

## ГЛАВА 5

### ГЕОМЕТРИЧЕСКИЕ ПРЕОБРАЗОВАНИЯ И ПРИВЯЗКА ИЗОБРАЖЕНИЙ

Во многих задачах тематического дешифрирования применяется взаимное сопоставление между собой изображений, сформированных с помощью датчиков различных физических полей. Ярким примером может служить развитие дистанционных методов контроля природных ресурсов и динамики экосистем (так называемого мониторинга), т.е. сопоставления снимков одной и той же территории, полученных в разное время и/или с помощью различных датчиков. Чаще всего используются датчики, регистрирующие оптическое, радиолокационное, радиотепловое, магнитное и другие поля. Совместное использование различных физических полей требует предварительной обработки соответствующих им изображений, например, с целью перевода изображений в одну спектральную область.

На практике изображения одного и того же объекта или участка местности, полученные в разное время или с помощью различных датчиков, могут значительно различаться один от другого. Отсюда вытекает ряд важных задач привязки, а также точной взаимной геометрической и амплитудной коррекции для последующего совместного анализа. В любом случае это требует установления соответствия между элементами исходных изображений, что сводится к выделению так называемых *опорных* (по другому, *реперных* или *сопряженных*) точек на изображениях, по которым можно осуществить координатную привязку снимков с одновременной геометрической коррекцией. (Точки на двух изображениях называются *сопряженными*, если они являются образами одной точки сцены [5.1, гл.13].) Например, аэрокосмический компьютерный мониторинг предполагает наличие дискретного по времени наблюдения с небольшим временным интерва-

лом, и поэтому, когда движущаяся камера фиксирует яркостный образ наблюдаемого объекта (оптическую поверхность) в виде последовательности изображений, этот образ от снимка к снимку деформируется вследствие *перспективных* искажений и изменения положения камеры. Геометрия соответствующих деформаций моделируется *проективными преобразованиями*, которые составляют более обширный класс, нежели известные преобразования евклидовой геометрии (достаточно сказать, что длины и углы в проективной геометрии не сохраняются, а параллельные линии могут пересекаться! [5.12]).

Восстановление пространственного рельефа по стереоснимкам приводит к проблеме идентификации: установлению точного координатного (поточечного) соответствия элементов стереоизображений. Решение этой задачи состоит в выделении пар реперных фрагментов и оценивании параметров «расхождения» соответствующих точек (это именуется в стереофотограмметрии бинокулярной диспаратностью), по которым можно восстановить функцию геометрического преобразования и оценить поверхность трехмерной сцены (рельеф) [5.5].

### 5.1. ГЕОМЕТРИЧЕСКИЕ ПРЕОБРАЗОВАНИЯ НА ПЛОСКОСТИ И В ПРОСТРАНСТВЕ

Геометрия является математическим базисом для решения многих задач машинного зрения и обработки изображений и содержит множество подобластей. Здесь мы рассмотрим лишь некоторые в охарактеризованном выше контексте привязки, преобразования и совмещения разновременных изображений одного и того же объекта.

При изучении геометрических преобразований плоских изображений (т.е. относящихся к двумерному случаю —  $2D$ ) будем предполагать, что мы работаем в евклидовом пространстве, где имеется ортонормированная декартова система координат, в которой координатные оси взаимно ортогональны, а соответствующие им единичные отрезки имеют одинаковую длину. Тогда каждой точке изображения ставится в соответствие упорядоченная пара чисел  $(x, y)$  декартовых координат: их можно интерпретировать как двумерный вектор  $x$ , геометрически представляемый отрезком прямой линии из точки  $(0, 0)$  в точку  $(x, y)$ .

Двумерные преобразования на плоскости будем интерпретировать как движения точек по отношению к фиксированному базису (а не как изменение базиса, оставляющее точки непод-

вижными). В частности, нас особенно будут интересовать *линейные преобразования*, представляемые матрицами, т. е. преобразования, при которых новые координаты точки линейно зависят от старых координат этой точки следующим образом:

$$\mathbf{x}' = \mathbf{T}\mathbf{x} . \quad (5.1)$$

Линейные преобразования могут быть различного типа, начиная от общего преобразования с произвольными элементами матрицы  $\mathbf{T}$  вплоть до специальных, случаев, когда на элементы матрицы накладываются те или иные ограничения. Интуитивно ясно, что каждому линейному преобразованию (или движению) на плоскости соответствует обратное, переводящее точки в первоначальное положение, и любым двум последовательно выполняемым преобразованиям точек плоскости соответствует некоторое третье преобразование, осуществляющее аналогичную (по результату) операцию. В таком случае принято говорить, что множество всех невырожденных линейных преобразований  $\mathbf{T}$  является замкнутым или, иначе, формирует группу, называемую здесь *общей линейной группой*. Интересно отметить, что само множество общих линейных преобразований может быть разбито на замкнутые подмножества или подгруппы. Прежде всего рассмотрим матрицы преобразования, связанные с наиболее важными подгруппами общей линейной (или *проективной*) группы, а именно, *евклидову* подгруппу, а также подгруппы *подобия* и *аффинную*. Это есть следствие того, что евклидова геометрия (так же как и аффинная) в действительности является подмножеством упомянутой нами выше проективной геометрии.

#### 5.1.1. ТОЧКИ И ПРЯМЫЕ ЛИНИИ НА ПЛОСКОСТИ – ДВОЙСТВЕННОСТЬ ОПИСАНИЙ

Прямая линия на плоскости, как известно из аналитической геометрии, состоит из всех точек, удовлетворяющих уравнению

$$ax + by + 1 = 0 .$$

Пусть две точки имеют координаты  $(x_1, y_1)$  и  $(x_2, y_2)$  соответственно. Каково уравнение линии, соединяющей их? Ясно, что, поскольку линия проходит через эти точки, она должна удовлетворять двум уравнениям:

$$ax_1 + by_1 + 1 = 0 ,$$

$$ax_2 + by_2 + 1 = 0 .$$

Данную систему из двух уравнений можно легко разрешить относительно неизвестных значений  $a$  и  $b$  и получить соответствующие выражения:

$$a = \frac{y_1 - y_2}{x_1 y_2 - x_2 y_1}, \quad b = \frac{x_2 - x_1}{x_1 y_2 - x_2 y_1}.$$

С другой стороны, предположим, что имеются две линии и нужно найти их точку пересечения  $(x, y)$ . Но две прямые должны соответствовать уравнениям

$$a_1 x + b_1 y + 1 = 0,$$

$$a_2 x + b_2 y + 1 = 0.$$

Отсюда для координат точки пересечения  $(x, y)$  получаем соотношения, аналогичные приведенным выше соотношениям для параметров линии  $(a, b)$ :

$$x = \frac{b_1 - b_2}{a_1 b_2 - a_2 b_1}, \quad y = \frac{a_2 - a_1}{a_1 b_2 - a_2 b_1}.$$

Здесь просматривается очень важная симметрия, или *двойственность* между проблемами пересечения двух прямых и (с другой стороны) линии, проходящей через две заданные точки. Координаты (параметры) пары линий и координаты пары точек в обоих случаях входят в формулы одинаковым образом. Далее мы увидим, что отмеченная двойственность распространяется и на другие соотношения между геометрическими объектами.

Имеется ряд проблем, связанных со специальными соотношениями выделенных пар точек и прямых. Предположим, что координаты двух точек отличаются лишь скалярным множителем:  $x_2 = \lambda x_1$ ,  $y_2 = \lambda y_1$ . Это означает, что  $x_1 y_2 - x_2 y_1 = 0$  и параметры прямой, соединяющей выделенные точки, определить невозможно. Прямая линия в данном случае проходит через начало координат  $(0, 0)$ , что, собственно, и создает проблему. Здесь нельзя непосредственно использовать уравнение прямой линии (проходящей через начало координат). Аналогичная проблема возникнет, когда мы попытаемся (формально, из приведенных выше уравнений) найти точку пересечения двух параллельных прямых, когда  $a_2 = \lambda a_1$ ,  $b_2 = \lambda b_1$ :

## 5.1.1.2. ОДНОРОДНЫЕ КООРДИНАТЫ

Для преодоления отмеченных проблем описания геометрических объектов, а также для решения задач преобразования 3D-пространства и 2D-плоскости в единообразном (матричном) виде вводится формализм так называемых *однородных координат*. Однородными координатами служат тройки чисел  $(\bar{x}, \bar{y}, w)$  (одновременно не равные нулю), связанные с обычными координатами точек плоскости соотношением:

$$\begin{pmatrix} \bar{x} \\ \bar{y} \\ w \end{pmatrix} = w \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}, \text{ так что } x = \frac{\bar{x}}{w}, \quad y = \frac{\bar{y}}{w}.$$

Совершенно очевидным свойством однородных координат является эквивалентность пары *однородных векторов*, если один в другой переводятся посредством скалярного множителя

$$(x, y, 1)' \equiv (\lambda x, \lambda y, \lambda)', \lambda \neq 0.$$

Поскольку скалярный множитель  $\lambda$  произвольный, однородные координаты в действительности представляют линию, проходящую через начало координат в евклидовом пространстве. Прямые линии на плоскости также можно представить трехмерными векторами в однородных координатах:

$$\begin{pmatrix} \bar{a} \\ \bar{b} \\ c \end{pmatrix} = c \begin{pmatrix} a \\ b \\ 1 \end{pmatrix}, \text{ где } c - \text{ произвольный скалярный множитель.}$$

Видно, что, как и для двух точек, однородные координаты двух линий эквивалентны, если отличаются лишь общим скалярным множителем. Однородные точки  $x = (\bar{x}, \bar{y}, w)'$ , лежащие на однородной линии  $p = (\bar{a}, \bar{b}, c)'$ , определяются уравнением

$$\bar{a}\bar{x} + \bar{b}\bar{y} + cw = 0, \text{ или } x^t p = 0.$$

Таким образом, точки и линии имеют здесь одинаковые представления. Нетрудно заметить, что прямым, проходящим через начало координат в данном представлении, соответствует значение  $c = 0$ . Двойственным образом точка пересечения двух параллельных прямых, лежащая в бесконечности, имеет множитель  $w = 0$ .

## 5.1.3. ЕВКЛИДОВЫ ПРЕОБРАЗОВАНИЯ

Сцену иногда можно рассматривать как твердое тело, когда взаимные деформации элементов сцены в трехмерном пространстве не допускаются. Аналогично и плоскость иногда можно считать жесткой (недеформируемой). Жестким движениям плоскости соответствует евклидова подгруппа, содержащая лишь преобразования сдвига и поворота (рис. 5.1), математически записываемыми в векторно-матричной форме как

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \mathbf{R} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \end{pmatrix}, \quad (5.2)$$

с матрицей поворота на угол  $\varphi$  вида  $\mathbf{R} = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}$  и вектором трансляции (сдвига)  $\mathbf{t} = \begin{pmatrix} t_x \\ t_y \end{pmatrix}$ .

