

УДК 007.001.362 : 681.3.019

А. А. Кадыров, Н. Г. Федотов

(Пе.иза)

**НОВЫЕ ПРИЗНАКИ ИЗОБРАЖЕНИЙ,
ИНВАРИАНТНЫЕ ОТНОСИТЕЛЬНО ГРУППЫ ДВИЖЕНИЙ
И АФФИННЫХ ПРЕОБРАЗОВАНИЙ***

Излагается теория формирования признаков, не зависящих от движений изображений, а также от поворотов и гомотетий. Отличительной чертой группы рассматриваемых признаков является представление каждого из них в виде последовательной композиции трех функционалов, действующих на функции одной переменной. Предлагаемая методика построения новых признаков имеет следующие черты: а) может быть построено большое число новых признаков; б) получающиеся признаки поддаются алгоритмизации с параллельными вычислениями. На примере задачи распознавания цветных изображений клеток крови построено множество признаков, решающих эту задачу удовлетворительно.

Введение. Для постановки диагноза в современной медицине используются микрофотографии тканей, в частности микрофотографии препаратов крови. Имеющиеся особенности клеток позволяют установить диагноз. Однако при исследовании множества фотографий неизбежны ошибки, связанные с утомлением исследователя. Тем более, что диагноз ставится не на основании анализа одной клетки, а на основании анализа статистики распределения особенностей клеток, т. е. требуется анализировать множество микрообъектов и на этом основании составить статистически значимые выборки и таблицы. Известно, что исследователю трудно анализировать большое количество информации вручную. Следовательно, актуальна задача машинного анализа микрофотографий крови и срезов тканей.

Особую значимость приобретают теоретические методы распознавания, привлекающие признаки, инвариантные к движениям изображений.

Представлена новая теория формирования признаков, не зависящих от движений изображений, а также от аффинных преобразований. Ее возможности продемонстрированы на выделении признаков клеток крови.

Общим свойством аффинных преобразований изображений является сохранение прямых линий на изображении. Можно даже определить аффинное преобразование как взаимно однозначное непрерывное преобразование плоскости, при котором прямые переходят в прямые.

Отличительной чертой группы рассматриваемых признаков является представление каждого из них в виде последовательной композиции трех функционалов, каждый из которых действует на функции одной переменной. Поскольку нас интересуют движения изображений, в данном случае скольжения по некоторым прямым, естественно, нас интересует также реакция функционалов на такое скольжение. Мы выделяем два типа реакции: независимость функционалов от сдвига и зависимость (сенситивность), при которой сдвиг может быть выделен отдельным слагаемым.

* Работа выполнена при поддержке Российского фонда фундаментальных исследований (№ 97-01-00106).

Для каждого из трех функционалов можно легко найти десятки разных конкретизаций, удовлетворяющих требуемым (см. далее) условиям. Следовательно, здесь сразу получаем тысячи новых признаков, инвариантных к движениям. Для распознавания 2^n объектов требуется порядка n признаков, следовательно, имеется возможность диагностировать большое число заболеваний.

Данная работа является продолжением [1]. Предлагаемая теория развита на базе работ [2—5]. По этой теории имеются публикации [6—8]. В данной статье подробно изложены основы теории и показано ее применение для распознавания классов изображений.

Основные обозначения. Рассмотрим двумерную плоскость без выделенной системы координат. Математической моделью такой плоскости является двумерное аффинное пространство. Предположим, что на плоскости есть изображение. Его можно задать некоторой функцией $F(M)$, зависящей от точки M плоскости. Значение функции может быть числом или же вектором, указывающим, например, цвет, яркость и другие характеристики изображения в точке M . Для нашей теории не имеет значения, как именно функция F будет описывать изображение; существенно лишь то, что функция F — это вся доступная нам информация об изображении. Функция F будет далее называться «функцией изображения».

Нас интересуют признаки изображений, которые не меняются при сдвигах, вращениях и других преобразованиях. На такие преобразования может быть две точки зрения. Согласно первой, мы, действительно, получаем новое изображение, т. е. изображение меняется. Согласно второй, изменяется позиция наблюдателя (например, фотокамеры): он сдвигается, наклоняется, удаляется, поворачивается. Вторую точку зрения математически можно реализовать вводом подходящей системы координат, которая представляется наблюдателю естественной.

Рассмотрим оба подхода и убедимся, что они дают одинаковый эффект. Вторая точка зрения, когда изображение неизменно, а меняется позиция наблюдателя, удобнее для математического анализа, тогда как результаты анализа изображений в конкретных случаях естественнее и привычнее толковать в терминах движений изображений (например, мы говорим: «знак повернут, сдвинут, увеличен»). Более точно будет дана формульная связь двух подходов.

Уточним сначала, какие именно преобразования изображений поддаются изучению предлагаемой методикой. Это сдвиги, повороты, гомотетии (т. е. равномерные сжатия и растяжения), сжатия (растяжения) в одном направлении, отражения относительно произвольной прямой, а также произвольные комбинации таких преобразований. Такие преобразования составляют *класс аффинных преобразований плоскости*. Мы рассматриваем только обратимые, т. е. невырожденные преобразования. Хорошо известно, что каждое аффинное преобразование можно представить в виде линейного невырожденного оператора с последующим сдвигом. Однако для этого нужно на плоскости ввести систему координат. Если не используются отражения относительно прямых, то определитель матричного оператора положителен и аффинное преобразование называется «сохраняющим ориентацию».

На числовой прямой аффинные (невырожденные) преобразования имеют вид $y = ax + w$, где $a \neq 0$ и w — постоянные для данного аффинного преобразования коэффициенты; x — независимая переменная; y — результат преобразования. Заметим, что при $w = 0$ имеем линейное преобразование прямой. Если $w \neq 0$, часто тоже употребляется термин «линейное преобразование», однако такое употребление термина не согласуется с понятием линейного преобразования в пространствах высших размерностей.

