

УДК 621.372.542

И.С.Грузман, В.С. Киричук, В.П. Косых,
Г.И. Перетягин, А.А.Спектор

ЦИФРОВАЯ ОБРАБОТКА ИЗОБРАЖЕНИЙ В ИНФОРМАЦИОННЫХ СИСТЕМАХ

УЧЕБНОЕ ПОСОБИЕ
для студентов V курса РЭФ
(специальности “Радиотехника” и
“Средства связи с подвижными объектами”)

Новосибирск
2000

УДК 621.372.542

Грузман И.С., Киричук В.С., Косых В.П., Перетягин Г.И., Спектор А.А.
ЦИФРОВАЯ ОБРАБОТКА ИЗОБРАЖЕНИЙ В ИНФОРМАЦИОННЫХ СИСТЕМАХ: Учебное пособие.- Новосибирск: Изд-во НГТУ, 2000. - 168.

Цифровая обработка изображений - интенсивно развивающаяся научная область, которая находит все более широкое применение в различных информационных технических системах: радиолокационных, связи, телевизионных и т.п. В данном учебном пособии рассмотрены следующие задачи обработки изображений: дискретизации и квантования, улучшения визуального качества путем поэлементного преобразования, подавления помех и восстановления, геометрического преобразования и привязки, фотограмметрии и стереовидения.

Учебное пособие предназначено для студентов старших курсов и магистрантов, изучающих цифровые методы обработки изображений.

Рецензенты д-р техн. наук, профессор Г.В. Мамчев
 канд. техн. наук, доцент В.Н. Васюков

Работа подготовлена
на кафедре теоретических основ радиотехники НГТУ и
кафедре автоматизации физического эксперимента НГУ

© Новосибирский государственный
технический университет 2000 г.

СОДЕРЖАНИЕ

	Стр.
Введение	5
1. ДИСКРЕТИЗАЦИЯ И КВАНТОВАНИЕ НЕПРЕРЫВНЫХ ИЗОБРАЖЕНИЙ	6
1.1. Дискретизация непрерывных изображений	6
1.2. Квантование изображений	14
2. УЛУЧШЕНИЕ ВИЗУАЛЬНОГО КАЧЕСТВА ИЗОБРАЖЕНИЙ ПУТЕМ ПОЭЛЕМЕНТНОГО ПРЕОБРАЗОВАНИЯ	22
2.1. Линейное контрастирование изображения	23
2.2. Соляризация изображения	24
2.3. Препарирование изображения	25
2.4. Преобразование гистограмм, эквалайзация	29
2.5. Применение табличного метода при поэлементных преобразованиях изображений	34
3. ФИЛЬТРАЦИЯ ИЗОБРАЖЕНИЙ	36
3.1. Оптимальная линейная фильтрация. Уравнение Винера-Хопфа	37
3.2. Масочная фильтрация изображений при наличии аддитивного белого шума	41
3.3. Рекуррентная каузальная фильтрация изображений	44
3.4. Применение фильтра Винера для некаузальной двумерной фильтрации	49
3.4.1. Двумерное дискретное преобразование Фурье	49
3.4.2. Циклическая свертка	53
3.4.3. Решение уравнения Винера-Хопфа в циклическом приближении	55
3.5. Байесовская фильтрация изображений	59
3.5.1. Сущность байесовской фильтрации	60
3.5.2. Марковская фильтрация одномерных последовательностей	61
3.5.3. Двухэтапная марковская фильтрация изображений	65
3.6. Медианная фильтрация	69
4. ВОССТАНОВЛЕНИЕ ИЗОБРАЖЕНИЙ	75
4.1. Модели изображений и их линейных искажений	76
4.1.1. Формирование изображений	76
4.1.2. Размытие вследствие движения (смаз)	79
4.1.3. Расфокусировка	81
4.2. Алгебраические методы восстановления изображений	85
4.3. Методы восстановления изображений на основе пространственной фильтрации	87
4.3.1. Инверсный фильтр	89
4.3.2. Фильтр Винера	92
4.3.3. Компенсация краевых эффектов при восстановлении линейно- искаженных изображений	96

4.4. Итерационные методы восстановления изображений	100
5. ГЕОМЕТРИЧЕСКИЕ ПРЕОБРАЗОВАНИЯ И ПРИВЯЗКА ИЗОБРАЖЕНИЙ	108
5.1. Геометрические преобразования на плоскости и в пространстве	109
5.1.1. <i>Точки и прямые линии на плоскости - двойственность описаний</i>	110
5.1.2. <i>Однородные координаты</i>	111
5.1.3. <i>Евклидовы преобразования</i>	112
5.1.4. <i>Аффинные преобразования</i>	113
5.1.5. <i>Проективные преобразования</i>	117
5.1.6. <i>Полиномиальное преобразование</i>	124
5.1.7. <i>Оценивание параметров преобразования</i>	125
5.2. Восстановление изображения в преобразованных координатах	127
5.3. Привязка изображений	133
5.3.1. <i>Корреляционный критерий сходства</i>	133
5.3.2. <i>Локальное уточнение сдвига</i>	137
5.3.3. <i>Кросс-спектральная мера сходства</i>	138
5.3.4. <i>Привязка по локальным неоднородностям</i>	141
6. ФОТОГРАММЕТРИЯ И СТЕРЕОВИДЕНИЕ	143
6.1. Модель регистрирующей камеры	143
6.2. Связь между различными системами координат	145
6.3. Стереоскопическая система	147
6.4. Калибровка камеры	151
6.5. Взаимное ориентирование	156
6.6. Поиск сопряженных точек	159
Заключение	166
Литература	166

ВВЕДЕНИЕ

Многие отрасли техники, имеющие отношение к получению, обработке, хранению и передаче информации, в значительной степени ориентируются в настоящее время на развитие систем, в которых информация имеет характер изображений. Изображение, которое можно рассматривать как двумерный сигнал, является значительно более емким носителем информации, чем обычный одномерный (временной) сигнал. Вместе с тем, решение научных и инженерных задач при работе с визуальными данными требует особых усилий, опирающихся на знание специфических методов, поскольку традиционная идеология одномерных сигналов и систем мало пригодна в этих случаях. В особой мере это проявляется при создании новых типов информационных систем, решающих такие проблемы, которые до сих пор в науке и технике не решались, и которые решаются сейчас благодаря использованию информации визуального характера.

В связи с этим, в вузовских программах появляются дисциплины, направленные на изучение принципов обработки изображений, причем, приоритетное внимание уделяется цифровым методам, привлекательным своей гибкостью. Отсутствие учебной литературы является сильным препятствием данному изучению, что и побудило авторов к написанию пособия. Следует отметить, что ограниченный объем не позволил охватить многие важные аспекты проблемы цифровой обработки изображений. Авторы пособия, читающие курс цифровой обработки изображений в НГТУ и НГУ, исходили из своих представлений о важности тех или иных разделов, а также опирались на многолетний научно-исследовательский и педагогический опыт.

Работа над пособием распределилась следующим образом. Глава 1 написана совместно И.С. Грузманом и А.А. Спектором, главы 2 и 3 - А.А. Спектором, глава 4 - И.С. Грузманом, глава 5 - В.С. Киричуком и Г.И. Перетягиным, глава 6 - В.С. Киричуком и В.П. Косых. При подготовке пособия авторы использовали известные издания в области цифровых методов обработки изображений, а также имеющиеся у них представления об аналогичных дисциплинах, присутствующих в учебных планах ряда высших учебных заведений страны. Среди них - школы в области цифровой обработки изображений Самарского государственного аэрокосмического университета под руководством В.А. Сойфера и В.В. Сергеева и Ульяновского технического университета под руководством К.К. Васильева. Авторы выражают им свою признательность за то сотрудничество, к числу результатов которого относится и данное учебное пособие.

Глава 1

ДИСКРЕТИЗАЦИЯ И КВАНТОВАНИЕ НЕПРЕРЫВНЫХ ИЗОБРАЖЕНИЙ

Очень редко изображения, получаемые в информационных системах, имеют цифровую форму. Поэтому их преобразование к этому виду является обязательной операцией, если предполагается использовать цифровую обработку, передачу, хранение. Как и при одномерных сигналах, данное преобразование включает в себя две процедуры. Первая состоит в замене непрерывного кадра дискретным и обычно называется *дискретизацией*, а вторая выполняет замену непрерывного множества значений яркости множеством квантованных значений и носит название *квантования*. При цифровом представлении каждому из квантованных значений яркости ставится в соответствие двоичное число, чем и достигается возможность ввода изображения в ЭВМ.

Двумерный характер изображения по сравнению с обычными сигналами содержит дополнительные возможности оптимизации цифрового представления с целью сокращения объема получаемых цифровых данных. В связи с этим изучался вопрос о наилучшем размещении уровней квантования, а также об использовании различных растротов [1.1...1.3], другие аспекты данной задачи. Следует, однако, сказать, что в подавляющем большинстве случаев на практике применяют дискретизацию, основанную на использовании прямоугольного раstra, и равномерное квантование яркости. Это связано с простотой выполнения соответствующих операций и относительно небольшими преимуществами от использования оптимальных преобразований. При использовании прямоугольного раstra в окончательном виде цифровое изображение обычно представляет собой матрицу, строки и столбцы которой соответствуют строкам и столбцам изображения.

1.1. Дискретизация непрерывных изображений

Замену непрерывного изображения дискретным можно выполнить различными способами. Можно, например, выбрать какую-либо систему ортогональных функций и, вычислив коэффициенты представления изображения по этой системе (по этому базису), заменить ими изображение. Многообразие базисов дает возможность образования различных дискретных представлений непрерывного изображения. Однако наиболее употребительной является периодическая дискретизация, в частности, как упоминалось выше, дискретизация с прямоугольным растром. Такой способ дискретизации может рассматриваться как один из вариантов применения ортогонального базиса, использующего в качестве своих элементов сдвинутые δ -функции. Далее,

следуя, в основном, [1.1], подробно рассмотрим основные особенности прямоугольной дискретизации.

Пусть $x_H(t_1, t_2)$ - непрерывное изображение, а $x(i_1, i_2)$ -соответствующее ему дискретное, полученное из непрерывного путем прямоугольной дискретизации. Это означает, что связь между ними определяется выражением:

$$x(i_1, i_2) = x_H(i_1 \Delta t_1, i_2 \Delta t_2), \quad (1.1)$$

где $\Delta t_1, \Delta t_2$ - соответственно вертикальный и горизонтальный шаги или интервалы дискретизации. Рис.1.1 иллюстрирует расположение отсчетов на плоскости (t_1, t_2) при прямоугольной дискретизации.

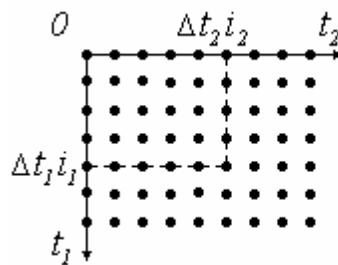


Рис. 1.1. Расположение отсчетов при прямоугольной дискретизации

Основной вопрос, который возникает при замене непрерывного изображения дискретным, состоит в определении условий, при которых такая замена является полноценной, т.е. не сопровождается потерей информации, содержащейся в непрерывном сигнале. Потери отсутствуют, если, располагая дискретным сигналом, можно восстановить непрерывный. С математической точки зрения вопрос, таким образом, заключается в восстановлении непрерывного сигнала в двумерных промежутках между узлами, в которых его значения известны или, иными словами, в осуществлении двумерной интерполяции. Ответить на этот вопрос можно, анализируя спектральные свойства непрерывного и дискретного изображений.

Двумерный непрерывный частотный спектр $X_H(\Omega_1, \Omega_2)$ непрерывного сигнала $x_H(t_1, t_2)$ определяется двумерным прямым преобразованием Фурье:

$$X_H(\Omega_1, \Omega_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_H(t_1, t_2) \exp(-j\Omega_1 t_1 - j\Omega_2 t_2) dt_1 dt_2, \quad (1.2)$$

которому отвечает двумерное обратное непрерывное преобразование Фурье:

$$x_H(t_1, t_2) = \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} X_H(\Omega_1, \Omega_2) \exp(j\Omega_1 t_1 + j\Omega_2 t_2) d\Omega_1 d\Omega_2. \quad (1.3)$$

Последнее соотношение верно при любых значениях t_1, t_2 , в том числе и в узлах прямоугольной решетки $t_1 = i_1 \Delta t_1, t_2 = i_2 \Delta t_2$. Поэтому для значений сигнала в узлах, учитывая (1.1), соотношение (1.3) можно записать в виде:

$$x(i_1, i_2) = \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} X_H(\Omega_1, \Omega_2) \exp(j\Omega_1 i_1 \Delta t_1 + j\Omega_2 i_2 \Delta t_2) d\Omega_1 d\Omega_2. \quad (1.4)$$

Обозначим для краткости через $S(k_1, k_2)$ прямоугольный участок в двумерной частотной области $\frac{-\pi + 2\pi k_1}{\Delta t_1} \leq \Omega_1 \leq \frac{\pi + 2\pi k_1}{\Delta t_1}, \frac{-\pi + 2\pi k_2}{\Delta t_2} \leq \Omega_2 \leq \frac{\pi + 2\pi k_2}{\Delta t_2}$.

Вычисление интеграла в (1.4) по всей частотной области можно заменить интегрированием по отдельным участкам $S(k_1, k_2)$ и суммированием результатов:

$$\begin{aligned} x(i_1, i_2) = & \frac{1}{4\pi^2} \sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} \iint_{S(k_1, k_2)} X_H(\Omega_1, \Omega_2) \times \\ & \times \exp(j\Omega_1 \Delta t_1 i_1 + j\Omega_2 \Delta t_2 i_2) d\Omega_1 d\Omega_2. \end{aligned}$$

Выполняя замену переменных по правилу $\Omega_1 \rightarrow \Omega_1 - 2\pi k_1 / \Delta t_1, \Omega_2 \rightarrow \Omega_2 - 2\pi k_2 / \Delta t_2$, добиваемся независимости области интегрирования от номеров k_1 и k_2 :

$$\begin{aligned} x(i_1, i_2) = & \frac{1}{4\pi^2} \int_{-\frac{\pi}{\Delta t_1}}^{\frac{\pi}{\Delta t_1}} \int_{-\frac{\pi}{\Delta t_2}}^{\frac{\pi}{\Delta t_2}} \sum_{k_1} \sum_{k_2} X_H(\Omega_1 + \frac{2\pi k_1}{\Delta t_1}, \Omega_2 + \frac{2\pi k_2}{\Delta t_2}) \times \\ & \times \exp(j\Omega_1 \Delta t_1 i_1 + j\Omega_2 \Delta t_2 i_2) d\Omega_1 d\Omega_2. \end{aligned}$$

Здесь учтено, что $\exp(-j2\pi k i) = 1$ при любых целых значениях k и i . Данное выражение по своей форме очень близко к обратному преобразованию Фурье. Отличие состоит лишь в неправильном виде экспоненциального множителя. Для придания ему необходимого вида введем нормированные частоты $\omega_1 = \Omega_1 \Delta t_1, \omega_2 = \Omega_2 \Delta t_2$ и выполним в соответствии с этим замену переменных. В результате получим:

$$x(i_1, i_2) = \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \frac{1}{\Delta t_1 \Delta t_2} \sum_{k_1} \sum_{k_2} X_H \left(\frac{\omega_1 + 2\pi k_1}{\Delta t_1}, \frac{\omega_2 + 2\pi k_2}{\Delta t_2} \right) \times \\ \times \exp(j\omega_1 i_1 + j\omega_2 i_2) d\omega_1 d\omega_2. \quad (1.5)$$

Теперь выражение (1.5) имеет форму обратного преобразования Фурье, следовательно стоящая под знаком интеграла функция

$$X(\omega_1, \omega_2) = \frac{1}{\Delta t_1 \Delta t_2} \sum_{k_1} \sum_{k_2} X_H \left(\frac{\omega_1 + 2\pi k_1}{\Delta t_1}, \frac{\omega_2 + 2\pi k_2}{\Delta t_2} \right) \quad (1.6)$$

является двумерным спектром дискретного изображения. В плоскости ненормированных частот выражение (1.6) имеет вид:

$$X(\Omega_1 \Delta t_1, \Omega_2 \Delta t_2) = \frac{1}{\Delta t_1 \Delta t_2} \sum_{k_1} \sum_{k_2} X_H \left(\Omega_1 + \frac{2\pi k_1}{\Delta t_1}, \Omega_2 + \frac{2\pi k_2}{\Delta t_2} \right) \quad (1.7)$$

Из (1.7) следует, что двумерный спектр дискретного изображения является прямоугольно периодическим с периодами $2\pi/\Delta t_1$ и $2\pi/\Delta t_2$ по осям частот Ω_1 и Ω_2 соответственно. Спектр дискретного изображения $X(\Omega_1, \Omega_2)$ образуется в результате суммирования бесконечного количества спектров $X_H(\Omega_1, \Omega_2)$ непрерывного изображения, отличающихся друг от друга частотными сдвигами $2\pi/\Delta t_1$ и $2\pi/\Delta t_2$. Рис.1.2 качественно показывает соотношение между двумерными спектрами непрерывного (рис.1.2.а) и дискретного (рис.1.2.б) изображений.

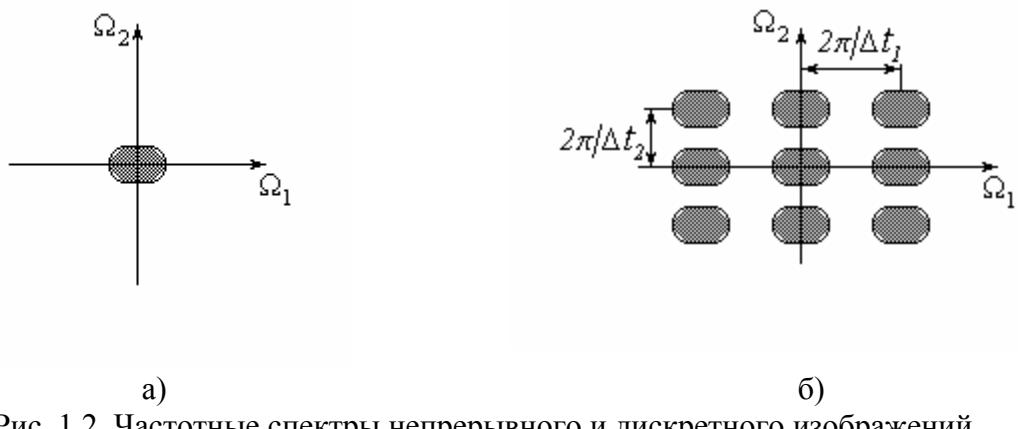


Рис. 1.2. Частотные спектры непрерывного и дискретного изображений

Сам результат суммирования существенно зависит от значений этих частотных сдвигов, или, иными словами, от выбора интервалов дискретизации $\Delta t_1, \Delta t_2$.

Допустим, что спектр непрерывного изображения $X_H(\Omega_1, \Omega_2)$ отличен от нуля в некоторой двумерной области в окрестности нулевой частоты, т. е. описывается двумерной финитной функцией. Если при этом интервалы дискретизации выбраны так, что $X_H(\Omega_1, \Omega_2) = 0$ при $|\Omega_1| \geq \pi/\Delta t_1$, $|\Omega_2| \geq \pi/\Delta t_2$, то наложения отдельных ветвей при формировании суммы (1.7) происходить не будет. Следовательно, в пределах каждого прямоугольного участка $S(k_1, k_2)$ от нуля будет отличаться лишь одно слагаемое. В частности, при $k_1 = 0, k_2 = 0$ имеем:

$$X(\Omega_1, \Omega_2) = \frac{1}{\Delta t_1 \Delta t_2} X_H(\Omega_1, \Omega_2) \text{ при } |\Omega_1| \geq \pi/\Delta t_1, |\Omega_2| \geq \pi/\Delta t_2. \quad (1.8)$$

Таким образом, в пределах частотной области $S(0,0)$ спектры непрерывного и дискретного изображений с точностью до постоянного множителя совпадают. При этом спектр дискретного изображения в этой частотной области содержит полную информацию о спектре непрерывного изображения. Подчеркнем, что данное совпадение имеет место лишь при оговоренных условиях, определяемых удачным выбором интервалов дискретизации. Отметим, что выполнение этих условий, согласно (1.8), достигается при достаточно малых значениях интервалов дискретизации $\Delta t_1, \Delta t_2$, которые должны удовлетворять требованиям:

$$\Delta t_1 \leq \pi/\Omega_{1\text{гр}}, \quad \Delta t_2 \leq \pi/\Omega_{2\text{гр}}, \quad (1.9)$$

в которых $\Omega_{1\text{гр}}, \Omega_{2\text{гр}}$ - граничные частоты двумерного спектра.

Соотношение (1.8) определяет способ получения непрерывного изображения $x_H(t_1, t_2)$ из дискретного $x(i_1, i_2)$. Для этого достаточно выполнить двумерную фильтрацию дискретного изображения низкочастотным фильтром с частотной характеристикой

$$K(j\Omega_1, j\Omega_2) = \begin{cases} \Delta t_1 \Delta t_2 & \text{при } |\Omega_1| \leq \pi/\Delta t_1, |\Omega_2| \leq \pi/\Delta t_2, \\ 0 & \text{при других } \Omega_1, \Omega_2. \end{cases} \quad (1.10)$$

Спектр изображения на его выходе содержит ненулевые компоненты лишь в частотной области $S(0,0)$ и равняется, согласно (1.8), спектру непрерывного изображения $X_H(\Omega_1, \Omega_2)$. Это означает, что изображение на выходе идеального фильтра низких частот совпадает с $x_H(t_1, t_2)$.

Таким образом, идеальное интерполяционное восстановление непрерывного изображения выполняется при помощи двумерного фильтра с прямоугольной частотной характеристикой (1.10). Нетрудно записать в явном

виде алгоритм восстановления непрерывного изображения. Двумерная импульсная характеристика восстанавливающего фильтра, которую легко получить при помощи обратного преобразования Фурье от (1.10), имеет вид:

$$h(t_1, t_2) = \frac{\sin(\pi t_1 / \Delta t_1)}{\pi t_1 / \Delta t_1} \frac{\sin(\pi t_2 / \Delta t_2)}{\pi t_2 / \Delta t_2}.$$

Продукт фильтрации может быть определен при помощи двумерной свертки входного изображения и данной импульсной характеристики. Представив входное изображение $x_{\text{вх}}(t_1, t_2)$ в виде двумерной последовательности δ -функций

$$x_{\text{вх}}(t_1, t_2) = \sum_{i_1} \sum_{i_2} x(i_1, i_2) \cdot \delta(t_1 - i_1 \Delta t_1) \delta(t_2 - i_2 \Delta t_2),$$

после выполнения свертки находим:

$$x(t_1, t_2) = \sum_{i_1} \sum_{i_2} x(i_1, i_2) \cdot \frac{\sin[\pi(t_1 - i_1 \Delta t_1) / \Delta t_1]}{\pi(t_1 - i_1 \Delta t_1) / \Delta t_1} \frac{\sin[\pi(t_2 - i_2 \Delta t_2) / \Delta t_2]}{\pi(t_2 - i_2 \Delta t_2) / \Delta t_2}. \quad (1.11)$$

Полученное соотношение указывает способ точного интерполяционного восстановления непрерывного изображения по известной последовательности его двумерных отсчетов. Согласно этому выражению для точного восстановления в роли интерполирующих функций должны использоваться двумерные функции вида $\sin x / x$. Соотношение (1.11) представляет собой двумерный вариант теоремы Котельникова-Найквиста.

Подчеркнем еще раз, что эти результаты справедливы, если двумерный спектр сигнала является финитным, а интервалы дискретизации достаточно малы. Справедливость сделанных выводов нарушается, если хотя бы одно из этих условий не выполняется. Реальные изображения редко имеют спектры с ярко выраженными граничными частотами. Одной из причин, приводящих к неограниченности спектра, является ограниченность размеров изображения. Из-за этого при суммировании в (1.7) в каждой из зон $S(k_1, k_2)$ проявляется действие слагаемых из соседних спектральных зон. При этом точное восстановление непрерывного изображения становится вообще невозможным. В частности, не приводит к точному восстановлению и использование фильтра с прямоугольной частотной характеристикой.

Особенностью оптимального восстановления изображения в промежутках между отсчетами является использование всех отсчетов дискретного изображения, как это предписывается процедурой (1.11). Это не всегда удобно, часто требуется восстанавливать сигнал в локальной области, опираясь на некоторое небольшое количество имеющихся дискретных значений. В этих случаях целесообразно применять квазипримимальное

восстановление при помощи различных интерполирующих функций. Такого рода задача возникает, например, при решении проблемы привязки двух изображений, когда из-за геометрических расстроек этих изображений имеющиеся отсчеты одного из них могут соответствовать некоторым точкам, находящимся в промежутках между узлами другого. Решение этой задачи более подробно обсуждается в последующих разделах данного пособия.

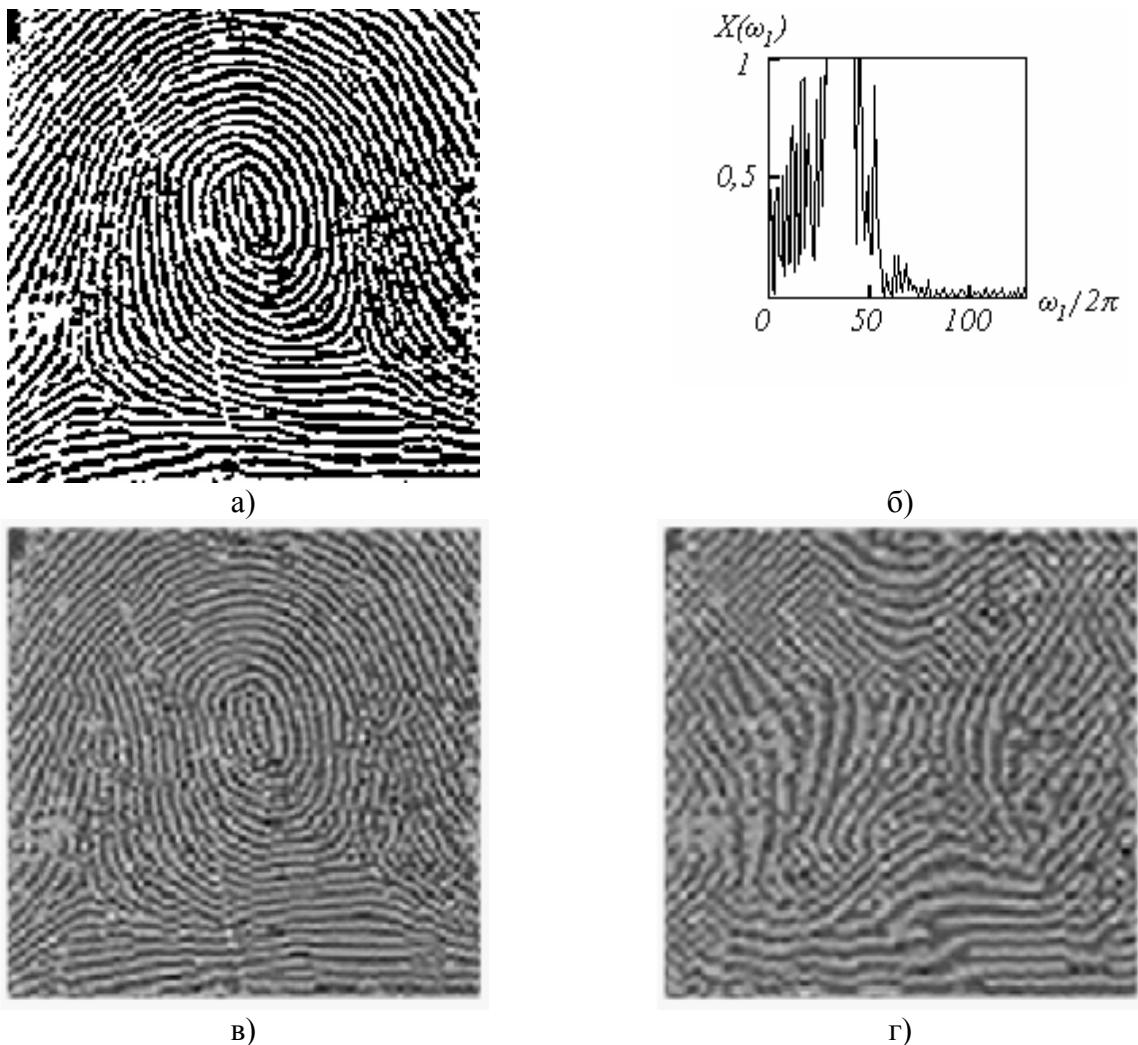


Рис. 1.3. Влияние интервала дискретизации на восстановление изображения «Отпечаток пальца»

Рис. 1.3 иллюстрирует влияние интервалов дискретизации на восстановление изображений. Исходное изображение, представляющее собой отпечаток пальца, приведено на рис. 1.3.а, а одно из сечений его нормированного спектра - на рис. 1.3.б. Данное изображение является дискретным, а в качестве граничной частоты использовано значение $\omega_{1\text{гр}} = 2\pi \cdot 128$. Как следует из рис. 1.3.б, значение спектра на этой частоте пренебрежимо мало, что гарантирует качественное восстановление. По сути дела, наблюдаемая на рис. 1.3.а картина является результатом восстановления непрерывного изображения, а роль восстанавливающего фильтра выполняет

устройство визуализации - монитор или принтер. В этом смысле изображение рис. 1.3.а может рассматриваться как непрерывное.

Рис. 1.3.в,г показывают последствия от неправильного выбора интервалов дискретизации. При их получении осуществлялась “дискретизация непрерывного” изображения рис. 1.3.а путем прореживания его отсчетов. Рис. 1.3.в соответствует увеличению шага дискретизации по каждой координате в три, а рис. 1.3.г - в четыре раза. Это было бы допустимо, если бы значения граничных частот были ниже в такое же число раз. В действительности, как видно из рис. 1.3.б, происходит нарушение требований (1.9), особенно грубое при четырехкратном прореживании отсчетов. Поэтому восстановленные при помощи алгоритма (1.11) изображения оказываются не только расфокусированными, но и сильно искажают текстуру отпечатка.

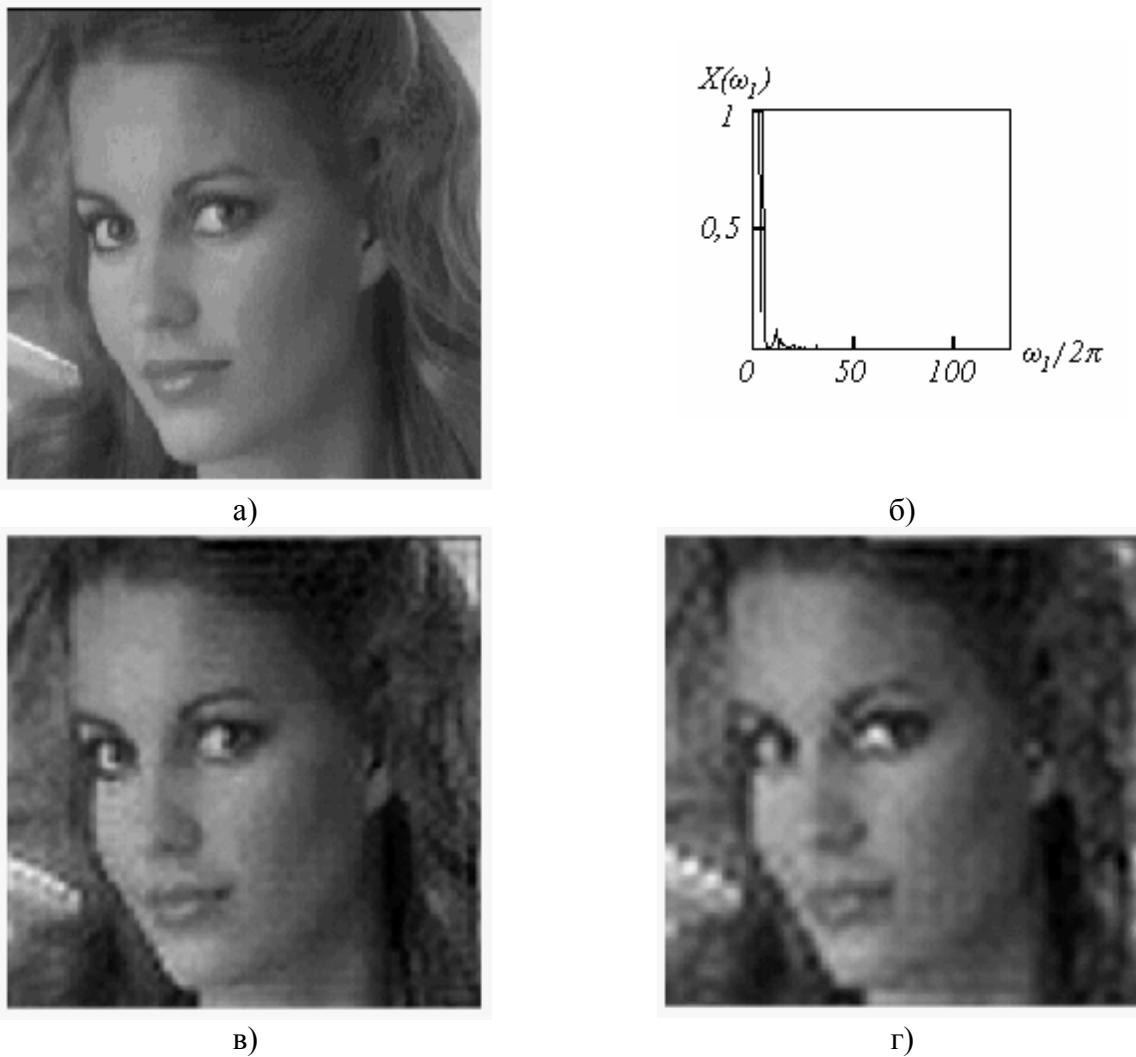


Рис. 1.4. Влияние интервала дискретизации на восстановление изображения «Портрет»

На рис. 1.4 приведена аналогичная серия результатов, полученных для изображения типа “портрет”. Последствия более сильного прореживания (в четыре раза на рис. 1.4.в и в шесть раз на рис. 1.4.г) проявляются в основном в потере четкости. Субъективно потери качества представляются менее значительными, чем на рис. 1.3. Это находит свое объяснение в значительно меньшей ширине спектра, чем у изображения отпечатка пальца. Дискретизация

исходного изображения соответствует граничной частоте $\omega_{\text{лгр}} = 2\pi \cdot 86$. Как видно из рис. 1.4.б, это значение намного превышает истинное значение $\omega_{\text{лгр}}$. Поэтому увеличение интервала дискретизации, иллюстрируемое рис. 1.3.в,г, хотя и ухудшает картину, все же не приводит к таким разрушительным последствиям, как в предыдущем примере.

1.2. Квантование изображений

При цифровой обработке изображений непрерывный динамический диапазон значений яркости делится на ряд дискретных уровней. Эта процедура называется квантованием. Квантователь преобразует непрерывную переменную x в дискретную переменную $x_{\text{кв}}$, принимающую конечное множество значений $\{r_1, \dots, r_L\}$. Эти значения называются уровнями квантования. В общем случае преобразование выражается ступенчатой функцией (рис. 1.5). Если яркость x отсчета изображения принадлежит интервалу $(d_j, d_{j+1}]$ (т.е., когда $d_j < x \leq d_{j+1}$), то исходный отсчет заменяется на уровень квантования r_j , где $d_j, j = \overline{1, L+1}$ - пороги квантования. При этом полагается, что динамический диапазон значений яркости ограничен и равен $[d_1, d_{L+1}]$.

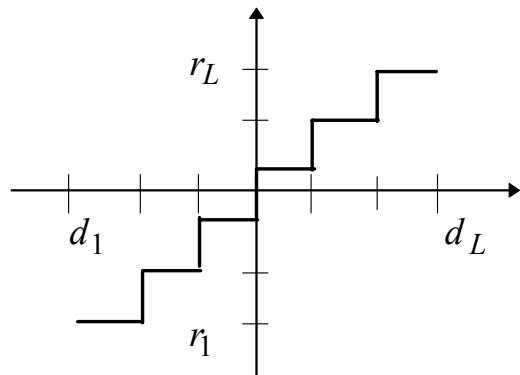


Рис.1.5.Функция, описывающая квантование

Задача построения квантователя состоит в определении значений порогов d_j и уровней r_j . Простейший способ решения этой задачи состоит в разбиении динамического диапазона на одинаковые интервалы. Однако такое решение не является наилучшим. Если значения яркости большинства отсчетов изображения сгруппированы, например, в «темной» области и число уровней L ограничено, то целесообразно квантовать неравномерно. В «темной» области следует квантовать чаще, а в «светлой» реже. Это позволит уменьшить ошибку квантования $\varepsilon = x - x_{\text{кв}}$.

Таким образом, задачу построения квантователя можно сформулировать как задачу нахождения оптимальных значений d_j и r_j , удовлетворяющих некоторому критерию оптимизации. Обычно при фиксированном числе уровней квантователь оптимизируется по критерию минимальной среднеквадратической ошибки

$$E\{\varepsilon^2\} = E\{(x - x_{\kappa\theta})^2\} = \min_{\{d_j, r_j\}}, \quad (1.12)$$

в предположении, что яркость x - случайная величина с известной плотностью вероятности $w(x)$.

Среднеквадратическая ошибка квантования (1.12) равна

$$E\{\varepsilon^2\} = \int_{d_1}^{d_{L+1}} (x - x_{\kappa\theta})^2 w(x) dx = \sum_{j=1}^L \int_{d_j}^{d_{j+1}} (x - r_j)^2 w(x) dx. \quad (1.13)$$

Дифференцируя (1.13) по переменным d_j , r_j и приравнивая производные нулю, получаем $2L - 1$ нелинейных уравнений

$$(d_j - r_j)^2 w(d_j) - (d_j - r_{j-1})^2 w(d_j) = 0, \quad j = \overline{2, L}, \quad (1.14)$$

$$\int_{d_j}^{d_{j+1}} (x - r_j) w(x) dx = 0, \quad j = \overline{1, L}.$$

Следует отметить, что крайние пороги d_1 и d_{L+1} определяются динамическим диапазоном яркости. Уравнения (1.14) нетрудно привести к виду

$$d_j = (r_j + r_{j-1})/2, \quad j = \overline{2, L}, \quad (1.15)$$

$$r_j = \frac{\int_{d_j}^{d_{j+1}} x w(x) dx}{\int_{d_j}^{d_{j+1}} w(x) dx}, \quad j = \overline{1, L}.$$

Из (1.15) следует, что пороги d_j должны располагаться по середине между двумя соседними уровнями r_{j-1} и r_j . Решение этих уравнений можно найти итеративным способом. Оптимальный квантователь, удовлетворяющий критерию (1.12), называется квантователем Ллойда-Макса [1.3, 1.5], а среднеквадратическая ошибка для такого квантователя равна

$$E\{\varepsilon^2\} = E\{x^2\} - \sum_{j=1}^L r_j^2 \frac{\int_{d_j}^{d_{j+1}} w(x) dx}{d_{j+1} - d_j} \quad (1.16)$$

При равномерном распределении яркости нелинейные уравнения (1.15) можно представить в виде [1.3]

$$d_j = d_1 + (j-1)\Delta, \quad j = \overline{2, L},$$

$$r_j = d_j + \Delta/2, \quad j = \overline{1, L},$$

$$\Delta = (d_{L+1} - d_1)/L,$$

$$\text{а среднеквадратическая ошибка равна } E\{\varepsilon^2\} = \frac{\Delta^2}{12}.$$

В системах цифровой обработки изображений стремятся уменьшить число уровней и порогов квантования, т.к. от их количества зависит длина двоичного кодового слова, которым представляются проквантованные отсчеты в ЭВМ. Однако при относительно небольшом числе уровней L на проквантованном изображении появляются ложные контуры. Они возникают вследствие скачкообразного изменения яркости проквантованного изображения (рис.1.6) и особенно заметны на пологих участках ее изменения.

Ложные контуры значительно ухудшают визуальное качество изображения, т.к. зрение человека особенно чувствительно именно к контурам. При равномерном квантовании типичных изображений требуется не менее 64 уровней. На рис.1.7.а и 1.7.б приведены результаты равномерного квантования изображения «Портрет» соответственно на 256 и 14 уровнях квантования.

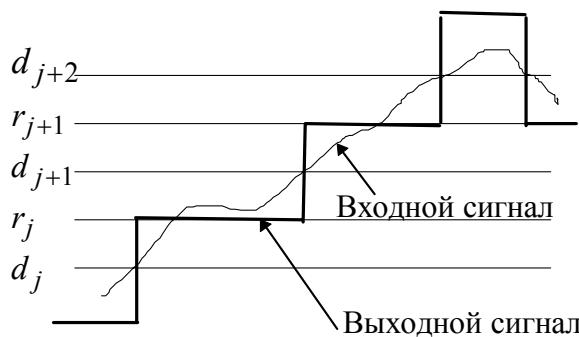


Рис.1.6. К механизму возникновения ложных контуров



a)



б)

Рис.1.7. Результаты равномерного квантования



Рис.1.8. Результат неравномерного квантования

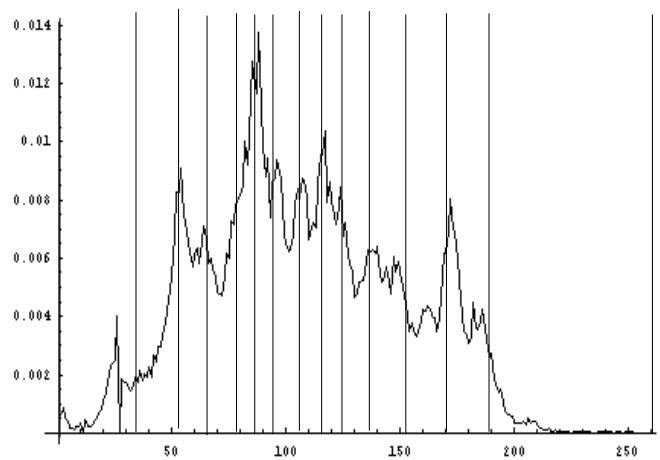


Рис.1.9. Гистограмма изображения “Портрет”

В темных частях изображения на рис. 1.7.б заметны ложные контуры. Использование квантователя Ллойда-Макса позволяет существенно снизить их уровень (см. рис. 1.8, где число уровней квантования также равно 14). На рис. 1.9 приведена гистограмма яркости изображения «Портрет» при 256 уровнях квантования и отмечены пороги d_j при $L+1=15$. Из рисунка следует, что чаще квантуются те области динамического диапазона, в которых сгруппированы значения яркости отсчетов.

Чтобы избежать неравномерного квантования, которое не может быть выполнено с помощью стандартного АЦП, используют нелинейные

преобразования (рис.1.10). Отсчет x исходного изображения подвергается нелинейному преобразованию, чтобы плотность распределения вероятностей преобразованных отсчетов y была равномерной, т.е. выполняется процедура эквализации, которая подробно описана в главе 2. Затем отсчеты y квантуются с равномерным шагом и подвергаются обратному нелинейному преобразованию.

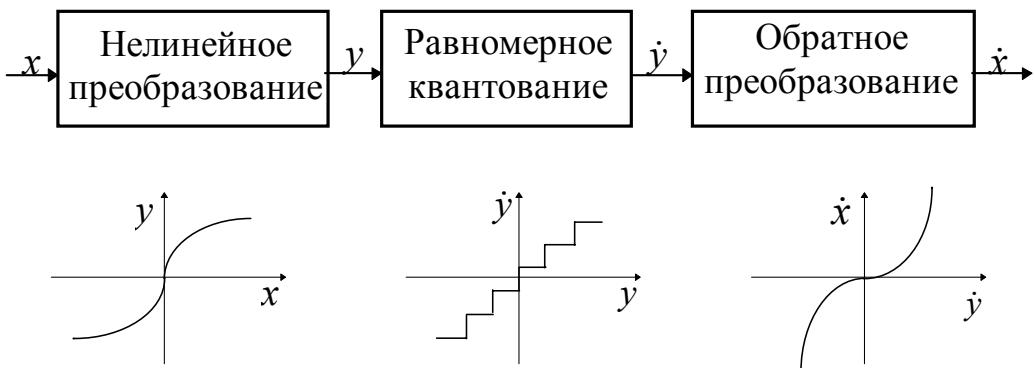


Рис.1.10. Квантование с предварительным нелинейным преобразованием

Для разрушения ложных контуров Робертс предложил перед равномерным квантованием к отсчетам яркости добавлять шум с равномерной плотностью распределения вероятностей. Добавленный шум переводит одни отсчеты изображения на уровень выше, а другие на уровень ниже. Тем самым разрушаются ложные контуры. Дисперсия добавляемого шума должна быть небольшой, чтобы не привести к искажениям, воспринимаемым как «снег» на изображении, и в то же время достаточной для разрушения ложных контуров. Обычно используют равномерно распределенный шум на интервале $[-\Delta/2, \Delta/2]$. Результаты равномерного квантования на 14 и 8 уровнях изображения «Портрет» с предварительным добавлением шума приведены на рис.1.11.а и 1.11.б. При 8-ми уровнях квантования добавляемый шум становится слишком заметным, однако ложные контуры разрушены практически полностью.



a)



б)

Рис.1.11. Результаты равномерного квантования с предварительным добавлением шума

Еще один метод квантования используется в полиграфии. Это метод формирования растровых бинарных (2-х уровневых) изображений из полутоноевых. При печати (например, газет или журналов) изображение формируется из белых и черных точек. Для этого все исходное изображение разбивается по пространственным координатам на одинаковые квадратные блоки. Обычно блок содержит 5×5 элементов. К каждому отсчету блока добавляется число с соответствующими координатами из матрицы возмущающего сигнала, размеры которой равны размерам блока. Например, в качестве матрицы возмущающего сигнала используют числа [1.5]:

$$\begin{bmatrix} 40 & 60 & 150 & 90 & 10 \\ 80 & 170 & 240 & 200 & 110 \\ 140 & 210 & 250 & 220 & 130 \\ 120 & 190 & 230 & 180 & 70 \\ 20 & 100 & 160 & 50 & 30 \end{bmatrix}.$$

Эта операция повторяется для всех блоков. Получаемое при этом изображение квантуется на два уровня. На рис. 1.12.а приведено полутононое изображение «Портрет» с добавленным возмущающим сигналом. На рис. 1.12.б,в приведены результаты бинарного квантования изображения «Портрет» с добавленным возмущающим сигналом (рис.1.13.б) и без него (рис.1.13.в).

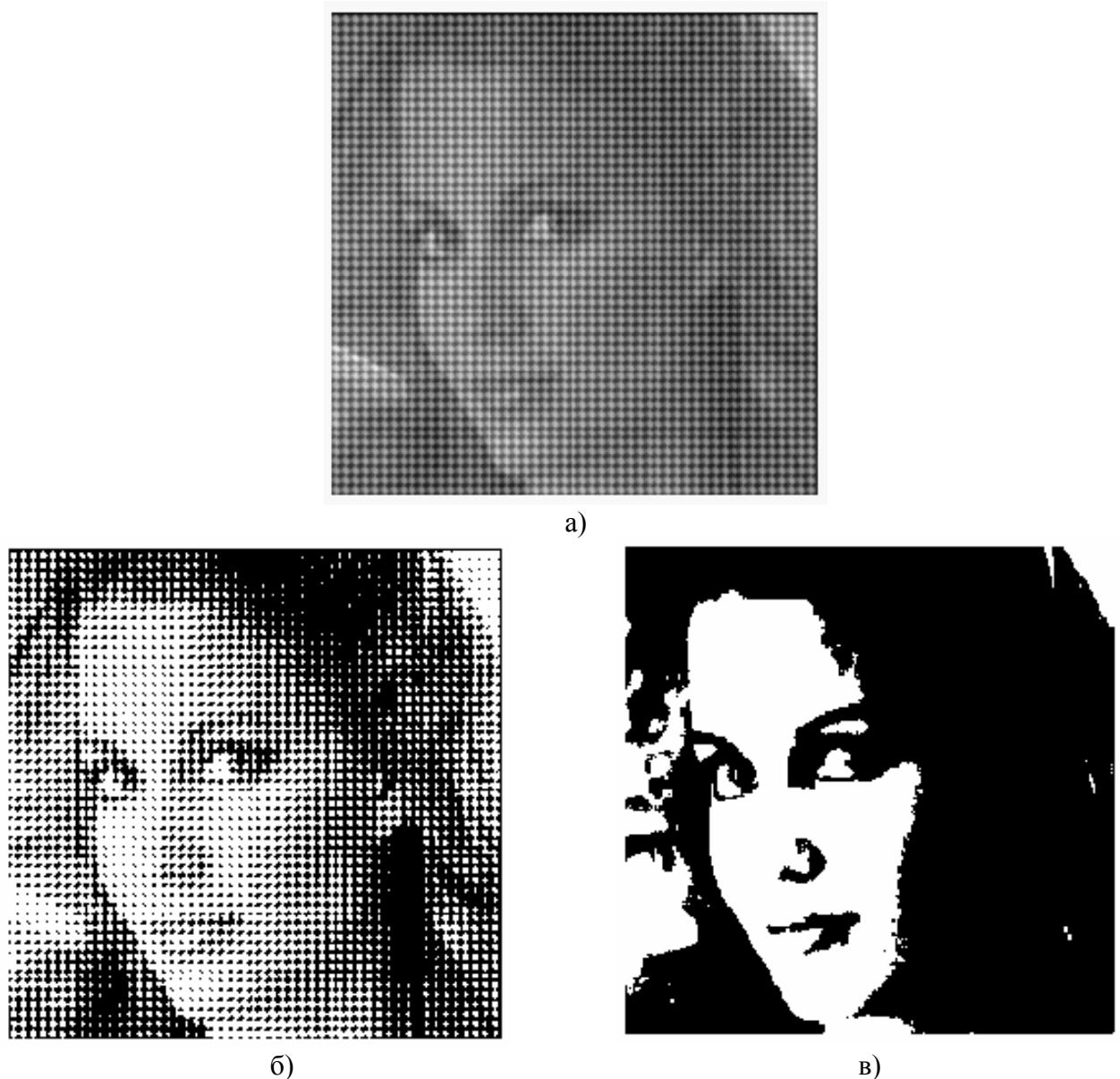


Рис.1.12.Растирование изображений

Бинарное растровое изображение обеспечивает значительно лучшее зрительное впечатление, чем обычное бинарное изображение. Передача шкалы яркости при растировании достигается благодаря изменению геометрических размеров белого пятна, наблюдаемого на черном фоне. Если в блоке сгруппировались «светлые» отсчеты, то геометрические размеры белого пятна максимальны и равны размеру блока. При уменьшении яркости его геометрические размеры также уменьшаются. Глаз человека выполняет локальное усреднение, создавая иллюзию наблюдения полутоонового изображения. Процедура растирования особенно эффективна при печати изображений с высоким разрешением, когда одиночное пятно едва различимо глазом.

