и 0,401. Эти значения вполне сопоставимы с погрешностью, вызванной неопределенностью описания гистограммой, построенной по окну ограниченного размера, истинного (ненаблюдаемого) распределения вероятностей пикселов изображения (см. параграф 4).

Рис. 11. Графики локальной энтропии для строки 511 изображения QB, вычисленной по базовому и предложенному алгоритмам

Заключение

В статье рассмотрен ряд вопросов, связанных с вычислением важной характеристики – локальной энтропии в скользящем окне обработки изображения. Предложен и исследован новый алгоритм приближенной оценки энтропии, обладающий в несколько раз меньшей вычислительной сложностью по сравнению с известными.

Для предложенного алгоритма осуществлен экспериментальный выбор ортогонального базиса для представления локальной гистограммы в виде усеченного ряда. Из числа рассмотренных базисов наилучшим оказался базис Хаара.

Рассмотрена методика построения иерархического аппроксиматора, осуществляющего быстрый пересчет множества коэффициентов разложения гистограммы в значение энтропии.

В качестве замечаний, дополняющих содержание статьи или указывающих направления дальнейших исследований и разработок, можно отметить следующее.

1. В ряде публикаций [7, 8, 12] отмечена нежелательность использования при обработке изображения прямоугольного окна, порождающего анизотропность (различие по направлениям в плоскости изображения) результатов обработки. В рамках предложенного алгоритма данный недостаток легко нивелируется применением многократного сглаживания. В статье [19] показано, что рекурсивное сглаживание, описанное в параграфе 2, будучи примененным трижды, делает результат обработки практически изотропным. При этом общая вычислительная сложность алгоритма оценки энтропии возрастает незначительно.

2. Предложенный быстрый алгоритм легко обобщается на построение других полей признаков изображения, основанных на локальной гистограмме, вычисляемой в режиме скользящего окна. Такое обобщение потребует всего лишь замены аппроксиматора, теперь настроенного на другой признак. Наибольший эффект в скорости вычислений будет достигнут при одновременном синтезе нескольких полей локальных гистограммных признаков, когда обобщенный алгоритм будет состоять из общей части (вычисления коэффициентов разложения гистограммы) и набора нужных аппроксиматоров.

3. Можно попытаться распространить предложенный подход к построению полей локальных признаков, основанных на использовании многомерных, в частности, двумерных гистограмм [2, 4]. Однако это является предметом будущих исследований.

Благодарности

Работа выполнена при поддержке гранта РНФ № 23-11-20013 (параграфы 1 – 3, 8) и по государственному заданию НИЦ «Курчатовский институт» (параграфы 4 – 7).

References

[1] Yaroslavsky LP. Introduction to digital image processing [In Russian]. Moscow: "Sovetskoe Radio" Publisher; 1979.
[2] Pratt WK. Digital image processing. New York: Willey; 1978. ISBN: 0471018880.
[3] Gashnikov MV, Glumov NI, Ilyasova NY, Myasnikov VV, Popov SB, Sergeyev VV, Soifer VA, Khramov AG, Chernov AV, Chernov VM, Chicheva MA, Fursov VA. Computer image processing methods [In Russian]. Moscow: "Fizmatlit" Publisher; 2003. ISBN: 5-9221-0270-2.
[4] Gonzalez RC, Woods RE. Digital image processing. London: Pearson; 2018. ISBN: 978-0-13-335672-4.
[5] Shapiro L, Stockman J. Computer vision [In Russian]. Moscow: Binom, Knowledge Laboratory; 2006. ISBN: 5-94774-384-1.
[6] Schowengerdt_RA. Remote sensing: models and methods for image processing. Academic Press; 2006. ISBN: 978-0123694072.
[7] Wei Y, Tao L. Efficient histogram-based sliding window. IEEE Computer Society Conf on Computer Vision and Pattern Recognition 2010: 3003-3010. DOI: 10.1109/CVPR.2010.5540049.
[8] Velichko A, Belyaev M, Wagner MP, Taravat A. Entropy approximation by machine learning regression: application for irregularity evaluation of images in Remote Sensing. Remote Sens 2022; 14(23): 5983. DOI: 10.3390/rs14235983.
[9] Huang TS, Yang GY, Tang GY. A fast two-dimensional median filtering algorithm. IEEE Trans Acoust Speech Signal Process 1979; ASSP-27(1): 13-18. DOI: 10.1109/TASSP.1979.1163188.
[10] Huang TS, Eklund JO, Nussbaumer GJ, et al. Fast algorithms in digital image processing [In Russian]. Moscow: “Radio i Svyaz” Publisher; 1984.
[11] Chentsov NN. Estimation of unknown distribution density from observations [In Russian]. Report USSR Academy of Sciences 1962; 147(1): 45-48.
[12] Kass M, Solomon J. Smoothed local histogram filters. ACM Trans Graph 2010; 29(426): 100. DOI: 10.1145/1778765.1778837.

