

Цифровые изображения на комплексном дискретном торе*

Мнухин В. Б.

mnukhin.valeriy@mail.ru

Южный Федеральный Университет

В работе предлагается алгебраический метод обработки цифровых изображений. Предлагается рассматривать изображения размера $p \times p$, где $p \geq 3$ — простое число вида $p = 4k - 1$, как функции на комплексном дискретном торе. Вводится понятие комплексного вращения и определяется новое обратимое преобразование, являющееся дискретным аналогом непрерывного преобразования Меллина. Строится модулярное преобразование Фурье–Меллина, инвариантное относительно циклических сдвигов, масштабирования и комплексных вращений цифровых изображений.

Ключевые слова: *Цифровое изображение, конечное поле, дискретный тор, дискретное преобразование, преобразование Фурье–Меллина, полярно-логарифмические координаты, модулярный логарифм.*

Digital images on a complex discrete torus*

Mnukhin V. B.

Southern Federal University

An algebraic method for digital images analysis and processing is considered. It is based on the presentation of digital images of size $p \times p$, where $p \geq 3$ is the prime number such that $p = 4k - 1$, as functions on a complex discrete torus. For such images, complex rotations are introduced and a new invertible linear transform, called the modular Mellin transform, is presented. The discrete modular Fourier–Mellin transform, invariant under circular shifts, scaling, and complex rotations, is also defined.

Keywords: *Digital image, finite field, discrete torus, discrete transform, Fourier–Mellin transform, log-polar coordinates, modular logarithm.*

Введение

В настоящее время разработка методов решения задач обработки, распознавания и классификации изображений проводится, как правило, в предположении непрерывности заданного изображения. Это позволяет использовать для решения поставленных задач мощный аппарат математического анализа и интегральных преобразований. Однако на практике применение полученных таким образом решений к цифровым изображениям неизбежно приводит к появлению систематических ошибок, связанных с дискретностью реальных цифровых изображений.

В качестве одного из примеров подобной ситуации укажем на хорошо известный метод преобразования Фурье–Меллина [1–3], активно используемый, в частности, при решении задач регистрации изображений [4, с. 674]. Метод основан на рассмотрении Фурье-спектра исходного изображения в полярно-логарифмической системе координат, после чего к преобразованному таким образом спектру вторично применяется преобразование Фу-

Работа выполнена при финансовой поддержке РФФИ, проекты № 11-07-00591 и № 13-07-00327.

рье. Нетрудно показать, что в *непрерывном случае* на базе такого преобразования легко определяются характеристики изображений, инвариантные относительно сдвигов, вращений и масштабирования. Вместе с тем в *дискретном случае* формальное использование этого метода сопряжено с существенными погрешностями [2, 5], связанными, в частности, с невозможностью адекватного преобразования цифровых изображений в полярную и полярно-логарифмическую системы координат.

В связи с этим возникает задача разработки методов, изначально ориентированных на цифровые изображения и использующих аппарат алгебры и теории чисел. В данной работе цифровые изображения размера $p \times p$, где $p \geq 3$ — простое число вида $p = 4k - 1$, предлагается рассматривать как функции на комплексных дискретных торах, которые можно считать «конечными комплексными плоскостями». Для таких изображений вводится понятие комплексного вращения, и строятся дискретные аналоги преобразований Меллина и Фурье–Меллина, позволяющие получать инварианты цифровых изображений относительно циклических сдвигов, масштабирования и комплексных вращений. Сохранение ортогональности при комплексных вращениях позволяет эффективно использовать их, в частности, в задачах анализа симметрии цифровых изображений [6, 7].

Комплексный дискретный тор

Кольцо классов вычетов по модулю целого $n > 1$ условимся обозначать как $\mathbb{Z}_n = \mathbb{Z}/n\mathbb{Z}$.

Определение 1. Пусть $m > 1$ и $n > 1$ — два целых числа. Множество

$$\mathbb{Z}_m \times \mathbb{Z}_n = \{(x, y) : x \in \mathbb{Z}_m, y \in \mathbb{Z}_n\}$$

будем называть *дискретным $m \times n$ -тором*. Пары $(x, y) \in \mathbb{Z}_m \times \mathbb{Z}_n$ будем называть *пикселями*, а классы вычетов x и y — координатами данного пикселя. Рисунок 1 иллюстрирует введенное понятие.

Очевидно, что любой тор $\mathbb{Z}_m \times \mathbb{Z}_n$ является кольцом относительно естественных операций $+$ и $\odot$ покомпонентного сложения и умножения: если $(a, b) \in \mathbb{Z}_m \times \mathbb{Z}_n$ и $(c, d) \in \mathbb{Z}_m \times \mathbb{Z}_n$, то

$$(a, b) + (c, d) = \left(\underset{m}{a + c}, \underset{n}{b + d} \right), \quad (a, b) \odot (c, d) = \left(\underset{m}{a \cdot c}, \underset{n}{b \cdot d} \right).$$

Вместе с тем нетрудно заметить, что некоторые торы обладают особыми свойствами. Например, для простого p тор $\mathbb{Z}_p^2 = \mathbb{Z}_p \times \mathbb{Z}_p$ является векторным пространством размерности 2 над конечным полем $\mathbb{Z}_p$.