Предположим, что имеется начало координат — точка O на плоскости и два неколлинеарных вектора e_1, e_2 . Каждая точка M на плоскости получает координаты x_1, x_2 :

$$M = O + e_1x_1 + e_2x_2,$$

используя матричные обозначения. Это равенство можно записать в виде

$$M = O + (e_1, e_2) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}. \quad (1)$$

Линейное преобразование сопоставит точку M с точкой с координатами, умноженными на некоторую матрицу A ; сдвиг на вектор w с координатами w_1, w_2 приведет к аффинному преобразованию. В общем виде аффинное преобразование в данной системе координат определяется невырожденной матрицей A и некоторым вектором w по формуле: каждая точка (1) переходит в точку

$$M' = O + (e_1, e_2) \left(A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \right). \quad (2)$$

Для исследования другой точки зрения на движение изображения выберем новую систему координат: начало координат — точка P и базисные неколлинеарные векторы v_1, v_2 . Точка M получит координаты y_1, y_2 :

$$M = P + (v_1, v_2) \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}. \quad (3)$$

Существует матрица (обозначим ее A) и вектор u , которые связывают обе системы координат следующим образом (применяется матричная запись):

$$(v_1, v_2) = (e_1, e_2)A^{-1}, \quad O - P = u = (v_1, v_2) \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}. \quad (4)$$

Приравнявая (1) и (3), получаем с учетом (4)

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}.$$

Две введенные системы координат для краткости обозначим тройками (O, e_1, e_2) и (P, v_1, v_2) .

Итак, можно сделать вывод. *При допущении, что изображение в некоторой системе координат (O, e_1, e_2) претерпело аффинное преобразование: каждая точка с координатами x_1, x_2 перешла в точку с координатами*

$$A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}.$$

Преобразованное изображение идентично тому, которое мы получим путем введения новых координат (P, v_1, v_2) по формулам (4). (При наблюдении изображения координатная система представляется наблюдателю прямоугольной, расстояние — декартовым.)

Будем рассматривать каждую прямую с фиксированным направлением на ней, т. е. ориентированную прямую. Зафиксируем систему (O, e_1, e_2) координат: начало координат — точка O на плоскости и два базисных вектора e_1, e_2 . Пусть l — прямая, φ — такой угол, что вектор $e_1 \cos \left(\varphi + \frac{\pi}{2} \right) + e_2 \sin \left(\varphi + \frac{\pi}{2} \right)$ — ее направляющий.

Для удобства изложения введем для любого угла θ вектор-столбец $\lambda(\theta)$ и матрицу $S(\theta)$ с помощью формул

$$\lambda(\theta) = \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix}, \quad S(\theta) = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}.$$

Матрица $S(\theta)$ — это матрица поворота на угол θ , при $\theta = 0$ она обращается в единичную матрицу: $S(0) = 1$. При различных θ эти матрицы перестановочны. Для любых θ_1, θ_2 выполняется $S(\theta_1)\lambda(\theta_2) = \lambda(\theta_1 + \theta_2)$.

В новых обозначениях направляющий вектор прямой l может быть записан в виде $(e_1, e_2)S(\pi/2)\lambda(\varphi)$. Ближайшая к началу O координат точка на прямой есть некоторая точка вида $O + (e_1, e_2)p\lambda(\varphi)$, поскольку вектор $(e_1, e_2)\lambda(\varphi)$ ортогонален прямой l , где p — некоторое число, его модуль равен расстоянию до прямой l . Это верно в выбранной системе координат, в другой системе координат ортогональность определится другими базисными векторами. Таким образом, уравнение прямой имеет вид

$$x_1\cos\varphi + x_2\sin\varphi = p,$$

единственное отличие от канонического нормального уравнения прямой — параметр p не обязательно положителен.

В параметрическом представлении прямая l может быть задана переменной точкой

$$M = O + (e_1, e_2)p\lambda(\varphi) + (e_1, e_2)t\lambda(\varphi + \pi/2),$$

пробегающей эту прямую при изменении параметра t в множестве всех вещественных чисел \mathbb{R} . Параметры φ и p будем называть полярными параметрами ориентированной прямой l в системе координат (O, e_1, e_2) , параметр t — аффинным параметром точки на прямой l , отвечающим данной системе координат.

Определим отображение в плоскость L , зависящее от шести параметров:

$$L(O, e_1, e_2, \varphi, p, t) = O + (e_1, e_2)(pI + tS(\pi/2))\lambda(\varphi), \quad (5)$$

при фиксированных первых пяти параметрах это параметризация прямой при условиях: $S(\pi/2)\lambda(\varphi)$ — координаты направляющего вектора прямой, задающего на ней ориентацию, модуль параметра p имеет смысл расстояния до прямой в системе координат (O, e_1, e_2) . В данном случае мы получили параметризацию прямой l .

В фиксированной системе координат (O, e_1, e_2) каждая ориентированная прямая единственным образом представляется в виде (5), т. е. ее параметры φ и p определены однозначно (φ с точностью до 2π). Та же прямая, но с противоположной ориентацией получает параметризацию

$$L(O, e_1, e_2, \varphi + \pi, -p, t) \quad (t \in \mathbb{R}).$$

Наиболее важное свойство параметризации (5) — непрерывная зависимость ориентированной прямой от параметров. Например, сразу видно, что в множестве всех ориентированных прямых на плоскости можно ввести структуру многообразия $S^1 \times \mathbb{R}$ (декартово произведение одномерной окружности на вещественную прямую, т. е. цилиндр).

Рассмотрим две системы координат (O, e_1, e_2) и (P, v_1, v_2) . Ориентированная прямая l в каждой из них получит свою параметризацию:

$$L(O, e_1, e_2, \varphi, p, t) \quad (t \in \mathbb{R}), \quad L(P, v_1, v_2, \varphi', p', t') \quad (t' \in \mathbb{R}). \quad (6)$$

Определим зависимости между полярными параметрами прямой в двух координатных системах (6), а также между аффинными параметрами одной точки на прямой, отвечающими двум этим системам. Иначе говоря, для данной прямой l выразим φ' и p' через φ, p и для данной конкретной точки M на этой прямой выразим t' через t . Результат алгебраических преобразований следующий:

пусть (O, e_1, e_2) и (P, v_1, v_2) — две системы координат на плоскости; l — некоторая ориентированная прямая; φ, p — ее полярные параметры в пер-

вой системе координат; φ' и p' — ее полярные параметры во второй системе координат. Тогда существуют числа $k > 0$ и b такие, что при всех t

$$L(O, e_1, e_2, \varphi, p, t) = L(P, v_1, v_2, \varphi', p', kt + b) \quad (t \in \mathbf{R}). \quad (7)$$

Поворот, гомотетия и сдвиг системы координат. Предположим, что две системы координат (O, e_1, e_2) и (P, v_1, v_2) на плоскости связаны равенствами

$$O = P, \quad (e_1, e_2) = (v_1, v_2) \mu S(\theta) \quad (\mu > 0), \quad (8)$$

где $S(\theta)$ — матрица поворота на некоторый угол θ . Например, если имеется изображение в первой системе координат, то во второй системе координат мы будем его наблюдать таким, как если бы оно было повернуто на угол θ и растянуто (увеличено) в μ раз при $\mu > 1$. Если же $0 < \mu < 1$, то изображение нам представится сжатым (т. е. уменьшенным). В случае $\mu = 1$ изображение будет лишь повернутым.