При помощи троек однородных координат и матриц третьего порядка можно описать любое линейное преобразование плоскости. Действительно, введением дополнительного единичного компонента уравнение (5.2) можно переписать следующим образом:

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}^t & 1 \end{bmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}, \quad \mathbf{0} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \quad (5.3)$$

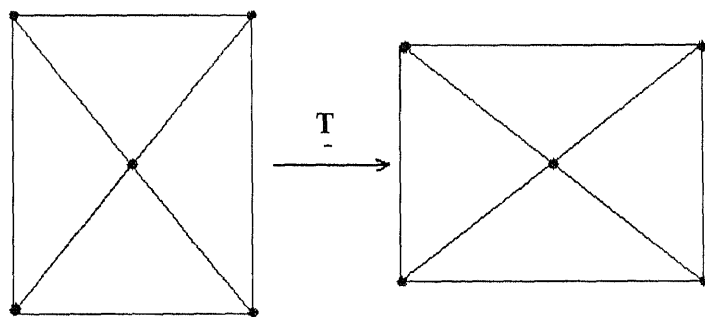


Рис. 5.1. Действие евклидова преобразования на пять точек плоскости (сдвиг, поворот)

Отметим далее, что два последовательно проведенных жестких движения плоскости могут быть представлены единственным движением:

$$\begin{pmatrix} x' \\ 1 \end{pmatrix} = \begin{bmatrix} \mathbf{R}_1 & \mathbf{t}_1 \\ \mathbf{0}' & 1 \end{bmatrix} \begin{bmatrix} \mathbf{R}_2 & \mathbf{t}_2 \\ \mathbf{0}' & 1 \end{bmatrix} \begin{pmatrix} \mathbf{x} \\ 1 \end{pmatrix} = \begin{bmatrix} \mathbf{R}_1 \mathbf{R}_2 & \mathbf{R}_1 \mathbf{t}_2 + \mathbf{t}_1 \\ \mathbf{0}' & 1 \end{bmatrix} \begin{pmatrix} \mathbf{x} \\ 1 \end{pmatrix}. \quad (5.4)$$

Комбинация двух последовательных вращений  $\mathbf{R}_1$  и  $\mathbf{R}_2$ , очевидно, сводится к вращению  $\mathbf{R}_1 \mathbf{R}_2$ . Кроме того, выбором вращения  $\mathbf{R}_2 = \mathbf{R}_1^{-1}$  и сдвига  $\mathbf{t}_2 = -\mathbf{R}_1^{-1} \mathbf{t}_1$  такое (второе) жесткое движение переводит точки плоскости в первоначальное положение. Отмеченной парой свойств, собственно говоря, и характеризуется группа, а класс матриц со структурой вида (5.3) известен как *евклидова группа* преобразований. (Она является, естественно, частным случаем линейных преобразований, у которых матрицы произвольные. Эти матрицы невырожденные и формируют общую линейную группу преобразований или *проективную группу*.) Интересно, что матрицы вращения  $\mathbf{R}$  сами по себе формируют так называемую *ортогональную* подгруппу с замечательным свойством  $\mathbf{R}^t \mathbf{R} = \mathbf{I}$ , где  $\mathbf{I}$  — единичная матрица.

#### 5.1.4. АФФИННЫЕ ПРЕОБРАЗОВАНИЯ

Если матрицу вращения в (5.2) заменить общей невырожденной матрицей  $\mathbf{A}$ , то получим преобразование

$$\begin{aligned} x' &= ax + by + c, \\ y' &= dx + ey + f, \end{aligned} \quad (5.5)$$

или в матричном виде

$$\mathbf{x}' = \mathbf{A}\mathbf{x} + \mathbf{c}, \quad \mathbf{x}' = \begin{pmatrix} x' \\ y' \end{pmatrix}, \quad \mathbf{A} = \begin{pmatrix} a & b \\ d & e \end{pmatrix}, \quad \mathbf{c} = \begin{pmatrix} c \\ f \end{pmatrix},$$

и в однородных координатах

$$\begin{pmatrix} x' \\ 1 \end{pmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{c} \\ \mathbf{0}^t & 1 \end{bmatrix} \begin{pmatrix} \mathbf{x} \\ 1 \end{pmatrix}, \quad \mathbf{0} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \quad (5.6)$$

Здесь также предполагается, что определитель матрицы преобразования не равен нулю:  $\det \mathbf{A} = \begin{vmatrix} a & b \\ d & e \end{vmatrix} = ae - bd \neq 0$ .

Уравнения (5.5), (5.6) определяют общую форму записи хорошо известного *аффинного преобразования* (рис. 5.2). Любое аффинное преобразование имеет *обратное*, которое также является аффинным. Произведение прямого и обратного преобразований дает *единичное преобразование*, оставляющее все на месте. Аффинное преобразование является самым общим взаимно однозначным отображением плоскости на плоскость, при котором сохраняются прямые линии. Сохраняются также отношения длин отрезков, лежащих на одной прямой (или на параллельных прямых), и отношения площадей фигур. Параллельные прямые сохраняют параллельность. В аффинных преобразованиях плоскости особую роль играют несколько важных частных случаев, имеющих простой и наглядный геометрический смысл, а также хорошо прослеживаемые геометрические характеристики [5.4, гл.15].

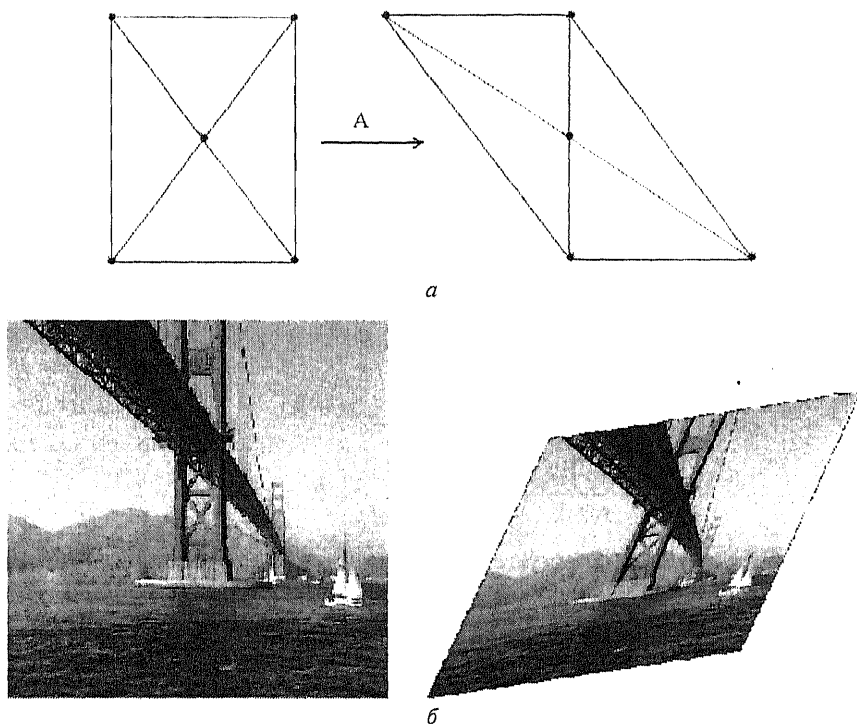


Рис. 5.2. Действие аффинного преобразования на пять точек (сдвиг, поворот, изменение масштабов вдоль осей, косоугольность с сохранением параллельных линий) (а); исходное изображение (слева) и его аффинно-преобразованная копия (справа); параметры аффинного преобразования:  $x' = 1.3x + 0.5y$ ;  $y' = 0.2x + 1.1y$  (б)



1. *Растяжение (сжатие)* вдоль координатных осей, задаваемое в виде

$$x' = \alpha x, \alpha > 0,$$

$$y' = \beta y, \beta > 0.$$

Растяжению вдоль определенной оси соответствует значение масштабного множителя, большего единицы. В однородных координатах матрица растяжения (сжатия) имеет вид

$$\mathbf{T}_{af} = \begin{bmatrix} \alpha & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

2. *Поворот* вокруг начальной точки на угол  $\varphi$ , описываемый формулой:

$$x' = x \cos \varphi - y \sin \varphi,$$

$$y' = x \sin \varphi + y \cos \varphi.$$

Матрица вращения (для однородных координат)

$$\mathbf{T}_{af} = \begin{bmatrix} \cos \varphi & -\sin \varphi & 0 \\ \sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

3. *Перенос*, задаваемый простейшими соотношениями:

$$x' = x + c,$$

$$y' = y + f.$$

Матрица переноса имеет вид

$$\mathbf{T}_{af} = \begin{bmatrix} 1 & 0 & c \\ 0 & 1 & f \\ 0 & 0 & 1 \end{bmatrix}.$$

4. *Отражение* (относительно какой-либо из осей, например оси абсцисс), задаваемое при помощи формулы:

$$x' = x,$$

$$y' = -y.$$

Матрица отражения, соответственно

$$\mathbf{T}_{af} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Из курса аналитической геометрии хорошо известно, что любое аффинное преобразование (5.5) всегда можно представить в виде композиции последовательно выполняемых простейших преобразований означенного вида. Более того, суперпозиция аффинных преобразований также является аффинным преобразованием. Ясно, что аффинные преобразования образуют *аффинную группу*. В частности, подгруппой аффинной группы преобразований является *группа подобия* (содержащая преобразования сдвига, поворота и изменения масштаба):

$$\begin{aligned}x' &= \alpha x \cos \varphi - \alpha y \sin \varphi + c, \\y' &= \alpha x \sin \varphi + \alpha y \cos \varphi + f.\end{aligned}$$

В то же время аффинная группа является подгруппой общей линейной (проективной) группы, а евклидова группа – частным случаем аффинной группы преобразований. Поэтому все отмеченные преобразования формируют иерархию в том смысле, что верно соотношение для их взаимной соподчиненности

$$\text{евклидово преобразование} \subset \text{аффинное} \subset \text{проективное преобразование}.$$

Зная параметры аффинного преобразования, можно вычислить непосредственно и параметры обратного преобразования:

$$\begin{aligned}x &= Ax' + By' + C, \\y &= Dx' + Ey' + F,\end{aligned}$$

решив систему уравнений (5.5) относительно  $(x', y')$ :

$$\begin{aligned}A &= \frac{e}{\det A}, \quad B = -\frac{b}{\det A}, \quad C = \frac{bf - ec}{\det A}, \\D &= -\frac{d}{\det A}, \quad E = \frac{a}{\det A}, \quad F = \frac{cd - af}{\det A}.\end{aligned}$$

Если параметры таковы, что  $(1 - a)(1 - e) \neq bd$ , то данное аффинное преобразование имеет *неподвижную точку*  $(X_0, Y_0)$ :

$$\begin{aligned}X_0 &= \frac{bf + c(1 - e)}{(1 - a)(1 - e) - bd}, \\Y_0 &= \frac{cd + f(1 - a)}{(1 - a)(1 - e) - bd}.\end{aligned}$$

Заметим, что при  $x' = Ax$ ,  $c = 0$  начало координат и будет являться неподвижной точкой.

