

ФЕДЕРАЛЬНОЕ АГЕНТСТВО ПО ОБРАЗОВАНИЮ
ГОСУДАРСТВЕННОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ
«САМАРСКИЙ ГОСУДАРСТВЕННЫЙ АЭРОКОСМИЧЕСКИЙ
УНИВЕРСИТЕТ имени академика С.П. КОРОЛЕВА»

В.В. МЯСНИКОВ, Н.И. ГЛУМОВ, В.В. СЕРГЕЕВ

МЕТОДЫ ОБНАРУЖЕНИЯ И РАСПОЗНАВАНИЯ ОБЪЕКТОВ НА ЦИФРОВЫХ ИЗОБРАЖЕНИЯХ

*Утверждено Редакционно-издательским советом университета
в качестве учебного пособия*

САМАРА
Издательство СГАУ
2006

УДК 681.3, 621.372.542
ББК 22.343
М994



**Инновационная образовательная программа
"Развитие центра компетенции и подготовка
специалистов мирового уровня в области аэро-
космических и геоинформационных технологий"**

Рецензенты: д-р физ.-мат. наук, проф. А. И. Жданов,
д-р техн. наук, проф. В. Г. Карташевский

Мясников В.В.

М994 **Методы обнаружения и распознавания объектов на цифровых изображениях:** учеб. пособие / В.В. Мясников, Н.И. Глумов, В.В. Сергеев. - Самара: Изд-во Самар. гос. аэрокосм. ун-та; 2006. - 168 с.: ил.

ISBN 5-7883-04-99-7

В данном пособии рассматриваются методы обнаружения и распознавания объектов на цифровых изображениях. Приведена общая постановка задачи распознавания образов, которая конкретизирована для задач обнаружения, локализации и распознавания объектов на изображениях. Описаны основные группы признаков, используемые при распознавании цифровых изображений, приведены алгоритмы их быстрого вычисления для ряда практически важных задач. Большое внимание в книге уделено задаче поиска и локализации объектов на изображении, а также методам совместной классификации.

Книга предназначена для специалистов, работающих в области прикладной математики, а также для аспирантов и студентов старших курсов.

УДК 681.3, 621.372.542
ББК 22.343

ISBN 5-7883-04-99-7

© В.В. Мясников, Н.И. Глумов, В.В. Сергеев, 2006
© Самарский государственный
аэрокосмический университет, 2006

ОГЛАВЛЕНИЕ

1. Постановка задачи распознавания образов	4
1.1. Задача распознавания образов.....	4
1.2. Вероятностный критерий качества классификации	6
1.3. Задачи распознавания на изображениях.....	8
2. Оптимальные стратегии распознавания	11
2.1. Оптимальные стратегии.....	11
2.2. Байесовский классификатор для нормально распределенных векторов признаков.....	16
3. Основные группы признаков, используемых при распознавании изображений	24
3.1. Геометрические признаки.....	24
3.2. Топологические признаки.....	29
3.3. Вероятностные признаки.....	30
3.4. Спектральные признаки.....	32
4. Формирование признаков по изображению	38
4.1. Основные требования к признакам, вычисляемым по изображениям.....	39
4.2. Нормализация изображений при вычислении признаков.....	41
4.3. Моментные инварианты как признаки изображения	66
5. Обнаружение и локализация объектов на изображении.....	80
5.1. Постановка задачи и анализ современного состояния.....	80
5.2. Критерии локализации объектов.....	83
5.3. Совместное обнаружение и локализация объектов.....	93
5.4. Распознавание объектов двух классов.....	105
6. Методы совместной классификация	119
6.1. Стратегии совместной классификации.....	120
6.2. Параллельная схема совместной классификации с минимальной информацией о решениях экспертов.....	124
6.3. Двухэтапная последовательная процедура классификации	134
7. Задачи, упражнения, контрольные вопросы.....	154
7.1. Задачи.....	154
7.2. Упражнения.....	158
7.3. Контрольные вопросы.....	163
Список литературы.....	165

1. ПОСТАНОВКА ЗАДАЧИ РАСПОЗНАВАНИЯ ОБРАЗОВ

1.1. Задача распознавания образов

Существует достаточно широкий круг задач, в которых изображения рассматриваются как источник информации, на основе которой необходимо вынести некоторое решение. Например, такого рода задачи возникают в медицинской диагностике, где изображение того или иного человеческого органа анализируется с целью определения возможного заболевания. В криминалистике для установления личности человека сравнивают изображения отпечатков пальцев. Современные приборы регистрации (фотоаппараты, видеокамеры и др.) анализируют получаемые изображения для автоматического определения оптимальных параметров съемки (дальности, экспозиции и др.). Этот список, несомненно, может быть продолжен. Все задачи такого рода относятся к задачам распознавания образов.

Распознавание образов - это научное направление, которое занимается разработкой теории, методов, алгоритмов и устройств обработки информации с целью классификации поступающих сигналов (в частности, изображений).

Заметим, что с задачей распознавания, в более широком ее понимании, сталкивается каждый человек. Читаем ли мы текст, переходим ли улицу, смотрим ли телевизор, слушаем ли музыку - в каждый момент времени мы решаем задачу распознавания. Для человека это стало настолько естественно, что не вызывает никаких затруднений. В то же время, автоматизировать этот процесс, научить техническое устройство или компьютер делать то, что с такой легкостью делает человек, наталкивается на ряд нетривиальных проблем. Обозначим их на примере.

Представим ситуацию, в которой необходимо обучиться узнавать некоторый наперед заданный объект (например, человека) в ситуации, если этот объект не наблюдается визуально. Подобная ситуация часто возникает в криминалистике и решается путем составления словесного описания (портрета) искомого объекта: фиксируется рост человека,

его комплекция, цвет волос и глаз и т.д. Сопоставляя полученное описание с параметрами человека, которого мы впоследствии видим, становится возможным утверждать - похож ли этот человек на составленный портрет или нет.

Приведенный пример позволяет заметить принятое в теории распознавания образов разделение процесса распознавания на два этапа. На *первом этапе* по некоторым априорным сведениям об объекте производится формирование определенного набора свойств или, в терминах распознавания образов, *признаков*. Сформированный набор признаков объекта, или его *образ*, на *втором этапе* используется для вынесения решения о *классе* объекта (в данном примере присутствует всего два класса: люди, похожие на составленное описание человека и не похожие). Таким образом, собственно *распознавание* или *классификация* осуществляется в некотором пространстве, называемом в теории распознавания образов *пространством признаков*. Процесс классификации может производиться по следующему правилу: если набор признаков классифицируемого объекта близок к набору признаков другого - значит и сами объекты похожи и, соответственно, относятся к одному классу. Если известен класс для одного из объектов, который выступает в качестве эталона, то по указанному правилу легко определить, какие из объектов относятся к тому же классу. Геометрическая интерпретация подобного правила приведена рис. 1.1.

Граница, которая разделяет образы объектов, относящиеся к разным классам, называется в распознавании образов *разделяющей границей*, а правило, производящее это разделение - *классификатором*.

Данный пример, а также приведенный комментарий, позволяют достаточно четко определить те основные группы проблем, которые возникают при решении задачи распознавания. Первая, очевидно, связана с *выбором свойств или признаков*, которые будут использоваться для распознавания. Вторая - с *построением* собственно *классификатора*, который бы оптимальным образом производил разделение образов на классы. Эти проблемы, несомненно, возникают не только в задачах распознавания изображений, а носят более глобальный характер.

В то же время в каждой прикладной области существуют свои специфические методы их решения.

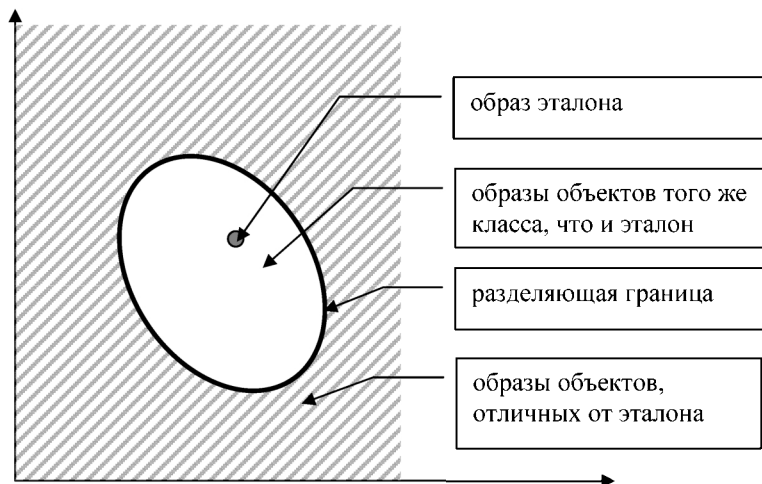


Рис.1.1. Пример классификатора объектов

К настоящему моменту разработан целый ряд процедур, ориентированных на решение задачи распознавания. Однако наиболее наглядной в рамках статистического подхода является задача построения оптимального классификатора, которая рассмотрена ниже.

1.2. Вероятностный критерий качества классификации

В идеале классификатор должен быть таким, чтобы области, выделяемые в пространстве признаков, соответствовали классам, то есть должно выполняться следующее условие: объект ω принадлежит классу Ω_l тогда и только тогда, когда соответствующий объекту вектор признаков $\mathbf{y}(\omega)$ принадлежит области D_l :

$$\forall \omega \in \Omega: \omega \in \Omega_l \Leftrightarrow \mathbf{y}(\omega) \in D_l, \quad l = \overline{0, L-1}. \quad (1.1)$$

Как правило, на практике данное условие выполняется не всегда и существует вероятность неверно классифицировать объект и допустить ошибку при распознавании.

Обозначим p_{lj} ($l, j = \overline{0, L-1}$) вероятностью того, что классификатор принимает решение об отнесении вектора признаков некоторого объекта к области D_j , в то время как сам объект принадлежит классу Ω_l :

$$p_{lj} = P(\mathbf{Y} \in D_j / \Omega_l). \quad (1.2)$$

При $l \neq j$ вероятности p_{lj} характеризуют ошибки распознавания и называются *вероятностями неверной или ошибочной классификации*, а при $l = j$ вероятности p_{ll} задают вероятности *верной (правильной) классификации* представителей соответствующего класса. Уменьшение вероятностей ошибочной классификации – это основная задача, которая возникает при построении классификатора.