Для данной прямой l запишем равенство двух параметризаций и, используя формулы (1) — (8), получим следующий результат:

если две системы (O, e_1, e_2) и (P, v_1, v_2) координат связаны преобразованиями гомотетии и поворота (8), то при всех p и φ

$$L(O, e_1, e_2, \varphi, p, t) = L(O, v_1, v_2, \varphi + \theta, \mu p, \mu t) \quad (t \in \mathbf{R}). \quad (9)$$

Предположим теперь, что две системы координат (O, e_1, e_2) и (P, v_1, v_2) на плоскости могут быть получены одна из другой сдвигом. Это значит, что базисные векторы одни и те же: $e_1 = v_1$ и $e_2 = v_2$, а начала координат разнятся на некоторый вектор $O - P = s$. В этом случае получаем следующее утверждение. Координаты вектора $O - P$ могут быть записаны в виде $s_0 \cos \psi_0$, $s_0 \sin \psi_0$, т. е.

$$s = (e_1, e_2) s_0 \lambda(\psi_0),$$

где s_0 и ψ_0 — некоторые числа.

Пусть две системы (O, e_1, e_2) и (P, v_1, v_2) координат отличаются лишь выбором начала координат: $O - P = (e_1, e_2) s_0 \lambda(\psi_0)$, тогда при всех p и φ

$$L(O, e_1, e_2, \varphi, p, t) = L(O + (e_1, e_2) s_0 \lambda(\psi_0), e_1, e_2, \varphi, p + s_0 \cos(\varphi - \psi_0), t + s_0 \sin(\varphi - \psi_0)) \quad (t \in \mathbf{R}). \quad (10)$$

Поворот, гомотетия и сдвиг изображения. Предположим, что изображение в системе (O, e_1, e_2) координат претерпело поворот на некоторый угол θ , гомотетию с коэффициентом $\mu > 0$ и затем сдвиг на вектор $w = w_1 e_1 + w_2 e_2$. Вид изображения будет таким же, как если бы мы ввели новую систему координат (P, v_1, v_2) по формулам (4), в которых следует положить $A = \mu S(\theta)$.

Рассмотрим промежуточную систему координат (O, v_1, v_2) . Пусть некоторая прямая l имеет параметризацию $L(O, e_1, e_2, \varphi, p, t)$ ($t \in \mathbf{R}$). Согласно (9), эта параметризация при каждом t совпадает с

$$L(O, v_1, v_2, \varphi + \theta, \mu p, \mu t). \quad (11)$$

Применим сейчас формулу (10). Пусть $s_0 \lambda(\psi_0)$ — произвольный вектор-столбец. Временно для применения формулы (10) в (11) переобозначим аргументы $v_1, v_2, \varphi + \theta, \mu p, \mu t$ соответственно в e_1, e_2, φ, p, t . Теперь применим (10) и получим, что (11) совпадает с

$$L(O + (v_1, v_2) s_0 \lambda(\psi_0), v_1, v_2, \varphi + \theta, \mu p + s_0 \cos(\varphi - \psi_0), \mu t + s_0 \sin(\varphi - \psi_0)) \quad (t \in \mathbf{R}). \quad (12)$$

Числа w_1, w_2 нам даны по допущенному предположению. Выберем s_0 и ψ_0 так, чтобы

$$\begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = s_0 \lambda(\psi_0),$$

тогда можно положить $P = O - (v_1, v_2)s_0\lambda(\psi_0)$ в соответствии с формулами (4). Параметризация (12) оказывается параметризацией ориентированной прямой l в системе координат (P, v_1, v_2) .

Сформулируем полученный результат.

Предположим, что две системы координат (O, e_1, e_2) и (P, v_1, v_2) на плоскости связаны равенствами

$$O - P = (v_1, v_2)s_0\lambda(\psi_0), \quad (e_1, e_2) = (v_1, v_2)\mu S(\theta) \quad (\mu > 0),$$

тогда при всех p и φ

$$\begin{aligned} L(O, e_1, e_2, \varphi, p, t) &= L(P, v_1, v_2, \varphi + \theta, \mu p + \\ &+ s_0\cos(\varphi - \psi_0), \mu t + s_0\sin(\varphi - \psi_0)) \quad (t \in \mathbf{R}). \end{aligned} \quad (13)$$

Вторая система (P, v_1, v_2) координат имеет следующий смысл: у произвольного изображения во второй системе координат такой вид, как если бы оно в первой системе координат претерпело конформное аффинное преобразование координат, а именно умножение на матрицу $\mu S(\theta)$ ($\mu > 0$) и сдвиг на вектор-столбец $s_0\lambda(\psi_0)$.

Триpletные признаки изображений. Для удобства записи примем следующее соглашение. Пусть три функционала T, P, Φ действуют на функции, зависящие от переменных t, p, φ соответственно. При этом t и p пробегают вещественную ось, а φ имеет смысл угла, т. е. можно считать функции от φ 2π -периодическими или же считать переменную φ пробегающей окружность — как удобнее. На каких именно функциональных пространствах или множествах заданы операторы, пока не имеет значения. Однако мы предполагаем встречающиеся далее формулировки имеющими смысл, в частности, функционалы T, P, Φ заданы на множествах функций, инвариантных к сдвигу графиков функций. Последнее означает, что если $f(\cdot)$ принадлежит одному такому множеству, то $f(\cdot + a)$ при любом числе a тоже принадлежит ему (здесь и далее точка означает место аргумента в соответствии с известной математической символикой).

Предположим, что имеется изображение на плоскости, описываемое функцией F . Предположим также, что даны три функционала $(\Phi, P$ и $T)$, свойства которых указаны выше. Для фиксированной системы (O, e_1, e_2) координат на плоскости можно ввести функцию f , тоже характеризующую изображение:

$$f(\varphi, p, t) = F(L(O, e_1, e_2, \varphi, p, t)),$$

где φ, p, t — любые числа; $f(\varphi, p, t)$ — значение функции изображения на ориентированной прямой с полярными параметрами φ, p в точке этой прямой с аффинным параметром t . Если зафиксировать φ и p , а затем φ , то можно последовательно определить еще две функции:

$$g(\varphi, p) = Tf(\varphi, p, t), \quad h(\varphi) = Pg(\varphi, p). \quad (14)$$

Полученная таким образом функция h 2π -периодична. Применим к ней функционал Φ и получим число, которое может служить признаком изображения. Итак, в системе координат (O, e_1, e_2) имеем признак ηF , или в более подробной записи

$$\eta F = \eta(O, e_1, e_2, \Phi, P, T, F) = \Phi \circ P \circ T(F \circ L) \quad (15)$$

(в соответствии с известной математической символикой знак « \circ » означает композицию отображений).

