

УДК 681.3.07

НОРМАЛИЗАЦИЯ ИЗОБРАЖЕНИЙ: ДОСТИЖЕНИЯ И НЕРЕШЕННЫЕ ЗАДАЧИ

Е. П. Путятин

ХНУРЭ, г. Харьков, Украина

Приведены новые результаты по нормализации аффинных, проективных и нелинейных преобразований. Обсуждаются проблемные вопросы нормализации общих преобразований двумерных изображений. В сравнительном аспекте рассматриваются основные математические методы нормализации: метод моментов, интегральной и одномерной последовательной корреляции. Описан подход к нормализации сложных групп, включая проективную, путем синтеза частичных нормализаторов, соответствующих подгруппам разложения.

КОМПЬЮТЕРНОЕ ЗРЕНИЕ, ИНТЕЛЛЕКТ, ИЗОБРАЖЕНИЕ, НОРМАЛИЗАЦИЯ, РАСПОЗНАВАНИЕ, ПРЕОБРАЗОВАНИЕ, АФФИННЫЕ, НЕЛИНЕЙНЫЕ, ЯРКОСТЬ, КОРРЕЛЯЦИЯ, МОМЕНТЫ.

Введение

Проблеме построения систем компьютерного зрения (СКЗ) уделяется все возрастающее внимание. Одним из основных направлений совершенствования таких систем является придание им свойств человеческого зрения*. Бионика зрительно-интеллекта как наука лежит в основе СКЗ транспортных и технологических роботов различного назначения. В широком смысле интеллектуальные СКЗ решают две взаимно дополняющие задачи: правильное восприятие образа и его анализ, понимание и распознавание. Задача нормализации заключается в обеспечении правильного восприятия, при котором создается некоторый стандартный вид изображения: его размеры, ориентация, распределение яркости. В принципе, стандарты или эталоны могут быть разными для глаза человека и робота, хотя в целом требования часто совпадают. Нормализация имеет важное практическое значение не только для распознавания изображений, облегчая сравнение с эталонами и увеличивая резерв времени на принятие решений на заключительном этапе. Существуют и другие области, непосредственно не связанные с распознаванием. Например, сопровождение подвижных объектов (охранная сигнализация, компьютерные игры и т. д.). Важное значение имеет «выравнивание» поля зрения при наблюдении реальных сцен из замкнутого пространства, например, из кабины пилота.

Ограничимся одиночными плоскими изображениями, заданными на однородном фоне, т. е. исключим задачу сегментации отдельных изображений из многоплановых сцен. Несмотря на это существенное упрощение, проблема нормализации, как мы

увидим ниже, остается сложной. Это связано с тем, что входные изображения отличаются от эталонных геометрическими преобразованиями: аффинными, проективными, нелинейными, преобразованиями яркости. Несмотря на такие отличия, человеческое зрение уверенно справляется со многими задачами нормализации изображений, чего еще нельзя сказать о современном компьютерном зрении.

Рассмотрим проблему последовательно. Вначале остановимся на моделях входных сигналов, затем исследуем возможности их нормализации.

1. Математические модели входных сигналов

В табл. 1 приведена классификация типов преобразований, с которыми сталкивается человек в повседневной жизни. Преобразования условно разделены на пять классов. Буквенные обозначения приведены в таблице справа, номера преобразований — слева. Преобразования из класса I являются базовым для других классов. Они образуют группы, причем преобразования 1–3 и 8–10 образуют аддитивные группы со значением единичного элемента, равным нулю, остальные (4–7) — мультипликативные группы с параметром единица.

Многопараметрические преобразования из класса II также образуют группы, элементами которых являются матрицы, составленные из параметров преобразований, так что единичному элементу соответствует единичная матрица. Удобно использовать матрицы (3×3) , так что аффинные преобразования 21 характеризуются матрицей вида A , множество которых образуют группу G_a .

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & 0 & 1 \end{pmatrix}, \det |A| \neq 0 \quad A, \quad (1)$$

где a_j — параметры преобразования ($i=1,2, j=1,3$).

* Исследования по математическому моделированию зрения, в том числе нормализации начаты в университете радиоэлектроники в 1963 г. [1] по инициативе моего учителя проф. Ю. П. Шабанова-Кушнарченко, которому посвящается данная работа.

Математические модели входных сигналов интеллектуальных СКЗ

I. Однопараметрические преобразования			Обозначение
1.	$x \rightarrow x + a \cos \gamma, y \rightarrow y + a \sin \gamma$, где: a — параметр	Смещения вдоль направления $(\cos \gamma, \sin \gamma)$	$G_{c,\gamma}$
2.	$x \rightarrow x + \alpha, y \rightarrow y$, где: α — параметр	Горизонтальные смещения	$G_{c,x}$
3.	$x \rightarrow x, y \rightarrow y + \beta$, где: β — параметр	Вертикальные смещения	$G_{c,y}$
4.	$x \rightarrow kx, y \rightarrow ky$, где: $k \neq 0$ — параметр	Подобия изменения масштаба	G_m
5.	$x \rightarrow kx, y \rightarrow y$, где: $k \neq 0$ — параметр	Горизонтальные растяжения	$G_{m,x}$
6.	$x \rightarrow x, y \rightarrow ky$, где: $k \neq 0$ — параметр	Вертикальные растяжения	$G_{m,y}$
7.	$x \rightarrow qx, y \rightarrow q^{-1}y$, где: $q \neq 0$ — параметр	Гиперболические растяжения	G_q
8.	$x \rightarrow x \cos \gamma + y \sin \gamma,$ $y \rightarrow -x \sin \gamma + y \cos \gamma$, где: γ — параметр	Повороты	G_u
9.	$x \rightarrow x + hy, y \rightarrow y$, где: h — параметр сдвига	Косые сдвиги (вдоль оси OX)	$G_{h,x}$
10.	$x \rightarrow x, y \rightarrow y + hx$, где: h — параметр сдвига	Косые сдвиги (вдоль оси OY)	$G_{h,y}$
II. Многопараметрические преобразования			
11.	$x \rightarrow x + \alpha, y \rightarrow y + \beta$, где: α, β — параметры смещения	Двумерные смещения	G_c
12.	$x \rightarrow kx + \alpha, y \rightarrow y$, где: k, α — параметры, $k \neq 0$	Одномерные смещения и растяжение по оси OX	$G_{c,m,x}$
13.	$x \rightarrow x, y \rightarrow ky + \beta$, где: k, β — параметры, $k \neq 0$	Одномерные смещения и растяжение по оси OY	$G_{c,m,y}$
14.	$x \rightarrow kx + \alpha, y \rightarrow ky + \beta$, где: k, α, β — параметры, $k \neq 0$	Смещения и изменения масштаба	$G_{c,m}$
15.	$x \rightarrow k_1x, y \rightarrow k_2y$, где: k_1, k_2 — параметры, $k_1 \neq 0, k_2 \neq 0$	Неоднородные изменения масштаба	G_d
16.	$x \rightarrow k_1x + \alpha, y \rightarrow k_2y + \beta$, где: k_1, k_2, α, β — параметры, $k_1 \neq 0, k_2 \neq 0$	Смещения и неоднородные изменения масштаба	$G_{c,d}$
17.	$x \rightarrow x \cos \gamma + y \sin \gamma + \alpha,$ $y \rightarrow -x \sin \gamma + y \cos \gamma + \beta,$ где: γ, α, β — параметры	Смещения и повороты (группа движений, или Евклидова группа)	$G_{c,u}$
18.	$x \rightarrow kx \cos \gamma + ky \sin \gamma,$ $y \rightarrow -kx \sin \gamma + ky \cos \gamma,$ где: γ, k — параметры	Изменения масштаба и повороты	$G_{m,u}$
19.	$x \rightarrow kx \cos \gamma + ky \sin \gamma + \alpha,$ $y \rightarrow -kx \sin \gamma + ky \cos \gamma + \beta,$ где: γ, k, α, β — параметры	Смещения, изменения масштаба, повороты (метрическая группа)	$G_{m,c,u}$
20.	$x \rightarrow a_{11}x + a_{12}y, y \rightarrow a_{21}x + a_{22}y,$ где: $a_{11}, a_{12}, a_{21}, a_{22}$ — параметры $\det \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \neq 0$	Центроаффинная группа (неоднородные изменения масштаба, косые сдвиги, повороты и отражения)	G_a^c
21.	$x \rightarrow a_{11}x + a_{12}y + \alpha, y \rightarrow a_{21}x + a_{22}y + \beta,$ где: $a_{11}, a_{12}, a_{21}, a_{22}$ — параметры $\det \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \neq 0$	Группа аффинных преобразований (неоднородные изменения масштаба, косые сдвиги, повороты, отражения и смещения)	G_a
III. Нелинейные преобразования			
22.	$x \rightarrow \frac{x}{(\lambda-1)y+1}, y \rightarrow \frac{\lambda(y)}{(\lambda-1)y+1},$ где: λ — параметр преобразования	Перспектива (гиперболическая гомология)	G_r
23.	$x \rightarrow \frac{x}{\lambda y+1}, y \rightarrow \frac{y}{\lambda y+1},$ где: λ — параметр преобразования	Перспектива (параболическая гомология)	G_p