Качество классификатора характеризуется величиной, называемой в теории статистических решений *условным средним риском*. Она задает среднюю величину потерь, связанных с принятием классификатором решения об отнесении данного вектора признаков \mathbf{y} к классу с номером j :

$$R_j(\mathbf{y}) = \frac{1}{p(\mathbf{y})} \sum_{l=0}^{L-1} c_{lj} P(\Omega_l) p(\mathbf{y} / \Omega_l). \quad (1.3)$$

В данном выражении:

- $P(\Omega_l)$ - *априорная вероятность* появления объектов из класса Ω_l , причем $\sum_{l=0}^{L-1} P(\Omega_l) = 1$;
- $p(\mathbf{y} / \Omega_l)$ - *условная плотность вероятностей* случайного вектора признаков \mathbf{Y} для объектов класса Ω_l (в теории распознавания образов ее называют *функцией правдоподобия* для соответствующего класса);
- $p(\mathbf{y})$ - *безусловная плотность вероятностей* случайного вектора \mathbf{Y} ;
- элементы квадратной матрицы

$$C = \left\| c_{lj} \right\|_{l,j=0}^{L-1} \quad (1.4)$$

характеризуют величины *штрафов* или *потерь* за ошибки классификатора. Матрица C может быть достаточно произвольной. Единственным ограничением на ее элементы является то, что штраф за ошибочное решение должен быть больше, чем штраф за решение правильное, то есть: $c_{lj} > c_{ll}$.

Интегральной величиной, характеризующей качество классификатора является *математическое ожидание потерь* или *общий риск*, который с учетом (1.2) и (1.3) имеет вид

$$R = \sum_{j=0}^{L-1} \int_{D_j} R_j(y) p(y) dy = \sum_{j=0}^{L-1} \sum_{l=0}^{L-1} c_{lj} P(\Omega_l) p_{lj}. \quad (1.5)$$

1.3. Задачи распознавания на изображениях

Весь спектр задач, которые приходится решать при распознавании на изображениях, можно подразделить на две группы:

- распознавание или классификация изображений;
- поиск и распознавание объектов (специфических локальных областей) на изображениях.

Это разделение связано с особенностями реализации процесса распознавания. В *первой группе* задач распознавание или классификация производится для всего изображения целиком. То есть, следуя принятой в п.1.1 терминологии, всё изображение целиком в процессе распознавания относят к одному из нескольких классов. Таким образом, решением задачи распознавания в этой группе является реализация отображения: *изображение – номер класса*. Следуя принятому в п.1.1 разбиению процесса распознавания на два этапа, указанное отображение реализуется в виде следующих двух отображений: отображения *изображение - признаки* и отображения *признаки - класс*. Это

позволяет представить процесс решения задачи классификации изображения в виде схемы, приведенной на рис.1.2, которая является традиционной и стандартной для задач распознавания образов.

Примером задач *первой группы* являются задачи распознавания лиц по фотографиям, распознавания дактилоскопических отпечатков, диагностики заболеваний по снимку того или иного человеческого органа и т.д.



Рис. 1.2. Схема решения задачи распознавания образов

В задачах *второй группы* процесс распознавания оказывается включенным в более общую технологию обработки изображения, связанную с поиском распознаваемых геометрических объектов на всей области наблюдения. *Объекты* в данной ситуации представляют собой относительно небольшие локальные области, появление которых может произойти в любой точке изображения. Причем информация о том: имеются ли объекты на изображении, каково их количество, ориентация, размеры и т.д., чаще всего отсутствует.

Результатом решения задачи распознавания в этой ситуации является не только класс найденного объекта, но также и его характеристики: положение, возможно размер, цвет, ориентация объекта в плоскости изображения и т.д.

Примером задач второй группы являются задачи дешифрирования аэрокосмических снимков, автоматического чтения текстов, нахождения локальных патологий на медицинских снимках и многие другие. Неопределенность в целом ряде характеристик объектов делает задачу их поиска и распознавания на изображении в математическом и вычислительном плане более сложной по сравнению с задачами первой группы. Это приводит к тому, что процесс ее решения не укладывается в приведенную схему, а производится в соответствии со схемой, в упрощенном виде

представленной на рис. 1.3, то есть включает в себя трудно формализуемую задачу выделения фрагментов («областей интереса»).

В соответствии с данной схемой анализу подвергается каждый фрагмент на изображении. По текущему фрагменту, выделенному *окном обработки*, производится формирование признаков и классификация. В зависимости от результатов классификации происходит расчет дополнительных параметров объекта.

Легко заметить, что схема решения задачи классификации изображения входит как составной элемент в схему решения задачи поиска и распознавания. Действительно, в более широком понимании задачи второй группы относятся к группе задач высокого уровня - задачам *анализа* наблюдаемого *изображения* или *сцены*. Под *анализом сцены* в данном случае понимается составление полного описания изображенных на снимке предметов с указанием их местоположения и взаимного расположения. В то же время следует заметить, что нередко наблюдается и противоположная ситуация, когда методы и технологии решения задач второй группы выступают в качестве составных элементов решения задач первой группы. Так, например, один из методов решения задачи распознавания людей по фотографиям их лиц заключается в нахождении на изображении ярко выраженных областей интереса: глаз, носа, губ и т.д., и их последующего описания [81, 65].

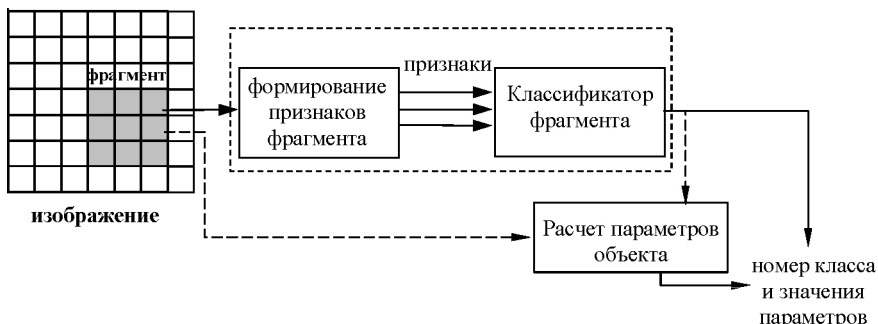


Рис. 1.3. Схема поиска и распознавания объектов на изображении

2. ОПТИМАЛЬНЫЕ СТРАТЕГИИ РАСПОЗНАВАНИЯ

2.1. Оптимальные стратегии

Процесс классификации аналогичен игре двух лиц, в которой выигрыш (проигрыш) одного из участников равен проигрышу (выигрышу) другого. Выбор оптимальной стратегии в игре зависит от количества исходной информации. Могут использоваться *байесовская*, *минимаксная* стратегии или *стратегия Неймана-Пирсона*. В зависимости от того, какая из стратегий используется для построения классификатора, последний называют, соответственно, *байесовским классификатором*, *минимаксным классификатором* или *классификатором Неймана-Пирсона*.

2.1.1 Байесовский классификатор

Байесовская стратегия используется при наличии полной априорной информации о классах, то есть когда известны:

- функции правдоподобия для каждого из классов;
- матрица штрафов;
- априорные вероятности для каждого из классов.

Стратегия решения выбирается таким образом, чтобы обеспечить минимум общего риска (1.5). Минимальный общий риск при этом называется *байесовским риском*. В соответствии с выражениями (1.3) и (1.5), минимум общего риска R будет обеспечен, если разбиение пространства признаков D будет осуществляться по следующему правилу: вектор $y \in D$ относится к области D_l только тогда, когда соответствующий условный средний риск $R_l(y)$ минимален:

$$\forall j \neq l \quad R_l(y) < R_j(y) \Rightarrow y \in D_l \quad (2.1)$$

Графическая иллюстрация байесовской стратегии приведена на рис. 2.1а.

Если матрица потерь (1.4) является простейшей, то есть, если ее элементы удовлетворяют равенству

$$c_{ij} = \begin{cases} 0, & i = j, \\ 1, & i \neq j, \end{cases}$$

то после подстановки в (2.1) выражения для условного среднего риска (1.3) имеем следующий явный вид байесовского классификатора (см. рис. 2.1а, б):

$$\forall j \neq l \ P(\Omega_l) p(\mathbf{y}/\Omega_l) \geq P(\Omega_j) p(\mathbf{y}/\Omega_j) \Rightarrow \mathbf{y} \in D_l. \quad (2.2)$$

Из (2.2), в частности, видно, что решающими функциями байесовского классификатора являются функции:

$$d_l(\mathbf{y}) = P(\Omega_l) p(\mathbf{y}/\Omega_l), \quad l = \overline{0, L-1}. \quad (2.3)$$

Часто используют также следующую форму записи байесовского классификатора:

$$\forall j \neq l \ \frac{p(\mathbf{y}/\Omega_l)}{p(\mathbf{y}/\Omega_j)} \geq \frac{P(\Omega_j)}{P(\Omega_l)} \Rightarrow \mathbf{y} \in D_l. \quad (2.4)$$

При этом функция $\Lambda_{lj}(\mathbf{y}) = \frac{p(\mathbf{y}/\Omega_l)}{p(\mathbf{y}/\Omega_j)}$ называется *отношением правдоподобия*, а величина $\lambda_{jl} = \frac{P(\Omega_j)}{P(\Omega_l)}$ - *пороговым значением*. Таким образом, байесовский классификатор основан на сравнении отношения правдоподобия с пороговым значением:

$$\forall j \neq l \ \Lambda_{lj}(\mathbf{y}) \geq \lambda_{jl} \Rightarrow \mathbf{y} \in D_j,$$

и называется поэтому *классификатором отношения правдоподобия*.