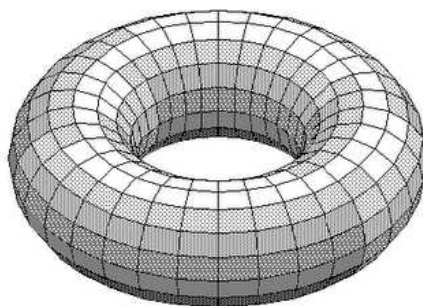


Рис. 1: Дискретный тор

Покажем, что в некоторых случаях на торе $\mathbb{Z}_p^2$ можно ввести структуру конечного поля со свойствами, аналогичными свойствам поля комплексных чисел $\mathbb{C}$. Вспоминая [8, с. 438]

построение кольца Гауссовых целых чисел $\mathbb{Z}[\sqrt{-1}]$, рассмотрим факторкольцо

$$\mathbb{C}_p = \mathbb{Z}_p[x]/(x^2 + \bar{1})$$

кольца многочленов над конечным полем $\mathbb{Z}_p$ по главному идеалу, порожденному двучленом $x^2 + \bar{1}$. Как очевидно, $\mathbb{C}_p$ содержит p^2 элементов вида $a + bi$, где $a, b \in \mathbb{Z}_p$, а i обозначает класс вычетов $\bar{x}$, так что $i^2 = \bar{-1} = \overline{p-1} \in \mathbb{Z}_p$.

Допустим, что простое число $p \geq 3$ имеет вид $p = 4k - 1$, (где $k \in \mathbb{Z}$), или, другими словами, что $p \equiv 3 \pmod{4}$. (Таковы, в частности, числа 71, 127, 251, 257 и 509.) Тогда, как известно ([9], с. 176), в поле $\mathbb{Z}_p$ элемент $\bar{-1} \in \mathbb{Z}_p$ не является квадратом, так что двучлен $x^2 + \bar{1}$ неприводим в кольце $\mathbb{Z}_p[x]$. Следовательно, кольцо $\mathbb{C}_p$ является полем [10, с. 293], и, согласно фундаментальной теореме [8, с. 428] теории конечных полей, $\mathbb{C}_p \simeq \mathbb{GF}(p^2)$. Остается заметить, что естественное отображение

$$\mathbb{C}_p \ni a + ib \longleftrightarrow (a, b) \in \mathbb{Z}_p^2$$

переносит структуру поля на тор $\mathbb{Z}_p^2$, превращая его в «конечную комплексную плоскость».

Определение 2. Если $p \geq 3$ — простое число такое, что $p \equiv 3 \pmod{4}$, то дискретный тор $\mathbb{Z}_p^2$ со структурой введенного выше конечного поля, будем называть комплексным дискретным тором (КДТ) и обозначать $\mathbb{C}_p$.

(Мы постараемся избегать словосочетания «дискретный комплексный тор», поскольку понятие «комплексный тор» хорошо известно в математике и используется в ином смысле.)

Учитывая аналогию $\mathbb{C}_p$ с комплексным полем $\mathbb{C}$ и кольцом $\mathbb{Z}[\sqrt{-1}]$, естественно называть элементы конечного поля $\mathbb{C}_p$ дискретными комплексными числами и использовать при работе с ними стандартную терминологию комплексного анализа. В частности, если $z = a + ib \in \mathbb{C}_p$, то будем писать $\operatorname{Re}(z) = a$ и $\operatorname{Im}(z) = b$. Число $z^* = a - ib$ будем называть сопряженным к $z = a + ib$, а величину $N(z) = zz^* = a^2 + b^2 \in \mathbb{Z}_p$ называть нормой числа $z \in \mathbb{C}(p)$. (Отметим, что в отличие от комплексных чисел модуль $|z| = \sqrt{N(z)}$, вообще говоря, в $\mathbb{C}_p$ не определен.) Непосредственно проверяется, что $N(z_1 z_2) = N(z_1)N(z_2)$; таким образом, норма является мультипликативной функцией. Кроме того, заметим, что

$$N(z) = 0 \Leftrightarrow z = 0.$$

(Действительно, если $N(z) = a^2 + b^2 = 0$ и $b \neq 0$, то $(a/b)^2 = -1$, что невозможно по предположению.)

Условимся далее считать, что число p удовлетворяет условию определения 2. Кроме того, там, где это не может привести к недопониманию, условимся отождествлять классы вычетов поля $\mathbb{Z}_p$ с их представителями и опускать верхнюю черту над обозначениями классов.