## 5.1.5. ПРОЕКТИВНЫЕ ПРЕОБРАЗОВАНИЯ

Как выше уже было сказано, общими линейными преобразованиями  $T$  (в представлении однородными координатами)

$$\begin{pmatrix} \bar{x}' \\ \bar{y}' \\ w' \end{pmatrix} = \begin{pmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ t_{31} & t_{32} & t_{33} \end{pmatrix} \begin{pmatrix} \bar{x} \\ \bar{y} \\ w \end{pmatrix} \quad (5.7)$$

формируется группа проективных преобразований (рис. 5.3). При представлении в обычных координатах очевидно, что соотношение (5.7) будет иметь нелинейный вид, связанный с перенормировкой:

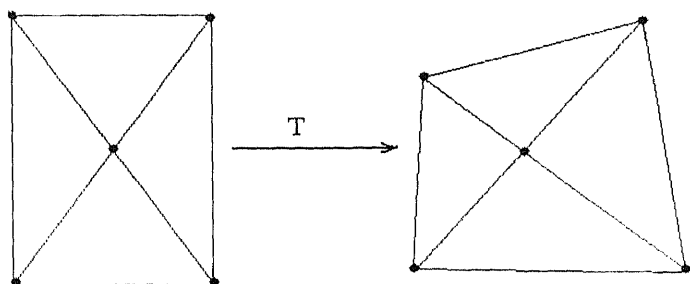
$$x' = \frac{t_{11}x + t_{12}y + t_{13}}{t_{31}x + t_{32}y + t_{33}}, y' = \frac{t_{21}x + t_{22}y + t_{23}}{t_{31}x + t_{32}y + t_{33}}.$$

Проективные преобразования в общем случае не сохраняют параллельности линий. Свойством, сохраняющимся при проективном преобразовании, является так называемая *коллинеарность* точек: три точки, лежащие на одной прямой (т.е. коллинеарные), после преобразования остаются лежать на одной прямой (рис. 5.3). Поэтому обратимое проективное преобразование принято называть еще *коллинеацией* (гомографией).

Проективное преобразование связано с отображением трехмерного изображения на двумерную плоскость. С математической точки зрения, удобно рассматривать мир, включенный в трехмерное проективное пространство  $P^3$ , а плоскость изображения, включенную в проективное двумерное пространство, —  $P^2$ . Точки на трехмерной сцене и на изображении представляются в проективных пространствах как векторы в однородных координатах.

Проективное преобразование из  $P^3$  в  $P^2$  (перспективная проекция), отображающее евклидову точку сцены  $\mathbf{p} = (X, Y, Z)^t$  в точку изображения  $\mathbf{x} = (x, y)^t$  и выраженное в однородных координатах, задается в виде

$$\begin{pmatrix} \bar{x} \\ \bar{y} \\ w \end{pmatrix} = \begin{pmatrix} wx \\ wy \\ w \end{pmatrix} = \mathbf{P} \begin{pmatrix} WX \\ WY \\ WZ \\ W \end{pmatrix} = \begin{pmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{pmatrix} \begin{pmatrix} \bar{X} \\ \bar{Y} \\ \bar{Z} \\ \bar{W} \end{pmatrix}. \quad (5.8)$$



а



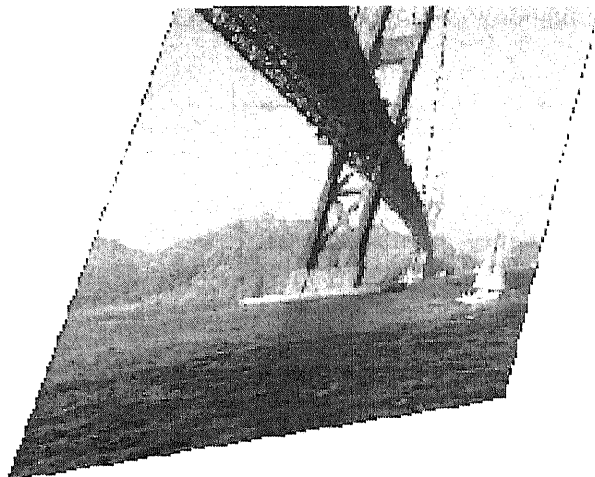
б

Рис. 5.3. Действие проективного преобразования на пять точек плоскости (а); исходное изображение (б); проективно преобразованные образы (в) и (г); параметры проективного преобразования, соответствующие (в) и (г):

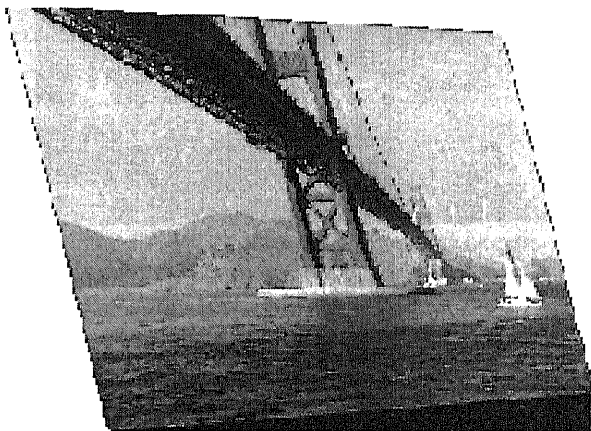
$$x' = \frac{1.1x + 0.35y}{0.00075x + 0.0005y + 1}, y' = \frac{0.2x + 1.1y}{0.00075x + 0.0005y + 1} \text{ и}$$

$$x' = \frac{1.1x - 0.2y}{0.00075x + 0.0005y + 1}, y' = \frac{0.1x + 0.9y}{0.00075x + 0.0005y + 1}$$

(см. также с. 142)



в



г

*Рис. 5.3. Окончание*

Однородные координаты векторов проективного пространства  $(\bar{X}, \bar{Y}, \bar{Z}, \bar{W})'$  и проективной плоскости  $(\bar{x}, \bar{y}, w)'$  соотносятся с неоднородными (евклидовыми) координатами векторов  $\mathbf{p}$  и  $\mathbf{x}$  следующим образом:

$$\mathbf{p} = \left( \frac{\bar{X}}{\bar{W}}, \frac{\bar{Y}}{\bar{W}}, \frac{\bar{Z}}{\bar{W}} \right)' \quad \text{и} \quad \mathbf{x} = \left( \frac{\bar{x}}{w}, \frac{\bar{y}}{w} \right)'.$$

Проективная геометрия составляет математический базис машинного зрения и компьютерной графики. Основные области применения связаны с описанием как процесса формирования изображений, так и их инвариантного представления, а именно: калибровка регистрирующей камеры, анализ движения по серии изображений, распознавание образов, реконструкция сцен по стереоснимкам, синтез изображений, анализ и восстановление формы по полутонам. Полезно отметить тот факт, что композиция двух перспективных проекций не обязательно является перспективной проекцией, но определяет проективное преобразование; т.е. (как мы знаем) проективные преобразования формируют группу, в то время как перспективные проекции — нет.

В связи с этим напомним, что изображение объектов на снимке, сформированном регистрирующей камерой, связано с чрезвычайно важной геометрической операцией — проектированием при помощи пучка прямых, поскольку каждая 2D точка является проекцией множества 3D точек вдоль некоторого направления («луча проектирования») в плоскость снимка (рис. 5.4). Предположим, что плоскость снимка камеры в системе координат  $(X, Y, Z)$  определяется соотношением  $Z = f$ . Если расстояние от плоскости изображения до центра проекции равно  $f$ , то координаты элементов изображения  $(x_i, y_i)$  соотносятся с пространственными координатами  $(X_i, Y_i, Z_i)$  объекта следующим образом:

$$x_i = \frac{fX_i}{Z_i}, \quad y_i = \frac{fY_i}{Z_i}. \quad (5.9)$$

Это нелинейные уравнения. Они могут быть сделаны линейными введением однородных координат. Заметим, что луч, проходящий через 2D точку  $(x_i, y_i)$ , является направляющим вектором

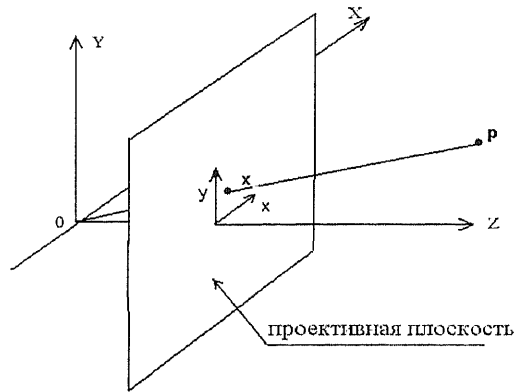


Рис. 5.4. Перспективная проекция

прямой, соединяющей точки  $(0, 0, 0)$  и  $(x_i, y_i, f)$ . В то же время 3D точка  $\lambda(x_i, y_i, f) = (\lambda x_i, \lambda y_i, \lambda f) = (X_i, Y_i, Z_i)$  также лежит на этом луче (и представляет его). При  $Z_i$  произвольном (пробегающем всевозможные значения) получим координаты всех 3D точек  $(\lambda x_i, \lambda y_i, Z_i) = (X_i, Y_i, Z_i)$  на этом луче, соответствующих единственной точке  $(x_i, y_i) = \left( \frac{fX_i}{Z_i}, \frac{fY_i}{Z_i} \right)$  на изображении в плоскости

$Z = f$ . По существу, каждая точка на изображении определяет луч, идущий от сцены в начало координат. Следовательно, однородные координаты точки (на снимке) в действительности представляют (и определяют) линию, проходящую через начало в евклидовом трехмерном пространстве  $E^3$ . Такой набор всевозможных линий, проходящих через начало  $E^3$ , и формирует *проективное* 2D-пространство  $P^2$  в плоскости  $Z = f$ . (Можно, конечно, положить фокусное расстояние  $f = 1$ , поскольку различным значениям  $f$  соответствует разный масштаб изображения.) В однородных координатах уравнения (5.9), естественно, имеют вид (5.8)

$$\begin{pmatrix} wx_i \\ wy_i \\ w \end{pmatrix} = \begin{pmatrix} \bar{x}_i \\ \bar{y}_i \\ \bar{w} \end{pmatrix} = \begin{pmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} X_i \\ Y_i \\ Z_i \\ 1 \end{pmatrix} \quad (5.9')$$

**Замечание.** При  $w = 0$ ,  $3D$  точка  $(\bar{x}, \bar{y}, 0)$  определяет линию, параллельную плоскости  $Z = 1$  (т.е. не имеющую с ней точек пересечения) и, следовательно, не имеет соответствия с какой-либо конечной точкой изображения. Такие линии, или однородные векторы, могут тем не менее иметь смысл, если считать, что соответствующая им точка  $(x, y)$  удаляется на бесконечность в направлении, задаваемом этими координатами:

$$\lim_{w \rightarrow 0} \left( \frac{\bar{x}}{w}, \frac{\bar{y}}{w}, 1 \right) = (x, y, 1) \Rightarrow \lim_{w \rightarrow 0} (\bar{x}, \bar{y}, w) = (\bar{x}, \bar{y}, 0).$$

Мы можем добавить все такие точки к проективной плоскости. Эти точки называются «идеальными» или точками на бесконечности. На изображениях проективной плоскости добавленные точки на бесконечности формируют «линию горизонта» (рис. 5.5). Существует разделение идеальных точек, обусловленное различными направлениями на плоскости; например, точки  $(1, 0, 0)$  и  $(0, 1, 0)$  связаны с горизонтальным и вертикальным направлениями (осями координат) соответственно. Можно также сказать, что все идеальные точки лежат на линии, называемой «идеальной линией» или *линией на бесконечности*, которая рассматривается тем не менее, как и обычная линия. Идеальная линия представляется в виде  $(0, 0, 1)$ .