ВОПРОСЫ К ГЛАВЕ 1

- 1.1. Поясните как проблема дискретизации изображения связана с двумерной интерполяцией?
- 1.2. Какова структура двумерного частотного спектра дискретного изображения?
- 1.3. При каких условиях, используя дискретное изображение, можно без потерь восстановить непрерывное?
- 1.4. Докажите, что двумерный фильтр с прямоугольной частотной характеристикой идеально восстанавливает непрерывное изображение из дискретного.
- 1.5. Учитывая, что спектры реальных изображений не являются финитными функциями, предложите метод расчета ошибок восстановления, вызванных нарушением требований (1.9).
- 1.6. Поясните в чем состоит задача поиска оптимального квантователя изображения?
- 1.7. Чем объясняется стремление использовать небольшое число уровней квантования и что препятствует этой тенденции?
- 1.8. Поясните механизм возникновения ложных контуров при малом числе уровней квантования изображения.
- 1.9. Каким образом добавление шума к изображению позволяет ослабить эффект ложных контуров, возникающих при квантовании? Как при этой процедуре следует выбирать характеристики шума?
- 1.10. С какой целью применяют нелинейное предискажение изображения при его равномерном квантовании? Какой должна быть характеристика нелинейного элемента?

Глава 2

УЛУЧШЕНИЕ ВИЗУАЛЬНОГО КАЧЕСТВА ИЗОБРАЖЕНИЙ ПУТЕМ ПОЭЛЕМЕНТНОГО ПРЕОБРАЗОВАНИЯ

В большом числе информационных систем применяется представление результатов обработки данных в виде изображения, выводимого на экран для использования наблюдателем. Процедуру, обеспечивающую такое представление, называют визуализацией. Желательно при помощи обработки придать выводимому изображению такие качества, благодаря которым его восприятие человеком было бы по возможности комфортным. Часто бывает полезным подчеркнуть, усилить какие-то черты, особенности, нюансы наблюданной картины с целью улучшения ее субъективного восприятия.

Последнее - субъективность восприятия - сильно усложняет применение формализованного подхода в достижении данных целей. Поэтому при обработке изображений для визуализации получили распространение методы, в которых часто отсутствуют строгие математические критерии оптимальности. Их заменяют качественные представления о целесообразности той или иной обработки, опирающиеся на субъективные оценки результатов.

Подавляющее большинство процедур обработки для получения результата в каждой точке кадра привлекает входные данные из некоторого множества точек исходного изображения, окружающих обрабатываемую точку. Однако имеется группа процедур, где осуществляется так называемая поэлементная обработка. Здесь результат обработки в любой точке кадра зависит только от значения входного изображения в этой же точке. Очевидным достоинством таких процедур является их предельная простота. Вместе с тем, многие из них приводят к очевидному субъективному улучшению визуального качества. Этим определяется внимание, которое уделяется поэлементным процедурам. Не преувеличивая их роли, отметим, что очень часто поэлементная обработка применяется как заключительный этап при решении более сложной задачи обработки изображения.

Сущность поэлементной обработки изображений сводится к следующему. Пусть $x(i, j) = x_{i,j}$, $y(i, j) = y_{i,j}$ - значения яркости исходного и получаемого после обработки изображений соответственно в точке кадра, имеющей декартовы координаты i (номер строки) и j (номер столбца). Поэлементная обработка означает, что существует функциональная однозначная зависимость между этими яркостями

$$y_{i,j} = f_{i,j}(x_{i,j}), \quad (2.1)$$

позволяющая по значению исходного сигнала определить значение выходного продукта. В общем случае, как это учтено в данном выражении, вид или параметры функции $f_{i,j}(\cdot)$, описывающей обработку, зависят от текущих

координат. При этом обработка является *неоднородной*. Однако в большинстве практически применяемых процедур используется *однородная* поэлементная обработка. В этом случае индексы i и j в выражении (2.1) могут отсутствовать. При этом зависимость между яркостями исходного и обработанного изображений описывается функцией:

$$y = f(x), \quad (2.2)$$

одинаковой для всех точек кадра.

Цель данной главы состоит в изучении наиболее распространенных процедур поэлементной обработки.

2.1. Линейное контрастирование изображения

Задача контрастирования связана с улучшением согласования динамического диапазона изображения и экрана, на котором выполняется визуализация. Если для цифрового представления каждого отсчета изображения отводится 1 байт (8 бит) запоминающего устройства, то входной или выходной сигналы могут принимать одно из 256 значений. Обычно в качестве рабочего используется диапазон 0...255; при этом значение 0 соответствует при визуализации уровню черного, а значение 255 - уровню белого. Предположим, что минимальная и максимальная яркости исходного изображения равны x_{\min} и x_{\max} соответственно. Если эти параметры или один из них существенно отличаются от граничных значений яркостного диапазона, то визуализированная картина выглядит как ненасыщенная, неудобная, утомляющая при наблюдении. Пример такого неудачного представления приведен на рис. 2.1.а, где диапазон яркостей имеет границы $x_{\min} = 180$, $x_{\max} = 240$.

При линейном контрастировании используется линейное поэлементное преобразование вида:

$$y = a \cdot x + b, \quad (2.3)$$

параметры которого a и b определяются желаемыми значениями минимальной y_{\min} и максимальной y_{\max} выходной яркости. Решив систему уравнений:

$$\begin{cases} y_{\min} = a \cdot x_{\min} + b \\ y_{\max} = a \cdot x_{\max} + b \end{cases},$$

относительно параметров преобразования a и b , нетрудно привести (2.3) к виду:

$$y = \frac{x - x_{\min}}{x_{\max} - x_{\min}} (y_{\max} - y_{\min}) + y_{\min}.$$

Результат линейного контрастирования исходного изображения, представленного на рис. 2.1.а, приведен на рис.2.1.б при $y_{\min} = 0$ и



а)



б)

Рис 2.1 Пример линейного контрастирования

$y_{\max} = 255$. Сравнение двух изображений свидетельствует о значительно лучшем визуальном качестве обработанного изображения. Улучшение связано с использованием после контрастирования полного динамического диапазона экрана, что отсутствует у исходного изображения.

2.2. Соляризация изображения

При данном виде обработки преобразование (2.2) имеет вид [2.1]:

$$y = k \cdot x \cdot (x_{\max} - x),$$

где x_{\max} - максимальное значение исходного сигнала, а k - константа, позволяющая управлять динамическим диапазоном преобразованного изображения. Функция, описывающая данное преобразование, является квадратичной параболой, ее график при $k = 1$ приведен на рис.2.2. При $y_{\max} = x_{\max}$ динамические диапазоны изображений совпадают, что может быть достигнуто при $k = 4/x_{\max}$. Другой путь нормализации динамического диапазона может состоять в применении одного из методов контрастирования, например, описанного выше линейного контрастирования.

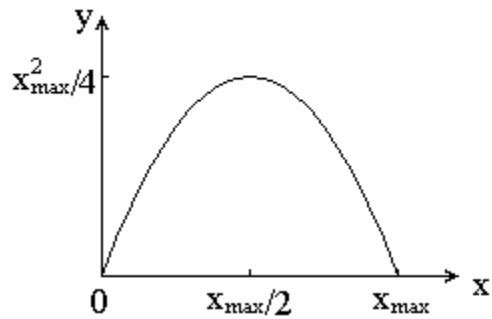


Рис. 2.2. Функция, описывающая соляризацию

Как следует из рис.2.2, смысл соляризации заключается в том, что участки исходного изображения, имеющие уровень белого или близкий к нему уровень яркости, после обработки имеют уровень черного. При этом сохраняют уровень черного и участки, имеющие его на исходном изображении. Уровень же белого на выходе приобретают участки, имеющие на входе средний уровень яркости (уровень серого). Пример применения соляризации приведен на рис.2.3.



Рис. 2.3. Пример соляризации

На рис.2.3.а показано исходное изображение, а на рис.2.3.б - результат его соляризации. На втором этапе обработки здесь применено линейное контрастирование при $y_{\min} = 0$ и $y_{\max} = 255$. Обработка приводит к повышению четкости деталей изображения: улучшены изображения глаз, повышен контраст на переходе “лицо - волосы” и т. д..

2.3. Препарирование изображения

Препарирование представляет собой целый класс поэлементных преобразований изображений. Характеристики применяемых на практике

процедур препарирования приведены на рис.2.4. Остановимся на описании некоторых из них.

Преобразование с пороговой характеристикой (рис.2.4.а) превращает полуточновое изображение, содержащее все уровни яркости, в бинарное, точки которого имеют яркости $y = 0$ или $y = y_{max}$. Такая операция, называемая иногда бинаризацией или бинарным квантованием, может быть полезной, когда для наблюдателя важны очертания объектов, присутствующих на изображении,

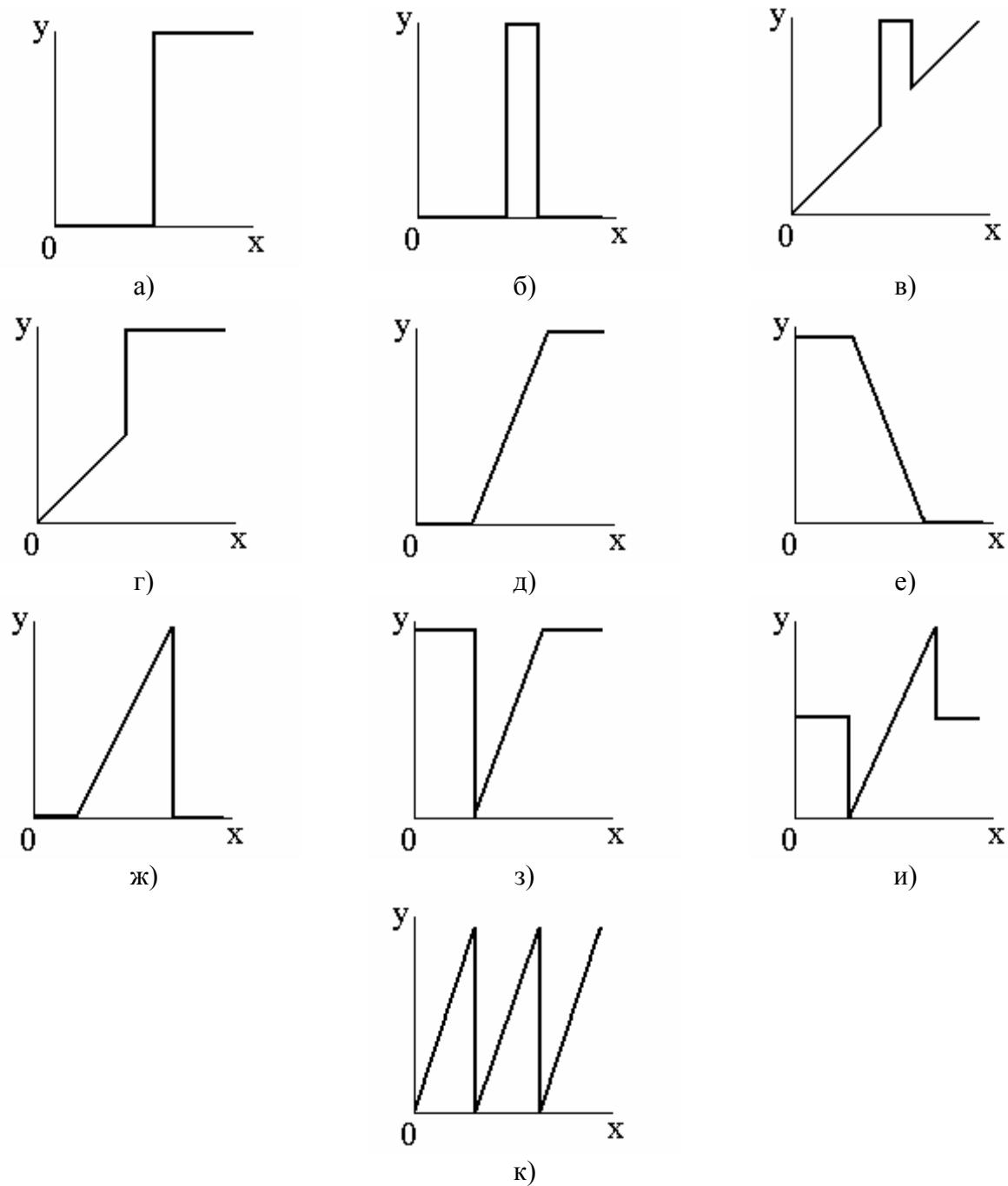


Рис.2.4.Примеры преобразований, используемых при препарировании

а детали, содержащиеся внутри объектов или внутри фона, не представляют интереса. Основной проблемой при проведении такой обработки является определение порога x_0 , сравнение с которым яркости исходного изображения позволяет определить значение выходного изображения в каждой его точке. Наиболее оправданным для математического описания изображения является применение теории вероятностей, случайных процессов и случайных полей. При этом определение оптимального порога бинарного квантования представляет собой статистическую задачу. Статистическому подходу к обработке изображений в последующих разделах уделяется значительное внимание, в том числе и при решении задачи разделения точек изображения на два класса - так

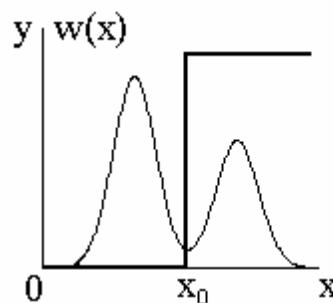


Рис.2.5. К выбору порога бинарного квантования

называемой бинарной сегментации. Здесь же ограничимся обсуждением частного, но практически важного случая. Иногда при обработке приходится иметь дело с изображениями, хранимыми как полутонаовые, но по своему содержанию мало отличающимися от бинарных. К ним относятся текст, штриховые рисунки, чертежи, изображение отпечатка пальца, пример которого приведен на рис.2.6.а. Плотность вероятности $w(x)$, описывающая распределение яркости такого изображения, может содержать два хорошо разделяющихся пика. Интуитивно понятно, что порог бинарного квантования следует выбирать посередине провала между этими пиками, как это показано на рис.2.5. Замена исходного полутонаового изображения *бинарным препаратом* решает две основные задачи. Во-первых, достигается большая наглядность при визуальном восприятии, чем у исходного изображения. Во-вторых, ощутимо сокращается объем запоминающего устройства для хранения изображения, поскольку бинарный препарат для записи каждой точки бинарного изображения требует лишь 1 бит памяти, в то время как полутонаовое изображение для решения той же задачи при наиболее часто применяемом формате представления - 8 бит. Пример бинаризации изображения отпечатка пальца приведен на рис.2.6.б.

Смысл других преобразований, представленных на рис.2.4, нетрудно понять, рассматривая их характеристики. Например, преобразование рис.2.4.б выполняет яркостный срез изображения, выделяя те его участки, где яркость соответствует выделенному интервалу. При этом остальные участки

оказываются полностью “погашенными” (имеют яркость, соответствующую уровню черного). Перемещая выделенный интервал по яркостной шкале и изменяя его ширину, можно детально исследовать содержание картины.

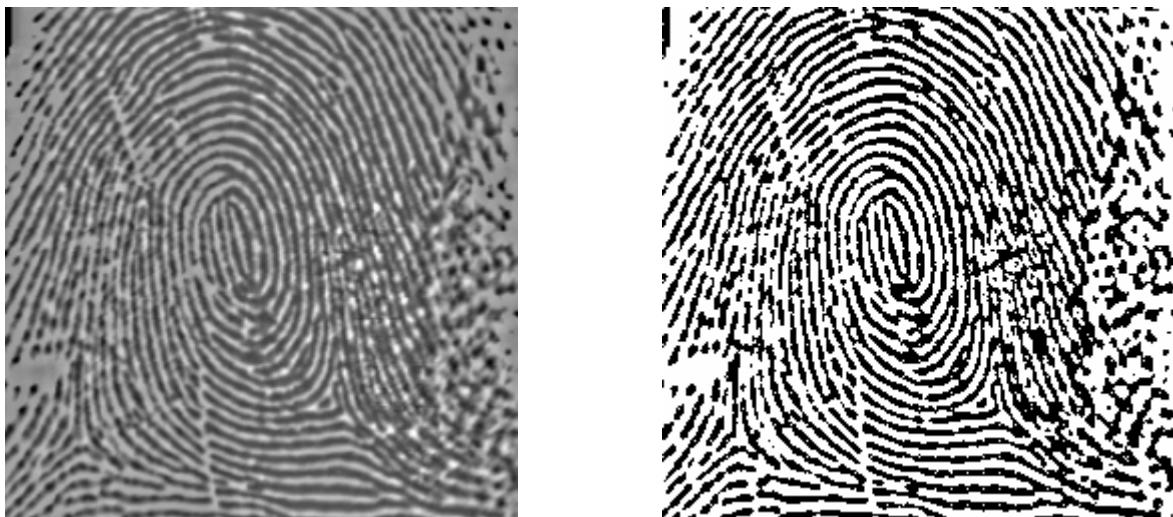


Рис.2.6. Пример бинаризации изображения

Преобразование, приведенное на рис.2.4.ж, также позволяет повысить детальность наблюдаемой картины в выбранном диапазоне яркостей, однако в отличие от предыдущего здесь выходное изображение использует полный динамический диапазон. По существу, это преобразование представляет собой линейное контрастирование, применяемое к *избранному диапазону* входного изображения. Как и в предыдущем варианте, участки, не попавшие в этот диапазон, образуют после препарирования черный фон.

Иногда наглядность изображения повышается применением преобразования типа пилообразного контрастирования (рис.2.4.к). При этом различные яркостные диапазоны одновременно подвергаются локальному яркостному контрастированию. Однако необходимо иметь в виду, что данное преобразование, как и некоторые другие, может сопровождаться появлением ложных контуров на получаемом препарате.

Аналогично можно качественно рассмотреть и остальные процедуры препарирования, представленные на рис.2.4.

На рис.2.7 приведены результаты эксперимента, в котором к аэроснимку участка земли (рис.2.7.а) применялись преобразования типа пороговая обработка (рис.2.7.б) и пилообразное контрастирование (рис.2.7.в). Первое приводит к выявлению границ отдельных участков, создавая общее интегральное представление о наблюдаемой сцене. Второе, наоборот, дает возможность наблюдения мелких деталей на всех участках изображения. Сочетание двух таких возможностей может оказаться полезным наблюдателю.

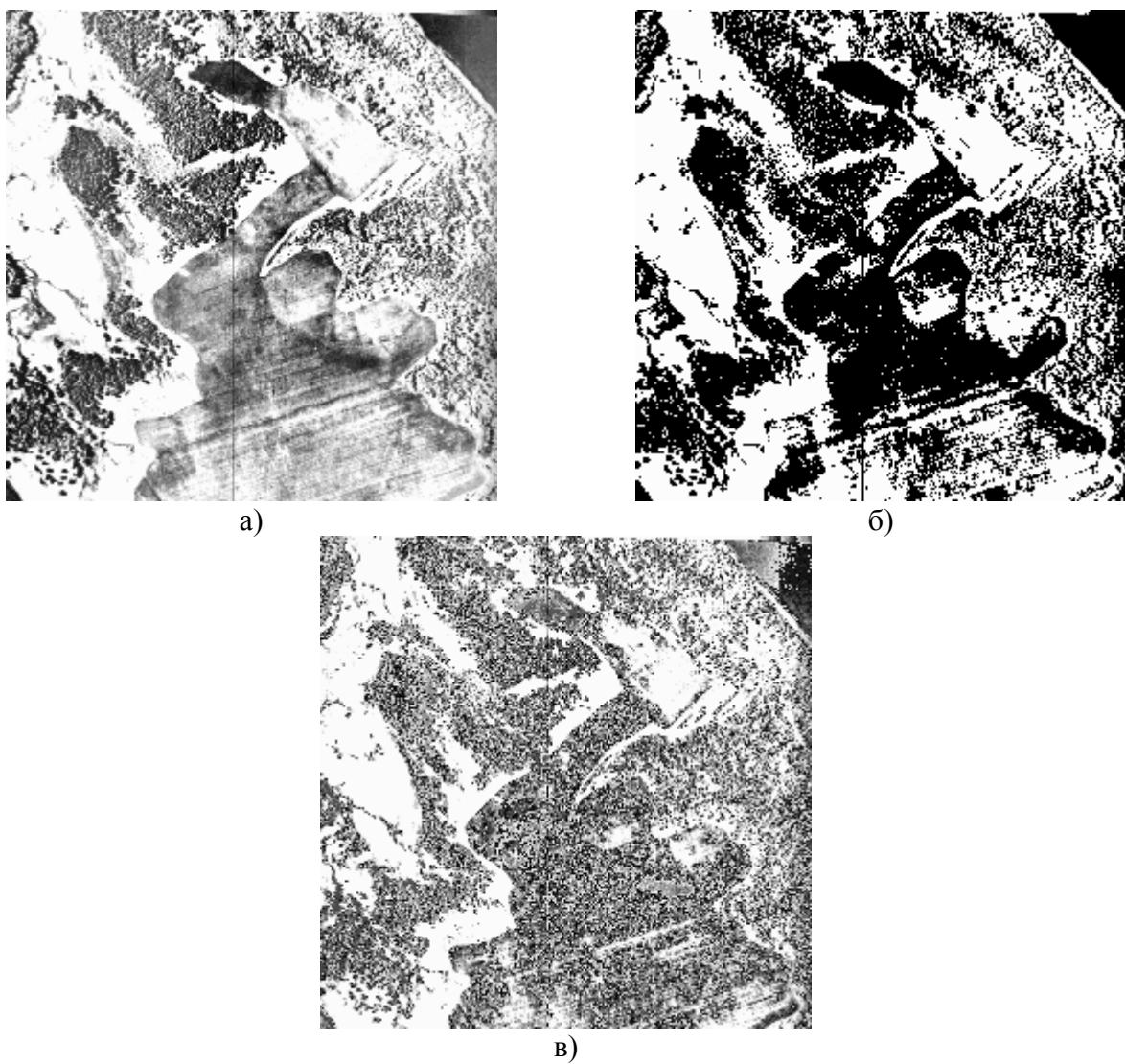


Рис.2.7. Примеры препарирования изображения

В заключение отметим, что препарирование часто используется и в автоматических системах обработки визуальной информации, поскольку подготовливаемый при этом препарат может содержать всю информацию, необходимую для последующей (вторичной) обработки. Например, если при наблюдении из космоса требуется автоматически обнаружить на изображении некоторый объект, имеющий известную конфигурацию, то для этого может быть достаточно бинарного препарата, передающего эту конфигурацию.

2.4. Преобразование гистограмм, эквализация

При всех поэлементных преобразованиях происходит изменение закона распределения вероятностей, описывающего изображение. Рассмотрим механизм этого изменения на примере произвольного преобразования с монотонной характеристикой, описываемой функцией $y = f(x)$ (рис.2.8),

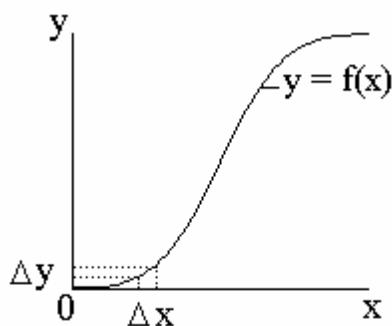


Рис.2.8. Нелинейное преобразование случайной величины

имеющей однозначную обратную функцию $x = \varphi(y)$. Предположим, что случайная величина x подчиняется плотности вероятности $w_x(x)$. Пусть Δx - произвольный малый интервал значений случайной величины x , а Δy - соответствующий ему интервал преобразованной случайной величины y . Попадание величины x в интервал Δx влечет за собой попадание величины y в интервал Δy , что означает вероятностную эквивалентность этих двух событий. Поэтому, учитывая малость обоих интервалов, можно записать приближенное равенство:

$$w_x(x)|\Delta x| \approx w_y(y)|\Delta y|,$$

где модули учитывают зависимость вероятностей от *абсолютных длин* интервалов (и независимость от *знаков* приращений Δx и Δy). Вычисляя отсюда плотность вероятности преобразованной величины, подставляя вместо x его выражение через обратную функцию и выполняя предельный переход при $\Delta x \rightarrow 0$ (и, следовательно, $\Delta y \rightarrow 0$), получаем :

$$w_y(y) = w_x(\varphi(y)) \cdot \left| \frac{d\varphi(y)}{dy} \right|. \quad (2.4)$$

Это выражение позволяет вычислить плотность вероятности продукта преобразования, которая, как видно из него, *не совпадает* с плотностью распределения исходной случайной величины. Ясно, что существенное влияние на плотность $w_y(y)$ оказывает выполняемое преобразование, поскольку в (2.4) входит его обратная функция и ее производная.

Соотношения становятся несколько сложнее, если преобразование описывается не взаимно-однозначной функцией [2.2]. Примером такой более сложной характеристики с неоднозначной обратной функцией может служить пилообразная характеристика рис.2.4.к. Однако, в общем, смысл вероятностных преобразований при этом не изменяется.

Все рассмотренные в данной главе поэлементные преобразования изображений можно рассмотреть с точки зрения изменения плотности вероятности, описываемого выражением (2.4). Очевидно, что ни при одном из них плотность вероятности выходного продукта не будет совпадать с плотностью вероятности исходного изображения (за исключением, конечно, тривиального преобразования $y = f(x) \equiv x$). Нетрудно убедиться, что при линейном контрастировании сохраняется вид плотности вероятности, однако в общем случае, т.е. при произвольных значениях параметров линейного преобразования, изменяются параметры плотности вероятности преобразованного изображения.

Определение вероятностных характеристик изображений, прошедших нелинейную обработку, является *прямой* задачей анализа. При решении практических задач обработки изображений может быть поставлена *обратная задача*: по *известному* виду плотности вероятности $w_x(x)$ и *желаемому* виду $w_y(y)$ определить *требуемое* преобразование $y = f(x)$, которому следует подвергнуть исходное изображение. В практике цифровой обработки изображений часто к полезному результату приводит преобразование изображения к равновероятному распределению [2.3]. В этом случае

$$w_y(y) = \begin{cases} 1/(y_{\max} - y_{\min}), & \text{при } y_{\min} \leq y \leq y_{\max}, \\ 0, & \text{при других } y, \end{cases} \quad (2.5)$$

где y_{\min} и y_{\max} - минимальное и максимальное значения яркости преобразованного изображения. Определим характеристику преобразователя, решающего данную задачу. Пусть x и y связаны функцией (2.2), а $F_x(x)$ и $F_y(y)$ - *интегральные законы распределения* входной и выходной величин. Учитывая (2.5), находим:

$$F_y(y) = \int_{y_{\min}}^y w_y(y) dy = \frac{y - y_{\min}}{y_{\max} - y_{\min}}.$$

Подставляя это выражение в условие вероятностной эквивалентности

$$F_x(x) = F_y(y),$$

после простых преобразований получаем соотношение

$$y = (y_{\max} - y_{\min}) \cdot F_x(x) + y_{\min}, \quad (2.6)$$

представляющее собой характеристику (2.2) в решаемой задаче. Согласно (2.6) исходное изображение проходит нелинейное преобразование, характеристика которого $F_x(x)$ определяется интегральным законом распределения самого исходного изображения. После этого результат приводится к заданному динамическому диапазону при помощи операции линейного контрастирования.

Аналогичным образом могут быть получены решения других подобных задач, в которых требуется привести законы распределения изображения к заданному виду. В [2.4] приведена таблица таких преобразований. Одно из них, так называемая *гиперболизация распределения*, предполагает приведение плотности вероятности преобразованного изображения к гиперболическому виду:

$$w_y(y) = \begin{cases} \frac{1}{y(\ln y_{\max} - \ln y_{\min})}, & \text{при } y_{\max} \leq y \leq y_{\min}, \\ 0, & \text{при других } y \end{cases} . \quad (2.7)$$

Если учесть, что при прохождении света через глаз входная яркость логарифмируется его сетчаткой, то итоговая плотность вероятности оказывается равномерной. Таким образом, отличие от предыдущего примера заключается в учете физиологических свойств зрения. Можно показать, что изображение с плотностью вероятности (2.7) получается на выходе нелинейного элемента с характеристикой:

$$y = y_{\min} \cdot \left(\frac{y_{\max}}{y_{\min}} \right)^{F_x(x)}, \quad (2.8)$$

также определяемой интегральным законом распределения исходного изображения.

Таким образом, преобразование плотности вероятности предполагает знание интегрального распределения для исходного изображения. Как правило, достоверные сведения о нем отсутствуют. Использование для рассматриваемых целей аналитических аппроксимаций также малопригодно, т.к. их небольшие отклонения от истинных распределений могут приводить к существенному отличию результатов от требуемых. Поэтому в практике обработки изображений преобразование распределений выполняют в два этапа.

На первом этапе измеряется *гистограмма* исходного изображения. Для цифрового изображения, шкала яркостей которого, например, принадлежит целочисленному диапазону 0...255, гистограмма представляет собой таблицу из 256 чисел. Каждое из них показывает количество точек в кадре, имеющих данную яркость. Разделив все числа этой таблицы на общий размер выборки, равный числу используемых точек изображения, получают *оценку распределения вероятностей яркости изображения*. Обозначим эту оценку

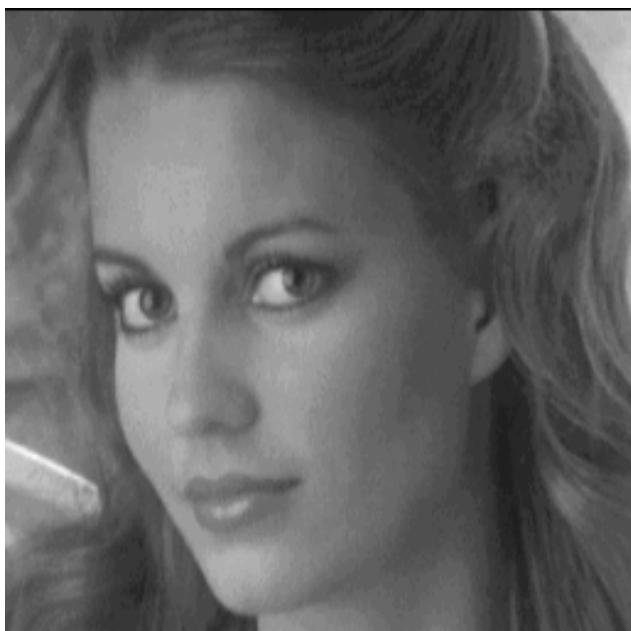
$w_x^*(j)$, $0 \leq j \leq 255$. Тогда оценка интегрального распределения получается по формуле:

$$F_x^*(j) = \sum_{i=0}^j w_x^*(i).$$

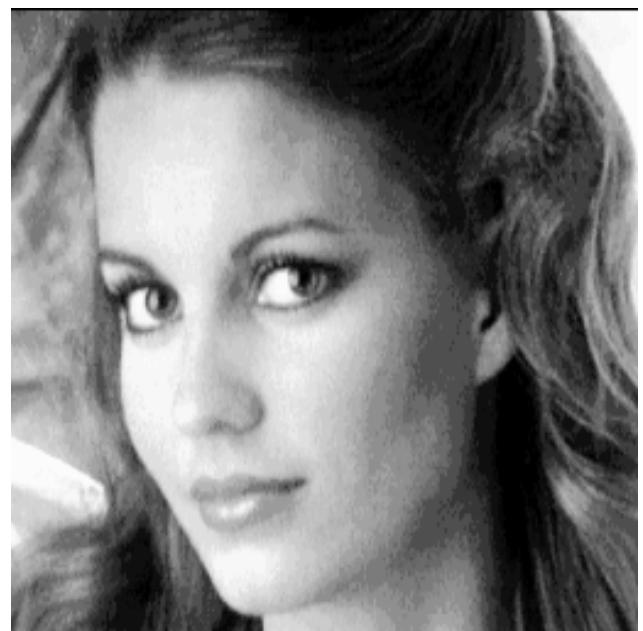
На втором этапе выполняется само нелинейное преобразование (2.2), обеспечивающее необходимые свойства выходного изображения. При этом вместо *неизвестного* истинного интегрального распределения используется его оценка, основанная на гистограмме. С учетом этого все методы поэлементного преобразования изображений, целью которых является видоизменение законов распределения, получили название *гистограммных методов*. В частности, преобразование, при котором выходное изображение имеет равномерное распределение, называется *эквализацией (выравниванием) гистограмм*.

Отметим, что процедуры преобразования гистограмм могут применяться как к изображению в целом, так и к отдельным его фрагментам. Последнее может быть полезным при обработке *нестационарных* изображений, содержание которых существенно различается по своим характеристикам на различных участках. В этом случае лучшего эффекта можно добиться, применяя гистограммную обработку к отдельным участкам.

Использование соотношений (2.4)-(2.8), справедливых для изображений с непрерывным распределением яркости, является не вполне корректным для цифровых изображений. Необходимо иметь в виду, что в результате обработки не удается получить идеальное распределение вероятностей выходного изображения, поэтому полезно проводить контроль его гистограммы.



а) исходное изображение



б) результат обработки

Рис. 2.9. Пример эквализации изображения

На рис.2.9 приведен пример эквализации, выполненной в соответствии с изложенной методикой. Характерной чертой многих изображений, получаемых в реальных изображающих системах, является значительный удельный вес темных участков и сравнительно малое число участков с высокой яркостью. Эквализация призвана откорректировать картину, выровняв интегральные площади участков с различными яркостями. Сравнение исходного (рис.2.9.а) и обработанного (рис.2.9.б) изображений показывает, что происходящее при обработке перераспределение яркостей приводит к улучшению визуального восприятия.

2.5. Применение табличного метода при поэлементных преобразованиях изображений

При поэлементных преобразованиях изображений вычисления по формуле (2.2) должны быть выполнены для всех точек исходного изображения. Даже при не очень трудоемких вычислениях в соответствии с видом применяемого преобразования общий объем вычислительной работы, выполняемой ЭВМ, может оказаться значительным. В тех же задачах, где функция $f(x)$ в (2.2) предполагает трудоемкие вычисления (умножения, деления, расчеты тригонометрических выражений, степенных и других функций), применение прямого метода преобразования яркостей может оказаться вообще неприемлемым. Неудобным является и то, что время выполнения обработки зависит от ее вычислительной сложности. От этих недостатков избавлен табличный метод, получивший широкое распространение в практике цифровой обработки изображений.

Сущность табличного метода состоит в том, что путем *предварительного расчета* создается таблица функции $y = f(x)$, общий вид которой показан в

табл.2.1.

x	x_1	x_2	x_{n-1}	x_n
$y=f(x)$	y_1	y_2	y_{n-1}	y_n

Табл.2.1. Табличное задание функции $y = f(x)$

При обработке изображения вместо вычислений используются готовые результаты путем обращения к этой таблице. При этом значение входной яркости x служит для определения *номера столбца*, из которого должна быть считана величина преобразованного сигнала y . Выполнение этой несложной операции по сравнению с вычислением на каждом шаге значения y позволяет сделать обработку достаточно технологичной, причем ее продолжительность оказывается независящей от сложности преобразования.

Необходимо иметь в виду, что все реальные таблицы, которые могут быть записаны в оперативной памяти ЭВМ, имеют ограниченную длину. Если множество значений входного сигнала превышает размеры таблицы, то при попадании значения x между точками, содержащимися в ней, приходится применять интерполяцию - приближенное определение недостающих значений функции $y = f(x)$ по имеющимся ее соседним значениям. Часто для этой цели используется *линейная интерполяция*, при которой на участке между заданными узлами незаданная функция заменяется отрезком прямой.

Вместе с тем, если исходное изображение представлено в целочисленной записи в диапазоне 0..255, то размеры полной таблицы, содержащей все эти значения, вполне приемлемы для ее хранения. Кроме того, в этом случае значение яркости исходного сигнала служит *адресом*, определяющим номер столбца в таблице. Обработка с использованием этого метода оказывается очень удобной и быстродействующей.

ВОПРОСЫ К ГЛАВЕ 2

- 2.1. В чем состоит сущность поэлементной обработки изображений ?
- 2.2. Как определяются параметры преобразования изображения при его линейном контрастировании ?
- 2.3. Поясните механизм действия поэлементных преобразований, применяемых при препарировании изображения (рис.2.4).
- 2.4. Каков механизм появления ложных контуров при применении пилообразного контрастирования? При каких еще процедурах могут возникать ложные контуры?
- 2.5. Докажите, что применение линейного контрастирования не изменяет *вида* плотности вероятности изображения. Как при этом изменяются параметры плотности вероятности?
- 2.6. Докажите справедливость преобразования (2.7) для получения гиперболического распределения (2.6).
- 2.7. Докажите, что применение гиперболизации распределения яркости приводит к равновероятному распределению сигнала на выходе сетчатки глаза, если учесть ее логарифмирующее воздействие на входной свет.
- 2.8. Поясните, почему при эквализации изображения не удается привести гистограмму к идеальному равномерному виду ?
- 2.9. В чем состоит сущность и каковы достоинства табличного метода поэлементного преобразования изображений ?

Глава 3

ФИЛЬТРАЦИЯ ИЗОБРАЖЕНИЙ

Обычно изображения, сформированные различными информационными системами, искажаются действием помех. Это затрудняет как их визуальный анализ человеком-оператором, так и автоматическую обработку в ЭВМ. При решении некоторых задач обработки изображений в роли помех могут выступать и те или иные компоненты самого изображения. Например, при анализе космического снимка земной поверхности может стоять задача определения границ между ее отдельными участками - лесом и полем, водой и сушей и т.п. С точки зрения этой задачи отдельные детали изображения внутри разделяемых областей являются помехой.

Ослабление действия помех достигается фильтрацией . При фильтрации яркость (сигнал) каждой точки исходного изображения, искаженного помехой, заменяется некоторым другим значением яркости, которое признается в наименьшей степени искаженным помехой. Что может послужить основой для таких решений ? Изображение часто представляет собой двумерную функцию пространственных координат, которая изменяется по этим координатам медленнее (иногда значительно медленнее), чем помеха, также являющаяся двумерной функцией. Это позволяет при оценке полезного сигнала в каждой точке кадра *принять во внимание* некоторое множество соседних точек, воспользовавшись определенной *похожестью* сигнала в этих точках. В других случаях, наоборот, признаком полезного сигнала являются резкие перепады яркости. Однако, как правило, частота этих перепадов относительно невелика, так что на значительных промежутках между ними сигнал либо постоянен, либо изменяется медленно. И в этом случае свойства сигнала проявляются при наблюдении его не только в локальной точке, но и при анализе ее окрестности. Заметим, что понятие окрестности является достаточно условным. Она может быть образована лишь ближайшими по кадру соседями, но могут быть окрестности, содержащие достаточно много и достаточно сильно удаленных точек кадра. В этом последнем случае, конечно, степень влияния далеких и близких точек на решения, принимаемые фильтром в данной точке кадра, будет совершенно различной.

Таким образом, идеология фильтрации основывается на рациональном использовании данных как из рабочей точки, так и из ее окрестности. В этом проявляется существенное отличие фильтрации от рассмотренных выше поэлементных процедур: фильтрация не может быть поэлементной процедурой обработки изображений.

Задача заключается в том, чтобы найти такую рациональную вычислительную процедуру, которая позволяла бы достигать наилучших результатов. Общепринято при решении этой задачи опираться на использование вероятностных моделей изображения и помехи, а также на

применение статистических критериев оптимальности. Причины этого понятны - это случайный характер как информационного сигнала, так и помехи и это стремление получить минимальное в *среднем* отличие результата обработки от идеального сигнала. Многообразие методов и алгоритмов связано с большим разнообразием сюжетов, которые приходится описывать различными математическими моделями. Кроме того, применяются различные критерии оптимальности, что также ведет к разнообразию методов фильтрации. Наконец, даже при совпадении моделей и критериев очень часто из-за математических трудностей не удается найти оптимальную процедуру. Сложность нахождения точных решений порождает различные варианты приближенных методов и процедур.

3.1. Оптимальная линейная фильтрация. Уравнение Винера-Хопфа

Пусть $x_{i,j}$ - значение яркости изображения - полезного сигнала на пересечении i -ой строки и j -го столбца, а наблюдаемое на входе фильтра изображение описывается моделью:

$$y_{i,j} = f(x_{i,j}, n_{i,j}), \quad i = \overline{0, I-1}, j = \overline{0, J-1}. \quad (3.1)$$

Здесь $n_{i,j}$ - значение помехи в точке с координатами (i, j) , $f(\cdot)$ - функция, описывающая взаимодействие сигнала и помехи, а I и J - соответственно число строк и столбцов в кадре.

В дальнейшем будем придерживаться принятой при цифровой обработке изображений декартовой системы координат с началом в левом верхнем углу кадра и с положительными направлениями из этой точки вниз и вправо. На рис. 3.1 показаны примеры окрестностей различных типов, изображенные в виде совокупностей точек. Центром окрестностей, рабочей точкой, в которой осуществляется обработка, является точка с координатами i, j (на рис. 3.1 не зачернена). В зависимости от типа окрестности различают *каузальную*, *некаузальную* и *полекаузальную* фильтрацию изображений. Понятие

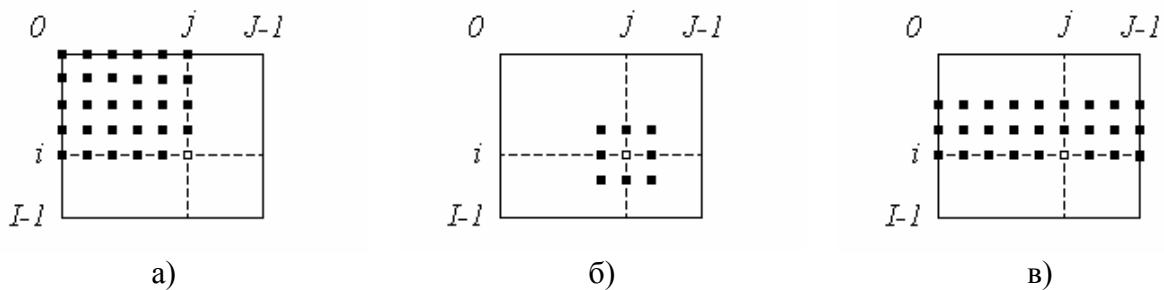


Рис. 3.1 Примеры окрестностей различных видов

каузальности (причинно-следственной зависимости) связывают с соотношением координат текущей точки i, j и точек, входящих в окрестность. Если *обе координаты* (номер строки и номер столбца) *всех* точек окрестности не превышают соответствующих координат текущей точки, то окрестность и использующая ее обработка называются каузальными. Пример такой окрестности представлен на рис. 3.1.а.

Некоторые точки окрестности, приведенной на рис. 3.1.б, удовлетворяют принципу каузальности. Вместе с тем, здесь имеются и такие точки, *обе координаты* которых превышают соответствующие координаты рабочей точки. Фильтрация, опирающаяся на использование окрестностей с сочетанием таких свойств, называется некаузальной.

Окрестности, показанной на рис. 3.1.в, соответствует полукаузальная фильтрация. Одна из координат *всех* точек окрестности - в данном примере номер строки - не превышает соответствующей координаты рабочей точки. Вторая же координата - в примере номер столбца - у некоторых точек также не превышает соответствующей координаты рабочей точки. Однако среди точек окрестности имеются и такие, у которых эта вторая координата превышает соответствующую координату рабочей точки.

Смысл, заложенный в данную классификацию, состоит в том, что, согласно принципу причинности, на формирование отклика физически осуществимого фильтра не могут оказывать влияния элементы входного сигнала, не поступившие к моменту формирования выходного отсчета. Этот принцип естественным образом «работает» в динамических системах, где все происходящие в них процессы являются временными процессами. При цифровой обработке изображений часто приходится иметь дело с ранее сформированными изображениями, уже хранящимися в памяти устройства обработки. В этом смысле соотношение координат, строго говоря, уже не играет такой принципиальной причинной роли, как при обработке сигналов в реальном масштабе времени. Вместе с тем, традиционно сложилась описанная выше классификация процедур обработки изображений, которой, в определенной мере, будем придерживаться и мы в последующем изложении.

При линейной фильтрации выходной эффект определяется линейной комбинацией входных данных:

$$x^*(i, j) = \sum_{(i_1, j_1) \in S} a(i_1, j_1) \cdot y(i - i_1, j - j_1). \quad (3.2)$$

В этом выражении $x^*(i, j) = x_{i,j}^*$ - результат фильтрации полезного сигнала $x(i, j) = x_{i,j}$ в точке кадра с координатами (i, j) ; S - множество точек (точнее - множество их координат), образующих окрестность; $a(i_1, j_1)$ - весовые коэффициенты, совокупность которых представляет собой *двумерную*

импульсную характеристику (ИХ). Если область S конечна, то импульсная характеристика имеет конечную длину и фильтр называется КИХ-фильтром. В противном случае импульсная характеристика имеет бесконечную длину, а фильтр называется БИХ-фильтром. В выражении (3.2) принято, что ИХ не зависит от координат точки (i, j) , в которой определяется выходной эффект. Процедуры обработки изображений, обладающие свойством независимости от координат, называются *однородными*.

Наиболее распространенным критерием оптимальности, применяемым для оценки качества обработки, является критерий минимума среднего квадрата ошибок. Применительно к фильтрации запишем его выражение в виде:

$$E \left\{ \left[x(i, j) - \sum_{(i_1, j_1) \in S} \sum a(i_1, j_1) \cdot y(i - i_1, j - j_1) \right]^2 \right\} = \min_{a(\cdot)}, \quad (3.3)$$

где $E\{\circ\}$ - символ математического ожидания. Согласно (3.3) отыскание оптимального фильтра заключается в определении его ИХ таким образом, чтобы средний квадрат ошибки $\varepsilon(i, j) = x(i, j) - x^*(i, j)$, выражающей различие между сигналом $x(i, j)$ и оценкой $x^*(i, j)$, формируемой фильтром, был минимальным. Математическое ожидание вычисляется по всем случайнм величинам, содержащимся в (3.3), что означает ориентацию критерия на учет *средних* ошибок.

Оптимизационную задачу (3.3) нетрудно свести к решению уравнения или системы уравнений. Для этого вычислим производную от левой части этого выражения по коэффициенту $a(k, l) = a_{k,l}$ и приравняем ее нулю. Учитывая, что операции дифференцирования, суммирования и математического ожидания являются линейными и поэтому перестановочны, приходим к выражению:

$$E\{x_{i,j} \cdot y_{i-k, j-l}\} = \sum_{i_1, j_1 \in S} \sum a_{i_1, j_1} \cdot E\{y_{i-i_1, j-j_1} \cdot y_{i-k, j-l}\}. \quad (3.4)$$

Входящие в него математические ожидания являются, как нетрудно видеть, отсчетами корреляционных функций, для которых введем следующие обозначения:

$$B_{xy}(k, l) = E\{x_{i,j} \cdot y_{i-k, j-l}\}, \quad B_y(k - i_1, l - j_1) = E\{y_{i-i_1, j-j_1} \cdot y_{i-k, j-l}\}.$$

С их учетом (3.4) примет более компактный вид:

$$B_{xy}(k, l) = \sum_{i_1, j_1 \in S} \sum a_{i_1, j_1} \cdot B_y(k - i_1, l - j_1) \quad (3.5)$$

Считая автокорреляционную $B_y(\circ)$ и взаимно корреляционную $B_{xy}(\circ)$ функции известными, замечаем, что (3.5) представляет собой линейное относительно искомых коэффициентов a_{i_1, j_1} алгебраическое уравнение. Число неизвестных в этом уравнении равняется числу точек n_S в окрестности S и является конечным в случае КИХ-фильтра и бесконечным при БИХ-фильтрации. Ограничимся в данном параграфе рассмотрением КИХ-фильтрации. Линейное алгебраическое уравнение со многими неизвестными имеет бесконечное множество решений. Если повторить дифференцирование (3.3) по остальным $n_S - 1$ неизвестным, то получим еще $n_S - 1$ уравнений, отличающихся друг от друга левыми частями $B_{xy}(\circ)$ и коэффициентами $B_y(\circ)$ в правых частях, т.к. определяющие их корреляции вычисляются каждый раз в различных точках. В результате образуется система n_S линейных алгебраических уравнений с n_S неизвестными, называемая в теории фильтрации *уравнением Винера-Хопфа*:

$$\left\{ \begin{array}{l} B_{xy}(k, l) = \sum_{i_1, j_1 \in S} a_{i_1, j_1} \cdot B_y(k - i_1, l - j_1), \\ k, l \in S. \end{array} \right. \quad (3.6)$$

Если разрешить ее относительно всех n_S неизвестных a_{i_1, j_1} , то будет найдена искомая импульсная характеристика линейного фильтра, минимизирующую средний квадрат ошибок фильтрации.

Определим средний квадрат ошибок оптимальной фильтрации. Для этого необходимо выполнить возведение в квадрат в выражении (3.3) и учесть в полученном выражении уравнение Винера-Хопфа (3.6). В результате нетрудно получить:

$$D_\varepsilon = D_x \left[1 - \sum_{i, j \in S} \sum a_{i, j} \cdot R_x(i, j) \right], \quad (3.7)$$

где $D_\varepsilon = E\{\varepsilon^2(\circ)\}$ - средний квадрат ошибок фильтрации.