Признаки изображений, имеющие вид (15), называем *триплетными*. Они легко поддаются исследованию на инвариантность, т. е. их можно исследовать с целью выявления признаков, которые не зависят от аффинных преобразований изображения. Ближайшей целью будет исследование условий на функционалы Φ , P , T , при которых признаки инвариантны. Для распознавания изображений могут быть использованы одновременно разные признаки вида (15), часть из которых не обязательно инвариантна, но они могут быть промежуточными для конструирования инвариантных признаков. Далее будет показано, что с помощью признаков вида (15) можно не только распознавать изображения, но и выяснять, какое именно преобразование претерпело изображение по сравнению с эталонным.

Признаки изображений, инвариантные к аффинным преобразованиям. Чтобы продемонстрировать возможности триплетных признаков вида (15), сформулируем теорему, которая позволяет конструировать инвариантные признаки. Эта теорема оказывается частным случаем более общих утверждений, приведенных далее.

Теорема 1. *Предположим, что функционалы Φ , P и T инвариантны к сдвигам, т. е. для всех чисел b :*

- 1) $T(u(t + b)) = T(u(t))$ для всех допустимых функций u ;
- 2) $P(u(p + b)) = P(u(p))$ для всех допустимых функций u ;
- 3) $\Phi(u(\varphi + b)) = \Phi(u(\varphi))$ для всех допустимых функций u .

В этих условиях признак (15) не зависит от сдвигов и поворотов изображения.

З а м е ч а н и е. В этой теореме не фиксируется заранее область задания функционалов и множество возможных функций изображений. Области задания должны быть такими, чтобы все свойства теоремы могли выполняться (т. е. чтобы результаты фигурирующих там операций не выходили за пределы соответствующих множеств) и мы могли вычислять признак по формуле (15). Подробное перечисление свойств было бы громоздко, тогда как естественные для решения задач распознавания функционалы и пространства почти наверное удовлетворяют всем требованиям наших теорем.

Функционалы, инвариантные и чувствительные к сдвигу. Экспериментирование с подбором функционалов для признака η в формуле (15) показывает, что большинство эффективных признаков обладает рядом общих свойств, которые рассмотрим отдельно.

Пусть x — вещественная переменная (для полной строгости следующих далее формул можно считать в них x тождественным отображением числовой прямой \mathbf{R} на себя), функционал Ξ определен на некотором множестве функций, замкнутом относительно сдвигов их графиков. Это требование связано с тем, что мы будем применять к ним операции сдвига. Кроме того, если потребуется рассматривать другие свойства, где к функциям применяются другие операции, то по умолчанию считаем, что и относительно этих операций рассматриваемое множество замкнуто.

Функционал Ξ называется инвариантным к сдвигу, если

$$(i1) \quad \Xi(\xi \circ (x + b)) = \Xi\xi \text{ при всех } b \in \mathbf{R}.$$

Инвариантный к сдвигу функционал Ξ может обладать дополнительными свойствами:

$$(i2) \quad \text{существует положительная функция } \alpha \text{ такая, что } \Xi(\xi \circ (ax)) = \alpha(a)\Xi\xi \text{ при всех } a > 0 \text{ из области } \text{Dom}\alpha \text{ задания функции } \alpha;$$

$$(i3) \quad \text{существует положительная функция } \gamma \text{ такая, что } \Xi(c\xi) = \gamma(c)\Xi\xi \text{ при всех } c > 0 \text{ из области } \text{Dom}\gamma \text{ задания функции } \gamma;$$

$$(i4) \quad \Xi(\xi \circ (-x)) = \Xi\xi; \text{ это свойство назовем симметричностью.}$$

Очевидно, инвариантный к сдвигу функционал, как и любой другой функционал, удовлетворяет равенствам (i2) и (i3) при $a = 1$, $c = 1$, тогда $\alpha(1) = 1$ и $\gamma(1) = 1$, т. е. $\text{Dom}\alpha = \{1\}$, $\text{Dom}\gamma = \{1\}$. Области определения функций α и γ содержат, по крайней мере, единицу.

Инвариантный к сдвигу функционал можно представить как операцию выбора точки на оси ординат по виду графика функции, не принимая во внимание положение начала координат. Типичные функционалы такого вида — среднее значение функции, полная вариация функции.

Рассмотрим теперь чувствительные («чувствительные») к сдвигам функционалы. В нашем изложении это понятие конкретизируется следующим образом. Пусть z — функционал, определенный на некотором множестве функций вещественной переменной. Относительно этого множества действуют принятые выше соглашения (см. определение инвариантных к сдвигу функционалов). Однако теперь это множество не может быть линейным пространством, как это будет видно из следующего определения.

Функционал z назовем чувствительным к сдвигу, если

$$(s1) \quad z(\zeta \circ (x + b)) = z\zeta - b \text{ при всех } b \in \mathbf{R}.$$

Дополнительные, необязательные свойства:

$$(s2) \quad z(\zeta \circ (ax)) = \frac{1}{a} z\zeta \text{ при всех } a > 0; \text{ вместе со свойством (s1) это дает}$$

$$z(\zeta \circ (a(x + b))) = \frac{1}{a} z\zeta - b \text{ или в другой форме } z(\zeta \circ (ax + b)) = \frac{1}{a} z\zeta - \frac{1}{a} b;$$

$$(s3) \quad z(c\zeta) = z\zeta \text{ при всех } c > 0;$$

$$(s4) \quad z(\zeta + d) = z\zeta \text{ при всех } d \in \mathbf{R}.$$

Чувствительный к сдвигу в смысле определения (s1) функционал z не может иметь в области определения нулевую функцию, ибо для функции $\zeta \equiv 0$ свойство (s1) не может выполняться. Аналогично свойство (s1) не может быть выполнено для любой периодической функции. Однако если функция не имеет ни одного периода, то чувствительный к сдвигу функционал на ней может быть определен непротиворечиво. Например, на множестве всех финитных ненулевых функций можно определить чувствительный к сдвигу функционал многими способами. Самый простой из них — наибольшая (или наименьшая) точка носителя функции (т. е. $z\zeta = \max \text{Close}\{x: \zeta(x) \neq 0\}$, Close — операция топологического замыкания). Неформально можно объяснить чувствительный к сдвигу функционал как операцию выбора точки на оси абсцисс по виду графика функции, не принимая во внимание положение начала координат. Если же мы не будем принимать во внимание также и масштаб, то удовлетворим свойству (s2).