**Определение.** Проективная плоскость  $P^2$  является аффинной плоскостью с присоединенными идеальной линией и множеством

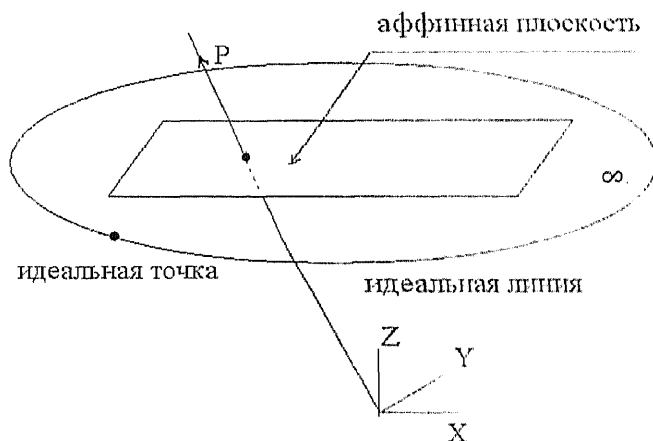


Рис. 5.5. Проективная плоскость = аффинная плоскость +  
+ идеальные точки (идеальная линия)



идеальных точек, которые не отличаются от обычных линий и точек [5.12]. (Аффинная плоскость, естественно, состоит из тех же точек, что и евклидова плоскость. Различие заключается в том, что в первой допускаются неоднородное масштабирование и косоугольность.)

**Замечание.** Найдём пересечение двух прямых  $\mathbf{u}_1 = (a_1, b_1, c_1)$  и  $\mathbf{u}_2 = (a_2, b_2, c_2)$ . Несложные вычисления показывают, что однородные координаты точки пересечения  $\mathbf{p}$  равны  $(b_1c_2 - b_2c_1, a_2c_1 - a_1c_2, a_1b_2 - a_2b_1)$ . Последняя формула легко запоминается, поскольку есть не что иное, как векторное произведение:  $\mathbf{p} = \mathbf{u}_1 \times \mathbf{u}_2$ . Если эти две прямые линии параллельны, т.е.  $\frac{a_1}{b_1} = \frac{a_2}{b_2}$ , то точка пересечения существует (идеальная точка!) и равна  $(b_1c_2 - b_2c_1, a_2c_1 - a_1c_2, 0)$ . Двойственным образом, задавая две точки  $\mathbf{p}_1$  и  $\mathbf{p}_2$ , можно непосредственно найти прямую, проходящую через них:  $\mathbf{u} = \mathbf{p}_1 \times \mathbf{p}_2$ .

Таким образом, идеализация процесса формирования изображения камерой может быть представлена как перспективная проекция из  $P^3$  в  $P^2$ . Допустим, что 3D координаты точек объекта известны. Тогда, зная элементы матрицы  $\mathbf{P}$ , относящиеся к данному проективному преобразованию, точки пространственного объекта можно связать с соответствующими им координатами на снимке в виде (5.8):

$$\begin{pmatrix} \lambda x_i \\ \lambda y_i \\ \lambda \end{pmatrix} = \mathbf{P} \begin{pmatrix} X_i \\ Y_i \\ Z_i \\ 1 \end{pmatrix} = \begin{pmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{pmatrix} \begin{pmatrix} X_i \\ Y_i \\ Z_i \\ 1 \end{pmatrix}. \quad (5.10)$$

Очевидно, что неизвестный масштабный множитель определяется как

$$\lambda = p_{31}X_i + p_{32}Y_i + p_{33}Z_i + p_{34},$$

так что координаты изображения объекта имеют вид отношения

$$\begin{aligned} x_i &= \frac{p_{11}X_i + p_{12}Y_i + p_{13}Z_i + p_{14}}{p_{31}X_i + p_{32}Y_i + p_{33}Z_i + p_{34}}, \\ y_i &= \frac{p_{21}X_i + p_{22}Y_i + p_{23}Z_i + p_{24}}{p_{31}X_i + p_{32}Y_i + p_{33}Z_i + p_{34}}. \end{aligned} \quad (5.11)$$

Пусть координаты характерных элементов изображения объекта  $(x_i, y_i), i = 1, 2, \dots, N$  соотнесены с пространственными координатами точек объекта и требуется вычислить элементы матрицы  $\mathbf{P}$ , т.е. осуществить так называемую *калибровку* камеры (см. главу 6). По крайней мере, 6 точек объекта нужно идентифицировать для этого на снимке ( $N = 6$ ), поскольку элементы матрицы  $\mathbf{P}$  определены с точностью до масштабного множителя (только лишь отношения элементов  $p_{ij}$  значимы) и существует 11 степеней свободы (неизвестных параметров). Обычно же требуется большее число точек, так как измерения сопровождаются помехами и оптимальное решение, минимизирующее их влияние на результат, находится методом наименьших квадратов.

*Проективным базисом* на плоскости  $P^2$  является множество из 4 точек, таких, что любые три из них не лежат на прямой. Соответственно  $3 \times 3$  матрица, сформированная векторами-столбцами однородных координат любых трех точек, должна иметь полный ранг. Легко проверить, например, что точки  $\{(1, 0, 0)', (0, 1, 0)', (0, 0, 1)', (1, 1, 1)'\}$  формируют так называемый *канонический базис*. В соответствии с приведенным выше *замечанием*, канонический базис содержит точки на бесконечности вдоль каждой координатной оси  $\{(1, 0, 0)', (0, 1, 0)'\}$ , начало  $(0, 0, 1)'$  и единичную точку  $(1, 1, 1)'$ . Очевидно, что коллинеация на  $P^2$ :  $\mathbf{x}' = \mathbf{T}\mathbf{x}$  целиком характеризуется геометрически ее действием на точки базиса.

Элементы  $3 \times 3$  матрицы  $\mathbf{T}$  определены с точностью до масштабного множителя, и из них лишь восемь значений независимы. Поэтому, поскольку каждая точка содержит две независимые координаты, четыре пары сопряженных точек (на двух снимках) позволяют определить  $\mathbf{T}$ .

Мы знаем, что аффинная плоскость, в отличие от проективной плоскости, не содержит идеальных точек. Тогда точка  $(x, y, 0)'$  должна быть трансформируема в точку  $(\lambda x, \lambda y, 0)'$  для произвольного масштабного множителя  $\lambda$ :

$$\lambda \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} = \mathbf{T} \begin{pmatrix} x \\ y \\ 0 \end{pmatrix},$$

что влечет  $t_{31} = t_{32} = 0$ . Матрица аффинного преобразования имеет вид

$$\mathbf{T} = \begin{pmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ 0 & 0 & t_{33} \end{pmatrix},$$

и здесь содержится лишь шесть независимых параметров, поскольку масштаб также не важен.

### 5.1.6. ПОЛИНОМИАЛЬНОЕ ПРЕОБРАЗОВАНИЕ

Выше мы рассмотрели геометрические преобразования, в общем-то, идеализированной модели камеры. В действительности формирование изображений сопровождается различного рода нелинейными искажениями (типа оптической дисторсии линзы). Приведение текущих снимков друг к другу или к некоторому эталонному в таком случае можно осуществить лишь нелинейной функцией преобразования. Кроме того, неравномерность движения носителя регистрирующей камеры также приводит к тому, что на практике геометрические искажения снимков не устраняются аффинным преобразованием координат элементов снимка. Поэтому используют полиномиальную аппроксимирующую функцию преобразования (рис. 5.6)

$$\begin{aligned} x' &= a_0 + a_1x + a_2y + a_3x^2 + a_4xy + a_5y^2 + \dots + a_p y^m, \\ y' &= b_0 + b_1x + b_2y + b_3x^2 + b_4xy + b_5y^2 + \dots + b_p y^m, \end{aligned} \quad (5.12)$$

где  $(x, y)$  — координаты точек эталонного снимка,  $(x', y')$  — соответствующие им координаты на текущем (сопоставляемом) снимке.

### 5.1.7. ОЦЕНИВАНИЕ ПАРАМЕТРОВ ПРЕОБРАЗОВАНИЯ

Параметры линейных и нелинейных преобразований (5.2), (5.5), (5.12) устанавливаются по парам взаимно соответствующих реперных точек, идентифицируемых в процессе поиска (см. далее). После этого каждой точке  $(x, y)$  эталонного снимка ставится в соответствие точка  $(x', y')$  текущего снимка. Будем считать, что точки цифрового изображения (пиксели) представлены на дискретной  $(M \times M)$  решетке с постоянным шагом так, что их целочисленные координаты имеют вид  $x_i = i - (M+1)/2$ ,  $y_j = j - (M+1)/2$ ;  $i, j = 1, 2, \dots, M$ . Коэффициенты  $(a_k, b_k)$  выбирают таким образом, чтобы минимизировать среднеквадратическую ошибку аппрок-

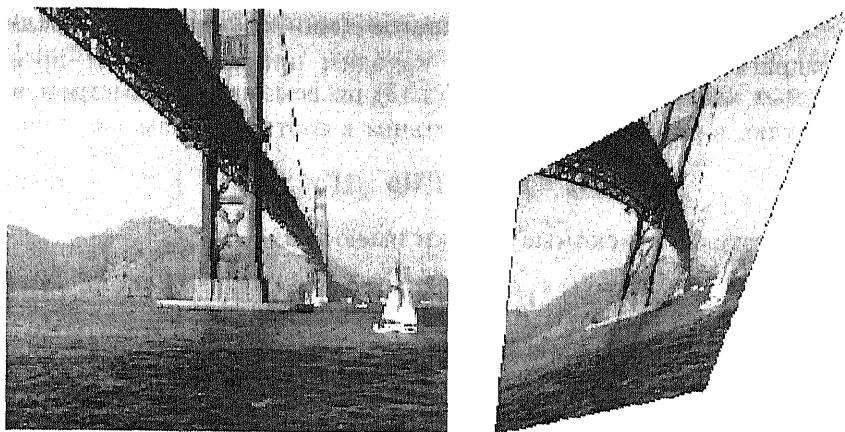


Рис. 5.6. Действие билинейного геометрического преобразования на исходный снимок; параметры преобразования:  
 $x' = 0.1x + 0.9y + 0.002xy$ ,  $y' = 0.2x + 1.1y + 0.0022xy$

симации фактически «наблюдаемых» координат  $(x'_i, y'_j)$  их полиномиальной оценкой  $(\hat{x}'_i, \hat{y}'_j)$  из (5.12) для набора заданных узловых точек  $(x_i, y_j)$ ,  $i, j \in (1, 2, \dots, N)$ . Координаты в плоскости наблюдаемого (текущего, или контролируемого) изображения удобно выразить в виде векторов

$$\mathbf{x}' = (x'_1, x'_2, \dots, x'_N)^t, \quad \mathbf{y}' = (y'_1, y'_2, \dots, y'_N)^t.$$

Аналогично коэффициенты полиномов можно представить в векторной форме:

$$\mathbf{b} = (b_1, b_2, \dots, b_p)^t,$$

$$\mathbf{a} = (a_1, a_2, \dots, a_p)^t.$$

Среднеквадратическую ошибку оценивания можно записать в компактной матричной форме

$$E^2 = (\mathbf{x}' - \mathbf{H}\mathbf{a})^t(\mathbf{x}' - \mathbf{H}\mathbf{a}) + (\mathbf{y}' - \mathbf{H}\mathbf{b})^t(\mathbf{y}' - \mathbf{H}\mathbf{b}), \quad (5.13)$$

где  $\mathbf{H} = \begin{pmatrix} 1 & x_1 & y_1 & \dots & y_1^m \\ 1 & x_2 & y_2 & \dots & y_2^m \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_N & y_N & \dots & y_N^m \end{pmatrix}$  — регрессионная матрица, соответствующая (5.12).

