

Пытьев Ю.П.
Чуличков А.И.

Методы морфологического анализа изображений



МОСКВА
ФИЗМАТЛИТ ®

УДК 519.25+519.8

ББК 22.18

П 95

Пытьев Ю.П., Чуличков А.И. **Методы морфологического анализа изображений.** — М.: ФИЗМАТЛИТ, 2010. — 336 с. — ISBN 978-5-9221-1225-3.

Рассмотрено *математическое понятие формы изображения* как (инвариантной относительно условий получения изображения) характеристики геометрической формы изображенного объекта.

Рассмотрены основанные на понятии формы изображения *методы морфологического анализа изображений*, ориентированные на компьютерные решения задач анализа и интерпретации изображенных объектов при априорной неопределенности условий регистрации их изображений, таких, в частности, как характер освещения, его спектральный состав и т. п.

Типичными являются задачи, в которых даны два изображения одной и той же местности, полученные в разное время при различных и неизвестных условиях освещения, и требуется выделить объекты, представленные на первом (втором) изображении и отсутствующие на втором (первом) изображении.

Для специалистов по математическому моделированию, а также для студентов старших курсов и аспирантов технических и физико-математических специальностей вузов.

Научное издание

ПЫТЬЕВ Юрий Петрович

ЧУЛИЧКОВ Алексей Иванович

МЕТОДЫ МОРФОЛОГИЧЕСКОГО АНАЛИЗА ИЗОБРАЖЕНИЙ

Редактор *И.Л. Легостаева*

Оригинал-макет: *К.А. Андреев*

Оформление переплета: *Н.В. Гришина*

Подписано в печать 04.05.10. Формат 60×90/16. Бумага офсетная. Печать офсетная.
Усл. печ. л. 21. Уч.-изд. л. 24. Тираж 100 экз. Заказ №

Издательская фирма «Физико-математическая литература»
МАИК «Наука/Интерпериодика»
117997, Москва, ул. Профсоюзная, 90
E-mail: fizmat@maik.ru, fmlsale@maik.ru;
<http://www.fml.ru>

Отпечатано в ООО «Чебоксарская типография № 1»
428019, г. Чебоксары, пр. И. Яковлева, 15

ISBN 978-5-9221-1225-3



9 785922 112253

ISBN 978-5-9221-1225-3

© ФИЗМАТЛИТ, 2010

© Ю. П. Пытьев, А. И. Чуличков, 2010

ОГЛАВЛЕНИЕ

Предисловие	9
Введение	13
Часть I. Форма полутонного изображения	
Глава 1. Основные понятия морфологического анализа	26
1.1. Математические модели изображения	26
1.1.1. Линейное пространство изображений (26). 1.1.2. Пространство изображений $\mathcal{L}_\mu^2(X)$ (27). 1.1.3. Пространство изображений $\mathcal{C}(X)$ (28).	
1.2. Понятие формы изображения. Сравнение изображений по форме . .	29
1.2.1. Модель и форма мозаичного изображения (30). 1.2.2. Операция сравнения изображений по форме (33).	
1.3. Форма изображения как оператор проецирования	37
1.3.1. Форма как оператор проецирования на множество \mathcal{V}_f в пространстве $\mathcal{L}_\mu^2(X)$ (37). 1.3.2. Аппроксимация в $\mathcal{L}_\mu^2(X)$ формы произвольного изображения (39). 1.3.3. Форма как оператор проецирования на множество \mathcal{V}_f в пространстве $\mathcal{C}(X)$ (40). 1.3.4. Форма как линейное множество в пространстве $\mathcal{C}(X)$ (41).	
1.4. Форма фрагмента изображения	43
1.4.1. Модель изображения объекта на произвольном фоне и его форма (43). 1.4.2. Форма деформированного фрагмента изображения (44).	
1.5. Примеры форм изображений	46
1.5.1. Форма кусочно гладкого изображения (46). 1.5.2. Форма размытого изображения (50). 1.5.3. Форма контурного изображения (51). 1.5.4. Форма изображения ламбертова объекта (53). 1.5.5. Форма изображения произвольного объекта, освещаемого конечным числом источников света произвольной интенсивности (62).	
1.6. Решение задач морфологического анализа изображений	63
1.6.1. Узнавание объекта по форме его изображения (63). 1.6.2. Классификация объектов по форме их изображений (63). 1.6.3. Выделение отличий по форме (65). 1.6.4. Оценивание параметров объекта по форме его изображения (65).	
1.7. Морфологические и корреляционные методы анализа изображений	67

1.8. Независимость по форме	69
1.8.1. Независимость изображений по форме, заданной как линейное подпространство $\mathcal{L}_\mu^2(X)$ (70). 1.8.2. \mathcal{C} -эквивалентные изображения (74). 1.8.3. Ранг корреляции форм изображений (78).	
Глава 2. Форма изображения с заданной упорядоченностью яркостей	80
2.1. Изображения с известной упорядоченностью яркостей в пространстве $\mathcal{L}_\mu^2(X)$	80
2.1.1. Мозаичные изображения с упорядоченными яркостями (80). 2.1.2. Сравнение по форме изображений с упорядоченной яркостью (84). 2.1.3. Аппроксимация проектора на конус (85). 2.1.4. Эквиализация и форма изображений (88). 2.1.5. Независимость изображений по форме, заданной в виде выпуклого замкнутого конуса (91).	
2.2. Изображения с упорядоченными яркостями в пространстве $\mathcal{C}(\mathcal{X})$. .	91
Глава 3. Морфологический анализ изображений, заданных с погрешностью	94
3.1. Анализ изображений, искаженных ограниченной погрешностью . .	94
3.2. Анализ изображений, искаженных неограниченной погрешностью . .	95
3.2.1. Близость изображения к форме, заданной как линейное подпространство евклидова пространства изображений (95). 3.2.2. Инвариантные свойства функционала близости изображения ξ к заданной форме (98). 3.2.3. Анализ формы изображения как элемента функционального пространства (101).	
3.3. Сравнение по форме двух изображений, регистрируемых с погрешностью	103

Часть II. Стохастические методы анализа формы

Глава 4. Морфологический анализ изображений, искаженных случайным шумом	108
4.1. Модель регистрации, искажающей изображение случайным шумом	108
4.2. Форма как линейное подпространство \mathcal{R}^n	110
4.2.1. Узнавание изображения, искаженного шумом ограниченной энергии (110). 4.2.2. Узнавание изображения, искаженного шумом неограниченной энергии (111). 4.2.3. Принцип максимума надежности при классификации изображений, искаженных гауссовым шумом (114). 4.2.4. Оценка параметров объекта по его зашумленному изображению (118). 4.2.5. Случайные множества, оценивающие значение параметра сцены по ее изображению (121).	
4.3. Форма как выпуклый замкнутый конус $\mathcal{V} \subset \mathcal{R}^n$	124
4.3.1. Близость изображения к форме \mathcal{V} (124). 4.3.2. Оценивающие множества. Форма как выпуклый замкнутый конус (126). 4.3.3. Оценка параметра формы изображения, минимизирующая	

максимальную погрешность при гарантированной надежности (129).	
Глава 5. Эмпирическое построение случайной формы изображения	130
5.1. Аппроксимация формы изображения, искаженного случайным шумом	130
5.1.1. Аппроксимация формы изображения формой мозаичного изображения (130). 5.1.2. Надежность мозаичной аппроксимации формы изображения при заданном разбиении поля зрения (133). 5.1.3. Надежность мозаичной аппроксимации формы изображения. Общий случай (134).	
5.2. Аппроксимация формы изображения параметрическим семейством подпространств	136
5.2.1. Форма изображения как семейство линейных подпространств (136). 5.2.2. Выбор наиболее надежного подпространства заданной размерности для аппроксимации входного сигнала (138). 5.2.3. Выбор размерности аппроксимации (139). 5.2.4. Связь с методом максимального правдоподобия (140). 5.2.5. Результаты вычислительных экспериментов (141).	
5.3. Аппроксимация формы множества изображений	145
5.3.1. Эффективная размерность множества изображений (145). 5.3.2. Эффективная размерность множества классов эквивалентности изображений (149).	

Часть III. Возможностные методы анализа формы

Глава 6. Вероятностные и возможностные модели формы изображений	153
6.1. Вероятностные и возможностные модели морфологического анализа изображений	153
6.2. Возможностная модель изображения и его интерпретации	156
6.2.1. Модель регистрации изображения (156). 6.2.2. Анализ изображений, искаженных аддитивным шумом (159).	
6.3. Примеры задания меры возможности на множестве изображений.	160
6.3.1. Форма как «четкое» множество (160). 6.3.2. Нечеткая форма изображения при деформации поля зрения (161). 6.3.3. Форма как множество изображений с известной упорядоченностью яркостей точек поля зрения (162).	
Глава 7. Стохастические модели возможности в задачах анализа изображений	165
7.1. Возможность в статистической теории проверки гипотез	165
7.1.1. Область принятия гипотезы и ее связь с оценивающим множеством (165). 7.1.2. Нечеткий параметр распределения случайного элемента. Простые гипотезы и альтернативы (167). 7.1.2. Нечеткий параметр распределения случайного элемента. Сложные гипотезы и альтернативы (168).	