Отдельно нужно рассмотреть вопрос о чувствительных к сдвигу функционалах, заданных на пространстве 2π -периодических функций. Определение (s2) здесь не подходит. Предположим, функция h 2π -периодична, а ее наименьший положительный период существует и равен τ , тогда τ (целое число) n раз укладывается в 2π . Анализируя (s1), можно заключить, что образом функционала здесь целесообразно считать дискретную равномерную сетку

$$\{a + j\tau: j = 0, 1, -1, 2, -2, 3, \dots\}. \quad (16)$$

Число a , определяющее сетку, само может быть задано с точностью до слагаемого, кратного τ .

Если некоторые числа A, B отличаются на кратное τ , то употребляется запись $A = B(\text{mod } \tau)$. Предположим теперь, что функция h 2π -периодична, о ее наименьшем периоде не делаем предположений. Пусть τ — положительное число (не обязательно период функции h). Функционал z назовем чувствительным по модулю τ к сдвигу, если

$$(s1. \tau) \quad z(h \circ (x + b)) = zh - b(\text{mod } \tau) \text{ при всех } b \in \mathbf{R}.$$

Если $\tau = 2\pi/n$, то будем употреблять также обозначение

$$(s1. 2\pi/n) = (s1. \tau) \text{ при } \tau = 2\pi/n.$$

Здесь, очевидно, функционал можно считать определенным с точностью до слагаемого, кратного τ . Строгая формулировка: образ z есть множество равномерных дискретных сеток вида (16). Очевидно, что данная формулиров-

ка — естественное обобщение понятия фазы для синусоидальных функций. Определение (sl. τ) будет применяться для анализа периодической функции h из (14), получающейся на предпоследнем шаге в процессе вычисления признака изображения (15). Определение фаз гармоник укладывается в схему (sl. τ), тогда как определение амплитуд — в схему (i1). Если сумма гармоник испытывает сдвиг, то разные гармоники сдвигаются на различные части своего периода, однако если за каждой гармоникой периода $\tau = 2\pi/n$ закреплена сетка вида (16), то все эти сетки испытывают один сдвиг. Таким образом, подход к фазам, как к сеткам, намеченный определением (sl. τ), позволит на основании спектрального анализа конструировать функционалы, чувствительные и инвариантные к сдвигам.

Триплетный признак η в случае поворота, гомотетии и сдвига изображения. Предположим, что две системы координат (O, e_1, e_2) и (P, v_1, v_2) на плоскости связаны равенствами

$$O - P = (v_1, v_2)s_0\lambda(\psi_0), \quad (e_1, e_2) = (v_1, v_2)\mu S(\theta) \quad (\mu > 0),$$

т. е. в иной трактовке мы имеем сдвиг и поворот изображения. Как было установлено в (10)—(13), параметры прямой и точки на прямой в одной системе координат и в другой (ниже помечены штрихами) связаны равенствами:

$$\varphi' = \varphi + \theta, \quad p' = \mu p + s_1(\varphi), \quad t' = \mu t + s_2(\varphi),$$

где $s_1(\varphi) = s_0\cos(\varphi - \psi_0)$, $s_2(\varphi) = s_0\sin(\varphi - \psi_0)$; положим также $\nu = 1/\mu$, тогда равенства можно записать в виде

$$\varphi = \varphi' - \theta, \quad p = \nu p' - \nu s_1(\varphi), \quad t = \nu t' - \nu s_2(\varphi).$$

Подставим эти формулы в (13) и затем, чтобы не усложнять обозначений, уберем штрихи. Получим равносильный вариант формулы (13): при всех p и φ

$$\begin{aligned} L(O, e_1, e_2, \varphi - \theta, \nu p - \nu s_1(\varphi), \nu t - \nu s_2(\varphi)) = \\ = L(P, v_1, v_2, \varphi, p, t) \quad (t \in \mathbb{R}). \end{aligned} \quad (17)$$

Наша цель — выразить признак η_2 , вычисленный во второй системе координат, через признак η_1 , вычисленный в первой. Имеем согласно (15) и (17)

$$\begin{aligned} \eta_2 = \eta(P, v_1, v_2, \Phi, P, T, F) = \Phi \circ P \circ T(F \circ L(P, v_1, v_2, \varphi, p, t)) = \\ = \Phi \circ P \circ T(F \circ L(O, e_1, e_2, \varphi - \theta, \nu p - \nu s_1(\varphi), \nu t - \nu s_2(\varphi))). \end{aligned}$$

Чтобы продолжить преобразования, нужно иметь какие-то допущения относительно функционалов. Ограничимся функционалами, инвариантными, а также чувствительными к сдвигу. Функционал T не следует брать чувствительным к сдвигу, поскольку если считать функцию изображения финитной, что вполне естественно, то функционал T должен будет применяться также и к нулевой функции. Выше замечено, что тогда мы не смогли бы считать T чувствительным к сдвигу в смысле наших определений. Итак, пусть функционал T удовлетворяет свойствам (i1) и число $\nu = 1/\mu$ входит в область задания функции $\alpha = \alpha_T$, тогда

$$\eta_2 = \Phi \circ P(\alpha_T(\nu)t(\varphi - \theta, \nu p - \nu s_1(\varphi))),$$

где g и следующая далее функция h определены в (16). В зависимости от свойств функционала P есть два варианта:

1) P удовлетворяет требованиям инвариантности (i1), тогда $\eta_2 = \Phi(\gamma_P(\alpha_T(\nu))\alpha_P(\nu)h(\varphi - \theta))$ при условии, что коэффициент имеет смысл, т. е. $\nu \in \text{Dom}\alpha_T, \nu \in \text{Dom}\alpha_P, \alpha_T(\nu)\alpha_P(\nu) \in \text{Dom}\gamma_P$;

2) P удовлетворяет требованиям чувствительности (s1), (s2), тогда $\eta_2 = \Phi(\mu h(\varphi - \theta) + s_1(\varphi))$.

