

Сравнение двумерных вариаций в контексте оценки сложности дискретных изображений

П. А. Чочиа, О. П. Милloкова

Институт проблем передачи информации им. А. А. Харкевича РАН, Москва, Россия
Поступила в редколлегию 27.04.2015

Аннотация—Исследуются вопросы оценки сложности дискретных изображений. Рассматриваются различные известные варианты двумерных вариаций в применении к дискретным изображениям. Изучаются возможности использования двумерных вариаций для оценки сложности изображений, которая интерпретируется как характеристика, отражающая число, размеры и заметность деталей. Предложена модификация оценок вариации, названная *показатель размеров объектов*. Проанализированы изменения значений оценок сложности при различных возможных трансформациях изображения. Теоретические выводы подтверждены экспериментальными исследованиями.

КЛЮЧЕВЫЕ СЛОВА: двумерные вариации, сложность изображения, обработка видеoinформации.

1. ВВЕДЕНИЕ

В теоретических исследованиях и при практическом применении методов обработки видеoinформации требуются оценки достигаемого результата. Для задач, которые удается сформулировать как задачи восстановления или фильтрации сигнала, часто используются различные варианты отклонений (MSE, PSNR, SSIM и др.), которые удобны для теоретических разработок или в тех случаях, когда исходный сигнал известен. В реальных же ситуациях неискаженный сигнал недоступен. При таких условиях хотелось бы найти численную оценку результата, не использующую оригинальный сигнал, которая была бы адекватна изображению как двумерной функции.

Как правило, для решения данной задачи предлагаются методы, использующие один из следующих двух подходов. Первый заключается в попытке решения задачи оценивания качества изображения [1, 2]. При этом отмечаются три уровня анализа, отличающиеся степенью доступности к неискаженному эталонному изображению для сравнения с ним результата. На первом уровне эталонное изображение доступно полностью, на втором лишь частично, на третьем полностью недоступно. Последний уровень назван «слепым» оцениванием. Внимание исследователей почти всегда фокусируется на первом, иногда на втором из указанных уровней, однако они оба предполагают использование оригинального сигнала. Другой подход чаще всего связан с независимым оцениванием сложности изображения. Обычно он используется при существенных ограничениях, например, в задаче сжатия изображений [3] или в задаче автоматического опознавания целей [4].

Численные оценки сигнала актуальны при решении различных задач, в частности, классической оптимизационной задачи восстановления искаженного сигнала методом регуляризации [5], которая формулируется следующим образом. Пусть исходный сигнал z искажен оператором A (линейным или нелинейным) и шумом ξ , в результате чего наблюдается сигнал

$u = A(z) + \xi$. Полагая оператор A с определенной точностью известным, строится восстанавливающий оператор \hat{A} (алгоритм) $\hat{A}(u) \Rightarrow \{z_\alpha\}$, отображающий u на множество допустимых результатов $z_\alpha \in Z$. Вычисляется функционал

$$\varphi_\alpha(u, \lambda) = \|Az_\alpha - u\| + \lambda\Omega(z_\alpha), \quad (1)$$

где $\Omega(z_\alpha)$ — неотрицательный стабилизирующий функционал, а λ — регуляризирующий параметр. Оптимальный на Z результат z^* находится путем минимизации функционала $\varphi_\alpha(u, \lambda)$ при выбранном значении λ :

$$\varphi_{z^*}(u, \lambda) = \inf_{z_\alpha \in Z} \varphi_\alpha(u, \lambda). \quad (2)$$

Восстанавливающий оператор \hat{A} строится исходя из предположений о свойствах искажающего оператора A и зависит от конкретных условий. В каких-то случаях, как в задаче линейного восстановления [5], он может быть обратным оператором $\hat{A} = A^{-1}$, в других, как, например, в задаче фильтрации шума на основе полной вариации [6], — единичным оператором $\hat{A} = E$, или каким-то другим.

Для нас важным в данном алгоритме является то, что в отличие от восстанавливающего оператора \hat{A} , стабилизирующий функционал $\Omega(z)$, вообще говоря, не зависит от искажающего преобразования и выбирается как некоторая характеристика сигнала z , которая может задаваться исходя из модели сигнала. Для одномерных функций обычно предлагается использовать оценки типа нормы или вариации [5]. В контексте исследования полезно отметить, что функционал $\Omega(z)$ можно интерпретировать как некоторую оценку сложности сигнала.

Применительно к изображению, которое является двумерным сигналом, также хотелось бы выбрать оценку, характеризующую его пространственную изменчивость. Плодотворным здесь представляется подход, основанный на рассмотрении изображения как функции с ограниченной вариацией [7] и оценивании ее изменчивости при помощи механизма двумерных и многомерных вариаций [8–11]. Вопрос возможности применения двумерных вариаций для оценки изображения ранее обсуждался в статьях [12, 13]; настоящая работа является продолжением исследований.

В работе будут рассмотрены различные известные варианты двумерных вариаций в применении к дискретным сигналам и исследованы возможности использования их для оценки сложности изображений. Сложность изображения будет интерпретироваться как характеристика, отражающая число, размеры и заметность (контраст) деталей на изображении. Разумно также полагать, что при увеличении числа и контраста деталей сложность должна возрастать, а при увеличении размеров деталей — уменьшаться.

2. О СЛОЖНОСТИ СИГНАЛОВ И ФУНКЦИЙ

Изучая вопрос сложности изображений первоначально рассмотрим, как трактуется понятие сложности для сигналов, функций и потока сообщений. Одномерный сигнал, являющийся действительной функцией времени, характеризуется энергией, длительностью и шириной спектра [14]. В теории могут использоваться различные определения длительности Δt и ширины спектра сигнала Δf , но согласно преобразованиям Фурье эти две величины в любом случае не могут быть конечными одновременно. Для реальных сигналов, имеющих начало и конец во времени, шириной спектра как правило считают полосу частот, в которой сосредоточена почти вся энергия сигнала (скажем, 99%); в теоретических исследованиях в качестве Δf часто принимают положение первого нуля спектра. Как известно из теории сигналов [15], независимо от выбранного определения длительности и ширины спектра сигнала, минимальное значение произведения $\Delta f \Delta t$, которое принято называть *базой*, для простых сигналов примерно равно единице. Соответственно, сигналы, для которых $\Delta f \Delta t \gg 1$, называют сложными [14]. Но к

изображениям данное определение применить затруднительно. Во-первых, изображение является существенно двумерной функцией, а во-вторых (отвлекаясь от сканерной технологии формирования), точки изображения, вообще говоря, соответствуют одному и тому же моменту или интервалу времени.