24.	$x \rightarrow \frac{a_{11}x + a_{12}y + a_{13}}{a_{31}x + a_{32}y + a_{33}}, y \rightarrow \frac{a_{21}x + a_{22}y + a_{23}}{a_{31}x + a_{32}y + a_{33}}$, где: a_{ij} — параметры ($i, j = 1, 2, 3$)	Проективные преобразования (общий случай)	G
25.	$x \rightarrow e^{ay}x, y \rightarrow y; x \rightarrow x, y \rightarrow e^{ax}y$, где: a — параметр преобразования	Преобразование масштаба с переменным коэффициентом	G_s
26.	$x \rightarrow a_1x^2 + a_2y^2 + a_3, y \rightarrow a_4x^2 + a_5y^2 + a_6$, где: a_i — параметры ($i = \overline{1,6}$)	Квадратичные преобразования (дисторсия линз)	K
27.	$x \rightarrow a_1x^2 + a_2xy + a_3y^2 + a_4x + a_5y + a_6$, $y \rightarrow a_7x^2 + a_8xy + a_9y^2 + a_{10}x + a_{11}y + a_{12}$, где: a_i — параметры ($i = \overline{1,12}$)	Полные квадратичные преобразования (кривизна Земли)	Z
IV. Непараметрические нелинейные преобразования			
28.	$x \rightarrow x, y \rightarrow y; x \rightarrow -x, y \rightarrow -y$; $x \rightarrow x, y \rightarrow -y; x \rightarrow -x, y \rightarrow y$	Зеркальные отражения	G_s
29.	$x \rightarrow \pm \frac{x}{y}, y \rightarrow \pm \frac{1}{y}; x \rightarrow \pm \frac{y}{x}, y \rightarrow \pm \frac{1}{x}$; $x \rightarrow \pm \frac{1}{x}, y \rightarrow \pm \frac{y}{x}; x \rightarrow \pm \frac{1}{y}, y \rightarrow \pm \frac{x}{y}$	Ортогональные проективные преобразования	G_o
30.	$x \rightarrow f_1(x), y \rightarrow f_2(y); (x \rightarrow x^2, y \rightarrow y^2)$; $x \rightarrow f_1(x, y), y \rightarrow f_2(x, y)$	Общий вид преобразования координат	
V. Преобразования яркости			
31.	$B(x, y) = kB_0(x, y) + c$, где $B(x, y)$ — функция яркости входного изображения, $B_0(x, y)$ — функция яркости эталона, k, c — параметры группы	Постоянный фильтр и подсветка	G_k, G_c, G_{kc}
32.	$B(x, y) = K(x, y)B_0(x, y) + C(x, y)$, где: $K(x, y), C(x, y)$ — мультипликативная и аддитивная составляющие искажений	Общий случай преобразований яркости	
33.	$B(x, y) = K(x, y)B_0[T(x, y)] + C(x, y)$, где: $T(x, y)$ — геометрические преобразования	Комбинированный случай (эквализация гистограмм, борьба с помехами в условиях геометрической нормализации)	

Преобразованиям 11–20 соответствуют различные значения параметров матрицы (1), например, для смещений 11: $a_{13} = \alpha, a_{23} = \beta, a_{11} = a_{22} = 1, a_{21} = a_{12} = 0$; для косоуго сдвига 9: $a_{12} = h, a_{11} = a_{22} = 1, a_{21} = a_{13} = a_{23} = 0$. Зеркальные отражения 28 (симметрия) образуют дискретную группу G_s , которая естественным образом входит в базовые преобразования аффинной группы в предположении, что параметры a_{11}, a_{22} могут иметь любой знак (+, -).

Наиболее сложными преобразованиями в рамках групповой модели G являются проективные преобразования 24, матрица которых имеет вид:

$$P = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}, \det(P) \neq 0, a_{ij} (i, j = \overline{1,3}). \quad (2)$$

Составляющими проективной группы являются аффинные, ортогональные проективные преобразования 29, 30 и преобразования перспективы 22, 23.

Человек может распознавать в определенных пределах изображения, спроектированные на нелинейную поверхность, например, Землю, компенсировать нелинейные искажения, вызванные дисторсией линз, и т. д. В какой-то степени они моделируются преобразованиями 25–27 (из них только преобразования 25 образуют группу). Нелинейные изменения пространства восприятия сцены описываются преобразованием 32. Простейшие преобразования яркости 31 хорошо описывают изменение яркости объекта, когда Солнце зашло за облако, а также наличие однородных фильтров и подсветки в СКЗ. Такие преобразования также образуют группы [2].

Последняя графа таблицы 1 описывает отличия входного изображения $B(x, y)$ от эталонного $B_0(x, y)$, включая шумы и локальные помехи $C(x, y)$, неравномерную чувствительность фотокамеры или сенсоров по пространству $K(x, y)$.