7.2. Возможность на множестве значений параметра формы изображения	168
7.2.1. Возможность, индуцированная результатом регистрации изображения с шумом заданной дисперсии (168). 7.2.2. Возможность, индуцированная результатом регистрации изображения с шумом неизвестной дисперсии (170).	
7.3. Эмпирическое построение нечеткой формы изображения	173

Часть IV. Цветные изображения

Глава 8. Морфологический анализ цветных изображений	177
8.1. Математическая модель цветного изображения	178
8.1.1. Цвет и яркость (178). 8.1.2. Цветное (спектрозональное) изображение (182).	
8.2. Форма цветного изображения	184
8.2.1. Сравнение цветных изображений по форме (185). 8.2.2. Форма мозаичного цветного изображения (187). 8.2.3. Общее понятие формы цветного изображения (192).	
8.3. Аппроксимация формы цветных изображений	193
8.3.1. Форма как оператор наилучшего приближения (193). 8.3.2. Приближение цветного изображения изображениями, цвет и яркость которых постоянны на заданных подмножествах разбиения $\{A_i\}$ поля зрения X (194). 8.3.3. Аппроксимация формы в широком смысле произвольного изображения $\vec{f}(\cdot)$ (197). 8.3.4. Приближение цветного изображения мозаичными изображениями с заданным конечным набором значений (200). 8.3.5. Приближение цветного изображения мозаичными изображениями с конечным числом значений вектора цвета (203). 8.3.6. Приближение цветного изображения изображениями, цвет которых постоянен на заданных подмножествах разбиения $\{A_j\}$ поля зрения X (204). 8.3.7. Приближение цветного изображения изображением с заданным набором значений вектора цвета (210). 8.3.8. Приближение цветного изображения изображением с конечным числом значений вектора цвета (213). 8.3.9. Форма изображения с заданным распределением цвета $\vec{\varphi}(x)$, $x \in X$ (214). 8.3.10. Случай, когда допускаются небольшие изменения цвета в пределах каждого A_i , $i = 1, \dots, N$ (215).	
8.4. Форма цветного изображения как оператор наилучшего приближения в чебышевской метрике	221
8.5. Задачи морфологического анализа цветных изображений	224
8.5.1. Задачи идентификации при произвольно меняющейся интенсивности освещения (224). 8.5.2. Задачи совмещения изображений и поиска фрагмента (225). 8.5.3. Задача анализа спектрозональных изображений (226).	

Часть V. Прикладные задачи

Глава 9. Примеры решения задач морфологического анализа изображений	228
9.1. Морфологическое подавление случайного шума	228
9.1.1. Подавление шума на кусочно постоянном изображении (228).	
9.1.2. Подавление шума на произвольных изображениях (233).	
9.2. Выделение неизвестного объекта на фоне, форма изображения которого известна	237
9.2.1. Морфологический метод (237). 9.2.2. Локальный морфологический алгоритм (239). 9.2.3. Локальный корреляционный метод (241). 9.2.4. Локальный метод ранговой корреляции (242). 9.2.5. Сравнение алгоритмов (242).	
9.3. Аппроксимация формы текстурнозначного изображения	244
9.4. Морфологический метод сжатия изображений текста	250
9.4.1. Математическая модель и форма изображения текста (251).	
9.4.2. Алгоритм морфологического сжатия изображения текста (254).	
9.5. Фильтрация гладкого фона	258
9.5.1. Математическая модель и форма изображения фона и сигнала (258). 9.5.2. Постановка и решение задачи фильтрации фона (259).	
9.6. Поиск области интерлейсинга	261
9.6.1. Модель изображения, искаженного интерлейсингом (262).	
9.6.2. Морфологический метод поиска области интерлейсинга (263).	
9.7. Измерение микрорельефа поверхности по набору его изображений	264
9.7.1. Математическая модель мультифокусного изображения (265). 9.7.2. Задача оценивания высоты рельефа поверхности (266). 9.7.3. Оценка высоты рельефа как оценка параметра формы (267).	
9.8. Классификация изображений и оценка параметров системы регистрации	270
9.8.1. Математическая модель формирования изображения (270).	
9.8.2. Классификация и оценка параметров измерительной аппаратуры (271).	
9.9. Цветовая сегментация на основе морфологического фильтра	273
9.10. Поиск отличий по форме в цветных изображениях	275
Приложения	276
А.1. Множества, подмножества, линейные пространства	276
А.1.1. Отношения. Частично упорядоченные множества. Решетки (276).	
А.2. Линейные операторы в пространствах изображений	278
А.2.1. Линейные операторы в евклидовых пространствах (278).	
А.2.2. Операторы ортогонального проецирования (278). А.2.3. Ко-	

нечномерная аппроксимация формы в широком смысле (280).	
А.2.4. Псевдообратный оператор (282).	
А.3. Операторы проецирования на выпуклые замкнутые множества . . .	284
А.3.1. Задача наилучшего приближения в нормированном пространстве (284).	
А.3.2. Выпуклое замкнутое множество евклидова пространства и проекторы на него (284).	
А.3.3. Аппроксимация проектора на выпуклое замкнутое множество (288).	
А.3.4. Проектор на выпуклый замкнутый конус евклидова пространства \mathcal{R}^n (290).	
А.4. Случайные элементы евклидова пространства.	296
А.4.1. Определение и свойства случайных элементов (296).	
А.4.2. Нормально распределенные случайные элементы евклидова пространства. Распределения Пирсона и Снедекора–Фишера (297).	
А.5. Методы проверки статистических гипотез	299
А.5.1. Нерандомизированный и рандомизированный критерии проверки статистической гипотезы (299).	
А.5.2. Симметрия задачи проверки гипотез. Инвариантные критерии (301).	
А.5.3. Симметрия задачи проверки гипотезы о независимости математического ожидания предъявляемого изображения от формы \mathcal{L}^N (303).	
А.5.4. Надежность статистической гипотезы. Простая гипотеза и простая альтернатива (307).	
А.5.5. Минимаксная надежность сложной гипотезы при сложной альтернативе (311).	
А.5.6. Оценки максимальной надежности (315).	
А.6. Возможность как альтернативная вероятности модель случайности	316
А.6.1. Вероятность: проблемы эмпирического построения и интерпретации (316).	
А.6.2. Возможность как мера предопределенности исходов стохастического эксперимента (317).	
А.6.3. Классы эквивалентных возможностей (318).	
А.6.4. Шкала значений возможности (318).	
А.6.5. Необходимость. Шкала значений необходимости (320).	
А.6.6. Возможность, максимально согласованная с вероятностью (321).	
А.6.7. Возможность: эмпирическая интерпретация и эмпирическое построение (322).	
А.7. Элементы теории возможностей	324
А.7.1. Интеграл. Определение и свойства (324).	
А.7.2. Мера возможности. Определение и свойства (326).	
А.7.3. Принцип относительности (327).	
А.7.4. Нечеткие множества (328).	
А.7.5. Нечеткие элементы (329).	
А.7.6. Нечеткие события (329).	
Список литературы.	331

Предисловие

Одним из самых эффективных инструментов получения информации об окружающем мире для человека является зрение. «Лучше один раз увидеть, чем сто раз услышать», — гласит народная поговорка. Действительно, с первого взгляда мы можем узнать знакомый предмет на своем рабочем столе или заметить, что кто-то нарушил знакомый порядок расположения предметов. Мы легко узнаем человека по его фотографии и выделяем в ряду изображений знакомые лица. «На глаз» мы можем оценить взаимное расположение предметов, их размеры и другие характеристики. Ко всему этому можно еще добавить, что подобные задачи мы решаем как при ярком солнечном свете, так и в сумерки, и при искусственном освещении, по фотографии или даже по рисунку. При этом качество изображений не играет заметной роли — даже малоконтрастные и сильно зашумленные фотографии подчас несут достаточно информации для узнавания изображенного объекта или сцены.

Эта легкость во многом определила интерес исследователей к изучению алгоритмов работы зрительного анализатора живых организмов; прогресс в этой области позволил бы конструировать автоматы, способные заменить человека во многих областях его деятельности. И хотя успехи в создании систем машинного зрения достигнуты впечатляющие, до сих пор такие системы не могут конкурировать с человеком при решении сложных задач распознавания, классификации и пр.

Возможно, одной из причин является то, что изображение, как правило, рассматривается в отрыве от модели его формирования. С формальной точки зрения изображением является скалярная или векторная функция, заданная на подмножестве плоскости, ее значения интерпретируются как яркость изображения. Стандартные методы анализа изображений состоят в применении к ним тех или иных преобразований, известных в математике или специально сконструированных для тех или иных целей [14, 15, 20, 27, 31, 35, 37, 65, 66, 85]. Результатом этих операций является некоторая система признаков, характерных для заданного класса изображений; далее методами теории распознавания образов [9–11, 19, 21, 61, 64, 90] решаются многие практические задачи. Однако как выбор признаков, так и способы оценки их значений — достаточно трудные задачи.