Сложность аналитической функции многих переменных, как отмечается в [16], характеризуется числом ее переменных n , а для s раз дифференцируемых функций — отношением n/s . Этот подход мало применим к изображению как функции яркости пространственных координат, поскольку в данном случае имеем всего $n = 2$.

Теоретико-информационный подход к оценке сложности потока сообщений, представимого в дискретном виде, основан на вычислении энтропии $H = -\sum_b p(z) \log(p(z))$, где $p(z)$ — вероятность элемента изображения со значением z , а $b \in [b_{min}, b_{max}]$ — диапазон значений z . Такой вариант по сути предполагает независимость значений в точках изображения, в то время как изображение любой сцены является сигналом с сильными корреляционными связями. Поэтому подходы, используемые при оценивании сложности потока независимых сообщений, например энтропия, оказываются несостоятельными. Использование в качестве значения z некоторой функции, вычисляемой по близлежащим элементам, также не дает существенного улучшения и ограничивает универсальность оценки.

Для оценивания сложности дискретного сигнала иногда используется коэффициент сокращения объема данных (количества бит), необходимых для его восстановления. Такой подход, идеологически близкий понятию колмогоровской сложности [17], в реальности связан с необходимостью выбора того или иного алгоритма сжатия и по сути является вариантом энтропийного оценивания.

Каждый из указанных способов оценки сложности по тем или иным причинам мало или даже совсем не применим к изображению как двумерной функции пространственных координат, отражающей некоторую физическую характеристику наблюдаемой сцены. В частности, существенный общий недостаток перечисленных вариантов оценок заключается в том, что они не являются метрическими характеристиками и не зависят от амплитуды сигнала.

3. ДВУМЕРНЫЕ ВАРИАЦИИ И ИХ ДИСКРЕТНЫЕ АНАЛОГИ

Распространенной метрической характеристикой изменчивости и сложности одномерной функции на отрезке $[a, b]$ является вариация. Считая функцию $z(x)$ отражением некоей реальной физической величины (например, энергии или плотности) можно полагать, что $z(x)$ на наблюдаемом отрезке ограничена и имеет конечное число точек разрыва первого рода. Тогда $z(x)$ является функцией с ограниченной полной вариацией, которая по определению есть величина

$$V_a^b(z(x)) \triangleq \sup_P \sum_{k=0}^{K-1} \|z(x_{k+1}) - z(x_k)\|, \quad (3)$$

т.е. точная верхняя грань по всем возможным разбиениям P отрезка $[a, b]$. В контексте кусочно-постоянной модели изображения [18] существенно отметить, что значение вариации не изменяется при «вставке» в функцию $z(x)$ участков вида $z(x) = \text{const}$, $\{x \in [c, d] \subset [a, b]\}$, если при этом не добавляются разрывы первого рода.

Полная вариация $V_a^b(z(x))$ может служить удовлетворительной оценкой сложности одномерной функции. Предложено много ее обобщений на случай функции многих переменных (вариации Витали, Арцела, Фреше, Тонелли и др.). Все они сводятся к тому, что определяется некоторый один функционал, ограниченность которого гарантирует наличие у функции ряда свойств, аналогичных свойствам одномерной функции с конечным изменением. Однако перечень таких свойств оказывается ощутимо беднее набора свойств вариации для одномерных

функций, хотя при этом и удается формулировать отдельные выводы и теоремы [10]. К тому же возникает определенная неоднозначность многомерных вариаций, связанная во многом с тем, что они существенно зависят от выбора системы координат.

Многомерные вариации (как и одномерные) формулируются для непрерывных функций через точные верхние грани по множеству допустимых разбиений носителя секущими гиперплоскостями на элементарные параллелепипеды. В случае дискретной функции, минимальное разбиение ограничено снизу пространственной точностью ее представления. Поэтому формулы вариации дискретных функций являются лишь соответствующими аналогами и приближениями формул вариации непрерывных функций, что в некоторых случаях может вызвать определенные неточности. Также необходимо отметить, что далеко не все операции могут быть применимы к дискретной функции без потери информации: например, операции поворота на произвольный угол, или операции сжатия как в области носителя, так и в пространстве значений.

Рассмотрим, как будут выглядеть наиболее известные определения вариаций [10, 11, 19] в применении к двумерной дискретной функции $\mathbf{F} = \{f(i, j)\}$ ($i = 1, \dots, I; j = 1, \dots, J$) на прямоугольном носителе D с соответствующим разбиением.

Арцела вариация [10, 19] по сути является суммой модулей градиентов по координатным осям и представляется формулой

$$A(\mathbf{F}, D) = \sum_{i=1}^{I-1} \sum_{j=1}^{J-1} (|f(i+1, j) - f(i, j)| + |f(i, j+1) - f(i, j)|). \quad (4)$$

Данный способ оценки изменчивости двумерной функции используется в большинстве теоретических и практических разработок.

Витали вариация [10, 19] формально записывается как

$$V(\mathbf{F}, D) = \sum_{i=1}^{I-1} \sum_{j=1}^{J-1} |f(i+1, j+1) + f(i, j) - f(i, j+1) - f(i+1, j)|. \quad (5)$$

В качестве существенного недостатка данного дискретного представления вариации следует отметить, что на участках функции $f(i, j)$, где наблюдаются перепады ее значений в направлении только одной из координатных осей (i или j), отклик оператора, соответствующего формуле (5), будет равен нулю.