Легко показать, что при произвольном виде матрицы штрафов в случае двух классов байесовский классификатор имеет вид

$$\frac{p(\mathbf{y}/\Omega_1) > P(\Omega_0)(c_{01} - c_{00})}{p(\mathbf{y}/\Omega_0) < P(\Omega_1)(c_{10} - c_{11})} \Rightarrow \mathbf{y} \in \begin{cases} D_1 \\ D_0 \end{cases}$$

с дискриминантными функциями:

$$d_j(\mathbf{y}) = P(\Omega_j)(c_{j(1-j)} - c_{jj})p(\mathbf{y}/\Omega_j), \quad j = \overline{0,1}.$$

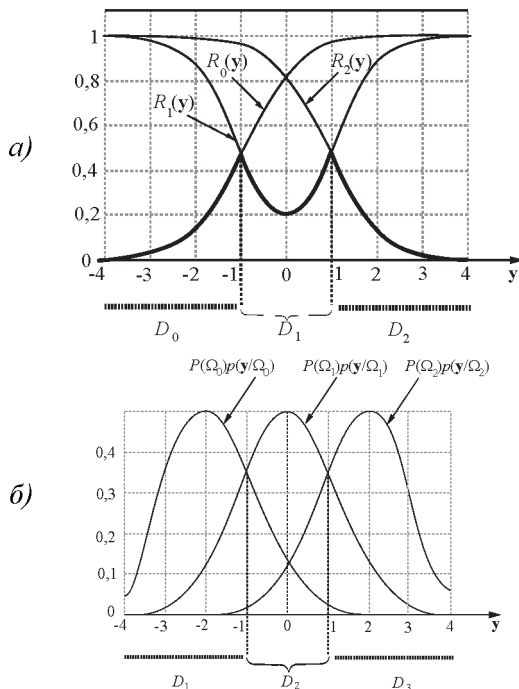


Рис. 2.1. Построение байесовского классификатора для простейшей матрицы штрафов:
 а) байесовская стратегия минимизации общего риска;
 б) байесовский классификатор

2.1.2. Минимаксный классификатор

Классификатор, основанный на минимаксной стратегии, используется для случая двух классов и если известны:

- функции правдоподобия для каждого из классов;
- матрица штрафов.

Минимизировать величину общего риска при отсутствии информации об априорных вероятностях классов, очевидно, невозможно. В то же время, предполагая возможность произвольного изменения значений априорных вероятностей классов, можно минимизировать максимально возможное значение риска. Действительно, общий риск (1.5) в случае двух классов может быть представлен в следующем виде:

$$R = (c_{11} + p_{10}(c_{10} - c_{11})) + P(\Omega_0) \cdot [(c_{00} + p_{01}(c_{01} - c_{00})) - (c_{11} + p_{10}(c_{10} - c_{11}))]. \quad (2.5)$$

При фиксированном классификаторе изменение априорной вероятности приводит к изменению величины общего риска, причем характер зависимости в (2.5) линейный (рис. 2.2). Поэтому поиск классификатора, минимизирующего максимально возможную величину общего риска, эквивалентен поиску такого байесовского классификатора, для которого величина (2.5) является постоянной, не зависящей от значения априорной вероятности $P(\Omega_0)$ величиной. Таким классификатором, очевидно, является байесовский классификатор, удовлетворяющий следующему дополнительному условию:

$$(c_{00} + p_{01}(c_{01} - c_{00})) - (c_{11} + p_{10}(c_{10} - c_{11})) = 0. \quad (2.6)$$

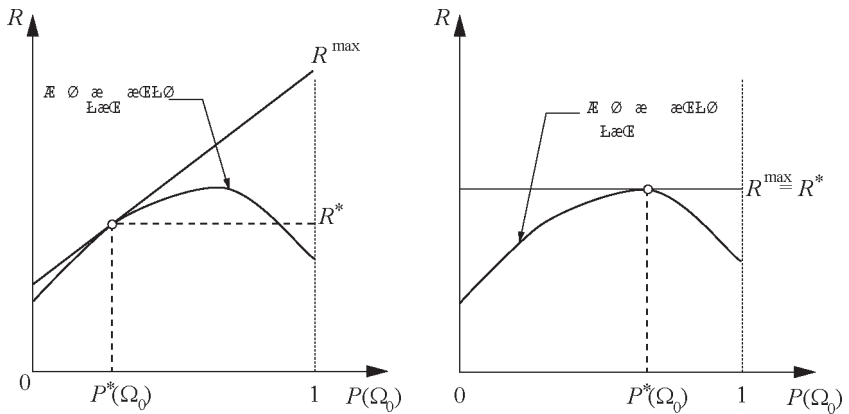


Рис. 2.2. Иллюстрация минимаксной стратегии построения классификатора

Из рисунка 2.2 видно, что значение величины общего риска для минимаксного классификатора равно максимальному значению байесовского (минимального) риска. Пара априорных вероятностей $(P^*(\Omega_0), 1 - P^*(\Omega_0))$, при которых байесовский риск принимает максимальное значение, называется *наименее благоприятным распределением априорных вероятностей*. Таким образом, *минимаксный классификатор* – это байесовский классификатор, полученный для пары наименее благоприятных априорных вероятностей.

В более простой ситуации, когда элементы матрицы штрафов таковы, что

$$c_{00} = c_{11} = 0, \quad c_{10} = c_1, \quad c_{01} = c_0,$$

условие (2.3) преобразуется в следующее:

$$p_{01}c_0 = p_{10}c_1. \quad (2.7)$$

Последнее выражение представляет собой условие выбора областей D_0 , D_1 в байесовском классификаторе.

2.1.3. Классификатор Неймана-Пирсона

Критерий Неймана-Пирсона в теории статистических решений используется для проверки гипотез. В классической постановке задачи используется только две возможные гипотезы и различают два типа ошибок:

ошибку первого рода p_0 - в контексте настоящего изложения $p_0 = p_{01}$,
ошибку второго рода p_1 - в контексте настоящего изложения $p_1 = p_{10}$.

Заметим, что в общем случае $p_1 + p_0 \neq 1$.

Классификатор, основанный на стратегии Неймана-Пирсона, используется для случая двух классов, и если известны только функции правдоподобия для каждого из классов. Суть стратегии Неймана-Пирсона

состоит в следующем: задается допустимое значение вероятности ошибки первого рода p_0 , а затем классификатор строится таким образом, чтобы обеспечить минимум вероятности ошибки второго рода p_1 :

$$\begin{cases} p_1 \rightarrow \min_{D_0, D_1}, \\ p_0 = p_0^*. \end{cases} \quad (2.8)$$

Решением задачи Неймана-Пирсона является классификатор вида:

$$\Lambda(\mathbf{y}) = \frac{p(\mathbf{y}/\Omega_1)}{p(\mathbf{y}/\Omega_0)} > \lambda \Rightarrow \begin{cases} \mathbf{y} \in D_1, \\ \mathbf{y} \in D_0, \end{cases} \quad (2.9)$$

где значение пороговой величины λ определяется, исходя из условия: $p_0 = p_0^*$ (рис.2.2). Из выражения (2.9) следует, что *классификатор Неймана-Пирсона – это классификатор отношения правдоподобия.*

2.2. Байесовский классификатор для нормально распределенных векторов признаков

2.2.1. Алгоритмы классификации

На практике часто возникает задача распознавания детерминированных объектов в условиях случайных помех. Ниже приведен достаточно типичный пример постановки подобной задачи и ее решения с использованием байесовской стратегии.

Пусть входной сигнал, задаваемый вектором $\mathbf{y} = (y_0, \dots, y_{K-1})^T$ и подлежащий распознаванию, представляет собой аддитивную смесь детерминированной и шумовой составляющих. Будем считать, что наблюдаемые вектора имеют нормальный закон распределения в каждом из L классов, то есть имеют плотность вероятностей вида

$$p(\mathbf{y}/\Omega_l) = \frac{1}{(2\pi)^{K/2} \sqrt{|\mathbf{B}_l|}} \exp\left(-\frac{1}{2}(\mathbf{y} - \mathbf{M}_l)^T \mathbf{B}_l^{-1} (\mathbf{y} - \mathbf{M}_l)\right), \quad (2.10)$$

$$l = \overline{0, L-1}.$$

Здесь

$$\mathbf{M}_l = E\{\mathbf{Y}/\Omega_l\}, \quad B_l = E\{(\mathbf{Y} - \mathbf{M}_l)(\mathbf{Y} - \mathbf{M}_l)^T / \Omega_l\}$$

математическое ожидание и ковариационная матрица вектора признаков из класса Ω_l , соответственно. Математические ожидания или средние характеризуют детерминированные составляющие распознаваемых сигналов, а ковариационные матрицы – характер шумовой составляющей. Считаются также известными априорные вероятности $P(\Omega_l)$ появления векторов из каждого класса. Требуется по реализации y случайного вектора \mathbf{Y} определить класс, к которому данный вектор принадлежит.

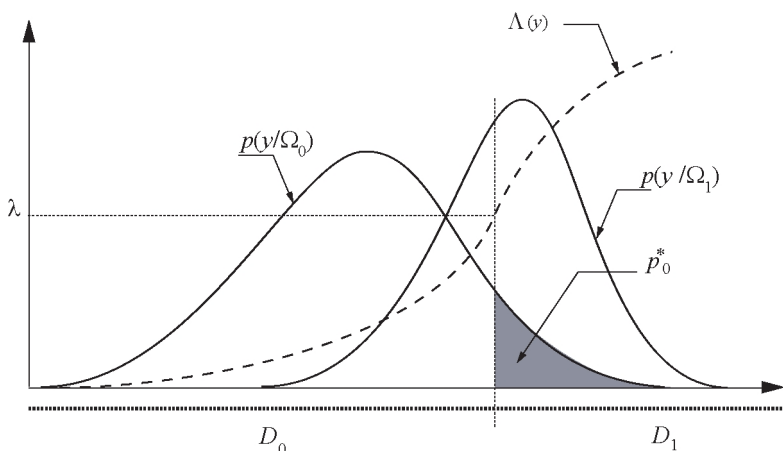


Рис. 2.3. Иллюстрация стратегии Неймана-Пирсона построения классификатора

Решением данной задачи является байесовский классификатор с дискриминантными функциями следующего вида:

$$d_l(y) = \ln P(\Omega_l) - \ln \sqrt{|B_l|} - \frac{1}{2}(y - \mathbf{M}_l)^T B_l^{-1} (y - \mathbf{M}_l), \quad l = \overline{0, L-1}. \quad (2.11)$$

Выражение (2.11) может быть существенно упрощено в некоторых частных случаях.

2.2.1.1. Случай 1

Предположим, что компоненты наблюдаемого вектора \mathbf{Y} являются независимыми и имеют одинаковую дисперсию σ_Y^2 , то есть $\mathbf{V}_l = \sigma_Y^2 \mathbf{I}$, где \mathbf{I} – единичная $K \times K$ матрица. Тогда законы распределения (2.10) отличаются только средними значениями, а решающие функции байесовского классификатора преобразуются к следующему виду:

$$d_l(\mathbf{y}) = 2\sigma_Y^2 \ln P(\Omega_l) - \|\mathbf{y} - \mu_l\|^2 \quad l = \overline{0, L-1}, \quad (2.12)$$

здесь $\|\dots\|$ – евклидова норма. При равных априорных вероятностях данное решающее правило приобретает очевидную трактовку: *вектор признаков \bar{x} относится к тому классу, расстояние до центра которого минимально.*

Классификатор в этом случае называют *классификатором по минимуму евклидова расстояния*. Пример разбиения пространства признаков при использовании подобного классификатора для случая трех классов приведен на рис. 2.4а.