Остановимся на анализе изменения средней яркости изображения при его фильтрации. Вычислив математическое ожидание от обеих частей (3.2), находим:

$$m_x^* = m_y \cdot \sum_{(i_1, j_1) \in S} \sum a(i_1, j_1), \quad (3.8)$$

где принято, что средняя яркость m_y входного изображения $y(\circ)$ не зависит от координат и, как результат, получено, что и средняя яркость m_{x^*} выходного изображения $x^*(\circ)$ также постоянна во всех точках кадра. Очень часто при обработке стремятся сохранить среднюю яркость изображения. Как следует из полученного выражения, достичь этого удается при выполнении равенства

$$\sum_{(i_1, j_1) \in S} a(i_1, j_1) = 1, \quad (3.9)$$

которое является дополнительным требованием к импульсной характеристике фильтра. Поэтому оптимизационную задачу (3.3) необходимо решать с учетом данного ограничения типа равенства.

Вместо этого часто перед фильтрацией осуществляют вычитание средней яркости m_y из входного изображения. Как следует из (3.8), среднее значение яркости на выходе фильтра при этом также равно нулю независимо от свойств импульсной характеристики. Это позволяет решать систему уравнений (3.6), игнорируя преобразование средней яркости. Желаемое же ее значение восстанавливается после фильтрации простым прибавлением к выходному эффекту.

3.2. Масочная фильтрация изображений при наличии аддитивного белого шума

Распространенным видом помехи является белый шум, аддитивно действующий на изображение. Наблюдаемое в этом случае изображение (3.1) имеет вид:

$$y_{i,j} = x_{i,j} + n_{i,j}, \quad i = \overline{0, I-1}, \quad j = \overline{0, J-1}, \quad (3.10)$$

а корреляционная функция шума $n_{i,j}$ описывается выражением:

$$B_n(i, j) = D_n \cdot \delta_i \cdot \delta_j.$$

Здесь D_n - дисперсия шума, а $\delta_k = \begin{cases} 1, & \text{при } k = 0 \\ 0, & \text{при } k \neq 0 \end{cases}$ - символ Кронекера.

Считаем, что входной сигнал центрирован, т.е. имеет нулевое математическое ожидание, а изображение $x(\circ)$ и шум $n(\circ)$ взаимно независимы, поэтому для корреляционной функции входного сигнала справедливо:

$$B_y(i, j) = D_x \cdot R_x(i, j) + D_n \cdot \delta_i \cdot \delta_j,$$

где D_x - дисперсия, а $R_x(i, j)$ - нормированная корреляционная функция полезного сигнала. Нетрудно видеть, что в этих условиях взаимная корреляционная функция $B_{xy}(\circ)$ совпадает с корреляционной функцией полезного сигнала $B_x(\circ)$. Поэтому уравнение Винера -Хопфа (3.6) приводится к виду:

$$\left\{ \begin{array}{l} q^2 \cdot R_x(k, l) = q^2 \cdot \sum_{i_1, j_1 \in S} \sum_{k, l \in S} a_{i_1, j_1} \cdot R_x(k - i_1, l - j_1) + a_{k, l}, \\ k, l \in S \end{array} \right. , \quad (3.11)$$

где $q^2 = D_x/D_n$ - отношение дисперсий сигнала и шума.

Преобразуем также выражение (3.7) для ошибок фильтрации, для чего запишем в явном виде то из уравнений в (3.11), которое соответствует значениям $k = l = 0$: $q^2 = q^2 \cdot \sum_{i_1, j_1 \in S} \sum_{k, l \in S} a_{i_1, j_1} \cdot R_x(i_1, j_1) + a_{0,0}$, откуда находим:

$a_{0,0}/q^2 = 1 - \sum_{i_1, j_1 \in S} \sum_{k, l \in S} a_{i_1, j_1} \cdot R_x(i_1, j_1)$. Сравнивая это соотношение с (3.7),

окончательно получаем:

$$\delta_\varepsilon^2 = a_{0,0}/q^2,$$

где $\delta_\varepsilon^2 = D_\varepsilon/D_x$ - относительный средний квадрат ошибок фильтрации. Таким образом, для определения ошибок фильтрации необходимо знать отношение сигнал/шум (которое входит также и в уравнение Винера-Хопфа) и значение оптимальной импульсной характеристики в точке (0,0).

Для того чтобы при решении уравнения (3.11) воспользоваться существующими программными средствами ЭВМ, необходимо выполнить его упорядоченное преобразование к каноническому векторно-матричному виду. Для этого требуется совокупность n_S неизвестных величин $a_{i,j}$ представить в виде вектора A . Точно также множество величин, образующих левые части (3.11), следует представить в виде вектора Q , а множество коэффициентов правой части в виде матрицы F размера $n_S \times n_S$. Тогда уравнение и его решение примут вид:

$$Q = F \cdot A, \quad A = F^{-1} \cdot Q.$$

В практике цифровой обработки изображений широко используется *масочная фильтрация*. Ее линейная разновидность является одним из вариантов двумерной КИХ-фильтрации. В качестве маски используется множество весовых коэффициентов, заданных во всех точках окрестности S , обычно симметрично окружающих рабочую точку кадра. Распространенным

видом окрестности, часто применяемым на практике, является квадрат 3×3 с рабочим элементом в центре, изображенный на рис. 3.1.б. Применяют различные разновидности масок, одним из эвристических вариантов является равномерная маска, все девять весовых коэффициентов которой равны $1/9$. Такой выбор коэффициентов отвечает условию сохранения средней яркости (3.9) и поэтому в процессе обработки центрировать изображение не требуется.

Визуально эффективность фильтрации можно оценить с помощью рис.3.2. На рис. 3.2.а показан зашумленный портрет (изображение без шума приведено на рис. 1.3.а) при отношении сигнал/шум равном -5дБ. Результат масочной фильтрации при оптимальном виде ИХ, найденной из (3.11), приведен на рис.3.2.б. Результат фильтрации, выполненной равномерным масочным оператором не приводится, поскольку с визуальной точки зрения он мало отличается от рис.3.2.б. При этом, однако, с количественной точки зрения различия достаточно заметны: если при оптимальной КИХ относительная ошибка $\delta_{\varepsilon \text{опт}}^2 = 0.309$, то при равномерной КИХ она возрастает почти на 30% и составляет $\delta_{\varepsilon \text{равн}}^2 = 0.395$. Различие резко возрастаёт при более высоком уровне шума. Так, например, при отношении сигнал/шум равном -10дБ имеем $\delta_{\varepsilon \text{опт}}^2 = 0.57$ и $\delta_{\varepsilon \text{равн}}^2 = 1.15$, т.е. применение равномерной КИХ вместо оптимальной приводит в этом случае к увеличению ошибок более чем вдвое.

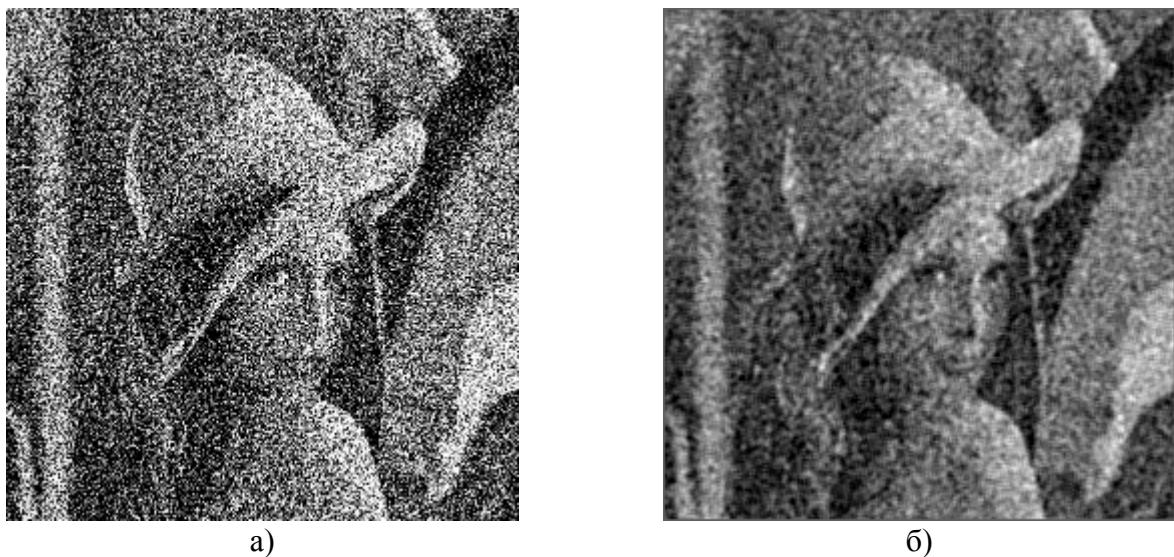


Рис. 3.2. Пример масочной фильтрации при $q^2 = -5$ дБ

Здесь полезно отметить определенное разногласие в оценках качества, даваемых человеческим глазом и применяемыми количественными показателями. Глаз является слишком совершенным изобретением природы, чтобы с ним могли соревноваться достаточно примитивные математические показатели типа среднего квадрата ошибок. Поэтому некоторые результаты, рассматриваемые с точки зрения математических показателей как

катастрофические, визуально могут быть вполне удовлетворительными. Означает ли это, что математические критерии вообще непригодны при цифровой обработке изображений? Конечно, нет. Цифровая обработка изображений находит применение в различных информационных системах с автоматическим принятием решений, основанным на этой обработке.

Функционирование таких систем, где отсутствует человеческий глаз, полностью подчинено математическим критериям и качество их работы оценивается только математическими показателями. Понятно, что и качество изображений, используемых в этих системах, также должно оцениваться только математическими критериями.

В заключение данного параграфа подчеркнем, что в целом применение описанных процедур фильтрации приводит к существенному снижению уровня шума на изображении. Количественно эффективность данной обработки можно охарактеризовать коэффициентом улучшения отношения сигнал/шум $\gamma = \delta_{\varepsilon}^{-2}/q^2$, где учтено, что величина $\delta_{\varepsilon}^{-2}$ определяет отношение сигнал/шум после фильтрации. Улучшение зависит от уровня шума на исходном изображении и составляет в приведенном эксперименте $\gamma = 10.2$ при $q^2 = -5$ дБ и $\gamma = 17.6$ при $q^2 = -10$ дБ. Коэффициент улучшения тем выше, чем сильнее шум на исходном изображении.

3.3. Рекуррентная каузальная фильтрация изображений

Проблема борьбы с шумом не решается полностью применением масочных фильтров по следующим причинам. Во-первых, ограниченность размера окрестности, используемой масочным фильтром, приводит к его потенциально ограниченной способности к подавлению шума. Это проявляется при значительном уровне шума на изображении - в меньшей степени при оптимальном выборе КИХ, сильнее при неоптимальной КИХ. Можно, конечно, увеличивать размер окрестности, прибегая к использованию КИХ-фильтров с более длинными импульсными характеристиками. Однако при этом усиливается второй недостаток масочного фильтра, состоящий в его и без того достаточно высокой вычислительной трудоемкости.

В настоящее время отсутствуют методы двумерной фильтрации, в которых сочетаются предельно достижимое качество фильтрации и низкие требования к вычислительным ресурсам ЭВМ, реализующей обработку. Существует много подходов к решению данной проблемы, но все они для достижения компромисса между точностью и реализуемостью прибегают к тем или иным приближениям. Рассмотрим один из них [3.1].

Идея заключается в использовании двумерного БИХ-фильтра с таким видом импульсной характеристики, при которой его практическая реализация была бы простой, и с такими параметрами этой импульсной характеристики, при которых эффективность фильтрации приближалась бы к потенциальному

возможной. Создать фильтр с такими свойствами удается на основе аналогии с одномерным фильтром Калмана.

Наиболее простым примером одномерной фильтрации является калмановская фильтрация однородной стационарной гауссовской последовательности, имеющей корреляционную функцию экспоненциального вида

$$B_x(r) = D_x \cdot \exp(-\alpha \cdot |r|) = D_x \cdot F^{|r|}, \quad F = \exp(-\alpha), \quad \alpha > 0.$$

Здесь D_x - дисперсия последовательности, а F - коэффициент ее одношаговой корреляции, определяемый параметром α , имеющим смысл ширины спектра последовательности. При ее наблюдении на фоне гауссовского белого шума оптимальный каузальный фильтр реализуется рекуррентным алгоритмом, который в стационарном (установившемся) режиме фильтрации имеет вид:

$$x_i^* = F \cdot x_{i-1}^* + A \cdot (y_i - F \cdot x_{i-1}^*). \quad (3.12)$$

Нетрудно установить, что импульсная характеристика этого фильтра имеет экспоненциальный вид:

$$a_i = A \cdot [(1 - A) \cdot F]^i, \quad i = 0, 1, 2, \dots, \quad (3.13)$$

где A - параметр, лежащий в пределах $0 < A < 1$, и получивший название коэффициента усиления фильтра Калмана. Первое слагаемое в алгоритме (3.12) определяет вклад в оценку сигнала на текущем i -м шаге фильтрации, вносимый оценкой предыдущего шага, и называется одношаговым прогнозом. Второе учитывает влияние текущего наблюдения y_i и называется новой информацией. Коэффициент усиления A определяет чувствительность фильтра к этой новой информации. При высоком уровне шума параметр A приближается к нулю. При этом, кроме общего снижения ИХ, увеличивается параметр $B = 1 - A$, приближаясь к единице. Это означает удлинение импульсной характеристики и, следовательно, сужение полосы пропускаемых фильтром частот. Очевидно, эти свойства ИХ способствуют эффективной фильтрации шума. При снижении шума, наоборот, A стремится к единице, B - к нулю, что соответствует расширению полосы частот до бесконечности. Здесь уместно отметить, что фильтрация не только ослабляет действие шума, но, к сожалению, и вносит так называемые динамические искажения в полезный сигнал. Механизм этих искажений очень прост и заключается в неравной передаче на выход фильтра различных спектральных компонент сигнала. Фильтр Калмана "ведет себя" вполне разумно, когда при исчезновении шума на входе расширяет полосу пропускаемых частот до бесконечности, поскольку именно при этом условии исчезают динамические ошибки фильтрации.

Отталкиваясь от (3.13) как от одномерного аналога, будем находить двумерную БИХ для каузальной фильтрации изображений от некоррелированного шума в виде двумерной экспоненты:

$$a_{i,j} = A^2 \cdot (B \cdot F)^{i+j}, \quad B = 1 - A, \quad i, j = 0, 1, 2, \dots \quad (3.14)$$

Здесь, как и в случае одномерного фильтра, F - коэффициент одношаговой корреляции изображения по строке и по столбцу, которые будем здесь считать одинаковыми. Для определения параметра A (или B), остающегося единственным неизвестным параметром двумерного фильтра, воспользуемся уравнением Винера-Хопфа в форме (3.4), переписав его в виде:

$$E \left\{ (x_{i,j} - \sum_{i_1, j_1 \in S} a_{i_1, j_1} \cdot y_{i-i_1, j-j_1}) \cdot y_{i-k, j-l} \right\} = 0, \quad k, l \in S.$$

Замечая, что выражение в круглых скобках является ошибкой фильтрации, представим эту формулу в виде:

$$E \left\{ \varepsilon_{i,j} \cdot y_{i-k, j-l} \right\} = 0, \quad k, l \in S. \quad (3.15)$$

Смысл данного выражения состоит в том, что при оптимальной линейной фильтрации ошибка ортогональна всем элементам наблюдаемых данных, используемых при фильтрации. Но тогда нетрудно убедиться и в ортогональности ошибки и результата фильтрации (получаемой оценки)

$$E \left\{ \varepsilon_{i,j} \cdot x_{i,j}^* \right\} = 0, \quad (3.16)$$

для чего достаточно вычислить левую часть этого выражения с учетом (3.2) и (3.15).

Для дальнейшего необходимо воспользоваться в (3.16) принятым представлением импульсной характеристики (3.14), в результате данное соотношение превращается в уравнение относительно искомого параметра A . В него входят корреляционная функция сигнала и взаимно-корреляционная функция сигнала и наблюдаемых данных. Поэтому необходимо конкретизировать корреляционную функцию сигнала, в качестве которой воспользуемся биэкспоненциальным представлением:

$$B_x(i, j) = D_x \cdot F^{|i|+|j|}. \quad (3.17)$$

С учетом этого, считая, что кадр имеет бесконечные размеры (это позволяет принять бесконечными соответствующие пределы суммирования в (3.2)), можно получить следующее алгебраическое уравнение

$$1 - \left[1 - \frac{B \cdot (1 - F^2)}{1 - B^2 \cdot F^2} \right]^2 - \frac{(1 - B)^2}{q^2} \cdot \frac{(1 - B \cdot F^2)^2}{(1 - B^2 \cdot F^2)^2} = 0 \quad (3.18)$$

относительно параметра B , численное решение которого не представляет проблемы. Анализируя (3.18), легко заметить, что при $q^2 \rightarrow \infty$ это уравнение удовлетворяется при $B = 0$, а при $q^2 \rightarrow 0$ его решением является $B = 1$. Эти предельные значения можно интерпретировать как изменения частотно-полосовых характеристик двумерного фильтра, аналогичные поведению параметров фильтра Калмана в подобных предельных ситуациях.

Подставив в (3.7) выражения ИХ (3.14) и корреляционной функции (3.17), можно получить следующую формулу для среднего квадрата ошибок фильтрации:

$$D_\varepsilon = D_x \cdot \left[1 - \frac{A^2}{(1 - B \cdot F^2)^2} \right].$$

Подставив далее выражение ИХ (3.14) в (3.2), можно привести выражение отклика фильтра к виду :

$$x_{i,j}^* = B \cdot F \cdot (x_{i-1,j}^* + x_{i,j-1}^*) - (B \cdot F)^2 \cdot x_{i-1,j-1}^* + A \cdot y_{i,j}. \quad (3.19)$$

Рекуррентный характер алгоритма (3.19) является важным положительным качеством рассматриваемого фильтра. Как следует из (3.19), его работа требует выполнения на каждом шаге обработки всего трех операций умножения и трех суммирования, причем структура алгоритма универсальна и, в частности, не зависит от отношения сигнал/шум. Для сравнения, масочный фильтр с размером окрестности 3×3 элементов требует выполнения на каждом шаге при общем виде КИХ девяти умножений и восьми суммирований. Таким образом, по количеству операций рекуррентный фильтр выигрывает у простейшего масочного практически в три раза. Очевидно, что попытка улучшить качество фильтрации масочным фильтром за счет увеличения размера применяемой окрестности приводит к увеличению числа операций и дальнейшему увеличению его проигрыша по этой характеристики.

При фильтрации реальных изображений ограниченного размера возникает граничная проблема получения оценок в точках нулевой строки и

нулевого столбца. Естественным решением является использование здесь обычной (одномерной) калмановской фильтрации.

Пример применения описанного двумерного фильтра показан на рис. 3.3, где представлен результат эксперимента с тем же портретом и при том же отношении сигнал/шум -5 дБ, что и при испытании масочного фильтра.

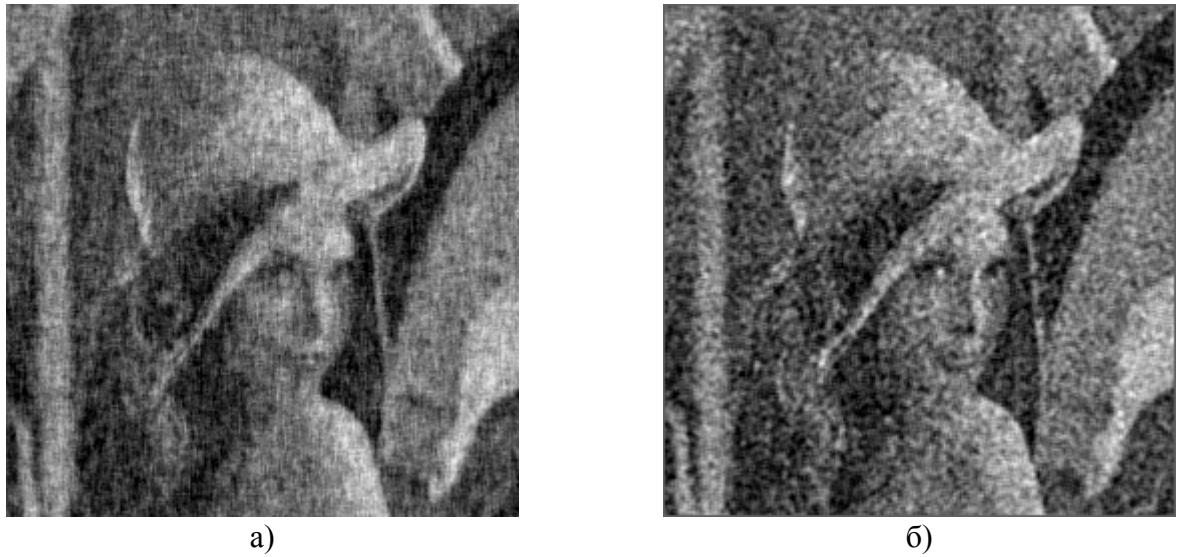


Рис. 3.3. Пример двумерной рекуррентной фильтрации

Поэтому здесь не приводится показанное на рис.3.2.а входное изображение с шумом. Результат двумерной рекуррентной фильтрации представлен на рис.3.3.а, а на рис.3.3.б для сравнения повторен результат оптимальной масочной фильтрации (рис.3.2.б). Визуальная оценка говорит в пользу двумерного рекуррентного фильтра, поскольку уровень остаточного шума на рис.3.3.а ниже. Сравнение по среднему квадрату ошибок совпадает с субъективной оценкой: величина δ_e^2 при масочной фильтрации составляет, как говорилось ранее, 0.309, а при двумерной рекуррентной - 0.29. Различие заметно усиливается при более высоком уровне шума. Так, при отношении сигнал/шум -10 дБ имеем соответственно δ_e^2 равным 0.57 и 0.43.

Необходимо отметить, однако, следующее. Вместе с уменьшением уровня шума при двумерной рекуррентной фильтрации наблюдается более значительная потеря резкости обработанного изображения. Это является проявлением упоминавшихся выше динамических искажений, более сильных при бесконечной импульсной характеристики, чем при конечной.

Во-вторых, рассмотренный двумерный фильтр не является абсолютно оптимальным, поскольку его структура определена волевым решением при выборе ИХ в виде (3.14). Поэтому и получаемое при его помощи ослабление шума не является предельным.

3.4. Применение фильтра Винера для некаузальной двумерной фильтрации

Потенциально наилучшие результаты обработки изображения, в частности, результаты фильтрации, достигаются при использовании некаузального принципа, поскольку этот принцип основан на применении абсолютно всех исходных данных при обработке каждой точки кадра. Понятно, что при рациональном использовании этих данных получаемый эффект максимальен. Одним из известных вариантов линейной некаузальной фильтрации изображений является фильтр Винера. Его применение связано с предположением о стационарности изображения. Поскольку наличие краев изображения служит нарушением стационарности, то винеровская фильтрация, не является строго оптимальной. Однако при размерах кадра, значительно превышающих интервал корреляции изображения, влияние границ является малым. Эти соображения служат важным стимулом к применению винеровской фильтрации для борьбы с шумом.

Технически фильтр Винера реализуется при помощи дискретного преобразования Фурье в частотной области. Поэтому, прежде чем рассматривать уравнение Винера-Хопфа, которое остается методологической основой фильтрации помех и в этом случае, нам необходимо познакомиться с двумерным дискретным преобразованием Фурье (ДПФ), некоторыми его свойствами и принципами линейной фильтрации на основе ДПФ.

3.4.1. Двумерное дискретное преобразование Фурье

Обозначим через

$$x(i_1, i_2) = x_{i_1, i_2}, \quad i_1 = \overline{0, I_1 - 1}, \quad i_2 = \overline{0, I_2 - 1} \quad (3.20)$$

двумерное поле (двумерный сигнал), описывающее дискретное изображение размера I_1 строк и I_2 столбцов. Вне указанных границ этот сигнал не определен. Выполним *периодическое продолжение* данного финитного сигнала, введя двумерный периодический сигнал

$$\tilde{x}_{i_1, i_2} = \sum_{m_1=-\infty}^{\infty} \sum_{m_2=-\infty}^{\infty} x_{i_1 - m_1 I_1, i_2 - m_2 I_2}. \quad (3.21)$$

Если сигнал x_{i_1, i_2} существует только внутри прямоугольника $R_{I_1, I_2} = \{(i_1, i_2): 0 \leq i_1 \leq I_1 - 1, 0 \leq i_2 \leq I_2 - 1\}$ со сторонами $I_1 \times I_2$ элементов

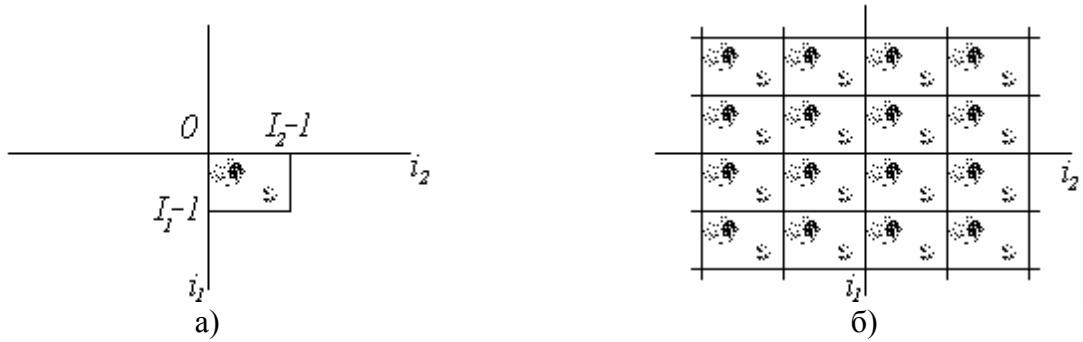


Рис. 3.4. Реальное (а) и периодически продолженное (б) изображения

(рис. 3.4.а), то сигнал \tilde{x}_{i_1, i_2} определен на всей плоскости (i_1, i_2) и является на ней *прямоугольно-периодическим* (рис. 3.4.б).

Любой периодический сигнал может быть представлен в виде *ряда Фурье*, но, в отличие от одномерных сигналов, двумерные описываются двумерным рядом Фурье, имеющим вид:

$$\tilde{x}_{i_1, i_2} = \frac{1}{I_1 I_2} \sum_{k_1=0}^{I_1-1} \sum_{k_2=0}^{I_2-1} \tilde{X}_{k_1, k_2} \exp\left(j \frac{2\pi}{I_1} i_1 k_1 + j \frac{2\pi}{I_2} i_2 k_2\right), \quad j = \sqrt{-1}. \quad (3.22)$$

Базисные функции этого двумерного представления - двумерные комплексные экспоненты (иногда называемые комплексными синусоидами)

$$\begin{aligned} \varphi_{k_1, k_2}(i_1, i_2) &= \frac{1}{I_1 I_2} W_1^{i_1 k_1} W_2^{i_2 k_2}, \\ W_1 &= \exp\left(j \frac{2\pi}{I_1}\right), \quad W_2 = \exp\left(j \frac{2\pi}{I_2}\right), \end{aligned} \quad (3.23)$$

имеющие, как и сигнал $\tilde{x}(\cdot)$, прямоугольную периодичность с тем же периодом $I_1 \times I_2$. Здесь (k_1, k_2) - двумерный номер базисной функции, а величины $\frac{2\pi k_1(2)}{I_1(2)}$ имеют смысл пространственных частот. Иногда пространственными

частотами называют целочисленные величины k_1 и k_2 .

Коэффициенты Фурье \tilde{X}_{k_1, k_2} ряда (3.22) образуют *двумерный частотный спектр* сигнала \tilde{x}_{i_1, i_2} и определяются формулой *прямого преобразования Фурье*:

$$\tilde{X}_{k_1, k_2} = \sum_{i_1=0}^{I_1-1} \sum_{i_2=0}^{I_2-1} \tilde{x}_{i_1, i_2} W_1^{i_1 k_1} W_2^{i_2 k_2} \quad (3.24)$$

Выражение (3.22), восстанавливающее сигнал \tilde{x}_{i_1, i_2} по его спектру \tilde{X}_{k_1, k_2} , является *обратным преобразованием Фурье*. В справедливости преобразований (3.22) и (3.24), называемых *двумерным ДПФ*, можно убедиться, подставив (3.24) в (3.22) и приведя правую часть полученного равенства к значению левой, т.е. к \tilde{x}_{i_1, i_2} .

Заметим, что для *точного представления* дискретного сигнала \tilde{x}_{i_1, i_2} с двумерным периодом $I_1 \times I_2$ элементов согласно формулам БПФ достаточно *конечного* числа базисных функций (3.23) - ряд (3.22) является *конечным*. Это и понятно, поскольку сам представляемый сигнал содержит в одном периоде конечное число точек, т.е. имеет конечное число степеней свободы. Ясно, что число степеней свободы в спектре не может отличаться от числа степеней свободы в самом сигнале.

Остановимся на наиболее существенных свойствах двумерного дискретного спектра Фурье. Вычислим спектральные коэффициенты (3.24) в частотных точках $k_1 + m_1 I_1, k_2 + m_2 I_2$:

$$\begin{aligned} \tilde{X}_{k_1+m_1 I_1, k_2+m_2 I_2} &= \sum_{i_1=0}^{I_1-1} \sum_{i_2=0}^{I_2-1} \tilde{x}_{i_1, i_2} W_1^{i_1 k_1} W_2^{i_2 k_2} \times \\ &\quad \times \exp(-j2\pi i_1 m_1 - j2\pi i_2 m_2) . \end{aligned} \quad (3.25)$$

Поскольку при любых целых значениях m_1 и m_2 последний множитель в полученном выражении равен единице, то отсюда имеем равенство:

$$\tilde{X}_{k_1+m_1 I_1, k_2+m_2 I_2} = \tilde{X}_{k_1, k_2},$$

означающее *прямоугольную периодичность двумерного ДПФ*. Следовательно, картина двумерного ДПФ подобна картине двумерного периодически продолженного сигнала, качественно показанной на рис. 3.4.б (если на ней пространственные координаты i_1, i_2 заменить частотными k_1, k_2). Однако необходимо иметь в виду, что спектральные коэффициенты \tilde{X}_{k_1, k_2} , как это следует из (3.24), являются *комплексными числами*, в том числе и при вещественном сигнале \tilde{x}_{i_1, i_2} . Но тогда возникает вопрос. Общее количество спектральных компонент, как установлено, равно $I_1 I_2$. Комплексное число эквивалентно паре вещественных чисел - действительной и мнимой частям при алгебраическом или модулю и фазе при экспоненциальном представлении. Следовательно, полный спектр описывается $2I_1 I_2$ вещественными числами, что вдвое превышает размерность самого сигнала \tilde{x}_{i_1, i_2} . В этом, на первый

взгляд, содержится противоречие. Оно находит свое разъяснение при дальнейшем изучении свойств двумерного ДПФ.

Преобразуем соотношение (3.25) следующим образом. Во-первых, вместо частот k_1, k_2 подставим частоты $-k_1, -k_2$. Во-вторых, выполним комплексное сопряжение обеих частей, что не нарушит равенства. В результате нетрудно получить выражение:

$$\tilde{X}_{k_1, k_2} = \tilde{X}_{m_1 I_1 - k_1, m_2 I_2 - k_2}^*,$$

которым устанавливается однозначная связь между спектральными коэффициентами в двух различных точках спектрального прямоугольника $R_{I_1, I_2} = \{(k_1, k_2): 0 \leq k_1 \leq I_1 - 1, 0 \leq k_2 \leq I_2 - 1\}$. Полученным соотношением и снимается противоречие, поскольку количество независимых спектральных коэффициентов уменьшается благодаря данной спектральной симметрии в два раза. Согласно установленному свойству, спектрально-сопряженной зависимостью связаны между собой спектральные коэффициенты, принадлежащие левому верхнему и правому нижнему углам спектрального прямоугольника R_{I_1, I_2} . Аналогично также связаны между собой коэффициенты Фурье из правого верхнего и левого нижнего участков спектрального прямоугольника R_{I_1, I_2} .

В заключение данного пункта укажем, что при практическом применении двумерного ДПФ - как прямого, так и обратного, совсем не требуется оперировать периодическими сигналами и спектрами, как это предполагается, казалось бы, преобразованиями (3.22) и (3.24). От этой необходимости избавляют сами соотношения (3.22) и (3.24). В самом деле, прямое преобразование Фурье (3.24) содержит в правой части значения периодически продолженного сигнала \tilde{x}_{i_1, i_2} лишь в пределах одного “главного” прямоугольника R_{I_1, I_2} . Но в этих пределах исходный x_{i_1, i_2} и периодически продолженный \tilde{x}_{i_1, i_2} сигналы полностью совпадают, что дает возможность использовать в формуле (3.24) исходный сигнал x_{i_1, i_2} . Аналогичные пояснения можно сделать и относительно обратного преобразования (3.22), откуда следует, что практически в процессе вычислений оперировать следует “основным” участком спектра, относящимся к спектральной области R_{I_1, I_2} .

Из сделанных пояснений, имеющих лишь исключительно вычислительное значение, не следует делать вывода об искусственности и ненужности рассмотренных математических моделей периодических полей. При обработке изображений возникают многочисленные задачи, правильное толкование и решение которых возможно только на основе этих математических интерпретаций. Одной из таких важнейших задач является

цифровая двумерная фильтрация в спектральной области, осуществление которой связано с выполнением так называемой *циклической свертки*.

3.4.2. Циклическая свертка

Достоинством обычного преобразования Фурье является то, что с его помощью очень просто выполняется стационарная фильтрация сигналов. Как известно, для этого требуется получить частотный спектр сигнала и частотный коэффициент передачи фильтра. Затем, перемножив их, определить частотный спектр выходного сигнала, а выполнив обратное преобразование Фурье - найти сам выходной сигнал. Возможно ли применение такой технологии для выполнения двумерной цифровой фильтрации ? Убедимся, что возможно, но с некоторыми оговорками.

Двумерный стационарный (инвариантный к сдвигу) фильтр характеризуют двумерной импульсной характеристикой a_{i_1, i_2} , позволяющей определить выходной сигнал при помощи двумерной свертки:

$$y_{i_1, i_2} = \sum_{r_1=0}^{I_1-1} \sum_{r_2=0}^{I_2-1} x_{r_1, r_2} \cdot a_{i_1-r_1, i_2-r_2}. \quad (3.26)$$

Это уравнение *обычной* свертки, в нем фигурируют обычные, непериодические сигналы и непериодическая ИХ фильтра. При цифровой обработке в ЭВМ не существует частотных спектров, соответствующих таким сигналам и ИХ. Для описания сигнала в частотной области привлекается, как установлено выше, периодически продолженный сигнал \tilde{x}_{i_1, i_2} , которому соответствует дискретный спектр \tilde{X}_{k_1, k_2} . По аналогии с (3.21) вводится и периодически продолженная ИХ:

$$\tilde{a}_{i_1, i_2} = \sum_{m_1=-\infty}^{\infty} \sum_{m_2=-\infty}^{\infty} a_{i_1-m_1 I_1, i_2-m_2 I_2},$$

двумерное ДПФ которой \tilde{A}_{k_1, k_2} имеет смысл частотного коэффициента передачи цифрового фильтра. Покажем, что перемножая \tilde{X}_{k_1, k_2} и \tilde{A}_{k_1, k_2} , мы находим спектр сигнала, определяемого *циклической сверткой*. Циклическая свертка отличается от обычной свертки (3.26) тем, что вместо функций x_{i_1, i_2} и a_{i_1, i_2} в ней представлены периодически продолженные функции \tilde{x}_{i_1, i_2} и \tilde{a}_{i_1, i_2} . Нетрудно установить, что при этом сигнал на выходе

$$\tilde{y}_{i_1, i_2} = \sum_{r_1=0}^{I_1-1} \sum_{r_2=0}^{I_2-1} \tilde{x}_{r_1, r_2} \cdot \tilde{a}_{i_1-r_1, i_2-r_2}. \quad (3.27)$$

также является периодическим. Покажем, что именно его спектр \tilde{Y}_{k_1, k_2} определяется выражением:

$$\tilde{Y}_{k_1, k_2} = \tilde{X}_{k_1, k_2} \cdot \tilde{A}_{k_1, k_2}. \quad (3.28)$$

Умножим для этого левую и правую части (3.27) на $W_1^{-i_1 k_1} W_2^{-i_2 k_2}$ и просуммируем по i_1 и i_2 в пределах области R_{I_1, I_2} . В левой части в результате имеем спектр \tilde{Y}_{k_1, k_2} . Преобразуем правую часть, предварительно умножив ее на величину $W_1^{r_1 k_1} W_2^{r_2 k_2} W_1^{-r_1 k_1} W_2^{-r_2 k_2}$, тождественно равную единице:

$$\tilde{Y}_{k_1, k_2} = \sum_{r_1=0}^{I_1-1} \sum_{r_2=0}^{I_2-1} \tilde{x}_{r_1, r_2} W_1^{-r_1 k_1} W_2^{-r_2 k_2} \sum_{r_1=0}^{I_1-1} \sum_{r_2=0}^{I_2-1} \tilde{a}_{i_1-r_1, i_2-r_2} W_1^{-(i_1-r_1) k_1} W_2^{-(i_2-r_2) k_2}.$$

Выполняя затем во внутренней сумме замену переменных $m_1 = i_1 - r_1$, $m_2 = i_2 - r_2$, учитывая, что при суммировании периодической функции в пределах периода границы суммирования могут быть синхронно сдвинуты на любую величину и оставляя их в этой связи неизменными, приходим к (3.28).

Таким образом установлено, что именно циклической свертке (3.27) соответствует удобный при практической реализации частотный метод определения выходного сигнала. Необходимо иметь в виду, что та фильтрация, к осуществлению которой обычно стремятся при решении конкретных задач, описывается обычной сверткой (3.26), а не циклической (3.27). Поэтому остается выяснить, в каком соотношении находятся эти две вычислительные процедуры. Не нарушая общности, для большей простоты и наглядности рассмотрим этот вопрос на примере одномерной фильтрации. Сигнал на выходе одномерного фильтра определяется сверткой

$$y_i = \sum_{r=0}^{I-1} x_r \cdot a_{i-r}, \quad (3.29)$$

$$\tilde{y}_i = \sum_{r=0}^{I-1} \tilde{x}_r \cdot \tilde{a}_{i-r} \quad (3.30)$$

в обычном и циклическом вариантах соответственно. Рис. 3.5 поясняет процесс

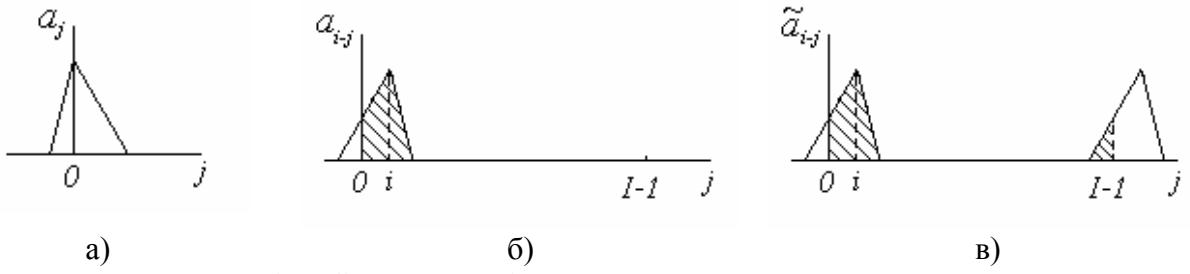


Рис. 3.5. Сравнение обычной и циклической свертки

вычисления выходного сигнала в обоих случаях. На рис. 3.5.а показана импульсная характеристика произвольного вида, соответствующая некаузальному фильтру (т.к. $a(j) \neq 0$ при $j < 0$). Рис. 3.5.б иллюстрирует образование суммы, вычисляемой при помощи обычной свертки (3.29), а рис. 3.5.в при помощи циклической (3.30). На рисунках штриховкой показаны области суммирования, выполняемого в соответствии с выражениями (3.29) и (3.30). Рисунки отражают определение реакции фильтра в точке i , расположенной вблизи границы рабочей области. В случае циклической свертки область суммирования является двухсвязной из-за периодичности ИХ, что приводит к различию результатов фильтрации. Очевидно, что эффекты, вызванные периодичностью, отсутствуют для точек, удаленных от границ. Поэтому для внутренних точек области $R_I = \{i: 0 \leq i \leq I - 1\}$, удаленность которых от границ превышает длину импульсной характеристики, результаты обычной и циклической сверток совпадают. Различия наблюдаются лишь для точек, примыкающих к границе. Если размеры этой приграничной области относительно невелики, то часто различиями пренебрегают. В тех случаях, когда граничные эффекты недопустимы, проблема может быть разрешена при помощи искусственного удлинения области R_I добавлением к ней на обоих концах такого количества нулевых элементов, при котором эффект циклическости проявляться не будет.

В заключение отметим, что характер различий обычной и циклической сверток при фильтрации двумерных цифровых сигналов остается таким же, как для одномерных сигналов. Аналогичны и меры, исключающие нежелательные эффекты циклической свертки.

3.4.3. Решение уравнения Винера-Хопфа в циклическом приближении

Вернемся к задаче некаузальной фильтрации шума на изображении. Оптимальный линейный фильтр определяется и в этом случае уравнением Винера-Хопфа (3.6), в котором для начала область существования S импульсной характеристики примем неограниченной. В результате имеем:

$$\left\{ \begin{array}{l} B_{xy}(k, l) = \sum_{i_1=-\infty}^{\infty} \sum_{j_1=-\infty}^{\infty} a_{i_1, j_1} B_y(k - i_1, l - j_1), \\ k, l \in (-\infty, \infty) \end{array} \right. . \quad (3.31)$$

Дискретный винеровский фильтр удается легко найти в циклическом приближении. Для этого требуется вместо реальных функций $B_{xy}(\cdot)$, $B_y(\cdot)$ и a_{i_1, j_1} подставить в уравнение (3.31) соответствующие периодически продолженные функции $\tilde{B}_{xy}(\cdot)$, $\tilde{B}_y(\cdot)$ и \tilde{a}_{i_1, j_1} , имеющие двумерный период R_{I_1, I_2} . При этом область определения ИХ также сужается до размеров прямоугольника R_{I_1, I_2} . Поэтому (3.31) принимает вид:

$$\left\{ \begin{array}{l} \tilde{B}_{xy}(k, l) = \sum_{i_1=0}^{I_1-1} \sum_{j_1=0}^{I_2-1} \tilde{a}_{i_1, j_1} \tilde{B}_y(k - i_1, l - j_1), \\ k, l \in R_{I_1, I_2} \end{array} \right. . \quad (3.32)$$

Периодичность функций, входящих в уравнение (3.32), позволяет применить к его обеим частям двумерное ДПФ, подобно тому, как это было сделано выше применительно к уравнению (3.27). В результате получаем:

$$\tilde{G}_{x,y}(k_1, k_2) = \tilde{A}_{k_1, k_2} \cdot \tilde{G}_y(k_1, k_2). \quad (3.33)$$

В этом выражении $\tilde{G}_{x,y}(k_1, k_2)$, $\tilde{G}_y(k_1, k_2)$ - спектральные плотности мощности, представляющие собой ДПФ соответствующих корреляционных функций, а \tilde{A}_{k_1, k_2} - ДПФ искомой импульсной характеристики, являющееся, таким образом, частотным коэффициентом передачи фильтра Винера. Все функции, входящие в (3.33), являются прямоугольно-периодичными с двумерным периодом R_{I_1, I_2} . Основное достижение, вызванное применением ДПФ, состоит в преобразовании сложного для решения уравнения Винера-Хопфа (3.32) в простейшее алгебраическое уравнение (3.33), решение которого, правда не для импульсной, а для частотной характеристики, имеет вид:

$$\tilde{A}_{k_1, k_2} = \frac{\tilde{G}_{x,y}(k_1, k_2)}{\tilde{G}_y(k_1, k_2)}. \quad (3.34)$$

Найденное решение дает эффективный способ осуществления оптимальной линейной фильтрации изображения. Его выполнение требует знания

спектральных плотностей мощности и опирается на применение к обрабатываемому изображению дискретного преобразования Фурье.

Не следует, однако, забывать, что переход от уравнения (3.31), определяющего истинно оптимальную характеристику фильтра, к уравнению (3.32), позволяющему найти ее периодически продолженный аналог, был выполнен без достаточного обоснования. Поэтому ничего нельзя пока сказать о том, в какой степени найденное решение близко к истинно оптимальному. Для ответа на этот вопрос рассмотрим снова для простоты одномерные аналоги уравнений (3.31) и (3.32), имеющие вид:

$$\begin{cases} B_{xy}(k) = \sum_{i=-\infty}^{\infty} a_i B_y(k-i), \\ k \in (-\infty, \infty) \end{cases} . \quad (3.35)$$

$$\begin{cases} \tilde{B}_{xy}(k) = \sum_{i=0}^I \tilde{a}_i \tilde{B}_y(k-i), \\ k \in [0, I-1] \end{cases} . \quad (3.36)$$

Рис. 3.6 иллюстрирует формирование сумм, входящих в правые части этих равенств при некотором произвольном значении сдвига k и достаточно большом значении интервала наблюдения I . Здесь показаны некаузальная ИХ и корреляционная функция входного сигнала, а штриховкой условно отмечена область суммирования. Из сравнения рис. 3.6.а, соответствующего (3.35), и рис. 3.6.б, соответствующего (3.36), видно, что, хотя во втором случае область суммирования является двухсвязной, это не приводит к различию результатов

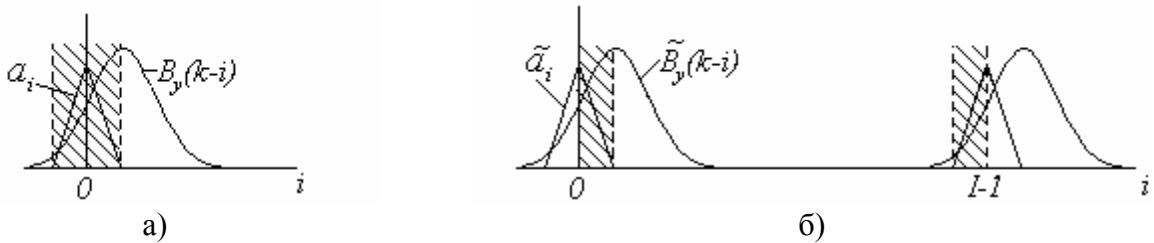


Рис. 3.6. Сравнение обычного и циклического уравнений Винера-Хопфа

суммирования. При большом значении длины интервала I соседние зоны на рис. 3.6.б не перекрываются, благодаря чему искажения результатов не происходит. Следовательно, уравнения (3.35) и (3.36) являются эквивалентными.

Если же интервал I будет соизмерим с размахом корреляционной функции $B_y(\cdot)$, то произойдет наложение соседних областей периодической картины, что в итоге приведет к изменению значений функций, стоящих под знаком суммы в (3.36), и исказит уравнение Винера-Хопфа. Таким образом,

условие применимости фильтра Винера, определяемого соотношением (3.34), состоит в его использовании для обработки изображений, имеющих достаточно большие размеры. Напомним также, что в данном пособии обсуждается уравнение Винера-Хопфа для стационарных сигналов и изображений. Поэтому вблизи границ обрабатываемого кадра, где само их существование приводит к нарушению этого условия, обработка отклоняется от оптимальной.

На рис. 3.7 приведен пример работы фильтра Винера. Как и ранее эксперимент выполнен при отношении сигнал/шум $q^2 = -5$ дБ. Относительный средний квадрат ошибок фильтрации δ_e^2 составляет в этом эксперименте величину 0,167, что является наилучшим показателем среди всех рассмотренных методов линейной фильтрации (напомним, что при масочной фильтрации выше было получено $\delta_e^2 = 0.309$, а при рекуррентной $\delta_e^2 = 0.29$). О наименьшем уровне остаточного шума на изображении говорит и визуальная оценка результата. Хотя нельзя не отметить, что это достигается ценой большей, чем при других методах, дефокусировки изображения. В этом проявляется общее диалектическое противоречие между борьбой с помехами и

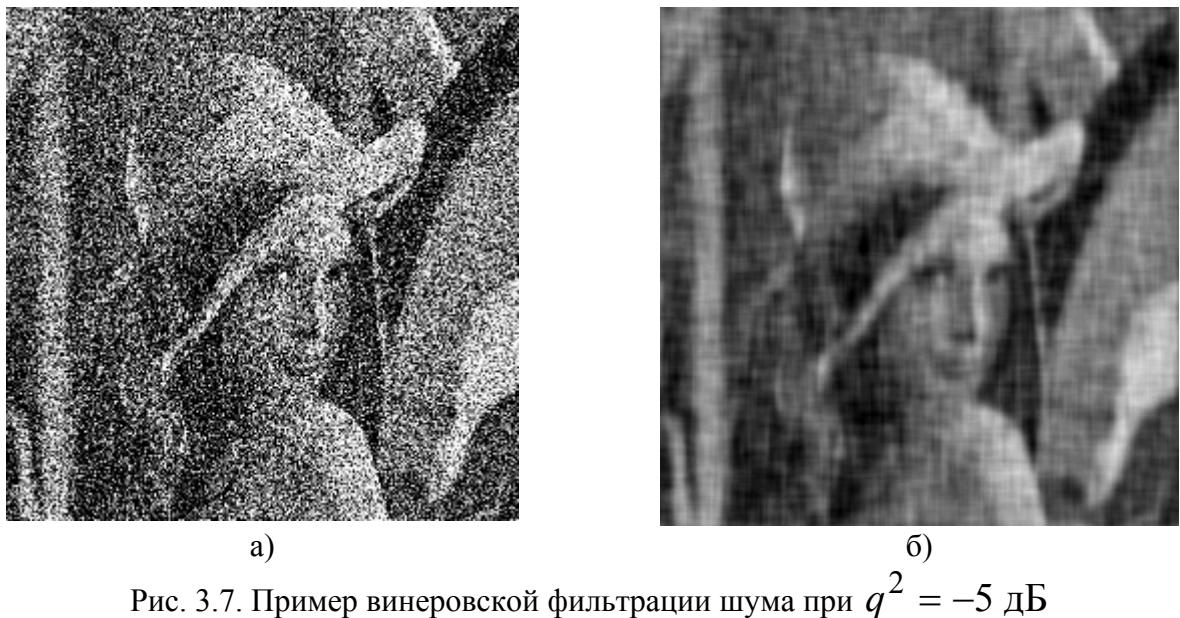


Рис. 3.7. Пример винеровской фильтрации шума при $q^2 = -5$ дБ

динамическимиискажениями обрабатываемого изображения, свойственное, как отмечалось и ранее, всем методам фильтрации.