Пьерпонта вариация [19, 20] предполагает равномерное разбиение каждой из сторон прямоугольника D на равные между собой отрезки, что идеологически соответствует равномерной дискретизации носителя. В каждом образуемом элементарном параллелепипеде $\delta(i, j)$ измеряется величина колебания $\omega(\mathbf{F}, \delta(i, j))$, а вариация вычисляется как

$$P(\mathbf{F}, D) = \sum_{i=1}^{I-1} \sum_{j=1}^{J-1} \omega(\mathbf{F}, \delta(i, j)). \quad (6)$$

Согласно [20], колебание ω есть $\omega(\mathbf{F}, \delta(i, j)) = \max\{f(x, y) \in \delta(i, j)\} - \min\{f(x, y) \in \delta(i, j)\}$, что в дискретном случае можно определить как

$$\omega(\mathbf{F}, \delta(i, j)) = \max\{|f(i+1, j) - f(i, j)|, |f(i, j+1) - f(i, j)|\}.$$

Тонелли вариация [10, 19] определяется как сумма одномерных вариаций (3) по всем значениям i и j каждой из координат при постоянном значении второй координаты. Пусть $\psi_1(i)$

— одномерная вариация функции $f(i, j)$ по j при $i = \text{const}$, а $\psi_2(j)$ — одномерная вариация функции $f(i, j)$ по i при $j = \text{const}$. Тогда

$$T(\mathbf{F}, D) = \sum_{i=1}^I \psi_1(i) + \sum_{j=1}^J \psi_2(j). \quad (7)$$

Нетрудно убедиться, что значения, вычисляемые по формулам (7) и (4) будут равны, т.е. вариации Арцела и Тонелли для дискретных функций совпадают.

Фреше вариация [19] в формулировке для двумерной дискретной функции совпадает с *Харди вариацией* [10, 19] и вычисляется следующим образом:

$$\begin{aligned} \Delta_1(f; (i, j)) &= f(i+1, j) - f(i, j), \\ \Delta_2(f; (i, j)) &= \Delta_2(\Delta_1(f; (i, j))) = (f(i+1, j+1) - f(i, j+1)) - (f(i+1, j) - f(i, j)), \end{aligned}$$

$$F(\mathbf{F}, D) = H(\mathbf{F}, D) = \sum_{i=1}^{I-1} \sum_{j=1}^{J-1} |\Delta_2(\Delta_1(f; (i, j)))|. \quad (8)$$

Как легко заметить, формулы (8) и (5) совпадают, т.е. вариации Витали, Фреше и Харди для дискретных функций будут между собой равны, в то время как для непрерывных функций они различаются.

Рассмотренные вариации так или иначе имеют в своей основе значение модуля градиента функции в точке, поэтому их величины и поведение близки между собой.

Кронрода вариация [9] в ряду вариаций многомерных функций стоит особняком по той причине, что не использует градиентные характеристики функции, а также дает в результате значения не одного, а нескольких функционалов. Ее рассмотрению посвящен следующий раздел.

4. ДВУМЕРНАЯ ВАРИАЦИЯ КРОНРОДА

Обобщения отдельных выводов и теорем, которые удавалось формулировать используя определения вышеперечисленных вариаций (для непрерывных функций), привели к выводу, что функция многих переменных должна характеризоваться не одним, а несколькими функционалами, которые в определенном смысле независимы. Данный тезис был обоснован А.С. Кронродом при изучении функций двух переменных [9]. Основываясь на понятии множеств уровня он предложил для функции двух переменных использовать два функционала, определяемых следующим образом:

$$w_1(f) = \int_{-\infty}^{\infty} v_0(e_t) dt, \quad w_2(f) = \int_{-\infty}^{\infty} v_1(e_t) dt. \quad (9)$$

Здесь множество e_t — это t -уровень функции $f(x, y)$, т.е. множество тех точек (x, y) , в которых $f(x, y) = t$; $v_0(e_t)$ — число компонент множества e_t ; $v_1(e_t)$ — длина e_t (по Хаусдорфу). Тем самым значение w_1 характеризует число локальных экстремумов функции.

А.Г. Витушкиным было найдено удачное определение вариации множества, причем данный подход, также основанный на рассмотрении множеств уровня, был расширен для функций многих переменных [10]. Для плоского множества e_t , которое является t -уровнем функции $f(x, y)$, вариация задается не одним, а тремя значениями: v_0 , v_1 и v_2 , которые определяются следующим образом.

Пусть E — замкнутое ограниченное множество на плоскости; тогда вариация $v_0(E)$ есть число компонент E , вариация $v_2(E)$ есть суммарная площадь компонент E , а значение вариации $v_1(E)$ определяется следующим выражением [10, 11]:

$$v_1(E) = c \int_0^{2\pi} v_0(E, L_\alpha) d\alpha, \quad \text{где} \quad v_0(E, L_\alpha) = \int_{L_\alpha} v_0(E \cap (L_\alpha)^\perp) dz.$$

Здесь L_α — прямая $x \cos \alpha + y \sin \alpha = 0$, а $(L_\alpha)^\perp$ — прямая, проходящая через точку $z \in L_\alpha$ перпендикулярно к L_α . Константа c выбирается таким образом, чтобы для отрезка I единичной длины на оси Ox выполнялось соотношение: $v_1(I) = 1$. Для случая конечной $v_0(e_t)$, вариацию $v_1(e_t)$ можно интерпретировать как длину границ компонент e_t . Вариации v_0 , v_1 и v_2 названы соответственно *нулевой*, *линейной* и *плоской* согласно тому, как изменяются их значения при гомеоморфизме: v_0 не меняется, v_1 меняется линейно с изменением коэффициента растяжения K , а v_2 меняется пропорционально K^2 , т.е. как площадь плоской фигуры.

Важно отметить, что вариация v_0 множества уровня e_t , входящая в определение вариации $w_1(f)$, является не метрической, а в определенном смысле топологической характеристикой функции, основанной на понятии связности и не изменяющейся при гомеоморфизме. Касательно значения вариации $w_2(f)$, в [7] было показано, что для непрерывно дифференцируемой на D функции $f(x, y)$ справедливо равенство

$$w_2(f) = \iint_D |\text{grad}(f(x, y))| dx dy,$$

т.е. $w_2(f)$ является интегралом градиента функции по области носителя. В нашем случае $f(x, y)$ есть значение сигнала (яркости) изображения, которое можно считать неотрицательным и ограниченным, а значит пределы интегрирования в (9) можно ограничить диапазоном $[0, T]$, где T — максимальное допустимое значение.

В отличие от v_0 и v_1 , характеристика $v_2(e_t)$ не представляет интереса для оценки сложности функции. Рассматривая $v_2(e_t)$ как сумму площадей всех компонент сечения уровня t функции яркости $f(x, y)$, разделенных границами, очевидно, что $v_2(e_t) = S(D)$, т.е. равно площади всего множества D , и тем самым дополнительной информации не несет.