Отметим, что дискретные комплексные числа допускают также матричное представление. Пусть $M_2(\mathbb{Z}_p)$ означает алгебру 2×2 -матриц над полем $\mathbb{Z}_p$. Рассмотрим следующее инъективное отображение $\mu : \mathbb{C}_p \rightarrow M_2(\mathbb{Z}_p)$:

$$\mu(a + bi) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}, \quad \text{и пусть} \quad P = \mu(\mathbb{C}_p) = \left\{ \begin{pmatrix} a & -b \\ b & a \end{pmatrix} : a, b \in \mathbb{Z}_p \right\} \subset M_2(\mathbb{Z}_p).$$

Утверждение 1. Биекция $\mu : \mathbb{C}_p \rightarrow P$ является изоморфизмом,

$$\mu(z + w) = \mu(z) + \mu(w), \quad \mu(zw) = \mu(z)\mu(w) \quad \text{для всех} \quad z, w \in \mathbb{C}_p.$$

При этом нормы дискретных комплексных чисел переходят в определители соответствующих матриц, $N(z) = \det \mu(z)$.

Доказательство полностью повторяет доказательство [8, с. 195] аналогичного утверждения для комплексных чисел.

Комплексные вращения цифровых изображений

Под *цифровым изображением* размера $m \times n$ условимся понимать комплекснозначную функцию $f(x, y) : \mathbb{Z}_m \times \mathbb{Z}_n \rightarrow \mathbb{C}$, определенную на всех пикселах тора $\mathbb{Z}_m \times \mathbb{Z}_n$. (Заметим, что такое определение позволяет считать изображениями и Фурье-образы. Поскольку тор является естественной областью определения дискретного преобразования Фурье [11, гл. 19], такой подход представляется полезным и будет использован ниже.)

На всяком торе $\mathbb{Z}_p^2$ определены следующие обратимые преобразования изображений:

— *циклический сдвиг* $\mathcal{T}_w$ на $w = (a, b) \in \mathbb{Z}_p^2$:

$$\mathcal{T}_w[f] = f(x - a, y - b);$$

— *масштабирование* $\mathcal{S}_a$ с коэффициентом $0 \neq a \in \mathbb{Z}_p$:

$$\mathcal{S}_a[f] = f(ax, ay);$$

— *отражения* $\mathcal{M}_x$ и $\mathcal{M}_y$:

$$\mathcal{M}_x[f] = f(-x, y), \quad \mathcal{M}_y[f] = f(x, -y).$$

В то же время задача введения для цифровых изображений математически корректного понятия *вращения* представляется нетривиальной. Обычно в этом качестве используют ([4], с. 390) преобразования вида

$$\widetilde{\mathcal{R}}_\alpha[f] = f([x \cos \alpha - y \sin \alpha], [x \sin \alpha + y \cos \alpha]),$$

где $[a]$ означает округление действительного числа a до ближайшего целого. К сожалению, преобразование $\widetilde{\mathcal{R}}_\alpha$ не коммутирует с дискретным преобразованием Фурье, что приводит к систематическим ошибкам при его использовании. Аналогичные проблемы возникают [5] при попытках введения на цифровых изображениях полярной и полярно-логарифмической систем координат.

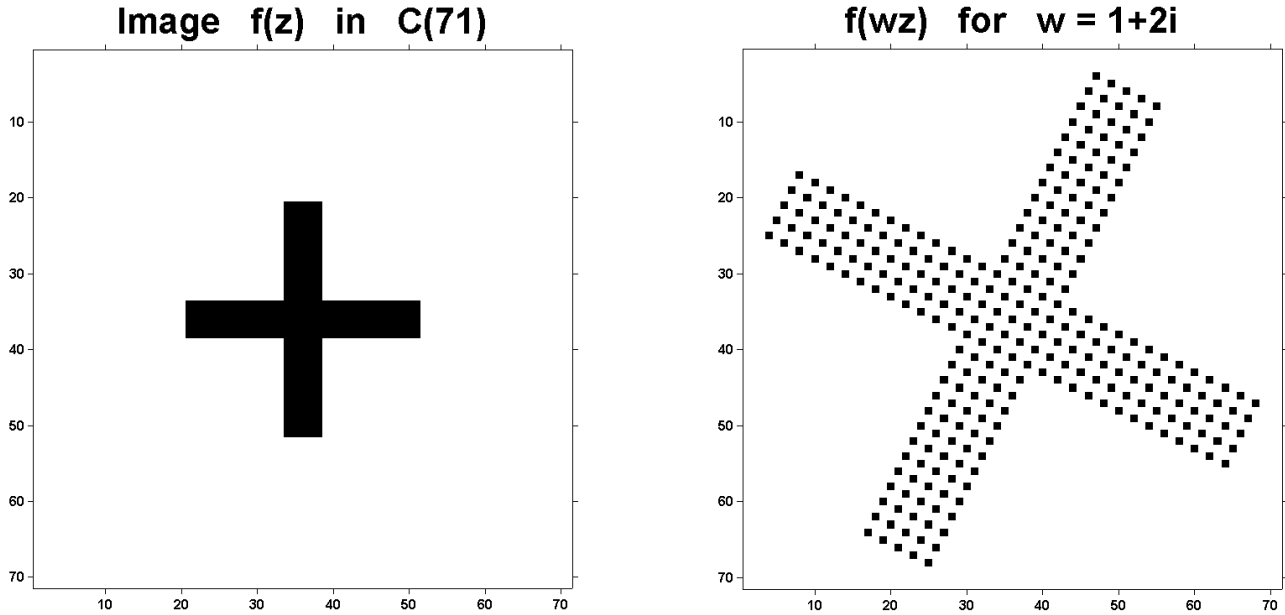
Рассмотрим теперь цифровые изображения на комплексном дискретном торе. Заметим, что вращение непрерывной функции комплексного переменного $f(z)$ на угол α в плоскости $\mathbb{C}$ естественно определяется как