[13] Ventzel ES. Probability theory [In Russian]. Moscow: "Nauka" Publisher; 1969.
[14] Shannon K. Works on information theory and cybernetics [In Russian]. Moscow: Publishing House of Foreign Literature; 1963.
[15] Oppenheim A, Schafer R. Digital signal processing [In Russian]. Moscow: "Tehnosphera" Publisher; 2012. ISBN: 978-5-94836-329-5.
[16] Zalmanzon LA. Fourier, Haar, Walsh transforms and their application in control, communications and other areas [In Russian]. Moscow: "Nauka" Publisher; 1989. ISBN: 5-02-014094-5.
[17] Dedus FF, Kulikova LI, Pankratov AN, Tetuev RK. Classical orthogonal bases in problems of analytical description and processing of information signals [In Russian]. Moscow: VMKMGU Publisher; 2004.
[18] Shvedov AS. Approximation of functions using neural networks and fuzzy systems [In Russian]. Problems of Control 2018; 1: 21-29.
[19] Sergeev VV, Myasnikov VV. Algorithm for fast implementation of the Gabor filter [In Russian]. Avtometria 1999; 6: 51-55.

Сведения об авторах

Сергеев Владислав Викторович, 1951 года рождения. В 1974 году окончил Куйбышевский авиационный институт (ныне Самарский университет). В 1978 г. защитил кандидатскую, а в 1993 году – докторскую диссертацию. В настоящее время является заведующим кафедрой геоинформатики и информационной безопасности Самарского университета, по совместительству – начальником лаборатории математических методов обработки изображений Института систем обработки изображений, НИЦ «Курчатовский институт». Область научных интересов: цифровая обработка сигналов, анализ изображений, распознавание образов, геоинформатика.
E-mail: *vserg@geosamara.ru*

Баврина Алина Юрьевна, 1980 года рождения, в 2003 году окончила Самарский государственный аэрокосмический университет имени академика С.П. Королёва (ныне Самарский университет), в 2006 году защитила кандидатскую диссертацию. Работает инженером-математиком в АО «Самара-Информспутник», а также в Самарском университете в должности старшего научного сотрудника по совместительству. Области научных интересов: обработка изображений, распознавание образов, дистанционное зондирование Земли, защита цифровых изображений. E-mail: *bavrina@mail.ru*

Зайцев Илья Дмитриевич, 2002 года рождения, студент 3 курса специалитета «Информационная безопасность автоматизированных систем» Самарского университета. Область интересов: обработка изображений, оптимизация вычислений. E-mail: *dp.io1866@yandex.ru*

Лазутов Марк Юрьевич, 2002 года рождения, студент 4 курса бакалавриата «Прикладная математика и информатика» Самарского университета. Область интересов: обработка изображений, распознавание образов. E-mail: *lazutov.mark@mail.ru*

Шапиро Давид Александрович, 1999 года рождения. В 2022 году окончил Самарский национальный исследовательский университет. Аспирант Самарского университета. Область интересов: анализ изображений, цифровые водяные знаки, стеганография. E-mail: *david-shapiro@mail.ru*

ГРНТИ: 28.23.15
Поступила в редакцию 16 февраля 2024 г. Окончательный вариант – 7 мая 2024 г.