Проведение обработки изображений при помощи фильтра Винера требует использования спектральной плотности мощности изображения. Существуют различные способы получения необходимой информации. Один из них основан на предварительном измерении требуемых характеристик по реальному изображению. Полученные при этом спектральные плотности вводятся в ЭВМ в виде таблиц, позволяя задать коэффициент передачи в численном виде. Другой способ, примененный и в представленном эксперименте, состоит в использовании некоторой математической модели изображения, вид

спектрально-корреляционных характеристик которой известен. В этом случае реальное изображение используется для измерения только отдельных параметров, входящих в используемую математическую модель. При проведении эксперимента, описанного выше, в частности, использовалась модель изображения в виде гауссовского двумерного поля с корреляционной функцией (3.17), а измерялись коэффициент одношаговой корреляции F и дисперсия D_x .

Анализ эффективности метода будет неполным, если не сделать оценки вычислительной эффективности реализующей его процедуры. Для вычисления ДПФ разработаны эффективные вычислительные методы, воплощенные в процедурах быстрого преобразования Фурье (БПФ). Количество комплексных умножений, составляющих основную трудоемкость двумерного БПФ, оценивают числом $I_1 I_2 \log_2 I_1 I_2$ [3.2]. Поскольку полный цикл обработки предполагает выполнение прямого и обратного БПФ, то это число следует удвоить. По отношению к одному элементу кадра число умножений, таким образом, составляет $2 \log_2 I_1 I_2$. При $I_1 = I_2 = 256$ число умножений в каждой точке кадра равно 32. Для сравнения напомним, что, например, рекуррентный двумерный фильтр, описанный выше, реализуется всего тремя вещественными умножениями в каждой точке кадра (при различных значениях одношагового коэффициента корреляции изображения по строкам и по столбцам - четырьмя умножениями).

3.5. Байесовская фильтрация изображений

При всех рассмотренных ранее методах фильтрации с самого начала закладывалось отыскание фильтра в классе линейных систем. Отсюда следует, что могут существовать нелинейные процедуры, обладающие более высокими качественными характеристиками, чем рассмотренные выше. Для их отыскания необходим более общий подход к фильтрации, чем тот, который опирается на решение уравнение Винера-Хопфа. Общепринятая достаточно универсальная идеология фильтрации использует байесовский принцип. Ее применение позволяет, по крайней мере теоретически, создавать как линейные, так и нелинейные алгоритмы фильтрации. Кроме того, этот принцип помогает выяснить, при каких условиях линейные процедуры фильтрации приводят к наивысшему качеству обработки и, следовательно, являются абсолютно оптимальными.

Отметим, однако, с самого начала основные недостатки байесовской фильтрации изображений. Первый является общим для байесовских методов вообще и заключается в очень высоких требованиях к объему и характеру данных, содержащихся в математических моделях сигналов и помех, удовлетворить которым на практике удается далеко не всегда. Второй связан со спецификой изображений как двумерных сигналов, что приводит к колоссальным вычислительным трудностям при попытке прямого

использования этого подхода. Последнее делает актуальной разработку таких методов, которые способны преодолеть данное ограничение. Достичь этого обычно удается ценой определенной потери качества обработки. Поэтому дело сводится к поиску таких методов, при которых потеря качества являлась бы приемлемой.

3.5.1. Сущность байесовской фильтрации

Полагаем, что на входе фильтра действует сигнал

$$y_{i_1, i_2} = f(x_{i_1, i_2}, n_{i_1, i_2}), \quad i_1 = \overline{0, I_1 - 1}, \quad i_2 = \overline{0, I_2 - 1}, \quad (3.37)$$

где x_{i_1, i_2} и n_{i_1, i_2} - полезный сигнал и помеха, а $f(\cdot)$ - функция, описывающая их взаимодействие. При байесовском методе считается, что сигнал и помеха - случайные процессы (случайные двумерные поля) с известными законами распределения вероятностей. Пусть X - вектор, элементы которого - все $I_1 I_2$ отсчетов, образующих кадр изображения, а $w_x(X)$ -их совместное распределение. Примем для простоты, что помеха и сигнал независимы, а распределение вектора помехи N равно $w_n(N)$. Воспользовавшись формулой Байеса, запишем *апостериорное распределение вероятностей* (APB) $w_{x|y}(X|Y)$:

$$w_{x|y}(X|Y) = \frac{w_x(X)w_{y|x}(Y|X)}{w_y(Y)}, \quad (3.38)$$

куда входит распределение $w_y(Y)$ наблюдаемых данных и условное распределение $w_{y|x}(Y|X)$ - называемое *функцией правдоподобия*. Смысл выражения (3.38) заключается в том, что оно дает возможность вычислить в устройстве обработки распределение вероятностей полезного сигнала, располагая входными данными Y и опираясь на вероятностную модель как самого полезного сигнала, так и наблюдаемых данных. APB является аккумулятором всех доступных сведений о полезном сигнале, которые содержатся в Y , а (3.38) указывает способ извлечения этих сведений.

Поскольку потребителя информации обычно интересует точечное значение сигнала X , то для его образования прибегают к вычислению либо математического ожидания APB, либо его координаты, обращающей это распределение в максимум. В математической статистике доказано, что эти способы получения результатов фильтрации соответствуют различным содержательным требованиям, предъявляемым к получаемым результатам [3.3].

Оперировать векторными величинами, входящими в (3.38), практически невозможно из-за громадной размерности векторов X и Y . Если, например, обрабатываемый кадр имеет размеры $I_1 = I_2 = 256$, то размерность этих векторов равна $256^2 = 65536$. Предположим, что изображение является простейшим с бинарными значениями элементов x_0 и x_1 . Общее число всевозможных изображений, имеющих всего две градации яркости, составляет $2^{65536} > 10^{19000}$. Задачей байесовского фильтра является вычисление распределения вероятностей $w_{x|y}(X|Y)$, которое можно представить себе в

данном случае в виде таблицы с размером, превышающим 10^{19000} . Явная нереальность этой задачи заставляет искать такие методы описания сигналов, которые приводили бы к резкому, качественному ее упрощению. В данном направлении предпринимаются усилия, разрабатываются различные подходы [3.4-3.6], но, к сожалению, универсальных эффективных методов двумерной байесовской обработки изображений, основанных на использовании всех данных Y , в настоящее время не найдено.

Отмеченная сложность байесовских процедур свойственна и фильтрации одномерных сигналов. Вместе с тем, в области одномерной фильтрации были получены блестящие решения проблемы, основанные на использовании марковских моделей сигналов и помех. В указанных работах [3.4-3.6] предпринимались разнообразные попытки распространить идеи марковской фильтрации на двумерные сигналы. Прежде чем остановиться на одном из методов, развитых в работах [3.6,3.8], рассмотрим кратко одномерную марковскую фильтрацию дискретных сигналов, поскольку она составляет основу двумерных процедур.

3.5.2. Марковская фильтрация одномерных последовательностей

Рассмотрим одномерную задачу фильтрации, когда входные данные представлены в виде одномерной последовательности наблюдений:

$$y_i = f(x_i n_i), \quad i = \overline{0, I-1} \quad (3.39)$$

Здесь все обозначения имеют тот же смысл, что и в (3.37) для двумерных сигналов. Для пояснения сути марковской фильтрации рассмотрим простейший вариант задачи: будем считать помеху независимым процессом (т.е. n_i -последовательность случайных независимых чисел), а сигнал - простой марковской последовательностью. На пояснении последнего остановимся подробнее. Последовательность является марковской, если ее совместное распределение вероятностей может быть представлено в виде:

$$w_x(X) = w_x(x_0) \prod_{j=1}^{I-1} q(x_j | x_{j-1}). \quad (3.40)$$

Данное выражение содержит в правой части одномерное распределение $w_x(x_0)$ для нулевого элемента последовательности и цепочку так называемых одношаговых распределений вероятностей перехода $q(x_j | x_{j-1})$, представляющих собой разновидность условных распределений. Соотношение (3.40) описывает свойство ограниченного последействия, проявляемое в том, что условное распределение $q(x_j | x_{j-1})$ элемента x_j зависит лишь от единственного соседнего элемента x_{j-1} . Последовательность как бы “прошита” цепочкой непосредственных соседних связей. Элементы, удаленные друг от друга более чем на один шаг, непосредственным вероятностным механизмом не связаны. Это, впрочем, совсем не означает их независимости, зависимость проявляется опосредованно, через цепочку прямых связей.

Часто индексы i, j , входящие в (3.40), ассоциируют с дискретным временем, а последовательность x_i называют случайным процессом. Тогда о соотношении (3.40) говорят, что оно описывает процесс в прямом времени. Известно, что марковский процесс x_i обладает марковским свойством и в обратном времени, что позволяет записать его распределение вероятностей в виде:

$$w_x(X) = w_x(x_{I-1}) \prod_{j=I-2}^0 \pi(x_j | x_{j+1}). \quad (3.41)$$

В соотношение (3.41) входит распределение последнего элемента и цепочка одношаговых распределений перехода в обратном времени $\pi(x_j | x_{j+1})$, не совпадающих с $q(x_j | x_{j-1})$.

Марковские процессы обладают разделяющим свойством, позволяющим представить их распределение еще в одной форме, полезной для разработки оптимальных процедур фильтрации. В соответствии с этим свойством любой элемент последовательности x_k разделяет ее на два условно независимых множества $X_{k-} = \{x_0, x_1, \dots, x_{k-1}\}$ и $X_{k+} = \{x_{k+1}, x_{k+2}, \dots, x_{I-1}\}$, которые при известном значении x_k условно независимы [3.7], т.е. имеют место равенства:

$$\begin{aligned}
w(X_{k-}, X_{k+} | x_k) &= w(X_{k-} | x_k) w(X_{k+} | x_k), \\
w_x(X) &= w(x_k) w(X_{k-} | x_k) w(X_{k+} | x_k) = \\
&= w(x_k) \prod_{r=k-1}^0 \pi(x_r | x_{r+1}) \prod_{j=k+1}^{I-1} q(x_j | x_{j-1}).
\end{aligned} \tag{3.42}$$

Последнее соотношение дает возможность построения некаузального фильтра, формирующего результат фильтрации при помощи очень удобных, экономичных вычислительных процедур. Это является результатом того, что апостериорное распределение вероятностей для произвольного k -го элемента последовательности может быть представлено в виде [3.7]:

$$w(x_k | Y) = c_k \frac{w(x_k | Y_{k-}) w(x_k | Y_{k+})}{w(x_k | y_k)}. \tag{3.43}$$

В правую часть (3.43) входят три частичных АРВ элемента x_k , различающиеся составом входных данных, на которых основаны эти АРВ. Здесь $Y_{k-} = \{y_0, y_1, \dots, y_{k-1}, y_k\}$, $Y_{k+} = \{y_k, y_{k+1}, \dots, y_{I-1}\}$ - векторы прошлых и будущих данных соответственно, причем оба содержат текущий элемент y_k . Основанные на них в отдельности АРВ могли бы послужить основой для образования каузальной и антикаузальной оценок полезного сигнала. В знаменателе стоит одноточечное АРВ, компенсирующее двукратное присутствие текущего наблюдения y_k в числителе, а коэффициент c_k подбирается так, чтобы обеспечивалась нормировка к единице получаемого АРВ.

Согласно (3.43) получение оценки складывается из двух этапов. На первом этапе из локальных входных данных формируются локальные АРВ, которые на втором этапе объединяются в окончательное АРВ, используемое далее для получения точечной оценки. Вычислительная сложность этого процесса в значительной степени определяется сложностью формирования локальных АРВ, главным образом находящихся в числителе формулы (3.43), т.к. получение одноточечного АРВ в знаменателе обычно является достаточно простой задачей.

Определение локальных АРВ очень сильно облегчается при использовании марковских свойств последовательностей. Оказывается, что они могут вычисляться при помощи рекуррентных соотношений. Так, например АРВ $w(x_k | Y_{k-})$ вычисляется на основе рекуррентного уравнения в прямом времени:

$$w(x_k | Y_{k-}) = c_{k-} \cdot \int_{\Omega_x} w(x_{k-1} | Y_{(k-1)-}) q(x_k | x_{k-1}) dx_{k-1}. \quad (3.44)$$

Здесь Ω_x - область интегрирования, определяемая областью значений последовательности x_i , а c_{k-} - нормирующий коэффициент. Для вычисления текущего АРВ $w(x_k | Y_{k-})$ необходимо, согласно (3.44), использовать АРВ $w(x_{k-1} | Y_{(k-1)-})$, являющееся одним из итогов работы фильтра в предыдущей точке. В удобстве этого выражения для практической реализации и состоит одно из основных преимуществ марковской фильтрации. Наряду с ограниченным объемом вычислений, предписанных (3.44), при работе на каждом шаге не требуется использовать полное множество входных данных, поскольку в выражение (3.44) входит только текущий элемент обработки y_k . Аналогично выглядит и рекуррентное соотношение для локального АРВ $w(x_k | Y_{k+})$ с тем лишь отличием, что оно развивается в обратном времени от конца интервала наблюдения к его началу. Оба рекуррентных соотношения должны быть дополнены граничными условиями, определяющими одноточечные АРВ $w(x_0 | y_0)$ и $w(x_{I-1} | y_{I-1})$, что, как упоминалось выше, не представляет сложной задачи.

В целом, процедура фильтрации, основанная на приведенных соотношениях, выглядит следующим образом. Сначала выполняется обработка последовательности в прямом времени, в результате чего во всех точках формируется АРВ $w(x_k | Y_{k-})$, затем осуществляется развертка в обратном времени, в ходе которой формируются локальные АРВ $w(x_k | Y_{k+})$. Далее снова осуществляется развертка в прямом (или в обратном) времени, в ходе которой вычисляются одноточечные $w(x_k | y_k)$ и полные $w(x_k | Y)$ АРВ. Одновременно в каждой точке определяется точечная оценка, например, как координата максимума АРВ. Состав вычислений очень удобен для параллельной реализации. При наличии параллельных вычислительных устройств можно одновременно в различных вычислителях определять различные локальные АРВ и затем в режиме последовательной обработки объединять все локальные результаты. Само это объединение происходит независимо для всех точек последовательности, что позволяет финальную процедуру формирования окончательных точечных оценок выполнять параллельно во всех точках.

3.5.3. Двухэтапная марковская фильтрация изображений

Рассмотрим подход к фильтрации изображений, основу которого составляет использование двумерных, но неполных входных данных, а также

наличие у них марковских свойств. Рассмотрим получение оценки изображения в произвольной точке кадра с координатами (i_1, i_2) , называя ее текущей точкой фильтрации. Осуществляя развертку изображения, можно последовательно в состояние текущей точки переводить все точки кадра, выполняя таким образом его полную обработку.

Будем считать, что для получения оценки x_{i_1, i_2}^* в точке (i_1, i_2) привлекаются лишь данные $Y = \{Y^{(1)}, Y^{(2)}, Y^{(3)}, Y^{(4)}, y\}$, образованные наблюдениями i_1 -й строки и i_2 -го столбца, на пересечении которых в кадре расположен оцениваемый элемент. Рис. 3.8 иллюстрирует геометрию задачи.

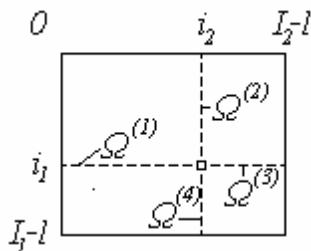


Рис. 3.8. Геометрия использования данных при двухэтапной фильтрации

Наблюдения $Y^{(k)}$, $k = \overline{1, 4}$ образованы входными данными, лежащими на горизонтальных и вертикальных лучах $\Omega^{(k)}$, расходящихся из текущей точки. Причем в каждый из этих векторов не входит текущее наблюдение $y = y_{i_1, i_2}$, которое в связи с этим представлено в Y отдельным элементом.

Будем, кроме того, рассматривать такие случайные поля X , которые обладают свойством условной независимости. Это означает, что совместное распределение всех его элементов X , расположенных на “кресте” Ω (рис. 3.8), можно представить в виде:

$$w(X) = w(x) \cdot w(X^{(1)}, X^{(3)}|x) \cdot w(X^{(2)}, X^{(4)}|x), \quad (3.45)$$

где верхние индексы также указывают на принадлежность векторов соответствующим лучам. Соотношение (3.45) означает, что значения сигнала на любой строке и на любом столбце изображения условно независимы, если известно значение сигнала x на пересечении этих строки и столбца. Если, кроме того, одномерные сигналы $X^{(1)}, x, X^{(3)}$ и $X^{(2)}, x, X^{(4)}$ являются марковскими последовательностями, для которых справедливо свойство условной независимости (3.42), то имеем:

$$w(X) = w(x) \cdot w(X^{(1)}|x) \cdot w(X^{(2)}|x) w(X^{(3)}|x) w(X^{(4)}|x). \quad (3.46)$$

Используя эту математическую модель изображения в случае независимой помехи n_{i_1, i_2} , можно одноточечное апостериорное распределение представить в виде [3.6]:

$$w(x|Y) = c \cdot \frac{\prod_{k=1}^4 w(x|Y^{(k)}, y)}{w^3(x|y)}. \quad (3.47)$$

Соотношение (3.47) служит теоретической базой для построения оптимальных двухэтапных процедур фильтрации, использующих неполные данные исходных наблюдений. Полное АРВ, основанное на всех *привлекаемых при фильтрации* данных Y , как и в одномерном случае, представляется в виде произведения частных АРВ, каждое из которых использует локальные данные одного из лучей $Y^{(k)}$ и текущий элемент y . Наличие в знаменателе третьей степени одноточечного АРВ служит компенсацией трехкратного “лишнего” участия текущего наблюдения в числителе. Константа c позволяет сделать АРВ нормированным.

Рассмотренный вариант обработки является разновидностью *некаузальной фильтрации*, поскольку в получении оценки участвуют элементы входного наблюдения, имеющие как большие, так и меньшие значения аргументов, чем у оцениваемого сигнала. Если в процессе обработки опираться только на данные двух лучей $\Omega^{(1)}$ и $\Omega^{(2)}$, то будет получен двумерный *каузальный фильтр*. При добавлении к ним третьего луча $\Omega^{(3)}$ реализуется полукаузальная фильтрация. Используя лучи $\Omega^{(1)}$ и $\Omega^{(3)}$, можем осуществить *одномерную некаузальную фильтрацию*. Коррекция, которой необходимо в этих случаях подвергнуть формулу (3.47), вполне очевидна.

Соотношение (3.47) дает возможность выполнить двумерную обработку изображения в виде некоторой совокупности одномерных процедур. Весь цикл вычислений можно представить следующим образом. Выполняется обработка всех строк изображения в прямом направлении (слева направо), в результате чего в каждой точке образуется распределение $w(x|Y^{(1)}, y)$. При этом используются одномерные рекуррентные процедуры, описанные выше. Далее происходит повторное сканирование строк, но в “обратном времени” - справа налево, в процессе которого вычисляются распределения $w(x|Y^{(3)}, y)$. Затем изображение аналогично дважды обрабатывается по столбцам - сверху вниз и снизу вверх, в результате чего определяются частные АРВ $w(x|Y^{(2)}, y)$ и $w(x|Y^{(4)}, y)$. Вычислением одноточечного АРВ $w(x|y)$ завершается первый этап обработки. На втором этапе происходит объединение всех частных АРВ в

каждой точке кадра в окончательное АРВ, а также на его основе вычисляются точечные оценки изображения x_{i_1, i_2}^* .

С точки зрения скорости вычислений данная технология обработки является очень привлекательной. Следует, вместе с тем, иметь в виду, что для ее реализации необходим достаточный запас оперативной памяти, чтобы хранить промежуточные результаты обработки, к числу которых относятся все частные АРВ. В этом отношении вычислительный процесс может быть существенно оптимизирован, поскольку ни одно из частных АРВ не представляет окончательной ценности. Это позволяет, например, не хранить отдельно пять различных распределений, входящих в правую часть (3.47), а по мере получения очередного сомножителя формировать произведение, именно которое и следует хранить в памяти до завершения вычислений. Очевидно, что структура вычислений, как и в одномерном случае, удобна для реализации при помощи многоканального вычислительного устройства.

Структура распределений очень сильно влияет на требуемые объем вычислений и ресурс памяти. Имеются очень “удобные” в этом смысле виды распределений. Например, если для описания изображения применима модель случайного поля с гауссовским распределением, то для представления каждого из частных и финального АРВ в (3.47) требуется наличие всего двух параметров - математического ожидания и дисперсии. Именно это и определяет конкретный характер и количество вычислений в процессе фильтрации, а также объем необходимой памяти.

Другим примером такого рода может служить математическая модель бинарного случайного поля, которое в различных точках принимает значения x_0 или x_1 . Такое описание также является очень экономичным, поскольку АРВ содержит всего две вероятности $w(x_0|Y)$ и $w(x_1|Y)$, непосредственное вычисление которых и выполняется при помощи (3.47).

Существует отдельный вопрос, связанный с применимостью марковских двумерных моделей (3.45), (3.46), позволяющих построить эффективные двухэтапные процедуры. Его изучение является достаточно непростой теоретической задачей. В частности, в работах [3.6., 3.8] установлено, что и для гауссовских, и для бинарных случайных полей необходимым и достаточным условием применимости (3.45) является возможность представления двумерных корреляционных функций этих полей в *разделимом виде*, т.е. в виде произведения двух множителей, один из которых описывает корреляцию изображения по строке, а второй - по столбцу. Дополнительные требования, вытекающие из (3.46), сводятся к существованию марковских свойств у одномерных последовательностей в горизонтальном и вертикальном сечениях изображения. В двух указанных примерах наличие таких свойств связано с *экспоненциальным видом* корреляционных функций этих одномерных сечений изображения.

На рис. 3.9 приведены результаты экспериментальной проверки двухэтапных алгоритмов фильтрации изображения. На рис. 3.9.а показано тестовое бинарное изображение “острова”, на рис. 3.9.б - изображение, искаженное белым гауссовским шумом (отношение сигнал/шум $q^2 = 0$ дБ). Рис.3.9.в иллюстрирует применение простой поэлементной пороговой обработки (рис. 1.4.а), при которой порог определялся так, чтобы реализовывалась одноточечная процедура максимума апостериорной вероятности. На рис. 3.9.г, 3.9.д и 3.9.е показаны различные результаты двухэтапной фильтрации. Первый из них соответствует одномерной каузальной фильтрации, второй - также одномерной, но некаузальной, а третий - двумерной некаузальной процедуре. Визуальное сравнение результатов говорит об очень низком качестве поэлементной обработки. При ее использовании вероятность

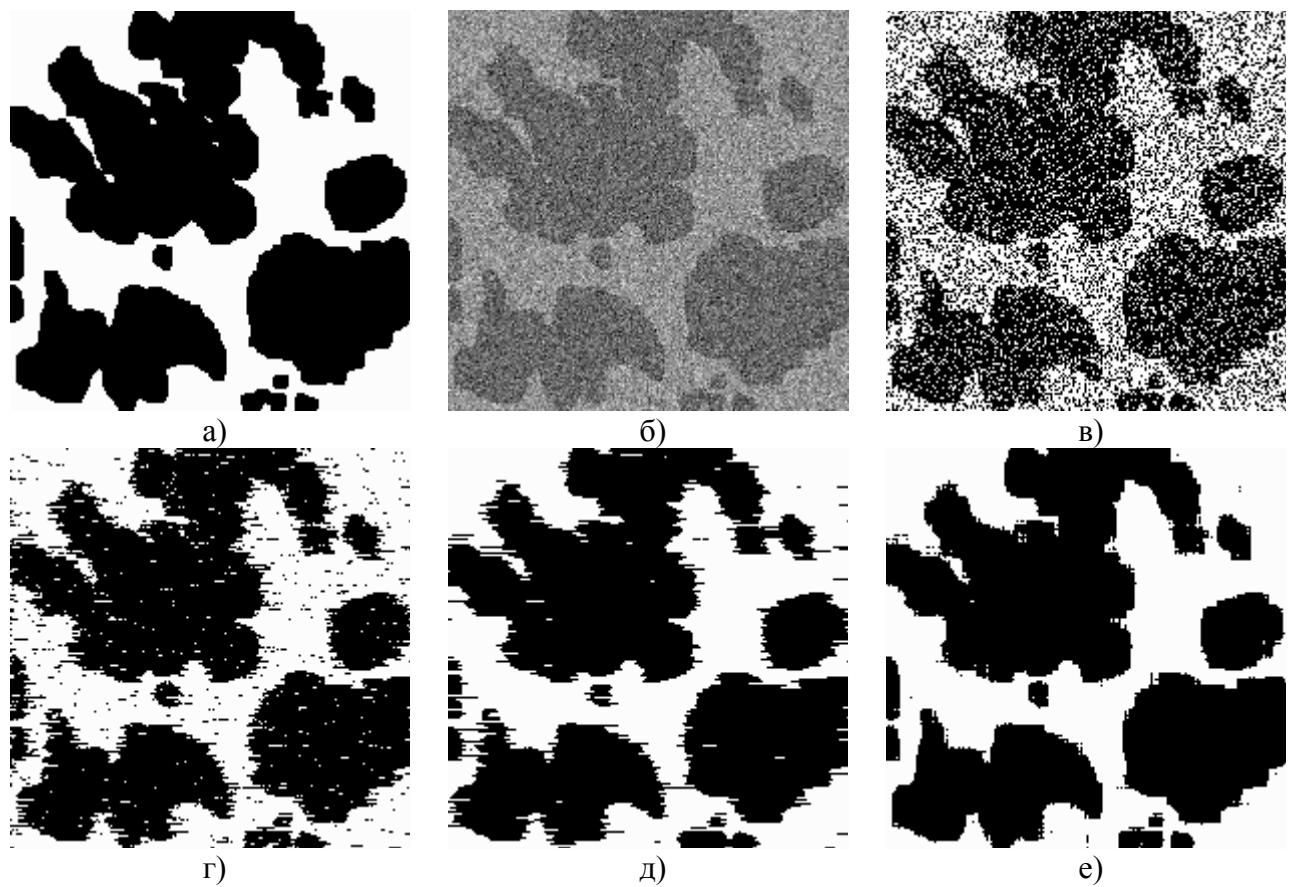


Рис. 3.9. Двухэтапная марковская фильтрация изображения

ошибки (т.е. события, состоящего в замене числа x_0 числом x_1 или наоборот) составила 0.23. Качество обработки улучшается при использовании фильтрации, причем оно повышается как при переходе от одномерной каузальной (при которой вероятность ошибки составляет 0.086) к одномерной некаузальной (вероятность ошибки 0.041), так и при переходе к двумерной обработке, при которой достигается вероятность ошибки равная 0.022. Таким образом, применение одномерной некаузальной фильтрации позволяет уменьшить вероятность ошибки в 5 раз по сравнению с поэлементной

пороговой обработкой, а двумерной некаузальной фильтрации - почти в 10 раз. Эти примеры говорят об очень высокой эффективности, которой может достигать фильтрация, и убеждают в полезности тех значительных усилий, которые необходимы для нахождения эффективных алгоритмов.

3.6. Медианная фильтрация

Все линейные алгоритмы фильтрации приводят к сглаживанию резких перепадов яркости изображений, прошедших обработку. Этот недостаток, особенно существенный, если потребителем информации является человек, принципиально не может быть исключен в рамках линейной обработки. Дело в том, что линейные процедуры являются оптимальными при гауссовском распределении сигналов, помех и наблюдаемых данных. Реальные изображения, строго говоря, не подчиняются данному распределению вероятностей. Причем, одна из основных причин этого состоит в наличии у изображений разнообразных границ, перепадов яркости, переходов от одной текстуры к другой и т. п.. Поддаваясь локальному гауссовскому описанию в пределах ограниченных участков, многие реальные изображения в этой связи плохо представляются как глобально гауссовые объекты. Именно это и служит причиной плохой передачи границ при линейной фильтрации.

Вторая особенность линейной фильтрации - ее оптимальность, как только что упоминалось, при гауссовском характере помех. Обычно этому условию отвечают шумовые помехи на изображениях, поэтому при их подавлении линейные алгоритмы имеют высокие показатели. Однако, часто приходится иметь дело с изображениями, искаженными помехами других типов. Одной из них является импульсная помеха. При ее воздействии на изображении наблюдаются белые или (и) черные точки, хаотически разбросанные по кадру. Применение линейной фильтрации в этом случае неэффективно - каждый из входных импульсов (по сути - дельта-функция) дает отклик в виде импульсной характеристики фильтра, а их совокупность способствует распространению помехи на всю площадь кадра.

Удачным решением перечисленных проблем является применение *медианной фильтрации*, предложенной Дж. Тьюки в 1971 г. для анализа экономических процессов. Наиболее полное исследование медианной фильтрации применительно к обработке изображений представлено в сборнике [3.9]. Отметим, что медианная фильтрация представляет собой эвристический метод обработки, ее алгоритм не является математическим решением строго сформулированной задачи. Поэтому исследователями уделяется большое внимание *анализу эффективности* обработки изображений на ее основе и сопоставлению с другими методами.

При применении медианного фильтра (МФ) происходит последовательная обработка каждой точки кадра, в результате чего образуется последовательность оценок. В идейном отношении обработка в различных

точках независима (этим МФ похож на масочный фильтр), но в целях ее ускорения целесообразно алгоритмически на каждом шаге использовать ранее выполненные вычисления.

При медианной фильтрации используется двумерное окно (апертура фильтра), обычно имеющее центральную симметрию, при этом его центр располагается в текущей точке фильтрации. На рис. 3.10 показаны два примера наиболее часто применяемых вариантов окон в виде креста и в виде квадрата. Размеры апертуры принадлежат к числу параметров, оптимизируемых в процессе анализа эффективности алгоритма. Отсчеты изображения, оказавшиеся в пределах окна, образуют *рабочую выборку* текущего шага.



Рис. 3.10. Примеры окон при медианной фильтрации

Двумерный характер окна позволяет выполнять, по существу, двумерную фильтрацию, поскольку для образования оценки привлекаются данные как из текущих строки и столбца, так и из соседних. Обозначим рабочую выборку в виде одномерного массива $Y = \{y_1, y_2, \dots, y_n\}$; число его элементов равняется размеру окна, а их расположение произвольно. Обычно применяют окна с нечетным числом точек n (это автоматически обеспечивается при центральной симметрии апертуры и при вхождении самой центральной точки в ее состав). Если упорядочить последовательность $\{y_i, i = \overline{1, n}\}$ по возрастанию, то ее *медианой* будет тот элемент выборки, который занимает центральное положение в этой упорядоченной последовательности. Полученное таким образом число и является продуктом фильтрации для текущей точки кадра. Понятно, что результат такой обработки в самом деле не зависит от того, в какой последовательности представлены элементы изображения в рабочей выборке Y . Введем формальное обозначение описанной процедуры в виде:

$$x^* = med(y_1, y_2, \dots, y_n). \quad (3.48)$$

Рассмотрим пример. Предположим, что выборка имеет вид: $Y = \{136, 110, 99, 45, 250, 55, 158, 104, 75\}$, а элемент 250, расположенный в ее центре, соответствует текущей точке фильтрации (i_1, i_2) (рис. 3.10). Большое значение яркости в этой точке кадра может быть результатом воздействия

импульсной (точечной) помехи. Упорядоченная по возрастанию выборка имеет при этом вид $\{45, 55, 75, 99, 104, 110, 136, 158, 250\}$, следовательно, в соответствии с процедурой (3.48), получаем $x^* = med(y_1, y_2, \dots, y_9) = 104$. Видим, что влияние “соседей” на результат фильтрации в текущей точке привело к “игнорированию” импульсного выброса яркости, что следует рассматривать как эффект фильтрации. Если импульсная помеха не является точечной, а покрывает некоторую локальную область, то она также может быть подавлена. Это произойдет, если размер этой локальной области будет меньше, чем половина размера апертуры МФ. Поэтому для подавления импульсных помех, поражающих локальные участки изображения, следует увеличивать размеры апертуры МФ.

Из (3.48) следует, что действие МФ состоит в “игнорировании” экстремальных значений входной выборки - как положительных, так и отрицательных выбросов. Такой принцип подавления помехи может быть применен и для ослабления шума на изображении. Однако исследование подавления шума при помощи медианной фильтрации показывает, что ее эффективность при решении этой задачи ниже, чем у линейной фильтрации [3.9].

Результаты экспериментов, иллюстрирующие работу МФ, приведены на рис. 3.11. В экспериментах применялся МФ, имеющий квадратную апертуру со стороной равной 3. В левом ряду представлены изображения, искаженные помехой, в правом - результаты их медианной фильтрации. На рис. 3.11.а и рис. 3.11.в показано исходное изображение, искаженное импульсной помехой. При ее наложении использовался датчик случайных чисел с равномерным на интервале $[0, 1]$ законом распределения, вырабатывающий во всех точках кадра независимые случайные числа. Интенсивность помехи задавалась вероятностью p ее возникновения в каждой точке. Если для случайного числа n_{i_1, i_2} , сформированного в точке (i_1, i_2) , выполнялось условие $n_{i_1, i_2} < p$, то яркость изображения x_{i_1, i_2} в этой точке замещалась числом 255, соответствующим максимальной яркости (уровню белого). На рис. 3.11.а действием импульсной помехи искажено 5 % ($p=0.05$), а на рис. 3.11.в - 10 % элементов изображения. Результаты обработки говорят о практически полном подавлении помехи в первом случае и о ее значительном ослаблении во втором.



а)



б)



в)



г)



д)



е)

Рис. 3.11. Примеры медианной фильтрации

Рис. 3.11.д показывает изображение, искаженное независимым гауссовским шумом при отношении сигнал/шум $q^2 = -5 \text{ дБ}$, а рис. 3.11.е - результат его фильтрации медианным фильтром. Условия данного

эксперимента позволяют сравнивать его результаты с результатами рассмотренной выше линейной фильтрации. В таблице 3.1 приведены данные, дающие возможность такого сравнения. Для различных методов фильтрации в этой таблице приводятся значения относительного среднего квадрата ошибок δ_ε^2 и коэффициента ослабления шума γ для случая, когда отношение сигнал/шум на входе фильтра составляет -5 дБ.

	масочный фильтр с оптимальн. КИХ	масочный фильтр с равномерн. КИХ	двумерный рекуррентн. фильтр	двумерный фильтр Винера	медианный фильтр
δ_ε^2	0.309	0.395	0.29	0.186	0.539
γ	10.2	8.0	10.9	17.0	5.86

Табл.3.1. Сравнение эффективности подавления шума при фильтрации изображений,

$$q^2 = -5 \text{ дБ}$$

Наибольшей эффективностью обладает двумерный фильтр Винера, уменьшающий средний квадрат ошибок в 17 раз. Медианный фильтр имеет наименьшую из всех рассмотренных фильтров эффективность, ему соответствует $\gamma=5.86$. Тем не менее, это число свидетельствует о том, что и при его помощи удается значительно снизить уровень шума на изображении.

Вместе с тем, как говорилось выше, и что демонстрирует рис. 3.11.е, медианная фильтрация в меньшей степени сглаживает границы изображения, чем любая линейная фильтрация. Механизм этого явления очень прост и заключается в следующем. Предположим, что апертура фильтра находится вблизи границы, разделяющей светлый и темный участки изображения, при этом ее центр располагается в области темного участка. Тогда, вероятнее всего, рабочая выборка будет содержать большее количество элементов с малыми значениями яркости, и, следовательно, медиана будет находиться среди тех элементов рабочей выборки, которые соответствуют этой области изображения. Ситуация меняется на противоположную, если центр апертуры смешен в область более высокой яркости. Но это и означает наличие чувствительности у МФ к перепадам яркости.

ВОПРОСЫ К ГЛАВЕ 3

- 3.1. Докажите, что при наблюдении аддитивной смеси независимых изображения и шума взаимная корреляционная функция $B_{xy}(i, j)$ совпадает с корреляционной функцией изображения $B_x(i, j)$.

- 3.2. Приведите уравнение Винера-Хопфа (3.6) к виду (3.11) применительно к случаю аддитивного взаимодействия изображения и белого шума.
- 3.3. Докажите справедливость выражения (3.7) для среднего квадрата ошибок фильтрации.
- 3.4. Поясните, почему при ограниченном размере окрестности, применяемой при КИХ-фильтрации, нельзя достичь предельного подавления шума ?
- 3.5. Докажите, что стационарный фильтр Калмана, работающий в соответствии с алгоритмом (3.12), имеет импульсную характеристику (3.13).
- 3.6. Докажите, что при оптимальной линейной фильтрации ошибка и результат фильтрации ортогональны.
- 3.7. Получите алгебраическое уравнение (3.18) для коэффициента усиления двумерного рекуррентного фильтра и качественно исследуйте тенденцию изменения его решения при вариации отношения сигнал/шум.
- 3.8. Приведите алгоритм фильтрации в двумерном фильтре с ИХ (3.14) к рекуррентному виду (3.19).
- 3.9. Докажите тождественность прямого и обратного двумерных ДПФ.
- 3.10. Покажите на примере, как может быть получено одноточечное апостериорное распределение вероятностей, входящее в формулу (3.43) для АРВ марковской последовательности.
- 3.11. По аналогии с рекуррентным уравнением прямого времени для локального АРВ $w(x_k | Y_{k-})$ запишите рекуррентное уравнение обратного времени для АРВ $w(x_k | Y_{k+})$ и поясните его.

Глава 4

ВОССТАНОВЛЕНИЕ ИЗОБРАЖЕНИЙ

Из-за несовершенства формирующих и регистрирующих систем записанное ими изображение представляет собой искаженную (нечеткую) копию оригинала. Основными причинами искажений, приводящих к ухудшению четкости, являются ограниченная разрешающая способность формирующей системы, расфокусировка, наличие искажающей среды (например, атмосферы), движение камеры по отношению к регистрируемому объекту и т.п. Устранение или ослабление искажений с целью повышения резкости относится к задаче восстановления изображений.

Наиболее общая схема формирования изображения представлена на рис. 4.1,

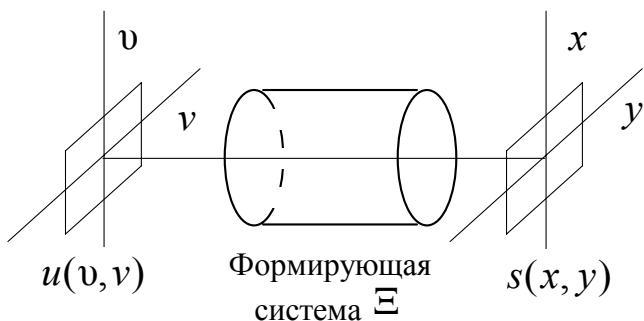


Рис.4.1. Схема формирования изображения

где $u(v, v)$ - неизвестная функция распределения яркости объекта, описываемая функцией двух переменных (v, v) ; $s(x, y)$ - наблюдаемое изображение, сформированное из $u(v, v)$ при помощи некоторого известного оператора искажений Ξ :

$$s(x, y) = \Xi u(v, v).$$

Вид оператора Ξ определяется свойствами формирующей системы. Двумерную функцию $u(v, v)$ в дальнейшем будем называть исходным изображением. Задача восстановления заключается в нахождении изображения $\hat{u}(v, v)$, являющегося оценкой исходного изображения $u(v, v)$ по наблюдаемому изображению $s(x, y)$, т.е. в устранении искажений, вносимых оператором Ξ .

4.1. Модели изображений и их линейных искажений

4.1.1. Формирование изображений

Большинство формирующих систем в первом приближении можно рассматривать как линейные и инвариантные к сдвигу. Изображения, сформированные такими системами, претерпевают *линейные пространственно-инвариантные искажения*, характеризующиеся тем, что механизм их возникновения один и тот же для всех точек (x, y) . Линейные искажения проявляются в ослаблении верхних частот исходного изображения. Визуально это приводит к ухудшению его резкости. В процессе записи изображения искажаются также шумами, присутствующими в любом реальном физическом устройстве. В ряде практически важных случаев шум можно считать аддитивным и независящим от исходного изображения.

С учетом вышеизложенного наблюдаемое нерезкое изображение $s(x, y)$ можно представить как выход линейной системы, показанной на рис. 4.2,

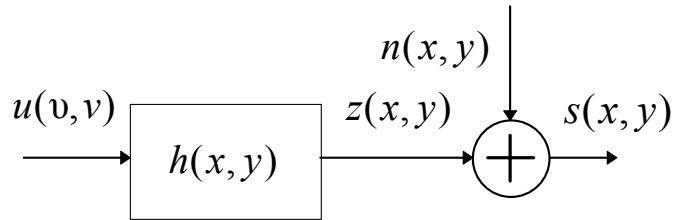


Рис.4.2. Линейная модель формирования изображения

а математическая модель процесса его формирования имеет вид:

$$s(x, y) = z(x, y) + n(x, y),$$

где $n(x, y)$ - аддитивный двумерный шум. Изображение $z(x, y)$, полученное путем линейного искажения исходного изображения при отсутствии шума, определяется интегралом свертки:

$$\begin{aligned}
 z(x, y) &= h(x, y) \otimes \otimes u(x, y) = \iint_{(v, v) \in \Theta^{(u)}} h(x - v, y - v) u(v, v) dv dv = \\
 &= \iint_{(v, v) \in \Theta^{(h)}} h(v, v) u(x - v, y - v) dv dv, \quad (x, y) \in \Theta^{(s)},
 \end{aligned} \tag{4.1}$$

где $\otimes \otimes$ - символ двумерной свертки; $h(x, y)$ - двумерная импульсная характеристика (или ФРТ - функция рассеяния точки) линейной искажающей системы. Таким образом, значение функции яркости $u(v, v)$ исходного изображения в точке с координатами (v, v) «размазывается» в соответствии с видом ФРТ $h(x, y)$ и искажается аддитивным шумом.

В выражении (4.1) учтено, что изображения, встречающиеся в практических задачах, имеют конечные размеры. Это означает, что яркость изображения полагается равной нулю всюду, кроме некоторой конечной области, которую будем называть кадром и обозначать через $\Theta^{(\cdot)}$. Верхний индекс в кадре $\Theta^{(\cdot)}$ соответствует символу изображения или ФРТ, для которого определен этот кадр. Например, кадр изображения $s(x, y)$ будем обозначать через $\Theta^{(s)}$. Мы будем рассматривать только прямоугольные кадры, стороны которых параллельны координатным осям. Под размерами кадра будем понимать совокупность длин вертикальной $L_x^{(s)}$ и горизонтальной $L_y^{(s)}$ сторон кадра $\Theta^{(s)}$. Часто оказывается удобным совместить центр кадра изображения с началом координат на плоскости изображения, тогда

$$s(x, y) = 0 \text{ при } |x| \geq L_x^{(s)} / 2, |y| \geq L_y^{(s)} / 2, \quad (4.2)$$

или когда координаты $(x, y) \notin \Theta^{(s)}$.

Относительные размеры кадров изображений и ФРТ в модели формирования (4.1) имеют важное значение. Как будет показано ниже, конечность их размеров значительно усложняет решение задачи восстановления. Размеры кадров $\Theta^{(s)}$, $\Theta^{(z)}$ и $\Theta^{(n)}$ равны между собой и обусловлены тем, что любая формирующая система имеет ограниченное поле зрения (рис.4.3).

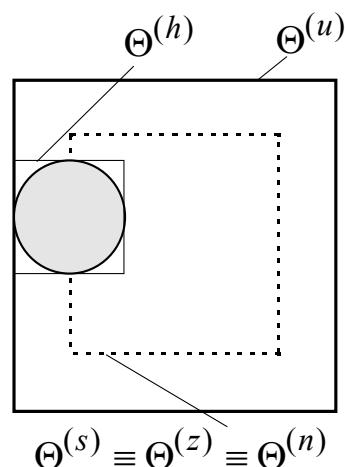


Рис.4.3. Относительные размеры изображения и ФРТ

Кадр $\Theta^{(h)}$ функции рассеяния точки представляет собой минимальный прямоугольник, содержащий все точки с координатами (x, y) , в которых величина абсолютного значения ФРТ заметно отлична от нуля, например, когда $|h(x, y)| \geq 10^{-10} \max(|h(x, y)|)$. Следует отметить, что кадр $\Theta^{(h)}$ ФРТ не всегда симметричен относительно начала координат. В частности, симметрией ФРТ не обладают каузальные линейные системы. Действие ФРТ сводится к тому, что каждая точка исходного изображения $s(v, v)$ «размазывается» в некоторую область, ограниченную кадром $\Theta^{(h)}$. Кадр $\Theta^{(u)}$ исходного изображения может быть построен путем перемещения $\Theta^{(h)}$ и представляет собой область всех точек, охватываемых кадром $\Theta^{(h)}$ при его перемещении по кадру $\Theta^{(s)}$ наблюдаемого изображения. Поэтому даже если регистрируемый объект имеет бесконечные размеры наблюдаемое изображение формируется лишь только за счет некоторой его части. Причем размеры кадра наблюдаемого изображения всегда меньше или равны размерам исходного. Размеры кадров исходного и наблюдаемого изображений равны лишь при отсутствии линейных искажений, т.е. когда импульсная характеристика искажающей системы равна дельта-функции. Нас будет интересовать восстановление изображения в пределах кадра $\Theta^{(s)}$. В некоторых случаях удается восстановить изображение в пределах кадра $\Theta^{(u)}$ исходного изображения, т.к. та его часть, которая лежит вне пределов кадра $\Theta^{(s)}$, также оказывает влияние на наблюдаемое изображение $s(x, y)$.

Для изображений, представленных в цифровой форме, двумерные функции $u(v, v)$, $z(x, y)$, $s(x, y)$, $h(x, y)$ и $n(x, y)$ с непрерывными аргументами в (4.1) заменяются двумерными массивами отсчетов, взятых на прямоугольных решетках с одинаковыми расстояниями $\Delta = \Delta x = \Delta y$ между узлами. В этом случае соотношение (4.1) принимает вид:

$$s(i_1, i_2) = z(i_1, i_2) + n(i_1, i_2), \quad i_1 = -\frac{L_1^{(s)}}{2}, \frac{L_1^{(s)}}{2}, \quad i_2 = -\frac{L_2^{(s)}}{2}, \frac{L_2^{(s)}}{2}. \quad (4.3)$$

где

$$z(i_1, i_2) = h(i_1, i_2) \otimes \otimes u(i_1, i_2) = \sum_{j_1 j_2 \in \Theta^{(h)}} h(j_1, j_2) u(i_1 - j_1, i_2 - j_2). \quad (4.4)$$

Аргументы с индексом 1 обозначают номер строки, а с индексом 2 - номер столбца. В дискретном случае размеры кадра (число отсчетов) $L_{i_1}^{(\cdot)}$ и $L_{i_2}^{(\cdot)}$

определяются отношением длин соответственно вертикальной и горизонтальной сторон кадра аналогового изображения к величине интервала дискретизации Δ .

Операция свертки, которая имеется в формулах (4.1) и (4.4), эквивалентна произведению в частотной области. Это позволяет выполнить быструю имитацию линейных искажений с помощью ДПФ, заменив обычную свертку циклической (смотри главу 3). Как правило, размеры кадра ФРТ много меньше размеров кадра исходного изображения, поэтому перед преобразованием массив $h(i_1, i_2)$ должен быть дополнен нулями. Кроме того, полагается, что изображения и ФРТ являются периодически продолженными, которые так же как и в главе 3 обозначаются волнистой линией.

Спектр линейно-искаженного изображения $\tilde{Z}(i_1, i_2)$ равен произведению спектра $\tilde{U}(k_1, k_2)$ исходного изображения $\tilde{u}(i_1, i_2)$ и передаточной функции $\tilde{H}(k_1, k_2)$ искажающей системы:

$$\tilde{Z}(k_1, k_2) = \tilde{U}(k_1, k_2) \tilde{H}(k_1, k_2), \quad (4.5)$$

где k_1 и k_2 - пространственные частоты. Размеры кадра изображения $\tilde{Z}(i_1, i_2)$, полученного после обратного ДПФ от $\tilde{Z}(k_1, k_2)$, равны размерам кадра $\Theta^{(u)}$ исходного изображения. Для завершения процедуры имитации необходимо «обрезать» края изображения $\tilde{Z}(i_1, i_2)$ до размеров кадра $\Theta^{(s)}$ и добавить аддитивную помеху $n(i_1, i_2)$. Здесь и далее результаты преобразования Фурье от функций в пространственной области будут обозначаться соответствующими прописными буквами.

Рассмотрим импульсные и частотные характеристики формирующих систем при смазе и расфокусировке.

4.1.2. Размытие вследствие движения (смаз)

Смаз изображения возникает при взаимном движении камеры и объекта относительно друг друга во время экспозиции. Наблюданное изображение окажется как бы результатом наложения со смещением множества исходных изображений. Мы рассмотрим только тот случай, когда камера перемещается с постоянной горизонтальной скоростью относительно снимаемого объекта. ФРТ и передаточная функция такой системы определяются выражениями:

$$h(x, y) = \begin{cases} \frac{1}{\alpha} \delta(x), & \text{при } 0 \leq y \leq \alpha; \\ 0 & \text{в остальных случаях,} \end{cases}$$

$$H(\omega_x, \omega_y) = \frac{\sin(\alpha \omega_y)}{\omega_y},$$

где длина смаза α - равна произведению скорости движения камеры на время экспозиции. Соответственно в дискретном случае ФРТ смаза равна

$$h(i_1, i_2) = \begin{cases} \frac{1}{L_{i_2}^{(h)}}, & \text{при } 0 \leq i_2 \leq L_{i_2}^{(h)} - 1 \text{ и } i_1 = 0; \\ 0 & \text{в остальных случаях,} \end{cases} \quad (4.6)$$

где размеры кадра $L_{i_2}^{(h)} = [\alpha / \Delta]$ и $L_{i_1}^{(h)} = 1$. Здесь квадратные скобки обозначают операцию округления до целого. Взаимное расположение кадров изображений и ФРТ при смазе показано на рис. 4.4.