В отличие от описанного выше подхода морфологические методы анализа изображений основаны на математических моделях, связывающих изображения с объектами изображаемой сцены и условиями их регистрации. Если речь идет об *анализе сцены*, а не об анализе *изображения* как такового (как объекта исследования), то все изменения изображения сцены, возникающие при изменении условий получения изображения (освещения, характеристик видеокамеры и т. п.), следует признать несущественными. С самых общих позиций анализатор сцены по ее изображениям можно представить как «черный ящик», на вход которого поступает изображение, а на выходе содержатся сведения, позволяющие высказываться о содержании сцены. В идеале выходной сигнал «черного ящика» — характеристики объектов сцены — остается прежним, если в широких пределах менять свойства входного изображения, связанные с изменением освещенности сцены, разрешающей способности системы формирования видеоинформации и др., но не обусловленные изменением «геометрии» сцены. С другой стороны, изменения в самой сцене — удаление или добавление предметов, изменение их взаимного расположения, т. е. такие, которые не могут быть вызваны изменением условий наблюдения, — приводят к изменениям на выходе анализатора.

Центральным понятием морфологических методов анализа служит понятие *формы изображения*, понимаемой как часть информации, сохраняющаяся при вариациях условий формирования изображения. Формально она может быть определена, например, как инвариант преобразований, моделирующих изменения условий формирования изображений сцены.

Рассмотрим примеры задач, которые не могут быть решены без использования морфологических методов анализа изображений.

На рис. П.1 приведен ряд изображений, на которых наблюдатель видит практически одну и ту же сцену, узнает знакомые предметы, может оценить их форму, взаимное расположение, размеры. Можно заметить, что в правом нижнем углу изображения П.1, *в*, отсутствует фрагмент («бочка»), имеющийся на изображениях П.1, *а* и *б* (см. цветную вклейку).

В то же время нельзя сказать, что эти изображения одинаковы — они отличаются яркостью и контрастом. Но эти отличия обусловлены не свойствами сцены, а условиями регистрации — характером освещения сцены, чувствительностью видеокамеры и т. п. Выделяя информацию о предметах сцены, человек даже не задумывается о том, каково время экспозиции, откуда падает свет и т. п. Формальная постановка задачи узнавания сцен по их изображениям является одной из целей морфологических методов анализа изображений, описанных в данной монографии.



Рис. П.2

Другой пример связан с проблемой поиска заданного знака, например буквы «А», на странице книги. Для человека эта задача требует напряженного внимания, особенно если страница частично залита краской и плохо освещена. Морфологические методы позволяют уверенно отыскивать знаки заданной формы даже в сложной помеховой обстановке, такой, например, какая изображена на рис. П.2.

Известно, что можно определять «на глаз» расстояние между предметами, их размер, причем при размытых сливающихся с фоном границах объектов это делать труднее. Однако именно в таких условиях приходится работать при исследовании наночастиц, см. рис. П.3. Морфологические методы построены таким образом, что гарантируют максимальную точность определения координат центров и радиусов частиц нанопорошка.

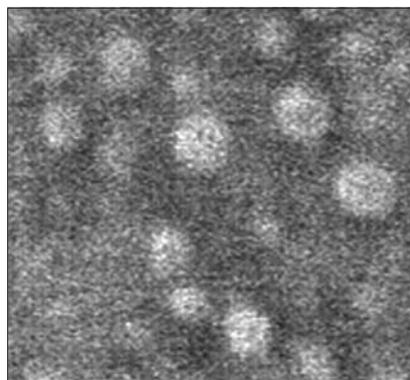


Рис. П.3

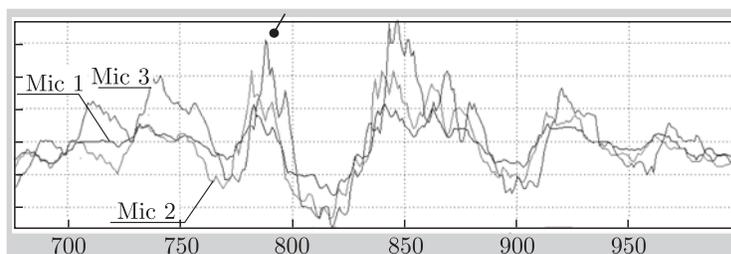


Рис. П.4

Еще один пример применения морфологических методов связан с проблемой совмещения сигналов, «в целом» достаточно схожих, но отличающихся деталями, см. рис. П.4; совмещение осуществляется путем сдвигов сигнала вдоль оси абсцисс. Эта задача возникает при определении временной задержки выходных сигналов трех микрофонов, регистрирующих акустический сигнал; значение относительного времени задержки сигналов на выходе микрофонов позволяет определить направление на источник звука. Различие трасс распространения звука обуславливает вариации регистрируемого звукового давления при сохранении общих особенностей сигнала. Морфологические методы позволяют дать формальное описание сходства сигналов и дают максимально точную оценку времени сдвига.

В настоящей книге описаны методы анализа и интерпретации изображений сцен, основанные на понятии формы изображения. Морфологический подход является еще одним шагом на пути анализа изображений с точки зрения содержащейся в нем информации, важной для решения поставленной задачи.

Издание этой книги стало возможным благодаря финансовой поддержке Российского фонда фундаментальных исследований (грант №09-07-07013) и помощи наших друзей и коллег. Авторы глубоко благодарны И. Богданову, В. Баскакову, С. Введенскому, Д. Делюкину, Д. Демину, Г. Животникову, А. Захарченко, В. Илюшину, П. Кобыльчаку, А. Корнильеву, И. Морозовой, Д. Устинину, И. Фаломкину, О. Фаломкиной, В. Шишакову и другим, предоставившим результаты расчетов, иллюстрирующих морфологические методы анализа данных.

Введение

Рассмотрим ряд изображений на рис. В.1 (см. цветную вклейку). Легко заметить, что это изображения одного и того же объекта — цветка горлицы. Наблюдатель воспринимает содержание изображенной сцены как неизменное, даже если яркость, контраст, цветовая гамма изображения изменяются в широких пределах. Из этого можно сделать вывод, что для решения таких задач, как узнавание и классификация объектов, оценивание их геометрической формы и т. п., важны не точные значения яркости и/или цвета изображения в каждой точке поля зрения, а знание некоторых структур на поле зрения, которые вызывают зрительное впечатление цветка и не меняются при изменении яркости, контраста и т. п.

С другой стороны, в реальных условиях нам, как правило, недоступна информация об условиях формирования изображения, позволяющая однозначно связать яркость и цвет изображения со свойствами изображаемых объектов сцены. Та совокупность математических характеристик изображений, которая независимо от условий их регистрации позволяет выделять, узнавать объекты и оценивать их геометрические характеристики, выделять эволюционирующие во времени элементы сцены на ее изображениях, полученных при произвольно меняющихся условиях регистрации и т. п., получила название *формы изображения*, а представленные в этой книге математические методы анализа и интерпретации изображенных сцен названы *морфологическими* [40, 47, 77–79]. Основу морфологических методов анализа изображений составляют математические модели и специально разработанный математический формализм, позволяющие охарактеризовать изображения в терминах инвариантов преобразований распределений их яркости и цвета по полю зрения, сохраняющих информацию, необходимую для решения поставленной задачи. Принципам морфологического анализа изображений и сигналов посвящены работы [40, 47, 77–79].

Проиллюстрируем понятие формы изображения на простом примере. Рассмотрим полутоновые монохромные (*gray-scale*) изображения однородно освещенного кубика, рис. В.2, полученные с помощью фотоаппарата. На нем представлены фрагменты, изображающие три видимые грани кубика, и фон. Благодаря однородности освещения и постоянству оптических свойств граней кубика и фона эти фрагменты

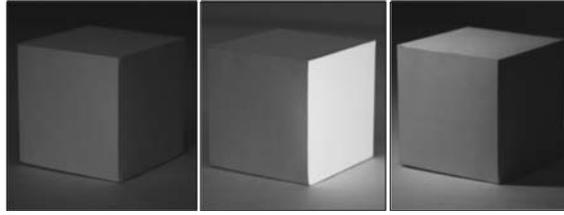


Рис. В.2. Изображения кубика при различных условиях наблюдения

изображения имеют примерно одинаковую яркость. Изменения условий однородного освещения приводят к изменению яркостей этих фрагментов, но геометрические свойства однородно освещенных и окрашенных областей поля зрения останутся прежними — они определяются формой объекта, геометрическими свойствами его поверхности и не зависят от условий регистрации изображений.

Следовательно, существенным для передачи геометрических характеристик кубика является разбиение поля зрения на множества примерно одинаковой яркости, а несущественным — яркость этих областей. В качестве модели каждого изображения кубика примем функцию, заданную на прямоугольной области (на поле зрения), разбитой на подмножества, соответствующие видимым граням кубика и фону.

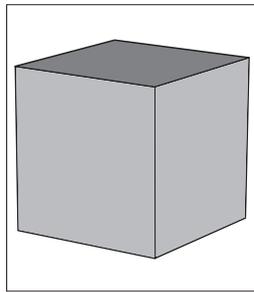


Рис. В.3. Модель изображения кубика как кусочно-постоянное изображение

Значения функции определяются яркостью изображения, они постоянны на каждом из подмножеств, см. рис. В.3. Изменяя яркость каждого из подмножеств в произвольных пределах, получим множество всех изображений кубика, которые могут быть получены при вариации условий их регистрации.