Нетрудно видеть, что решающие функции (2.12) можно преобразовать к линейной форме:

$$d_l(\mathbf{y}) = \mathbf{M}_l^T \mathbf{y} - \frac{1}{2} \mathbf{M}_l^T \mathbf{M}_l + \sigma_Y^2 \ln P(\Omega_l), \quad l = \overline{0, L-1}.$$

В этом случае *разделяющие границы* между различными областями D_l , задаваемые соотношениями вида:

$$d_{ij}(\mathbf{y}) \equiv d_l(\mathbf{y}) - d_j(\mathbf{y}) = 0, \quad 0 \leq l < j \leq L-1,$$

также являются линейными:

$$d_{ij}(\mathbf{y}) = (\mathbf{M}_l - \mathbf{M}_j)^T \mathbf{y} - \frac{1}{2} (\mathbf{M}_l + \mathbf{M}_j)^T (\mathbf{M}_l - \mathbf{M}_j) + \sigma_Y^2 \ln \frac{P(\Omega_l)}{P(\Omega_j)},$$

$$0 \leq l < j \leq L-1,$$

и говорят о линейном классификаторе.

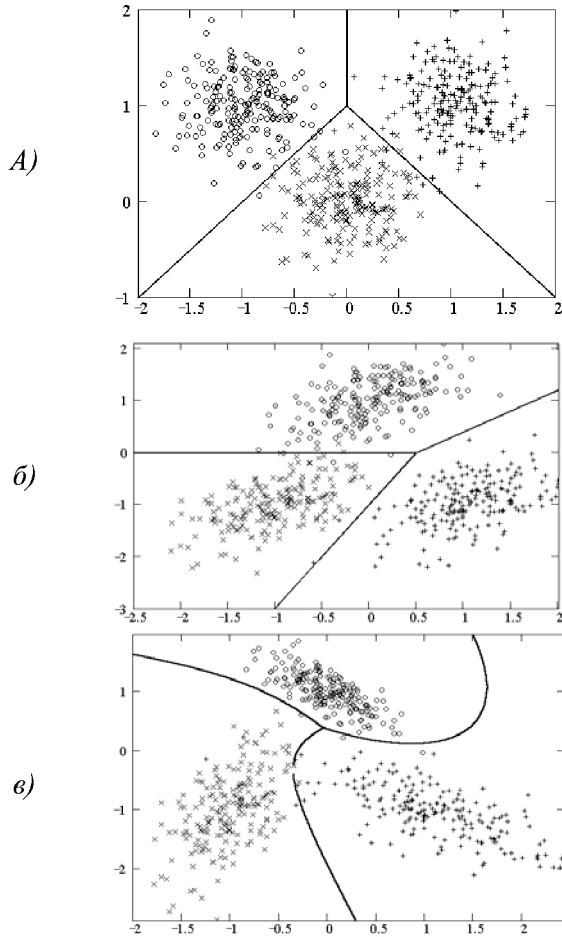


Рис. 2.4. Байесовский классификатор в случае нормально распределенных векторов признаков: а) признаки статистически независимых и одинаково распределены; б) корреляционные матрицы одинаковы; в) корреляционные матрицы различны

2.2.1.2. Случай 2

Предположим, что все корреляционные матрицы одинаковы: $\mathbf{V}_l = \mathbf{V}$. Тогда решающие функции байесовского классификатора представляются в виде

$$d_l(\mathbf{y}) = 2 \ln P(\Omega_l) - (\mathbf{y} - \mathbf{M}_l)^T \mathbf{V}^{-1} (\mathbf{y} - \mathbf{M}_l), \quad l = \overline{0, L-1}.$$

Величина

$$\rho(\mathbf{y}, \mathbf{M}_l) = (\mathbf{y} - \mathbf{M}_l)^T \mathbf{B}^{-1} (\mathbf{y} - \mathbf{M}_l) \quad (2.13)$$

называется *расстоянием Махаланобиса* между векторами \mathbf{y} и \mathbf{M}_l и является мерой близости вектора \mathbf{y} к центру класса Ω_l , учитывающей как дисперсии компонент вектора \mathbf{Y} , так и их взаимную корреляцию. Очевидно, что в данной ситуации классификатор снова оказывается классификатором по минимуму расстояния Махаланобиса (см. рис. 2.4б). Кроме того, и решающие функции, и разделяющие границы снова являются линейными:

$$d_l(\mathbf{y}) = \mathbf{M}_l^T \mathbf{B}^{-1} \mathbf{y} - \frac{1}{2} \mathbf{M}_l^T \mathbf{B}^{-1} \mathbf{M}_l + \ln P(\Omega_l), \quad l = \overline{0, L-1}, \quad (2.14)$$

$$d_{ij}(\mathbf{y}) = (\mathbf{M}_i - \mathbf{M}_j)^T \mathbf{B}^{-1} \mathbf{y} - \frac{1}{2} (\mathbf{M}_i + \mathbf{M}_j)^T \mathbf{B}^{-1} (\mathbf{M}_i - \mathbf{M}_j) + \ln \frac{P(\Omega_i)}{P(\Omega_j)},$$

$$0 \leq l < j \leq L-1.$$

А, следовательно, линейным является и классификатор.

2.2.1.3. Случай 3

В ситуации, когда все корреляционные матрицы различны, необходимо пользоваться выражением (2.11) для дискриминантных функций. Разделяющие границы в этом случае представляются в следующем виде:

$$d_{ij}(\mathbf{y}) = \mathbf{y}^T (\mathbf{B}_j^{-1} - \mathbf{B}_i^{-1}) \mathbf{y} + 2(\mathbf{M}_i^T \mathbf{B}_i^{-1} - \mathbf{M}_j^T \mathbf{B}_j^{-1}) \mathbf{y} + \left[\ln \frac{|\mathbf{B}_i|}{|\mathbf{B}_j|} + 2 \ln \frac{P(\Omega_i)}{P(\Omega_j)} - \mathbf{M}_i^T \mathbf{B}_i^{-1} \mathbf{M}_i + \mathbf{M}_j^T \mathbf{B}_j^{-1} \mathbf{M}_j \right],$$

$$0 \leq l < j \leq L-1 \quad (2.15)$$

и являются, очевидно, квадратичными функциями. Такие границы называются *гиперквадриками* (гиперсферы, гиперпараболы и т.д., пример их приведен на рис. 2.4в), а сам классификатор называется *квадратичным*.

2.2.2. Вычисление вероятностей ошибочной классификации для нормально распределенных векторов признаков

Эффективность любого классификатора характеризуется вероятностями ошибок. Однако их нахождение в общем случае оказывается достаточно сложной задачей, поскольку требует вычисления многомерных интегралов:

$$p_{ij} = \int_{D_j} p(\mathbf{y}/\Omega_l) d\mathbf{y}, \quad l \neq j, \quad l, j = \overline{0, L-1}. \quad (2.16)$$

При использовании байесовского классификатора, который является классификатором отношения правдоподобия, многомерный интеграл (2.16) может быть заменен одномерным от плотности вероятностей отношения правдоподобия Λ_{ij} в каждом из классов. В частности, в случае двух классов для вероятностей ошибок имеем следующие выражения:

$$p_0 = \int_{\lambda}^{+\infty} p_{\Lambda}(u/\Omega_0) du, \quad p_1 = \int_{-\infty}^{\lambda} p_{\Lambda}(u/\Omega_1) du, \quad (2.17)$$

где $\Lambda = \Lambda(\mathbf{Y}) = \frac{p(\mathbf{Y}/\Omega_1)}{p(\mathbf{Y}/\Omega_0)}$, а $\lambda = \frac{P(\Omega_0)}{P(\Omega_1)}$ - пороговое значение. Плотность вероятностей отношения правдоподобия удастся найти далеко не всегда. Однако, когда случайный вектор \mathbf{Y} имеет нормальный закон распределения, это может быть сделано.

Пусть вектор признаков в каждом из двух классов характеризуется нормальным законом распределения, причем все корреляционные матрицы являются равными $\mathbf{V}_l = \mathbf{V}$ ($l=0,1$). Тогда случайная величина $\tilde{\Lambda} = \ln(\Lambda(\mathbf{Y}))$ имеет нормальный закон распределения с параметрами:

$$E\{\mathcal{X}/\Omega_0\} = E\{\ln(\Lambda(\mathbf{Y}))/\Omega_0\} = -\frac{1}{2}\rho(\mathbf{M}_0, \mathbf{M}_1),$$

$$E\{\mathcal{X}/\Omega_1\} = E\{\ln(\Lambda(\mathbf{Y}))/\Omega_1\} = \frac{1}{2}\rho(\mathbf{M}_0, \mathbf{M}_1),$$

$$E\left\{\left(\mathcal{X}/\Omega_l - E\{\mathcal{X}/\Omega_l\}\right)^2\right\} = \rho(\mathbf{M}_0, \mathbf{M}_1), \quad l = \overline{0,1},$$

где $\rho(\mathbf{M}_0, \mathbf{M}_1)$ - расстояние Махаланобиса между векторами средних \mathbf{M}_0 и \mathbf{M}_1 . Таким образом законы распределения $\tilde{\Lambda}$ в каждом из классов отличаются лишь математическими ожиданиями, причем $E\{\tilde{\Lambda}/\Omega_1\} = -E\{\tilde{\Lambda}/\Omega_0\}$. В этом случае плотности вероятностей искомой случайной величины выглядит так, как показано на рис. 2.5.