$$\mathcal{R}_\alpha[f] = f(e^{i\alpha}z).$$

Поэтому, рассматривая изображения на КДТ как функции $f(z) : \mathbb{C}_p \rightarrow \mathbb{C}$ «дискретного комплексного переменного», для всякого ненулевого $w \in \mathbb{C}_p$ введем следующее обратимое преобразование изображения $f(z)$:

$$\mathcal{R}_w[f] = f(wz), \quad \mathcal{R}_w^{-1}[f] = f(w^{-1}z). \quad (1)$$

Определение 3. Преобразование $\mathcal{R}_w[f]$ будем называть *комплексным вращением* цифрового изображения f .

Рис. 2: Комплексное вращение на торе $\mathbb{C}_{71}$

Отметим, что масштабирование является частным случаем комплексного вращения, а сдвиги и отражения на КДТ принимают вид

$$\mathcal{T}_w[f] = f(z + w), \quad \mathcal{M}_y[f] = f(z^*), \quad \mathcal{M}_x[f] = f(-z^*).$$

Проиллюстрируем введенное понятие. Рисунок 2 показывает изображение креста на КДТ $\mathbb{C}_{71}$ и результат комплексного вращения этого изображения при $w = 1 + 2i$. Как нетрудно заметить, при этом «площадь увеличилась в 5 раз». В целом, преобразование можно рассматривать как непрерывное вращение на угол $\arctg 2$, совмещенное с масштабированием на $\sqrt{5}$.

Несмотря на то что на КДТ понятия угла, расстояния, площади и отношений порядка утрачивают привычный смысл, понятие ортогональности на нем сохраняется.

Определение 4. Два вектора $z_1 = (a, b)$ и $z_2 = (c, d)$ на комплексном дискретном торе $\mathbb{Z}_p^2$ назовем ортогональными, если $ac + bd = 0$ или, что равносильно, $\operatorname{Re}(z_1 z_2^*) = 0$.

Поскольку

$$\operatorname{Re}((wz_1)(wz_2)^*) = N(w) \operatorname{Re}(z_1 z_2^*),$$

немедленно получаем

Утверждение 2. Комплексные вращения сохраняют ортогональность: если z_1 и z_2 ортогональны, то wz_1 и wz_2 тоже ортогональны.

На рис. 3 показаны преобразования известного изображения из базы данных BRODATZ, редуцированного из исходного размера 512×512 в формат 251×251 и рассматриваемого на комплексном торе $\mathbb{C}_{251}$. Оригинал показан в левой верхней части рисунка.

В правой верхней части показан результат вращения при $w = 1 + i$. Поскольку при этом «площадь увеличивается» в 2 раза, части преобразованного изображения накладываются друг на друга, что визуально выглядит как его искажение. Важно отметить, что в действительности преобразованное изображение полностью сохраняет всю информацию об оригинале и является просто перестановкой его пикселей.

Результат вращения при $w_1 = (1 + 2i)^{-1} = -50 + 100i$ показан в левой нижней части рисунка. В этом случае «площадь уменьшается» в 5 раз, что выглядит как появление четырех уменьшенных версий исходного изображения и ряда фрагментов, в совокупности составляющих пятую. Заметим, что эти уменьшенные версии, визуально представляющие себя идентичными, в действительности составлены из различных пикселей оригинала.

В правой нижней части показан результат вращения на $w_2 = 9 + 13i$, когда $N(w_2) = 9^2 + 13^2 = 250 = -1 \in \mathbb{Z}_{251}$. Изображение состоит из 250 фрагментов оригинала, что объясняет визуальные искажения. Вместе с тем хорошо заметно, что фрагменты повернуты на угол $\arctg(13/9) \approx 55^\circ$.

Подводя итоги, назовем угол α *допустимым*, если $\operatorname{tg}(\alpha)$ является рациональным числом, $\operatorname{tg}(\alpha) \in \mathbb{Q}$. Тогда комплексные вращения можно рассматривать как повороты на допустимые углы с соответствующим масштабированием и с учетом периодичности на торе.