Можно предположить, что полезным окажется использование подмножества $\tilde{v}_2(e_t)$ ($\tilde{v}_2(e_t) \subset v_2(e_t)$), состоящего, скажем, только из тех точек (x, y) , для которых выполняется $f(x, y) \geq t$. По аналогии с (9) определим третий функционал как интеграл плоской вариации

$$w_3(f) = \int_0^T \tilde{v}_2(e_t) dt,$$

значение которого будет равно объему трехмерной фигуры, ограниченной функцией $f(x, y)$ и плоскостью $z(x, y) = 0$. Поскольку $f(x, y)$ задана на ограниченном двумерном множестве $D \ni (x, y)$, справедливо выражение $\tilde{v}_2(e_t)/S(D) = 1 - F_f(t)$, где $S(D)$ — площадь D , а $F_f(t)$ — функция распределения значений $f(x, y)$. Отсюда

$$\tilde{v}_2(e_t) = S(D) \left(1 - \int_0^t h(z) dz \right) = S(D) \int_t^T h(z) dz,$$

где $h(z) = p\{f(x, y) = z\}$ есть плотность вероятности значений функции $f(x, y)$. Легко видеть, что при этом соотношение $w_3(f)/S(D)$ будет равно среднему значению $f(x, y)$ на D , или, в интерпретации изображения, его средней яркости. Это означает, что функционал $w_3(f)$ также

не содержит информации о структуре $f(x, y)$. Тем самым, для оценки сложности функции $f(x, y)$ остаются только две вышеприведенные вариации $w_1(f)$ и $w_2(f)$.

В случае дискретной функции $f(i, j)$, двумерные вариации (9) выражаются следующими формулами:

$$w_1(f) = \sum_{x=0}^{T-1} v_0(e_t)/T \quad \text{и} \quad w_2(f) = \sum_{x=0}^{T-1} v_1(e_t)/T. \quad (10)$$

где T — общее число возможных значений функции $f(i, j)$ (для изображения — число градаций яркости). При вычислении вариаций, в качестве множества e_t выступает бинарная матрица $\mathbf{B} = \{b(i, j)\}$, строящаяся следующим образом:

$$b(i, j) = \begin{cases} 0, & \text{если } f(i, j) < t, \\ 1, & \text{если } f(i, j) \geq t. \end{cases}$$

Дискретизация изображения осуществляется, как правило, на квадратной решетке. При этом значение $v_0(e_t)$ определяется как сумма числа связных компонент, составленных из соседствующих (в смысле 4-соседства [21]) элементов изображения со значением $b(i, j) = 0$ и числа аналогичных связных компонент из элементов с $b(i, j) = 1$. Значение $v_1(e_t)$ есть суммарная длина границ компонент $v_0(e_t)$. В случае дискретной функции $f(i, j)$ значение $v_1(e_t)$ зависит от способа вычисления длины границы. На квадратной решетке длину границ чаще всего измеряют в метрике L_1 , определяя $v_1(e_t)$ как суммарное число сегментов решетки, разделяющих элементы с несовпадающими значениями.

Формальное вычисление значений вариаций согласно (10) на ограниченном носителе D предполагает, что сама область D также учитывается в качестве отдельной компоненты. Это приводит к следующему противоречию. Вариации w_1 и w_2 для функции постоянного значения $f(i, j) = \text{const}$, для которой они очевидно должны быть равны нулю, оказываются ненулевыми: $w_1(f) = 1$, а $w_2(f) = P(D)$ — периметру области D . Для устранения данного противоречия следует модифицировать формулы (10) следующим образом:

$$w_1(f) = \left(\sum_{x=0}^{T-1} v_0(e_t)/T \right) - 1 \quad \text{и} \quad w_2(f) = \left(\sum_{x=0}^{T-1} v_1(e_t)/T \right) - P(D). \quad (11)$$

В дальнейшем значения w_1 и w_2 для дискретной функции $f(i, j)$ на ограниченном носителе будут пониматься именно в смысле (11). Будучи нормированным на величину диапазона T , значение w_1 является характеристикой числа и амплитуды объектов изображения, образующих в сечениях отдельные компоненты. Поэтому удобно назвать значение w_1 *показателем числа объектов* изображения. Можно также показать, что значение второй вариации $w_2(f)$ в (11) будет совпадать со значением дискретной вариации Тонелли (7).

5. ПОКАЗАТЕЛЬ РАЗМЕРОВ ОБЪЕКТОВ ИЗОБРАЖЕНИЯ

Получаемые значения вариаций Кронрода позволяют оценить некоторые параметры изображения, которые отражают как амплитудные, так и морфологические его характеристики. Значение первой вариации w_1 отражает число и контраст деталей на изображении, а значение второй, w_2 , — сумму периметров деталей. Наряду с этим, важным выглядит соотношение вариаций $q_w(f) = w_2(f)/w_1(f)$, которое отражает средний периметр деталей на изображении (здесь предполагаем, что $w_1 > 0$, иначе $q_w(f) = 0$).

В дискретном изображении наименьшим возможным объектом является один элемент (пиксель). Это значит, что минимально возможная компонента, из числа входящих в множество

уровня e_t , имеет линейный размер равный одному шагу дискретизации. Периметр такой компоненты из одного элемента будет равен 4 — значению, которое есть аналог числа π в метрике L_1 и является минимально возможным для соотношения периметр/площадь объекта. Предполагая, что в дискретном случае длина границ компонент, составляющих $v_1(e_t)$, измеряется в метрике L_1 , очевидно, что и для $q_w(f)$ также существует минимально возможное значение, которое тоже равно 4, причем минимум достигается в случае, когда все компоненты множеств уровня e_t для всех t являются одноэлементными. В связи со сказанным удобно ввести характеристику

$$d(f) = q(f)/4, \quad (12)$$

которую по аналогии с показателем числа объектов w_1 в (11) естественно назвать *показатель размеров объектов* изображения; для нее будет выполняться соотношение $d(f) \geq 1$.