Выражения для вероятностей ошибок (2.17) преобразуются к следующему виду:

$$p_0 = 1 - \Phi\left(\frac{\mathcal{X}^{\text{о}} + \frac{1}{2}\rho(\mathbf{M}_0, \mathbf{M}_1)}{\sqrt{\rho(\mathbf{M}_0, \mathbf{M}_1)}}\right), \quad p_1 = \Phi\left(\frac{\mathcal{X}^{\text{о}} - \frac{1}{2}\rho(\mathbf{M}_0, \mathbf{M}_1)}{\sqrt{\rho(\mathbf{M}_0, \mathbf{M}_1)}}\right), \quad (2.18)$$

где $\Phi(\dots)$ - функция Лапласа, а

$$\mathcal{X}^{\text{о}} = \ln \lambda = \ln\left(\frac{P(\Omega_0)(c_{01} - c_{00})}{P(\Omega_1)(c_{10} - c_{11})}\right)$$

- новая пороговая величина (рис.2.5). В частном случае, когда матрица штрафов является простейшей и априорные вероятности классов совпадают, имеем:

$$\lambda = 1, \quad \mathcal{X}^{\text{о}} = 0, \quad p_0 = 1 - \Phi\left(\frac{1}{2}\sqrt{\rho(\mathbf{M}_0, \mathbf{M}_1)}\right), \quad p_1 = \Phi\left(-\frac{1}{2}\sqrt{\rho(\mathbf{M}_0, \mathbf{M}_1)}\right).$$

Общий риск при этом определяется формулой:

$$R = 1 - \Phi\left(\frac{1}{2}\sqrt{\rho(\mathbf{M}_0, \mathbf{M}_1)}\right) \quad (2.19)$$

и монотонно убывает с ростом расстояния Махаланобиса между векторами средних.

Минимаксный классификатор

Предположим, что матрица штрафов имеет следующие элементы: $c_{01} = c_{10} = c$, $c_{00} = c_{11} = 0$. Тогда соотношение (2.7) для выбора разделяющей границы байесовского классификатора, соответствующего минимаксной стратегии, превращается в равенство вероятностей ошибочной классификации: $p_0 = p_1$. С учетом выражений (2.18) получаем, что пороговое значение для минимаксного классификатора: $\lambda = 1$, $\tilde{\lambda} = 0$.

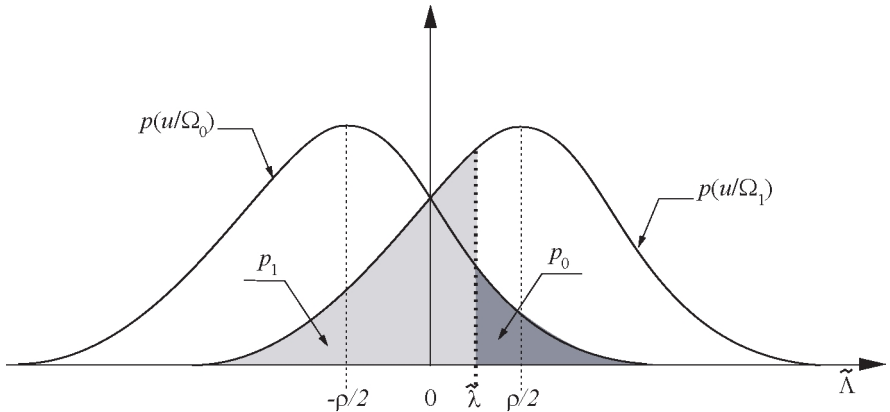


Рис. 2.5. Плотности вероятностей логарифма отношения правдоподобия для нормально распределенных векторов признаков с равными корреляционными матрицами

Классификатор Неймана-Пирсона

Используя условие (2.8) и равенства (2.18), получаем, что пороговое значение классификатора Неймана-Пирсона определяется по формуле:

$$\lambda = e^{\%_0}, \quad \%_0 = -\frac{1}{2} \rho(\mathbf{M}_0, \mathbf{M}_1) + \sqrt{\rho(\mathbf{M}_0, \mathbf{M}_1)} \Phi^{-1}(1 - p_0^*), \quad (2.20)$$

где p_0^* - заданная величина вероятности ошибки первого рода.

3. ОСНОВНЫЕ ГРУППЫ ПРИЗНАКОВ, ИСПОЛЬЗУЕМЫХ ПРИ РАСПОЗНАВАНИИ ИЗОБРАЖЕНИЙ

Построение описания изображения на основе его представления с использованием признаков - едва ли не самая сложная задача в процессе построения любой системы распознавания. При этом, если в рамках некоторых математических моделей удалось формализовать процесс классификации, то процесс выбора признаков до сих пор остался процедурой эвристической и зависимой как от предметной области, так и от разработчика. В то же время определенный опыт, накопленный за годы использования средств распознавания образов и обработки изображений для решения практических задач, позволяет выделить ряд основных групп признаков, которые успешно используются для описания и распознавания изображений. Принятая ниже классификация признаков на группы отражает специфику подходов, на основании которых производится их построение.

3.1. Геометрические признаки

К этой группе относятся те признаки, расчет которых основан на использовании геометрических характеристик, представленных на изображении объектов. Это могут быть, например, следующие признаки:

- геометрические размеры изображенного объекта по вертикали или горизонтали,
- расстояние между наиболее удаленными точками на изображенном объекте,
- периметр и площадь изображенного объекта,
- компактность объекта (как соотношение между его площадью и периметром),
- числовые характеристики описанных или вписанных в изображение объекта геометрических фигур, таких как окружности, многоугольники, и т.д.

Достаточно специфическими, но широко используемыми в геометрической группе являются признаки, связанные с представлением *геометрии контура объекта*. Как известно контур - одна из наиболее важных характеристик изображенного объекта при его восприятии человеком. Поэтому описание контура - одна из популярных задач, решаемых в обработке изображений. К наиболее типичным методам описания контура, используемым также и для решения задач распознавания, относятся следующие.

3.1.1 Цепные коды

Метод цепного кодирования для представления контура объекта был предложен Фриманом. Он заключается в том, чтобы границу объекта, расположенного на дискретной сетке, представить в виде набора элементарных отрезков. Тогда полной характеристикой границы в каждой точке является направление требуемого отрезка. Рисунок 3.1а иллюстрирует данный метод. В данном случае предполагается, что точки на границе являются только 4-х связными (отрезок откладывается лишь в 4-х направлениях). Иногда применяют модификацию данного метода, использующую 8-связную модель, как на рис. 3.1б.

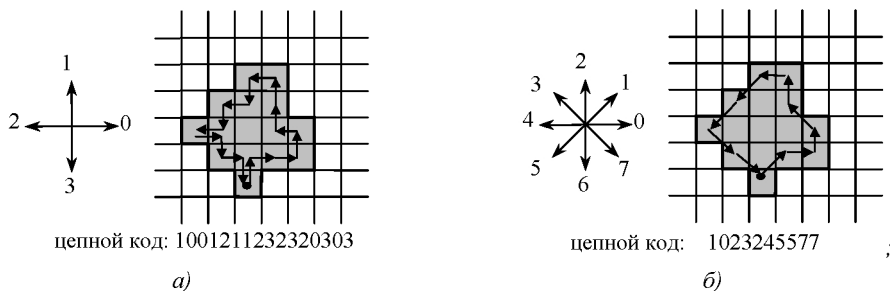


Рис. 3.1. Примеры построения цепного кода:
4-связная модель границы (а); 8-связная модель границы (б)

Несомненным достоинством представления границы изображаемого объекта цепным кодом является простота реализации алгоритма его описания, простота получения на основе этого описания некоторых других геометрических характеристик объекта (например: периметр, площадь, линейные размеры по вертикали и горизонтали), воз-

возможность достижения инвариантности описания к преобразованиям подобия – масштабированию изображения, его переносу и повороту. Основным недостатком является высокая неустойчивость получаемых описаний к искажениям в изображениях.

Приложение непрерывных моделей к задаче представления дискретного контура позволило получить ряд практически полезных описаний.

3.1.2 Ортогональные представления функции кривизны

Рассматривая угол наклона к касательной контура как периодическую функцию некоторой переменной s : $\kappa(s) = \kappa(x_1(s), x_2(s))$, можно получить представление контура в виде ряда, коэффициенты которого вычисляются:

$$c_\kappa = \frac{1}{L} \int_0^L \kappa(s) \varphi_k(s) ds, \quad (3.1)$$

где $\{\varphi_k(s)\}_{k=0}^{K-1}$ набор ортонормированных или ортогональных функций на $[0, L]$:

$$\frac{1}{P} \int_0^P \varphi_k(s) \varphi_j(s) ds = \lambda_k \delta(k - j),$$

здесь L - периметр объекта. В качестве функций $\varphi_k(s)$ чаще всего используют комплекснозначную экспоненциальную функцию:

$\varphi_k(s) = \exp\left(-i \frac{2\pi}{L} sk\right)$. В этом случае говорят о представлении контура

в виде ряда Фурье, а коэффициенты λ_k называют *Фурье-дескрипторами* данного контура.

Описание (3.1) обладает свойствами инвариантности к преобразованиям подобия. Однако для его использования необходимо получить промежуточное представление контура в виде функции $\kappa(s)$, которое

само по себе является его характеристикой. Это может быть сделано различными способами, в частности с помощью ψ -кривых.

3.1.3. ψ -кривые

Представление контура объекта с использованием ψ -кривой основано на аппроксимации прямолинейных участков границы объекта в виде отрезков ломаных, а области изменения направления границы - в виде дуг окружностей (см. рис. 3.2).

В этом случае периодическая функция $\kappa(s)$ оказывается представлена прямолинейными горизонтальными участками и короткими кривыми, соответствующими областям изменения направления границы. При надлежащем построении ψ -кривых можно также добиться инвариантности к преобразованиям подобия (например, выбирая в качестве начального отрезка самый длинный или самый короткий и т.д.).

3.1.4. Слайн - аппроксимация границы

Аппарат сплайнов достаточно широко используется для практических приложений, в частности, для решения задачи описания границы. Суть метода сплайн аппроксимации - представления границы объекта в виде кусочно-полиномиальной функции часто с выполнением ряда ограничений, накладываемых на её гладкость. Наиболее популярными являются сплайны третьего порядка, поскольку обладают наименьшей степенью, при которой полиномиальная функция может менять знак кривизны. Последнее позволяет добиться определенной гладкости в точках соединения сплайнов за счет равенства первых производных в этих точках.

При использовании наиболее простого задания $\kappa(s)$ в виде горизонтальных прямых с разрывом в точках изменения направления границы, получаем представление контура в виде полилиний - соединяющихся между собой отрезков прямых (см. рис. 3.3). Точки разбиения полилинии выбираются, исходя из соображений наилучшего представления контура.

Надо отметить, что изложенные методы и алгоритмы далеко не исчерпывают весь комплекс методов представления контура объекта для его расположения, но являются наиболее часто используемыми в практических приложениях.