Отметим, что сохранение ортогональности при комплексных вращениях позволяет эффективно использовать их, в частности, в задачах анализа симметрии цифровых изображений [6, 7].

Модулярные логарифмы и преобразование Меллина

Воспользуемся аналогией между $\mathbb{C}_p$ и $\mathbb{C}$ для переноса понятия комплексного логарифма в смысле главного значения (или, что равносильно, полярно-логарифмической системы координат) на комплексный дискретный тор. Понятно, что это невозможно сделать с помощью «наивного» подхода, основанного на понятиях полярного угла и полярного радиуса. Поэтому обратимся к алгебраическим методам введения полярных координат на комплексной плоскости.

Вспомним [10], что мультипликативная группа $\mathbb{C}^*$ поля комплексных чисел $\mathbb{C}$ допускает разложение в прямое произведение

$$\mathbb{C}^* \simeq \mathbb{R}^+ \times U(1) \quad (2)$$

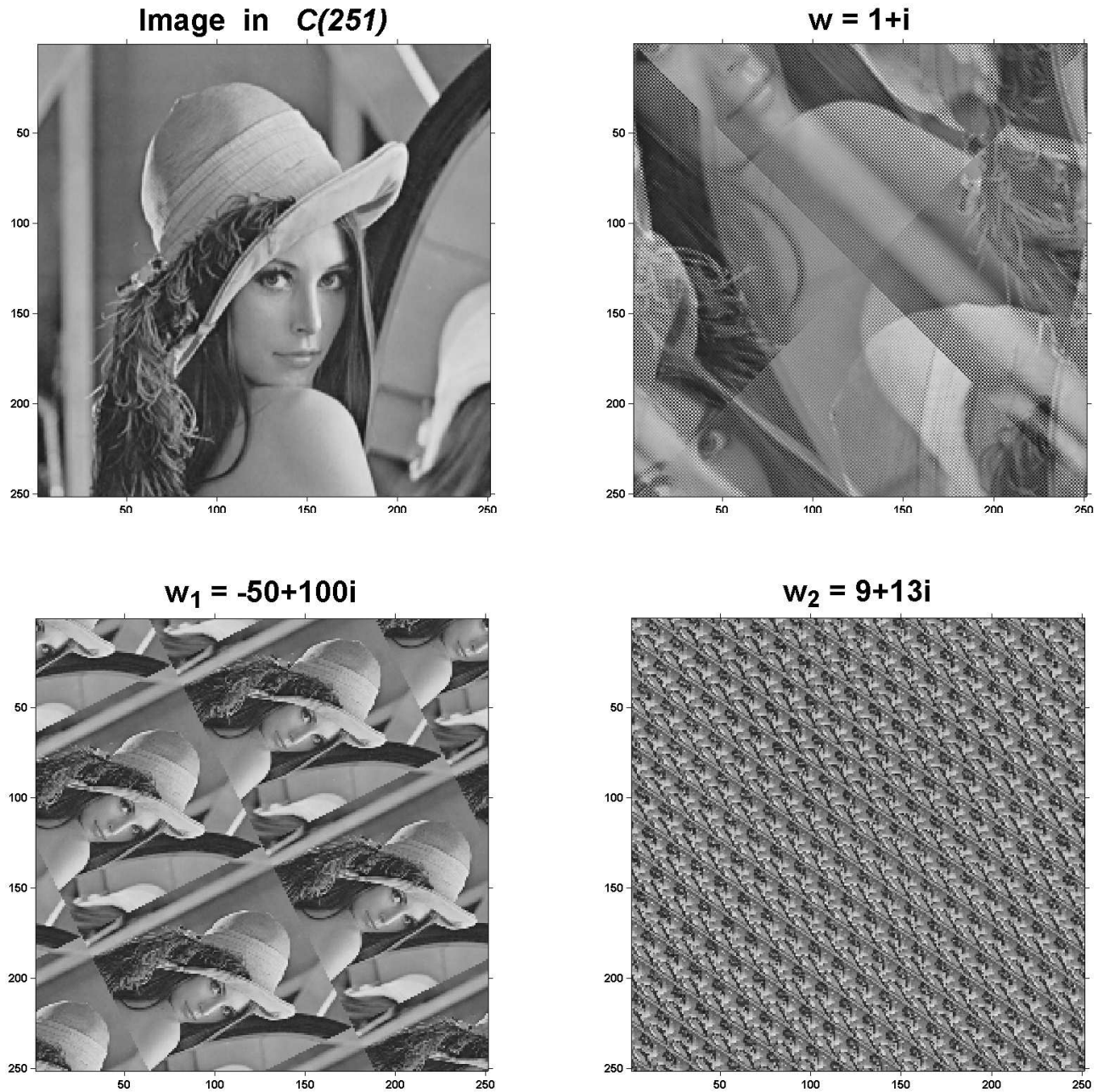
где $U(1) = \{ e^{i\theta} : 0 \leq \theta < 2\pi \}$ — одномерная унитарная группа, которую геометрически можно рассматривать как единичную окружность на комплексной плоскости, а $\mathbb{R}^+ = \{ r > 0 : r \in \mathbb{R} \}$ — мультипликативная группа неотрицательных действительных чисел, геометрически представляющая собой на $\mathbb{C}$ луч с выколотым началом. Тем самым устанавливается взаимно-однозначное соответствие