6. ДВУМЕРНАЯ ВАРИАЦИЯ КАК ОЦЕНКА СЛОЖНОСТИ ИЗОБРАЖЕНИЯ

Для изображений с преобладанием мелких деталей, которые на множествах уровня отображаются компонентами небольших размеров, значения d будут малыми, а для изображений с крупными деталями — большими. Отметим, что к деталям изображения относятся также и шумы, являющиеся локальными выбросами сигнала, и проявляющиеся на множествах уровня как самостоятельные компоненты с минимально возможным периметром. В связи с этим значение d должно убывать при увеличении числа и амплитуды шумовых выбросов. Ниже этот факт подтверждается экспериментально.

Особый интерес представляет поведение значений $w_1(f)$ и $d(f)$ при возможных преобразованиях функции $f(x, y)$. Для некоторых важных частных случаев можно сформулировать следующие утверждения.

1. Значения w_1 и d не зависят от выбора системы координат или ее поворота относительно носителя D .
2. При линейных амплитудных преобразованиях вида $Cf(x, y)$, где C — константа, значение w_1 увеличивается пропорционально C , а значение d не изменяется.
3. При линейном пространственном растяжении носителя D в K раз, значение w_1 не изменяется, а значение d увеличивается также в K раз.
4. Расширим D добавлением области U , на которой $f(x, y)$ имеет постоянное значение; это означает, что $f(x, y) = \text{const} \{ (x, y) \in U \subset \tilde{D} = D \cup U \}$. Условием выполнения операции расширения является то, что она не должна добавлять новых участков разрыва первого рода. Такое расширение не приводит к изменению значений w_1 и d .
5. При увеличении количества деталей на изображении (но при сохранении необходимых статистических соотношений и распределений) значение w_1 пропорционально увеличивается, а d не изменяется.
6. Пусть область D прямоугольна, и по выбранной оси, скажем оси Y , ограничена отрезком $[a, b]$. Расширим ее областью U , заданной на полуотрезке $(b, 2b - a]$, на котором $f(x, y)$ имеет зеркальное продолжение, т.е. $f(x, b + y) = f(x, b - y)$. При таком расширении значение w_1 увеличивается пропорционально изменению площади, тогда как значение d не меняется.

Таким образом, при указанных изменениях двумерной функции $f(x, y)$, как минимум одно из значений w_1 и d оказывается инвариантом, второе же значение является предсказуемой функцией преобразования. Применительно к изображениям утверждение 2 соответствует линейному изменению контраста; утверждение 3 — линейной геометрической трансформации; утверждение 4 — добавлению/удалению участков изображения с ровным фоном без деталей; утверждение 5 близко предыдущему и соответствует концентрации/разреженности деталей на изображении; утверждение 6 — вариант зеркальной пролонгации изображения за рамки

области D , часто используемый алгоритмами локального анализа. Кроме того, повторение свойств части изображения на всем изображении по существу означает однородность (стационарность), т.е. достаточность определения значений w_1 и d на доверительном участке и интерпретации их для изображения в целом.

Отметим важную особенность введенного показателя размеров объектов $d(f)$, имеющую следствие в утверждениях 2 и 5. Характеристика $d(f)$ обладает тем свойством, что отражает лишь средние размеры объектов на изображении и не зависит от количества и контраста самих объектов — эту часть информации несет вариация $w_1(f)$.

Из сказанного ясно, что первая вариация $w_1(f)$ вместе с характеристикой $d(f)$ отражают пространственную и яркостную изменчивость изображения и в совокупности могут служить показателями *сложности изображения* в соответствии с требованиями, сформулированными во введении. Различное поведение указанных характеристик подтверждает ту изначальную гипотезу А.С. Кронрода, что для описания изменчивости двумерных функций какой-то одной вариации недостаточно и необходимы два независимых функционала.

7. О СРАВНЕНИИ И ИНТЕРПРЕТАЦИИ ОЦЕНОК

Поскольку характеристик сложности оказывается две, причем они вообще говоря независимы, то возникает естественный вопрос, как с их помощью сравнивать различные изображения между собой. Достаточно очевидно, что если характеристика $d(f)$ постоянна, то с увеличением $w_1(f)$ сложность изображения будет возрастать; также сложность будет возрастать и с уменьшением $d(f)$ при постоянном $w_1(f)$. Тогда, если для пары изображений f_1 и f_2 выполняются соотношения $w_1(f_1) < w_1(f_2)$ и $d(f_1) > d(f_2)$, то естественно полагать, что изображение f_2 является более сложным, чем f_1 . Остается вопрос, как сравнивать другие случаи.

Рассмотрим изображение с малым значением w_1 и большим значением d ; такие соотношения означают, что на изображении имеется малое число деталей больших размеров (с большими периметрами). Наоборот, большое значение w_1 и малое значение d свидетельствуют, что изображение содержит много мелких деталей. В подобных случаях обычно принято говорить, что оценка сложности зависит от конкретной задачи.

В некоторых случаях оптимизация в многомерном пространстве затруднительна и приходится из нескольких функционалов тем или иным образом формировать обобщенное значение. Одним из простейших вариантов является следующая линейная комбинация:

$$w(z) = w_1(z) + Cw_2(z) = w_1(z)(1 + Cd(z)), \quad (13)$$

которая может, например, применяться в (1) в качестве стабилизирующего функционала $\Omega(z)$.

8. ЭКСПЕРИМЕНТАЛЬНЫЕ ИССЛЕДОВАНИЯ

Для экспериментов был выбран набор из реальных и синтезированных изображений различной детальности и тематического содержания. Часть изображений представлена на Рис. 1: (а) — компьютерный рисунок с протяженными кусочно-постоянными объектами; (б) — ровный фон с малоразмерными объектами случайной яркости; (в), (г) — портреты с мелкими и крупными деталями; (д), (е), (ж) — городской ландшафт, аэрофотоснимок и натюрморт; (з) — компьютерная томограмма.

В Таблицах 1 и 2 представлены значения двумерных вариаций для тестовых изображений на Рис. 1. Двумерные вариации Арцела (A), Тонелли (T), Пьерпонта (P), Витали (V), Фреше (F), Харди (H), Кронрода ($K-w_1$ и $K-w_2$) измерялись согласно формулам (4)–(8) и (11). Приведены также соответствующие показатели размеров объектов (d) согласно (12) и значения

энтропии (E). Таблица 1 демонстрирует, как изменяются указанные значения для изображения на Рис. 1,а при нанесении гауссова шума $N(0, \sigma^2)$ с СКО $\sigma = 1; 2; 3$ и 5% диапазона значений яркостей.