С геометрической точки зрения аффинные преобразования первого и второго классов не требуют

особых комментариев. На рис. 1. приведена иллюстрация косо́го сдвига вдоль оси абсцисс. Проективные преобразования G (рис. 2) возникают в результате поворота объекта относительно оси, перпендикулярной оси наблюдения. Перспектива G_p обычно ассоциируется с уходящими вдаль рельсами железной дороги (рис. 3). Понятно, что численное моделирование проективной группы будет не совсем адекватно теоретическим моделям ввиду невозможности реализации точек на бесконечности. Проблема восприятия и нормализации возникает также при анализе ортогональной подгруппы G_0 проективной группы G . На рис. 4, 5 показано, когда изменяется положение точек при таких преобразованиях до неузнаваемости. С этой группой связаны дополнительные ограничения при нормализации. На рис. 6 проиллюстрировано преобразование 25 нелинейной группы G_3 .



Рис. 1. Косой сдвиг $B_0(x, y) = B(x + y \operatorname{tg}(\alpha), y)$, где α — угол сдвига

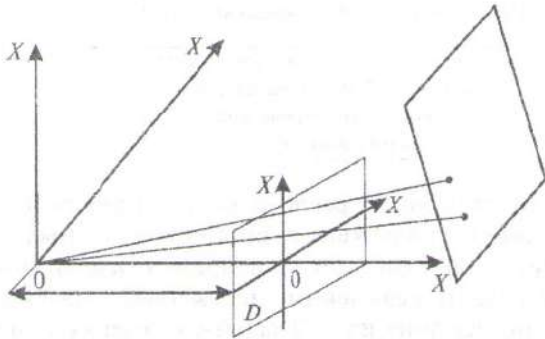


Рис. 2. Возникновение проективных преобразований

При нормализации важное значение имеет изучение свойств преобразований, причем внимание уделяется не только известным геометрическим свойствам типа инвариантов, скажем, для аффинной группы и ее подгрупп. Новым является изучение возможности применения конкретного преобразования в составе суперпозиции, что важно при синтезе нормализаторов. С этой точки зрения преобразования могут быть коммутативными в данной суперпозиции, перестановочными или неперестановочными. Для перестановочных, в отличие от коммутативных, допускается изменение величины параметра, но не самого вида преобразования.

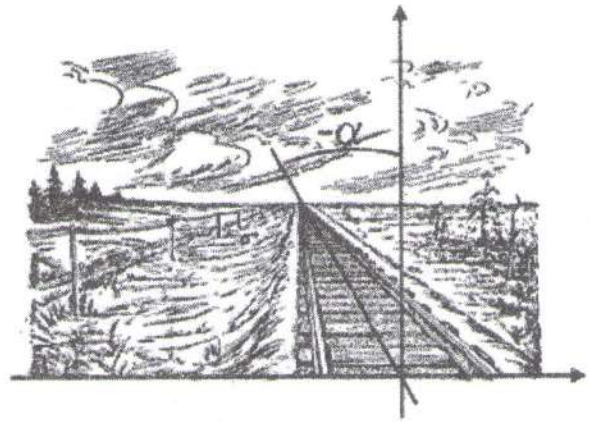


Рис. 3. Пример преобразования перспективы вдоль оси с углом наклона α

$$B_0(x, y) = B \left(\frac{x}{(\lambda \sin(-\alpha) + \lambda \cos(-\alpha))y + 1}, \frac{y}{(\lambda \sin(-\alpha) + \lambda \cos(-\alpha))y + 1} \right),$$

где λ — параметр перспективы;
 B, B_0 — входное и эталонное изображения

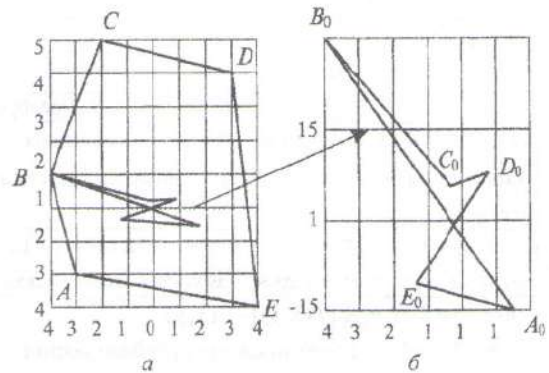


Рис. 4

a — пример искажения фигуры $ABCDE$, внутри которой находится начало координат, преобразованием $X = x/y, Y = 1/y$ в фигуру $A_0B_0C_0D_0E_0$;
 b — фигура $A_0B_0C_0D_0E_0$, увеличенная в 5 раз

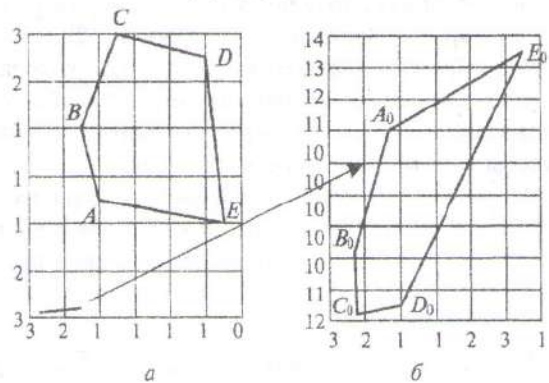


Рис. 5

a — пример искажения фигуры $ABCDE$, расположенной в первой четверти, преобразованием $X = x/y, Y = 1/y$ в фигуру $A_0B_0C_0D_0E_0$;
 b — та же фигура, увеличенная в 70 раз

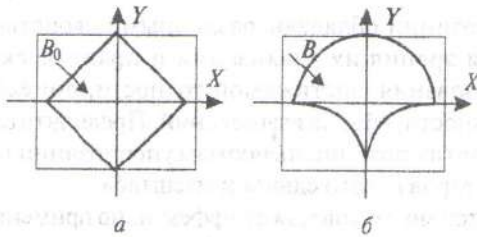


Рис. 6. Нелинейное масштабирование

В нижних рядах таблицы классификации приведены входные сигналы общего вида, соответствующие реальным сценам, для которых нормализация — это получение стандартного фона и контраста (группы G_k, G_c, G_{kc}). Более того, процедуры подавления помех, выравнивания гистограмм и т. д. кажутся вполне естественными нормализаторами.

2. Математические модели нормализации

Прежде всего обратим внимание на интегральный подход, основанный на преобразовании всего поля изображения. Например, модуль преобразования Фурье обеспечивает автоматическое центрирование, т. е. компенсацию параметров α, β группы G_c , преобразование Меллина компенсирует преобразование масштаба.

Однако этот подход имеет ограниченную область применения, т. к. не вычисляет величины параметров отличия эталонных от входных изображений. Кроме того, цифровая реализация преобразований Меллина весьма затруднительна.

Другой подход к построению нормализаторов основан на непосредственном вычислении параметров и действия обратных преобразований, построенных на их базе. Несмотря на то, что обратные преобразования формально существуют для всех моделей входных сигналов из табл. 1, полностью обоснованные модели нормализации возможны только для преобразований, образующих группы.