$$z = (a, b) \leftrightarrow (r, \theta) \in \mathbb{R}^+ \times [0, 2\pi),$$

между ненулевыми комплексными числами $z \in \mathbb{C}^*$ и парами действительных чисел (r, θ) , называемых *полярными координатами* точки z . Полагая $l = \ln r$, приходим к *полярно-логарифмическим координатам* (l, θ) .

Пусть $\mathbb{C}_p^* = \mathbb{C}_p \setminus \{0\}$ — мультипликативная группа тора $\mathbb{C}_p$. Поскольку [8, стр. 429] мультипликативная группа всякого конечного поля является циклической, $\mathbb{C}_p^* = \langle g \rangle$, где g — некоторый примитивный элемент поля $\mathbb{C}_p$. Например, как нетрудно проверить, $\mathbb{C}_{71}^* = \langle 2 + 7i \rangle$, а $\mathbb{C}_{251}^* = \langle 1 + 5i \rangle$. Нам понадобится следующий элементарно проверяемый результат:

Утверждение 3. Для всякого целого $p = 4k + 3 \geq 3$, числа $m = (p - 1)/2$ и $n = 2(p + 1)$ являются взаимно-простыми: $\gcd(m, n) = 1$.

Рис. 3: Комплексные вращения на $\mathbb{C}_{251}$

Заметим, что $mn = p^2 - 1 = |\mathbb{C}_p^*|$. Поэтому для комплексного дискретного тора возникает следующий аналог разложения (2):

Теорема 1. Мультипликативная группа $\mathbb{C}_p^* = \langle g \rangle$ комплексного дискретного тора разлагается в прямое произведение циклических групп порядков $m = (p-1)/2$ и $n = 2(p+1)$,

$$\mathbb{C}_p^* \simeq \mathbb{Z}_m \times \mathbb{Z}_n, \quad (3)$$

причем

$$\mathbb{Z}_n = \langle g^m \rangle \simeq \{z \in \mathbb{C}_p^* : N(z) = \pm 1\} \quad \text{и} \quad \mathbb{Z}_m = \langle g^n \rangle \simeq \{a^2 : 0 \neq a \in \mathbb{Z}_p\}.$$

Таким образом, если $(l, \theta) \in \mathbb{Z}_m \times \mathbb{Z}_n$ соответствует $z \in \mathbb{C}_p^*$, то $l \in \mathbb{Z}_m$ может рассматриваться как «логарифм модуля», а $\theta \in \mathbb{Z}_n$ — как «аргумент» дискретного комплексного числа z .

Определение 5. Дискретный тор $\mathbb{Z}_m \times \mathbb{Z}_n$ будем называть *полярной областью* для комплексного дискретного тора $\mathbb{C}_p$. (Заметим, что сама полярная область комплексным дискретным тором не является.)

Зафиксируем примитивный элемент g и определим отображение

$$\text{Exp}_g : \mathbb{Z}_m \times \mathbb{Z}_n \rightarrow \mathbb{C}_p^*$$

следующим образом:

$$\text{если } (l, \theta) \in \mathbb{Z}_m \times \mathbb{Z}_n, \quad \text{то } \text{Exp}_g(l, \theta) = g^{nl+m\theta} \in \mathbb{C}_p^*.$$

Как легко видеть, Exp_g изоморфно отображает аддитивную группу кольца $\mathbb{Z}_m \times \mathbb{Z}_n$ на мультипликативную группу $\mathbb{C}^*(p)$. Назовем его *модулярной экспонентой*, а обратное к нему отображение $\text{Ln}_g : \mathbb{C}^*(p) \rightarrow \mathbb{Z}_m \times \mathbb{Z}_n$ — *модулярным логарифмом*. Непосредственно из определения вытекает «основное логарифмическое тождество»

$$\text{Ln}_g(z_1 z_2) = \text{Ln}_g(z_1) + \text{Ln}_g(z_2).$$

Понятие модулярного логарифма позволяет однозначно сопоставить каждому цифровому изображению $f(z)$ размера $p \times p$ цифровое изображение $\psi(l, \theta)$ размера $m \times n$ в полярной области $\mathbb{Z}_m \times \mathbb{Z}_n$ следующим образом:

$$f(z) = \psi(\text{Ln}_g(z)).$$

Определение 6. Преобразование $\mathcal{P}[f] = \psi$ назовем *полярным преобразованием* изображения f , а изображение ψ будем называть *полярным образом* для f .