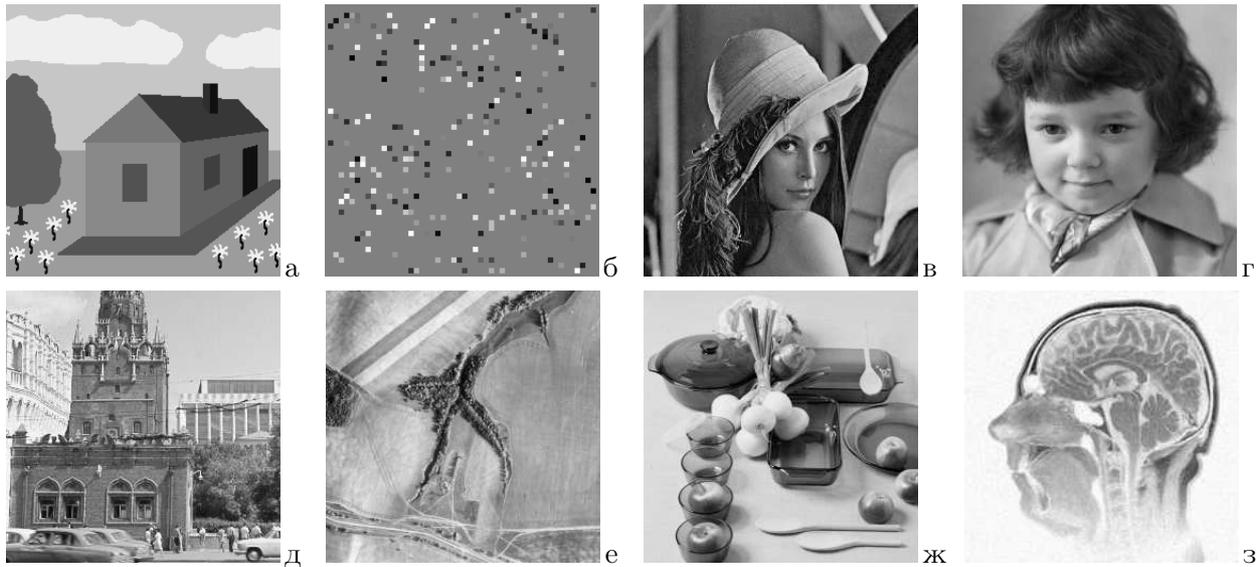


Рис. 1. Тестовые изображения.

Для удобства анализа полученных данных, значения из Таблицы 1 также представлены графически на Рис. 2 и 3. Из приведенных данных и графиков можно сделать несколько выводов: а) в подтверждение сказанному выше, значения вариаций Арцела / Тонелли (A/T) и второй вариации Кронрода ($K-w_2$) с большой степенью точности совпадают; б) с увеличением уровня шума значения всех вариаций растут практически линейно; в) энтропия асимптотически возрастает до предельного значения 8, равного числу битов в записи значений градаций яркости, а показатель размеров объектов (d) убывает сначала резко, затем медленно, стремясь к предельному значению 1.

Таблица 1. Значения двумерных вариаций, показателя размеров объектов (d), энтропии (E), а также их зависимости от уровня добавленного шума для изображения на Рис. 1,а.

Шум, %	A/T	P	$V/F/H$	$K-w_1$	$K-w_2$	d	E
$\sigma = 0$	1080,58	945,73	503,35	9,79	1081,64	27,62	2,97
$\sigma = 1$	1641,15	1327,58	909,12	110,27	1644,71	3,73	5,06
$\sigma = 2$	2198,37	1696,72	1303,17	210,29	2204,56	2,62	5,94
$\sigma = 3$	2758,17	2072,65	1697,21	309,89	2766,76	2,23	6,43
$\sigma = 5$	3867,95	2816,74	2487,14	511,75	3892,58	1,90	7,01

Таблица 2. Значения двумерных вариаций, показателя размеров объектов (d) и энтропии (E) для реальных изображений на Рис. 1,в-з.

Изображение	A/T	P	$V/F/H$	$K-w_1$	$K-w_2$	d	E
в	5242,72	3715,70	2859,43	469,85	5260,25	2,80	7,59
г	2583,56	1825,14	1038,43	129,23	2589,58	5,01	7,60
д	7928,75	6072,46	3497,91	618,79	7947,70	3,21	7,59
е	4178,20	2954,15	1981,67	332,67	4184,39	3,16	7,19
ж	3256,29	2389,90	1562,20	256,84	3260,76	3,18	7,34
з	5107,64	3642,48	2499,84	419,62	5118,65	3,05	6,76

Таблица 2 является продолжением Таблицы 1; в ней представлены значения двумерных вариаций для реальных изображений, показанных на Рис. 1, в-з. Видно, что значения одних и тех же вариаций для разных изображений отличаются в 4 и более раз, тогда как значения энтропии изменяются очень слабо. Это показывает, что энтропия не может являться подходящей оценкой сложности изображения. Сравнение значений еще раз подтверждает достаточно точное совпадение вариаций Арцела / Тонелли (A/T) и второй вариации Кронрода ($K-w_2$) для всех изображений.

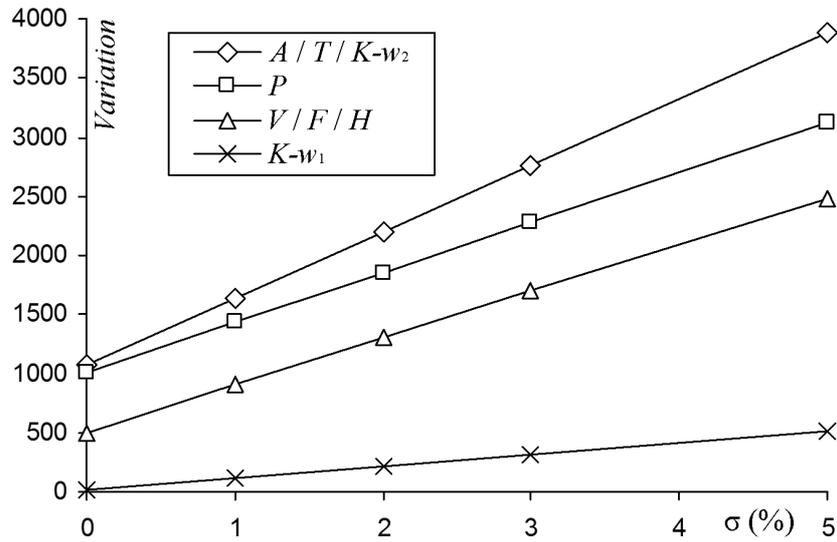


Рис. 2. Зависимость значений вариаций от СКО нанесенного гауссова шума для изображения на Рис. 1, а.