В зависимости от свойств функционала Φ каждый из этих вариантов разветвляется. Рассмотрим первый вариант.

1i) Функционал Φ удовлетворяет свойству (i1), и следующий далее коэффициент имеет смысл. Тогда

$$\eta_2 = \gamma_\Phi(\gamma_P(\alpha_T(\nu))\alpha_P(\nu))\eta_1.$$

Заметим сразу, что это равенство доказывает приведенную выше теорему, ибо в ее случае $\nu = 1/\mu = 1$.

1s) Функционал Φ удовлетворяет свойству (s1. $2\pi/n$) и (s3). Тогда $\eta_2 = \eta_1 + \theta(\text{modd}2\pi/n)$.

Вариант 2 более сложен, но в нем можно получить параметры аффинного преобразования μ, s_0, ψ_0 . Гармоника $s_1(\varphi) = s_0 \cos(\varphi - \psi_0)$, входящая в выражение $\eta_2 = \Phi(\mu h(\varphi - \theta) + s_1(\varphi))$, полностью определяет сдвиг изображения и может быть выделена средствами гармонического анализа, если функция $\mu h(\varphi - \theta)$ уже известна. Опишем этот процесс. После вычисления функционалов T , а затем P будет получена 2π -периодическая функция

$$h'(\varphi) = \mu h(\varphi - \theta) + s_0 \cos(\varphi - \psi_0), \quad (18)$$

где h — периодическая функция, которая получилась бы на предпоследнем шаге вычисления признака стандартного изображения. Однако в рассматриваемой системе координат изображение не обязательно имеет стандартный вид, и нам на предпоследнем шаге известна функция (18), но неизвестны параметры аффинного преобразования μ, θ, s_0, ψ_0 . Их надлежит определить, также надлежит определить функцию h для распознавания образа. Полагая, что функция (18) может быть разложена в ряд Фурье, получим

$$\begin{aligned} h'(\varphi) &= \mu \sum_{i=0}^{\infty} h^{(i)}(\varphi - \theta) + s_0 \cos(\varphi - \psi_0) = \\ &= \mu \sum_{i \neq 1} h^{(i)}(\varphi - \theta) + \mu h^{(1)}(\varphi - \theta) + s_0 \cos(\varphi - \psi_0) = \\ &= \mu h^\perp(\varphi - \theta) + \mu h^{(1)}(\varphi - \theta) + s_0 \cos(\varphi - \psi_0), \end{aligned} \quad (19)$$

где $h^{(i)}$ — i -я гармоника в разложении h , т. е. $h^{(i)}$ — линейная комбинация $\cos(i\varphi)$ и $\sin(i\varphi)$ ($i = 0, 1, 2, \dots$); функция h^\perp — проекция h на подпространство, ортогональное первой гармонике. Отсюда видно, что нам доступны две функции:

$$(h')^\perp = \mu h^\perp(\varphi - \theta), \quad \mu h^{(1)}(\varphi - \theta) + s_0 \cos(\varphi - \psi_0). \quad (20)$$

Применяя к первой из них некоторое множество функционалов, инвариантных к сдвигу (например, амплитуды гармоник и их отношения), по полученным данным мы сможем распознать изображение и тем самым определить h и μ . Чувствительные к сдвигу функционалы (например, фазы) позволят нам затем определить θ . Из второй доступной нам функции можно найти вектор сдвига изображения. Таким образом, задача распознавания решается одновременно с определением деформаций эталонных изображений.

Результаты проведенных рассуждений (17)—(19) сведены в следующую таблицу (операция $u \rightarrow u^\perp$ определена выше — это элиминирование первой гармоники).

Доказательство теоремы 1. Теорему 1 доказывает первая строчка таблицы. Утверждение, похожее на теорему 1, можно получить из третьей строчки таблицы. Еще две строчки таблицы могут понадобиться при конструкции целого множества инвариантных функционалов.

Описание расчета признаков. Рассмотрим прямую с уравнением $x \cos \varphi + y \sin \varphi = p$. Направление прямой задается вектором с координатами $(-\sin \varphi, \cos \varphi)$. Центр декартовой системы координат расположим в центре изображения. Параметр p в отличие от обычно принятой записи может принимать также отрицательные значения. Это связано с тем, что мы хотим, чтобы каждая прямая считалась дважды: с учетом ее возможных направлений.

На прямой бинарное изображение оставляет след — функцию изображения вдоль прямой. На прямой имеется естественная одномерная система координат, поэтому можно вычислять функционал T (grace-функционал). Результат вычисления зависит от двух параметров: p и φ . Направим ось Ox горизонтально, а ось Oy вертикально. В каждой точке (p, φ) интенсивностью цвета отметим значение функционала T . Получим grace-трансформанту изображения.

Главное свойство grace-трансформанты состоит в том, что она «хорошо» реагирует на движения объекта на изображении. Если объект поворачивается, то трансформанта лишь сдвигается вправо или влево. Если объект смещается на вектор, то столбцы трансформанты сдвигаются вверх или вниз, т. е. трансформанта получает «волнообразное» возмущение. При изучении движущихся объектов и для численного анализа трансформанты гораздо удобнее исходных изображений. Отсюда видно, что если мы обработаем столбцы, то получим периодическую функцию, сдвиг которой пропорционален повороту эталонного изображения. Эта обработка столбцов осуществляется «диаметральным» оператором. Результатом является периодическая кривая, назовем ее «циркком».

И наконец, на последнем шаге по цирку с помощью circus-функционала определяем число — признак изображения.

Признаки, использованные в численных экспериментах. Grace-функционалы. Grace-функционал использовался в шести вариантах. Их номера 1, 2, 3, 5, 6, 7 (отсутствует номер 4). Это связано с тем, что в работающей программе имеется стандартная классификация функционалов. Для распознавания клеток крови — эритроцитов — эмпирическим путем определено, что из имеющегося в программе списка grace-функционалов наиболее пригодны именно эти.