On estimating the local entropy of an image in a sliding window

V.V. Sergeyev [1,2], A.Y. Bavrina [1,3], I.D. Zaitsev [1], M.Y. Lazutov [1], D.A. Shapiro [1]
[1] Samara National Research University, 443086, Samara, Russia, Moskovskoye Shosse 34;
[2] Image Processing Systems Institute, NRC «Kurchatov Institute», 443001, Samara, Russia, Molodogvardeyskaya 151;
[3] JSC «Samara-Informsputnik», 443080, Samara, Russia, Karl Marx ave. 192, off. 717

Abstract

The paper considers a number of issues related to the calculation of the local entropy in a sliding window for image processing tasks. A new algorithm for entropy estimation is proposed, which has several times lower computational complexity than the known ones. The algorithm is based on the approximate representation of the histogram in the form of a truncated series expansion over some system of orthogonal basis functions, recursive calculation of the coefficients of this expansion in a sliding window and subsequent recalculation of the coefficients into the entropy estimate. Questions of an orthogonal basis selection for representing the local histogram as a truncated series are considered and the appropriateness of using the Haar basis is shown. A technique of constructing hierarchical approximator, which realizes fast recalculation of histogram decomposition coefficients into entropy value, is described. The theoretical statements of the paper are verified by experiments on the Earth remote sensing images.

Keywords: image processing, local feature generation, sliding window, local histogram, entropy estimation, basis function decomposition, recursive processing, hierarchical approximation, estimation error, computational complexity.

Citation: Sergeyev VV, Bavrina AY, Zaitsev ID, Lazutov MY, Shapiro DA. On estimating the local entropy of an image in a sliding window. Computer Optics 2024; 48(5): 714-725. DOI: 10.18287/2412-6179-CO-1509.

Acknowledgements: This work was financially supported by the Russian Science Foundation under project No. 23-11-20013 (Sections 1-3, 8) and the RF Ministry of Science and the State assignment of NRC "Kurchatov Institute" (Sections 4-7).

Authors' information

Vladislav Victorovich Sergeyev (b. 1951) graduated from Kuibyshev Aviation Institute (presently, Samara National Research University) in 1974. In 1978 he received his PhD and in 1993 – DrSc. Currently he works as the head of GIS and Information Security department at Samara National Research University, also holding a part-time position as the head of a laboratory at the Image Processing Systems Institute, NRC «Kurchatov Institute». Research interests are digital signals and image processing, GIS and pattern recognition. E-mail: *vserg@geosamara.ru*

Alina Yurievna Bavrina (b. 1980) graduated from Samara State Aerospace University (presently, Samara National Research University) in 2003, received her PhD in 2006. Currently she is an engineer-mathematician at JSC «Samara-Informsputnik» and has a part-time position as senior researcher at Samara National Research University. Area of interests: digital image processing, pattern recognition, Earth remote sensing, watermarking. E-mail: *bavrina@mail.ru*

Ilya Dmitrievich Zaitsev (b. 2002) is the 3-rd year student of a specialist program «Information Security of Automated Systems» at Samara National Research University. Area of interest: image processing, computational optimization. E-mail: *dp.io1866@yandex.ru*

Mark Yurievich Lazutov (b. 2002) is the 4-th year student of a bachelor's program «Applied Mathematics and Informatics» at Samara National Research University. Area of interests: digital image processing, pattern recognition.
E-mail: *lazutov.mark@mail.ru*

David Aleksandrovich Shapiro (b. 1999) graduated from Samara National Research University in 2022. Postgraduate student at Samara National Research University. Area of interest: image analysis, digital watermarks, steganography. E-mail: *david-shapiro@mail.ru*

Code of State Categories Scientific and Technical Information (in Russian – GRNTI)): 28.23.15
Received February 16, 2024. The final version – May 7, 2024.