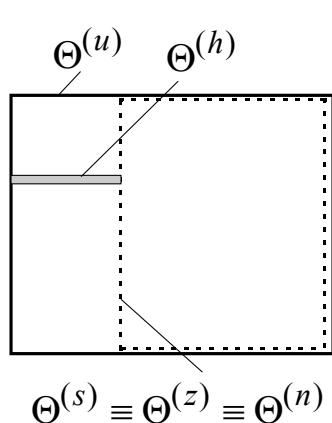


Рис.4.4. Взаимное расположение изображения и ФРТ при смазе

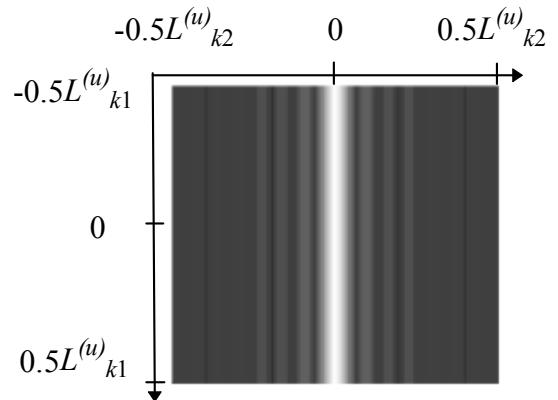


Рис.4.5. Изображение модуля частотной характеристики искажающей системы

Дополнив ФРТ (4.6) нулями до размеров кадра исходного изображения и применяя двумерное ДПФ, получим частотную характеристику искажающей системы:

$$\tilde{H}(k_1, k_2) = \frac{1}{L_{i_2}^{(h)} \sqrt{L_{i_2}^{(u)}}} e^{j(\pi(L_{i_2}^{(h)} - 1)k_2 / L_{i_2}^{(u)})} \frac{\sin(\pi L_{i_2}^{(h)} k_2 / L_{i_2}^{(u)}))}{\sin(\pi k_2 / L_{i_2}^{(u)}))} \quad (4.7)$$

для всех k_1 .

Изображение модуля $\tilde{H}(k_1, k_2)$ приведено на рис.4.5 при $L_{i_2}^{(h)} = 15$ и размерах исходного изображения 185×170 элементов.

На рис.4.7 приведен искаженный вариант исходного изображения «Сатурн» (рис.4.6). Горизонтальный смаз составляет 15 элементов. Исходное изображение содержит 185×170 элементов, а искаженное - 170×170 элементов.



Рис.4.6.Исходное изображение
“Сатурн”

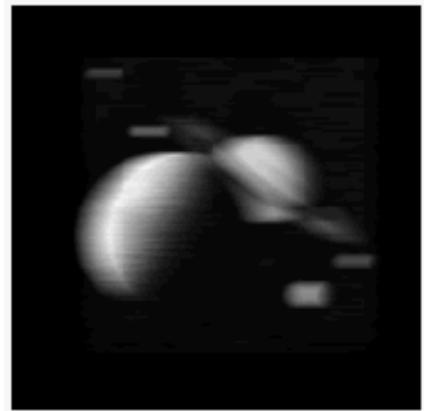


Рис.4.7. Смазанное
изображение “Сатурн”

4.1.3. Расфокусировка

Четкость изображения характеризуется воспроизведением мелких деталей и определяется разрешающей способностью формирующей системы. Разрешающая способность, например, оптической системы численно выражается количеством пар черно-белых линий на 1 мм изображения, которое формируется объективом системы. Если плоскость формируемого изображения находится в фокусе объектива, то пучок лучей, исходящий от точки на объекте, сходится в точку на изображении. При расфокусировке точка воспроизводится в виде некоторого пятна (кружка размытия), и две близко расположенные точки на исходном изображении сливаются в одну на наблюдаемом. Величина кружка размытия зависит от фокусного расстояния объектива, а также от расстояний от объектива до объекта и до плоскости формируемого изображения [4.1]. Дискретное изображение будет четким (сфокусированным), если диаметр кружка размытия не превышает шага дискретизации Δ наблюдаемого изображения. В противном случае линейные искажения становятся заметными.

При расфокусировке распределение интенсивности на изображении точечного источника, формируемого тонкой линзой с круговой апертурой, постоянно в пределах кружка размытия радиусом r и равно нулю за его пределами. Это соответствует цилиндрической ФРТ

$$h(x, y) = \begin{cases} \frac{1}{\pi r^2}, & x^2 + y^2 \leq r^2; \\ 0 & \text{в остальных случаях.} \end{cases} \quad (4.8)$$

Из (4.8) следует, что размеры кадра $L_x^{(h)} = L_y^{(h)} = 2r$. Взяв двумерное преобразование Фурье от (4.8), получим передаточную функцию оптической системы

$$H(\omega_x, \omega_y) = J_1(r\rho)/r\rho, \quad \rho^2 = \omega_x^2 + \omega_y^2, \quad (4.9)$$

где $J_1(\cdot)$ - функция Бесселя первого порядка. В дискретном случае ФРТ (4.8) имеет вид:

$$h(i_1, i_2) = \begin{cases} \frac{1}{\pi r^2}, & i_1^2 + i_2^2 \leq r^2; \\ 0 & \text{в остальных случаях.} \end{cases} \quad (4.10)$$

На рис.4.8 и 4.9 показаны ФРТ для тонкой линзы (4.10) и модуль ее передаточной функции при радиусе кружка размытия $r=10$ и размерах кадра изображения 200×200 элементов.

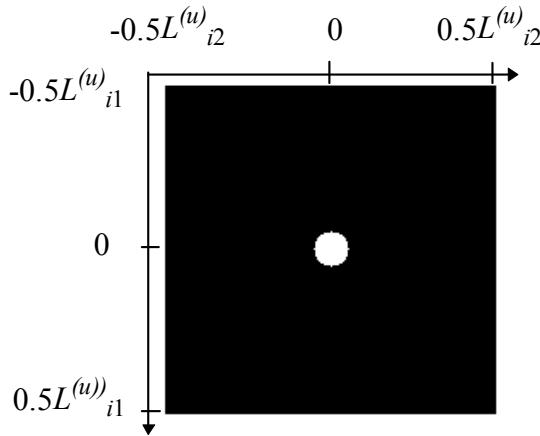


Рис.4.8. ФРТ тонкой линзы

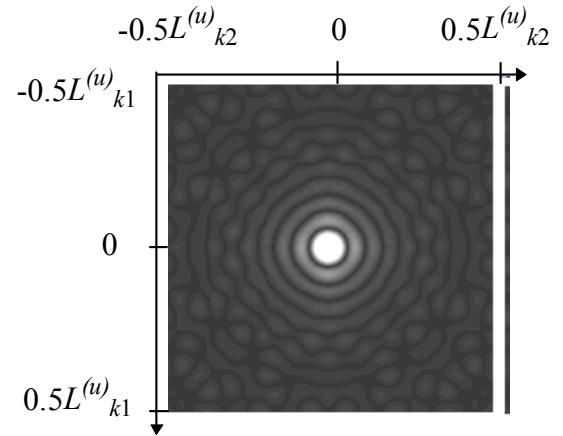


Рис.4.9. Изображение модуля частотной характеристики тонкой линзы

Земную атмосферу также можно рассматривать как оптическую систему. В качестве приближенной модели ФРТ такой системы используется двумерный гауссовский импульс

$$h(x, y) = K \exp[-\alpha (x^2 + y^2)], \quad (4.11)$$

который в дискретном случае имеет вид

$$h(i_1, i_2) = K \exp[-\alpha (i_1^2 + i_2^2)], \quad (4.12)$$

где K - нормирующий коэффициент, α - коэффициент пространственной нерезкости. Передаточная функция, соответствующая ФРТ (4.11), определяется выражением

$$H(\omega_x, \omega_y) = \exp[-(\omega_x^2 + \omega_y^2)/4\alpha]. \quad (4.13)$$

Очевидно, что точки, для которых выполняется условие (4.2), образуют круг радиусом

$$r = \sqrt{-\frac{\ln(10^{-10})}{\alpha}}. \quad (4.14)$$

Следовательно, чем больше α , тем меньше расфокусировка наблюдаемого изображения. ФРТ для земной атмосферы и соответствующая ей передаточная функция при $\alpha = 0.1$ приведены на рис. 4.10 и 4.11. Радиус кружка размытия примерно равен 15. Размеры пятна ФРТ на рис. 4.10 визуально кажутся меньше чем размеры пятна для тонкой линзы (рис. 4.8), т.к. гауссовский импульс является быстро убывающей функцией.

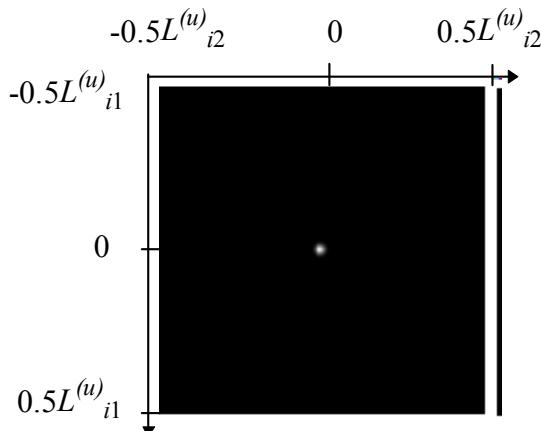


Рис.4.10. ФРТ атмосферы Земли

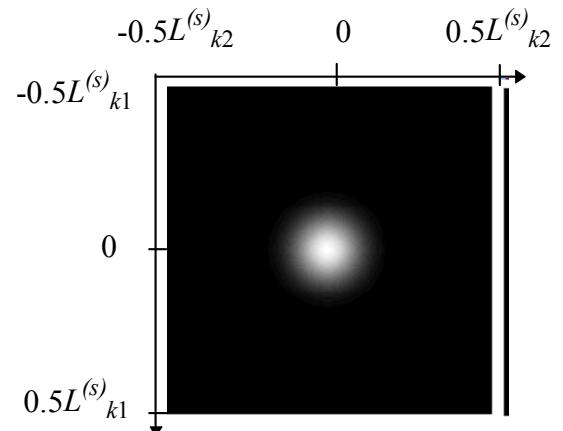


Рис.4.11. Изображение модуля частотной характеристики атмосферы Земли

Соотношения между различными кадрами изображений при расфокусировке соответствуют тем, что приведены на рис. 4.3, поскольку кадр ФРТ симметричен относительно центра координат. На рис. 4.12 приведен

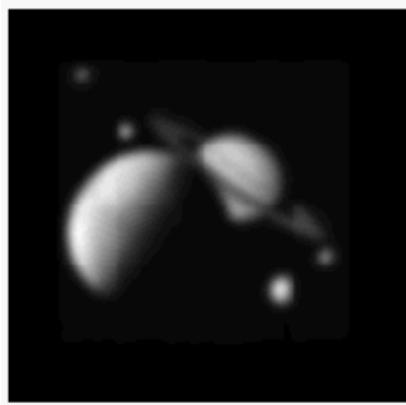


Рис.4.12. Дефокусированное изображение “Сатурн”

искаженный вариант изображения «Сатурн» (рис.4.6). Свертка исходного изображения производилась с гауссовским импульсом при $\alpha = 0.1$. Искаженное изображение содержит 170×170 элементов.

Таким образом, можно выделить три основных фактора, которые существенно усложняют решение проблемы восстановления изображений.

1. Искажения типа расфокусировок или смаз проявляются в ослаблении верхних пространственных частот изображения, т.к. формирующие системы представляют собой фильтры нижних частот. При этом отношение сигнал/шум на верхних частотах, определяющих четкость изображения, будет значительно хуже, чем для изображения в целом. Если система, формирующая изображение, ослабляет сигнал на каких-то пространственных частотах, то при восстановлении он должен быть усилен в той мере, в какой был ослаблен. Вместе с сигналом будут усиливаться и шумы. Поэтому улучшение качества изображения по резкости может привести к ухудшению его качества по зашумленности.
2. Яркость на краях кадра искаженного изображения зависит от яркости объектов, расположенных вне кадра, за счет свертки исходного изображения с ФРТ. При восстановлении изображений из-за неполной информации о сигнале вне кадра возникают краевые эффекты. Влияние краевых эффектов на качество восстановления в ряде случаев оказывается даже более существенным, чем зашумленность изображения.
3. При искажениях, вызванных движением или расфокусировкой камеры, передаточные функции (4.7) и (4.9) имеют нули, наличие которых обусловлено осциллирующим характером передаточных функций. Поскольку спектр искаженного изображения равен произведению спектра исходного изображения и передаточной функции (см. (4.5)), то наличие нулей приводит к полной утрате данных об исходном изображении на соответствующих частотах. По этой причине не удается абсолютно точно восстановить исходное изображение по наблюдаемому изображению, даже если отсутствуют шумы наблюдения и размеры кадров неограничены.

При решении задач восстановления изображений используются различные алгоритмы, как имеющие строгое математическое обоснование, так и эмпирические. Для искажений, описываемых уравнением свертки, эти алгоритмы условно можно разделить на три основные группы: алгоритмы решения системы алгебраических уравнений, алгоритмы фильтрации изображений в частотной области и итерационные алгоритмы.

4.2. Алгебраические методы восстановления изображений

Соотношение (4.4) для цифровых изображений фактически представляет собой систему линейных алгебраических уравнений относительно $u(i_1, i_2)$. Поэтому задача восстановления исходного изображения $u(i_1, i_2)$ при известной ФРТ $h(i_1, i_2)$ может быть сведена к решению такой системы.

Удобно представить соотношения (4.3) и (4.4) в матричной форме, используя лексикографическое упорядочивание. Для этого двумерный массив наблюдаемого изображения $s(i_1, i_2)$ размера $L^{(s)} = (L_1^{(s)} + 1) \times (L_2^{(s)} + 1)$ преобразуем в вектор-столбец s размера $L^{(s)} \times 1$. Преобразование осуществляется разверткой массива $s(i_1, i_2)$ по строкам. Аналогичным образом преобразуются в вектор-столбцы z , u и n искаженное при отсутствии шума изображение $z(i_1, i_2)$, исходное изображение $u(i_1, i_2)$ и шум $n(i_1, i_2)$. Размеры векторов z , u и n равны $L^{(u)} = L^{(z)} = L^{(n)} = (L_1^{(u)} + 1) \times (L_2^{(u)} + 1)$. Связь между лексикографически упорядоченными изображениями определяется соотношением

$$s = z + n, \quad (4.15)$$

где искаженное изображение

$$z = hu. \quad (4.16)$$

Символ h обозначает прямоугольную матрицу размером $L^{(z)} \times L^{(u)}$, с помощью которой вектор исходного изображения u преобразуется в искаженное изображение z . Матрица h имеет блочную структуру [4.2], элементы которой представляют собой отсчеты ФРТ. Задачи восстановления изображений алгебраическими методами при наличии и отсутствии шумов наблюдения имеют качественные различия.

Если шумами наблюдения можно пренебречь, то задача восстановления изображения сводится к нахождению оценки (решения) \hat{u} матричного уравнения (4.16), удовлетворяющей условию

$$z = h\hat{u}. \quad (4.17)$$

Если бы h была квадратной матрицей и существовала бы обратная матрица $(h)^{-1}$, то, очевидно, что решение системы имело бы вид

$$\hat{u} = (h)^{-1} z. \quad (4.18)$$

Однако матричное уравнение (4.16) представляет собой *недоопределенную* систему линейных алгебраических уравнений, т.к. количество неизвестных $L^{(u)}$ больше числа уравнений $L^{(z)}$ (размеры исходного изображения всегда больше размеров искаженного изображения). Поэтому матрица h является прямоугольной матрицей размером $L^{(z)} \times L^{(u)}$. В этом случае для отыскания решения используются различные методы псевдообращения матриц, которые описаны в [4.3]. Если недоопределенная система (4.16) разрешима, то она имеет несколько решений. Возникает проблема выбора единственного решения из множества возможных, которое и будет принято в качестве оценки \hat{u} . Среди всех возможных решений недоопределенной разрешимой системы (4.16) в качестве оценки \hat{u} выбирается решение, минимизирующее норму ошибки восстановления

$$\varepsilon_1^2 = e_1(\hat{u})^T e_1(\hat{u}) = \sum_{i=1}^{L^{(u)}} (u_i - \hat{u}_i)^2 = \min_{\hat{u}}, \quad (4.19)$$

где T - символ транспонирования; $e_1(\hat{u}) = u - \hat{u}$ - вектор ошибки восстановления. Критерий (4.19) называется критерием наименьших квадратов. Доказано [4.3], что норма ошибки будет минимальной, если оценка

$$\hat{u} = h^+ z,$$

где $h^+ = h^T (hh^T)^{-1}$ - обобщенная обратная матрица. В общем случае норма ошибки не равна нулю.

Точное восстановление исходного изображения при отсутствии шумов возможно, во-первых, когда искаженное изображение получено в результате циклической свертки исходного изображения и ФРТ. Во-вторых, когда объекты исходного изображения расположены в центре кадра и наблюдаются на фоне постоянной яркости, причем расстояние от объектов до границ кадра больше апертуры ФРТ. В том и другом случаях число неизвестных будет равно числу уравнений, т.к. объекты, расположенные вне кадра, не будут влиять на яркость наблюдаемого изображения. Иными словами, точное восстановление при отсутствии шума возможно тогда, когда ограничение размеров кадра наблюдаемого изображения не приводит к потере информации об исходном изображении.

Для искаженных изображений, наблюдаемых в присутствии шумов, к элементам вектора-столбца z добавляются отсчеты вектора-столбца n . Это делает систему уравнений, как правило, неразрешимой. Неразрешимость системы означает, что не существует оценки исходного изображения, при которой она перейдет в тождество. Можно найти лишь приближенное решение неразрешимой системы, которое определяется из условия минимума нормы ошибки [4.4, 4.5]

$$\varepsilon_2^2 = (e_2(\hat{u}))^T e_2(\hat{u}) = (s - h\hat{u})^T (s - h\hat{u}) = \min_{\hat{u}} L^{(s)}. \quad (4.21)$$

В этом случае оптимальным оператором (в смысле критерия наименьших квадратов (4.21)), формирующим оценку \hat{u} , также является обобщенная обратная матрица h^+ . Причем этот оператор является единственным оператором, обеспечивающим минимум нормы оценки $(\hat{u})^T \hat{u} = \sum_{i=1}^{L^{(s)}} \hat{u}_i^2$.

Таким образом, в обоих рассмотренных случаях обобщенное обращение матриц дает оптимальное решение, удовлетворяющее критериям наименьших квадратов (4.19) или (4.21). Следует подчеркнуть, что, несмотря на одинаковые названия, по сути это два разных критерия. Для разрешимой недоопределенной системы (4.16)) (когда выбирается одно решение из множества возможных) ошибка $e_2(\hat{u})$ равна нулю. В противном случае ошибка $e_2(\hat{u})$ всегда отлична от нуля, т.к. точное решение системы отсутствует.

Основным недостатком алгебраических алгоритмов восстановления изображений является необходимость выполнения трудоемких операций обращения, умножения и транспонирования матриц огромных размеров. Напомним, что размер матрицы h равен произведению числа отсчетов исходного и наблюдаемого изображений. Кроме того, обращение матриц больших размеров представляет собой трудную задачу. Эта задача значительно упрощается, если искаженное изображение формируется из исходного путем циклической свертки с ФРТ. К сожалению, для реальных задач восстановления изображений это условие не выполняется. Альтернативой алгебраическим методам являются методы линейной фильтрации изображений.

4.3. Методы восстановления изображений на основе пространственной фильтрации

Методы восстановления изображений, которые будут рассмотрены в данном разделе, реализуются с помощью ДПФ в частотной области. При этом обычная свертка заменяется циклической как в модели формирования искаженного изображения (4.4), так и в процедуре восстановления методом пространственной фильтрации. Все изображения $\tilde{s}(i_1, i_2)$, $\tilde{z}(i_1, i_2)$, $\tilde{n}(i_1, i_2)$,

$\tilde{u}(i_1, i_2)$ и ФРТ полагаются периодически продолженными и имеют одинаковые размеры кадра $L = L_1 \times L_2$. Замена обычной свертки циклической в модели формирования приводит к тому, что при синтезе восстанавливающего фильтра не учитывается факт влияния объектов, расположенных вне поля зрения объектива, на значение яркости на краях искаженного изображения. При восстановлении реально искаженных изображений таким фильтром возникают краевые эффекты, компенсация которых является одной из основных задач при реализации алгоритмов восстановления изображения на основе пространственной фильтрации.

При циклической свертке модель (4.4) формирования искаженного изображения определяется соотношением

$$\begin{aligned} \tilde{s}(i_1, i_2) &= \tilde{z}(i_1, i_2) + \tilde{n}(i_1, i_2) = \\ &= \tilde{h}(i_1, i_2) \otimes \otimes \tilde{u}(i_1, i_2) + \tilde{n}(i_1, i_2), \quad (i_1, i_2) \in \Theta, \end{aligned} \quad (4.22)$$

где Θ - кадр изображения, одинаковый для всех изображений и ФРТ, входящих в (4.22). Размеры кадра равны периоду повторения изображений и ФРТ.

Применяя к (4.22) ДПФ, получим

$$\begin{aligned} \tilde{S}(k_1, k_2) &= \tilde{Z}(k_1, k_2) + \tilde{N}(k_1, k_2) = \\ &= \tilde{H}(k_1, k_2) \tilde{U}(k_1, k_2) + \tilde{N}(k_1, k_2), \quad (k_1, k_2) \in \Theta. \end{aligned} \quad (4.23)$$

Система восстановления изображений на основе пространственной фильтрации представляет собой линейный пространственно-инвариантный двумерный фильтр. На выходе этого фильтра формируется оценка

$$\hat{u}(i_1, i_2) = \tilde{h}_e(i_1, i_2) \otimes \otimes \tilde{s}(i_1, i_2) \quad (4.24)$$

исходного изображения $u(i_1, i_2)$. В пространственно-частотной области спектр оценки с учетом (4.24) можно записать как

$$\begin{aligned} \hat{U}(k_1, k_2) &= \tilde{H}_e(k_1, k_2) \tilde{S}(k_1, k_2) = \\ &= \tilde{H}_e(k_1, k_2) (\tilde{H}(k_1, k_2) \tilde{U}(k_1, k_2) + \tilde{N}(k_1, k_2)). \end{aligned} \quad (4.25)$$

4.3.1. Инверсный фильтр

Простейшим способом восстановления четкости изображения является обработка наблюдаемого изображения в пространственно-частотной области инверсным фильтром [4.5]. Передаточная функция инверсного восстанавливающего фильтра определяется соотношением

$$\tilde{H}_{\text{inv}}(k_1, k_2) = 1 / \tilde{H}(k_1, k_2). \quad (4.26)$$

Она выбирается из условия $\tilde{H}_{\text{inv}}(k_1, k_2)\tilde{H}(k_1, k_2) = 1$, обеспечивающего компенсацию искажений, вносимых ФРТ формирующей системы. При этом спектр оценки исходного изображения равен

$$\hat{U}(k_1, k_2) = \tilde{U}(k_1, k_2) + \tilde{N}(k_1, k_2) / \tilde{H}(k_1, k_2). \quad (4.27)$$

Таким образом, восстановленное изображение равно сумме исходного изображения и шума наблюдения, прошедшего через инверсный фильтр. При отсутствии шума достигается точное восстановление инверсным фильтром исходного изображения $\tilde{u}(i_1, i_2)$ по искаженному изображению $\tilde{z}(i_1, i_2)$. При восстановлении изображений инверсным фильтром возникают краевые эффекты, которые проявляются в виде осциллирующей помехи большой мощности, полностью маскирующей восстановленное изображение. Краевые эффекты возникают даже при отсутствии шума наблюдения.

На рис. 4.13. и 4.14 приведены результаты восстановления изображений «Часы» и «Сатурн» инверсным фильтром: а) исходные изображения размером 200×200 элементов; б) дефокусированные изображения, полученные в результате свертки с гауссовским импульсом при $\alpha = 0.1$ с последующим «обрезанием» краев до размеров 170×170 элементов; в) изображения, восстановленные инверсным фильтром. Восстановить изображение «Часы» инверсным фильтром не удается из-за краевых эффектов. Практически идеальное восстановление изображения «Сатурн» объясняется тем, что объекты наблюдаются на фоне постоянной яркости и расположены в центре кадра. В этом случае изображения $z(i_1, i_2)$ и $\tilde{z}(i_1, i_2)$, полученные в результате обычной и циклической свертки с ФРТ, равны друг другу. Отметим, что при этих условиях алгебраический метод также позволяет точно восстановить изображение. Однако при инверсной фильтрации процедура обращения матриц заменяется на более простую процедуру перемножения массивов в частотной области.

На рис. 4.15 и 4.16 приведены сечения типичных частотных характеристик ФРТ и соответствующих им инверсных фильтров, из которых следует, что модуль передаточной функции формирующей системы, как правило, стремится к нулю на высоких частотах. Кроме того, нули в передаточной функции имеются в рабочей полосе частот при расфокусировке камеры (4.10) и смазе (4.6). В этом случае инверсный фильтр является сингулярным, т.к. модуль его

передаточной функции становится бесконечно большим на пространственных частотах, соответствующих нулевым значениям модуля передаточной функции искажающей системы. Причем наличие даже относительно слабого шума приводит к появлению интенсивных шумовых составляющих на выходе инверсного фильтра, полностью разрушающих изображение. Этот факт иллюстрируется рис.4.17. К дефокусированному изображению «Сатурн» (рис. 4.14.б) был добавлен аддитивный дельта-коррелированный шум. Из восстановленного изображения видно, что даже при пренебрежимо малом уровне шума (отношение сигнал/шум $q^2 = \frac{D_z}{D_n} = 10^{27}$) метод инверсной фильтрации дает очень плохой результат.

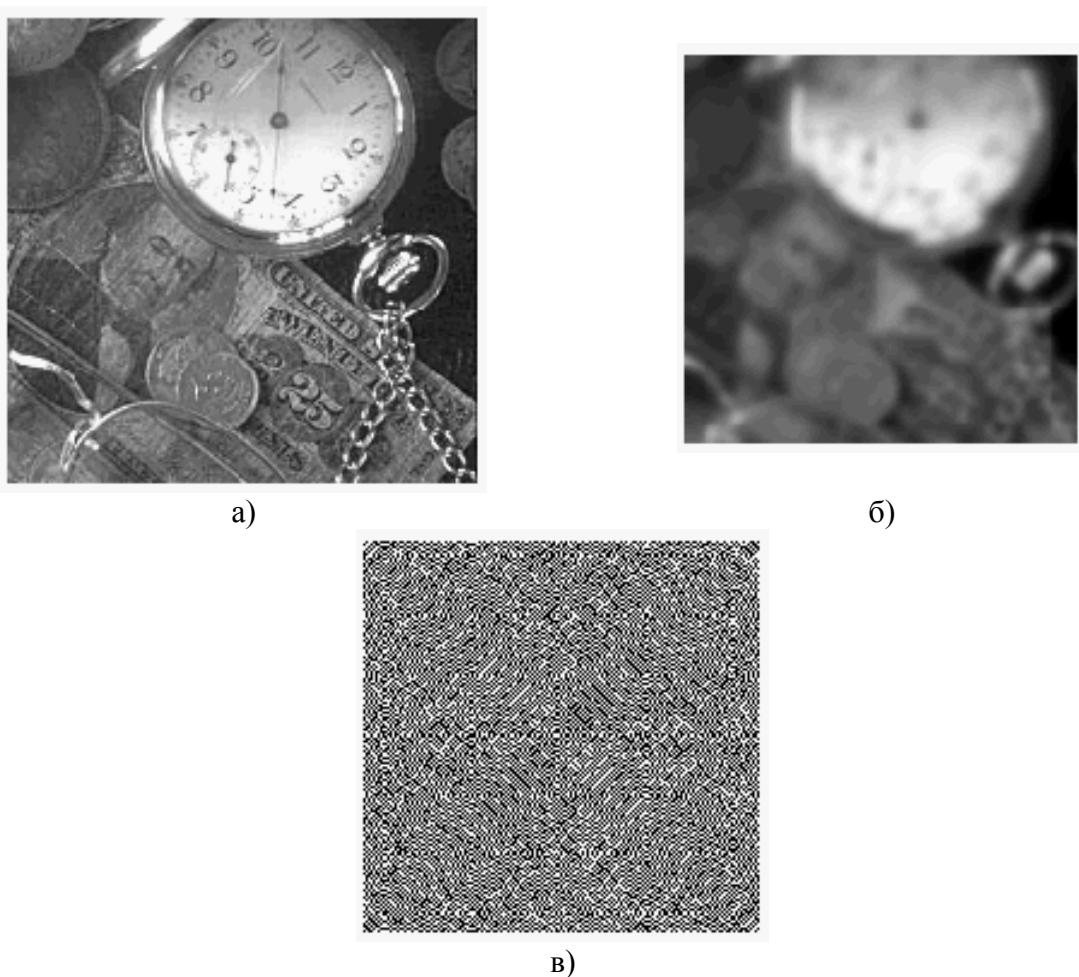


Рис.4.13. Результаты восстановления изображения “Часы”

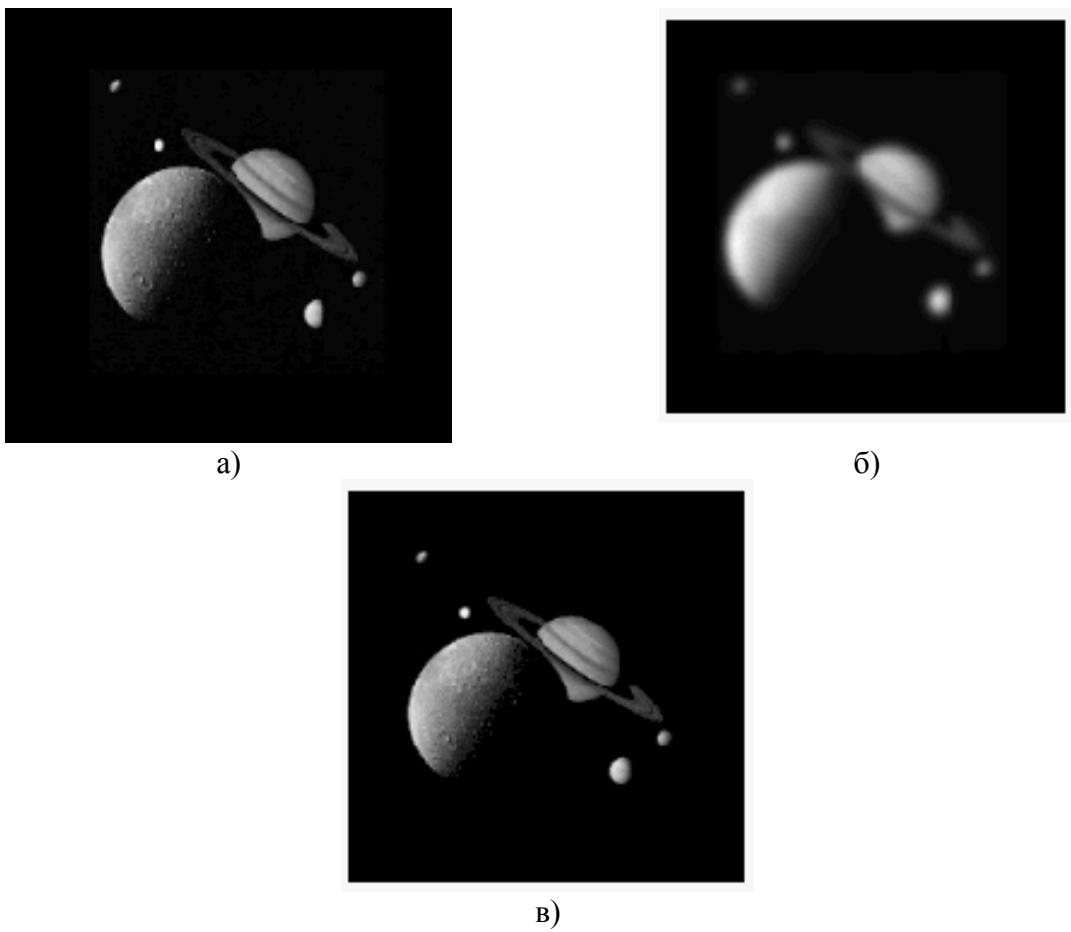


Рис.4.14. Результаты восстановления изображения “Сатурн”

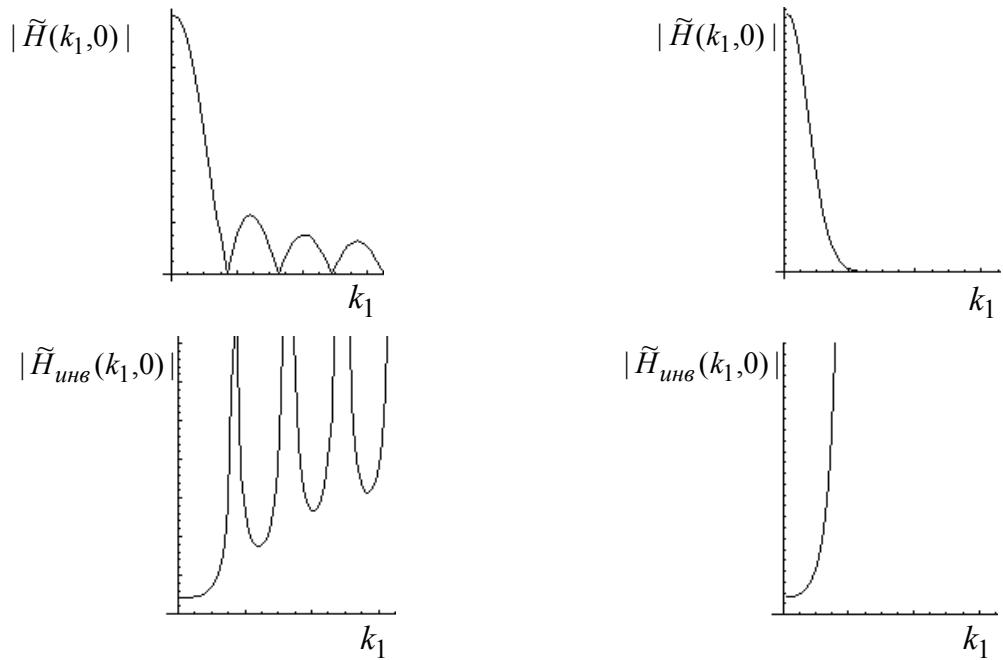


Рис.4.15. Частотные характеристики искажающей системы с цилиндрической ФРТ и инверсного фильтра

Рис.4.16. Частотные характеристики искажающей системы с гауссовской ФРТ и инверсного фильтра

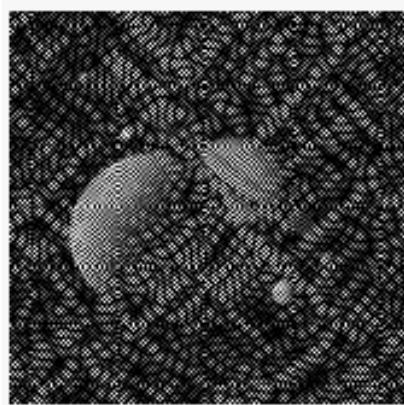


Рис.4.17. Результат восстановления изображения “Сатурн” при $q^2 = 10^{27}$

Существуют частные методы ослабления шумов, которые заключаются в ограничении полосы инверсного фильтра. Последовательно с инверсным фильтром включается корректирующее звено, модуль передаточной функции которого стремится к нулю за пределами некоторой наперед заданной граничной частоты. При этом граничная частота выбирается из компромисса между снижением уровня шума и четкостью восстановленного изображения. Однако эти методы не решают проблем краевых эффектов и наличия нулей передаточной функции формирующей системы в рабочем диапазоне частот.

Таким образом, несмотря на очевидную простоту метода инверсной фильтрации, он может успешно использоваться для восстановления ограниченного класса изображений, у которых уровень фона на краях постоянен. Кроме того, метод инверсной фильтрации обладает чрезвычайно низкой помехоустойчивостью.

4.3.2. Фильтр Винера

Инверсная фильтрация обладает низкой помехоустойчивостью, потому что этот метод не учитывает зашумленность наблюдаемого изображения. Значительно менее подвержен влиянию помех и сингулярностей, обусловленных нулями передаточной функции искажающей системы, фильтр Винера (смотри главу 3), т.к. при его синтезе наряду с видом ФРТ используется информация о спектральных плотностях мощности изображения и шума. При этом полагается, что изображение является реализацией случайного двумерного поля. Частотная характеристика восстанавливющего фильтра Винера, полученная для периодически продолженных изображений, с учетом (2.34) имеет вид [4.6]

$$\tilde{H}_{\text{Вин}}(k_1, k_2) = \frac{\tilde{G}_{u,s}(k_1, k_2)}{\tilde{G}_s(k_1, k_2)} = \frac{\tilde{H}^*(k_1, k_2)\tilde{G}_u(k_1, k_2)}{|\tilde{H}(k_1, k_2)|^2 \tilde{G}_u(k_1, k_2) + \tilde{G}_n(k_1, k_2)}, \quad (4.28)$$

где $\tilde{G}_n(k_1, k_2)$, $\tilde{G}_s(k_1, k_2)$, $\tilde{G}_u(k_1, k_2)$ - спектральные плотности мощности периодически продолженных шума, наблюдаемого и исходного изображений; $\tilde{G}_{u,s}(k_1, k_2)$ - взаимная спектральная плотность мощности периодически продолженных исходного и наблюдаемого изображений; $(\cdot)^*$ - символ комплексного сопряжения. Как и при инверсной фильтрации, обработка производится в частотной области.

Преобразуем передаточную функцию фильтра Винера (4.28) следующим образом:

$$\tilde{H}_{B_{in}}(k_1, k_2) = \frac{\tilde{H}^*(k_1, k_2)}{|\tilde{H}(k_1, k_2)|^2 + \frac{\tilde{G}_n(k_1, k_2)}{\tilde{G}_u(k_1, k_2)}} \quad (4.29)$$

Анализируя соотношения (4.28) и (4.29), можно отметить следующее:

1. При отсутствии шума фильтр Винера переходит в инверсный фильтр. Следовательно, в области низких частот, где, как правило, отношение сигнал/шум велико, передаточные функции инверсного и винеровского фильтров практически совпадают.
2. При уменьшении спектральной плотности мощности исходного изображения передаточная функция фильтра Винера стремится к нулю. Для изображений это характерно на верхних частотах.
3. На частотах, соответствующих нулям передаточной функции формирующей системы, передаточная функция фильтра Винера также равна нулю. Таким образом решается проблема сингулярности восстанавливающего фильтра.

На рис. 4.18 приведены одномерные сечения типичных передаточных функций винеровских фильтров (сплошная линия). Здесь же для сравнения приведены сечения передаточных функций инверсных фильтров (4.15) и (4.16), которые обозначены штриховой линией.

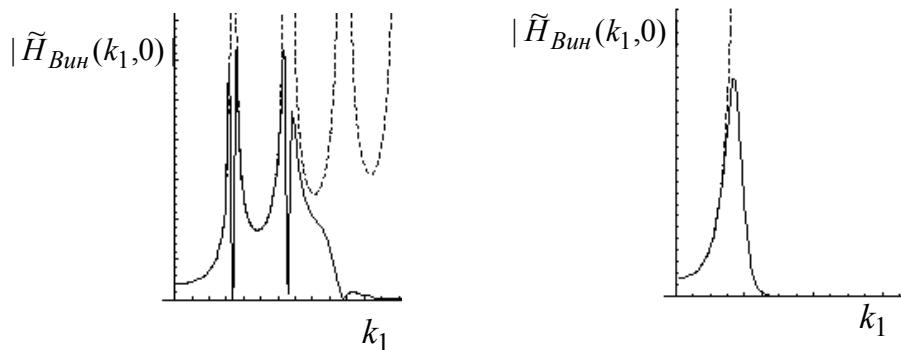


Рис.4.18. Частотный характеристики фильтра Винера при цилиндрической и гауссовской ФРТ

Рассмотрим результаты моделирования винеровского алгоритма восстановления. На рис. 4.19.а и 4.21.а приведены результаты искажения

изображений «Сатурн» и «Часы» сверткой с гауссовской ФРТ ($\alpha = 0.1$) с последующим «обрезанием» краев и добавлением аддитивного дельта-коррелированного шума ($q^2 = 10^4$). На рис. 4.20.а и 4.22.б приведены изображения, полученные в результате смазы ($L_{i_2}^{(h)} = 15$) изображений «Сатурн» (рис. 4.6) и «Часы» (рис. 4.22.а) ($L_{i_2}^{(h)} = 15$) также с последующим «обрезанием» краев и добавлением аддитивного дельта-коррелированного шума ($q^2 = 10^4$).

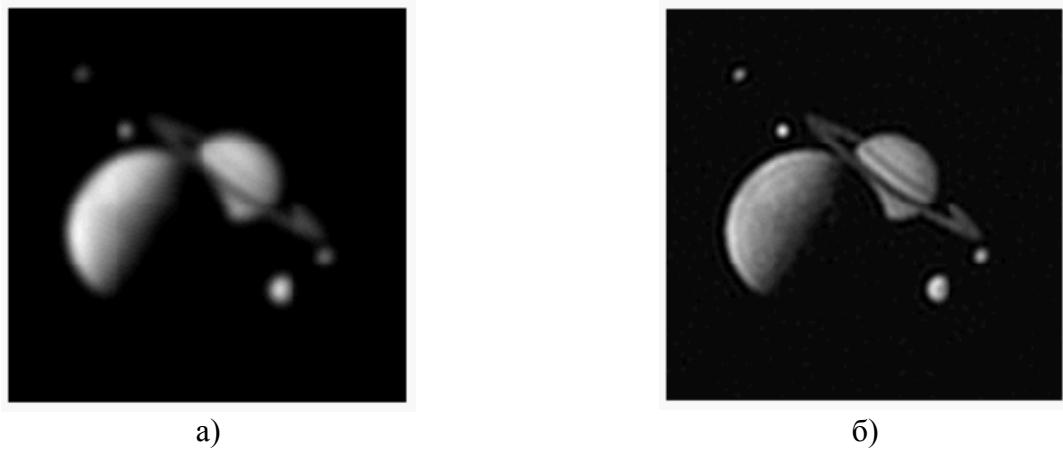


Рис.4.19. Восстановление дефокусированного изображения “Сатурн” при $q^2 = 10^4$

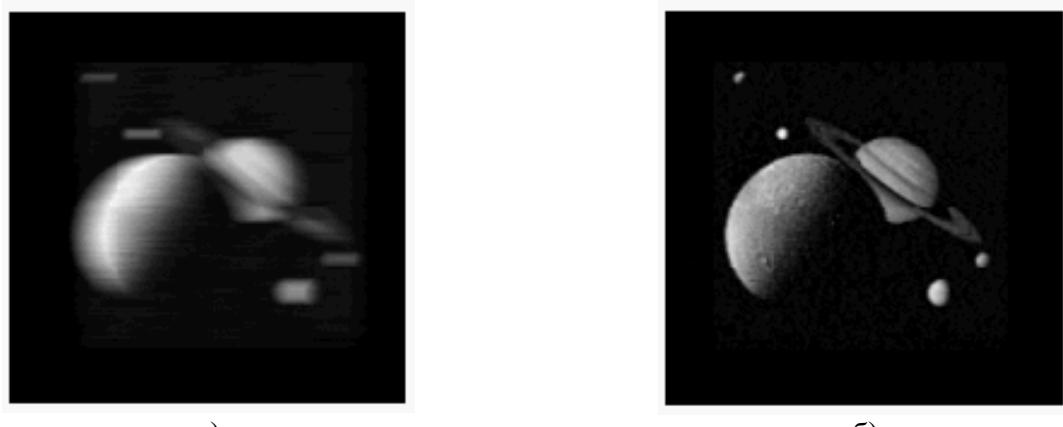


Рис.4.20. Восстановление смазанного изображения “Сатурн” при $q^2 = 10^4$

Размеры всех наблюдаемых и восстановленных изображений равны 170×170 элементов. Результаты восстановления винеровским фильтром изображения «Сатурн» (рис. 4.19.б и рис.4.20.б) свидетельствуют о том, что фильтр Винера значительно лучше подавляет шумы. Осциллирующая помеха на результатах восстановления изображения «Часы» (рис. 4.21.б и рис.4.22.в) вызвана краевыми эффектами. Очевидно, что ее уровень существенно меньше, чем при инверсной фильтрации (см. рис.4.13.в). Однако винеровский фильтр лишь

частично компенсирует краевые эффекты, которые делают качество восстановления неудовлетворительным.

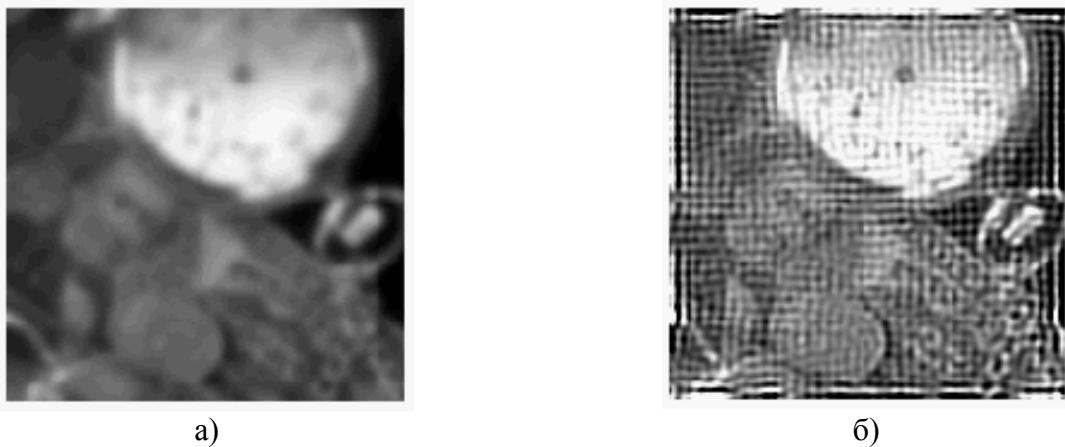


Рис.4.21. Восстановление дефокусированного изображения “Часы” при $q^2 = 10^4$

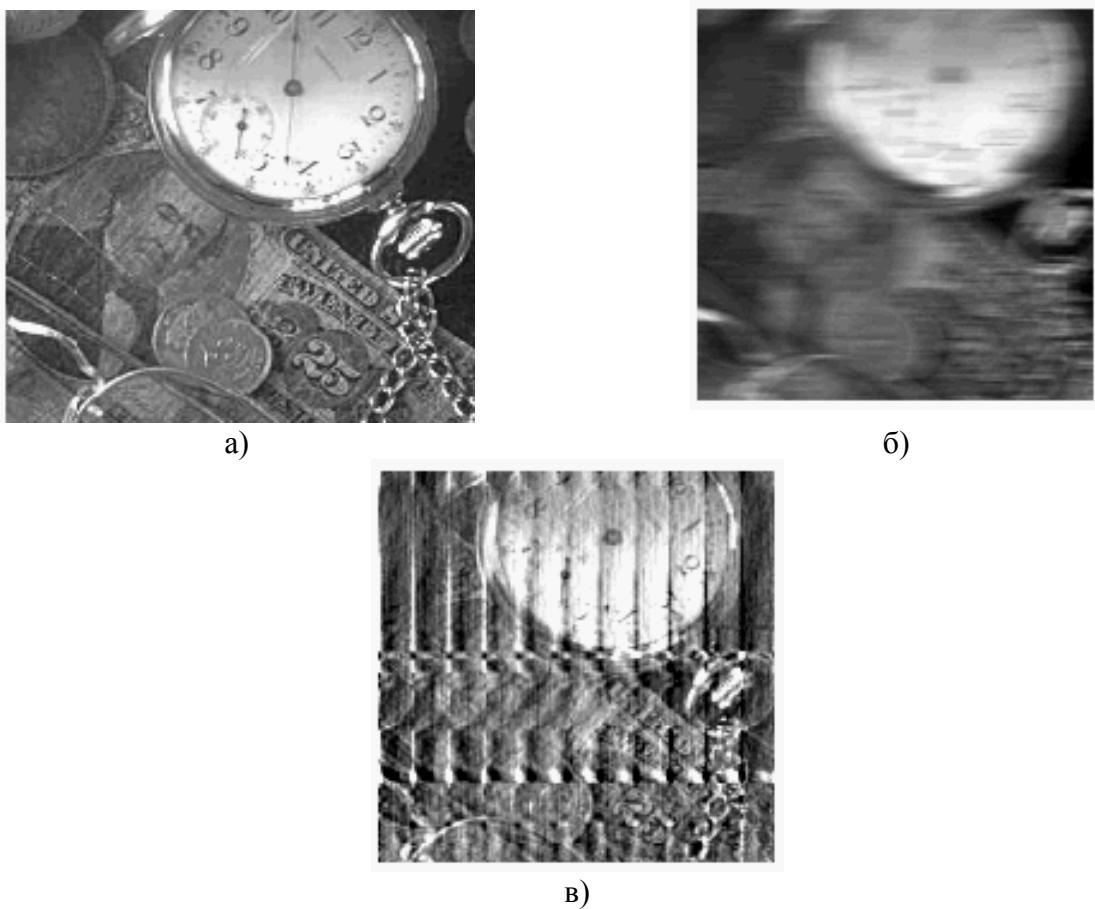


Рис.4.22. Восстановление смазанного изображения “Часы” при $q^2 = 10^4$

Таким образом, за счет использования информации о спектральных характеристиках изображения и шума, фильтр Винера обладает относительно высокой помехоустойчивостью и у него отсутствует сингулярность, обусловленная нулями передаточной функции формирующей системы. Основным недостатком фильтра Винера остается наличие краевых эффектов, которые проявляются в виде осциллирующей помехи, маскирующей восстановленное изображение.

4.3.3. Компенсация краевых эффектов при восстановлении линейно-искаженных изображений

На восстановленных изображениях, приведенных на рис.4.13.в, 4.21.б и 4.22.в, присутствует осциллирующая помеха большой интенсивности, которая возникает из-за того, что инверсный фильтр и фильтр Винера были синтезированы без учета ограниченных размеров наблюдаемых изображений. Вследствие того, что искаженное изображение записывается в кадре конечного размера, в усеченном изображении происходит потеря информации, содержащейся в исходном изображении вблизи границ. Поэтому при коррекции линейных искажений усеченного изображения возникают ложные детали в виде ряби или полос, интенсивность которых особенно велика при цилиндрической форме ФРТ и равномерном смазе.