В самом деле, каким бы ни было сложным преобразование входного изображения, если множество этих преобразований образует группу, то всегда есть тождественное преобразование, характеризующее эталон, и существует обратное преобразование, с помощью которого можно получить этот эталон. Результат последовательного выполнения двух преобразований также принадлежит той же группе. Такими свойствами обладают все преобразования из табл. 1, помеченные справа буквой G (с индексом и без).

Насколько обоснованными являются эти рассуждения, видно из такого примера. Преобразования G_d с неодинаковым изменением масштаба по осям и преобразования поворота G_u сами по себе образуют группы, однако их совместное действие группы не образует. Следовательно, эталон, опре-

деленный единицами групп ($\gamma=0, k_1=k_2=1$) не существует, так как последовательное выполнение нескольких таких преобразований выведет итоговое изображение за пределы трех параметров как основной характеристики исходных групп. Понадобится еще один параметр (например, косой сдвиг или другой поворот), но тогда получается группа уже с четырьмя параметрами (центроаффинная G_a^c). Заметим, что суперпозиция поворота и одинакового масштаба образует группу, и можно говорить об эталоне и нормализации.

Аппарат теории групп позволяет с единых позиций рассмотреть задачу нормализации входных сигналов, приведенных в табл. 1.

Пусть M — множество входных изображений B , M_0 — множество эталонов B_0 , G — группа преобразований, $g \in G$ — элемент группы, т. е. конкретное преобразование. Рассматривается действие группы G на множестве входных изображений.

Оператором нормализации F относительно действия группы G называется отображение $F: M \rightarrow M_0$, для которого существует $g \in G$ такое, что $F(B) = B_0$ для любого $B \in M$.

При этом учитывается, что различные изображения $B \in M$ образуют классы эквивалентности, если $B = gB_0, g \in G$. Общий вид оператора нормализации обоснован в [2, 4] и имеет вид:

$$F(B) = \Phi(B)B, \quad \Phi: M \rightarrow G, \Phi(B) = g^{-1}, g \in G, B \in M. \quad (3)$$

Отображение $\Phi: M \rightarrow G$ представляет собой вещественный функционал для однопараметрических групп и вектор-функцию для многопараметрических.

Например, для параллельных смещений (группа G_c), когда входной сигнал определяется классом эквивалентности

$$B(x, y) = gB_0(x, y) = B_0(x - \alpha, y - \beta),$$

имеем:

$$F(B) = \Phi(B)B = B(x + \Phi_1(B), y + \Phi_2(B)), \quad (4)$$

где $\Phi: M \rightarrow G_c$ определяется двумя вещественными функционалами $\Phi_1(B)$ и $\Phi_2(B)$.

Физический смысл нормализатора (4) заключается в компенсации смещений α, β , т. е. центрировании. При этом $\Phi_1(B) = \alpha^{-1}, \Phi_2(B) = \beta^{-1}$.

Для преобразований перспективы G_p (параболическая гомология) входной сигнал определяется классом эквивалентности

$$B(x, y) = gB_0(x, y) = B_0\left(\frac{x}{\lambda y + 1}, \frac{y}{\lambda y + 1}\right),$$

а нормализатор имеет вид:

$$F(B) = \Phi(B)B = B \left(\frac{x}{\Phi(B)y+1}, \frac{y}{\Phi(B)y+1} \right), \quad (5)$$

где Φ — функционал, вычисляющий параметр λ .

Аналогично можно выписать модели нормализаторов и для всех других типов входных сигналов, образующих группы, однако понятно, что основная трудность ложится на поиск подходящих отображений $\Phi(B)$.

3. Аффинная нормализация

В принципе существует два подхода к построению вектор-функции $\Phi(B)$: параллельный и последовательный. Параллельный подход основан на использовании интегральных моментов в форме линейных и нелинейных функционалов. Зависимость между моментами для эталонных изображений и преобразованных (искаженных) выражается в виде системы уравнений, решение которой позволяет определить параметры a_{ij} матрицы (1). Параллельная нормализация хорошо зарекомендовала себя в режиме определения малых приращений параметров с целью управления, например, рамкой слежения в системах динамического видеонаблюдения [5].

Последовательный подход обеспечивает вычисление параметров преобразований, а не их приращений, и лежит в основе целого ряда технических устройств, именуемых координаторами положения объектов в пространстве. Теоретической предпосылкой вычисления параметров являются найденные различные способы представления аффинной группы в виде суперпозиции подгрупп. Если ранее из классической геометрии для центроаффинной группы следовало только разложение: поворот (U_1) — преобразование масштаба (D) — поворот (U_2), то есть $G_a^c = G_u G_d G_{u_2}$, то теперь найден ряд других возможных разложений центроаффинной группы:

$$\begin{aligned} &G_u G_d G_{h,y}, G_u G_d G_{h,x}, G_d G_u G_{h,x}, \\ &G_d G_u G_{h,y}, G_d G_{h,x} G_u, G_d G_{h,y} G_u, \quad (6) \\ &G_d G_u G_{m,x}, G_d G_u G_{m,y}, G_d G_{h,x} G_{h,y}, \end{aligned}$$

$$G_d G_{h,y} G_{h,x}, G_{h,x} G_d G_{h,y}, G_{h,x} G_d G_{h,y}, G_{h,y} G_d G_{h,x},$$

где обозначения подгрупп соответствуют обозначениям, приведенным в табл. 1.

В основе последовательной нормализации полной аффинной группы лежит синтез частичных нормализаторов, соответствующих подгруппам разложений (6) и нормализатора смещений F_c , согласно (4). Как сами частичные нормализаторы, так и их

суперпозиции обладают различными свойствами с точки зрения их реализации и практического использования: достижимой точности, помехозащищенности, объема вычислений. Предпочтительными с этих позиций являются суперпозиции нормализаторов косоугольного сдвига и масштаба.

Метод моментов может эффективно применяться не только для параллельной, но и для последовательной нормализации, однако его сравнительно низкая помехозащищенность вынуждает искать другие подходы.

Разработана модификация метода частных корреляций, именно корреляция вдоль прямой, образованной пересечением полутонового изображения и перпендикулярной к нему плоскости. В этом случае корреляция по плоскому изображению заменяется корреляцией вдоль прямой. Задача состоит в поиске максимального сходства распределения яркости вдоль прямых на эталонном и входном изображениях путем перебора с некоторым шагом всех положений прямых [6].

Для центрированных изображений, когда начало координат совпадает с обобщенным центром тяжести, поиск параметров аффинных преобразований сводится к последовательной корреляции двух пар прямых эталонного и входного изображений по масштабу.

Сравнительный анализ экспериментальных данных центроаффинной нормализации позволил сделать следующие выводы по трем методам: интегральной корреляции (ИК) (по всем четырем параметрам), одномерной корреляции (ОК) и применении моментов. Так, метод одномерной корреляции превосходит по быстродействию метод интегральной корреляции на три порядка, но на порядок уступает методу моментов. По помехозащищенности правильное распознавание объектов с вероятностью 0,96 достигалось для случайного шума методом ИК при $\mu > 3$, методом ОК при $\mu > 7$, методом моментов при $\mu > 10$, где μ — отношение сигнал/шум \bar{B}/δ , \bar{B} — средняя яркость изображения, δ — среднеквадратическое отклонение.