В этом случае решение задачи узнавания сцены (кубик в заданном ракурсе на однородном фоне при однородном освещении) состоит в том, чтобы проверить, принадлежит ли предъявленное изображение, обозначим его f , этому множеству изображений или нет. Факт принадлежности изображения f указанному множеству означает, что можно указать условия регистрации, при которых полученная фотография с приемлемой точностью совпадет с предъявленным изображением, и, следовательно, нет причин утверждать, что f не изображает заданную сцену с кубиком. В противном случае f не может быть порождено кубиком ни при каких условиях наблюдения — его следует признать изображением другой сцены.

Инвариант условий регистрации — в рассматриваемом примере множество всех изображений кубика — называется *формой изображения*¹⁾.

Описанный здесь подход может быть применен для узнавания произвольной сцены, представленной полутоновым изображением. Рассмотрев изображения сцены, полученные при всех возможных условиях их регистрации, получим множество \mathcal{V} всех изображений рассматриваемой сцены. Это множество называется *формой изображения сцены*. Охарактеризовать его можно, задав конструктивную процедуру проверки принадлежности любого изображения этому множеству. Например, если изображения рассматриваются как элементы некоторого метрического пространства \mathcal{R} , так что определено расстояние между любыми двумя изображениями, а множество \mathcal{V} замкнуто в \mathcal{R} , то такой процедурой является вычисление расстояния от предъявленного изображения до множества всех возможных изображений сцены. (Расстояние от некоторого элемента f до заданного множества \mathcal{V} определяется как точная нижняя грань расстояний от f до элементов множества \mathcal{V} ; для замкнутых множеств точная нижняя грань достигается на некотором элементе Pf множества \mathcal{V} , его называют проекцией f на \mathcal{V} .) Равенство нулю этого расстояния означает, что $f \in \mathcal{V}$, т. е. что предъявлено изображение рассматриваемой сцены, отвечающее некоторым условиям регистрации.

В рассмотренном примере инвариантным относительно условий регистрации является равенство или неравенство нулю расстояния между предъявленным изображением и его проекцией на множество всех изображений сцены, и форма изображения сцены в этом случае может быть отождествлена с операцией вычисления проекции. В этой книге методы анализа изображений в терминах их формы называются *методами морфологического анализа*.

Очертим круг задач, для решения которых используются методы морфологического анализа.

Задача узнавания объекта по его изображению. С формальной точки зрения, узнать заданную сцену или объект на изображении — значит, определить, может ли эта сцена или объект при некоторых условиях регистрации дать предъявленное для анализа изображение.

Рассмотрим для примера изображение листа календаря, приведенное на рис. В.4, а. Несмотря на весьма низкое качество изображения, можно заметить, что в каждой таблице числа, означающие дни месяца, содержат три изображения цифры «пять». Но как указать формальное

¹⁾ В данном случае форма изображения кубика содержит все, что может сообщить о геометрической форме кубика его изображение.

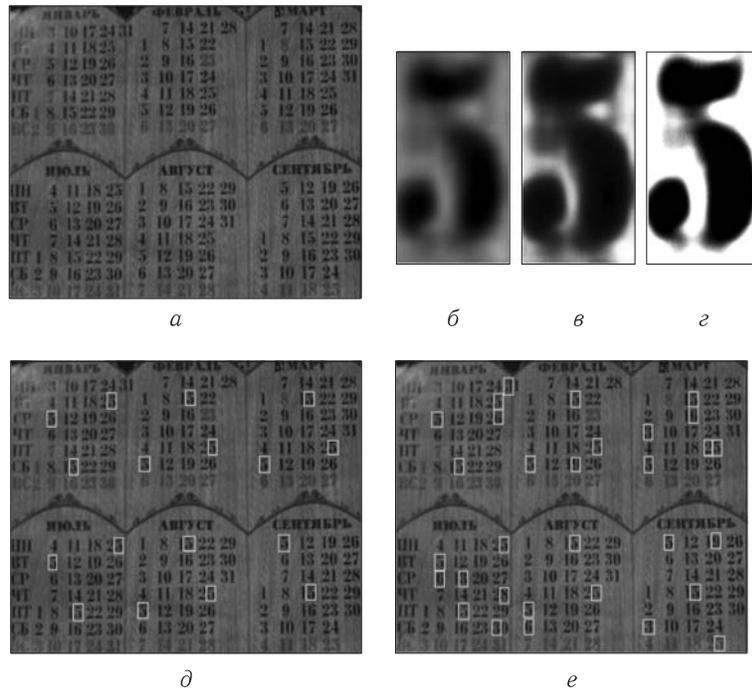


Рис. В.4. а) Изображение календаря. б)–г) Изображения цифры «пять». д) Результат узнавания; есть пропущенные фрагменты. е) Результат узнавания; имеются ложные срабатывания

правило, которое позволит узнать изображение этой цифры «автоматически», независимо от условий получения изображения В.4, а?

В простейшей ситуации для решения этой задачи морфологически-ными методами следует определить форму изображения цифры «пять». Рассмотрим для этого увеличенное изображение цифры «пять», взятое из фрагмента изображения календаря, соответствующего дате «25 февраля», см. рис. В.4, б. Будем считать, что все другие изображения цифры «пять» получены из него путем нелинейного преобразования его яркости (варианты таких изображений приведены на рис. В.4, в, г). Множество \mathcal{V} таких изображений будем считать формой изображения фрагмента, содержащего изображение цифры «пять». Изображения из множества \mathcal{V} будем считать *сравнимыми по форме* с изображением цифры «пять» на рис. В.4, б. Все изображения, не содержащиеся в множестве \mathcal{V} , будут не сравнимы по форме с изображением, приведенным на рис. В.4, б. В этом смысле формой изображения цифры «пять» можно считать множество изображений, сравнимых с ним по форме.

В идеальном варианте узнавание цифры «пять» на предъявленном фрагменте сводится к выяснению, принадлежит ли указанный фрагмент определенной таким образом форме изображения. Однако фрагменты изображения, содержащие другие цифры «5», отличаются от рис. В.4, б, не только преобразованием яркости: на изображении присутствуют шумы, знаки могут быть плохо пропечатаны и т. п.

Поэтому будем считать, что в предъявленном фрагменте можно узнать цифру «пять», если его яркость можно «с достаточной точностью» приблизить изображениями из формы изображения цифры «пять», т. е. полученными из изображения на рис. В.4, б, нелинейными преобразованиями его яркости (определения понятий «точность приближения», «нелинейное преобразование яркости» и др. будут даны ниже в зависимости от рассматриваемых математических моделей изображений).

Результат поиска участков поля зрения, содержащих изображение цифры «пять», приведен на рис. В.4, д, е. Найденные участки выделены белыми прямоугольниками. Фрагменты изображения на выделенных участках признавались изображением цифры «пять», если отличие (например, норма разности) между приближаемым фрагментом и его проекцией на множество всех изображений этой цифры (проекцией «должным образом» сдвинутого фрагмента на форму изображения цифры «пять») не превосходило заданный порог.

Изменяя значение порога, можно задавать более или менее жесткие условия узнавания. Так, малое значение порога привело к результату, изображенному на рис. В.4, д. Здесь нет ни одного ложного срабатывания, однако оказалась не узнанной цифра в дате «25 сентября», отличающаяся малой контрастностью. Увеличение порога приводит к тому, что пропущенный ранее фрагмент признается теперь как содержащий цифру «пять», однако есть и ошибочные узнавания — в датах «26 января», «31 января», «19 февраля» и др.

Поиск фрагмента, сравнимого по форме с изображением знака «шесть», иллюстрирует рис. В.5. В верхней его части приведено изображение ряда цифр. Изображения искажены небольшим шумом. Область прямоугольной формы, в которую целиком помещается изображение одного знака, движется по полю зрения в горизонтальном направлении, и рассматривается фрагмент изображения, вырезаемый этой областью. Этот фрагмент приближается изображениями, сравнимыми по форме с изображением цифры «шесть». График точности этого приближения приведен на нижней части рис. В.5.

Как видно из рисунка, фрагменты изображения, содержащие цифры «три», «пять», «восемь», «девять» и «ноль», достаточно близки к форме изображения цифры «шесть» — об этом свидетельствуют довольно глубокие локальные минимумы на приведенном графике, однако наи-

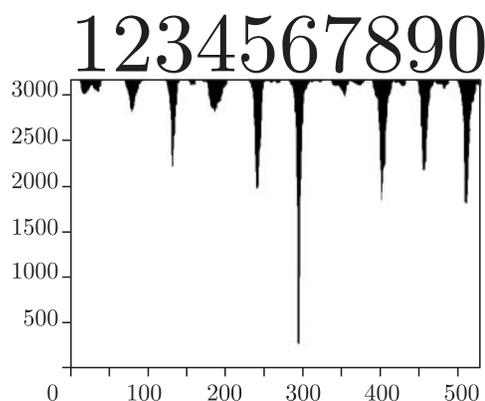


Рис. В.5. Вверху — изображение цифр, внизу — точность приближения подвижного фрагмента изображениями цифры «шесть» в зависимости от его расположения

более глубокий минимум соответствует предъявленному фрагменту, действительно содержащему цифру «шесть».