Некоторые из перечисленных ниже функционалов могут быть рассчитаны только для бинарных изображений (а именно функционалы 2, 3, 5), поэтому изображение сначала преобразуется в бинарное по следующему правилу. Все-

T	Свойство		Связь η_2 и η_1
	p	Φ	
(i1) $\alpha_T(v)$ имеет смысл	(i1) $\alpha_p(v)$, $\gamma_p(\alpha_T(v))$ имеют смысл	(i1)	$\eta_2 = \gamma_\Phi(\gamma_p(\alpha_T(v))\alpha_p(v))\eta_1$, если коэффициент имеет смысл
		(s1. $2\pi/n$), (s3)	$\eta_2 = \eta_1 + \theta(\text{modd}2\pi/n)$
	(s1), (s2)	(i1) $\forall u \Phi u = (\Phi, \perp)$	$\eta_2 = \gamma_\Phi(\mu)\eta_1$, если коэффициент имеет смысл
		(s1. $2\pi/n$), (s3) $\forall u \Phi u = (\Phi, \perp)$	$\eta_2 = \eta_1 + \theta(\text{modd}2\pi/n)$

го на каждом изображении имеем 16 оттенков (от черного до белого). Бинарное изображение получается с использованием порога Color Triger = 5.5.

1. Интеграл вдоль прямой (это преобразование ведет к преобразованию Радона).

2. Длина максимального сегмента при пересечении прямой l и образа, при этом пересечение считается, если только тон изображения больше числа Color Triger.

3. Число сегментов при пересечении прямой и образа (используется Color Triger).

5. Длина отрезка между первым касанием образа и последним. Математически это есть длина выпуклой оболочки суппорта (носителя) функции (используется Color Triger).

6. Дисперсия функции, предварительно нормированной на ее интеграл. Нормирование делается для того, чтобы использовать само понятие дисперсии, которое определено только для неотрицательных функций, интеграл от которых равен единице. Если функция была тождественным нулем, то считаем, что дисперсия равна нулю. Это необходимо для непрерывности trase- преобразования (хотя обычно считают, что дисперсия нулевой функции равна бесконечности).

7. Дисперсия функции, вычисленная, как указано в предыдущем пункте, затем умноженная на интеграл от функции. Это делается для того, чтобы учесть малость функции и придать малый вес случайному шуму и искажениям.

Матрица для вывода трансформанты (trase- матрица) имеет следующие размеры. По переменной φ ось горизонтальна, область изменения $0, \dots, 2\pi$, число дискрет 70. По переменной p ось направлена вверх, область изменения $-100, \dots, 100$, число дискрет 50.

Диаметральные функционалы. Для диаметрального функционала использовались пять вариантов:

1) гильбертова норма функции — это корень квадратный из интеграла от функции в квадрате;

2) максимум функции;

5) мера носителя функции (для таблично заданной функции это число ненулевых компонент, умноженное на шаг дискретизации);

6) максимум абсолютной величины первой производной;

9) вариация функции.

Цирковые функционалы. Для циркового функционала использовались четыре варианта:

7) амплитуда второй гармоники фурье-функции, деленная на максимум функции;

10) евклидова норма (т. е. гильбертова норма пространства L^2);

12) евклидова норма, деленная на вариацию функции;

14) амплитуда четвертой гармоники фурье-функции, деленная на евклидову норму функции.

Итого мы имеем $6 \times 5 \times 4 = 120$ разных функционалов.

Число обрабатываемых рисунков равно 35. Таким образом, должно получиться $120 \times 35 = 4200$ чисел, которые являются признаками. Некоторые из этих признаков не информативны, другие — слабо информативны для задачи различения патологических эритроцитов. Однако использование их в совокупности решило задачу распознавания классов полностью.

Каждое изображение, таким образом, характеризуем набором 120 чисел, которые образуют вектор. Далее вычисляется среднее расстояние в 120-мерном пространстве до представителей каждого класса изображений. Класс, расстояние до которого минимально, считается результатом процесса распознавания данного изображения.

Описание представленного расчетного материала. *Эритроциты.* На рис. 1 даны эритроциты и другие клетки крови. Из них выбрано пять различ-

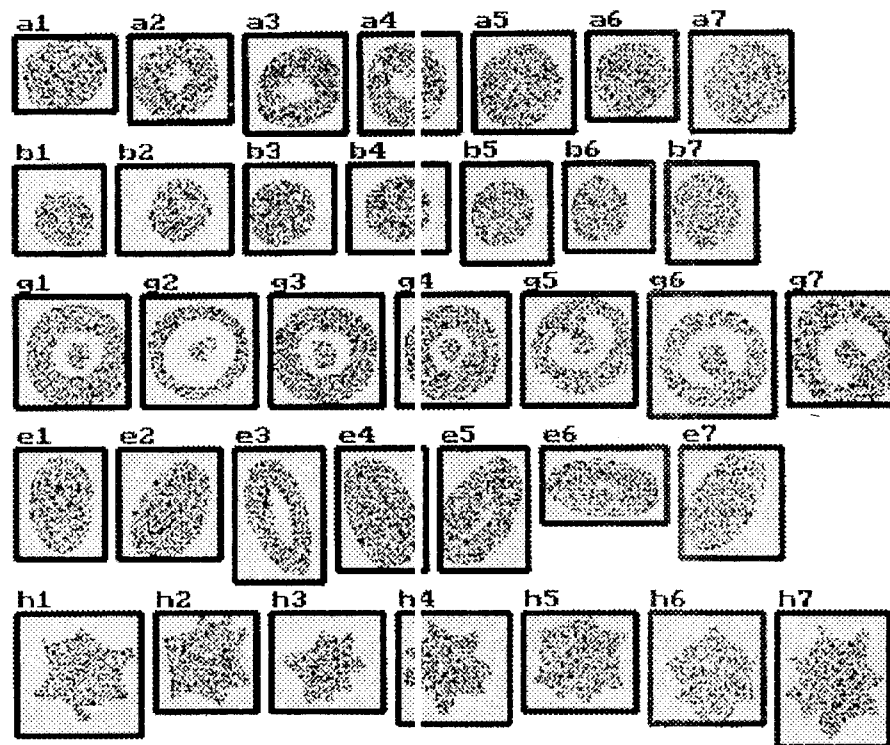


Рис. 1

ных классов: а, b, g, е, h. Каждый класс представлен семью образцами, т. е. имеем изображения:

a1, a2, a3, a4, a5, a6, a7 — класс а,

b1, b2, b3, b4, b5, b6, b7 — b,

g1, g2, g3, g4, g5, g6, g7 — g,

e1, e2, e3, e4, e5, e6, e7 — e,

h1, h2, h3, h4, h5, h6, h7 — h.

Кроме того, выделяем серии. Это изображения:

a1, b1, g1, e1, h1 — серия 1,

a2, b2, g2, e2, h2 — 2,

a3, b3, g3, e3, h3 — 3,

a4, b4, g4, e4, h4 — 4,

a5, b5, g5, e5, h5 — 5,

a6, b6, g6, e6, h6 — 6,

a7, b7, g7, e7, h7 — 7.