Иная картина наблюдается при действии локальных помех. Пусть $\mu_l = \frac{S_n}{S_c} \cdot 100$, где S_n — площадь помехи на изображении, S_c — площадь полезного сигнала. Метод ОК позволяет производить распознавание с вероятностью 0,96 при $\mu_l \leq 24\%$, метод ИК работает только до 9%, а метод моментов — до 5%. Экспериментальные данные подтверждают целесообразность применения метода моментов в системах нормализации реального времени при слежении за объектами. В то же время метод одномерных корреляций применим для решения задач распознавания изображений в условиях аффинных

преобразований и действия случайных и даже локальных помех.

Большой интерес представляют методы математической морфологии для определения параметров преобразований в задачах нормализации.

На данный момент успешным оказался опыт применения морфологии для аффинных преобразований. В основу положена возможность представления полутоновых изображений сечениями, причем с позиций функционального анализа сечения можно конструировать по диапазонам яркости с любой точностью приближения к исходному изображению и даже с переменным порогом. Сечения — бинарные образы, к ним применимы операции алгебры Минковского. С помощью таких операций разработаны методы выделения особых точек: крайних точек и скелетов для силуэтных изображений [7]. На рис. 7, 8 показаны примеры работы программ на реальных изображениях.

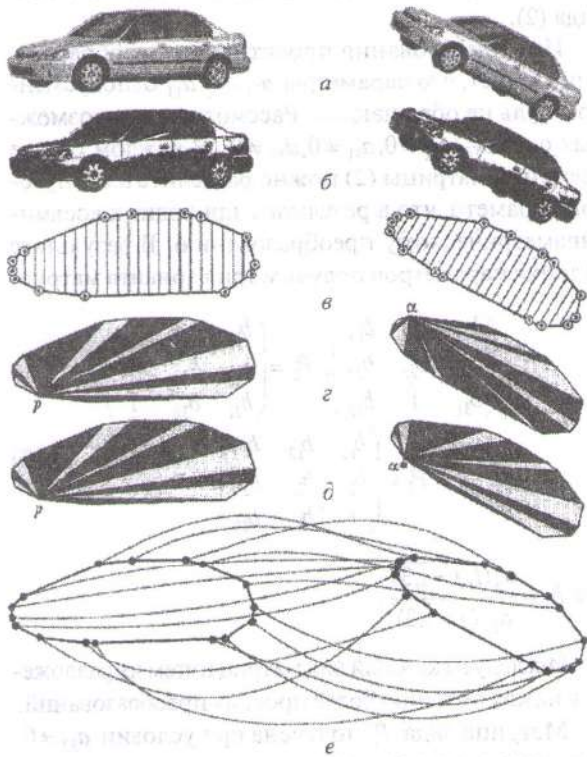


Рис. 7. Метод крайних точек:
 а — эталон и исходное изображение; б — их бинарные сечения; в — выделенные крайние точки;
 г, д — иллюстрация поиска соответствующих точек;
 е — результат поиска

После выделения особых точек (крайних на рис. 7 и точек скелетона на рис. 8) возникает задача поиска соответствий особых точек эталона и входного изображения. Решается она путем анализа возможных подстановок из множеств особых точек и с

учетом известного из геометрии инварианта постоянства площади объекта при аффинных преобразованиях. В конечном счете, переопределенная система линейных уравнений относительно параметров аффинных преобразований позволяет с высокой точностью (методом наименьших квадратов) находить параметры аффинных преобразований.

Еще предстоит отыскать возможности применения методов морфологии для нормализации проективных преобразований.

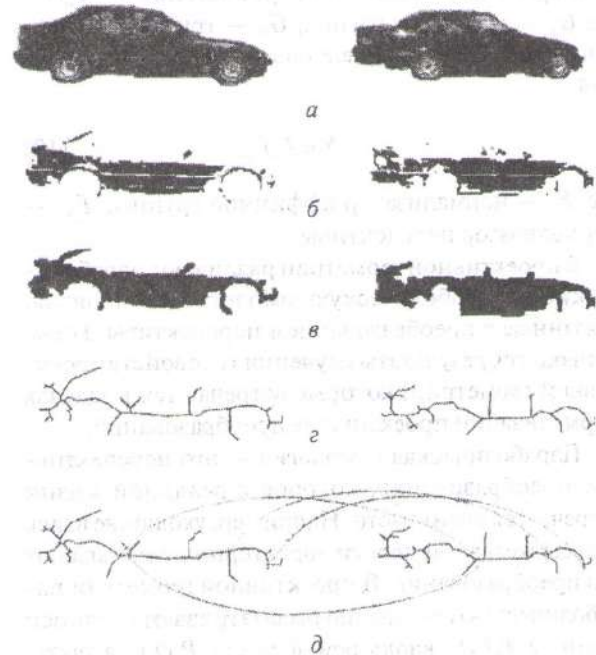


Рис. 8. Метод скелетов:
 а — эталон и исходное изображение; б — сечения;
 в — улучшенные сечения; г — скелеты;
 д — результат нахождения соответствующих точек

4. Нормализация проективных преобразований

Проективную группу G также можно представить в виде композиции более простых, основными из которых являются аффинная, ортогональная и группа перспективы. Класс эквивалентности входных сигналов задается соотношением:

$$B(x, y) = B_0 \left(\frac{a_{11}x + a_{12}y + a_{13}}{a_{31}x + a_{32}y + a_{33}}, \frac{a_{21}x + a_{22}y + a_{23}}{a_{31}x + a_{32}y + a_{33}} \right). \quad (7)$$

Частным случаем (подгруппой) проективной группы является центропроективная группа G_0 , определяемая классом эквивалентности:

$$B(x, y) = B_0 \left(\frac{a_{11}x + a_{12}y}{a_{31}x + a_{32}y + 1}, \frac{a_{21}x + a_{22}y}{a_{31}x + a_{32}y + 1} \right). \quad (8)$$

Существуют два подхода к синтезу нормализатора полной проективной группы [8]. Один из них основан на разложении $G = \bar{G}C'$, где \bar{G} — центро-проективная группа, C' — группа центрирования ($a_{13} = 0, a_{23} = 0$). Соответствующий нормализатор имеет вид:

$$F = F_0 F'_C, \quad (9)$$

где F_0 — нормализатор группы \bar{G} , F'_C — нормализатор проективного центрирования.

Другой подход использует разложение $G = G_a G_p$, где G_a — аффинная группа, G_p — группа перспективы. Нормализатор представляется как суперпозиция:

$$F = F_a F_p, \quad (10)$$

где F_a — нормализатор аффинной группы, F_p — нормализатор перспективы.

В проективной геометрии различают параболическую и гиперболическую гомологии как понятия, связанные с преобразованием перспективы. Ниже приводятся результаты изучения тех свойств проективной геометрии, которые встречаются в задачах нормализации проективных преобразований.