Создание математических методов решения задачи узнавания объектов по их изображениям весьма актуальны для разработки охранных систем, для систем контроля, когда, например, на изображении, формируемом видеокамерой, требуется узнать автомобиль заданной марки или в отпечатке пальца узнать характерный участок папиллярного узора, и др.

Здесь описан простейший морфологический метод узнавания объекта по его изображению. Заметим, что форму изображения можно построить, исходя из математической модели изображения. Например, если известны геометрическая форма объектов сцены, оптические свойства их поверхностей, условия освещения и модель устройства, формирующего изображения, то можно указать, какое изображение данной сцены может быть получено или как будет выглядеть фрагмент изображения заданного объекта. Так, в примере с изображением кубика, см. рис. В.2, можно считать, что все его изображения — кусочно-постоянные, принимающие значения, равные константе на областях поля зрения, соответствующих трем граням кубика и фону. Варьируя неизвестные параметры модели, можно получить все множество изображений объекта или сцены. На практике для того, чтобы указать множество изображений сцены, как правило, достаточно указать лишь самые общие характеристики объекта и системы регистрации и на этой основе построить множество изображений, которые может породить

данный объект. Остается лишь проверить, входит ли предъявленное для анализа изображение в это множество.

Однако в примере с узнаванием цифр форма изображения задавалась без обращения к математической модели, описывающей физические процессы его формирования. Достаточно было указать, что, например, при изменении контраста изображения знаков сохраняют свои характерные особенности. Точно так же не произойдет перепутывания изображений цифр, если изменять в широких пределах яркость изображений. В рассмотренном примере был задан класс возможных преобразований $\mathcal{F} : f \rightarrow F \circ f$, изменяющих яркость исходного изображения f по следующему правилу: яркость $f(x)$ точки x преобразуется в яркость $F \circ f(x) = F(f(x))$ независимо от расположения точки x на поле зрения (т.е. функция $F(\cdot)$ не зависит от x). Точки поля зрения, различающиеся по яркости на изображении f , после преобразования $F \circ f$ могут получить одинаковую яркость. Если же на изображении f яркости точек x и y равны, $f(x) = f(y)$, то это равенство сохранится и для преобразованного изображения: $F \circ f(x) = F \circ f(y)$. Таким образом, в результате преобразования F изображение может стать «более простым» по своей форме, так как области различной яркости изображения f могут получить одну и ту же яркость на изображении $F \circ f$, при этом соответствующая область постоянной яркости изображения $F \circ f$ будет объединением областей постоянной яркости изображения f .

Это наблюдение позволяет определить сравнение изображений по форме. Пусть задан некоторый класс \mathcal{F} преобразований изображения f . Будем говорить, что *форма g не сложнее, чем форма f* , если изображение g может быть представлено в виде $g = F \circ f$ при некотором $F \in \mathcal{F}$. Множество изображений, форма которых не сложнее формы f , назовем *формой изображения f* . На практике форму изображения можно определить, указав максимальный инвариант преобразований яркости изображения f .

Задача классификации объектов по их изображениям.

Рассмотрим изображение, содержащее цифры от нуля до девяти, см. рис. В.6, а. В простейшей задаче классификации требуется указать, какой объект изображен на фрагменте поля зрения, содержащем, например, цифру «семь», см. рис. В.6, б.

В отличие от предыдущего случая здесь заранее известно, изображения каких объектов могут быть предъявлены для анализа. В этом смысле классификация является более простой задачей, чем узнавание, так как в последнем случае множество альтернатив неопределено. Так же, как и при решении предыдущей задачи, для каждой цифры определяется класс всех ее возможных изображений (форма). Далее

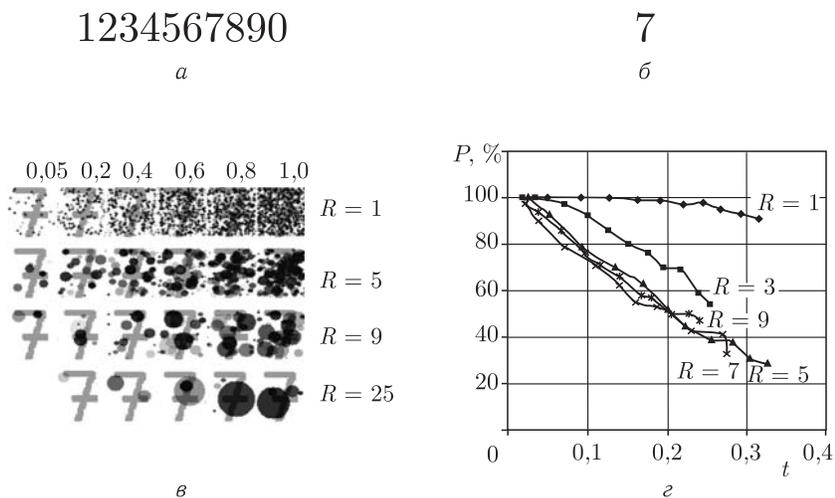


Рис. В.6. а) Изображения цифр. б) Изображение цифры «7». в) Искажённые шумом изображения цифры «семь». г) Частота правильной классификации в зависимости от отношения «шум/сигнал»

принимается решение, с какой из полученного набора форм наиболее схоже предъявленное изображение. Для этого находится ближайшая к нему форма изображений, для чего решается задача его наилучшего приближения изображениями из заданной формы.

Если есть несколько множеств, наиболее близких к заданному, то для выбора одного из них требуется привлечь дополнительные соображения.

Морфологические методы классификации обладают высокой помехозащищённостью. Для иллюстрации этого свойства на рис. В.6, в, приведен набор изображений цифры «семь», искажённых аддитивным шумом в виде пятен в форме кругов различных радиусов; яркость пятен однородна и случайно изменяется от пятна к пятну. Изображения на рис. В.6, в, расположены в виде таблицы. Каждая строка таблицы содержит изображения цифры «семь», искажённые пятнами одного и того же радиуса, столбцы таблицы содержат изображения с одинаковым отношением «шум/сигнал», понимаемого как отношение квадрата нормы шумового изображения к квадрату нормы «полезного» изображения. Значение этого отношения приведено над каждым из столбцов.

Графики частоты ошибочных решений в зависимости от величины отношения «шум/сигнал» для различных способов искажения (для радиусов кругов $R = 1, 3, 5, 7, 9$) приведены на рис. В.6, г. Видно, что чем меньше радиус круга, тем более помехоустойчив метод классификации.

Задача выделения объектов, представленных на изображениях.

На рис. В.7 сверху приведены изображения f и g двух сцен. Изображения получены при разных условиях освещения, а сцены отличаются тем, что на одной из них (изображенной на g) появился новый объект — кошка. Требуется по изображениям указать отличия в сценах.

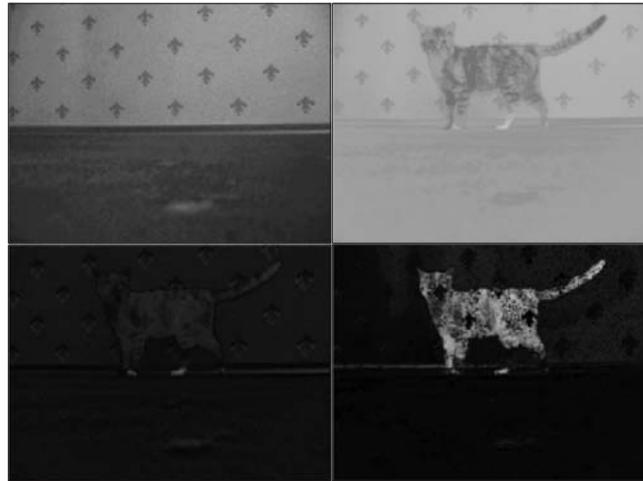


Рис. В.7. Изображение двух сцен (вверху), разность изображений (внизу слева) и отличие по форме (внизу справа)

Изображения получены при различных условиях, из-за этого яркости областей, соответствующих одинаковым объектам сцен, могут существенно различаться. Поэтому не имеет смысла вычитать поточечно яркость одного изображения из яркости другого — об этом свидетельствует результат, приведенный на рис. В.7 слева внизу. «Серое» изображение показывает, что яркости соответствующих точек заметно отличны друг от друга.

Определим форму изображения f , представленного на рис. В.7 слева вверху, как множество изображений сцены, полученных при всех возможных условиях его регистрации, при этом изменение условий регистрации будем моделировать преобразованием яркости исходного изображения f .

Можно заметить, что появление кошки в комнате приводит к нарушению формы изображения f , поскольку никаким (нелокальным) изменением яркости невозможно привести изображение f к изображению g . Морфологический метод выделения отличия предъявленного изображения g от заданного f по форме сводится к нахождению тех особенностей изображения g , которые не могут быть получены из f

путем преобразования яркости f : отличия в форме содержатся в разности изображения g и его наилучшего приближения изображениями из формы f . Эта разность приведена на рис. В.7 справа внизу. В результате такого морфологического вычитания на темном фоне появляется достаточно яркий объект (кошка), показывающий отличие по форме изображения g от f .

Оценивание параметров объекта по его изображению.