К сожалению, решить уравнение Винера-Хопфа для сигналов и изображений, наблюдаемых на ограниченном интервале, не удается. Поэтому отсутствуют оптимальные пространственно-инвариантные фильтры, учитывающие краевые эффекты. Для компенсации краевых эффектов используются различные эвристические алгоритмы. Некоторые из них будут рассмотрены в данном подразделе.

Если нас интересует центральная часть изображения и его размеры значительно больше размеров кадра ФРТ, то для компенсации краевых эффектов применяют умножение наблюдаемого изображения на функцию окна $w(i_1, i_2)$, которая плавно уменьшается до нуля на краях кадра $\Theta^{(s)}$ наблюдаемого изображения и равны нулю всюду за его пределами. После этого изображение восстанавливается фильтром Винера.

При дефокусировке функция окна является разделимой относительно пространственных координат: $w(i_1, i_2) = w(i_1)w(i_2)$. При вертикальном или горизонтальном смазе используется одномерная функция окна, на которую умножаются соответственно столбцы или строки наблюдаемого изображения. Известно много одномерных функций окна $w(i)$, которые могут быть использованы при восстановлении изображений, например, окна Бартлетта, Кайзера, Блэкмана и др [4.7].

Хорошие результаты дает функция окна [4.8]

$$w_r(i) = (i | \gamma, \beta) = 0.5(\operatorname{th}((i + \frac{\gamma}{2})/\beta) - \operatorname{th}((i - \frac{\gamma}{2})/\beta)), \quad (4.30)$$

форма которой определяется двумя независимыми параметрами γ и β . Параметр γ влияет на размеры окна, а β - на скорость спада краев окна к нулю.

На рис 4.23 и 4.24 приведены результаты восстановления изображения «Часы» при горизонтальном смазе, где а - результаты умножения строк

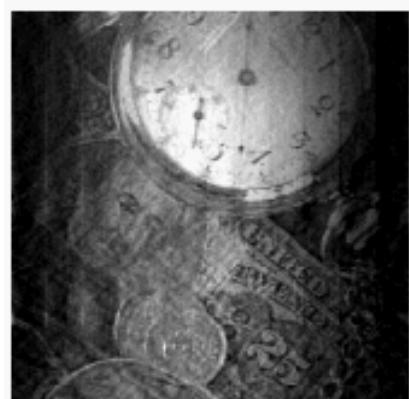
искаженного изображения, приведенного на рис.4.22.б, на окно Кайзера и окно (4.30); б - результаты восстановления фильтром Винера. Параметры окон подбирались, исходя из визуального качества восстанавливаемых изображений.

Уровень яркости на краях изображений, умноженных на окно, стремится к нулю, поэтому вместе с уменьшением краевых эффектов сужаются границы восстанавливаемого изображения. Кроме того, оптимальные параметры окон зависят от параметров искажающей системы и определяются опытным путем, что затрудняет практическое применение алгоритмов восстановления.

Учесть ограниченные размеры наблюдаемого изображения можно на этапе синтеза фильтра Винера, который использует информацию о спектрально-корреляционных характеристиках изображения. Получение изображения ограниченных размеров эквивалентно умножению бесконечного изображения на окно единичной яркости, размеры которого равны размерам кадра $\Theta^{(s)}$. Очевидно, что спектрально-корреляционные характеристики такого усеченного изображения будут отличаться от аналогичных характеристик бесконечных изображений. Корреляционная функция усеченного изображения может быть получена путем умножения на окно



а)

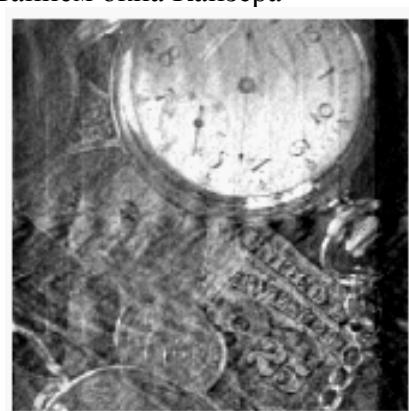


б)

Рис. 4.23. Восстановление с использованием окна Кайзера



а)



б)

Рис. 4.24. Восстановление с использованием окна (4.30)

$$\psi(i_1, i_2) = \begin{cases} \prod_{j=1}^2 \left(1 - \frac{|i_j|}{L_j^{(s)} + 1}\right), & \text{при } |i_j| \leq L_j^{(s)}, \\ 0, & \text{при } |i_j| \geq L_j^{(s)}, \end{cases} \quad (4.31)$$

корреляционной функции неограниченного изображения [4.9]. В этом случае спектральная плотность мощности усеченного изображения равна свертке спектральной плотности мощности неограниченного изображения и спектральной плотности окна (4.31). Подставляя соответствующие спектральные плотности мощности в уравнение Винера-Хинчина и решая его, получим коэффициент передачи фильтра для усеченного изображения [4.10]

$$H_{\overline{Bih}}(\omega_1, \omega_2) = \frac{G_{u,s}(\omega_1, \omega_2) \otimes \otimes \Psi(\omega_1, \omega_2)}{G_s(\omega_1, \omega_2) \otimes \otimes \Psi(\omega_1, \omega_2)}, \quad (4.32)$$

где $\Psi(\omega_1, \omega_2)$ - спектральная плотность окна (4.31). Следует подчеркнуть, что импульсная характеристика фильтра (4.32) не сводится к произведению импульсной характеристики фильтра Винера и регуляризирующего двумерного треугольного окна (4.31).

На рис. 4.25 приведен результат восстановления изображения «Часы» фильтром (4.32), откуда следует, что фильтр (4.32) практически полностью компенсирует краевые эффекты. Это позволяет отказаться от предварительной обработки. Качество восстановления изображения в центре и на краях почти одинаковое. Параметры фильтра (4.32) полностью определяются исходными данными и не требуют дополнительной подстройки. При использовании быстрого преобразования Фурье для обработки изображений объем вычислений при реализации фильтра (4.32) такой же, как и для фильтра Винера (4.28).

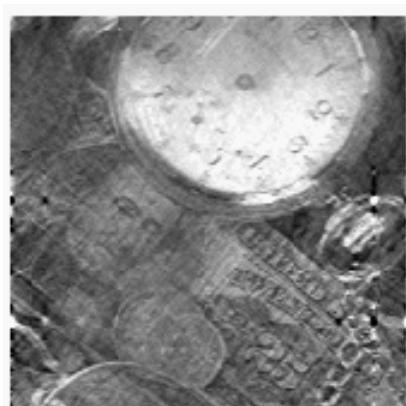


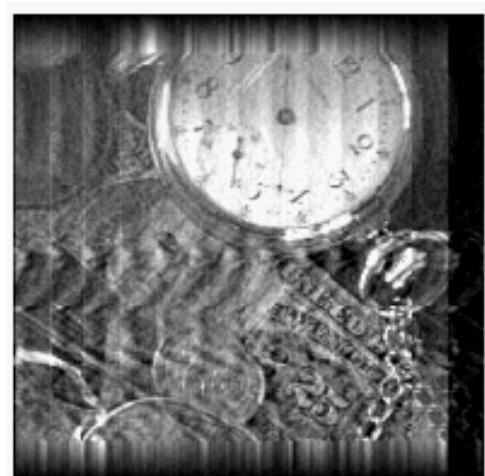
Рис.4.25. Результат восстановления с компенсацией краевых эффектов

Иногда наряду с компенсацией краевых эффектов требуется расширить границы восстанавливаемого изображения, чтобы извлечь больший объем информации об исходном изображении. Для этого используется процедура экстраполяции, которая состоит в том, что двумерную функцию яркости $s(i_1, i_2)$ наблюдаемого изображения продолжают с границ кадра $\Theta^{(s)}$ на кадр больших размеров $\Theta^{(\vartheta)}$ так, чтобы функция яркости была гладкой и на границах кадра $\Theta^{(\vartheta)}$ равнялась нулю.

Простейшей является процедура одномерной экстраполяции. Яркость изображения вдоль строк и столбцов за пределами кадра наблюдаемого изображения $\Theta^{(s)}$ задается в виде полинома

$$f(i_j) = a_0 + a_1 i_j + a_2 i_j^2 \dots, \quad j = \overline{1, 2}.$$

Коэффициенты $\{a_k\}$ определяются исходя из требований, предъявляемых к свойствам функции яркости. Например, на границе кадра $\Theta^{(\vartheta)}$ функция яркости должна равняться нулю, она должна быть неотрицательной, максимальное значение экстраполирующей функции не должно превышать максимального значения наблюдаемого изображения и т.п. Метод экстраполяции иллюстрируется рис.4.25, где а - экстраполированное изображение; б - результат восстановления. Размер наблюдаемого изображения «Часы» (см. рис. 4.22.б) равен 170×170 элементов, экстраполированного - 200×200 элементов. В качестве экстраполирующей функции использовался полином первой степени.



а) б)
Рис. 4.26. Восстановление с применением экстраполяции

Следует обратить внимание на то, что применение процедуры экстраполяции позволило восстановить изображение в пределах кадра исходного изображения, размеры которого больше размеров кадра

наблюдаемого изображения на величину смаза. Очевидно, что в центральной части качество восстановления при экстраполяции хуже, чем при умножении на окно. Однако эта процедура позволяет увеличить размеры кадра восстановленного изображения.

Улучшить качество восстановления можно, используя одновременно экстраполяцию наблюдаемого изображения и фильтра (4.32). Результат такой комбинированной процедуры приведен на рис 4.27. Рассмотренные методы восстановления являются линейными. Их широкое использование обусловлено достаточно простыми методами синтеза и анализа линейных систем, а также высокой вычислительной эффективностью. Однако эти методы не являются оптимальными и не всегда обеспечивают эффективную компенсацию искажений. Линейная обработка является лишь приближением к оптимальной обработке, т.к. статистические характеристики подавляющего большинства изображений являются негауссовскими. Кроме того, линейные методы не учитывают априорные данные о восстанавливаемых изображениях. Поэтому интерес представляют нелинейные методы обработки изображений. Синтез оптимальных нелинейных алгоритмов, как правило, значительно сложнее, чем линейных. Однако существуют линейные методы восстановления, которые достаточно просто могут быть преобразованы в нелинейные, учитывающие априорные данные об изображениях и помехах. Ярким примером таких методов являются итерационные методы (методы последовательных приближений).

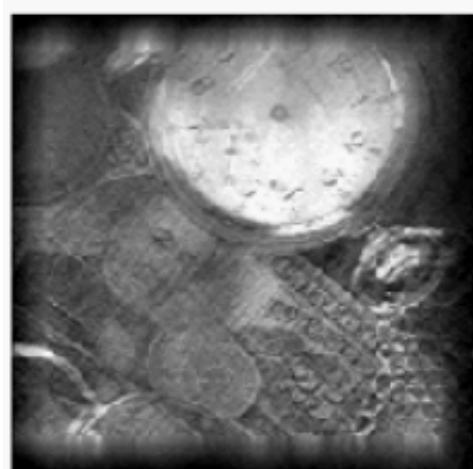


Рис.4.27. Восстановление с использованием экстраполяции и компенсации краевых эффектов

4.4. Итерационные методы восстановления изображений

Итерационными методами называют способы решения задач, в которых, выбирая некоторое начальное приближенное решение, вычисляют следующие, более точные приближения, используя предыдущие.

Рассмотрим один из способов построения итерационных процедур, основанный на разложении в ряд частотной характеристики инверсного фильтра [4.6]. Спектр оценки исходного изображения при инверсной фильтрации определяется соотношением

$$\hat{U}(\omega_1, \omega_2) = \frac{1}{H(\omega_1, \omega_2)} S(\omega_1, \omega_2). \quad (4.33)$$

Представим передаточную функцию инверсного фильтра $\frac{1}{H(\omega_1, \omega_2)}$ в виде геометрической прогрессии:

$$\frac{1}{H(\omega_1, \omega_2)} = \sum_{l=0}^{\infty} (1 - H(\omega_1, \omega_2))^l. \quad (4.34)$$

Подставляя (4.34) в (4.33), получим

$$\begin{aligned} \hat{U}(\omega_1, \omega_2) &= S(\omega_1, \omega_2) + (1 - H(\omega_1, \omega_2))S(\omega_1, \omega_2) + \\ &\quad + (1 - H(\omega_1, \omega_2))^2 S(\omega_1, \omega_2) + \dots \end{aligned} \quad (4.35)$$

Соотношение (4.35) позволяет представить процедуру нахождения оценки $\hat{U}(\omega_1, \omega_2)$ в виде последовательных приближений:

$$\hat{U}^{(0)}(\omega_1, \omega_2) = S(\omega_1, \omega_2); \quad (4.36)$$

$$\begin{aligned} \hat{U}^{(1)}(\omega_1, \omega_2) &= S(\omega_1, \omega_2) + (1 - H(\omega_1, \omega_2))S(\omega_1, \omega_2) \\ &= S(\omega_1, \omega_2) + (1 - H(\omega_1, \omega_2))\hat{U}^{(0)}(\omega_1, \omega_2); \end{aligned}$$

$$\begin{aligned} \hat{U}^{(2)}(\omega_1, \omega_2) &= S(\omega_1, \omega_2) + (1 - H(\omega_1, \omega_2))S(\omega_1, \omega_2) + (1 - H(\omega_1, \omega_2))^2 S(\omega_1, \omega_2) \\ &= S(\omega_1, \omega_2) + (1 - H(\omega_1, \omega_2))(S(\omega_1, \omega_2) + (1 - H(\omega_1, \omega_2))^2 S(\omega_1, \omega_2)) \\ &= S(\omega_1, \omega_2) + (1 - H(\omega_1, \omega_2))U^{(1)}(\omega_1, \omega_2); \end{aligned}$$

.....

$$\hat{U}^{(n)}(\omega_1, \omega_2) = S(\omega_1, \omega_2) + (1 - H(\omega_1, \omega_2))\hat{U}^{(n-1)}(\omega_1, \omega_2),$$

где каждое последующее приближение вычисляется по предыдущему. Взяв преобразование Фурье от соотношений (4.36), получим итерационную процедуру Ван Циттера [4.11]:

$$\begin{aligned}\hat{u}^{(0)}(i_1, i_2) &= s(i_1, i_2); \\ \hat{u}^{(n)}(i_1, i_2) &= s(i_1, i_2) + (1 - h(i_1, i_2)) \otimes \otimes \hat{u}^{(n-1)}(i_1, i_2),\end{aligned}\tag{4.37}$$

которую можно интерпретировать как процедуру последовательного нахождения поправок $\sigma^{(n)} = (1 - h(i_1, i_2)) \otimes \otimes \hat{u}^{(n-1)}(i_1, i_2)$ к искаженному изображению $s(i_1, i_2)$. Если в результате последовательных приближений на n -м шаге будет найдено точное решение : $\hat{u}^{(n)}(i_1, i_2) = u(i_1, i_2)$, то на последующих шагах, как нетрудно убедиться, оценка изменяться не будет. В итерационном алгоритме (4.37) нахождение обратного оператора заменяется на многократное вычисление свертки.

При использовании итерационных алгоритмов необходимо знать ответы на два вопроса - сходится ли он и, если сходится, то к какому решению. Сходимость алгоритма (4.37) к решению (4.33) определяется сходимостью ряда бесконечной геометрической прогрессии (4.34). Этот ряд сходится при $|1 - H(\omega_1, \omega_2)| < 1$, т.е. когда передаточная функция искажающей системы удовлетворяет условию

$$0 < H(\omega_1, \omega_2) < 1 .\tag{4.38}$$

Условие (4.38) выполняется для гауссовской ФРТ. При цилиндрической ФРТ и равномерном смазе соотношение (4.33) заменяют на эквивалентное соотношение

$$\hat{U}(\omega_1, \omega_2) = \frac{1}{|H(\omega_1, \omega_2)|^2} H^*(\omega_1, \omega_2) S(\omega_1, \omega_2).$$

Тогда итерационный алгоритм (4.37) имеет вид [4.6]

$$\begin{aligned}\hat{u}^{(0)}(i_1, i_2) &= h_1(i_1, i_2) \otimes \otimes s(i_1, i_2); \\ \hat{u}^{(n)}(i_1, i_2) &= h_1(i_1, i_2) \otimes \otimes s(i_1, i_2) + (1 - h_2(i_1, i_2)) \otimes \otimes \hat{u}^{(n-1)}(i_1, i_2),\end{aligned}\tag{4.39}$$

где $h_1(i_1, i_2)$ и $h_2(i_1, i_2)$ - импульсные характеристики фильтров с передаточными функциями $H^*(\omega_1, \omega_2)$ и $|H(\omega_1, \omega_2)|^2$ соответственно. Свертка в (4.37) и (4.39) может быть выполнена с помощь БПФ в предположении, что изображения и импульсные характеристики являются периодически продолженными.

Очевидно, что рассмотренный итерационный алгоритм является линейным и не имеет никаких преимуществ по сравнению с линейными алгоритмами. Однако этот метод позволяет эффективно бороться с краевыми эффектами и чрезмерным усилением шумов при восстановлении изображений. Итеративный процесс всегда можно остановить, если шум и осциллирующая помеха на изображении резко усиливаются. Остановка итеративного процесса означает усечение ряда (4.34), что приводит к ограничению коэффициента усиления за пределами некоторой граничной частоты. С увеличением длины ряда возрастают граничная частота и коэффициент усиления фильтра. Этот эффект иллюстрируется рис. 4.28, где приведены одномерные сечения частотных характеристик фильтров при 10-ти и 15-ти слагаемых в ряде (4.34) (сплошные линии). Здесь же для сравнения приведено одномерное сечение частотной характеристики инверсного фильтра (штриховая линия).

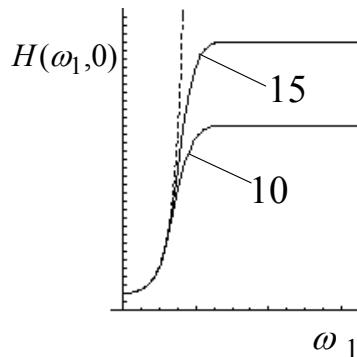


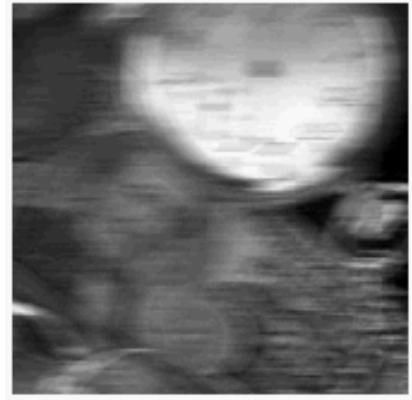
Рис. 4.28. Частотные характеристики итерационного фильтра на разных шагах

На рис 4.29 приведены результаты восстановления изображения «Часы», где а и б - повторно приведенные исходное (рис.4.22.а) и искаженное в результате смаза (рис 4.22.б) изображения; в - восстановленное изображение итерационным алгоритмом (4.37) (число итераций $n = 105$); г - результат восстановления по экстраполированному наблюдаемому изображению 4.26.а (число итераций $n = 135$). В качестве критерия остановки итеративного процесса использовался критерий минимума нормированной среднеквадратической ошибки оценивания:

$$\varepsilon^2 = \frac{\frac{1}{L_1^{(\varepsilon)} L_2^{(\varepsilon)}} \sum_{(i_1, i_2) \in \Theta^{(\varepsilon)}} (u(i_1, i_2) - \hat{u}(i_1, i_2))^2}{\frac{1}{L_1^{(\varepsilon)} L_2^{(\varepsilon)}} \sum_{(i_1, i_2) \in \Theta^{(\varepsilon)}} u(i_1, i_2)^2 - \left(\frac{1}{L_1^{(\varepsilon)} L_2^{(\varepsilon)}} \sum_{(i_1, i_2) \in \Theta^{(\varepsilon)}} u(i_1, i_2) \right)^2}, \quad (4.40)$$



а)



б)



в)



г)

Рис. 4.29. Восстановление изображения “Часы” итерационным алгоритмом

где кадр $\Theta^{(\varepsilon)}$ размером 150×150 элементов расположен в центре кадра $\Theta^{(s)}$ наблюдаемого изображения. Для рис. 4.29.в $\varepsilon^2 = 11\%$, а для рис.4.29.г $\varepsilon^2 = 6\%$. Таким образом, итерационный алгоритм практически полностью компенсирует краевые эффекты. Кроме того, применение экстраполяции позволяет снизить среднеквадратическую ошибку оценивания почти в 2 раза.

Наряду с описанными выше свойствами итерационные алгоритмы могут быть легко преобразованы в нелинейные путем введения нелинейных ограничений для восстанавливаемого изображения [4.6, 4.11]. Ограничения формулируются на основе априорных данных о форме или структуре объектов на исходном изображении. К априорным данным относятся такие свойства изображения, как неотрицательность яркости, ее верхний и нижний пределы,

минимальная мощность сигнала, ограниченная пространственная и спектральная протяженность и. т.п.

Даже учет такого простейшего ограничения как верхний и нижний пределы значений яркости приводит к значительному улучшению качества восстановления, т.к. среди всех возможных решений выбирается то, которое не имеет сильных осциляций яркости.

Итерационный алгоритм, например (4.39), с ограничением имеет вид

$$\hat{u}^{(0)}(i_1, i_2) = h_1(i_1, i_2) \otimes \otimes s(i_1, i_2); \quad (4.41)$$

$$\hat{u}^{(n)}(i_1, i_2) = h_1(i_1, i_2) \otimes \otimes s(i_1, i_2) + (1 - h_2(i_1, i_2)) \otimes \otimes \mathfrak{J}\{\hat{u}^{(n-1)}(i_1, i_2)\},$$

где $\mathfrak{J}\{\cdot\}$ оператор ограничения.

Например, если используется оператор ограничения на неотрицательность $\mathfrak{J}\{u(i_1, i_2)\} = \begin{cases} u(i_1, i_2); & u(i_1, i_2) \geq 0, \\ 0; & u(i_1, i_2) < 0, \end{cases}$, то из (4.41) следует,

что в тех областях, где яркость оценки $\hat{u}^{(n-1)}(i_1, i_2)$ меньше нуля, изменение оценки не происходит. Нелинейный итерационный алгоритм (4.41) будет сходится, если сходится линейный алгоритм (4.39) и оператор $\mathfrak{J}\{\cdot\}$ является нерасширяющимся оператором. Для пространства сигналов L_2 это означает, что действие оператора на изображение не должно приводить к увеличению его энергии. Очевидно, что к нерасширяющимся операторам относится оператор ограничения на неотрицательность, а также оператор ограничения диапазона, который определяется следующим соотношением

$$\mathfrak{J}\{u(i_1, i_2)\} = \begin{cases} u(i_1, i_2); & a \leq u(i_1, i_2) \leq b, \\ a; & u(i_1, i_2) < a, \\ b; & u(i_1, i_2) > b. \end{cases} \quad (4.42)$$

Для большинства цифровых изображений диапазон изменения яркости равен $[0,255]$. Использование нелинейного алгоритма с ограничением диапазона для восстановления изображения «Часы» (рис. 4.29.б) обеспечивает уменьшение среднеквадратической ошибки до 4%. Особенно эффективен этот алгоритм при восстановлении изображений с распределением яркости, близким к бинарному. На рис.4.30 приведены результаты восстановления изображения «Текст», где а - исходное изображение размером 98×113 эл.; б - часть исходного изображения, попадающая в кадра $\Theta^{(s)}$ размером 98×98 эл.; в - наблюдаемое изображение размером 98×98 эл., полученное в результате смазы исходного изображения ($L_{i_2}^{(h)} = 15$, $q^2 = 10^3$) ; г - экстраполированное

изображение размером 128×128 эл.; д и е - изображения, восстановленные по экстраполированному изображению итерационным линейным алгоритмом ($\varepsilon^2 = 11\%$) и итерационным алгоритмом с ограничением диапазона яркости ($\varepsilon^2 = 4.5\%$). Ошибка вычислялась по кадру $\Theta^{(\varepsilon)}$ размером 68×68 эл. На рис.4.30 размеры изображений увеличены в полтора раза.

Решение (4.20)
Но запись итер
Рассмотрим од
Этот оператор
Выражение (4

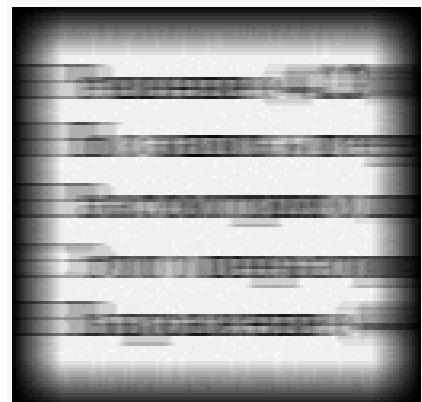
a)



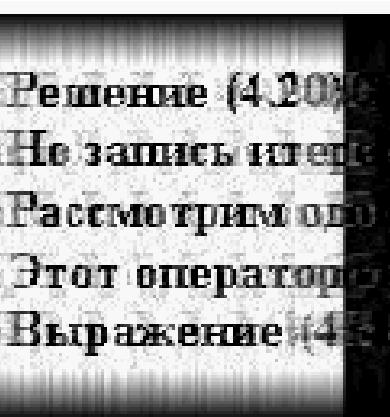
в)

Решение (4.20)
Но запись итер
Рассмотрим од
Этот оператор
Выражение (4

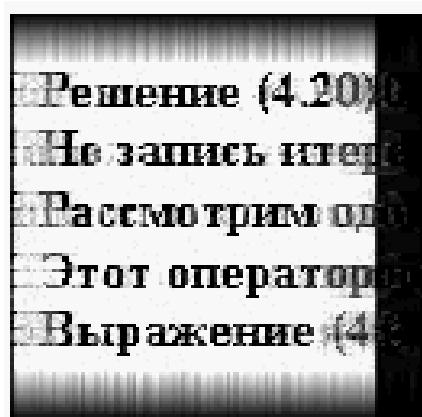
б)



г)



д)



е)

Рис. 4.30. Восстановление изображения “Текст” нелинейным итерационным алгоритмом

Таким образом, использование априорных данных о диапазоне изменения яркости позволило уменьшить среднеквадратическую ошибку почти в 2.5 раза.

Экстраполяция обеспечила восстановление изображения в пределах кадра исходного изображения, размеры которого больше размеров кадра $\Theta^{(s)}$ наблюдаемого изображения.

ВОПРОСЫ К ГЛАВЕ 4

- 4.1. Почему ухудшение четкости изображения описывается интегралом свертки исходного изображения и импульсной характеристики искажающей системы?
- 4.2. В чем принципиальное различие между расфокусировкой и смазом ?
- 4.3. Почему нельзя абсолютно точно восстановить изображения, сформированные реальными системами ?
- 4.4. Чем отличается задача восстановления изображений алгебраическим методам при наличии шума наблюдения от аналогичной задачи при его отсутствии ?
- 4.5. Назовите условия, при выполнении которых инверсная фильтрация обеспечивает высокое качество восстановления изображений.
- 4.6. Почему помехоустойчивость фильтра Винера выше чем инверсного ?
- 4.7. Назовите причины возникновения краевых эффектов и методы борьбы с ними.
- 4.8. Каковы достоинства и недостатки итерационных методов ?
- 4.9. Почему итерационный алгоритм восстановления при гауссовской ФРТ отличается от итерационного алгоритма при цилиндрической ФРТ ?

Глава 5

ГЕОМЕТРИЧЕСКИЕ ПРЕОБРАЗОВАНИЯ И ПРИВЯЗКА ИЗОБРАЖЕНИЙ

Многие задачи тематического дешифрирования сводятся к взаимному сопоставлению 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), что сводится к сопоставлению снимков одной и той же территории, полученных в разное время и/или с помощью различных датчиков. Чаще всего используются 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 интервалом, и поэтому, когда движущаяся камера фиксирует яркостный образ наблюдаемого объекта (оптическую поверхность) в виде последовательности изображений, то 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: установления точного координатного (поточечного) соответствия элементов стереоизображений. Решение этой задачи состоит в выделении пар реперных фрагментов и оценивании параметров «расхождения» соответственных точек (это именуется в стереофотограмметрии бинокулярной диспарантностью), 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 (то есть 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 \mathbf{x} , геометрически представляемый отрезком прямой линии из точки $(0,0)$ в точку (x, y) .

Двумерные преобразования на плоскости мы будем интерпретировать как движения точек по отношению к фиксированному базису (а не как изменение базиса, оставляющее точки неподвижными). В частности, нас особенно будут интересовать *линейные преобразования*, представляемые матрицами, то есть 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 случая произвольных элементов 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 всегда найдется 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, мы рассмотрим матрицы преобразования, связанные с наиболее важными подгруппами общей линейной

(или *проективной*) группы, а именно *евклидову* подгруппу, а также подгруппы *подобия* и *аффинную*. Это является следствием того, что евклидова геометрия (также как и аффинная) в действительности является подмножеством выше упомянутой нами 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, то она должна удовлетворять двум уравнениям

$$\begin{aligned} ax_1 + by_1 + 1 &= 0, \\ ax_2 + by_2 + 1 &= 0. \end{aligned}$$

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

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

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

$$\begin{aligned} a_1 x + b_1 y + 1 &= 0, \\ a_2 x + b_2 y + 1 &= 0. \end{aligned}$$

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

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

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 $a_2 = \lambda a_1, b_2 = \lambda b_1$.

5.1.2. 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:

$$\begin{pmatrix} \bar{x} \\ \bar{y} \\ w \end{pmatrix} = w \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}, \text{ так что } x = \bar{x}/w, y = \bar{y}/w. \text{ Совершенно очевидным свойством}$$

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

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

Поскольку скалярный множитель λ произвольный, то однородные координаты в действительности представляют линию, проходящую через начало координат в евклидовом пространстве. Прямые линии на плоскости также можно представить 3-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 $\mathbf{x} = (\bar{x}, \bar{y}, w)^t$, лежащие на однородной линии $\mathbf{p} = (\bar{a}, \bar{b}, c)^t$ определяются уравнением

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

Таким образом, точки и линии имеют здесь одинаковые представления. Нетрудно заметить, что прямым, проходящим через начало 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 движением плоскости соответствует евклидова подгруппа, содержащая лишь преобразования сдвига и поворота (рис.5.1), математически записываемых 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 φ 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 дополнительной единичной компоненты уравнение (5.2) можно переписать следующим образом:

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

Отметим далее, что два последовательно проведенные жесткие движения плоскости могут быть представлены единственным движением:

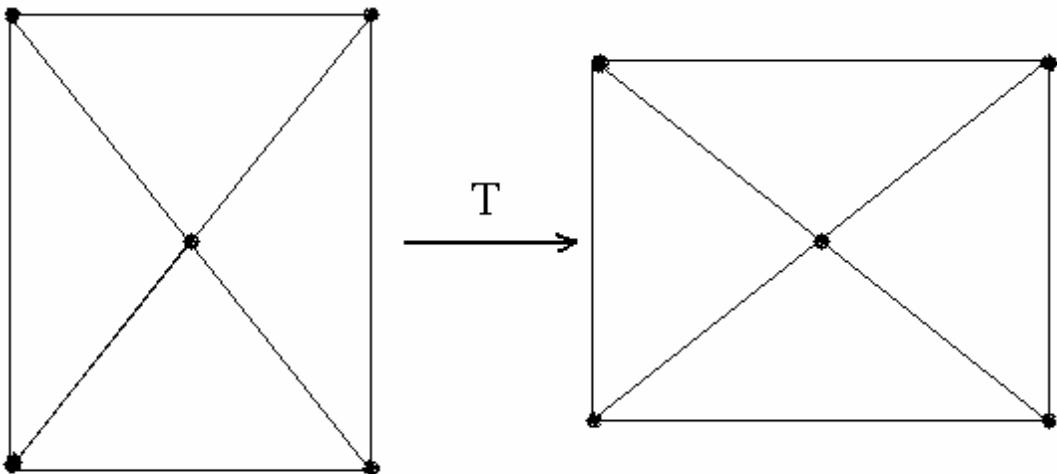


Рис.5.1. Действие евклидова преобразования на пять точек плоскости
(сдвиг, поворот)

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

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 $\mathbf{R}^t\mathbf{R} = I$, где I - единичная матрица.

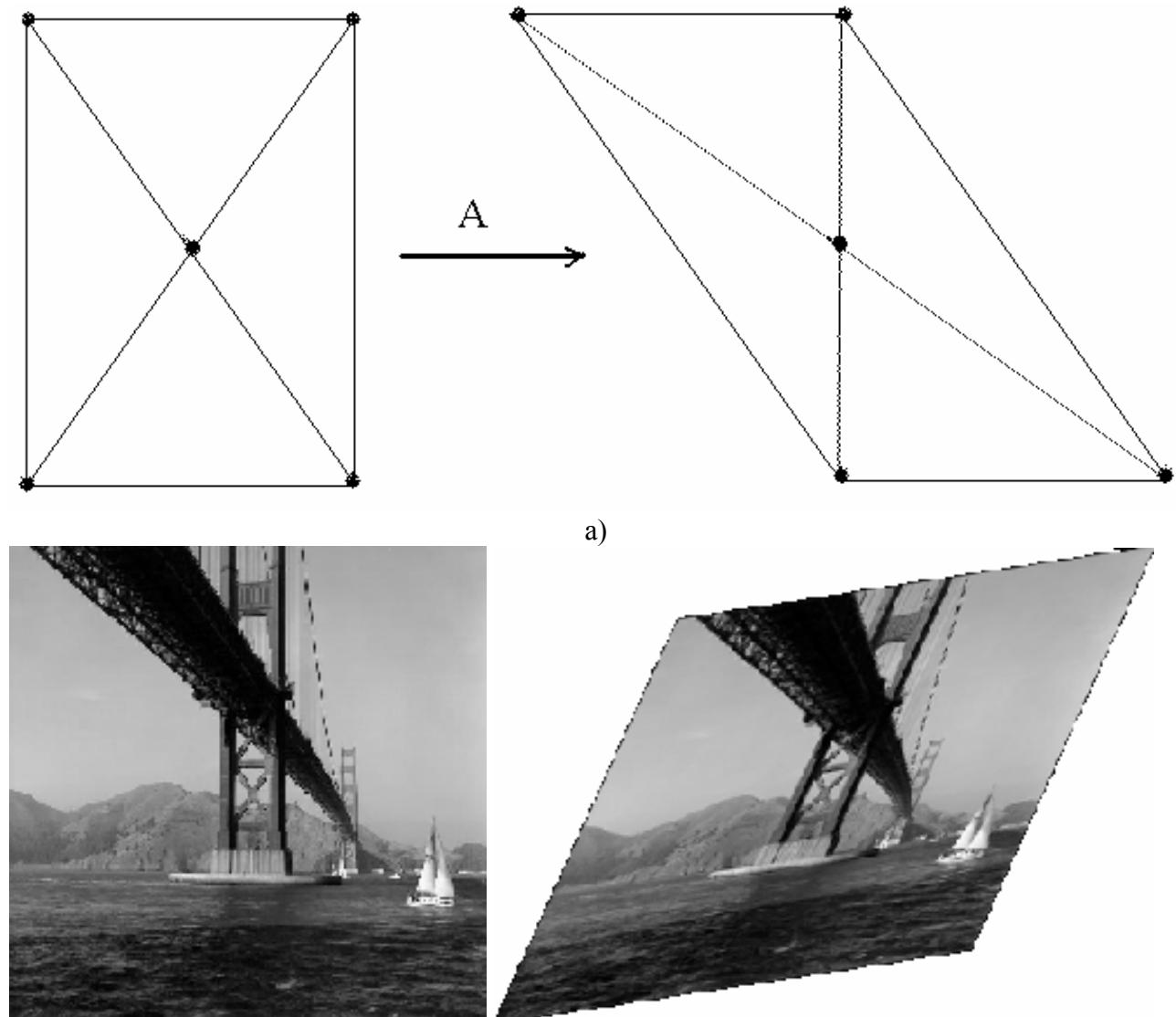
5.1.4. Аффинные преобразования

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

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



б)
Рис.5.2.

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

и в однородных координатах

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

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* (смотрите 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 переходят в параллельные.

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].

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

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

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

Растяжению вдоль соответствующей 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 φ , 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:

$$\begin{aligned}x' &= x + c, \\y' &= y + f.\end{aligned}$$

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) задается при помощи формулы:

$$\begin{aligned}x' &= x, \\y' &= -y.\end{aligned}$$

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 композции последовательно выполняемых простейших преобразований описанного 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 является частным случаем аффинной группы преобразований. Поэтому все отмеченные преобразования формируют иерархию в том смысле, что верно соотношение для их взаимной соподчиненности

евклидово преобразование \subset *аффинное* \subset *проективное преобразование*.

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

$$x = Ax' + By' + C, \\ y = Dx' + Ey' + F,$$

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

$$A = \frac{e}{\det \mathbf{A}}, B = -\frac{b}{\det \mathbf{A}}, C = \frac{(bf - ec)}{\det \mathbf{A}}, \\ D = -\frac{d}{\det \mathbf{A}}, E = \frac{a}{\det \mathbf{A}}, F = \frac{(cd - af)}{\det \mathbf{A}}.$$

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

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

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

5.1.5. Проективные преобразования

Как выше уже было сказано, общими линейными преобразованиями \mathbf{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 соотношение (5.7) будет иметь нелинейный вид, связанный с перенормировкой

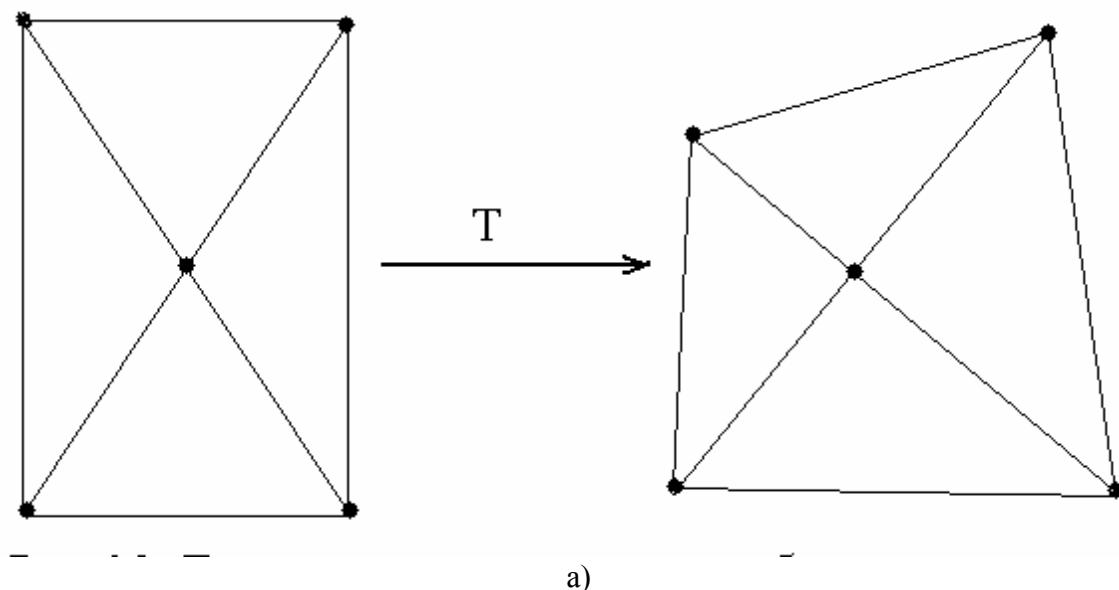
$$x' = \frac{t_{11}x + t_{12}y + t_{13}}{t_{31}x + t_{32}y + t_{33}}, y' = \frac{t_{21}x + t_{22}y + t_{23}}{t_{31}x + t_{32}y + t_{33}}.$$

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

Проективное преобразование связано с отображением трехмерной визуальной информации на двумерную плоскость. С математической точки зрения удобно рассматривать мир, включенным в трехмерное проективное пространство P^3 , а плоскость изображения, включенной в проективное пространство размерности два - P^2 . Точки на трехмерной сцене и на изображении представляются в проективных пространствах как векторы в однородных координатах.

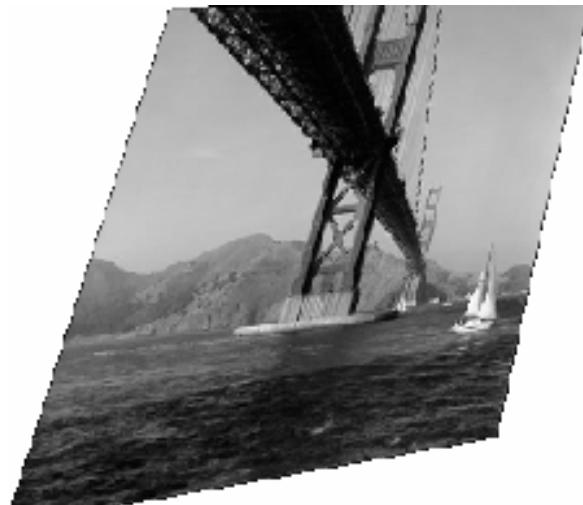
Проективное преобразование из P^3 в P^2 (перспективная проекция), отображающее евклидову точку сцены $\mathbf{p} = (X, Y, Z)^t$ в точку изображения $\mathbf{x} = (x, y)^t$ и выраженное в однородных координатах, задается в виде:

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





б)



в)



г)

Рис.5.3.

- а) действие проективного преобразования на пять точек плоскости;
 б) исходное изображение; в) и г) проективно преобразованные образы; параметры проективного преобразования соответственно:

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

и

$$x' = \frac{1.1x - 0.2y}{0.00075x + 0.00005y + 1}, y' = \frac{0.1x + 0.9y}{0.00075x + 0.00005y + 1}$$

Однородные координаты векторов проективного пространства $(\bar{X}, \bar{Y}, \bar{Z}, W)^t$ и проективной плоскости $(\bar{x}, \bar{y}, w)^t$ соотносятся с неоднородными (евклидовыми) координатами векторов \mathbf{p} и \mathbf{x} заданным выше образом: $\mathbf{p} = (\bar{X}/W, \bar{Y}/W, \bar{Z}/W)^t$ и $\mathbf{x} = (\bar{x}/w, \bar{y}/w)^t$.

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 является с необходимостью перспективной проекцией, но определяет проективное преобразование; то есть (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 $Z = f$. Простая геометрия показывает, что 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

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

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 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)

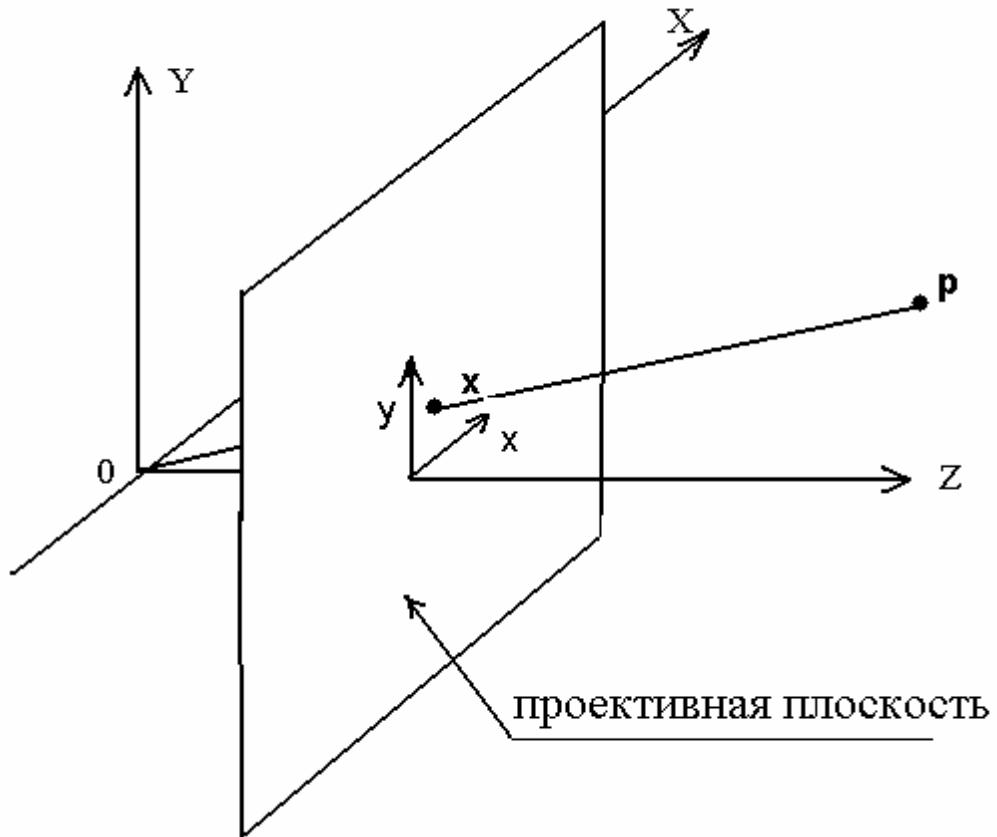


Рис.5.4. Перспективная проекция

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

Замечание. При $w=0$, 3D точка $(\bar{x}, \bar{y}, 0)$, в общем-то, определяет линию, параллельную плоскости $Z=1$, 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» (см. 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 (осами 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 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 состоит в том, что в первой допускаются неоднородное масштабирование и косоугольность.)

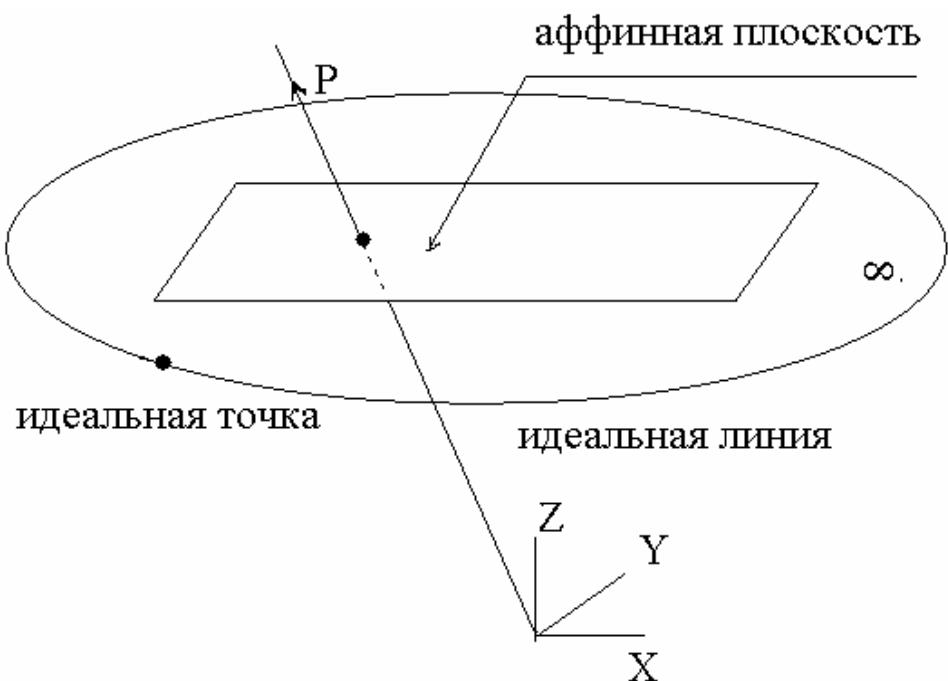


Рис.5.5. Проективная плоскость = аффинная плоскость + идеальные точки
(идеальная линия)

Замечание. Найдем 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, то есть $\frac{a_1}{b_1} = \frac{a_2}{b_2}$, 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

$$x_i = \frac{p_{11}X_i + p_{12}Y_i + p_{13}Z_i + p_{14}}{p_{31}X_i + p_{32}Y_i + p_{33}Z_i + p_{34}}, \quad (5.11)$$

$$y_i = \frac{p_{21}X_i + p_{22}Y_i + p_{23}Z_i + p_{24}}{p_{31}X_i + p_{32}Y_i + p_{33}Z_i + p_{34}}.$$

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} (то есть 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, 3×3 матрица, сформированная вектор-столбцами однородных координат любых 3 точек, должна иметь полный ранг. Легко проверить, например, что точки $\{(1,0,0)^t, (0,1,0)^t, (0,0,1)^t, (1,1,1)^t\}$ формируют так называемый *канонический базис*. В соответствии с приведенным выше *замечанием*, канонический базис содержит точки на бесконечности вдоль каждой координатной оси $\{(1,0,0)^t, (0,1,0)^t\}$, начало $(0,0,1)^t$ и единичную точку $(1,1,1)^t$. Очевидно, что коллинеация на P^2 : $\mathbf{x}' = \mathbf{T}\mathbf{x}$ целиком характеризуется геометрически ее действием на точки базиса.