Параболическая гомология — это перспективное преобразование, которое в реальной жизни встречается очень часто. Например, уходящие вдаль рельсы железной дороги характерно представляют это преобразование. В проективной геометрии параболическую гомологию рассматривают вдоль оси ординат $P_y(\lambda)$, вдоль оси абсцисс $P_x(\lambda)$, а также вдоль прямой с угловым коэффициентом $P_\alpha(\lambda)$ (рис. 3). Очевидно, что

$$P_x(\lambda) = U(90)P_y(\lambda),$$

где $U(90)$ — поворот на угол 90 градусов.

Между гиперболической и параболической гомологиями существует прямая зависимость по параметрам сжатия $G_a(1, \lambda)$. Сказанное позволяет рассматривать только параболическую гомологию, которую будем называть просто перспективой.

Исследования перспективных преобразований позволили сформулировать критерий присутствия проективных преобразований на изображении.

Если энергия изображения изменяется пропорционально изменению площади, то на изображении присутствуют преобразования, отличные от аффинных. Если эти преобразования вписываются в проективную модель, то перераспределение энергии происходит вдоль оси преобразования перспективы.

На основании этого критерия функционал Φ , определяющий параметр перспективы, для нормализатора (5) можно определить в виде:

$$\lambda = \Phi = \frac{E_B}{S} - \frac{E_{B_0}}{S_0} = 0, \quad (11)$$

где $E_B = \iint B(x, y) dx dy$ — энергия изображения $B(x, y)$, S — площадь изображения $B(x, y)$,

$E_{B_0} = \iint B_0(x, y) dx dy$ — энергия изображения $B_0(x, y)$, S_0 — площадь изображения $B_0(x, y)$.

В (11) параметр λ задан неявным образом. Процедура нормализации заключается в многократном преобразовании перспективы, и при этом после каждого шага вычисляется функционал Φ . Преобразования продолжают до тех пор, пока функционал не будет равен нулю. Этот принцип осуществления нормализации соответствует построению нормализатора следящего типа [2].

Вернемся теперь к общей постановке вопроса, а именно, разработке математической модели нормализации в рамках модели восприятия, с матрицей вида (2).

Из существования проективных преобразований следует, что параметры a_{31}, a_{32}, a_{33} одновременно в нуль не обращаются. Рассмотрим три возможных случая: $a_{32} \neq 0, a_{31} \neq 0, a_{33} \neq 0$. В каждом случае элементы матрицы (2) можно разделить на ненулевой параметр, что в результате приводит к восьми-параметрическому преобразованию. В результате деления параметров получим три варианта матриц:

$$P_1 = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & 1 & b_{33} \end{pmatrix}, P_2 = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & 1 \end{pmatrix}, \quad (12)$$

$$P_3 = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ 1 & b_{32} & b_{33} \end{pmatrix},$$

$$\text{где } b_{ij} = \frac{a_{ij}(i, j = \overline{1, 3})}{a_{3j}(j = \overline{1, 3})}.$$

Исследуем каждый вид матриц путем их разложения на композицию более простых преобразований.

Матрица вида P_1 получена при условии $a_{32} \neq 0$. Она допускает разложение:

$$P_1 = ANK, \quad (13)$$

$$\text{где } A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & 0 & 1 \end{pmatrix}, K = \begin{pmatrix} 1 & 0 & 0 \\ a_{31} & 1 & a_{33} \\ 0 & 0 & 1 \end{pmatrix} \text{ — матри-$$

$$\text{цы аффинных преобразований; } N = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \text{ —}$$

матрица ортогонального непараметрического преобразования вида $B(x, y) = B_0 \left(\frac{x}{y}, \frac{1}{y} \right)$ (группа G_0 табл. 1).

Исследования показали, что это преобразование вносит столь значительные искажения в изображение, что нормализация практически невозможна. Подобное разложение можно привести и для матрицы P_3 , которое также будет содержать ортогональное проективное преобразование. В итоге — те же выводы, что и для матрицы P_1 .

Обратим внимание на отличительные свойства вычислительной проективной геометрии. Проективные преобразования в задачах нормализации чаще всего рассматриваются в некоторой замкнутой области. Это накладывает свою специфику на свойства проективных преобразований, отличных от классических. Так, из проективной геометрии известно, что любые собственные точки могут перейти при преобразовании в несобственные. В задачах же нормализации изучают в основном конечномерные объекты. Следовательно, изучаемый объект не должен выходить из поля зрения.

Рассмотрим координатное представление полной проективной группы. В зависимости от выбора начала координат восьмипараметрическое представление проективного искажения будет иметь один из трех видов (12). Из определения преобразований P_1 и P_3 следует, что они не существуют в точке $(0, 0)$, откуда вытекает следующее утверждение: для того, чтобы проективное преобразование в некоторой замкнутой области зрения имело вид P_2 , достаточно, чтобы начало координат находилось внутри этой области. Далее обратим внимание на то, как искажаются параллельные и пересекающиеся прямые реальных объектов. Так, параллельные рельсы железной дороги при восприятии глазом человека кажутся пересекающимися вдали, то есть теряют свою параллельность. Однако в машинном поле зрения они не будут иметь общей точки.

Другой случай: пусть исходные прямые не параллельны, но и не пересекаются в реальном поле зрения. При проективных преобразованиях они могут пересекаться, и фигуры, вообще говоря, переходят в другие фигуры, что сильно усложняет нормализацию и распознавание. В связи с этим необходимо вводить ограничения на проективные преобразования внутри поля зрения. Так, увеличение размеров поля зрения D и объекта B на порядок приводит к тому, что интервал возможных значений параметра перспективы уменьшается на порядок. Например, для поля зрения D размера $2a \times 2a$ и объекта B размера $2b \times 2b$ интервал допустимых значений параметра перспективы λ при изменении поля зрения D

и объекта B (чтобы ни одна точка объекта не вышла за пределы поля зрения) должен лежать в пределах:

- 1) $a = 10, b = 5, \quad \lambda \in [-10^{-1}, 10^{-1}]$,
- 2) $a = 100, b = 50, \quad \lambda \in [-10^{-2}, 10^{-2}]$,
- 3) $a = 1000, b = 500, \quad \lambda \in [-10^{-3}, 10^{-3}]$.

Обратимся к матрице P_2 . Нами получены два возможных разложения матрицы:

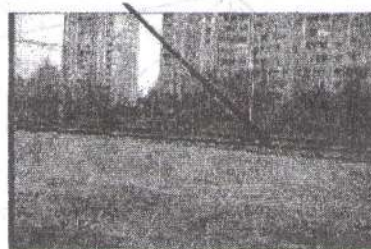
$$\begin{aligned} P_2 &= AP_x(h_1)P_y(h_2), \\ P_2 &= AP_\alpha(h), \end{aligned} \quad (14)$$

где A — аффинная матрица; $P_x(h_1), P_y(h_2), P_\alpha(h)$ — преобразования перспективы с параметрами h_1, h_2, h .

Именно эти разложения позволяют представить нормализатор проективной группы в виде (9) и (10).