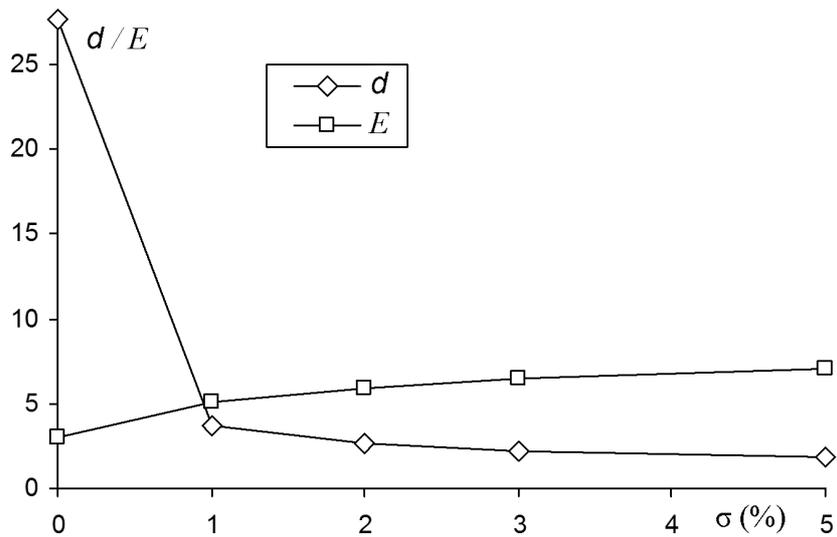


Рис. 3. Зависимость значений показателя размеров объектов (d) и энтропии (E) от СКО (σ , %) нанесенного гауссова шума для изображения на Рис. 1, а.

На Рис. 4 и 5 показаны зависимости значений двумерных вариаций от размеров деталей и их поворота. Для этого было сформировано несколько изображений с постоянным фоном и одинаковым числом прямоугольных объектов случайной яркости, имеющих размеры от 1x1 до 5x5 элементов; последнее из них показано на Рис. 1,б. Значения вариаций и показателей размеров объектов измерялись как для изображений с нормальным положением объектов, когда их стороны параллельны строкам и столбцам, так и для изображений с повернутыми на 45° объектами без интерполяции значений элементов, но при соответствующем приближенном проведении прямых контурных линий.

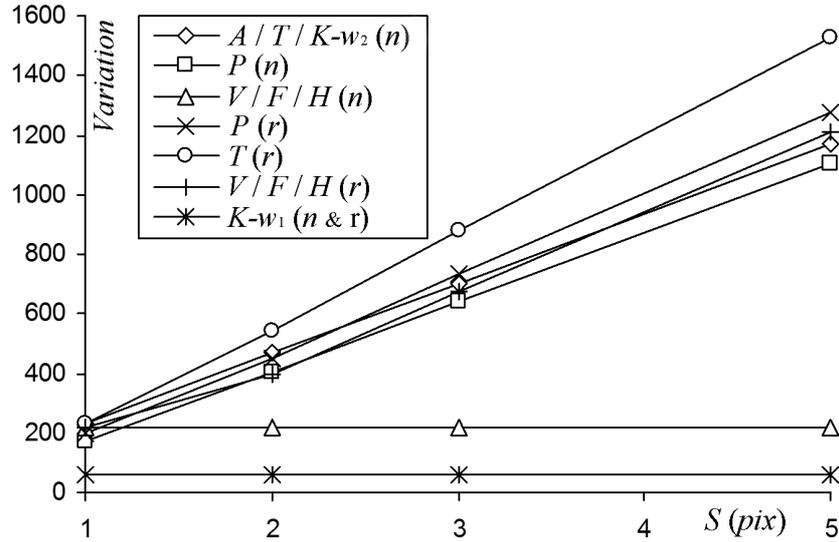


Рис. 4. Зависимость значений вариаций от размеров объектов S (в элементах) в нормальном положении (n) и после поворота (r) для изображения на Рис. 1,б.

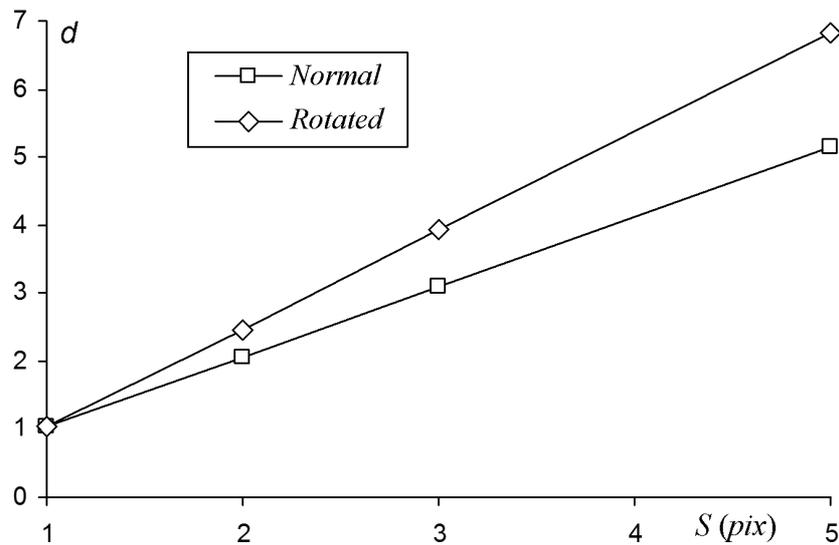


Рис. 5. Зависимость показателя размеров объектов d от размеров объектов S (в элементах) и от их поворота для изображения на Рис. 1,б.