Оценки наименьших квадратов (минимизирующих средне-квадратическую ошибку  $E^2$ ) находим приравниванием производных квадратичной формы (5.13) по векторам параметров  $\mathbf{a}$  и  $\mathbf{b}$  нулю, в результате чего приходим к соотношениям

$$\mathbf{H}^t \mathbf{H} \mathbf{a} = \mathbf{H}^t \mathbf{x}', \quad \mathbf{H}^t \mathbf{H} \mathbf{b} = \mathbf{H}^t \mathbf{y}'. \quad (5.14)$$

Следовательно, искомые оценки имеют вид

$$\hat{\mathbf{a}} = (\mathbf{H}^t \mathbf{H})^{-1} \mathbf{H}^t \mathbf{x}', \quad \hat{\mathbf{b}} = (\mathbf{H}^t \mathbf{H})^{-1} \mathbf{H}^t \mathbf{y}'. \quad (5.15)$$

При больших значениях  $m$  регрессионная матрица  $\mathbf{H}$ , соответствующая (5.12), становится неустойчивой, что приводит к большим ошибкам в определении коэффициентов преобразования. Одним из способов уменьшения этого эффекта является использование полиномов Чебышева. Полиномиальное преобразование (5.12) в этом случае представляется как

$$\widehat{\mathbf{x}}' = \sum_{l=0}^m \sum_{k=0}^{m-l} \alpha_{lk} P_N^{(k)}(x) P_N^{(l)}(y), \quad (5.16)$$

$$\widehat{\mathbf{y}}' = \sum_{l=0}^m \sum_{k=0}^{m-l} \beta_{lk} P_N^{(k)}(x) P_N^{(l)}(y), \quad (5.17)$$

где  $P_N^{(l)}(x)$  — попарно ортогональные на заданном множестве точек многочлены Чебышева, получаемые из последовательности  $1, x, x^2, x^3, \dots$  методом ортогонализации Грама-Шмидта [5.3, гл.20]. Ортогональные многочлены до третьей степени имеют вид

$$P_N^{(0)}(x) = 1, P_N^{(1)}(x) = \lambda_N^{(1)} x, P_N^{(2)}(x) = \lambda_N^{(2)} [x^2 - (N^2 - 1)/12],$$

$$P_N^{(3)}(x) = \lambda_N^{(3)} [x^3 - (3N^2 - 7)x/20],$$

$$\lambda_N^{(1)} = [12/(N^2 - 1)]^{1/2}, \lambda_N^{(2)} = [4 \cdot 9 \cdot 5 / (N^2 - 1)(N^2 - 4)]^{1/2},$$

$$\lambda_N^{(3)} = [16 \cdot 25 \cdot 7 / (N^2 - 1)(N^2 - 4)(N^2 - 9)]^{1/2}.$$

Коэффициенты  $(\alpha_{kl}, \beta_{kl})$  снова выбирают таким образом, чтобы минимизировать среднеквадратическую ошибку аппроксимации фактически «наблюдаемых» координат  $(x'_i, y'_j)$  их полиномиальной оценкой  $(\widehat{x}'_i, \widehat{y}'_j)$  из (5.16), (5.17) для набора заданных узловых точек  $(x_i, y_j)$ ,  $i, j \in (1, 2, \dots, N)$ .

Пользуясь этой методикой, можно без труда вычислить коэффициенты  $(\alpha_{kl}, \beta_{kl})$  по методу наименьших квадратов [аналогично формулам (5.13)...(5.15)], где регрессионная матрица имеет вид

$$H' = \begin{pmatrix} 1 & P_N^{(1)}(x_1) & P_N^{(1)}(y_1) & \dots & P_N^{(m)}(y_1) \\ 1 & P_N^{(1)}(x_2) & P_N^{(1)}(y_2) & \dots & P_N^{(m)}(y_2) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & P_N^{(1)}(x_N) & P_N^{(1)}(y_N) & \dots & P_N^{(m)}(y_N) \end{pmatrix}. \quad (5.18)$$

## 5.2. ВОССТАНОВЛЕНИЕ ИЗОБРАЖЕНИЯ В ПРЕОБРАЗОВАННЫХ КООРДИНАТАХ

После оценивания параметров геометрического преобразования встает задача собственно геометрической коррекции или, по другому, восстановления изображения в преобразованных координатах.

Будем считать, что заданы два снимка ( $U$  и  $V$ ) одной и той же местности, полученные с некоторыми отклонениями точек съемки и условий освещенности. Вследствие этого изображения на снимках отличаются друг от друга геометрическими искажениями. Будем также считать, что на изображениях выделены сопряженные точки, по которым произведено оценивание параметров геометрического преобразования. Зная коэффициенты линейного (или полиномиального) преобразования, можно вычислить в плоскости корректируемого изображения  $V$  координаты всех точек  $(x'_i, y'_j)$ , соответствующие точкам с целочисленными координатами  $(x_i, y_j)$  на эталонном снимке  $U$  (рис. 5.7).

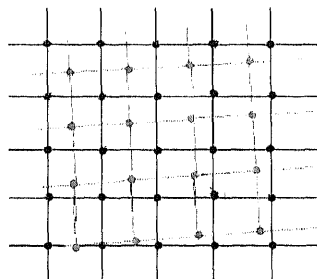


Рис. 5.7. Вычисленные координаты  $(x', y')$  (выделены серым тоном), наложенные на исходную дискретную целочисленную решетку

Восстановив уровни яркости наблюдаемых элементов в вычисленных точках на корректируемом снимке, т. е. осуществив «передискретизацию», полученные значения также можно поместить на дискретном растре размером  $N \times N$ , приведя тем самым искаженное изображение  $V$  в формат эталонного снимка  $U$ .

Поскольку координаты  $(x'_i, y'_j)$  не попадают чаще всего в узлы дискретной решетки (рис. 5.7), возникает задача восстановления соответствующего значения яркости по ближайшим отсчетам. Она решается с помощью методов двумерной интерполяции [5.2, разд. 5.3]. Интерполированное непрерывное изображение в плоскости снимка  $V$  можно описать сверткой

$$\hat{v}(x, y) = \sum_i \sum_j v(i, j) g(x - ih, y - jh), \quad (5.19)$$

где  $g(t, s)$  — интерполирующая функция (называемая также интерполяционным ядром);  $h$  — шаг дискретизации исходного изображения;  $v(ih, jh)$  — известные отсчеты яркости в точках дискретного раstra. Оценка непрерывного изображения позволяет осуществить его передискретизацию на новом множестве точек.

Интерполяционное ядро имеет значительное влияние на численное поведение интерполированных функций. Теоретически оптимальную интерполяцию обеспечивает известная sinc-функция, в одномерном случае имеющая вид

$$\sin c(x - ih) = \frac{\sin 2\pi\Omega(x - ih)}{2\pi(x - ih)}, \quad (5.20)$$

где  $\Omega$  есть ширина полосы частот  $v(x, y)$ . Из теоремы отсчетов следует (см. главу I), что sinc-функция дает наилучшую реконструкцию  $v(x, y)$ , если последняя имела ограниченный спектр и была первоначально оцифрована вблизи частоты Найквиста.

Поскольку интерполяция противоположна дискретизации, интерполирующая функция (5.20) по существу является идеальным низкочастотным фильтром, вырезающим основной участок ограниченного спектра из множества его повторяющихся копий. Однако этот теоретический метод практически невозможно реализовать в контексте обработки изображений. В частности, ограничение области суммирования в (5.19) приводит к тому, что осцилляции, известные как феномен Гиббса, будут проникать в восстанавливаемый образ  $\hat{v}(x, y)$ . Поэтому на практике используют интерполяционные ядра, реализация которых сопряжена с меньшими трудностями. В одномерном случае это прямоугольные, треугольные, B-сплайн функции и т.п. [5.2, разд. 4.3]. При выборе соответствующего ядра исходят из соображений как необходимой точности интерполяции, так и вычислительной эф-

фективности. Понятно, что здесь одномерные функции должны быть преобразованы в двумерные функции. Общий подход состоит во введении так называемых «сепарабельных» интерполяционных функций в виде произведения двух одномерных функций. Сепарабельность во многих отношениях достаточно привлекательна в приложениях, хотя и влечет неизотропность (за исключением гауссовых функций). Однако данные на квадратной решетке дискретизованы также неизотропно.

С вычислительной точки зрения предпочтителен алгоритм, известный как интерполятор по ближайшему соседу, где значение в точке  $(x, y)$  приравнивается значению ближайшего отсчета дискретного раstra. Этот метод соответствует прямоугольному интерполирующему ядру (рис. 5.8). Свертка с прямоугольной функцией в пространственной области эквивалентна умножению сигнала в области частот на sinc-функцию. Последняя является плохим приближением к низкочастотному фильтру, поскольку имеет бесконечное множество боковых лепестков. Алгоритм ближайшего соседа приводит к локальным сдвигам относительно первоначального изображения на величины разностей между вычисленной точкой и ближайшей точкой дискретного раstra (т.е. вплоть до  $\frac{h}{2}$ ). Треугольное ядро (рис. 5.9) в двумерном случае приводит к билинейной интерполяции по четырем ближайшим соседям точки  $(x, y)$ :  $\{[ih, jh], [(i+1)h, jh], [ih, (j+1)h], [(i+1)h, (j+1)h]\}$ .

Справа — график модуля фурье-образа ядра. (Пунктирной линией показан идеальный низкочастотный фильтр с частотой среза  $(-\pi, +\pi)$ .)

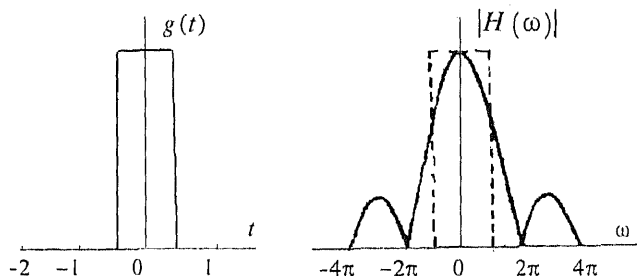


Рис. 5.8. Интерполятор по ближайшему соседу с прямоугольным ядром



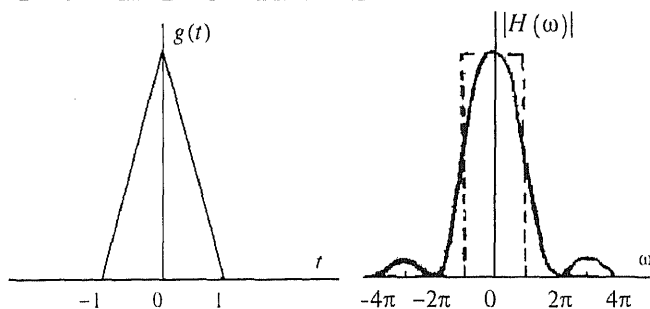


Рис. 5.9. Линейная интерполяционная функция и модуль ее фурье-образа (на правом рисунке пунктиром отмечен идеальный низкочастотный фильтр)

Здесь интерполированный сигнал представляется в виде

$$\hat{v}(x, y) = a + b(x - ih) + c(y - jh) + d(x - ih)(y - jh), \quad (5.21)$$

где

$$\begin{aligned} a &= v(i, j), b = v((i+1), j) - a, c = v(i, (j+1)) - a, \\ d &= v((i+1), (j+1)) + a - b - c. \end{aligned}$$

Приближение к низкочастотному фильтру в данном случае еще далеко от идеального, и к тому же производная интерполированного сигнала терпит разрывы в узлах интерполяции (тем не менее формула (5.21) часто применяется на практике, поскольку удовлетворяет одновременно требованиям приемлемой точности и приемлемым затратам вычислительных ресурсов).