Часто на практике оказываются неизвестными параметры объектов сцены, определяющие, например, их расположение, размеры и т. п. Информация об этих параметрах содержится в изображении. На рис. В.8 приведено изображение двух кубиков. Несмотря на то, что неизвестны многие параметры модели, связывающей рассматриваемую сцену с его изображением, предоставленной информации достаточно для того, чтобы оценить, насколько отличаются размеры изображенных кубиков,

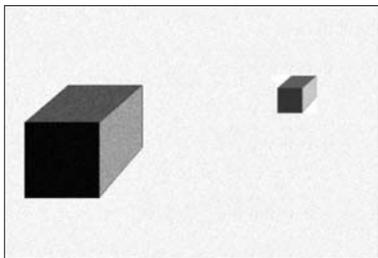


Рис. В.8. Изображение несет информацию о размере и/или удаленности объекта

если известно, что они находятся на одинаковом расстоянии от фотокамеры, или насколько отличаются расстояния до них, если размеры кубиков одинаковы.

Другой пример связан с определением координат заданного знака на поле зрения. В электронной литографии перед напылением слоя вещества на поверхность пластины кремния требуется расположить эту пластину в строго определенном положении. Для контроля положения пластины на ее поверхность наносится реперный знак, изображенный на рис. В.9 слева сверху. Изображение этого знака в электронном микроскопе позволяет определить координаты пластины и скорректировать ее положение. В рассматриваемом примере ситуация несколько усложняется тем, что при напылении репер (крест) оказывается под слоем вещества, что приводит к размытию его изображения, причем контролировать это размытие достаточно сложно. Тем не менее информация о его положении на поле зрения присутствует в изображении. На рис. В.9 справа сверху показано, как в электронном микроскопе выглядит под слоями напыленного резиста; это изображение получено компьютерным моделированием процесса рассеяния электронов в материале слоев методом Монте-Карло, [67].

Для определения координат репера по его размытому изображению задается форма изображения репера как множество его размытых изображений со всевозможными способами размытия; форма зависит

от двумерного вектора параметров, которыми являются координаты креста на поле зрения. Далее вычисляется расстояние от предъявленного изображения до формы размытого изображения репера для различных значений координат центра репера на поле зрения. График зависимости этого расстояния от значений координат центра репера приведен на рис. В.9 слева внизу. Ярко выраженный минимум этой зависимости позволяет с точностью, достаточной для технологии изготовления микросхем, определять положение репера: цена деления на шкале координат равна 0,1 мкм. Найденное расположение репера на фрагменте предъявленного изображения приведено на рис. В.9 справа внизу. Заметим, что визуальное совмещение в этом случае практически невозможно.

Все эти примеры связаны с необходимостью оценивать по изображению те или иные параметры сцен. Для решения этой задачи с каждым фиксированным значением параметра сцены свяжем множество изображений этой сцены. С формальной точки зрения следует найти такое значение параметров, для которого множество определяемых им изображений включает предъявленное изображение. Если такой набор не единственен, то для выбора нужной оценки следует привлекать дополнительные соображения, а если предъявленное изображение не включено ни в какое множество, следует искать наиболее похожие по форме.

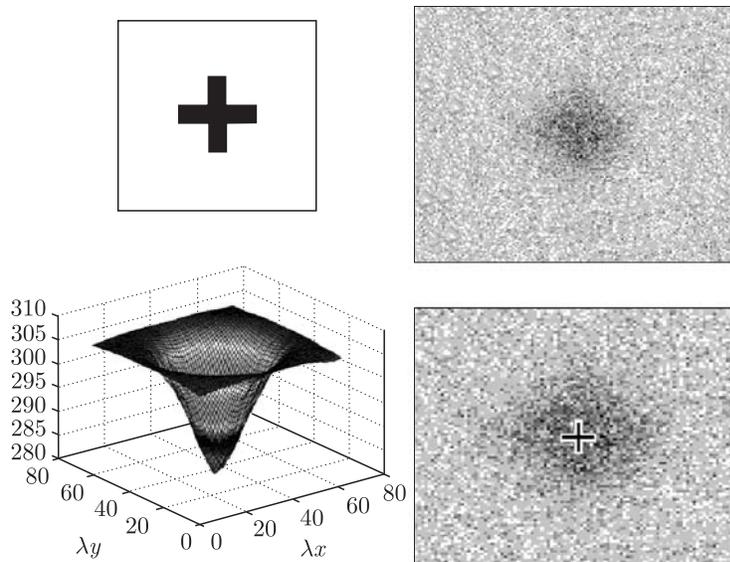


Рис. В.9. Оценка положения репера по его изображению

Однако если нас интересует не столько сцена, на которую в наибольшей степени похоже предъявленное изображение, а *наиболее точная оценка* параметра λ (в рассмотренном примере λ — вектор из координат репера), следует поставить задачу так, чтобы оценка $\hat{\lambda}$ параметра λ выбиралась из условий минимума погрешности оценки. Для решения этой задачи требуется уточнить модель формирования предъявляемого изображения.

Пусть, например, известно, что координаты λ изображенного репера могут принимать значение из некоторого заданного множества Λ на плоскости. Множество изображений репера с координатами $\lambda \in \Lambda$ обозначим $\mathcal{V}(\lambda)$. Рассмотрим схему формирования изображения f репера в виде равенства

$$\xi = f + \nu, \quad (\text{B.1})$$

в котором «неискаженное» изображение f принадлежит множеству $\mathcal{V}(\lambda)$, $\lambda \in \Lambda$, а «шумовое изображение» ν , моделирующее погрешности, возникающие при формировании изображения f , принадлежит заданному подмножеству \mathcal{N} множества всех изображений. Задача состоит в выборе оценки $\hat{\lambda}$ для $\lambda \in \Lambda$, удовлетворяющей условию

$$\|\lambda - \hat{\lambda}\| = \inf_{\lambda' \in \Lambda} \max\{\|\lambda' - \lambda\| \mid \lambda : \xi = f + \nu, f \in \mathcal{V}_\lambda, \nu \in \mathcal{N}\}. \quad (\text{B.2})$$

Согласно (B.2) оценка $\hat{\lambda}$ минимизирует максимально возможную погрешность оценивания параметров координат знака $\lambda \in \Lambda$ и называется минимаксной [45].

Для решения задачи (B.2) построим множество $\Lambda_\xi \subset \Lambda$ значений параметра λ , при которых возможно равенство (B.1) при некоторых $\nu \in \mathcal{N}$ и $f \in \mathcal{V}_\lambda$. Это множество содержит те и только те значения вектора λ , для которых отличие предъявленного изображения ξ от некоторого изображения из множества $\mathcal{V}(\lambda)$ изображений, порожденных репером с координатами λ , может быть объяснено наличием шума $\nu \in \mathcal{N}$.

Решением задачи (B.2) в этом случае является центр шара минимального радиуса, содержащего множество Λ_ξ . Радиус этого шара дает оценку погрешности определения искоемых координат.

Продемонстрируем работу метода минимаксного оценивания параметров объектов на примере анализа размеров и положений двух наночастиц по их изображению в растровом электронном микроскопе, см. рис. В.10, а. Известно, что частица в идеальных условиях изображается областью поля зрения в виде круга, яркость которой отличается от яркости фона, а радиус круга изменяется в некоторых заданных пределах. При регистрации изображение искажается аддитивным шумом ν , принадлежащим заданному ограниченному множеству. Требуется по

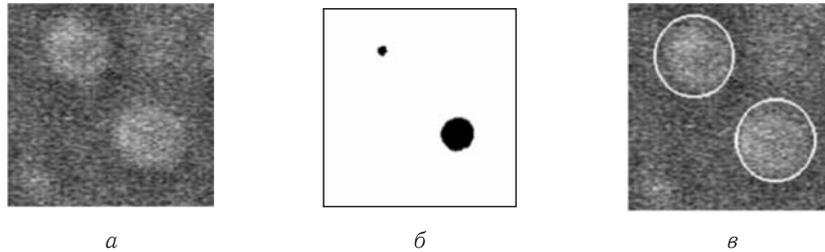


Рис. В.10. Оценка размеров и расположения наночастиц по их изображению

данному изображению определить координаты центра и радиус круга, изображающего наночастицу. Положение частиц априори произвольно.

Существует целое множество значений центров кругов с радиусами, лежащими в заданных пределах, при которых их изображения будут отличаться от предъявленного на рис. В.10, *a*, не более, чем на элемент из множества \mathcal{N} . Это множество центров показано на рис. В.10, *б*, черным цветом. Оно состоит из двух односвязных областей, первая из них оценивает возможное положение центра первой частицы, вторая — второй. Минимаксной оценкой положения наночастицы является центр круга минимального радиуса, содержащего соответствующую область на рис. В.10, *б*; его радиус задает величину минимаксной погрешности оценивания.

Точно так же найдется несколько значений радиусов двух кругов, лежащих в поле зрения изображения рис. В.10, *a*, для которых изображение этих кругов отличается от рис. В.10, *a*, не более, чем на элемент из \mathcal{N} . Минимаксной оценкой радиуса частицы является середина отрезка, содержащего возможные значения радиуса соответствующего круга. На рис. В.10, *в*, изображены круги, радиусы и положения центров которых равны минимаксным оценкам параметров соответствующих наночастиц.