Задача состоит в отнесении каждого изображения к одному из классов.

Результаты сравнения серий. Результаты распознавания серий кровяных клеток представлены на рис. 2, а—ж. При расчетах использовано $1 \times 5 \times 4 = 20$ признаков. Результаты серии коррелируют примерно на 80 % при использовании 20 признаков при фиксированном trace-функционале.

Анализ результатов эксперимента. Можно сделать вывод, что большинство приведенных в этой работе признаков удовлетворительно решают поставленную задачу различения эритроцитов. Некоторые признаки хорошо работают, даже если исследователь не может увидеть явных различий. Кроме того, показано, что имеются признаки, более тонкие, чем требует данная конкретная задача. Это доказывает, что у предлагаемой системы различения есть перспективы.

В результате эксперимента по определению классов 35 представленных изображений мы не получили ни одной ошибки.

Суммы расстояний образов а, b, g, e, h серии 1 до всех других образов классов а, b, g, e, h						
	а	б	g	е	h	а
a1	1,03#	1,18	1,34	1,83	1,24	
b1	0,98	0,82#	1,17	2,17	1,04	
g1	0,94	1,06	0,80#	2,36	1,10	
e1	1,67	1,69	1,86	1,21#	1,66	
h1	1,32	1,22	1,24	1,87	0,93#	
Суммы расстояний образов а, b, g, e, h серии 2 до всех других образов классов а, b, g, e, h						
	а	б	g	е	h	б
a2	0,83#	0,85	0,90	2,29	1,11	
b2	1,01	0,82#	1,18	2,19	1,04	
g2	1,16	1,16	0,99#	2,29	1,21	
e2	1,67	1,83	1,92	1,17#	1,60	
h2	1,13	1,15	1,20	2,01	0,97#	
Суммы расстояний образов а, b, g, e, h серии 3 до всех других образов классов а, b, g, e, h						
	а	б	g	е	h	в
a3	0,79#	0,90	1,02	1,84	1,08	
b3	0,83	0,78#	0,96	2,20	1,03	
g3	0,94	1,06	0,80#	2,21	1,10	
e3	3,29	3,22	3,29	1,72#	2,90	
h3	1,29	1,10	1,32	1,80	0,86#	
Суммы расстояний образов а, b, g, e, h серии 4 до всех других образов классов а, b, g, e, h						
	а	б	g	е	h	г
a4	0,82#	0,86	0,93	2,08	1,10	
b4	1,07	0,91#	1,12	2,37	1,22	
g4	0,99	1,04	0,81#	2,30	1,19	
e4	1,93	2,05	2,13	0,95#	1,78	
h4	1,17	0,98	1,07	2,05	0,97#	
Суммы расстояний образов а, b, g, e, h серии 5 до всех других образов классов а, b, g, e, h						
	а	б	g	е	h	д
a5	0,76#	0,89	1,00	2,06	1,11	
b5	0,87	0,72#	1,00	2,16	0,98	
g5	0,92	1,00	0,79#	2,37	1,14	
e5	2,11	2,27	2,21	1,11#	1,90	
h5	1,07	1,12	1,15	1,97	0,91#	
Суммы расстояний образов а, b, g, e, h серии 6 до всех других образов классов а, b, g, e, h						
	а	б	g	е	h	е
a6	0,79#	0,93	0,92	2,28	1,24	
b6	0,89	0,82#	1,13	2,08	1,12	
g6	1,19	1,21	1,00#	2,25	1,29	
e6	2,65	2,81	2,91	1,34#	2,46	
h6	1,16	1,03	1,10	1,98	0,90#	
Суммы расстояний образов а, b, g, e, h серии 7 до всех других образов классов а, b, g, e, h						
	а	б	g	е	h	ж
a7	0,82#	0,90	0,96	2,17	1,12	
b7	0,88	0,83#	1,03	2,10	1,08	
g7	1,06	1,20	0,92#	2,18	1,16	
e7	2,06	2,15	2,27	1,07#	1,91	
h7	1,07	1,05	1,14	1,80	0,92#	

Рис. 2

Заключение. Установлено, что предлагаемые признаки, действительно, хорошо различают предложенные для анализа классы эритроцитов. Этим подтверждена полезность для медицинской практики данной теории. Рассмотренные признаки удовлетворяют условиям теоремы, приведенным в таблице, следовательно, эти признаки инвариантны к движениям изображенных объектов.

Существенным является то, что предложено сразу много — порядка сотни — хороших признаков. Подход открывает новые возможности автоматического распознавания биологических микрообъектов.

СПИСОК ЛИТЕРАТУРЫ

1. Федотов Н. Г., Кадыров А. А. Новые методы формирования признаков распознавания образов с позиций стохастической геометрии // Автометрия. 1996. № 1. С. 88.
2. Федотов Н. Г. Методы стохастической геометрии в распознавании образов. М.: Радио и связь, 1990.
3. Fedotov N. G., Larin M. Computer vision and stochastic geometry // 4th Int. Conf. Artificial Intelligence: Methodology, Systems, Applications AIMS'A'90 (Varna, Bulgaria) /Ed. P. Jorrand, V. Sgurev. Amsterdam — Oxford — New York — Tokyo, 1990. P. 270.
4. Fedotov N. G. Stochastic geometry techniques in pattern recognition // Proc. Latvian Signal Proc. Int. Conf. LISP'90. Riga, USSR, 1990. Vol. 1. P. 256.
5. Fedotov N. G. Perception of vision information in AI systems in view of stochastic systems // Proc. 3d Conf. Artificial Intelligence'92, Former Soviet Union, Tver., 1992. Vol. 1. P. 160.
6. Kadyrov A. A., Saveleva M. V., Fedotov N. G. Image scanning leads to alternative understanding of image // Proc. of the Third Int. Conf. on Automation, Robotics and Computer Vision ICARCV'94. Singapore, 1994. Vol. 3. P. 2029.
7. Fedotov N. G., Kadyrov A. A. Image scanning in machine vision leads to new understanding of image // Proc. of 5th Int. Workshop on Digital Image Proc. and Computer Graphics. Samara, Russia: SPIE, 1994.
8. Kadyrov A. Triple features for linear distorted images // The 6th Int. Conf. "Computer Analysis of Images and Patterns (CAIP'95)". Prague, Czech. Republic, 1995 (Springer Verlag Lectures Notes on Computer Science).

Поступила в редакцию 30 октября 1995 г.