Наиболее подходящим для интерполяции изображений является кубический *B*-сплайн (рис. 5.10), поскольку в результате его применения получается функция, непрерывная и гладкая в узлах интерполяции. Ядро кубической свертки составляется из кусков кубических полиномов, определенных на подынтервалах  $(-2, -1)$ ,  $(-1, 0)$ ,  $(0, 1)$ ,  $(1, 2)$  по каждой из координат. Вне интервала  $(-2, 2)$  интерполяционное ядро равно нулю. Двумерный кубический *B*-сплайн может быть записан как произведение двух одномерных интерполяционных функций по каждой из координат

$$g(s, t) = g(s)g(t),$$

где  $(s, t) = \left( \frac{x}{h} - i, \frac{y}{h} - j \right)$ , так что

$$g(s) = \begin{cases} \frac{|s^3|}{2} - s^2 + \frac{2}{3}, & |s| \in (0, 1); \\ -\frac{|s^3|}{6} + s^2 - 2|s| + \frac{4}{3}, & |s| \in (1, 2), \end{cases} \quad (5.22)$$

и  $g(s) = g(-s)$ , т.е. ядро симметричное.

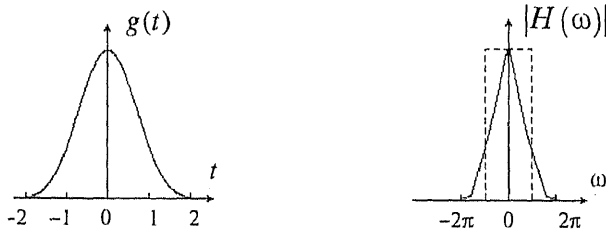


Рис. 5.10. Кубический В-сплайн и модуль его фурье-образа

Как показали непосредственные исследования, кубический В-сплайн имеет тенденцию к сглаживанию передискретизованного изображения по сравнению с его первоначальной копией. Поэтому были предприняты определенные усилия для выбора кубического сплайна, более подходящего задачам обработки изображений. Общий кубический сплайн задается в виде

$$\begin{aligned} g(s) &= a_{30} |s^3| + a_{20} s^2 + a_{10} |s| + a_{00}, & |s| \in (0, 1); \\ g(s) &= a_{31} |s^3| + a_{21} s^2 + a_{11} |s| + a_{01}, & |s| \in (1, 2). \end{aligned} \quad (5.23)$$

Есть несколько естественных ограничений на данное интерполяционное ядро. Так требуется, чтобы значение интерполирующей функции в нуле было равно 1, а в точках с координатами 1 и 2 равно 0. Кроме того, необходимо, чтобы ядро было непрерывным в точках 0 и 1, чтобы наклон в точках 0 и 2 был равен 0, и первая производная была непрерывной. В совокупности это дает семь ограничений, в то время как неизвестных параметров восемь и, следовательно, нужно еще одно условие для однозначного определения интерполяционного ядра. В частности, если интерполяционную функцию привести в соответствие с первыми тремя членами ее разложения в ряд Тейлора, то неизвестный параметр  $a$  должен быть равен  $\left(-\frac{1}{2}\right)$ . Для практических

задач удобнее все семь коэффициентов определить через неизвестный параметр  $a$  и интерполяционное ядро представлять в виде [5.10]

$$g(s) = \begin{cases} (a+2)|s|^3 - (a+3)|s|^2 + 1, & |s| \leq 1; \\ a|s|^3 - 5a|s|^2 + 8a|s| - 4a, & 1 < |s| \leq 2; \\ 0, & |s| > 2. \end{cases} \quad (5.24)$$

Когда константа  $a$  отрицательная, ядро (5.24) положительное в интервале от 0 до 1 и отрицательное в интервале от 1 до 2. Когда  $a$  возрастает, глубина боковых лепестков в интервале от 1 до 2 также увеличивается. Таким образом, с отрицательным значением свободной константы  $a$  интерполяционное ядро имеет вид усеченной sinc-функции. Выяснилось [5.11], что эта функция имеет более предпочтительные высокочастотные свойства, нежели кубический  $B$ -сплайн, и было предложено называть данную функцию *высокоразрешающим интерполяционным кубическим сплайном*. Варьируя значение параметра  $a$  в пределах от  $-1$  (рис. 5.11, *а*) до  $-1/2$  (рис. 5.11, *б*), в каждом конкретном случае можно добиться приемлемой точности при передискретизации.

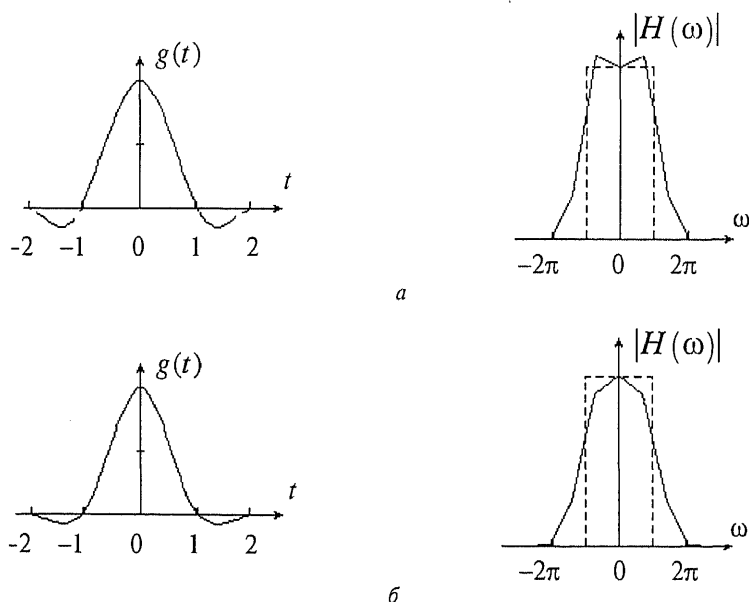


Рис. 5.11. Высокоразрешающий интерполяционный кубический сплайн и его частотная характеристика:  $a = -1$  (а);  $a = -0.5$  (б)

### 5.3. ПРИВЯЗКА ИЗОБРАЖЕНИЙ

В практике обработки изображений задача поиска соответствия получила большое распространение и известна как проблема «поиска по образцу». Формально ее можно рассматривать как процесс отождествления эталонного изображения (образа фрагмента) на первом снимке с одним из множества образов фрагментов, лежащих в некоторой (задаваемой) области (зоне поиска) второго снимка. Алгоритмы установления сходства в своих основополагающих вариантах в той или иной степени связаны с получением характеристик стохастической взаимосвязи сравниваемых фрагментов изображений [5.2, гл.19]. Все они основываются на идеях корреляционной и спектральной теории сигналов, и для соответствующих критериев получены экспериментальные характеристики основных процедур поиска по образцу.

#### 5.3.1. КОРРЕЛЯЦИОННЫЙ КРИТЕРИЙ СХОДСТВА

Будем считать, что изображение эталонного фрагмента (выбранного на снимке  $A$  и представляемого матрицей  $U_0$  размером  $n \times n$ ), сравнивается с изображениями фрагментов снимка  $B$  в «зоне поиска»  $\Omega$  размером  $L \times L$ ,  $L = n + m$ . Перекрытие между фрагментами определяется шагом  $h$  дискретной решетки  $hZ^2$  (в плоскости  $P^2$ ), на которой заданы наблюдаемые переменные  $\{u_0(x), x = (x, y)\}$  на  $A$  или  $\{u(x)\}$  на  $B$ . В процессе скользящего поиска (когда каждый очередной фрагмент получается из предыдущего простым сдвигом на один шаг) вычисляется «функция сходства» между изображением эталонного фрагмента  $\{u_0(x), x \in \Gamma_A\}$  и изображениями текущих (контролируемых) фрагментов  $\{u(x), x \in \Gamma_B\}$ . Здесь требуется найти функцию сходства, которая бы с максимально возможной точностью и достоверностью позволяла локализовать фрагмент, соответствующий изображению эталонного фрагмента, фиксируя таким образом сопряженные точки на снимках.

Взаимно соответствующие элементы изображений одного объекта на снимках должны, очевидно, удовлетворять соотношению

$$u_0(x, y) = (au(x + k, y + l) + b) \operatorname{rect}\left(\frac{x}{n}, \frac{y}{n}\right) + \varepsilon(x, y), \quad (5.25)$$

где  $a$  и  $b$  — параметры контраста и средней освещенности;  $k, l$  — параметры относительного сдвига образца и его аналога на контролируемом снимке;  $\varepsilon(x, y)$  — шум;

$$\text{rect}\left(\frac{x}{n}, \frac{y}{n}\right) = \begin{cases} 1, & x \leq n, y \leq n; \\ 0, & \text{иначе.} \end{cases}$$

В такой формулировке процедура селекции образца должна найти параметры  $k$  и  $l$ , характеризующие сдвиг реперных фрагментов.