Элементы 3×3 матрицы \mathbf{T} определены с точностью до масштабного множителя, и из них лишь восемь значений независимы. Поэтому, поскольку каждая точка содержит две независимых координаты, то 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 $(x, y, 0)^t$ должна быть трансформируема в точку $(\lambda x, \lambda y, 0)^t$ для произвольного масштабного множителя λ :

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

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 привлекли к рассмотрению геометрических преобразований, в общем-то, идеализированную модель 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 привлекают полиномиальную аппроксимирующую функцию преобразования (рис.5.6)

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

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



Рис.5.6. Действие билинейного геометрического преобразования на исходный снимок;
параметры преобразования:

$$x' = 0.1x + 0.9y + 0.002xy, y' = 0.2x + 1.1y + 0.0022xy$$

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 среднеквадратичную ошибку аппроксимации фактически «наблюдаемых» координат (x'_i, y'_j) их полиномиальной оценкой (\hat{x}'_i, \hat{y}'_j) из (5.12) для набора заданных узловых точек

$(x_i, y_j), i, j \subset (1, 2, \dots, N)$. 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

$$\mathbf{a} = (a_1, a_2, \dots, a_p)^t, \quad \mathbf{b} = (b_1, b_2, \dots, b_p)^t.$$

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

$$\mathbf{H} = \begin{pmatrix} 1 & x_1 & y_1 & .. & y_1^m \\ 1 & x_2 & y_1 & .. & y_1^m \\ 1 & x_2 & y_2 & .. & y_2^m \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_N & y_N & .. & y_N^m \end{pmatrix}.$$

Оценки наименьших квадратов (минимизирующих среднеквадратическую ошибку E^2) находятся 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

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

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

где $P_N^{(i)}(x)$ -попарно ортогональные на заданном множестве точек ортогональные 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

$$\begin{aligned} P_N^{(0)}(x) &= 1, P_N^{(1)}(x) = \lambda_N^{(1)} x, P_N^{(2)}(x) = \lambda_N^{(2)} [x^2 - (N^2 - 1)/12], \\ P_N^{(3)}(x) &= \lambda_N^{(3)} [x^3 - (3N^2 - 7)x/20], \\ \lambda_N^{(1)} &= [12/(N^2 - 1)]^{1/2}, \lambda_N^{(2)} = [4 \cdot 9 \cdot 5 / (N^2 - 1)(N^2 - 4)]^{1/2}, \\ \lambda_N^{(3)} &= [16 \cdot 25 \cdot 7 / (N^2 - 1)(N^2 - 4)(N^2 - 9)]^{1/2}. \end{aligned}$$

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

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

$$\mathbf{H}' = \left(\begin{array}{ccc|c} 1 & P_N^{(1)}(x_1) & P_N^{(1)}(y_1) & \dots & P_N^{(m)}(y_1) \\ 1 & P_N^{(1)}(x_2) & P_N^{(1)}(y_1) & \dots & P_N^{(m)}(y_1) \\ 1 & P_N^{(1)}(x_2) & P_N^{(1)}(y_2) & \dots & P_N^{(m)}(y_2) \\ \hline \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & P_N^{(1)}(x_N) & P_N^{(1)}(y_N) & \dots & P_N^{(m)}(y_N) \end{array} \right). \quad (5.18)$$

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

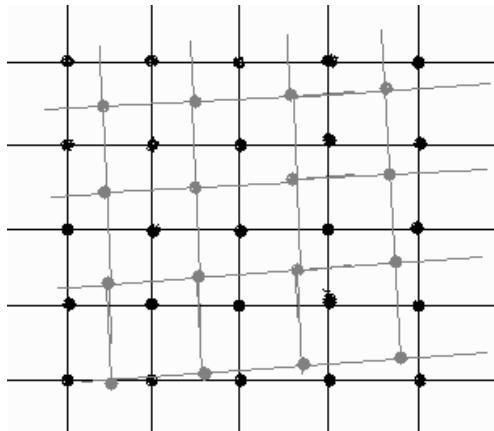


Рис.5.7. Вычисленные координаты (x', y') (выделены серым тоном), наложенные на исходную дискретную целочисленную решетку

освещенности. Вследствие этого изображения на снимках отличаются друг от друга геометрическимиискажениями. 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).

Восстановив уровни яркости наблюдаемых элементов в вычисленных точках на корректируемом снимке, то есть 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 (см. рис.5.7), то 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 функцией свертки

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

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 Ω есть ширина полосы частот $v(x, y)$. Из теоремы отсчетов следует (см. главу 1), 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, то 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, В-сплайн функции и т.п. [5.2, разд.4.3]. При выборе соответствующего ядра исходят из соображений как необходимой точности интерполяции, так и вычислительной эффективности. 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 не изотропно.

С вычислительной точки зрения предпочтителен алгоритм, известный как интерполятор по ближайшему соседу, где значение в точке (x, y) приписывается равным величине ближайшего отсчета дискретного растра. 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 растра (то есть вплоть до $\frac{h}{2}$). Треугольное ядро (рис.5.9) в двумерном случае приводит к билинейной интерполяции по четырем ближайшим соседям точки (x, y) : $\{(ih, jh)\}, [(i+1)h, jh], [ih, (j+1)h], [(i+1)h, (j+1)h]\}$.

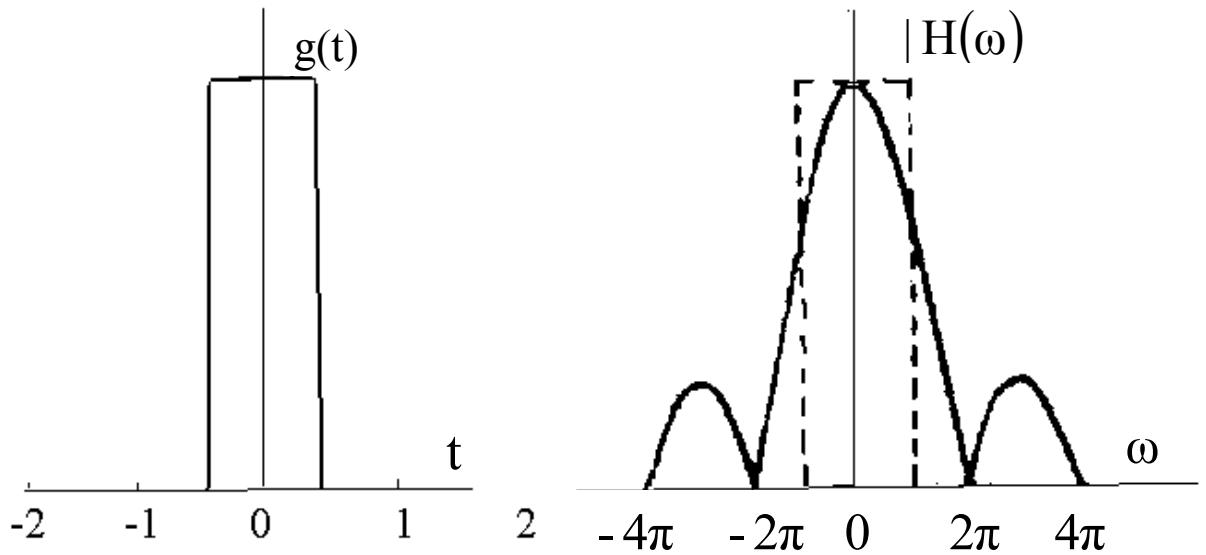


Рис.5.8. Интерполятор по ближайшему соседу с прямоугольным ядром. Справа график модуля Фурье-образа ядра. Пунктирной линией показан идеальный низкочастотный фильтр с частотой среза $(-\pi, +\pi)$.

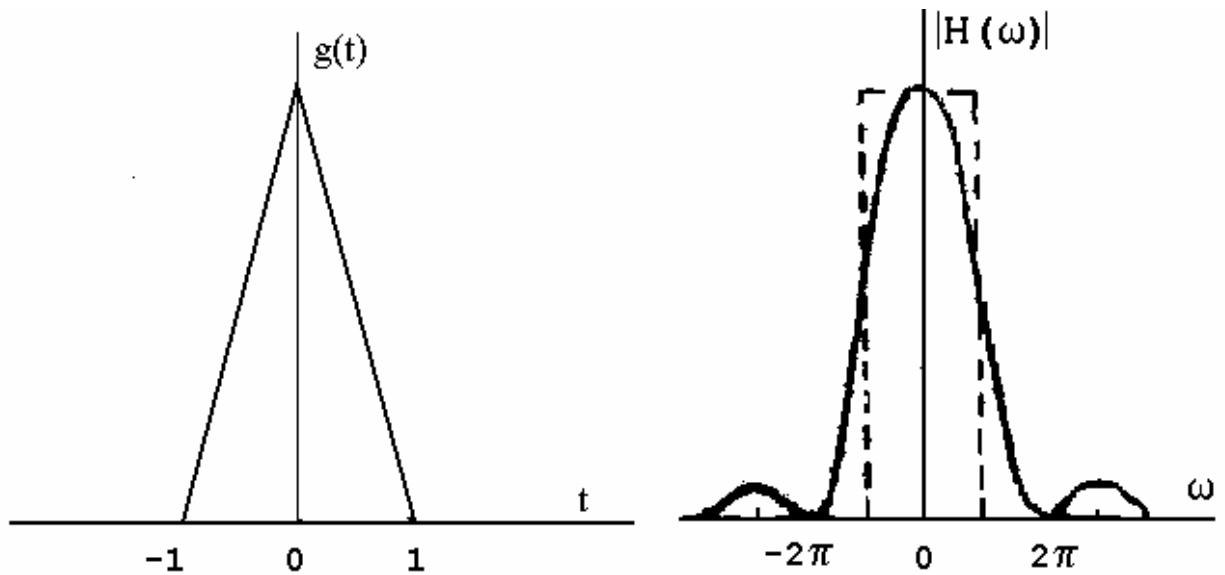


Рис.5.9. Линейная интерполяционная функция и модуль ее Фурье-образа (на правом рисунке пунктиром отмечен идеальный низкочастотный фильтр)

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 здесь 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 В-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 подинтервалах (-2,-1), (-1,0), (0,1), (1,2) по каждой из координат. Вне интервала (-2,2) интерполяционное ядро равно нулю. Двумерный кубический В-сплайн может быть записан как произведение двух одномерных интерполяционных функций по каждой из координат

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

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

и $g(s) = g(-s)$, то есть ядро симметричное.

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

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

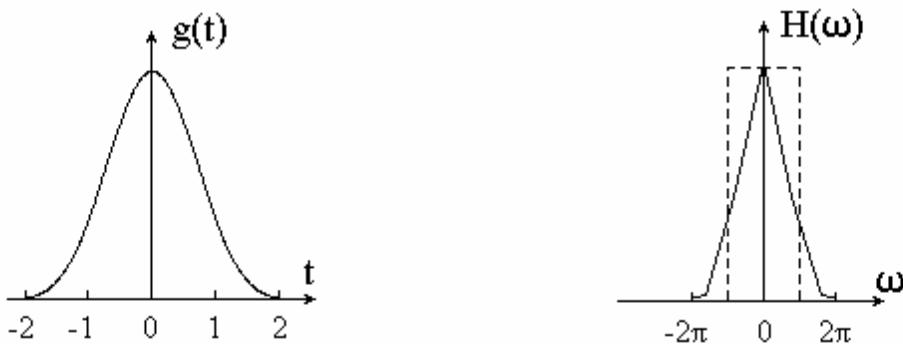


Рис.5.10. Кубический В-сплайн и модуль его Фурье-образа.

Имеется 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, тогда неизвестный параметр a должен быть равен $(-\frac{1}{2})$. 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 отрицательна, 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 В-сплайн, и было предложено называть данную функцию *высокоразрешающим интерполяционным кубическим сплайном*. Варьируя значением 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.

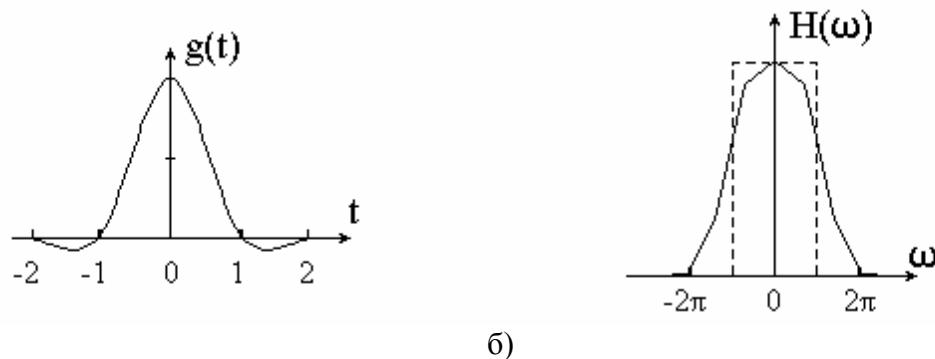
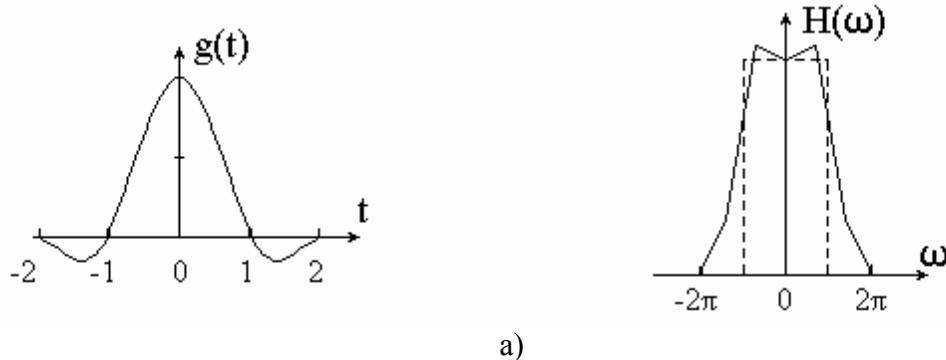


Рис.5.11. Высокоразрешающий интерполяционный кубический сплайн и его частотная характеристика : а) - $a=-1$; б) - $a= - 0.5$

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 А и представляемого матрицей U_0 размером $n \times n$), сравнивается с изображениями фрагментов снимка В в «зоне поиска» Ω 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 $\{u_0(\mathbf{x}), \mathbf{x} = (x, y)\}$ на А или $\{u(\mathbf{x})\}$ на В. В процессе скользящего поиска (когда каждый очередной фрагмент получается из предыдущего простым сдвигом на один дискрет) вычисляется «функция сходства» между изображением эталонного фрагмента $\{u_0(\mathbf{x}), \mathbf{x} \in \Gamma_A\}$ и изображениями текущих (контролируемых) фрагментов $\{u(\mathbf{x}), \mathbf{x} \in \Gamma_B\}$. 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

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

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;

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

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

$$\begin{aligned} \bar{u}(x, y) &= u(x, y) - \bar{u}, \bar{u} = \frac{1}{L^2} \sum_{(x, y) \in \Omega} u(x, y); \\ \bar{u}_0(x, y) &= u_0(x, y) - \bar{u}_0, \bar{u}_0 = \frac{1}{n^2} \sum_{x, y=1}^n u_0(x, y). \end{aligned}$$

В качестве меры различия в точке (k, l) будем брать среднеквадратичную ошибку

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

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

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

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 нормируем среднеквадратичную ошибку к энергии эталона

$$\bar{\varepsilon}_{\hat{a}}^2(k,l) = 1 - \frac{\left[\sum_x \sum_y \tilde{u}_0(x,y) \tilde{u}(x,y) \right]^2}{\sum_x \sum_y [\tilde{u}(x,y)]^2 \sum_x \sum_y [\tilde{u}_0(x,y)]^2},$$

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

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

Соблюдение условий достоверности обнаружения также приводит к необходимости установления порога для величины взаимной корреляции $\max \hat{r}(k,l)$: если $\max \hat{r}(k,l) \geq r_{nop}$, 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: $u_0(\mathbf{x}) = u(\mathbf{Ax} + \mathbf{t})$, 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 θ ; α - 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 θ и $|1 - \alpha|$ имеет вид

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

Было практически продемонстрировано, что серьезным недостатком корреляционной меры сходства является ее чувствительность к геометрическим искажениям видимых размеров сопряженных фрагментов при изменении ракурса съемки.

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

1. 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.
2. 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

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

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 $u_0(\mathbf{x})$: $u_0(\mathbf{x}) \approx u(\mathbf{A} \mathbf{x} + \mathbf{a})$. Это дает возможность повторным просмотром найти

уже существенно большее число пар сопряженных точек на исходном $u_0(\mathbf{x})$ и аффинно-преобразованном $u(\mathbf{A}^{(2)}\mathbf{x} + \mathbf{a}^{(2)})$ 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:

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

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

$$\begin{aligned}\frac{\partial \hat{r}_{0p}}{\partial x} |_{\hat{x}_{\text{ль}}, \hat{y}_{\text{ль}}} = & - \frac{\partial^2 \hat{r}_{0p}}{\partial x^2} |_{\hat{x}, \hat{y}} \delta x - \frac{\partial^2 \hat{r}_{0p}}{\partial x \partial y} |_{\hat{x}, \hat{y}} \delta y, \\ \frac{\partial \hat{r}_{0p}}{\partial y} |_{\hat{x}_{\text{ль}}, \hat{y}_{\text{ль}}} = & - \frac{\partial^2 \hat{r}_{0p}}{\partial y^2} |_{\hat{x}, \hat{y}} \delta y - \frac{\partial^2 \hat{r}_{0p}}{\partial x \partial y} |_{\hat{x}, \hat{y}} \delta x.\end{aligned}$$

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 $\mathbf{x}_1, \mathbf{x}_2$ в целом связываются соотношением

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

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 центрированных переменных $\tilde{u}_0(x, y), \tilde{u}(x, y)$ к спектральным компонентам (фурье-образам) изображений $\check{\mathbf{U}}, \check{\mathbf{U}}_0$:

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

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

Поскольку при смещении «образца» $\{\tilde{u}_0(x, y), x \leq n, y \leq n\}$ в области поиска Ω меняется только его фазовый спектр $\{\mathcal{F}(\check{\mathbf{U}}_0(x+k, y+l))\}_{pq} = s_{pq}(\check{\mathbf{U}}_0) \exp(i \frac{2\pi}{L} (pk + ql))$, то будем минимизировать функционал

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

по параметрам сдвига k, l , где звездочкой обозначена операция комплексного сопряжения, а функция h_{pq} осуществляет «взвешивание» разностной меры спектральных компонент. Так как «энергия» изображений фрагментов $|s_{pq}(\check{\mathbf{U}})|^2, |s_{pq}(\check{\mathbf{U}}_0)|^2$ не зависит от параметров k, l , то положение минимума

функционала F_{kl} соответствует точке экстремума перекрестного члена при раскрытии скобок в (5.32)

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

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

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

так как в этом случае обратное дискретное фурье-преобразование

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

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 $C_{kl}(\check{\mathbf{U}}, \check{\mathbf{U}}_0)$ есть фурье-образ функции когерентности

$$\gamma_{pq} = \frac{s_{pq}^*(\check{\mathbf{U}}_0) s_{pq}(\check{\mathbf{U}})}{\left| s_{pq}^*(\check{\mathbf{U}}_0) s_{pq}(\check{\mathbf{U}}) \right|}.$$

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.

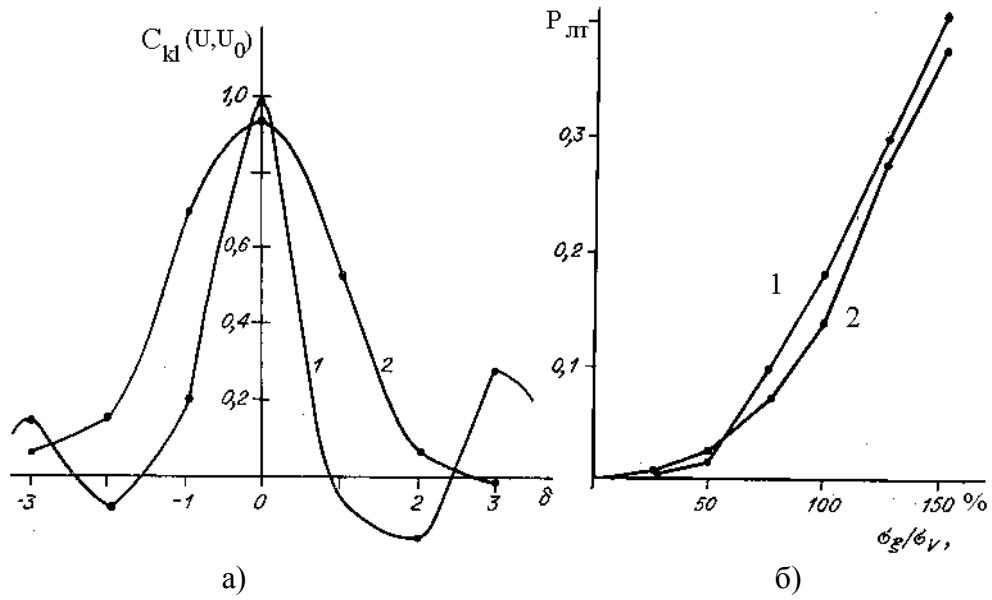


Рис.5.12. Меры эффективности алгоритмов идентификации сходства:

- 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 σ_ξ к амплитуде фонового сигнала (в %): 1 - ВЛП в фазовом алгоритме, 2 - в корреляционном алгоритме.

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 кросскорреляционной функции сходства $C_{kl}(\bar{U}, \bar{U}_0)$, 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° ($n = 18, L = 25$). Алгоритм фазовой корреляции здесь нормально функционирует, если угол поворота не превышает 5° . 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 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 здесь представлен результат привязки радиолокационных снимков.

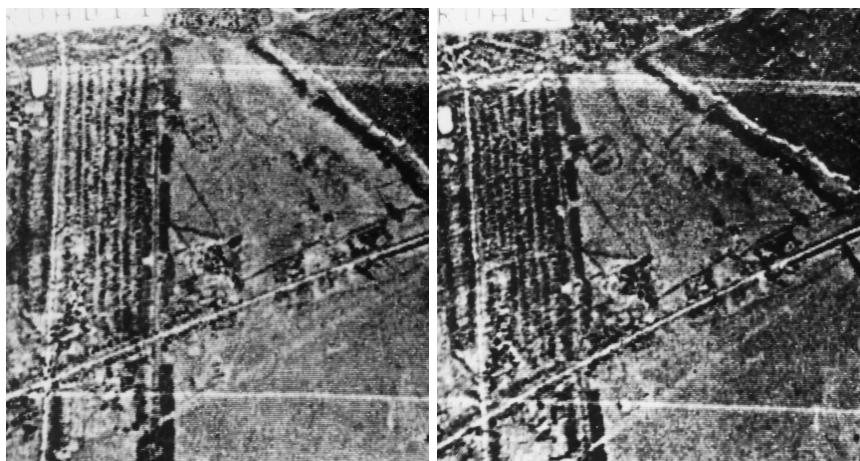


Рис.5.13. Радиолокационные изображения местности, снятые при двух пролетах самолета (разные ракурсы съемки)

5.3.4. Привязка по локальным неоднородностям

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

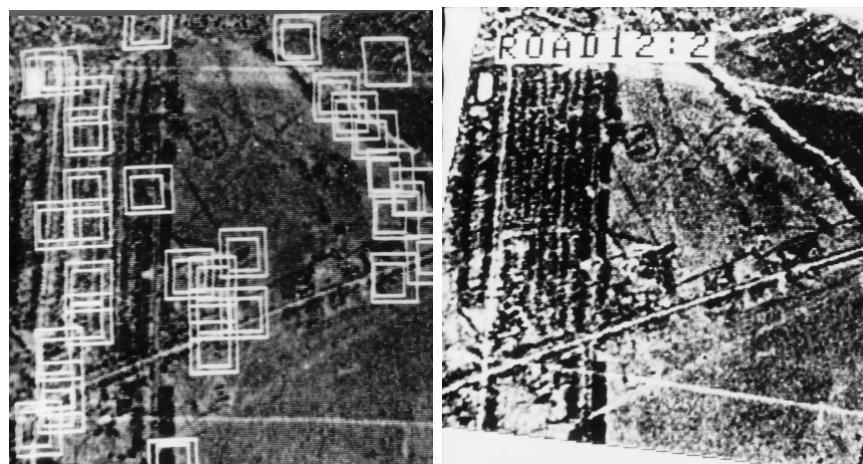


Рис.5.14. Результаты привязки изображений

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 параметрами, 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 критерий сходства не регистрировал 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. Выделение краев [5.2, гл.17] можно рассматривать как средство быстрого предварительного просмотра эталонного изображения и отбора информативных областей. В полученных областях далее выбираются 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 φ (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.

Глава 6

ФОТОГРАММЕТРИЯ И СТЕРЕОВИДЕНИЕ

В обычном изображении трехмерного пространства информация о расстоянии до различных элементов сцены проявляется только в виде косвенных признаков: через относительные размеры объектов, затенение одних объектов другими, различной освещенностью и т.д. Один из способов получения информации о глубине состоит в регистрации нескольких изображений сцены под различными ракурсами. В этом случае точки сцены дают изображения, относительное положение которых зависит от расстояния до точки наблюдения. Оказывается, сопоставляя эти изображения, в ряде случаев можно реконструировать трехмерную структуру сцены. Основы теории определения положения объектов в пространстве по их перспективным изображениям были положены еще в средние века, а в XVIII веке перспективные рисунки стали использоваться в топографических целях. Появление фотографии положило начало фотограмметрии – науке об определении формы, размеров и пространственного положения различных объектов посредством измерения их фотографических изображений. В начале XX века был изобретен стереокомпаратор – прибор для измерения пространственного положения объектов по паре перекрывающихся фотографических изображений. Интерес к методам восстановления трехмерной структуры сцен по их плоским изображениям возобновился в середине XX столетия в связи с исследованиями в области искусственного интеллекта, а практическая потребность в робототехнических устройствах, способных ориентироваться в трехмерном пространстве, постоянно поддерживает этот интерес в последние десятилетия.

В этой главе мы рассмотрим процесс формирования изображений, соотношения между координатами точек сцены и их изображениями и методы оценивания параметров системы регистрации и трехмерной структуры сцены.

6.1. Модель регистрирующей камеры

Строго говоря, различные точки пространства предметов отображаются оптической системой камеры в пространстве изображений на различных расстояниях от фокальной плоскости. Однако, если расстояние между камерой и наблюданной сценой значительно превышает фокусное расстояние оптической системы, можно считать, что изображение строится в ее фокальной плоскости. В этом случае можно воспользоваться проективной моделью камеры, в которой изображение трехмерного объекта получается проектированием его в фокальную плоскость (плоскость изображения) через единственную точку, называемую оптическим центром. Прямая линия, перпендикулярная плоскости изображения и проходящая через эту точку,

называется оптической осью камеры, а точка пересечения оптической оси с плоскостью изображения – главной точкой.

Определим в трехмерном пространстве ортогональную правую систему координат OXYZ, начало которой совпадает с оптическим центром, ось OZ - с оптической осью камеры. Такая система называется *стандартной системой координат* камеры. Пусть плоскость изображения находится на расстоянии f от оптического центра. В этой плоскости зададим систему координат oxy с началом в главной точке и осями ox и oy, параллельными осям OX и OY соответственно (рис. 6.1). Легко убедиться, что в стандартной системе координат проекцией точки трехмерного пространства M с координатами (X, Y, Z) является точка m в плоскости изображения с координатами (x, y) , причем

$$x = fX / Z, \quad y = fY / Z.$$

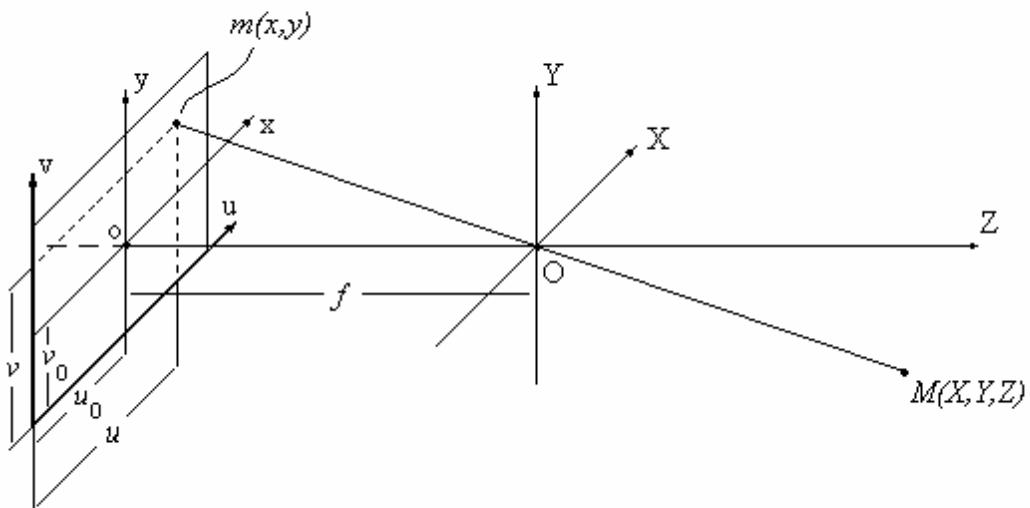


Рис.6.1. Система координат проективной камеры

Для полного описания камеры следует учесть, что для регистрации изображения в плоскости изображения камеры помещается какой-либо фотоприемник. В фотокамере это фотопластиинка или фотопленка, в видеокамеру это видикон или фотоприемная полупроводниковая матрица, в некоторых специальных камерах – электромеханический сканер. В общем случае измерение координат в фотоприемнике осуществляется в единицах, отличных от единиц, задающих координаты в стандартной системе. Например, если используется матричный фотоприемник, его естественными координатами являются номер строки и номер столбца фотоприемной ячейки. Поэтому для полного описания камеры необходимо выразить координаты точки m в

естественных единицах фотоприемника. В достаточно общем для любых фотоприемников виде (рис. 6.1) это может выглядеть как

$$u = x/w + u_0, \quad v = y/h + v_0,$$

где (u_0, v_0) - координаты главной точки относительно начала координат фотоприемника (в естественных координатах фотоприемника); w и h - масштабы вдоль осей ох и oy (например, расстояния между ячейками матричного фотоприемника вдоль строк и столбцов).

В новой системе координаты проекции точки m примут вид

$$u = \frac{fX}{wZ} + u_0, \quad v = \frac{fY}{hZ} + v_0. \quad (6.1)$$

Для последующего изложения введем трехмерный вектор $\mathbf{M} = (X, Y, Z)^T$, соответствующий точке M , и двумерный вектор $\mathbf{m} = (x, y)^T$, соответствующий точке m . Определим также вектор однородных (см. гл. 5) внутренних координат камеры $\mathbf{v} = (u, v, 1)^T$. Используя эти обозначения, соотношения (6.1) можно представить в компактной векторно-матричной записи:

$$Z\mathbf{v} = \mathbf{AM}, \quad (6.2)$$

где $\mathbf{A} = \begin{bmatrix} f/w & 0 & u_0 \\ 0 & f/h & v_0 \\ 0 & 0 & 1 \end{bmatrix}$ - матрица, известная под названием матрицы

внутренних параметров камеры, поскольку она содержит только параметры оптической системы и фотоприемника камеры.

6.2. Связь между различными системами координат

В общем случае трехмерные координаты точки могут быть заданы в системе, не совпадающей со стандартной системой координат камеры (назовем ее *глобальной*). Пусть OXYZ – глобальная система координат, а O'X'Y'Z' – стандартная система координат камеры. Переход от системы OXYZ к системе O'X'Y'Z' можно осуществить поворотом координатных осей к системе OX"Y"Z" и последующим смещением начала координат. Тогда связь между координатами точки M в глобальной и стандартной системе может быть представлена как

$$\mathbf{M}' = \mathbf{RM} + \mathbf{t}, \quad (6.3)$$

где \mathbf{M} и \mathbf{M}' - векторы пространственных координат точки M в глобальной и стандартной системах, соответственно; \mathbf{R} - матрица размерности 3×3 , описывающая поворот стандартной системы координат относительно глобальной; компонентами матрицы являются направляющие косинусы осей глобальной системы в стандартной системе координат ; \mathbf{t} - трехмерный вектор смещения начала координат глобальной системы относительно начала координат стандартной.

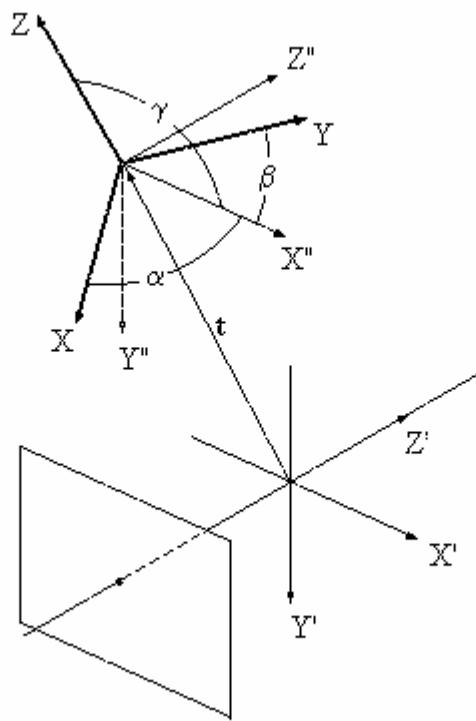


Рис.6.2. Переход от глобальной системы координат к стандартной системе координат камеры.

На рис. 6.2. схематически показано преобразование координат. Здесь α, β и γ - углы, образованные осью OX'' с осями OY и OZ соответственно. Элементы первой строки матрицы \mathbf{R} [6.1, п.14.10] содержат косинусы этих углов: $r_{11} = \cos\alpha$, $r_{12} = \cos\beta$, $r_{13} = \cos\gamma$. Аналогично, вторая и третья строки матрицы содержат косинусы углов, образованных соответственно осями $O'Y'$ и $O'Z'$ с осями глобальной системы координат.

Особенность матрицы \mathbf{R} состоит в том, что она зависит только от трех параметров, поскольку все девять ее элементов связаны шестью уравнениями связи и, следовательно, не являются независимыми. Обозначив строки матрицы в виде векторов \mathbf{r}_1^T , \mathbf{r}_2^T и \mathbf{r}_3^T , эти уравнения можно представить в виде:

$$\mathbf{r}_1^T \mathbf{r}_1 = 1, \quad \mathbf{r}_2^T \mathbf{r}_2 = 1, \quad \mathbf{r}_3^T \mathbf{r}_3 = 1, \quad \mathbf{r}_1^T \mathbf{r}_2 = 0, \quad \mathbf{r}_1^T \mathbf{r}_3 = 0, \quad \mathbf{r}_2^T \mathbf{r}_3 = 0, \quad (6.4)$$

Уравнения (6.4) являются условиями взаимной ортогональности векторов \mathbf{r}_i . Матрица, построенная из таких векторов, называется *ортогональной*. Для ортогональной матрицы справедливо соотношение $\mathbf{R}^{-1} = \mathbf{R}^T$. Условие взаимной ортогональности векторов \mathbf{r}_i в трехмерном пространстве можно выразить в другой удобной форме, которая понадобится нам позже:

$$\mathbf{r}_3 = \pm(\mathbf{r}_1 \times \mathbf{r}_2), \quad \mathbf{r}_2 = \pm(\mathbf{r}_3 \times \mathbf{r}_1), \quad \mathbf{r}_1 = \pm(\mathbf{r}_2 \times \mathbf{r}_3). \quad (6.5)$$

Верхний знак соответствует случаю, когда матрица \mathbf{R} представляет преобразование, не изменяющее взаимной ориентации осей системы, а нижний – преобразование, изменяющее правую систему координат на левую и наоборот.

Смысл вектора $\mathbf{t} = (t_x, t_y, t_z)^T$ ясен непосредственно из рисунка.

6.3. Стереоскопическая система

Рассмотрим ситуацию, когда две камеры, находящиеся в разных точках, регистрируют одну и ту же сцену. Пара изображений, получаемых при этом, называется *стереопарой*. Обратимся сначала к простейшему случаю. Пусть одинаковые камеры расположены так, что их оптические оси параллельны, а прямая, проходящая через оптические центры, перпендикулярна оптическим осям (эта прямая называется *базовой линией*, а ее отрезок, заключенный между оптическими центрами – *базой*). Положим длину базы равной b . Выберем такую глобальную систему координат, начало которой О расположено на базовой линии посередине между оптическими центрами камер, ось OZ параллельна оптическим осям, а ось OX направлена вдоль базовой линии (рис. 6.3). Пусть начала координат в плоскостях изображений камер совпадают с главными точками ($u_0 = v_0 = 0$), а единицы измерения координат в глобальной системе и в плоскостях изображения камеры одинаковы ($w = h = 1$).

Выберем точку M с глобальными координатами (X, Y, Z) . Координаты ее проекции в плоскости изображения первой (левой) камеры обозначим через (x', y') , а в плоскости изображения второй (правой) камеры – через (x'', y'') . (Проекции одной и той же точки M в плоскостях изображений разных камер называются *сопряженными точками*.) Нетрудно проверить, что

$$x' = f(X + b/2)/Z, \quad x'' = f(X - b/2)/Z, \quad y' = y'' = fY/Z.$$

Заметим, что в направлении, перпендикулярном направлению базовой линии, координаты сопряженных точек (y -координаты) совпадают. Это обстоятельство имеет большое значение при автоматизированном поиске сопряженных точек на стереопаре, позволяя существенно сократить размеры зоны поиска. Из первых двух соотношений следует, что

$$Z = fb/(x' - x''). \quad (6.6)$$

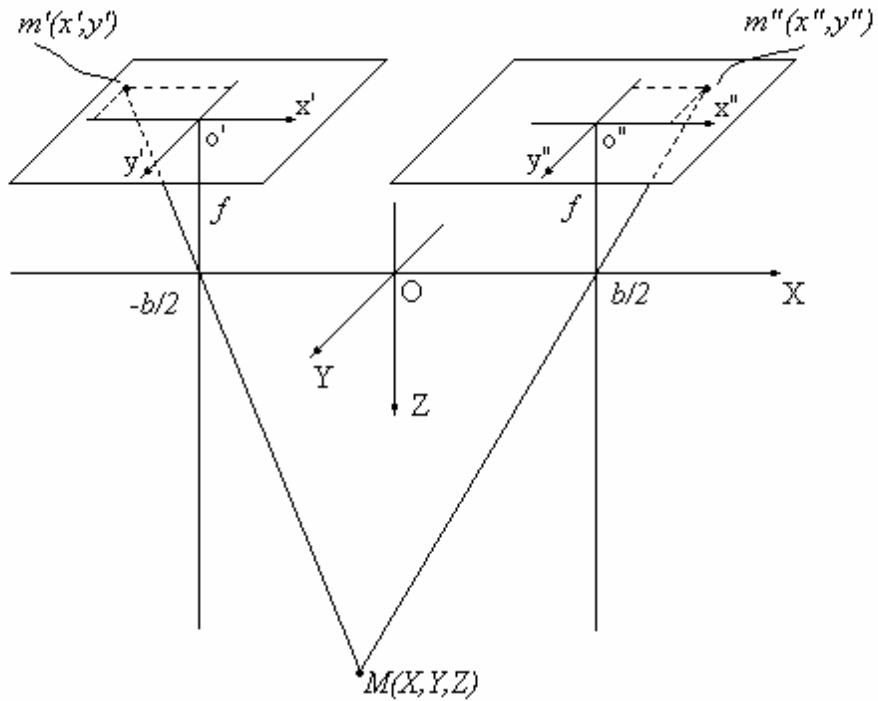


Рис.6.3. Простейшая стереоскопическая система

Это означает, что, зная геометрию съемки и выполнив измерения координат проекций одной и той же точки в плоскостях изображения камер, можно вычислить глубину (координату Z) этой точки. Более того, полученные соотношения позволяют вычислить полностью трехмерные координаты точки:

$$X = b \frac{(x' + x'')}{2(x' - x'')}, \quad Y = b \frac{(y' + y'')}{2(x' - x'')}. \quad (6.7)$$

Разность $d = x' - x''$ называется *диспарантностью*. Из (6.6) и (6.7) следует, что ошибки в координатах проекций сильнее сказываются при малой диспарантности и, следовательно, расстояния до далеких объектов измеряются менее точно, чем до близких. С другой стороны, при фиксированной дальности диспарантность пропорциональна размеру базы, следовательно, точность измерений повышается с увеличением базы. Далее мы, однако, увидим, что увеличение базы может привести к ошибкам, которые не компенсируются увеличением точности измерений.

Теперь рассмотрим общий случай, когда оптические оси камер не параллельны, и направление смещения оптического центра одной камеры относительно оптического центра другой произвольно (рис.6.4). Введем для каждой камеры свою стандартную систему координат, так как это было сделано

в разделе 6.1. Пусть первой камере соответствует система координат $O'X'Y'Z'$, а второй – $O''X''Y''Z''$ (рис. 6.4). Пусть вектор $\mathbf{M}' = (X', Y', Z')^T$ характеризует координаты некоторой точки M трехмерного пространства в системе первой камеры, а вектор $\mathbf{M}''(X'', Y'', Z'')^T$ – в системе второй. Переход от глобальной системы координат к стандартным системам первой и второй камер осуществляется с помощью преобразований $\mathbf{M}' = \mathbf{R}'\mathbf{M} + \mathbf{t}'$ и $\mathbf{M}'' = \mathbf{R}''\mathbf{M} + \mathbf{t}''$ соответственно. Учитывая это, легко показать, что связь между векторами \mathbf{M}' и \mathbf{M}'' задается соотношением

$$\mathbf{M}'' = \mathbf{R}\mathbf{M}' + \mathbf{t}, \quad (6.8)$$

где $\mathbf{R} = \mathbf{R}''\mathbf{R}'^T$ – ортогональная матрица, описывающая ориентацию системы координат второй камеры относительно первой, а $\mathbf{t} = -\mathbf{R}''\mathbf{R}'^T\mathbf{t}' + \mathbf{t}''$ – вектор трансляции, определяющий положение оптического центра второй камеры в системе координат первой. Матрицу \mathbf{R} и вектор \mathbf{t} принято называть *внешними параметрами* системы регистрации.

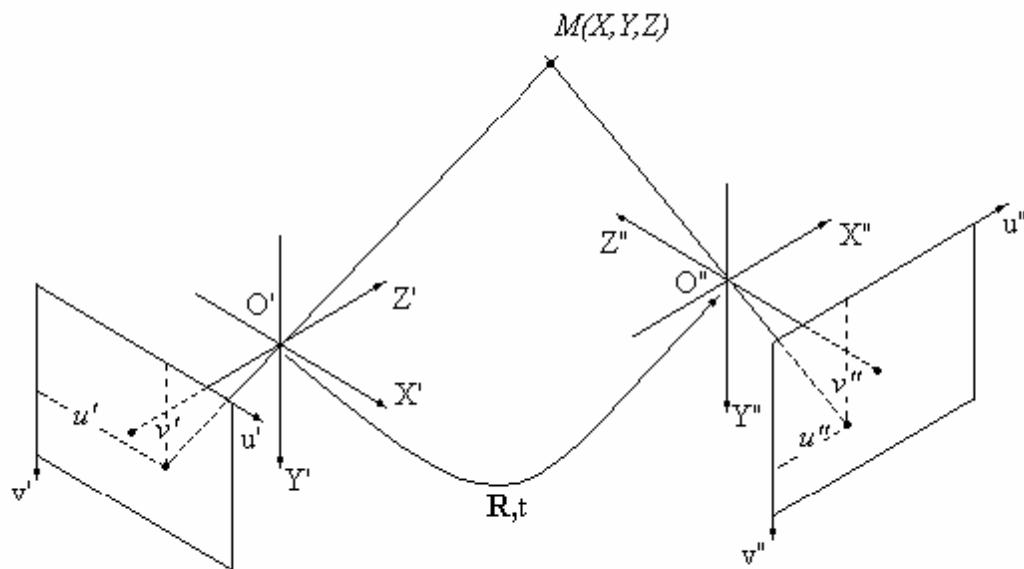


Рис.6.4. Система двух произвольно ориентированных камер

Используя (6.2) из (6.8) можно получить соотношение, связывающее координаты сопряженных точек (в координатах фотоприемника) :

$$Z''\mathbf{A}_2^{-1}\mathbf{v}'' = Z'\mathbf{R}\mathbf{A}_1^{-1}\mathbf{v}' + \mathbf{t}, \quad (6.9)$$

где (см. (6.2)) $\mathbf{v}' = \mathbf{A}_1\mathbf{M}'/Z'$, $\mathbf{v}'' = \mathbf{A}_2\mathbf{M}''/Z''$. (6.10)

(Предполагается, что регистрация может выполняться двумя различными камерами, внутренние параметры которых определяются матрицами \mathbf{A}_1 и \mathbf{A}_2).

Уравнения (6.8) и (6.2) позволяют оценить трехмерные координаты точки M в системе координат любой из камер, если известны внешние параметры системы камер и удается измерить координаты изображений этой точки в плоскостях изображения камеры (т.е. оценить векторы \mathbf{v}' и \mathbf{v}''). Поскольку компоненты векторов \mathbf{v}' и \mathbf{v}'' могут содержать ошибки, реально соотношение (6.9) принимает вид:

$$Z''\mathbf{A}_2^{-1}\mathbf{v}'' - Z'\mathbf{R}\mathbf{A}_1^{-1}\mathbf{v}' - \mathbf{t} = \mathbf{e},$$

где \mathbf{e} - вектор невязки, обусловленный наличием ошибок измерений.

Учитывая это, для оценивания неизвестных Z' и Z'' можно воспользоваться методом наименьших квадратов (МНК), суть которого состоит в том, чтобы найти такие оценки Z' и Z'' , которые бы минимизировали сумму квадратов компонент (норму) вектора невязки: $J_1(Z', Z'') = \mathbf{e}^T \mathbf{e}$. Приравнивая частные производные J_1 по Z' и Z'' нулю (условие достижения экстремума) получим систему, состоящую из двух скалярных уравнений

$$\left. \begin{aligned} Z'\mathbf{v}'^T \mathbf{A}_1^{-T} \mathbf{A}_1^{-1} \mathbf{v}' - Z'\mathbf{v}'^T \mathbf{A}_1^{-T} \mathbf{R}^T \mathbf{A}_2^{-1} \mathbf{v}'' + \mathbf{v}'^T \mathbf{A}_1^{-T} \mathbf{R}^T \mathbf{t} &= 0 \\ -Z'\mathbf{v}'^T \mathbf{A}_1^{-T} \mathbf{R}^T \mathbf{A}_2^{-1} \mathbf{v}'' + Z''\mathbf{v}''^T \mathbf{A}_2^{-T} \mathbf{A}_2^{-1} \mathbf{v}'' - \mathbf{v}''^T \mathbf{A}_2^{-T} \mathbf{t} &= 0 \end{aligned} \right\},$$

разрешая которую относительно Z' и Z'' получим

$$\begin{bmatrix} Z' \\ Z'' \end{bmatrix} = \begin{bmatrix} \mathbf{v}'^T \mathbf{A}_1^{-T} \mathbf{A}_1^{-1} \mathbf{v}' & -\mathbf{v}'^T \mathbf{A}_1^{-T} \mathbf{R}^T \mathbf{A}_2^{-1} \mathbf{v}'' \\ -\mathbf{v}'^T \mathbf{A}_1^{-T} \mathbf{R}^T \mathbf{A}_2^{-1} \mathbf{v}'' & \mathbf{v}''^T \mathbf{A}_2^{-T} \mathbf{A}_2^{-1} \mathbf{v}'' \end{bmatrix}^{-1} \begin{bmatrix} -\mathbf{v}'^T \mathbf{A}_1^{-T} \mathbf{R}^T \\ \mathbf{v}''^T \mathbf{A}_2^{-T} \end{bmatrix} \mathbf{t}. \quad (6.11)$$

Теперь, используя (6.2), можно вычислить вектор трехмерных координат точки M в системе любой из камер:

$$\mathbf{M}' = Z'\mathbf{A}_1^{-1}\mathbf{v}', \quad \mathbf{M}'' = Z''\mathbf{A}_2^{-1}\mathbf{v}''. \quad (6.12)$$

Из всего изложенного выше следует, что для оценивания трехмерных координат некоторой точки по стереопаре необходимо: а) знать внутренние параметры камер (задача калибровки), б) знать параметры взаимного расположения камер (задача взаимного ориентирования), в) найти и определить на изображениях координаты соответствующих данной точке сопряженных точек (задача поиска сопряженных точек).

6.4. Калибровка камеры

Заметим сразу, что поскольку фокусное расстояние f и масштабные коэффициенты w и h входят в матрицу \mathbf{A} только в комбинациях f/w и f/h , то раздельно оценить все три этих параметра невозможно. Введем новые параметры $\alpha_u = f/w$ и $\alpha_v = f/h$, которые и будут подлежать оцениванию.

Пусть камера регистрирует сцену, содержащую N опорных точек, и трехмерные координаты точек известны в глобальной системе координат. Задача состоит в том, чтобы по трехмерным координатам опорных точек (X_i, Y_i, Z_i) и координатам их проекций в плоскости изображений камеры (u_i, v_i) , $i = \overline{1, N}$ оценить элементы матрицы \mathbf{A} .

Если положение камеры относительно этой системы известно (т.е. известны матрица \mathbf{R} и вектор \mathbf{t} в выражении (6.3)), можно сразу пересчитать координаты опорных точек в стандартную систему камеры и для оценивания внутренних параметров воспользоваться соотношениями (6.1), подставив вместо f/w и f/h α_u и α_v соответственно. Перепишем (6.1) в виде

$$\alpha_u X + u_0 Z = uZ, \quad \alpha_v Y + v_0 Z = vZ.$$

Для N опорных точек получим две системы, состоящие из N линейных уравнений каждая, относительно неизвестных u_0 , α_u и v_0 , α_v :

$$\begin{aligned} \mathbf{B}_u \mathbf{a}_u &= \mathbf{w}_u, \\ \mathbf{B}_v \mathbf{a}_v &= \mathbf{w}_v, \end{aligned} \tag{6.13}$$

$$\text{где } \mathbf{B}_u = \begin{bmatrix} X_1 & Z_1 \\ \cdot & \cdot \\ \cdot & \cdot \\ X_N & Z_N \end{bmatrix}, \quad \mathbf{a} = \begin{bmatrix} \alpha_u \\ u_0 \end{bmatrix}, \quad \mathbf{w}_u = \begin{bmatrix} u_1 Z_1 \\ \cdot \\ \cdot \\ u_N Z_N \end{bmatrix},$$

$$\mathbf{B}_v = \begin{bmatrix} Y_1 & Z_1 \\ \cdot & \cdot \\ \cdot & \cdot \\ Y_N & Z_N \end{bmatrix}, \quad \mathbf{a}_v = \begin{bmatrix} \alpha_v \\ v_0 \end{bmatrix}, \quad \mathbf{w}_v = \begin{bmatrix} v_1 Z_1 \\ \cdot \\ \cdot \\ v_N Z_N \end{bmatrix}.$$

Очевидно, достаточно двух опорных точек, чтобы получить оценки внутренних параметров камеры:

$$\alpha_u = \frac{Z_1 Z_2 (u_1 - u_2)}{X_1 Z_2 - X_2 Z_1}, \quad \alpha_v = \frac{Z_1 Z_2 (v_1 - v_2)}{Y_1 Z_2 - Y_2 Z_1},$$

$$u_0 = \frac{X_2 Z_1 u_1 - X_1 Z_2 u_2}{X_2 Z_1 - X_1 Z_2}, \quad v_0 = \frac{Y_2 Z_1 v_1 - Y_1 Z_2 v_2}{Y_2 Z_1 - Y_1 Z_2}.$$

Обратим внимание на необходимость выполнения условий $X_1 Z_2 - X_2 Z_1 \neq 0$ и $Y_1 Z_2 - Y_2 Z_1 \neq 0$, которые формально выражают требование, чтобы в плоскости изображения камеры проекции опорных точек не лежали на прямых, параллельных осям координат.