Нормализатор вида (9) целесообразно применять, когда есть возможность центрировать входное и эталонное изображения, а нормализатор (10) — когда возможно установить ось перспективного преобразования. На рис. 9 приведен пример реальной ситуации, когда эффективен нормализатор (10).

Направление оси преобразования перспективы



Объект, за которым ведется наблюдение

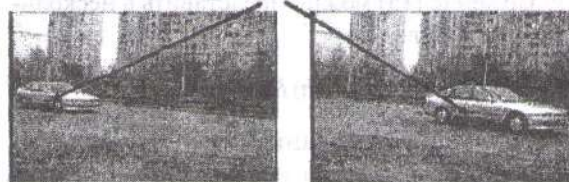


Рис. 9. Пример проективных преобразований

Рассмотрим нормализацию группы центропроекттивных преобразований \bar{G} , т. е. преобразований, оставляющих неподвижной точку, расположенную в начале координат. Очевидно, что центропроекттивное преобразование относится к виду P_2 .

Совокупность центропроекттивных преобразований образует группу относительно операции умножения. Матрица \bar{P}_2 этой группы представима в виде:

$$\bar{P}_2 = D(m, n)P_y(t)P_x(k)H_x(\beta)U(\alpha), \quad (15)$$

где $D_m(m, n)$, $P_y(t)$, $P_x(k)$, $H_x(\beta)$, $U(\alpha)$ — матрицы преобразований масштаба, перспективы вдоль осей x , y косоугольного сдвига вдоль оси абсцисс и поворота (со своими параметрами).

Рассмотрим геометрическую интерпретацию разложения (15).

Пусть векторы ℓ и η (рис. 10) совпадают по направлению с координатными осями и при центропроективном преобразовании переходят, соответственно, в ℓ' и η' . Последовательная нормализация векторов ℓ' и η' представляет собой геометрическую интерпретацию разложения (15), где α — угол между ℓ и ℓ' ; β — угол между η и η' ; \vec{n} — вектор нормали к прямой ℓ' ; m, t — соответственно, параметры сжатия и перспективы вдоль вектора ℓ' ; n, k — параметры сжатия и перспективы вдоль вектора η' ; т. е. $\ell' = \frac{m\ell}{t\ell+1}$ и $\eta' = \frac{n\eta}{k\eta+1}$.

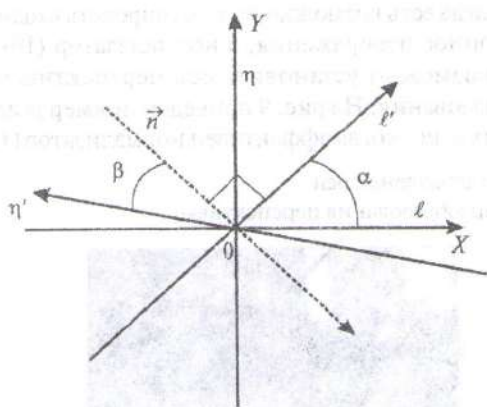


Рис. 10. Геометрическая интерпретация разложения

$$\ell_0 = D(m, n)P_y(t)P_x(k)H_x(\beta)U(\alpha)$$

Разложение (15) можно представить в несколько упрощенном виде:

$$P_0 = D(m, n)P_\alpha(h)H_x(\beta)U(\alpha), \quad (16)$$

где $P_\alpha(h)$ — преобразование перспективы вдоль прямой с углом наклона α и параметром h .

На основе разложений (15), (16) нормализатор (9) проективной группы представляется в следующем виде:

$$F = F_0 F_C',$$

где

$$F_0 = F_D F_{P_\alpha} F_{H_x} F_U$$

или

$$F_0 = F_D F_{P_y} F_{P_x} F_{H_x} F_U.$$

Для нормализатора (9) важным моментом является центрирование изображений. Легко показать, что функционалы обобщенного центра тяжести, пригодные для аффинных преобразований, не под-

ходят для проективных. Конечно, эта задача может быть легко решена, если заранее заданы соответствующие метки (вешки) на эталонном и входном изображениях, что часто делается при аэрофотосъемке. Возникает вопрос, существует ли общий метод нахождения соответствующих точек для произвольных полутоновых изображений. Эта задача, в принципе, решена методом последовательных одномерных корреляций относительно двух пар точек, заданных произвольно на эталоне и входном изображении. При этом затраты машинного времени весьма велики, поскольку метод вначале предполагает поиск соответствующих линий на эталоне и входном изображении. Метод универсален и не предполагает наличия хотя бы одной пары заранее заданных соответствующих точек.

На рис. 11 приведены примеры нормализации проективных изображений двух типов транспортных средств.

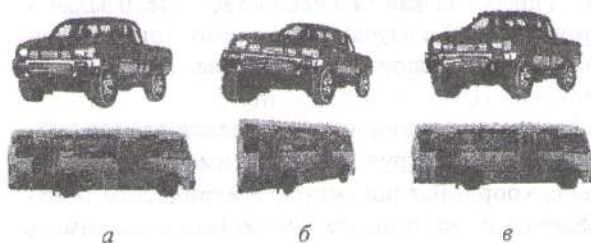


Рис. 11. Примеры нормализации проективных изображений: а — эталонные изображения; б — входные изображения, проективно искаженные; в — нормализованные изображения

5. Нелинейные преобразования

Преобразования классов III и IV можно разделить дополнительно на параметрические и непараметрические. Для параметрических преобразований эталон естественно определяется заданием значений параметров (единица или нуль в зависимости от того, мультипликативную или аддитивную функцию они выполняют).

Примеры таких преобразований приведены в табл. 1 под номерами 25, 26, 27. Это очень широкий класс преобразований, включающий некоторые группы Ли, за исключением аффинных и проективных, а также нелинейные преобразования типа квадратичных форм, которые вообще не образуют групп. Нормализация целого ряда групп Ли представляется сложной и пока нерешенной задачей, хотя попытки использовать их для распознавания изображений предпринимались [15].

В работе [9] рассматривалась возможность нормализации квадратичных преобразований, связывающих эталонное и входное изображения, начиная с одномерных вида $B(x) = B(a_1 x^2 + a_2)$ и заканчивая достаточно сложными двумерными:

$$B(x) = B(a_1x^2 + a_2y^2 + a_3, a_4x^2 + a_5y^2 + a_6).$$

В основу исследований положен подход, основанный на подборе моментов для эталонных и входных изображений типа:

$$\iint_D B(x, y) K_i(x, y) dx dy, \quad i = 1, 2, \dots, n.$$

Например, если связь между эталонным и преобразованным изображением задается в виде $B(x) = B(a_1x^2 + a_2y^2, a_3x^2 + a_4y^2)$, то нормализатор выглядит следующим образом:

$$\begin{aligned} F &= F_2 F_1 [B(x, y)] = F_2 F_1 \left[B_0 \left(A \begin{pmatrix} x^2 \\ y^2 \end{pmatrix} \right) \right] = \\ &= F_2 F_1 \left[B_0 \left(A^{-1} A \begin{pmatrix} x^2 \\ y^2 \end{pmatrix} \right) \right] = \\ &= F_2 [B(x^2, y^2)] \Rightarrow B_0(x, y), \quad A = \begin{vmatrix} a_1 & a_2 \\ a_3 & a_4 \end{vmatrix} \neq 0. \end{aligned}$$

Для нахождения неизвестных $a_i, i = \overline{1, 4}$ возможно применение пяти моментов с функциями $K_i(x, y) = (xu, xy^3, x^3y, xy^5, x^5y)$. При этом моменты для эталонных изображений имеют следующие функции: $K_i^0(x, y) = (1, y, x, xy, y^2, x^2)$. Нелинейная система пяти уравнений, связывающая моменты входных изображений и эталона, решается однозначно при условии $x, y > 0$.