Ради простоты будем считать, что параметр  $b$  не меняется по полю снимков, что позволяет перейти к центрированным переменным

$$\bar{u}(x, y) = u(x, y) - \bar{u}, \bar{u} = \frac{1}{L^2} \sum_{(x, y) \in \Omega} u(x, y);$$

$$\bar{u}_0(x, y) = u_0(x, y) - \bar{u}_0, \bar{u}_0 = \frac{1}{n^2} \sum_{x, y=1}^n u_0(x, y).$$

В качестве меры различия в точке  $(k, l)$  будем брать средне-квадратическую ошибку

$$\epsilon_a^2(k, l) = \sum_x \sum_y [\bar{u}_0(x, y) - a\bar{u}(x + k, y + l)]^2, \quad (5.26)$$

которая минимизируется перебором всех допускаемых сдвигов эталона по заданной области контролируемого снимка. Считается, что в точке экстремума реализуется сходство, если  $\epsilon_a^2(k, l) \leq \lambda$ , где  $\lambda$  — некоторый установленный порог. Из требования минимума ошибки  $(\epsilon_a^2(k, l))'_a = 0$  находим оценку  $a$ , подставляем ее в формулу (5.26) и приходим к выражению

$$\epsilon_a^2(k, l) = \sum_x \sum_y [\bar{u}_0(x, y)]^2 - \frac{\left[ \sum_x \sum_y \bar{u}_0(x, y) \bar{u}(x, y) \right]^2}{\sum_x \sum_y [\bar{u}(x, y)]^2}. \quad (5.27)$$

Первый член выражения (5.27) — «энергия» эталонного сигнала — является величиной постоянной, не зависящей от параметров сдвига  $(k, l)$ . Поэтому точка экстремума не изменится, если мы будем нормировать среднеквадратическую ошибку к энергии эталона

$$\bar{\epsilon}_a^2(k, l) = 1 - \frac{\left[ \sum_x \sum_y \bar{u}_0(x, y) \bar{u}(x, y) \right]^2}{\sum_x \sum_y [\bar{u}(x, y)]^2 \sum_x \sum_y [\bar{u}_0(x, y)]^2},$$

и вместо минимума нормированной среднеквадратической ошибки будем искать максимум *коэффициента корреляции* текущего фрагмента с эталоном

$$\hat{r}(k, l) = \frac{\left[ \sum_x \sum_y \bar{u}_0(x, y) \bar{u}(x, y) \right]}{\left\{ \sum_x \sum_y [\bar{u}(x, y)]^2 \sum_x \sum_y [\bar{u}_0(x, y)]^2 \right\}^{\frac{1}{2}}}. \quad (5.28)$$

Соблюдение условий достоверности обнаружения также приводит к необходимости устанавливать порог для величины взаимной корреляции  $\max \hat{r}(k, l)$ : если  $\max \hat{r}(k, l) \geq r_{\text{пор}}$ , то с заданной вероятностью гарантируется действительное сходство найденной пары фрагментов. Величина порога определяется функцией распределения коэффициента корреляции (при случайных выборках) и задаваемой доверительной вероятностью принятия решения о действительном сходстве фрагментов.

Функционирование данного (по существу классического) алгоритма при наличии искажений в изображениях рассмотрено в работе [5.9]. Различия между эталонным и текущим ( $u_0(x, y)$  и  $u(x, y)$ ) изображениями были обусловлены аддитивным шумом и геометрическими искажениями, которые моделировались аффинными преобразованиями координат изображений:

$$u_0(\mathbf{x}) = u(\mathbf{A}\mathbf{x} + \mathbf{t}),$$

где  $\mathbf{x} = (x, y)$ ;  $\mathbf{A} = \alpha \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix}$  — матрица относительного поворота изображений на угол  $\theta$ ;  $\alpha$  — коэффициент изменения масштаба. В работе показано, что среднее значение основного пика корреляционной функции геометрически искаженных изображений, нормированное к средней величине пика при отсутствии искажений, зависит от интенсивности искажений и при малых  $\theta$  и  $|1 - \alpha|$  имеет вид

$$d \approx \sqrt{(1 - \alpha)^2 + \theta^2}.$$

Было практически продемонстрировано, что серьезным недостатком корреляционной меры сходства является ее чувстви-

тельность к геометрическим искажениям видимых размеров сопряженных фрагментов при изменении ракурса съемки.

Обычно за критерии эффективности процедур идентификации сходства принимают точность совмещения фрагментов и вероятность ложной привязки, когда экстремум функционала сходства значимо смещен относительно истинного положения. Анализ результатов имитационных экспериментов позволил сделать следующие выводы.

При наличии геометрических искажений существует оптимальный размер фрагмента эталонного изображения, позволяющий минимизировать вероятность ложной привязки. Оптимальный размер фрагмента пропорционален эффективному радиусу корреляции (полуширине графика автокорреляционной функции) и уменьшается с увеличением геометрических искажений.

При заданном уровне искажений размер эталонного изображения, при котором погрешность совмещения минимальна, меньше, чем размер изображения, необходимый для минимизации вероятности ложной привязки.

Здесь можно порекомендовать использовать полезную модификацию метода идентификации сходства, заключающуюся в том, что искажения геометрии на втором снимке (относительно первого) предварительно компенсируются аффинной (или полиномиальной) «подгонкой». Например, параметры аффинного преобразования

$$A = \begin{bmatrix} a & b \\ d & e \end{bmatrix}, \quad \mathbf{a} = \begin{bmatrix} g \\ h \end{bmatrix}$$

можно оценивать адаптивно (в несколько «проходов»), когда на первом этапе задается достаточно большая зона поиска по образцу, что позволяет на искаженном (по отношению к исходному) снимке находить сопряженные точки. Даже трех пар опорных точек достаточно, чтобы оценить (в первом приближении) параметры аффинного преобразования и осуществить аффинную подгонку геометрии изображения  $u(\mathbf{x})$  к геометрии изображения  $u_0(\mathbf{x})$ :  $u_0(\mathbf{x}) \approx u(A\mathbf{x} + \mathbf{a})$ . Это дает возможность повторным просмотром найти уже существенно большее число пар сопряженных точек на исходном  $u_0(\mathbf{x})$  и аффинно-преобразованном  $u(A^{(2)}\mathbf{x} + \mathbf{a}^{(2)})$  снимках и уточнить по ним параметры аффинной аппроксимации. Дальнейшее повторение этой процедуры позволяет, в принципе, идентифицировать любое (допустимое данной

аппроксимацией) число пар сопряженных точек и, следовательно, добиться заданной точности в оценивании параметров геометрического преобразования.

Поиск по образцу в данном методе сводится к вычислению нормированной взаимной корреляции распределения яркости (двумерного сигнала) на текущем фрагменте первого снимка с распределениями яркостей фрагментов, лежащих в некоторой предполагаемой окрестности образа этого фрагмента, на аффинно-преобразованном втором снимке и определению целочисленных параметров взаимного смещения исходного фрагмента и его образа, устанавливаемого по экстремуму корреляционного функционала.

### 5.3.2. ЛОКАЛЬНОЕ УТОЧНЕНИЕ СДВИГА

После определения целочисленных параметров смещения чаще всего требуется локальное уточнение сдвига фрагментов в пределах дискрета (шага целочисленной решетки) [5.8, гл.15]. Чтобы найти соответствующий вектор сдвига  $(\delta x, \delta y)$ , будем считать, что кросс-корреляционная функция  $\hat{r}_{0p}(x, y)$  в окрестности точки экстремума  $(\hat{x}, \hat{y})$  разлагается в ряд Тейлора:

$$\begin{aligned} \hat{r}_{0p}(x, y) \approx \hat{r}_{0p}(\hat{x}, \hat{y}) + \frac{1}{2} \frac{\partial^2 \hat{r}_{0p}}{\partial x^2} \Big|_{\hat{x}, \hat{y}} (x - \hat{x})^2 + \\ + \frac{1}{2} \frac{\partial^2 \hat{r}_{0p}}{\partial y^2} \Big|_{\hat{x}, \hat{y}} (y - \hat{y})^2 + \frac{\partial^2 \hat{r}_{0p}}{\partial x \partial y} (x - \hat{x})(y - \hat{y}) \end{aligned} \quad (5.29)$$

(первые производные в точке экстремума равны нулю). Дифференцируя данное выражение, непосредственно получаем

$$\frac{\partial \hat{r}_{0p}}{\partial x} \Big|_{\hat{x}, \hat{y}} = -\frac{\partial^2 \hat{r}_{0p}}{\partial x^2} \Big|_{\hat{x}, \hat{y}} \delta x - \frac{\partial^2 \hat{r}_{0p}}{\partial x \partial y} \Big|_{\hat{x}, \hat{y}} \delta y,$$

$$\frac{\partial \hat{r}_{0p}}{\partial y} \Big|_{\hat{x}, \hat{y}} = -\frac{\partial^2 \hat{r}_{0p}}{\partial y^2} \Big|_{\hat{x}, \hat{y}} \delta y - \frac{\partial^2 \hat{r}_{0p}}{\partial x \partial y} \Big|_{\hat{x}, \hat{y}} \delta x.$$

В результате приходим к матричному уравнению для параметров  $(\delta x, \delta y)$ :



$$\begin{bmatrix} (\hat{r}_{0p})''_{xx} & (\hat{r}_{0p})''_{xy} \\ (\hat{r}_{0p})''_{yx} & (\hat{r}_{0p})''_{yy} \end{bmatrix} \begin{pmatrix} \delta x \\ \delta y \end{pmatrix} = - \begin{pmatrix} (\hat{r}_{0p})'_x \\ (\hat{r}_{0p})'_y \end{pmatrix}. \quad (5.30)$$

Таким образом, в результате всех вычислений, проведенных над элементами изображений пары снимков, координаты пар сопряженных точек  $x_1$ ,  $x_2$  в целом связываются соотношением

$$\begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \hat{\mathbf{A}} \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + \begin{pmatrix} \hat{x} \\ \hat{y} \end{pmatrix} + \begin{pmatrix} \delta \hat{x} \\ \delta \hat{y} \end{pmatrix} + \hat{\mathbf{a}}. \quad (5.31)$$

### 5.3.3. КРОСС-СПЕКТРАЛЬНАЯ МЕРА СХОДСТВА

Здесь мы разовьем далее охарактеризованный метод решения задачи привязки. Для этого будем исходить из исследования обобщенной кросс-спектральной меры сходства. Для единообразия запишем матрицу эталонного фрагмента в левом верхнем углу нулевой матрицы порядка  $L \times L$ . Это позволяет перейти от централизованных переменных  $\tilde{u}_0(x, y), \tilde{u}(x, y)$  к спектральным компонентам (Фурье-образам) изображений  $\tilde{\mathbf{U}}, \tilde{\mathbf{U}}_0$ :

$$s_{pq}(\tilde{\mathbf{U}}) = \{ F(\tilde{\mathbf{U}}) \}_{pq} = |s_{pq}(\tilde{\mathbf{U}})| \exp(i\varphi_{pq}(\tilde{\mathbf{U}})), \quad p, q = 0, 1, \dots, L-1;$$

$$|s_{pq}(\tilde{\mathbf{U}})| = \sqrt{\operatorname{Re}^2 s_{pq}(\tilde{\mathbf{U}}) + \operatorname{Im}^2 s_{pq}(\tilde{\mathbf{U}})}, \quad \varphi_{pq}(\tilde{\mathbf{U}}) = \operatorname{arctg} \left( \frac{\operatorname{Im} s_{pq}(\tilde{\mathbf{U}})}{\operatorname{Re} s_{pq}(\tilde{\mathbf{U}})} \right).$$

Поскольку при смещении «образца»  $\{\tilde{u}_0(x, y), x \leq n, y \leq n\}$  в области поиска  $\Omega$  меняется только его фазовый спектр

$$\{ F(\tilde{\mathbf{U}}_0(x+k, y+l)) \}_{pq} = s_{pq}(\tilde{\mathbf{U}}_0) \exp \left( i \frac{2\pi}{L} (pk + ql) \right),$$

будем минимизировать функционал

$$\begin{aligned} F_{kl} = \sum_{p=0}^{L-1} \sum_{q=0}^{L-1} h_{pq} & \left( s_{pq}(\tilde{\mathbf{U}}) - e^{i \frac{2\pi}{L} (pk+ql)} s_{pq}(\tilde{\mathbf{U}}_0) \right)^* \times \\ & \times \left( s_{pq}(\tilde{\mathbf{U}}) - e^{i \frac{2\pi}{L} (pk+ql)} s_{pq}(\tilde{\mathbf{U}}_0) \right) \end{aligned} \quad (5.32)$$