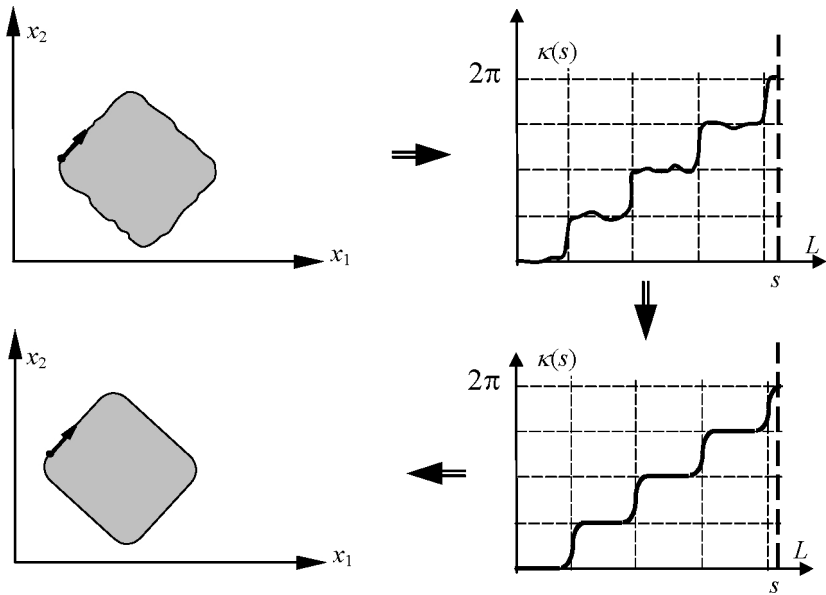


Рис. 3.2. Представление контура объекта ψ -кривой

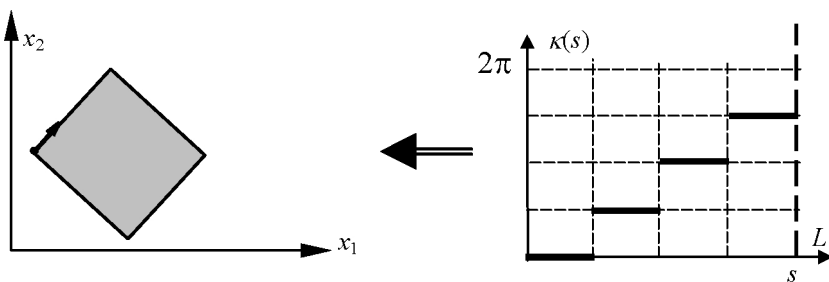


Рис. 3.3. Представление контура объекта полилинией

В заключение рассмотрения группы геометрических признаков укажем их основные свойства применительно к задаче распознавания. К ним относятся:

- возможность обеспечения инвариантности к преобразованиям подобия изображенного объекта;
- низкая помехоустойчивость.

Эти два основных свойства позволяют говорить о том, что геометрические признаки могут использоваться в задачах распознавания, которые характеризуются:

- объектами с ярко выраженной геометрической структурой (границей);
- малым уровнем шумов и динамических искажений.

3.2. Топологические признаки

К данной группе относятся те признаки, которые характеризуют топологические свойства изображенного объекта. Под топологическими свойствами понимают те свойства, которые остаются инвариантными относительно *топологических* или *гомеоморфных отображений*. Последние подразумевают под собой взаимно однозначные непрерывные (прямые и обратные) отображения. В связи с этим топологические признаки оказываются априорно инвариантными к широкому спектру геометрических преобразований: преобразованиям подобия, аффинным преобразованиям и многим другим. В то же время подобная общность признаков делает их достаточно неинформативными, так как многие обладающие различной геометрией объекты оказываются неразличимыми в рамках топологического описания. Такими, например, являются объекты, изображенные на рис. 3.4.

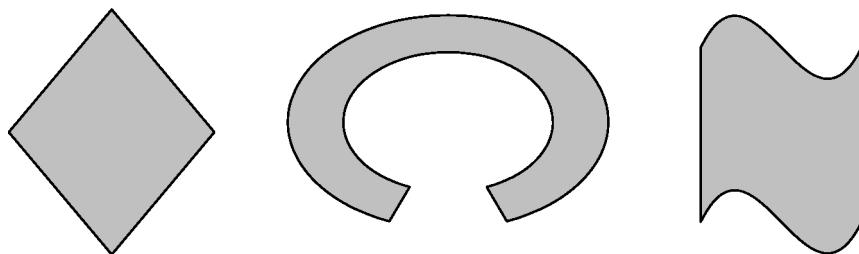


Рис. 3.4. Пример топологически неразличимых объектов

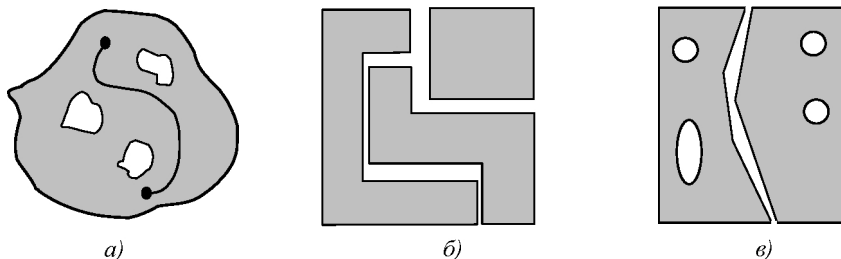


Рис. 3.5. Примеры топологически различных объектов: однокомпонентный объект (а); трехкомпонентный объект (б); двухкомпонентный объект (в)

Кроме того, топологических признаков относительно немного, и их расчет, как правило, достаточно сложен. Приведем некоторые из них.

Число связных компонентов объекта - это такое минимальное число компонент, составляющих объект, в каждом из которых любые две точки могут быть соединены линией, полностью содержащейся в том же компоненте. Примеры объектов, состоящих из различного числа компонентов, приведены на рис. 3.5.

Число «дыр» в объекте - характеризует число связных компонентов, не принадлежащих объекту, но находящихся внутри него. Так на рис. 3.5а число «дыр» равно трем.

Число Эйлера - вычисляется как разность между числом связных компонентов объекта и числом «дыр» на нем. Для объекта на рис. 3.5в число Эйлера равно (-2).

Суммируя все вышесказанное, можно утверждать о том, что топологические признаки применительно к задаче распознавания на изображениях могут быть использованы лишь в редких случаях и лишь при малом уровне искажений, связанных с исчезновением или появлением частей объектов.

3.3. Вероятностные признаки

Название этой группы признаков отражает характер модели, используемой для описания функции яркости на изображении. А именно, на функцию яркости смотрят как на реализацию (стационарного) случайного процесса или процессов (для цветных изображений). В этом

случае признаками изображения являются числовые характеристики случайного процесса. К наиболее часто используемым относятся:

- *яркостные характеристики*, такие как гистограмма распределения значений яркости на изображении, начальные μ_k и центральные η_k моменты функции яркости:

$$\mu_k = E \left\{ f^k(n) \right\}, \quad \eta_k = E \left\{ (f(n) - \alpha_1)^k \right\},$$

где $\mu_1 \equiv \mu_f$ - среднее функции яркости на изображении;

- *текстурные характеристики изображения*, к которым относятся характеристики случайного процесса, определяющие его корреляционные свойства, такие как коэффициент корреляции на изображении, корреляционную функцию изображения и др;
- *энергетические характеристики изображения*, к которым относятся отсчеты его энергетического спектра;
- *признаки стохастической геометрии*.

Данные признаки характеризуют случайные величины, связанные с наступлением каких-либо геометрических событий. Наиболее часто используемой схемой их генерации является случайное бросание прямой линии на плоскость изображения и вычисление некоторых характеристик, связанных с событием пересечения этой линией области объекта. Например, это может быть число пересечений объекта линией, максимальная или минимальная длина отрезка линии, содержащейся в объекте, суммарная длина отрезков линии, находящихся в области объекта, и т.д. (см. рис.3.6). Вероятностные характеристики таких случайных величин некоторым образом описывают форму объекта. Причем, в силу случайности положения линии, достигается инвариантность к смещениям и поворотам изображения. Более того, при надлежащем построении характеристик может быть достигнута инвариантность к масштабу. Несомненным достоинством таких признаков являются относительно высокая их помехоустойчивость и возможность достаточно простой автоматизации процесса наращивания числа таких признаков.

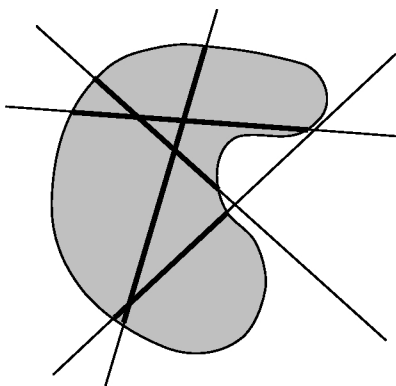


Рис. 3.6. Иллюстрация способа генерации признаков
стохастической геометрии

В то же время, несмотря на целый ряд преимуществ, признаки стохастической геометрии редко используются в распознавании изображений. Связано это, в основном, с двумя причинами:

- высокой вычислительной сложностью расчета признаков;
- отсутствием явной геометрической интерпретации получаемых значений признаков.

3.4. Спектральные признаки

К данной группе относятся те признаки, процесс получения которых использует спектральную модель преобразования изображения:

$$g(m_1, m_2) = \sum_{n_1=0}^{M_1-1} \sum_{n_2=0}^{N_2-1} f(n_1, n_2) W(n_1, n_2, m_1, m_2).$$

Характер функций $W(\cdot)$, называемых *ядрами* преобразования, может быть достаточно произвольным. Наиболее типичными являются следующие ядра и соответствующие им преобразования.

3.4.1 Разложение Карунена-Лоэва (в форме Хотеллинга)

Пусть для описания функции яркости исходного изображения используется модель дискретного случайного поля $f(n_1, n_2)$ с нулевым математическим ожиданием и известной корреляционной функцией

$$R(m_1, m_2, n_1, n_2) = M[f(m_1, m_2) \cdot f(n_1, n_2)].$$

Тогда ядро преобразования Карунена-Лоэва имеет следующий вид:

$$W(n_1, n_2, m_1, m_2) = \Psi_{m_1, m_2}(n_1, n_2),$$

где $\Psi_{m_1, m_2}(n_1, n_2)$ - собственные векторы корреляционной матрицы $R(m_1, m_2, n_1, n_2)$, вычисляемые из соотношения

$$\sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} \Psi_{m_1, m_2}(n_1, n_2) R(k_1, n_1, k_2, n_2) = \lambda_{m_1, m_2} \Psi_{m_1, m_2}(k_1, k_2),$$

λ_{m_1, m_2} - собственные значения корреляционной матрицы $R(m_1, m_2, n_1, n_2)$, равные дисперсиям соответствующих признаков.