Если «шумовое изображение» ν в (В.1) — случайный элемент пространства изображений с заданным законом его распределения вероятностей, то можно построить для λ оценивающие множества минимального размера, [45].

Часть I
ФОРМА ПОЛУТОНОВОГО ИЗОБРАЖЕНИЯ

Глава 1
**ОСНОВНЫЕ ПОНЯТИЯ
МОРФОЛОГИЧЕСКОГО АНАЛИЗА**

1.1. Математические модели изображения

1.1.1. Линейное пространство изображений. С математической точки зрения под изображением будем понимать числовую функцию $f(\cdot)$, заданную на ограниченном подмножестве X плоскости \mathcal{R}^2 . Область X будем называть *полем зрения*, а значение $f(x)$ функции $f(\cdot)$ в точке $x \in X$ — *яркостью в точке x* поля зрения X .

На практике для представления изображений в памяти компьютера в качестве поля зрения X выбирают набор узлов прямоугольной сетки, а яркость в каждом узле задают целым числом в интервале от 0 до 255. Однако такое множество объектов лишено многих привлекательных свойств: при сложении или вычитании яркостей двух изображений результат может выйти за рамки множества яркостей $\{0, 1, \dots, 255\}$. Для теоретического анализа проблем, возникающих при создании методов морфологического анализа, условимся считать, что поле зрения X — произвольное ограниченное подмножество плоскости \mathcal{R}^2 , а яркость изображения в каждой точке поля зрения X принимает конечное числовое значение: $f(x) \in (-\infty, \infty)$, $x \in X$. Определим линейные операции сложения изображений и умножения изображения на число следующим образом:

$$(f + g)(x) = f(x) + g(x), \quad (a \cdot f)(x) = a \cdot f(x), \quad x \in X.$$

В этом случае множество всех изображений, заданных на поле зрения X , представляет собой линейное пространство.

Заметим, что линейность пространства изображений приводит к тому, что функции, принимающие отрицательные значения на поле зре-

ния X , тоже считаются изображениями. Отрицательные яркости могут возникнуть, если пользоваться логарифмической шкалой; в этом случае $f(x) = k \ln I(x)$, где $I(x)$ — световая энергия, попадающая на элементарную площадку, расположенную в точке $x \in X$, k — калибровочный коэффициент. Логарифмическая шкала удобна тем, что экспоненциальный рост яркости воспринимается зрительным аппаратом человека как линейный.

Обозначим $\mu(\cdot)$ некоторую меру на σ -алгебре борелевских подмножеств поля зрения X . Будем считать, что существует интеграл от квадрата изображения по полю зрения X :

$$\int_X f^2(x) d\mu(x) < \infty. \quad (1.1)$$

В качестве меры μ подмножества A поля зрения X будем использовать либо его площадь (меру Лебега), либо так называемую считающую меру, когда на поле зрения X задано конечное множество точек (узлов сетки), и мера множества $A \subset X$ равна числу узлов сетки, принадлежащих множеству A . Форму записи (1.1) будем использовать только для того, чтобы единым образом представить два случая: когда поле зрения X представляет собой ограниченное подмножество плоскости \mathcal{R}^2 , интеграл понимается «в обычном смысле» (как интеграл Лебега):

$$\int_X f(x) d\mu(x) = \int_X f(x) dx.$$

Если же заданы сетка узлов $\{x_1, \dots, x_n\} \in X$ и считающая мера, то

$$\int_X f(x) d\mu(x) = \sum_{i=1}^n f(x_i).$$

В зависимости от того, как определено расстояние между изображениями, будем рассматривать различные метрические пространства изображений.

1.1.2. Пространство изображений $\mathcal{L}_\mu^2(X)$. Если выполнено (1.1), то для любых двух изображений f и g можно определить скалярное произведение по формуле

$$(g, f) = \int_X f(x)g(x) d\mu(x) \quad (1.2)$$

и норму изображения

$$\|f\| = \left(\int_X f^2(x) d\mu(x) \right)^{1/2}.$$

Расстояние между изображениями f и g при этом определяется нормой разности $f - g$:

$$\rho(f, g) = \|f - g\| = \left(\int_X (f(x) - g(x))^2 d\mu(x) \right)^{1/2}.$$

Таким образом определенное линейное пространство изображений со скалярным произведением (1.2) называется *евклидовым пространством* $\mathcal{L}_\mu^2(X)$.¹⁾ Если какое-либо соотношение выполнено для всех точек множества X , кроме точек множества нулевой меры, будем говорить, что оно выполнено почти всюду на X , и обозначать это символом $(\text{mod } \mu)$. В частности, для эквивалентных функций $f, g \in \mathcal{L}_\mu^2(X)$ можно записать $f(x) = g(x)$, $x \in X \pmod{\mu}$.

В том случае, когда на X заданы сетка узлов и считающая мера, для вычисления нормы, скалярного произведения и расстояния между изображениями достаточно знать значения изображений в узлах сетки. Тогда случае любое изображение $f \in \mathcal{L}_\mu^2(X)$ можно представить как вектор n -мерного пространства с координатами $(f(x_1), \dots, f(x_n)) \in \mathcal{R}^n$ и

$$(g, f) = \sum_{i=1}^n f(x_i) g(x_i),$$

$$\|f\| = \left(\sum_{i=1}^n f^2(x_i) \right)^{1/2},$$

$$\rho(f, g) = \left(\sum_{i=1}^n (f(x_i) - g(x_i))^2 \right)^{1/2}.$$

1.1.3. Пространство изображений $\mathcal{C}(X)$. Рассмотрим класс $\mathcal{C}(X)$ кусочно непрерывных изображений, заданных на ограниченном подмножестве плоскости — на поле зрения $X \subset \mathcal{R}^2$: изображение f принадлежит классу $\mathcal{C}(X)$ тогда и только тогда, когда существует конечное измеримое разбиение A_1, \dots, A_N поля зрения X , такое, что на каждом из A_i , $i = 1, \dots, N$, функция $f(\cdot)$ непрерывна, причем

$$\varphi(f) = \sup_{x \in X} |f(x)| < \infty.$$

¹⁾ Более аккуратное определение пространства $\mathcal{L}_\mu^2(X)$ и его свойства можно найти в [28]. Здесь отметим лишь, что если две функции $f, g \in \mathcal{L}_\mu^2(X)$ совпадают всюду, кроме, быть может, подмножества X , мера μ которого равна нулю, то расстояние $\rho(f, g) = 0$. Такие функции называются эквивалентными, и элементами пространства $\mathcal{L}_\mu^2(X)$ являются классы эквивалентных функций.

Функционал $\varphi(f)$, определенный на $\mathcal{C}(X)$, является нормой элемента $f \in \mathcal{C}(X)$; обозначим его $\|f\|_C$. Расстояние между изображениями f и g в этом случае определяется нормой разности $f - g$:

$$\rho_C(f, g) = \|f - g\|_C = \sup_{x \in X} |f(x) - g(x)|.$$

Для любого изображения $f \in \mathcal{C}(X)$ определен интеграл

$$\int_X |f(x)| d\mu(x) < \infty.$$

Линейное нормированное пространство изображений $\mathcal{C}(X)$ не является евклидовым: в нем нельзя определить скалярное произведение, согласованное с нормой $|\cdot|_C$.

В дальнейшем в общем случае для обозначения линейного пространства всех изображений будем использовать символ \mathcal{L} , указывая в нужных случаях, о каком из пространств, $\mathcal{L}_\mu^2(X)$ или $\mathcal{C}(X)$, идет речь. При этом если $\mathcal{L} = \mathcal{L}_\mu^2(X)$, равенства типа $f(x) = g(x)$, $x \in X$, понимаются как равенства почти всюду на $X \pmod{\mu}$.

1.2. Понятие формы изображения. Сравнение изображений по форме

В этом пункте перейдем к построению формы изображения как инварианта преобразований изображения, моделирующих изменение условий его регистрации. Как уже упоминалось, знание яркости изображения в каждой точке поля зрения является избыточным для решения множества задач анализа изображенной сцены. Значительно более важной является «структура» (называемая в морфологическом анализе *формой* изображения), не изменяющаяся при вариации условий его формирования. Примером такой структуры (*формы*) может служить разбиение поля зрения на множества постоянной яркости и/или цвета: на изображениях горицвета, приведенных на рис. В.1, область, соответствующая лепесткам цветка, всегда имеет одну и ту же геометрическую форму, сохраняющуюся при различных условиях регистрации изображения горицвета. Изображения кубика на рис. В.2, полученные при меняющихся в достаточно широких пределах условиях наблюдений, будут состоять из четырех областей поля зрения, на каждой из которых яркость изображения постоянна. Эти области соответствуют трем видимым граням кубика и фону. Заметим, что при некоторых условиях освещения яркости некоторых из этих областей могут оказаться одинаковыми, при этом поле зрения разобьется на области постоянной яркости, число которых будет меньше четырех.

В этом случае естественно считать, что форма такого изображения будет менее сложной, чем у исходного.

1.2.1. Модель и форма мозаичного изображения. Остановимся на последнем примере подробнее и построим достаточно простую математическую модель, позволяющую получить множество изображений, форма которых не сложнее формы изображения кубика на однородном фоне (или, иными словами, — множество изображений, сравнимых по форме с изображением кубика).