Как показывает следующий результат, преобразование $\mathcal{P}[f]$ может рассматриваться как дискретный аналог перехода в полярно-логарифмическую систему координат.

Утверждение 4. Комплексным вращениям изображения на КДТ соответствуют циклические сдвиги его образа в полярной области: если $\mathcal{P}[f(z)] = \psi(l, \theta)$, то для всякого $w \in \mathbb{C}_p^*$ выполняется соотношение

$$\mathcal{P}[f(wz)] = \psi(l - l_0, \theta - \theta_0), \quad \text{где } (l_0, \theta_0) = \text{Ln}_g(w). \quad (4)$$

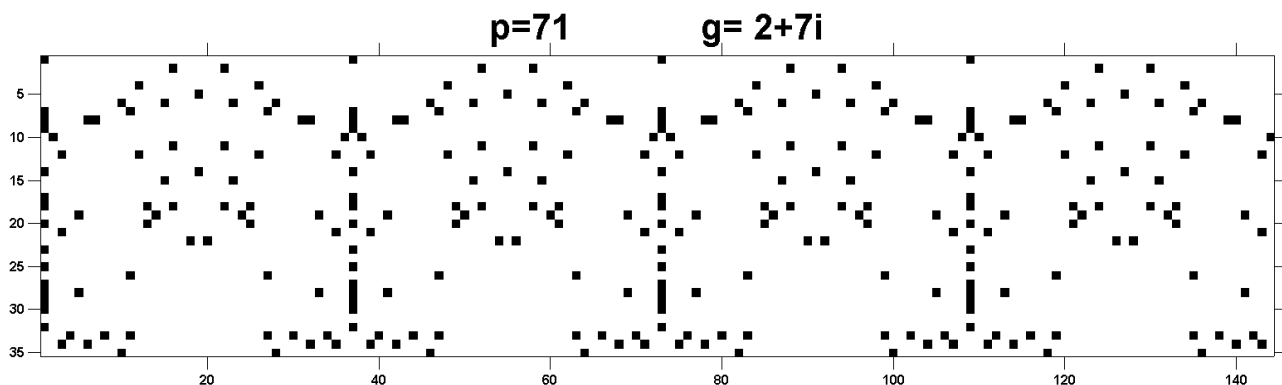


Рис. 4: Полярный образ изображения на рис. 2

В качестве примера рассмотрим изображение f на торе $\mathbb{C}_{71}$, показанное в левой части рис. 2. Его полярный образ имеет размер 35×144 и для $g = 2 + 7i$ показан на рис. 4. Поскольку $\text{Ln}_g(1 + 2i) = (14, 22) \in \mathbb{Z}_{35} \times \mathbb{Z}_{144}$, полярный образ изображения $f(wz)$ из правой части рис. 2 получается из полярного образа оригинала f циклическим сдвигом на 22 единицы вдоль оси «углов» θ и на 14 единиц вдоль оси «логарифмов радиусов» l . Хорошо заметный периодический характер полярного образа обусловлен симметриями оригинала.

Понятно, что полярный образ изображения зависит от выбора примитивного элемента g , выступающего в качестве «основания» модулярного логарифма $\text{Ln}_g(z)$. Следующий результат можно рассматривать как аналог «формулы замены основания».

Утверждение 5. Если g и h — два различных примитивных элемента группы $\mathbb{C}_p^*$, а z — произвольный элемент из $\mathbb{C}_p^*$, то

$$\text{Ln}_g(h) \odot \text{Ln}_h(z) = \mathbb{I}_p \odot \text{Ln}_g(z), \quad (5)$$

где $\odot$ означает умножение в кольце $\mathbb{Z}_m \times \mathbb{Z}_n$, а $\mathbb{I}_p := \text{Ln}_g(g)$ является константой, зависящей только от p , и равной $\mathbb{I}_p = (4, m)^{-1} \in \mathbb{Z}_m \times \mathbb{Z}_n$. В частности, $\mathbb{I}_{71} = (9, 107)$, $\mathbb{I}_{251} = (94, 125)$.

Отметим, что полярное преобразование $\mathcal{P}[f]$, являясь обратимым на $\mathbb{C}_p^*$, не будет обратимым в $\mathbb{C}_p$ только потому, что пиксел $(0, 0)$ не принадлежит полярной области. Формально добавляя к ней этот пиксел и считая его значение $f(0, 0)$ неизменным, можем говорить об обратимости $\mathcal{P}$ в расширенной полярной области.