Основным достоинством разложения Карунена-Лоэва является наличие ряда свойств оптимальности рассчитываемых признаков. А именно:

- концентрация мощности (дисперсии) в минимально возможном числе признаков,
- минимальная среднеквадратичная погрешность восстановления исходного изображения при заданном числе признаков,
- некоррелированность, а в случае нормального распределения яркости исходного изображения и независимость, рассчитываемых признаков.

Отмеченные свойства позволяют решить задачу выбора минимального набора признаков при заданной погрешности распознавания/классификации.

Главным недостатком признаков Карунена-Лоэва является отсутствие быстрых алгоритмов их формирования. Кроме того, для получения оптимальных результатов необходимо точное знание корреляционной функции исходного изображения.

Определенную сложность представляет также проблема нахождения собственных векторов и собственных значений корреляционных матриц больших размеров. Так при размерах исходного изображения

$N_1 \times N_2$ в общем случае необходимо решать проблему собственных значений для матрицы размером $N_1 N_2 \times N_1 N_2$.

Однако известно, что для стационарных в широком смысле полей изображений асимптотически (при увеличении размеров исходного изображения) оптимальной аппроксимацией преобразования Карунена-Лоэва являются преобразование Фурье и косинусное преобразование, рассмотренные ниже.

3.4.2 Фурье-преобразование (Фурье-признаки изображения)

$$W(n_1, n_2, m_1, m_2) = e^{-i2\pi \left(\frac{m_1 m_1}{N_1} + \frac{n_2 m_2}{N_2} \right)}.$$

Данное преобразование позволяет получить спектр изображения, характеризующий его на различных пространственных частотах. К достоинствам Фурье-признаков относятся существование быстрых алгоритмов их формирования - БПФ, полнота описания изображения (преобразование обратимо), помехоустойчивость. К недостаткам – отсутствие инвариантности к целому ряду геометрических преобразований, в том числе к преобразованиям подобия.

3.4.3 Косинусное преобразование (косинусные признаки)

$$W(n_1, n_2, m_1, m_2) = \cos \left(2\pi \frac{(2n_1 + 1)m_1}{N_1} \right) \cos \left(2\pi \frac{(2n_2 + 1)m_2}{N_2} \right).$$

Достоинства и недостатки данного представления те же, что и для Фурье-признаков.

3.4.4 Преобразование Радона

Наиболее часто встречающийся вид преобразования:

$$W(n_1, n_2, m_1, m_2) = \delta(m_2 - n_1 \cos m_1 - n_2 \sin m_1).$$

Суть преобразования – перевести линейные объекты, присутствующие на изображении $f(n_1, n_2)$, в точечные объекты в пространстве признаков – образе $g(m_1, m_2)$. Действительно, для непрерывного аналога преобразования

$$g(\xi_1, \xi_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x_1, x_2) \delta(\xi_2 - x_1 \cos \xi_1 - x_2 \sin \xi_1) dx_1 dx_2 \quad (3.2)$$

линия, задаваемая в координатах входного изображения (x_1, x_2) уравнением

$$\xi_2 = x_1 \cos \xi_1 + x_2 \sin \xi_1,$$

перейдет в точку (ξ_1, ξ_2) на изображении образа. При этом параметры ξ_1 и ξ_2 играют роль угла наклона прямой (ξ_1) и расстояния от начала координат (ξ_2), как показано на рис. 3.7.

Легко заметить, что преобразование (3.2) эквивалентно интегрированию изображения вдоль соответствующей прямой:

$$g(\xi_1, \xi_2) = \int_{-\infty}^{\infty} f\left(x_1, \frac{\xi_2}{\cos \xi_1} - x_1 \operatorname{ctg} \xi_1\right) dx_1.$$

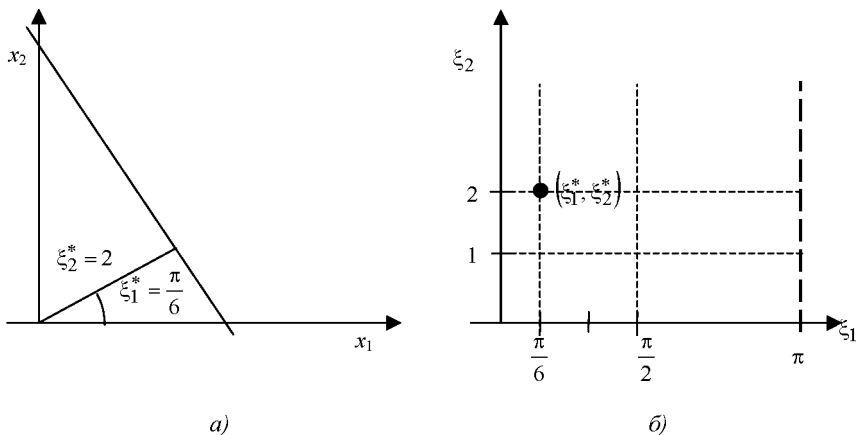


Рис. 3.7. Соответствие изображения линии и ее образа в преобразовании Радона: изображение линии (а); образ линии (б)

Иллюстрация преобразования Радона для изображения представлена на рис. 3.8.

Признаки, вычисляемые с использованием преобразования Радона, за счет нескольких дополнительных несложных преобразований над образом $g(m_1, m_2)$ могут быть сделаны инвариантными к преобразованиям подобия. Более того, они достаточно помехоустойчивы и существуют алгоритмы их быстрого вычисления, базирующиеся на алгоритме БПФ. Все это позволяет утверждать о достаточной перспективности их использования в задачах распознавания изображений.

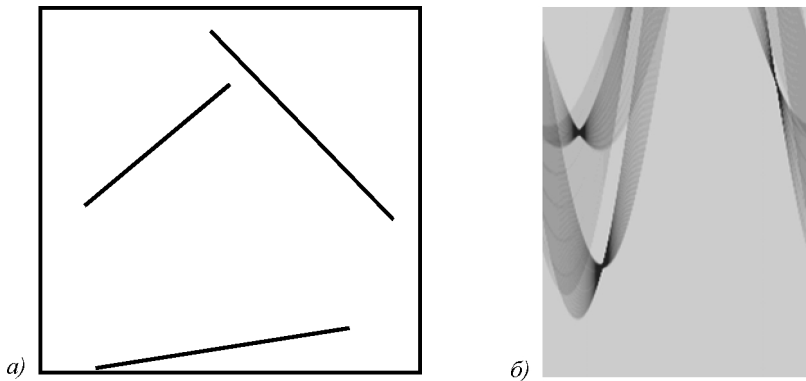


Рис. 3.8. Иллюстрация преобразования Радона: исходное изображение (а); результат преобразования (б)

3.4.5. Полиномиальные моменты

$$W(m_1, m_2, n_1, n_2) = \sum_{i=0}^{m_1} \sum_{j=0}^{m_2} a_{ij} n_1^i n_2^j.$$

Здесь a_{ij} - некоторые коэффициенты полиномиального ядра. Чаще всего при распознавании на изображениях ограничиваются степенными моментами $\mu_{m_1 m_2}$, для которых

$$a_{ij} = \begin{cases} 0, & i \neq m_1 \text{ или } j \neq m_2, \\ 1, & i = m_1 \text{ и } j = m_2. \end{cases}$$

Они рассчитывается по формуле:

$$\mu_{ij} = \sum_{i=-N_1/2}^{N_1/2} \sum_{j=-N_2/2}^{N_2/2} n_1^i n_2^j f(n_1, n_2). \quad (3.3)$$

Популярность степенных моментов обусловлена тем, что на их основе формируются так называемые *моментные инварианты*, имеющие вполне определенный физический смысл. К ним, например, относятся следующие:

$$\Phi_1 = \mu_{20} + \mu_{02}, \quad \Phi_2 = (\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2,$$

$$\Phi_3 = (\mu_{30} - 3\mu_{12})^2 + (3\mu_{21} - \mu_{03})^2, \quad \Phi_4 = (\mu_{30} + \mu_{12})^2 + (\mu_{21} + \mu_{03})^2.$$

В частности, Φ_1 задает момент инерции двумерного объекта, Φ_2 с точностью до масштабного множителя $\left(\frac{1}{\mu_{00}}\right)$ определяет эксцентриситет (удлинение) объекта и т.д.

Инвариантность указанных характеристик распространяется на преобразования типа поворота. Также можно добиться инвариантности моментных характеристик к более общим преобразованиям, в частности к аффинным преобразованиям. Более подробно эти вопросы рассмотрены в п. 4.3.

4. ФОРМИРОВАНИЕ ПРИЗНАКОВ ПО ИЗОБРАЖЕНИЮ

Формирование признаков - первый этап в любой системе распознавания образов. И качество всей системы оказывается жестко зависимо от того, насколько хорошо подобраны признаки для описания объекта - в данном случае изображения.

Задача формирования признаков достаточно сложна, потому что процесс описания изображения, или построения набора признаков, до настоящего времени остается процедурой эвристической, во многом зависимой от опыта и квалификации разработчика. В связи с этим при разработке каждой новой системы распознавания чаще всего даже специалистам в распознавании образов приходится решать её заново, ориентируясь на специфику обрабатываемых изображений и изображенных на них объектов.

В рамках данной главы представлены математические методы и алгоритмы, используемые для формирования описания изображений и эффективно применяемые для решения целого ряда задач распознавания на изображениях. Более подробно, в п. 4.1 представлены основные требования, которые обычно предъявляются к формируемым признакам изображения, а также выделены методы и алгоритмы, используемые для удовлетворения этим требованиям. В п. 4.2. описан ряд алгоритмов, используемых для «нормализации» изображения. Процесс нормализации, как будет показано, позволяет привести само изображение и признаки к виду, более удобному для распознавания. Обзор основных групп признаков изображений и математических основ их формирования, включающий также и принципиальные для задачи распознавания свойства признаков, был приведен в главе 3. В настоящей главе, в п. 4.3 представлен алгоритм формирования признаков (моментных инвариантов), значения которых оказываются инвариантными (то есть неизменными) к наиболее частым искажениям: изменению масштаба изображенного объекта и его ориентации.

4.1. Основные требования к признакам, вычисляемым по изображениям

Основным требованием к используемой системе признаков анализируемого изображения является требование эффективности процесса распознавания. Это требование часто имеет ряд противоречивых аспектов, на которых мы остановимся ниже.