Вообще говоря, для создания математической модели изображения сцены необходимы сведения об излучении, падающем на объекты сцены, нужно знать их геометрическую форму, оптические свойства их поверхностей, а также фона, на котором они рассматриваются, и, наконец, модель системы, формирующей изображение. Понятно, что на практике столь подробные сведения вряд ли доступны, однако для решения описанных выше задач достаточно лишь самых общих представлений о том, как устроены объекты сцены, фон и система формирования их изображения.

Пусть выполнены следующие условия.

1. Освещение сцены однородно, но может меняться его интенсивность, направление, а также чувствительность устройства формирования изображений.
2. Изображаемая сцена состоит из расположенного на оптически однородном фоне кубика с оптически однородными гранями, так или иначе, что интенсивность отраженного света определяется только взаимным положением нормали к поверхности, направлением на наблюдателя, а также интенсивностью и направлением падающего на сцену светового излучения.
3. Кубик изображается при определенном ракурсе и на фиксированном расстоянии от системы формирования изображения.
4. Размеры объекта и изображаемого участка фона достаточно малы по сравнению с расстоянием до регистрирующей системы.

В первом приближении (если не учитывать переотражения от граней, затенения и другие подобные эффекты) эти условия означают, что из любой элементарной площадки однородной плоской грани объекта или фона в видеосистему приходит излучение одинаковой интенсивности, определяемое взаимным расположением нормали к поверхности, направлением на видеосистему и, вообще говоря, направлением падающего излучения, если поверхность не является диффузно отражающей (ламбертовой). Тогда каждая грань кубика будет изображаться областью поля зрения X , все точки которой имеют одинаковую яркость. То же самое можно сказать и об изображении фона: соответствующая ему область поля зрения на изображении сцены также будет характе-

ризоваться одинаковой яркостью. В результате изображение кубика на однородном фоне можно представить как кусочно постоянную функцию, заданную на поле зрения X ; области постоянной яркости соответствуют изображениям граней объекта и фона, см. рис. 1.1, справа. При изменении условий освещения (при сохранении его однородности), изменении коэффициента усиления видеокамеры и т. п. яркости областей могут меняться в широких пределах, однако сама геометрическая форма этих областей останется неизменной.

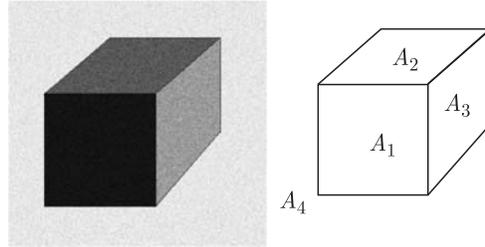


Рис. 1.1. Изображение кубика и области постоянной яркости

Следовательно, «часть информации», ответственная за геометрическую форму изображаемого объекта, содержится в форме областей постоянной яркости изображения, а их яркость относится к «части информации», несущественной с точки зрения геометрических свойств объекта, так как в рамках используемой модели формирования изображения эти яркости могут быть произвольными.

Модель изображения кубика на однородном поле зрения запишем в виде

$$f_{\text{cub}}(x) = \sum_{i=1}^4 c_i \chi_i(x), \quad x \in X. \quad (1.3)$$

Действительно, здесь все поле зрения X разбито на области $A_i \subset X$, $i = 1, \dots, 4$; все точки области A_i имеют одинаковую яркость c_i ,

$$\chi_i(x) = \begin{cases} 1, & x \in A_i, \\ 0, & x \notin A_i, \end{cases}$$

— индикаторная функция множества A_i , $A_i \cap A_j = \emptyset$ при $i \neq j$, $i = 1, \dots, 4$; $\bigcup_{i=1}^4 A_i = X$. Изменение условий наблюдения — перемещение и изменение яркости источника освещения, чувствительности видеокамеры и т. п. — приведет к изменению значений яркости c_i подмножеств A_i , $i = 1, \dots, 4$, поля зрения X .

Если объект представляет собой выпуклый многогранник с оптически однородными поверхностями граней, освещенный однородным

потоком света и изображаемый на однородном фоне, то его изображение запишется в виде

$$f(x) = \sum_{i=1}^{N+1} c_i \chi_i(x), \quad x \in X. \quad (1.4)$$

Здесь N — число граней, видимых из точки наблюдения, c_i — яркость изображения грани A_i , $i = 1, \dots, N$; c_{N+1} — яркость изображения фона. Множества одинаковой яркости, изображающие грани объекта, имеют индикаторные функции $\chi_i(\cdot)$ соответственно, $i = 1, \dots, N$, а фон — индикаторную функцию $\chi_{N+1}(\cdot)$. Изображения вида (1.4) являются частным случаем так называемых мозаичных изображений; в общем случае для мозаичного изображения яркость области A_i является непрерывной функцией точки $x \in A_i$, $i = 1, \dots, N$, см. ниже, п. 1.5.1.

Если в результате изменения условий наблюдения яркости c_i , $i = 1, \dots, 4$, могут принимать любые числовые значения, то множество всех изображений кубика можно записать в виде

$$\mathcal{V}_{f_{\text{cub}}} = \left\{ f(x) = \sum_{i=1}^4 c_i \chi_i(x), \quad x \in X, \quad c_i \in (-\infty, \infty), \quad i = 1, \dots, 4 \right\}.$$

Подмножество $\mathcal{V}_{f_{\text{cub}}}$ линейного пространства \mathcal{L} всех изображений представляет собой множество всех линейных комбинаций четырех линейно независимых функций χ_i , $i = 1, \dots, 4$, т. е. четырехмерное линейное подпространство в линейном пространстве всех изображений, и обладает тем свойством, что любое изображение кубика, полученное в результате изменений условий наблюдения, принадлежит этому множеству, а если некоторое изображение ему не принадлежит, то оно не может быть изображением кубика ни при каких условиях наблюдения.

В более общем случае, когда рассматривается объект с плоскими гранями на однородном фоне при однородном освещении, множество всевозможных его изображений

$$\mathcal{V}_f = \left\{ f(x) = \sum_{i=1}^{N+1} c_i \chi_i(x), \quad x \in X, \quad c_i \in (-\infty, \infty), \quad i = 1, \dots, N+1 \right\} \quad (1.5)$$

является $(N+1)$ -мерным линейным подпространством.

Определение 1.1. Множество \mathcal{V}_f , определенное в (1.5), называется *формой изображения объекта с плоскими оптически однородными гранями на однородном фоне при однородном освещении*, или, короче, *формой изображения $f(\cdot)$* .

Определенная таким образом форма изображения может рассматриваться как множество изображений фиксированной сцены при всех

возможных условиях регистрации. Заметим, что для задания этого множества достаточно указать разбиение поля зрения X на множества A_1, \dots, A_{N+1} ; все точки каждого из этих множеств на любом изображении из \mathcal{V}_f имеют одинаковую яркость.

Отметим следующий факт. Если для некоторого изображения яркости двух или большего числа множеств совпадают, то разбиение поля зрения на множества равной яркости содержит менее, чем $N + 1$, элементов. Как уже упоминалось, форма такого изображения считается менее сложной, чем форма изображения (1.4) с несовпадающими яркостями c_1, \dots, c_{N+1} . Множество \mathcal{V}_f в (1.5) можно теперь рассматривать как множество изображений, форма которых не сложнее, чем форма изображения

$$f(x) = \sum_{i=1}^{N+1} c_i \chi(x), \quad x \in X, \quad (1.6)$$

для которого яркости множеств A_1, \dots, A_{N+1} различны: $c_i \neq c_j$, $i \neq j$, $i, j = 1, \dots, N + 1$.

1.2.2. Операция сравнения изображений по форме. В этом пункте опишем формальный подход, позволяющий строить множества изображений с заданной или более простой формой, не обращаясь к модели мозаичного изображения.

Заметим, что множество изображений \mathcal{V}_f , определенное в (1.5), может быть получено из одного-единственного изображения (1.6). Для этого надо лишь для каждого $i = 1, \dots, N + 1$ произвольным образом изменять значения яркости c_i области A_i постоянной яркости заданного изображения (1.6). Формально это можно сделать, рассматривая множество изображений вида

$$g(x) = F(f(x)) \equiv \sum_{i=1}^{N+1} F(c_i) \chi(x), \quad x \in X, \quad (1.7)$$

где функция $F(\cdot)$ — любая из некоторого класса \mathbf{F} числовых функций, заданных на числовой прямой \mathcal{R}^1 . Далее для таких изображений с преобразованной яркостью будем использовать обозначение $g = F \circ f \in \mathcal{L}$:

$$g(x) = (F \circ f)(x) \equiv F(f(x)), \quad x \in X.$$

Естественно рассматривать в качестве множества всех возможных преобразований яркости класс \mathbf{F}_f всех таких функций, для которых результирующее изображение $F \circ f$, $F \in \mathbf{F}_f$, тоже является элементом пространства \mathcal{L} (в частности, если рассматривается евклидово пространство изображений $\mathcal{L}_\mu^2(X)$, то $F \circ f$ должна быть функцией, определенной на X , квадрат которой интегрируем — для этого достаточно, чтобы $F_f \in \mathbf{F}$ были ограниченными борелевскими функциями). В част-