Очевидно, что на всяком дискретном торе $\mathbb{Z}_m \times \mathbb{Z}_n$ естественно определяется дискретное преобразование Фурье $\mathcal{F}$: под Фурье-образом изображения $f : \mathbb{Z}_m \times \mathbb{Z}_n \rightarrow \mathbb{C}$ понимается функция $\mathcal{F}[f] = F : \mathbb{Z}_m \times \mathbb{Z}_n \rightarrow \mathbb{C}$, такая, что

$$F(\bar{u}, \bar{v}) = \frac{1}{\sqrt{mn}} \sum_{(\bar{x}, \bar{y}) \in \mathbb{Z}_m \times \mathbb{Z}_n} f(\bar{x}, \bar{y}) e^{-2\pi i(x \frac{u}{m} + y \frac{v}{n})}, \quad (6)$$

где $x, y, u, v \in \mathbb{Z}$ являются произвольными представителями соответствующих классов вычетов $\bar{x}, \bar{u} \in \mathbb{Z}_m$, $\bar{y}, \bar{v} \in \mathbb{Z}_n$, а $i \in \mathbb{C}$ — обычная мнимая единица поля комплексных чисел.

Определение 7. Зафиксировав примитивный элемент g , будем называть дискретное преобразование Фурье полярного образа $\mathcal{P}[f]$ модулярным преобразованием Меллина изображения f , обозначая его как

$$\mathcal{M}[f] = \mathcal{F}[\mathcal{P}[f]].$$

Нетрудно заметить, что $\mathcal{M}$ является линейным и обратимым на расширенной полярной области. Из утверждения 4 и теоремы о сдвиге для преобразования Фурье [11] немедленно получаем

Следствие 1. Спектр модулярного преобразования Меллина цифрового изображения инвариантен относительно комплексных вращений (а значит, и масштабирования): для любого ненулевого $w \in \mathbb{C}_p$,

$$|\mathcal{M}[f(wz)]| = |\mathcal{M}[f(z)]|.$$

По аналогии с $\mathcal{M}[f]$ можно ввести такие преобразования, как $\mathcal{P}^{-1}\mathcal{F}\mathcal{P}$, $\mathcal{F}\mathcal{P}\mathcal{F}^{-1}$ и $\mathcal{P}^{-1}\mathcal{F}\mathcal{P}\mathcal{F}^{-1}$. В частности, пусть

$$\widetilde{\mathcal{H}}[f] = \mathcal{M}[|\mathcal{F}^{-1}[f]|] = \mathcal{F}\mathcal{P}[|\mathcal{F}^{-1}[f]|].$$

Как легко видеть, $|\widetilde{\mathcal{H}}[f]|$ инвариантно относительно сдвигов, масштабирования и комплексных вращений цифрового изображения f . Это позволяет считать $\widetilde{\mathcal{H}}$ дискретным аналогом преобразования Фурье-Меллина.

Заключение

Можно надеяться, что представленные выше результаты предоставят дополнительные возможности для решения различных проблем качественной обработки цифровых изображений. Вместе с тем вне рамок настоящей работы остались такие вопросы, как детальное исследование свойств введенных дискретных преобразований, изучение возможностей их 3D-обобщений, а также детали их практического применения, в частности, для анализа симметрии изображений. Автор рассчитывает вернуться к рассмотрению указанных вопросов в дальнейшем.

Литература

- [1] Reddy S., Chatterji B. An FFT-based technique for translation, rotation, and scale invariant image registration // *IEEE Trans. on Image Processing*. 1996. Vol. 5, no. 8. Pp. 1266–1271.
- [2] Derrode S., Ghorbel F. Robust and efficient Fourier–Mellin transform approximations for gray-level image reconstruction and complete invariant description // *Computer Vision and Image Understanding*. 2001. Vol. 83, no. 1. Pp. 57–78.
- [3] Derrode S., Ghorbel F. Shape analysis and symmetry detection in gray-level objects using the analytical Fourier–Mellin representation // *Signal Processing*. 2004. Vol. 84, no. 1. Pp. 25–39.
- [4] Pratt W. K. Digital image processing. — John Wiley and Sons, 2007. 782 p.
- [5] Averbuch A., Coifman R. R., Donoho D. L., Israeli M., Shkolnisky Y. A framework for discrete integral transformations I: The pseudo-polar Fourier transform // *SIAM J. Sci. Comput.* 2008. Vol. 30. Pp. 764–784.
- [6] Каркищенко А. Н., Мнухин В. Б. Преобразование симметрии периодических структур в частотной области // *ММРО-15*. М.: МАКС Пресс, 2011. С. 386–389.
- [7] Каркищенко А. Н., Мнухин В. Б. Распознавание симметрии изображений в частотной области // *Интеллектуализация обработки информации (ИОИ-2012)*. М.: ТОРУС ПРЕСС, 2012. С. 426–429.
- [8] Кострикин А. И. Введение в алгебру. М.: Наука, 1977. 496 с.
- [9] Burton D. M. Elementary number theory. McGraw Hill, 2007. 434 p.
- [10] Dummit D. S., Foote R. M. Abstract algebra. John Wiley and Sons, 2004. 932 p.
- [11] Poularikas A. D. The transform and applications handbook. CRC Press, 2010. 1336 p.