Для увеличения точности оценок внутренних параметров следует использовать большее количество опорных точек. Системы (6.13) в этом случае становятся переопределенными, и для их решения можно применить МНК. Минимизируя норму вектора невязки $\mathbf{e} = \mathbf{Ba} - \mathbf{w}$ получим оценки внутренних параметров

$$\mathbf{a}_u = (\mathbf{B}_u^T \mathbf{B}_u)^{-1} \mathbf{B}_u^T \mathbf{w}_u, \quad \mathbf{a}_v = (\mathbf{B}_v^T \mathbf{B}_v)^{-1} \mathbf{B}_v^T \mathbf{w}_v.$$

Рассмотрим теперь задачу калибровки в более общей постановке. Пусть неизвестны не только матрица \mathbf{A} , но и положение камеры относительно глобальной системы координат (т.е. матрица \mathbf{R} и вектор \mathbf{t} в выражении (6.3)). Такая постановка обусловлена тем, что технически довольно сложно выполнить точные измерения положения камеры и особенно ее ориентации относительно произвольной системы координат. Используя (6.2) и (6.3) и вводя вектор $\tilde{\mathbf{M}} = (X, Y, Z, 1)$, связь между глобальными координатами опорной точки M и координатами ее изображения можно представить в виде

$$Z' \mathbf{v}' = \mathbf{ARM} + \mathbf{t} = \begin{bmatrix} \alpha_u \mathbf{r}_1^T + u_0 \mathbf{r}_3^T & \alpha_u t_x + u_0 t_z \\ \alpha_v \mathbf{r}_2^T + v_0 \mathbf{r}_3^T & \alpha_v t_y + v_0 t_z \\ \mathbf{r}_3^T & t_z \end{bmatrix} \tilde{\mathbf{M}} = \mathbf{CM} \quad (6.14)$$

Матрица \mathbf{C} имеет размер 3×4 и называется *калибровочной*. Обозначим

$$\begin{aligned} \mathbf{c}_1^T &= (c_{11}, c_{12}, c_{13})^T = \alpha_u \mathbf{r}_1^T + u_0 \mathbf{r}_3^T, \quad c_{14} = \alpha_u t_x + u_0 t_z, \\ \mathbf{c}_2^T &= (c_{21}, c_{22}, c_{23})^T = \alpha_v \mathbf{r}_2^T + v_0 \mathbf{r}_3^T, \quad c_{24} = \alpha_v t_y + v_0 t_z, \\ \mathbf{c}_3^T &= (c_{31}, c_{32}, c_{33})^T = \mathbf{r}_3^T, \quad c_{34} = t_z. \end{aligned} \quad (6.15)$$

Отметим два важных свойства этой матрицы, которые легко получить из (6.15), учитывая (6.4) и (6.5):

$$\begin{aligned} \mathbf{c}_3^T \mathbf{c}_3 &= 1, \\ (\mathbf{c}_1 \times \mathbf{c}_3)^T (\mathbf{c}_2 \times \mathbf{c}_3) &= 0. \end{aligned} \quad (6.16)$$

Оказывается, что если элементы калибровочной матрицы c_{ij} известны, через них можно вычислить матрицы \mathbf{A} , \mathbf{R} и вектор \mathbf{t} . Действительно, учитывая (6.4), легко показать, что

$$\begin{aligned} u_0 &= \mathbf{c}_1^T \mathbf{c}_3, \\ v_0 &= \mathbf{c}_2^T \mathbf{c}_3, \\ \alpha_u &= \mathbf{c}_1 \times \mathbf{c}_3, \\ \alpha_v &= \mathbf{c}_2 \times \mathbf{c}_3. \end{aligned} \tag{6.17}$$

Кроме того, непосредственно из (6.14) следует, что

$$\begin{aligned} t_z &= c_{34}, \\ \mathbf{r}_3^T &= \mathbf{c}_3^T, \\ t_x &= (c_{14} - u_0 t_z) / \alpha_u, \\ t_y &= (c_{24} - v_0 t_z) / \alpha_v, \\ \mathbf{r}_1^T &= (\mathbf{c}_1^T - u_0 \mathbf{c}_3^T) / \alpha_u, \\ \mathbf{r}_2^T &= (\mathbf{c}_2^T - v_0 \mathbf{c}_3^T) / \alpha_v. \end{aligned} \tag{6.18}$$

Задача калибровки, следовательно, сводится к предварительному оцениванию элементов матрицы \mathbf{C} .

Рассмотрим сначала линейный метод оценивания матрицы \mathbf{C} . Запишем матричное уравнение (6.14) как систему трех обычных уравнений

$$\begin{aligned} Z'u' &= c_{11}X + c_{12}Y + c_{13}Z + c_{14}, \\ Z'v' &= c_{21}X + c_{22}Y + c_{23}Z + c_{24}, \\ Z' &= c_{31}X + c_{32}Y + c_{33}Z + c_{34}, \end{aligned}$$

или, подставляя в два первых уравнения значение Z' из третьего,

$$\begin{aligned} c_{11}X + c_{12}Y + c_{13}Z + c_{14} - c_{31}Xu' - c_{32}Yu' - c_{33}Zu' - c_{34}u' &= 0, \\ c_{21}X + c_{22}Y + c_{23}Z + c_{24} - c_{31}Xv' - c_{32}Yv' - c_{33}Zv' - c_{34}v' &= 0. \end{aligned}$$

Зная координаты (X_i, Y_i, Z_i) N опорных точек в трехмерном пространстве и координаты их проекций (u_i, v_i) в плоскости изображения камеры, получим однородную систему из $2N$ линейных уравнений относительно 12 неизвестных элементов калибровочной матрицы c_{ij} :

$$\left. \begin{array}{l} c_{11}X_1 + c_{12}Y_1 + c_{13}Z_1 + c_{14} - c_{31}X_1u'_1 - c_{32}Y_1u'_1 - c_{33}Z_1u'_1 - c_{34}u'_1 = 0 \\ c_{21}X_1 + c_{22}Y_1 + c_{23}Z_1 + c_{24} - c_{31}X_1v'_1 - c_{32}Y_1v'_1 - c_{33}Z_1v'_1 - c_{34}v'_1 = 0 \\ \dots \dots \dots \dots \dots \dots \dots \dots \\ c_{11}X_N + c_{12}Y_N + c_{13}Z_N + c_{14} - c_{31}X_Nu'_N - c_{32}Y_Nu'_N - c_{33}Z_Nu'_N - c_{34}u'_N = 0 \\ c_{21}X_N + c_{22}Y_N + c_{23}Z_N + c_{24} - c_{31}X_Nv'_N - c_{32}Y_Nv'_N - c_{33}Z_Nv'_N - c_{34}v'_N = 0 \end{array} \right\}. \quad (6.19)$$

Представим эту систему в матрично-векторном виде:

$$\mathbf{L}\mathbf{c} = \mathbf{0}, \quad (6.20)$$

где

$$\mathbf{L} = \begin{bmatrix} X_1 & Y_1 & Z_1 & 1 & 0 & 0 & 0 & -u'_1X_1 & -u'_1Y_1 & -u'_1Z_1 & -u'_1 \\ 0 & 0 & 0 & 0 & X_1 & Y_1 & Z_1 & 1 & -v'_1X_1 & -v'_1Y_1 & -v'_1Z_1 & -v'_1 \\ \cdot & \cdot \\ \cdot & \cdot \\ X_N & Y_N & Z_N & 1 & 0 & 0 & 0 & -u'_N X_N & -u'_N Y_N & -u'_N Z_N & -u'_N \\ 0 & 0 & 0 & 0 & X_N & Y_N & Z_N & 1 & -v'_N X_N & -v'_N Y_N & -v'_N Z_N & -v'_N \end{bmatrix}$$

$$\mathbf{c} = (c_{11}, c_{12}, c_{13}, c_{14}, c_{21}, c_{22}, c_{23}, c_{24}, c_{31}, c_{32}, c_{33}, c_{34})^T, \quad \mathbf{0} = \underbrace{(0, 0, \dots, 0, 0)^T}_{2N}.$$

Сначала рассмотрим некоторые общие особенности этой системы. Очевидно, что одним из решений этой системы является тривиальное $\mathbf{c} = \mathbf{0}$, которое не имеет физического смысла. Известно [6.2, с.153], что если однородная линейная система имеет хотя бы одно ненулевое решение, то она имеет бесконечное множество решений, причем, если $\hat{\mathbf{c}}$ - решение, то и $a\hat{\mathbf{c}}$, где a - произвольное число, тоже является решением. Здесь необходимо различать два случая.

Первый – когда ранг матрицы \mathbf{L} на единицу меньше размера вектора \mathbf{c} . Тогда существует только одно (с точностью до произвольного скалярного множителя) решение. Именно этот случай и представляет практический интерес. Для реализации этого условия необходимо (но недостаточно), чтобы количество уравнений в (6.20) было не менее 11, следовательно, количество опорных точек должно быть не менее шести. Ограничить множество решений можно, воспользовавшись первым из условий (6.16). Действительно, определив некоторое решение $\hat{\mathbf{c}}$, в качестве оценки компонент калибровочной матрицы выберем $\mathbf{c} = a\hat{\mathbf{c}}$ такое, чтобы $(a\hat{c}_{32})^2 + (a\hat{c}_{32})^2 + (a\hat{c}_{33})^2 = 1$. Такая нормировка

определяет калибровочную матрицу с точностью до знака. Выбрать правильный знак матрицы можно, например, зная, с какой стороны от плоскости OXY глобальной системы координат находится камера, и учитывая первое из соотношений (6.18). Знак c_{34} должен совпадать со знаком компоненты t_z вектора трансляции.

Второй случай реализуется, если ранг L меньше размера вектора \mathbf{c} на два и более. Здесь может существовать множество различных решений системы, среди которых осуществить правильный выбор без привлечения дополнительных данных невозможно. Показано [6.3], что такая ситуация возникает, в частности, если все опорные точки лежат в одной плоскости. Чтобы избежать этой ситуации, в качестве тестового объекта часто используют объект, приведенный на рис. 6.5.

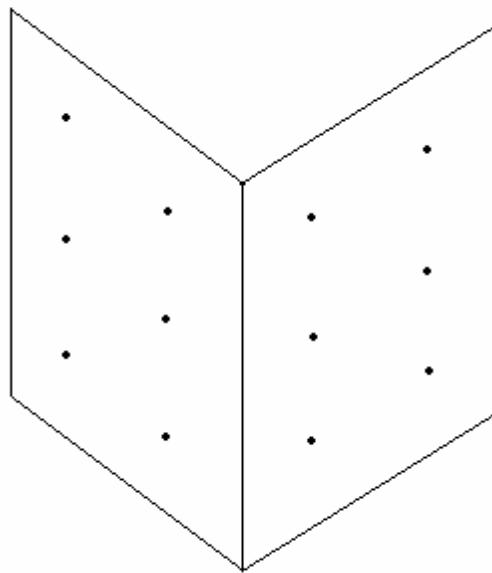


Рис.6.5. Калибровочный объект

Рассмотрим теперь непосредственно метод решения системы (6.20). Обычно, чтобы уменьшить влияние ошибок измерений трехмерных координат опорных точек и координат их изображений в камере, используют больше, чем шесть опорных точек. Тогда система (6.20) становится переопределенной. Кроме того, как и в разделе 6.2, из-за ошибок в измерениях координат реально система (6.20) принимает вид

$$L\mathbf{c} = \mathbf{e},$$

где \mathbf{e} , как и в п.6.2, - неизвестный вектор невязки, обусловленный наличием ошибок измерений.

В этой ситуации можно снова воспользоваться МНК, согласно которому в качестве оценки вектора \mathbf{c} следует принять такой, который минимизирует значение функционала $J_2 = \mathbf{e}^T \mathbf{e} = \mathbf{c}^T \mathbf{L}^T \mathbf{L} \mathbf{c}$ при условии $\mathbf{c}_3^T \mathbf{c}_3 = 1$ (см.(6.16)). Поскольку функционал J_2 представляет собой квадратичную форму с неотрицательно определенной симметричной матрицей $\mathbf{L}^T \mathbf{L}$, то минимума по \mathbf{c} при условии $\mathbf{c}^T \mathbf{c} \neq 0$ он достигает, если \mathbf{c} - собственный вектор матрицы $\mathbf{L}^T \mathbf{L}$, соответствующий ее минимальному собственному числу λ_{\min} [6.4, с.64]. Заметим, что $J_{2\min} = \lambda_{\min}$. (Собственным вектором матрицы \mathbf{A} называется вектор \mathbf{b} , который удовлетворяет уравнению $\mathbf{Ab} = \lambda \mathbf{b}$, где λ - скалярный множитель, называемый собственным числом матрицы \mathbf{A}).

Другой подход к оцениванию матрицы \mathbf{C} основан на минимизации расстояний между измеренными проекциями опорных точек и вычисленными в соответствии с (6.14). Определим величину

$$E = \sum_{i=1}^N \left\{ \left(\frac{c_{11}X_i + c_{12}Y_i + c_{13}Z_i + c_{14}}{c_{31}X_i + c_{32}Y_i + c_{33}Z_i + c_{34}} - u'_i \right)^2 + \left(\frac{c_{21}X_i + c_{22}Y_i + c_{23}Z_i + c_{24}}{c_{31}X_i + c_{32}Y_i + c_{33}Z_i + c_{34}} - v'_i \right)^2 \right\},$$

которая представляет сумму квадратов этих расстояний, и минимизируем ее по \mathbf{c} с учетом ограничений (6.16). Решить эту задачу можно посредством методов условной минимизации [6.5, гл.VI], изложение которых выходит за рамки данного учебника. Следует отметить только, что такой подход обычно дает результаты оценивания, более устойчивые к ошибкам измерений, чем рассмотренный выше линейный.

6.5. Взаимное ориентирование

В предыдущем параграфе показано, как с помощью тестовой сцены не только оценить внутренние параметры камеры, но и определить ее положение относительно системы координат, в которой задана тестовая сцена. Последнее позволяет нам, откалибровав две камеры по одной тестовой сцене, определить их взаимное положение (см. (6.8)). Однако если внутренние параметры камер определены их конструкцией и могут считаться постоянными, то взаимное положение камер во многих практических случаях может изменяться в процессе регистрации сцены. В связи с этим возникает задача взаимного ориентирования, не предполагающая наличия тестовой сцены.

Рассмотрим снова пару камер, внутренние параметры которых известны, но неизвестны внешние параметры (матрица \mathbf{R} и вектор \mathbf{t}). Умножив обе части выражения (6.8) слева сначала векторно на \mathbf{t} , а затем скалярно на \mathbf{M}'' , получим

$$\mathbf{M}''^T (\mathbf{t} \times \mathbf{RM}') = 0.$$

Это соотношение формально выражает тот факт, что векторы \mathbf{M}' , \mathbf{M}'' и \mathbf{t} лежат в одной плоскости, проходящей через три точки: оптические центры камер O' и O'' и точку наблюдения M . Выражая \mathbf{M} через \mathbf{v} из (6.10) получим:

$$\mathbf{v}''^T \mathbf{A}_2^{-T} (\mathbf{t} \times \mathbf{R} \mathbf{A}_1^{-1} \mathbf{v}') = 0 \quad (6.21)$$

или, учитывая свойства смешанного произведения векторов,

$$(\mathbf{v}''^T \mathbf{A}_2^{-T} \times \mathbf{R} \mathbf{A}_1^{-1} \mathbf{v}')^T \mathbf{t} = 0. \quad (6.22)$$

Эквивалентные соотношения (6.21) и (6.22) являются основой для оценивания матрицы \mathbf{R} и вектора \mathbf{t} . Предположим, что известны координаты n пар сопряженных точек и, соответственно, n пар векторов \mathbf{v}'_i и \mathbf{v}''_i $i=1, 2, \dots, n$.

Рассмотрим метод оценивания \mathbf{R} и \mathbf{t} , использующий (6.22). Так как это соотношение справедливо для любой пары сопряженных точек, мы имеем систему из n уравнений относительно неизвестных \mathbf{R} и \mathbf{t} , которую можно представить в виде:

$$\mathbf{B}_n \mathbf{t} = \mathbf{0}, \text{ где } \mathbf{B}_n = \begin{bmatrix} (\mathbf{v}_1''^T \mathbf{A}_2^{-T} \times \mathbf{R} \mathbf{A}_1^{-1} \mathbf{v}_1')^T \\ \dots \\ (\mathbf{v}_n''^T \mathbf{A}_2^{-T} \times \mathbf{R} \mathbf{A}_1^{-1} \mathbf{v}_n')^T \end{bmatrix}. \quad (6.23)$$

Система (6.23) является однородной линейной по \mathbf{t} . Это означает, что вектор трансляции можно оценить только с точностью до постоянного множителя. Вводя условие нормировки $\|\mathbf{t}\|^2 = 1$, количество возможных решений можно ограничить двумя, отличающимися знаком. Вопрос о выборе знака будет рассмотрен позже. Система (6.23) содержит пять неизвестных, так как матрица \mathbf{R} в силу условий нормировки и ортогональности зависит от трех параметров, а вектор \mathbf{t} с учетом введенной нормировки – от двух. Поэтому число уравнений в системе, следовательно и число пар известных сопряженных точек n должно быть не менее пяти.

Поскольку на практике в матрицу \mathbf{B}_n входят не точные значения координат сопряженных точек, а результаты их измерений, которые могут содержать ошибки, реально система (6.23) имеет ненулевую правую часть, т.е.

$$\mathbf{B}_n \mathbf{t} = \mathbf{e},$$

где \mathbf{e} , как и в п.6.2, - вектор невязки, обусловленный наличием ошибок измерений.

Согласно МНК в качестве оценок матрицы вращения и вектора трансляции следует выбрать такие \mathbf{R} и \mathbf{t} , которые минимизируют значения

функционала $J_2(\mathbf{R}, \mathbf{t}) = \mathbf{e}^T \mathbf{e} = \mathbf{t}^T \mathbf{B}_n^T \mathbf{B}_n \mathbf{t}$. Как упоминалось ранее, при условии $\|\mathbf{t}\|^2 = 1$ квадратичная форма $\mathbf{t}^T \mathbf{B}_n^T \mathbf{B}_n \mathbf{t}$ достигает минимума $J_2 = \lambda_{\min}$ по \mathbf{t} (λ_{\min} - минимальное собственное число матрицы $\mathbf{B}_n^T \mathbf{B}_n$), если \mathbf{t} - собственный вектор матрицы, соответствующий λ_{\min} . Поэтому процедуру оценивания \mathbf{R} и \mathbf{t} можно разбить на два этапа. На первом находится матрица \mathbf{R} , минимизирующая λ_{\min} . На втором оценивается собственный вектор матрицы $\mathbf{B}_n^T \mathbf{B}_n$, соответствующий λ_{\min} . Существует множество алгоритмов и их программных реализаций для вычисления собственных векторов, поэтому второй этап не вызывает трудностей.

Значительно более сложной задачей является задача оценивания матрицы \mathbf{R} . Один из возможных алгоритмов состоит в следующем [6.6]. Известно [6.1, п.14.10], что матрица \mathbf{R} может быть представлена в виде $\mathbf{R} = \mathbf{R}_x(\omega_x) \mathbf{R}_y(\omega_y) \mathbf{R}_z(\omega_z)$, где

$$\mathbf{R}_x = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \omega_x & -\sin \omega_x \\ 0 & \sin \omega_x & \cos \omega_x \end{bmatrix}, \quad \mathbf{R}_y = \begin{bmatrix} \cos \omega_y & 0 & \sin \omega_y \\ 0 & 1 & 0 \\ -\sin \omega_y & 0 & \cos \omega_y \end{bmatrix},$$

$$\mathbf{R}_z = \begin{bmatrix} \cos \omega_z & -\sin \omega_z & 0 \\ \sin \omega_z & \cos \omega_z & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Углы ω_x , ω_y и ω_z есть те три неизвестных параметра, от которых зависит матрица \mathbf{R} . На практике всегда известен диапазон, в котором они могут лежать. Выполняя в этом диапазоне полный перебор по всем углам с достаточно грубым шагом (например, 1°) можно приблизиться к значениям, удовлетворяющим требованиям минимизации функционала J_2 по \mathbf{R} . Затем в окрестности этих значений для уточнения положения минимума можно воспользоваться одним из известных методов минимизации [6.5, гл.V] (например, наискорейшего спуска, Ньютона, Маркуардта).

Наконец, получив оценки \mathbf{R} и \mathbf{t} , можно, используя (6.11), оценить и Z -координаты наблюдаемых точек. Из способа задания систем координат (см. рис.6.4) следует, что Z' и Z'' должны быть положительными. Этим условием и определяется выбор правильного знака вектора трансляции \mathbf{t} .

В заключение необходимо сказать, что развитием темы взаимного ориентирования является задача *самокалибровки* системы камер, целью которой является оценивание как внутренних так и внешних параметров. Не останавливаясь на этой задаче подробно, отметим только, что в системе, состоящей из двух, даже одинаковых, камер, данных для самокалибровки недостаточно. Добавление третьей камеры с теми же внутренними параметрами

делает самокалибровку возможной. Подробное исследование этого вопроса можно найти в [6.7].

Отметим, что задача определения взаимного положения камер может иметь и другую трактовку. Предположим, что движущаяся камера непрерывно регистрирует некоторую сцену. Тогда, анализируя последовательность изображений и решая эту задачу, можно определить характер движения камеры в пространстве.

6.6. Поиск сопряженных точек

Центральной проблемой компьютерного стереоизрения является поиск сопряженных точек. Задача автоматизированного поиска сопряженных точек состоит в следующем: на одном из изображений выбрана точка m' , являющаяся проекцией некоторой точки M трехмерного пространства, необходимо на втором изображении найти точку m'' – проекцию той же точки. Достаточно очевидно, что фактически при решении этой задачи речь идет об отождествлении не отдельных точек, а фрагментов изображений, лежащих в окрестностях этих точек. Здесь возникает два вопроса. Первый – если на одном изображении указана точка, где на втором изображении искать сопряженную? Второй – что является критерием «похожести» фрагментов?

Оказывается, на первый вопрос можно дать достаточно строгий ответ. Обратимся к соотношению (6.21). Известно [6.8, Приложение], что векторное произведение $\mathbf{a} \times \mathbf{b}$ можно представить как $[a]_x \mathbf{b}$, где матрица $[a]_x$ имеет вид:

$$[a]_x = \begin{bmatrix} 0 & -a_z & a_y \\ a_z & 0 & -a_x \\ -a_y & a_x & 0 \end{bmatrix}.$$

Тогда (6.21) можно представить как

$$\mathbf{v}''^T \mathbf{F} \mathbf{v}' = 0 \quad (6.24)$$

или

$$\mathbf{v}'^T \mathbf{F}^T \mathbf{v}'' = 0,$$

где

$$\mathbf{F} = \mathbf{A}_2^{-T} [t]_x \mathbf{R} \mathbf{A}_1^{-1}. \quad (6.25)$$

С другой стороны, любая прямая линия на плоскости может быть задана уравнением $ax + by + c = 0$ или, в векторном виде, $\mathbf{a}^T \mathbf{m} = 0$, где $\mathbf{m} = (x, y, 1)$, $\mathbf{a} = (a, b, c)$. Сравнивая уравнение прямой с (6.24) и с (6.25) приходим к выводу, что (6.24) задает в плоскости изображения первой камеры прямую

$$\mathbf{a}'^T \mathbf{v}' = 0 \quad (6.26)$$

с вектором коэффициентов $\mathbf{a}' = \mathbf{F}^T \mathbf{v}''$,
а в плоскости изображения второй камеры - прямую

$$\mathbf{a}''^T \mathbf{v}'' = 0 \quad (6.27)$$

с вектором коэффициентов $\mathbf{a}'' = \mathbf{F} \mathbf{v}'$.

Это значит, что если в плоскости изображения первой камеры указана точка с вектором внутренних координат \mathbf{v}' , то сопряженная точка в плоскости изображения второй камеры может лежать только на прямой, заданной уравнением (6.26). Аналогичное правило справедливо и для точек, заданных в плоскости изображения второй камеры.

Этот результат имеет наглядную геометрическую иллюстрацию (рис.6.6).

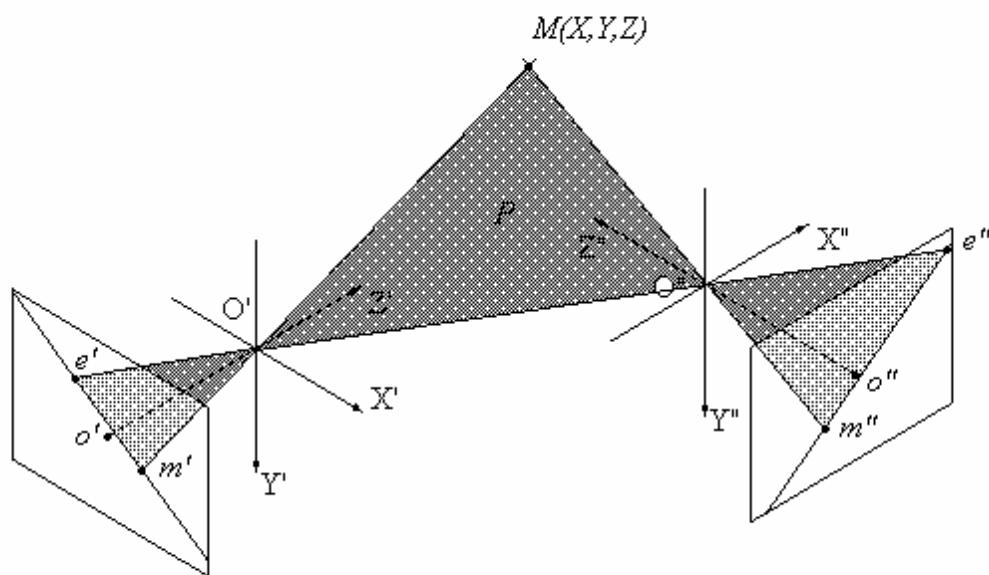


Рис. 6.6. Эпиполярные линии

Оптические центры камер O' и O'' и наблюдаемая точка M образуют плоскость P , которая пересекает плоскости изображения камер по прямым $e'm'$ и $e''m''$, где точка e' является изображением оптического центра O'' второй камеры в плоскости первой, а e'' - изображением центра O' в плоскости второй. С другой стороны, если задана точка m' , то точка M трехмерного пространства может лежать только на прямой, проходящей через m' и оптический центр O' . Следовательно, ее проекция в плоскости изображения второй камеры должна лежать в плоскости P , лежащей на этой прямой и

оптическом центре C'' , а значит, на линии пересечения плоскости P и плоскости изображения второй камеры. Линии пересечения плоскости P с плоскостями изображений камер называются *эпиполярными линиями*.

Практическая ценность полученного результата заключается в том, что его использование при поиске сопряженных точек позволяет существенно сократить размер зоны поиска, выполняя поиск не на всем изображении, а только вдоль эпиполярной линии, и тем самым снизить вероятность ложной идентификации фрагментов.

В отличие от рассмотренных ранее вопросов, сопоставление окрестностей сопряженных точек не поддается строгой формализации, поскольку в его основе лежит проблема идентификации по изображениям фрагментов реального трехмерного мира, которые с трудом поддаются адекватному формальному описанию. Возникающие при съемке и зависящие от ракурса проективные и яркостные искажения приводят к тому, что в изображениях одних и тех же участков сцены, снятых под разными ракурсами, могут появиться значительные отличия. Принципиально важно то, что эти отличия зависят не только от геометрии съемки, но и от геометрических и физических характеристик самой сцены. Расположение источника света по отношению к поверхности сцены влияет на распределение освещенности. Положение элементов сцены, их отражательные и рассеивающие свойства определяют количество энергии, попадающей в объективы камер, а следовательно, и локальные отличия в яркости сопряженных фрагментов изображений.

На рис.6.7 схематично показаны некоторые причины возникновения различий в изображениях. Пунктиром показаны индикаторы излучения (рассеяния) в точках А и В наблюдаемой поверхности.

Видно, что интенсивность излучения из точки А в направлении левой камеры (с оптическим центром C') больше, чем в направлении правой камеры (с оптическим центром C''). Точка В, наоборот, излучает в камеру C' слабее, чем в камеру C'' . Жирными столбиками показана интенсивность сигнала в изображениях этих точек в левой (a' и b') и правой (a'' и b'') камерах. Более того, точка С вообще не видна в левой камере, поскольку ее загораживает участок поверхности ВС.

Величина различий зависит от разницы в ракурсах съемки. Как правило, чем больше эта разница (в частности, чем больше база), тем менее похожими становятся изображения. Поэтому все методы отождествления окрестностей сопряженных точек в большей или меньшей степени опираются не на формальный подход, а на здравый смысл.

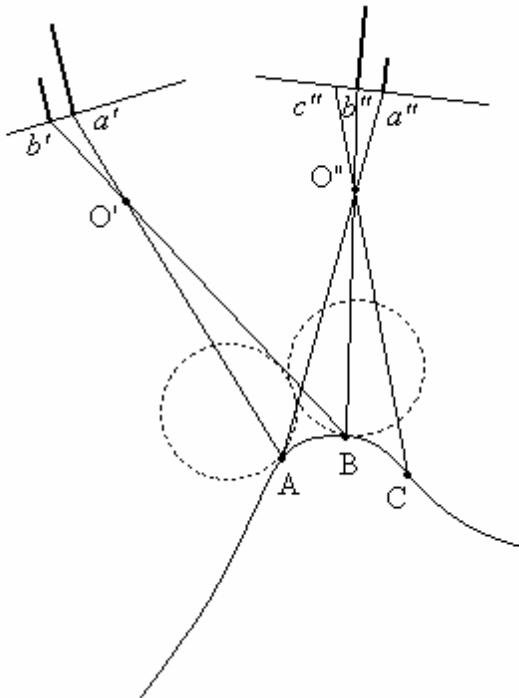


Рис.6.7. Проективные и яркостные искажения

Одним из наиболее распространенных методов идентификации является корреляционный. Детальное описание корреляционного метода приведено в гл.6. При незначительных отличиях в ракурсах съемки и на достаточно гладких поверхностях от него можно ожидать хороших результатов. На рис.6.8 (а и б) приведена пара вертолетных снимков района озера Карымское (Камчатка), на которые нанесена прореженная сетка сопряженных точек, найденных корреляционным методом, а на рис.6.8.в показана поверхность, «натянутая» на трехмерные точки, координаты которых вычислены методом, изложенным в п.6.2. Следует заметить, что построение поверхности, лежащей на заданном множестве трехмерных точек, представляет собой самостоятельную нетривиальную задачу, рассмотрение которой выходит за рамки темы, обсуждаемой в этой главе.

На снимках есть области, где сопряженные точки не удалось найти. Причин этому несколько. Часть сцены, видимая в правой части правого снимка, просто не попала в поле зрения левой камеры. Кроме этого, на снимках есть участки с мало изменяющейся яркостью. На таких участках корреляционный критерий не дает хорошо выраженного максимума, поэтому здесь координаты сопряженных точек определяются с большими погрешностями. Более того, здесь велика вероятность ложных отождествлений. Пространственное представление о таких участках получить без привлечения дополнительной информации невозможно. Если размеры этих участков малы, а по косвенным признакам можно предположить, что поверхность сцены изменяется плавно, можно интерполировать трехмерную структуру участков с их границ вправо, как это сделано для участков a', a'' и b', b'' .

Если их размеры велики, или сопряженные точки на границах не определены, восстановить трехмерную структуру участка не удается (участки c', c'' и d', d'').

Для того, чтобы восстановить мелкую структуру сцены, необходимо, чтобы сопряженные точки лежали достаточно часто. В приведенном примере для построения рельефа найдено более 7000 точек.

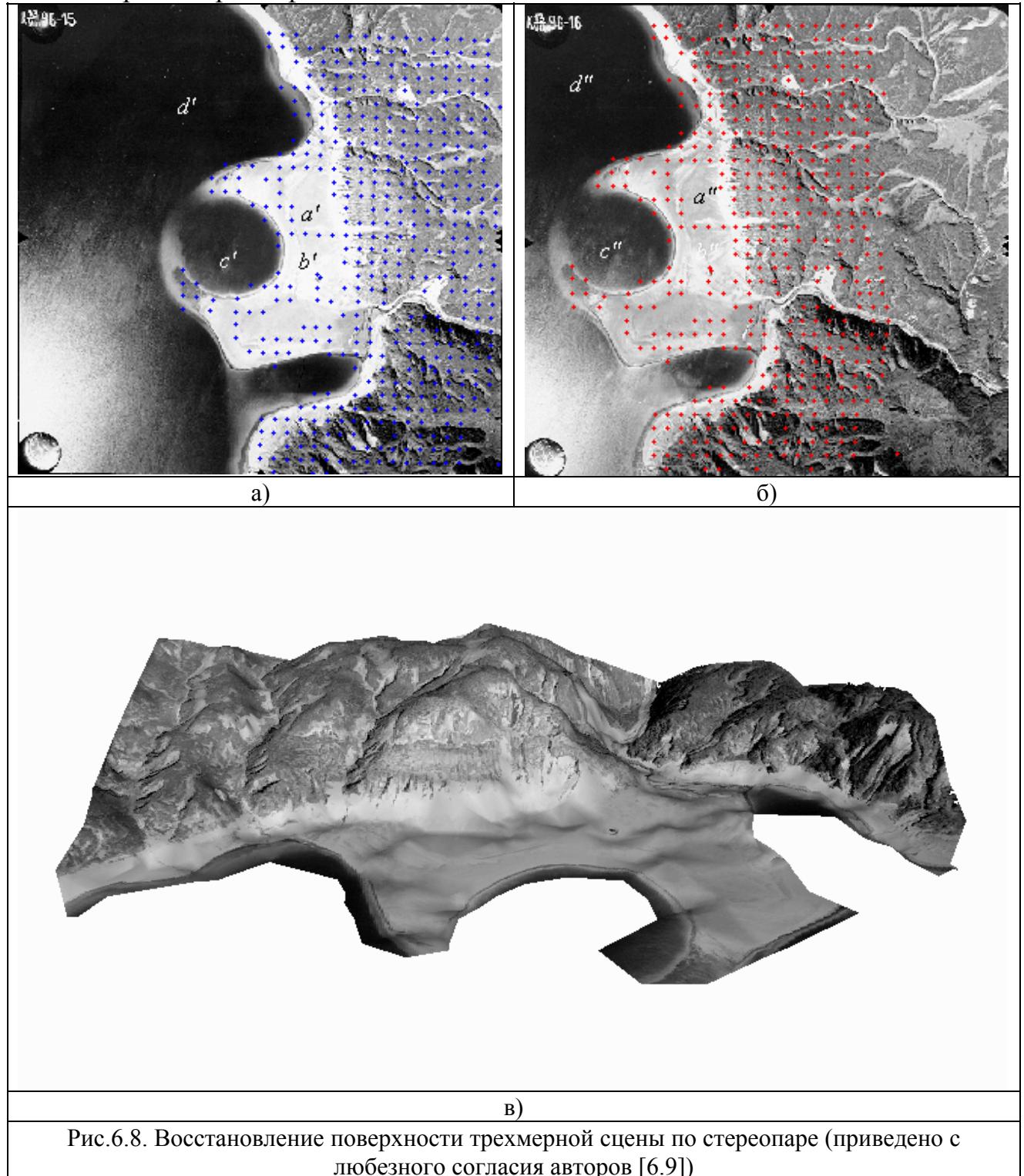


Рис.6.8. Восстановление поверхности трехмерной сцены по стереопаре (приведено с любезного согласия авторов [6.9])

Сложность применения корреляционного метода заключается в том, чтобы подобрать такие размеры сопоставляемых фрагментов, при которых

отличия в тождественных фрагментах еще невелики (для этого нужно уменьшать размеры), а оценка коэффициента корреляции остается достоверной (для этого размеры надо увеличивать). Но самым серьезным его недостатком является чувствительность к масштабным искажениям, присутствующим в отождествляемых фрагментах. Простейший способ уменьшения масштабных различий в сопряженных фрагментах состоит в использовании предварительной аффинной подстройки изображений. Этот способ, однако, оказывается недейственным при вариациях высот в сцене, сравнимых с расстоянием от сцены до системы камер, поскольку в этом случае появляются значительные локальные масштабные искажения.

Влияние яркостных искажений можно попытаться устраниить, отказавшись от сопоставления фрагментов исходных изображений, и перейдя к сопоставлению фрагментов контурных изображений, где выделены края, т.е. переходы между областями с постоянной яркостью. На таких переходах, как известно, градиент яркости имеет максимум, а лапласиан обращается в нуль. Это обстоятельство и используется для выделения краев [6.10, п.17.4]. Существенным моментом является то, что контуры на изображениях непрерывных поверхностей должны располагаться в одинаковом порядке, как это видно на рис. 6.9 (изображения точек *A*, *B* и *D*). Это ограничение позволяет в некоторой степени ослабить влияние масштабных искажений. Однако и здесь, как и в корреляционном методе, существует проблема ложного отождествления. Связана она, в частности, с тем, что видимая часть поверхности зависит от ракурса наблюдения (поэтому, например, на рис.6.9 изображение точки *C* в правой камере отсутствует), и на разных изображениях могут пропадать разные края. Кроме того, сама задача выделения краев только на первый взгляд кажется такой простой. В действительности неизбежное присутствие шумов в изображениях может привести к частичному исчезновению контуров или появлению ложных. Идея поиска сопряженных точек с помощью отождествления краев была особенно популярна у биологов и психологов, исследовавших механизмы зрения [6.11], но, как нам кажется, в таком идеальном виде практического применения не нашла.

Еще один подход к поиску сопряженных точек опирается на предварительный раздельный анализ изображений, позволяющий выделить на них некоторые характерные объекты или особенности. Так на аэроснимках городских территорий можно попытаться выделить перекрестки, отдельные дома или их фрагменты, деревья и т.п., а затем выполнить их отождествление между снимками.

По-видимому, наиболее удачных результатов можно ожидать от комбинированного использования всех этих методов и применения итеративной схемы, когда по найденным сопряженным точкам строится оценка поверхности, с учетом которой изображения подвергаются масштабной коррекции и выполняется поиск новых сопряженных точек.

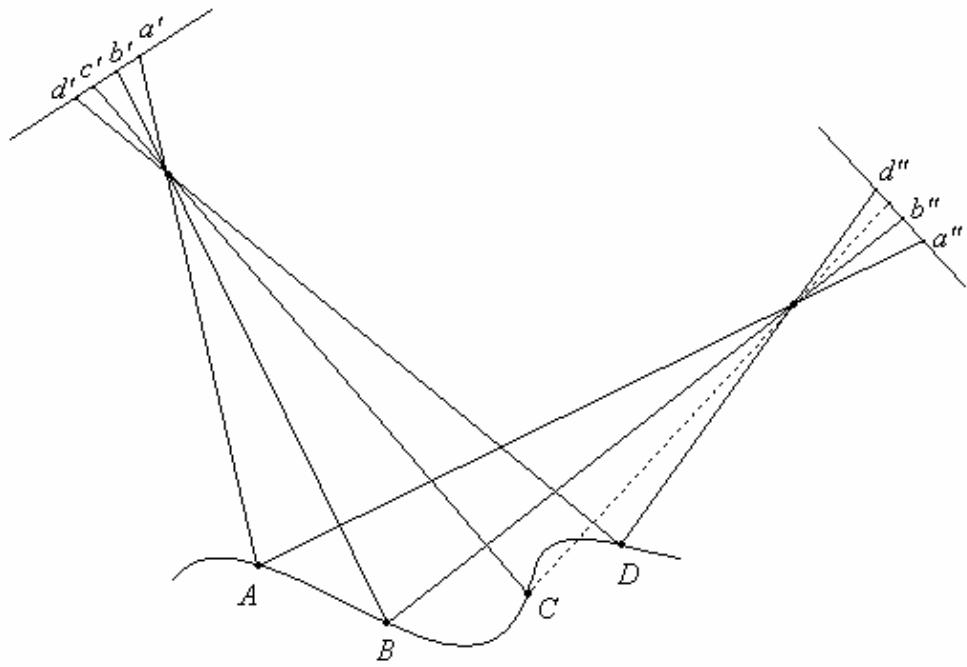


Рис.6.9. Упорядочение сопряженных точек в случае непрерывных поверхностей

ВОПРОСЫ К ГЛАВЕ 6

- 6.1. Какой смысл имеет знак при f в выражении (6.1)?
- 6.2. Какой вид примет выражение (6.3), если вектор трансляции будет задан в глобальной системе координат?
- 6.3. Обратимся к рис. 6.3. Какой вид будут иметь матрицы \mathbf{A}_1 , \mathbf{A}_2 , \mathbf{R} и вектор \mathbf{t} в выражении (6.9) для ситуации, приведенной на рисунке?
- 6.4. Пусть в ситуации, изображенной на рис. 6.3, векторы $\mathbf{v}' = (x', y, 1)$ и $\mathbf{v}'' = (x'', y, 1)$ внутренних координат проекций точки M в плоскостях изображений левой и правой камер известны. Получите оценки трехмерных координат точки M в системах координат правой и левой камер, пользуясь выражениями (6.11) и (6.12). Сравните полученный результат с (6.6) и (6.7). Объясните отличия.
- 6.5. Почему в предыдущем вопросе векторы $\mathbf{v}' = (x', y, 1)$ и $\mathbf{v}'' = (x'', y, 1)$ имеют одинаковые y -компоненты?
- 6.6. Докажите справедливость соотношений (6.17).
- 6.7. Докажите справедливость соотношений (6.18).
- 6.8. На рис. 6.6 точки e' и e'' являются изображениями оптических центров камер. Полагая, что матрицы \mathbf{A}_1 , \mathbf{A}_2 , \mathbf{R} и вектор \mathbf{t} известны, найдите координаты этих точек в плоскостях изображений соответствующих камер.
- 6.9. Где будут находиться точки e' и e'' , если камеры расположены, как показано на рис. 6.3?

ЗАКЛЮЧЕНИЕ

Проблемам обработки изображений, представленным в данном учебном пособии, специалистами уделяется очень большое внимание. Вместе с тем, описанные в пособии задачи и методы их решения далеко не исчерпывают всего многообразия результатов, имеющихся в области цифровой обработки визуальной информации. Они лишь служат иллюстрацией, впрочем, по мнению авторов, достаточно характерной, тех направлений, в которых ведутся исследования и разработки. Авторы пособия выражают уверенность, что глубокое изучение методов, представленных в пособии, позволит читателю в дальнейшем свободно ориентироваться в имеющемся море публикаций и при необходимости самому принимать участие в соответствующих работах.

Литература

Литература к главе 1

- 1.1. *Даджион Д., Мерсеро Р.* Цифровая обработка многомерных сигналов. - М.: Мир, 1988.
- 1.2. *Ярославский Л.П.* Введение в цифровую обработку изображений. - М.: Сов. радио, 1979.
- 1.3. *Прэтт У.* Цифровая обработка изображений. Кн.1. - М.: Мир, 1982.
- 1.4. Методы передачи изображений. Сокращение избыточности / У.К.Прэтт, Д.Д. Сакрисон, Х.Г.Д. Мусманн и др. Под ред. У.К.Прэтта. -М.: Радио и связь, 1983.
- 1.5. *Джайн А.К.* Сжатие видеоинформации // ТИИЭР.- Т.69. - № 3.

Литература к главе 2

- 2.1. *Горбунов Б.А., Дементьев В.Н., Пяткин В.П.* Распознавание изображений в дистанционном зондировании //Автоматизированная обработка изображений природных комплексов Сибири. - Новосибирск: Наука, 1988.
- 2.2. *Левин Б.Р.* Теоретические основы статистической радиотехники. Кн.1. - М.: Сов. радио, 1966.
- 2.3. *Ярославский Л.П.* Введение в цифровую обработку изображений. - М.: Сов. радио, 1979.
- 2.4. *Прэтт У.* Цифровая обработка изображений. Кн.2. - М.: Мир, 1982.

Литература к главе 3

- 3.1. *Спектор А.А.* Рекуррентная фильтрация гауссовских дискретных полей, наблюдаемых в гауссовых шумах //Радиотехника и электроника, 1994, №7.
- 3.2. *Даджион Д., Мерсеро Р.* Цифровая обработка многомерных сигналов. - М.: Мир, 1988.

- 3.3. *Левин Б.Р.* Теоретические основы статистической радиотехники. Кн.2. - М.: Сов. радио, 1975.
- 3.4. *Woods J.W.* Two-Dimensional Digital Signal Processing 1. Berlin e.a., 1981.
- 3.5. *Васильев К.К., Крашенинников В.Р.* Методы фильтрации многомерных случайных полей. - Саратов: Сарат. гос. ун-т, 1990.
- 3.6. *Грузман И.С., Микерин В.И., Спектр А.А.* Двухэтапная фильтрация изображений на основе использования ограниченных данных //Радиотехника и электроника, 1995, №5.
- 3.7. *Грузман И.С., Спектр А.А.* Применение свойства условной независимости для симметричного сглаживания марковских процессов //Радиотехника и электроника, 1997, №5.
- 3.8. *Грузман И.С.* Двухэтапная фильтрация бинарных изображений //Автометрия, 1999, №3.
- 3.9. Быстрые алгоритмы в цифровой обработке изображений /Под ред. Т.С. Хуанга. - М.: Радио и связь, 1984.

Литература к главе 4

- 4.1. *Борн М., Вольф Э.* Основы оптики.- М.: Наука, 1970.
- 4.2. *Прэйтт У.* Цифровая обработка изображений. Кн.1. - М.: Мир, 1982.
- 4.3. *Беллман Р.* Введение в теорию матриц. / Пер. С англ.; Под ред. В.Б. Лидского . - М.: Наука, 1976.
- 4.4. Реконструкция изображений: Пер. с англ. / Под ред. Г.Старка. - М.: Мир, 1992.
- 4.5. *Прэйтт У.* Цифровая обработка изображений. Кн.2. - М.: Мир, 1982.
- 4.6. *Василенко Г.И., Тараторин А.М.* Восстановление изображений. - М.: Радио и связь, 1986.
- 4.7. *Рабинер Л., Гоулд Б.* Теория и применение цифровой обработки сигналов. - М.: Мир, 1978.
- 4.8. *Бейтс Р., Мак-Донелл М.* Восстановление и реконструкция изображений. - М.: Мир, 1989.
- 4.9. *Бендат Дж., Пирсол А.* Прикладной анализ случайных данных: Пер. с англ. - М.: Мир, 1989.
- 4.10. *Грузман И.С.* Компенсация краевых эффектов при коррекции линейных искажений изображений // Автометрия. -1995.- № 2.
- 4.11. *Бъемон Ж., Лагендейк Л., Мерсеро Р.М.* Итерационные методы улучшения изображений // ТИИЭР.-1990.-№5.

Литература к главе 5

- 5.1. *Хорн Б.К.П.* Зрение роботов.- М.: Мир, 1989.
- 5.2. *Прэйтт У.* Цифровая обработка изображений.- М.: Мир,1982.
- 5.3. *Корн Г., Корн Т.* Справочник по математике для научных работников и инженеров.- М.: Наука, 1988.

- 5.4. Тышкевич Р.И., Феденко А.С. Линейная алгебра и аналитическая геометрия.- Минск:Вышешшая школа, 1968.
- 5.5. Абрамов В.В., Киричук В.С. и др. Реконструкция трехмерных поверхностей по двум проекциям при отслеживании камерой заданной точки сцены.//Автометрия, 1998, №5, с.3.
- 5.6. Киричук В.С., Перетягин Г.И. Об установлении сходства фрагментов с эталоном//Автометрия. 1986.№4.с.83.
- 5.7. Битюцкий О.И., Перетягин Г.И. Поиск и локализация реперных фрагментов при совмещении повторных снимков//Автометрия.1988.№3.
- 5.8. Jahne B. Digital Image Processing: Concepts, Algorithms, and Scientific Applications.- Berlin: Springer-Verlag, 1993.
- 5.9. Mostafavi H., Smith F.W. Image correlation with geometric distortion.- IEEE Transactions, 1978, v.AES-14,p.487-500.
- 5.10. Keys R.G. Cubic convolution interpolation for Digital Image Processing.- IEEE Trans. Acoust., Speech, Signal Processing, 1981, vol.ASSP-29, p.1153-1160.
- 5.11. Parker A.J., Kenyon R., Troxel D.E. Comparison of interpolating methods for Image Resampling.- IEEE Trans. Medical Imaging, 1983, vol.MI-2, No.1, p.31-39.
- 5.12. Birchfield S. An Introduction to Projective Geometry (for computer vision).1998. <http://www.dai.ed.ac.uk/Cvonline/geom.htm>

Литература к главе 6

- 6.1. Корн Г., Корн Т. Справочник по математике.- М.: Наука, 1973.
- 6.2. Бронштейн Н.Н., Семеняев К.А. Справочник по математике.- М.: Наука, 1966.
- 6.3. Faugeras O. Tree Dimensional Computer Vision.- MIT Press, 1993.
- 6.4. Rao C.R. Линейные статистические методы и их применения.- М.: Наука, 1968.
- 6.5. Бард Й. Нелинейное оценивание параметров.- М.: Статистика, 1979.
- 6.6. Hu X., Ahuja N. Motion and Structure Estimation Using Long Sequence motion models.// Image and Vision Computing. -1993. -V.11. -№9. -P.549 – 569.
- 6.7. Maybank S., Faugeras O. A. Theory of Self-calibration of a Moving Camera.// International Journal of Computer Vision. -1992. -V.8. -№2. -P.123 - 151.
- 6.8. Хорн Б.К.П. Зрение роботов. -М.: Мир, 1989.
- 6.9. Popov S.A., Kirichuk V.S. Algorithm of Estimation of the Geometric Parameters of the System of Two Projection Cameras by the Method of the Least Squares (MLS). //Pattern Recognition and Image Analysis, -1999, -N2, -p.304.
- 6.10. Прэтт У. Цифровая обработка изображений. -М.: Мир. 1982.
- 6.11. Marr Д. Зрение: информационный подход к изучению представления и обработки зрительных образов. -М.: Радио и связь, 1987.