Можно строго и конструктивно определить задачу нормализации в форме (3) для нелинейных экспоненциальных шкал преобразований, поскольку они образуют группу G_3 .

Несмотря на определенные перечисленные успехи в нахождении параметров нормализации, многие вопросы остаются нерешенными. Это прежде всего стабильность эталона, неоднозначность результатов нормализации. Особенно это наглядно проявляется в задачах автоматического масштабирования шкал на осях координат.

Например, нелинейные преобразования 25 из табл. 1 образуют, как нетрудно показать, группу. Тем не менее вопросы автоматической нормализации подобных преобразований (включая, например, и логарифмические шкалы с параметрами) еще не решены.

Возможности нормализации нелинейных преобразований при отсутствии параметров выглядят вообще неясными, особенно с точки зрения дистанционного зондирования объектов через их изображения. Речь идет о преобразованиях 29, 30 из табл. 1. Правда, зеркальные отражения (как дискретная группа с четырьмя преобразованиями G_4) состав-

ляют исключение и могут однозначно автоматически приводиться к эталону нормализатором (3) с функционалами, описанными в [2].

Обратим внимание на тот факт, что нелинейные преобразования 26, 27 после нормализации их параметров приводят входные изображения не к эталону $B_0(x, y)$, а к нелинейной зависимости типа $B(x, y) = B_0(x^2, y^2)$. Разумеется, если заранее известен вид функциональных преобразований координат, то нормализация возможна применением необходимых обратных преобразователей. Это осуществимо не только для преобразований вида $x^2 \rightarrow x, y^2 \rightarrow y$, но, в принципе, и для преобразований G_6 . Вопрос о том, как отыскать функционалы $\Phi(B)$ из (3), которые бы позволили идентифицировать каждое из 16 типов преобразований 29 табл. 1, например, по образцу зеркальных отражений, пока не ясен.

Разумеется, остается открытым вопрос о распознавании типа преобразований, которым подвергается входное изображение. Если соотношение (11) позволяет отличить аффинное преобразование от неаффинного, то отличить, скажем, проективное преобразование от преобразований типа квадратичных форм пока невозможно.

Очевидно, что задачи улучшения и реставрирования изображений уместно рассматривать с позиций общей теории нормализации в совокупности с преобразованиями носителя. Успешные попытки в этом направлении уже сделаны для преобразований 31 [2]. Возможности изучения комбинированного случая (преобразования 33) явно не исчерпаны.

Заключение

Более чем сорокалетний период развития теории и практики нормализации изображений, включая, например, прогресс в области создания систем слежения за подвижными объектами, подтверждает высокую интеллектуальную сложность проблемы, сравнимую с задачами распознавания образов. Несмотря на успехи первых исследований в этой области [10–15], в дальнейшем серьезные результаты стали появляться реже. Связано это как с математическими трудностями решения задач «обратной» геометрии, так и с проблемой восприятия сцен реального мира как человеком, так и компьютерным зрением. На первый взгляд, может представиться нецелесообразной постановка задачи о нормализации ряда преобразований, например, нелинейных или ортогональных (номера 25–30 табл. 1), поскольку человеческое зрение уже не может правильно воспринимать изображения с такими сильными искажениями. Однако нам представляется, что интеллектуальное компьютерное зрение не должно ограничиваться возможностями зрения биологических объектов, включая человека.

Если компьютер считает быстрее человека, а «космическое зрение» спутников уже работает в пяти диапазонах электромагнитных волн (вместо трех человеческих), то, естественно, возникает вопрос о возможном превосходстве в будущем бионических систем нормализации и распознавания образов. Успешное решение поставленных в настоящей работе задач и практическая реализация полученных результатов будут способствовать более широкому пониманию проблемы зрительного восприятия внешнего мира.

Список литературы: 1. Шабанов—Кушнаренко Ю.П., Рвачев В.Л., Мурашко А.Г. Математичні моделі зору. — Київ: Наукова Думка, 1967. — 250 с. 2. Путьятин Е.П., Аверин С.И. Обработка изображений в робототехнике. — М.: Машиностроение, 1990. — 320 с. 3. Путьятин Е.П., Шульгин И.В., Юрченко В.П. Построение инвариантов смещения и поворота зрительных картин // Биологическая и медицинская кибернетика и бионика. — Киев: ИК АН УССР, 1970. — Вып. 3. — С. 57–67. 4. Путьятин Е.П. Теоретические предпосылки нормализации изображений // Проблемы бионики. — Харьков: Вища школа, 1973. — Вып. 10. — С. 82–89. 5. Липанов В.А., Путьятин Е.П. Построение алгоритмов нормализации следящей рамки // Ракетно-космична техніка. — Харків: ХВУ, 1999. — С. 189–193.

6. Путьятин Е.П., Яковлева Е.В. Исследование свойств центральных прямых для нормализации аффинных преобразований // АСУ и приборы автоматики. — Харьков: ХНУРЭ, 2002. — Вып. 121. — С.35–42. 7. Маиталир С.В., Мегель Ю.Е., Путьятин Е.П. Анализ изображений биологических объектов по бинарным сечениям // Радиоэлектроника и информатика. — 2004. — №2. — С. 116–121. 8. Любченко В.А., Путьятин Е.П. Математические модели зрительного восприятия и нормализации изображений // Проблемы бионики. — 2002. — № 56. — С. 44–48. 9. Путьятин Е.П., Долженкова Т.Г. Вопросы нормализации нелинейных преобразований. Сообщение 1 и 2. — Сб. Проблемы бионики. — Харьков, 1980. — Вып. 24. — С. 111–120. 10. Ху М.К. Pattern recognition by moment invariants. Proc. IRE, v. 49, 1428, 1961. 11. Васильев В.И., Реуцкий В.Ю. Инвариантные описания плоских зрительных образов. «Автоматика», 1966. — № 4. 12. Amari S. Theory of normalization of pattern signals in feature spaces. Electronic and Communications in Japan, vol. 49, № 7, 1966. 13. Махонин В.А. Об аффинном опознании плоских фигур. // Изв. АН СССР, сер. Техническая кибернетика, 1963, — №1; 14. Шмидт А.А., Якубович В.А. Алгоритмы математической фильтрации и их применение в задачах распознавания сложных изображений // Автоматика и телемеханика. — 1978. — № 2. — С. 161–176; 15. Файн В.С. Опознавание изображений. — М.: Наука, 1970. — 296 с.

Поступила в редколлегию 12.01.07