На Рис. 4 хорошо видно, что график $K-w_1(n&r)$ горизонтален и совпадает для нормальных и повернутых объектов. Это подтверждает, что показатель числа объектов w_1 инвариантен как к изменению размеров объектов, так и к их повороту. Постоянство графика значений $V/F/H(n)$, который соответствует объектам в нормальной ориентации, подтверждает сделанный выше вывод, что локальные операторы, используемые в дискретных аналогах вариаций Витали / Фреше / Харди, дают нулевые отклики на тех участках функции, где она имеет перепад значений в направлении лишь одной из координатных осей. Если же контурные линии не вертикальны или не горизонтальны, то такой особенности не наблюдается (график $V/F/H(r)$). Графики всех остальных вариаций расположены достаточно плотным пучком, что говорит о том, что соответствующие значения вариаций различаются слабо.

Как видно из графиков на Рис. 5, показатель размеров объектов d изменяется пропорционально размерам объектов, тем самым иллюстрируя утверждение 3, что значение этой характеристики линейно зависит от коэффициента растяжения носителя. Тот факт, что график, соответствующий повернутым объектам, имеет несколько большую крутизну, противоречит тем самым утверждению 1, объясняется тем, что длина контура повернутого дискретного объекта, измеряемая в метрике L_1 , оказывается больше, чем длина контура объекта с вертикальным и горизонтальным расположением краев.

9. ВЫВОДЫ

Исследована возможность оценки сложности изображения при помощи двумерных вариаций. Показана редукция многомерных вариаций непрерывных функций к двумерным дискретным функциям. В данном контексте проведено сравнение различных двумерных вариаций; подтверждено преимущество использования двумерной вариации Кронрода. На основе двумерных вариаций введена характеристика, соответствующая среднему значению периметров деталей на изображении. Предложена модификация данной характеристики, названная *показатель размеров объектов*.

Проанализированы изменения значений вариаций при различных возможных трансформациях изображения. Теоретические выводы подтверждены экспериментальными исследованиями. Продемонстрировано влияние шума и простейших геометрических трансформаций на сложность изображения, которая интерпретируется как характеристика, отражающая число, размеры и заметность деталей на изображении.

Предложен вариант редукции двух значений вариаций к единственному обобщенному функционалу, который может использоваться в классических оптимизационных схемах.

Проведенные исследования подтверждают, что комбинация показателя числа объектов w_1 и показателя размеров объектов d отражает морфологическую структуру изображения и поэтому может служить оценкой его сложности.

Работа выполнена за счет гранта Российского научного фонда (проект № 14-50-00150).

СПИСОК ЛИТЕРАТУРЫ

1. Wang Z., Bovik A.C., Sheikh H.R., Simoncelli E.P. Image quality assessment: from error visibility to structural similarity. // IEEE Transactions on Image Processing, 2004, vol. 13, no. 4, pp. 600–612.
2. Wang Z., Bovik A.C. Modern image quality assessment. N.Y.: Morgan and Claypool, 2006.
3. Yu H., Winkler S. Image complexity and spatial information. // Proc. 5th International Workshop on Quality of Multimedia Experience (QoMEX), 2013, Klagenfurt, Austria, pp. 12–17.
4. Peters R.A., Strickland R.N. Image Complexity Metrics for Automatic Target Recognizers // Automatic Target Recognizer System and Technology Conference, 1990, Silver Spring, MD, pp. 30–31.

5. Тихонов А.Н., Арсенин В.Я. Методы решения некорректных задач. М.: Наука, 1979.
6. Rudin L.I., Osher S., Fatemi E. Nonlinear total variation based noise removal algorithms. // *Physica D: Nonlinear Phenomena*, 1992, vol. 60, pp.259–268.
7. Милюкова О.П. Изображение как функция с ограниченной полной вариацией. Иконика. Цифровая обработка видеоинформации. М.: Наука, 1989, С. 19–25.
8. Adams C.R., Clarkson J.A. On definitions of bounded variation for functions of two variables, *Trans, of Am. Math. Soc.*, 1933, vol. 36, pp. 824.
9. Кронрод А.С. О функциях двух переменных. УМН, 1950, т. 5, № 1, с. 24–134.
10. Витушкин А.Г. О многомерных вариациях. М.: Гостехиздат, 1955.
11. Иванов Л.Д. Вариации множеств и функций. М.: Наука, 1975.
12. Chochia P.A., Milukova O.P. Two–Dimensional Variation and Image Decomposition // 6th International Workshop „Image Processing and Computer Optics (DIP-97)“. *Proceedings SPIE*, 1998, vol. 3346, pp. 329–339.
13. Милюкова О.П., Чочиа П.А. Об оценке сложности изображений с помощью двумерных вариаций. // *Информационные процессы*, 2012, Т. 12, № 4, С. 362–371.
14. Варакин Л.Е. Теория сложных сигналов. М.: Сов. радио, 1970.
15. Харкевич А.А. Спектры и анализ. М.: Гостехиздат, 1957.
16. Витушкин А.Г. Оценка сложности задачи табулирования. М.: Физ–мат. лит., 1959.
17. Колмогоров А.Н. Три подхода к определению понятия «количество информации». // *Проблемы передачи информации*, 1965, Т. 1, вып. 1, С. 3–11.
18. Чочиа П.А. Двухмасштабная модель изображения // *Кодирование и обработка изображений*. М.: Наука, 1988, С. 69–87.
19. Математическая энциклопедия. М.: Советская энциклопедия, 1977, Т. 1–5.
20. Pierpont J. *Lectures on the theory of functions of real variables*. Vol. 1. N.Y.: Dover Publications, 1959.
21. Гонсалес Р., Вудс Р. *Цифровая обработка изображений*. М.: Техносфера, 2012.

Two–dimensional variations comparison in the context of digital image complexity assessment

Chochia P. A. Milukova O. P.

The questions of complexity of digital images are investigated. Different known two–dimensional variation methods are considered in their application to digital images. The possibilities of using two–dimensional variations for estimating the *image complexity* are studied. The image complexity is interpreted as some attribute which is specified by the quantity, the sizes and the visibility of image details. The modified assessment is proposed, that is named as *component size index*. The variation changes under different image transformations are analyzed. Theoretical conclusions are confirmed by experimental explorations.

KEYWORDS: two–dimensional variation, image complexity, image processing.