по параметрам сдвига  $k, l$ , где звездочкой обозначена операция комплексного сопряжения, а функция  $h_{pq}$  осуществляет «взвешивание» разностной меры спектральных компонентов. Так как «энергия» изображений фрагментов  $|s_{pq}(\bar{U})|^2$ ,  $|s_{pq}(\bar{U}_0)|^2$  не зависит от параметров  $k, l$ , то положение минимума функционала  $F_{kl}$  соответствует точке экстремума перекрестного члена при раскрытии скобок в (5.32):

$$C_{kl}(\bar{U}, \bar{U}_0) = \sum_{p=0}^{L-1} \sum_{q=0}^{L-1} h_{pq} s_{pq}^*(\bar{U}_0) s_{pq}(\bar{U}) \exp\left(i \frac{2\pi}{L} (pk + ql)\right). \quad (5.33)$$

Точность локализации определяется крутизной данной кросс-спектральной функции вблизи экстремума, характеризующего положение образа эталонного фрагмента. С этой точки зрения наилучшей моделью для  $C_{kl}(\bar{U}, \bar{U}_0)$  служит дискретный вариант дельта функции  $\delta(k - \hat{k}, l - \hat{l})$ . Нетрудно убедиться, что последнему условию удовлетворяет весовая функция  $h_{pq}$  вида

$$\hat{h}_{pq} = |s_{pq}^*(\bar{U}_0) s_{pq}(\bar{U})|^{-1}, \quad (5.34)$$

так как в этом случае обратное дискретное фурье-преобразование

$$\begin{aligned} F^{-1} \left\{ \exp \left[ i(\varphi_{pq}(\bar{U}(x+k, y+l)) - \varphi_{pq}(\bar{U}_0(x, y))) \right] \right\} = \\ = \delta(k - \hat{k}, l - \hat{l}) = \begin{cases} 1, k = \hat{k}, l = \hat{l}; \\ 0, (k, l) \neq (\hat{k}, \hat{l}). \end{cases} \end{aligned}$$

Полученный результат можно интерпретировать следующим образом. Обнаружитель с передаточной функцией (5.34) представляет собой «фазовый» фильтр, т.е. фильтр, который фазовую часть комплексного спектра изображений оставляет без изменений, а амплитудный спектр нормализует. Нетрудно усмотреть, что в данном варианте кросс-спектральная мера  $C_{kl}(\bar{U}, \bar{U}_0)$  есть Фурье-образ функции когерентности

$$\gamma_{pq} = \frac{s_{pq}^*(\bar{U}_0) s_{pq}(\bar{U})}{|s_{pq}^*(\bar{U}_0) s_{pq}(\bar{U})|}.$$

Тем самым можно отнести все признаки оптимальности когерентного функционала [5.6] и к кросс-спектральной мере сходства, допускающей к тому же эффективную реализацию на основе быстрого преобразования Фурье.

Следует отметить, что выражение (5.33) при  $h_{pq} \equiv 1$  соответствует классическому корреляционному алгоритму, реализуемому в спектральной области. Функционирование данного алгоритма при наличии геометрических искажений и выводы, сделанные на основе моделирования, выше были нами приведены.

Анализ результатов имитационных экспериментов показал [5.7], что аналогичные выводы справедливы и для когерентного (фазового) алгоритма. Однако рабочая зона идентификации фазового фильтра (по величине допустимых геометрических искажений) меньше аналогичной зоны для корреляционного алгоритма. Это и понятно, так как чем меньше эффективный радиус кросс-корреляционной функции сходства  $C_{kl}(\vec{U}, \vec{U}_0)$ , тем более чувствителен алгоритм к изменению геометрии идентифицируемых фрагментов. В частности, алгоритм нормированной корреляции ведет удовлетворительный поиск по образцу для снимков, развернутых на угол до  $10^\circ$  ( $n = 18$ ,  $L = 25$ ). Алгоритм фазовой корреляции здесь нормально функционирует, если угол поворота не превышает  $5^\circ$ . На рис. 5.12, *а* представлено типичное поведение кросс-спектральной меры сходства в рассматриваемых вариантах, а на рис. 5.12, *б* даны графики оцененных вероятностей

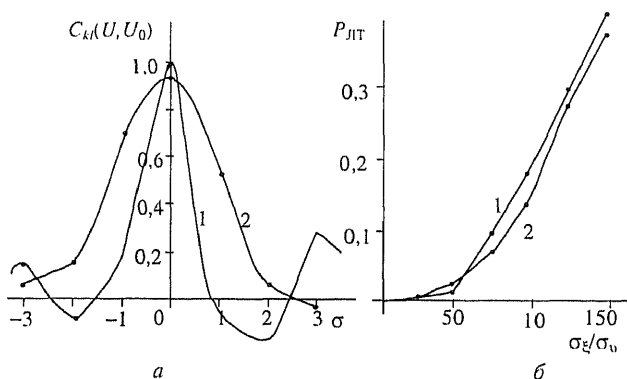


Рис. 5.12. Меры эффективности алгоритмов идентификации сходства: *а* — поведение кросс-спектральной меры сходства вблизи ее экстремума; 1 — отклик фазового фильтра, 2 — отклик корреляционного обнаружителя; *б* — оценки вероятностей ложной привязки (ВЛП) в зависимости от отношения амплитуды шума  $\sigma_z$  к амплитуде фонового сигнала (в %): 1 — ВЛП в фазовом алгоритме, 2 — в корреляционном алгоритме

ложной привязки в зависимости от уровня случайной составляющей сигнала на контролируемом снимке. Видно, что фазовый фильтр в достаточно широком диапазоне мощностей помехи (до 55 %) имеет меньшую вероятность ложной привязки фрагментов, нежели классический корреляционный алгоритм. Для иллюстрации изложенных методов на рис. 5.14 показан результат привязки радиолокационных снимков (рис. 5.13).

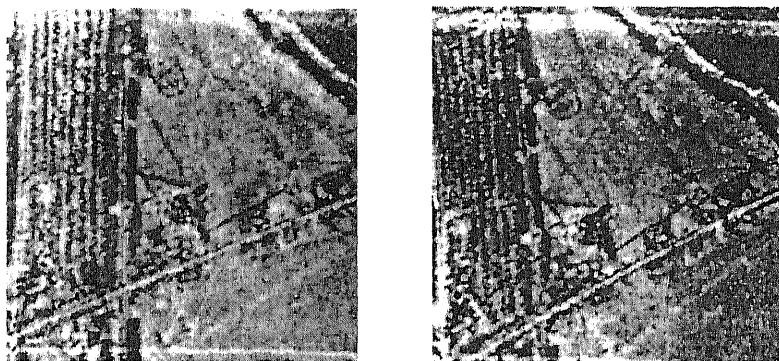


Рис. 5.13. Радиолокационные изображения местности, снятые при двух пролетах самолета (разные ракурсы съемки)

#### 5.3.4. ПРИВЯЗКА ПО ЛОКАЛЬНЫМ НЕОДНОРОДНОСТЯМ

Если в некоторой части изображения уровень яркости более или менее постоянен или на одном из снимков искажен ино-родными включениями, то сопряженные точки искать достаточно трудно. Кросскорреляционные методы, использующие фрагменты меньшие, чем данная область однородной яркости, не дадут ярко выраженного максимума либо максимум не превысит заданный пороговый уровень.

На левом снимке рис. 5.14 большие светлые рамки соответствуют зонам поиска для правого снимка рис. 5.13, малые рамки характеризуют размер эталонных фрагментов, выбираемых на левом снимке рис. 5.13 (по изображениям соответствующих фрагментов вычислялась кросс-спектральная мера сходства). Смещение малых рамок от центра больших рамок указывает на величину найденных локальных геометрических деформаций между снимками. Правый снимок соответствует правому снимку на рис. 5.13, геометрически преобразованному в формат левого снимка (полиномиальное преобразование третьей степени с па-

раметрами, оценки которых мало отличаются от аффинного преобразования; передискретизация осуществлялась на основе высокоразрешающей бикубической сплайн-интерполяции). Это достаточно хорошо видно на рис. 5.14, где рамками обозначены фрагменты, для которых взаимная мера сходства между изображениями находится в доверительном интервале.

В то же время в большей части областей снимков не регистрировалось действительного сходства в выбранных фрагментах зоны поиска. Поэтому может оказаться более разумным осуществлять поиск реперных фрагментов только в информационно насыщенных областях, где яркость быстро меняется, например, на краях между более или менее однородными областями.

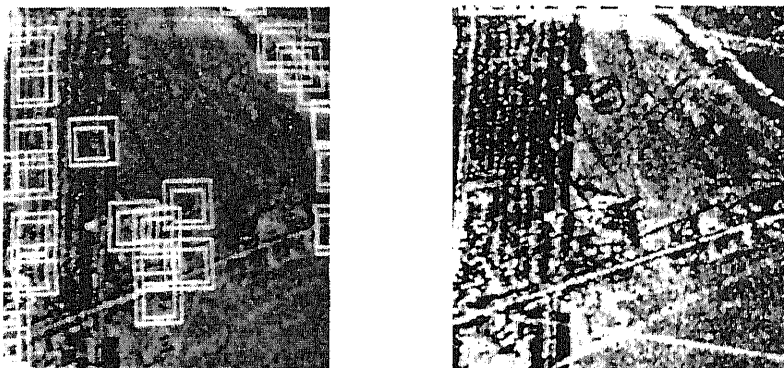


Рис. 5.14. Результаты привязки изображений

Выделение краев [5.2, гл.17] можно рассматривать как средство быстрого предварительного просмотра эталонного изображения и отбора информативных областей. В полученных областях далее выбирают эталонные реперные фрагменты и осуществляется поиск корреляционными методами соответствующих им фрагментов на контролируемых изображениях. В этом случае привязка будет осуществляться по заведомо информативным областям, что повышает точность локализации сопряженных точек и уменьшает вероятность ложного отождествления.

## ВОПРОСЫ К ГЛАВЕ 5

5.1. Охарактеризуйте круг проблем, решение которых приводит к привязке последовательности изображений и их взаимной геометрической коррекции.

5.2. Выстройте иерархию геометрических преобразований. Какие из них сохраняют параллельность прямых?

5.3. Постройте матрицу поворота (в однородных координатах) вокруг точки плоскости  $(x_0, y_0)$  на угол  $\varphi$  (указание: совместить центр поворота с началом координат, повернуть, вернуть центр поворота в прежнее положение и перемножить полученные матрицы элементарных преобразований).

5.4. Чем отличается аффинная плоскость от евклидовой плоскости и от проективной плоскости?

5.5. В каких ситуациях рекомендуется применять полиномиальную аппроксимацию (в том числе и полиномы Чебышева) для описания геометрических деформаций?

5.6. Опишите процедуру восстановления изображения в преобразованных координатах. Какие используются методы интерполяции и в чем их различие?

5.7. Назовите меры сходства изображений и охарактеризуйте их относительные свойства (смысл их оптимальности, точность локализации, достоверность, устойчивость к геометрическим деформациям).

5.8. Опишите процедуру привязки изображений и методы ее ускорения.