Во-первых, оно предъявляет к признакам требования вычислительного характера. Они заключаются в том, чтобы существовал алгоритм расчета признаков, и этот алгоритм был вычислительно эффективен. Необходимость первого требования обусловлена тем, что разработка любой системой распознавания происходит как правило в рамках некоторого лимита «ценового» ресурса. Это ограничивает возможности использования ряда средств формирования признаков. Например, подобное ограничение в области распознавания изображений может отразиться на возможности использования видеоаппаратуры, регистрирующей цветные изображения, либо обеспечивающей повышенную разрешающую способность. Второе требование вытекает из необходимости удовлетворения определенным временным ограничениям, накладываемым на процесс распознавания в целом. Это требование достаточно типично для систем распознавания в реальном времени, например, для бортовых систем дистанционного зондирования, систем оперативного контроля и других.

Во-вторых, требование эффективности распознавания накладывает определенные требования на значения признаков. А именно, для объектов различных классов значения признаков должны отличаться сильнее, чем для объектов одного класса. Это требование иногда интерпретируют как требование компактности описания класса в пространстве признаков, когда образы объектов одного класса в пространстве признаков образуют компактные области - кластеры или таксоны. Это не совсем корректно, так как на требование «различия» или «близости» значений признаков следует смотреть с точки зрения используемого при распознавании классификатора - решающего правила. В то же время компактность образов, в случае её достижимости, позволяет существенно упростить процедуру классификации и сделать её более устойчивой.

В-третьих, требование эффективности распознавания приводит к необходимости удовлетворения требования устойчивости или инвариантности образа (описания) к ряду возможных искажений объекта. Действительно, классифицируемые на изображении объекты в реальной жизни подвергаются целому ряду изменений. Кроме того, процесс регистрации сцены приносит дополнительные искажения, связанные с неидеальностью приборов регистрации, изменением освещенности, шумами и т.д. На практике это приводит к изменению изображения объекта и, следовательно, изменению его образа в пространстве признаков. Последнее в общем случае может повлиять на результаты классификации и, следовательно, эффективность системы в целом.

Следует отметить, что если характер требований эффективности признаков, существования методов их расчета и наличия быстрых алгоритмов их вычисления часто зависит от специфики решаемой задачи, то требование инвариантности для целого ряда практических задач распознавания на изображениях оказывается достаточно общим. В частности, оно подразумевает все или некоторые из следующих требований:

- инвариантность к шумовым и динамическим искажениям;
- инвариантность к яркостным искажениям (изменению яркости и контраста);
- инвариантность к изменению местоположения объекта;
- инвариантность к изменению масштаба объекта;
- инвариантность к изменению ориентации объекта (к повороту объекта в плоскости изображения);
- инвариантность к произвольным аффинным преобразованиям;
- инвариантность к изменению ракурса съемки объекта (для трехмерных объектов).

Инвариантность к указанным искажениям в общем случае достигается за счет предварительного преобразования исходного изображения. В частности, первая группа искажений, связанная с шумовыми и динамическими изменениями изображения, устраняется за счет использования методов и алгоритмов восстановления. Подробно математические аспекты процесса восстановления и соответствующие алгоритмы описаны в [22].

Яркостные искажения, относящиеся ко второй группе, эффективно могут быть устранены за счет приведения изображения к «нормализованному» виду. Соответствующий математический аппарат и алгоритмы представлены в п. 4.2. В том же разделе приведены методы нормализации изображений объектов для искажений, включающих в себя изменения положения, масштаба и ориентации изображенного объекта (преобразования подобия).

4.2. Нормализация изображений при вычислении признаков

Под нормализацией в дальнейшем будем понимать такое преобразование изображения, которое позволяет привести его к виду, удобному для распознавания. Последнее подразумевает некий стандарт для нормализованного изображения, в качестве которого могут выступать средняя яркость, разброс или дисперсия яркости на изображении, ориентация изображенного объекта, его размеры и т.д.

В соответствии с принятым в п. 1.3 подразделением задач распознавания на две группы наблюдается соответствующее разделение на группы для алгоритмов нормализации. Связано это с особенностями реализации вычислений при нормализации изображения в целом – глобальная нормализация, либо его фрагментов – локальная нормализация. В то же время математические основы и методы нормализации оказываются одинаковыми. Они, в свою очередь, подразделяются в соответствии с требованиями инвариантности к признакам, которые представлены в п. 4.1. А именно, основными являются следующие:

- яркостная нормализация,
- нормализация положения объекта,
- нормализация масштаба объекта,
- нормализация ориентации объекта.

4.2.1 Яркостная нормализация

Одним из важных этапов решения задачи обнаружения и распознавания объектов на цифровых изображениях является нормализация яркостных искажений. Такие искажения в системах дистанционного формирования изображений возникают вследствие как природных, так и технических факторов (условий наблюдения, освещенности, влияния атмосферы, времени экспозиции и т.д.). В результате их действия яркостные статистические характеристики фона и объектов на изображении могут существенно изменяться при сохранении неизменной формы объектов. Из предположения, что основную информацию об объектах несут их геометрические характеристики, следует необходимость яркостной нормализации, то есть преобразования изображения к виду, не зависящему от случайного характера яркостных искажений.

При отсутствии априорной информации об искажениях вид и параметры нормализующего яркостного преобразования определяются непосредственно по обрабатываемому изображению. Обычно вид преобразования задается, а параметры вычисляются на основе яркостных статистических характеристик (среднего, дисперсии, гистограммы) [23,27,41,44], однако параметры могут быть определены и с помощью других признаков изображения (например, моментных инвариантов [21]).

Известные алгоритмы нормализации обеспечивают выравнивание («эквиализацию») гистограммы на изображении [26,41,44], приведение к заданным значениям («стандартизацию») среднего и дисперсии [44].

Следует отметить, что несмотря на большое количество работ по обработке изображений, в которых рассматривается нормализация яркостных искажений [21,23,26,27,41,44], данную проблему нельзя считать окончательно решенной вследствие ряда причин.

Во-первых, алгоритмы яркостной нормализации оказываются существенно различными по эффективности для различных классов изображений. Так, для нормализации изображений крупномасштабных объектов оказывается необходимым учитывать специфику математической модели такого изображения.

Во-вторых, анализ реальных изображений показывает, что на них всегда присутствует шумовая составляющая яркости с характеристиками, часто различающимися на объекте и фоне. При этом в упомянутых работах не уделено достаточного внимания влиянию шумов на яркостную нормализацию.

В-третьих, для обнаружения (распознавания) объектов на изображении иногда более важной является «стандартизация» среднего значения не по всему изображению, а на объекте и фоне в отдельности.

В данном разделе рассматриваются вопросы построения и применения нормализующего преобразования с учетом обозначенных вопросов.

Рассмотрим математическую модель изображения, в котором значительную часть поля яркости $f(n_1, n_2)$ занимает пространственно-протяженный объект. Тогда область изображения можно разбить на две, соответствующие объекту D_o и фону D_b : $D = D_o \cup D_b$. Области D можно поставить в соответствие эталонное изображение, каждый элемент которого определяется в зависимости от его принадлежности области D_o или D_b :

$$\bar{f}(n_1, n_2) = \begin{cases} 0, & \text{при } (n_1, n_2) \in D_b, \\ 1, & \text{при } (n_1, n_2) \in D_o. \end{cases} \quad (4.1)$$

Используя данное представление, запишем модель нормализуемого изображения:

$$f(n_1, n_2) = (\bar{\zeta}_o(n_1, n_2) + A_o) \bar{f}(n_1, n_2) + \bar{\zeta}_b(n_1, n_2) (1 - \bar{f}(n_1, n_2)) + A + v(n_1, n_2). \quad (4.2)$$

Будем считать, что $\bar{\zeta}_o(n_1, n_2)$, $\bar{\zeta}_b(n_1, n_2)$ - центрированные однородные случайные поля с заданными автоковариационными функциями, $v(n_1, n_2)$ - шумовая составляющая наблюдаемого изображения, A и A_o - параметры яркостного преобразования изображения относительно эталона. Величина A_o является контрастом между объектом и фоном (разность между их средними значениями), а величина A - средней яркостью фона.

Вначале, определим основные требования к искомому нормализующему преобразованию, которые естественным образом вытекающие из самой постановки задачи.

Преобразование G устранения яркостных искажений изображения должно обеспечивать:

- минимизацию некоторого критерия сходства нормализованного изображения в окне с эталонным:

$$\|y(v(f(n_1, n_2))) - y(\bar{f}(n_1, n_2))\| \rightarrow \min, (n_1, n_2) \in D, \quad (4.3)$$

где y - вектор признаков, вычисляемых на области изображения D (в качестве признаков можно использовать отсчеты изображения, тогда $y(f(n_1, n_2)) = f(n_1, n_2)$); $\|y\|$ - норма вектора y в заданном пространстве функций дискретного аргумента;

- инвариантность к линейному яркостному преобразованию изображения:

$$y(v(A \cdot f(n_1, n_2) + B)) = y(v(f(n_1, n_2))), \quad (4.4)$$

где A, B - параметры преобразования;

- для эталонного изображения (4.1) должно выполняться соотношение:

$$y(v(\bar{f}(n_1, n_2))) = y(\bar{f}(n_1, n_2)).$$

При построении преобразования могут задаваться и дополнительные требования, например, условие несмещенности математических ожиданий признаков на нормализованном изображении относительно признаков, определяемых на эталонном изображении:

$$E\{y(v(f(n_1, n_2)))\} = y(\bar{f}(n_1, n_2)). \quad (4.5)$$

4.2.1.1. Построение нормализующего преобразования

Для построения нормализующего преобразования рассмотрим некоторые свойства изображения, удовлетворяющего моделям (4.1) и (4.2). С учетом этих выражений нетрудно получить статистические характеристики изображения - среднее и дисперсию

$$\mu_f = \mu_b + (\mu_o - \mu_b)k, \quad \sigma_f^2 = A_o^2 k(1-k) + \sigma_b^2(1-k) + \sigma_o^2 k, \quad (4.6)$$

где $\mu_b = A$, $\mu_o = A_o + A$, σ_b^2 , σ_o^2 - соответственно, средние значения и дисперсии случайной составляющей на участках фона и объекта;

$$k = \frac{|D_o|}{|D|} = \frac{|D_o|}{|D_o| + |D_b|} - \text{доля площади изображения, занимаемая объектом;}$$

$|D_o|$, $|D_b|$ - количество элементов в областях D_o и D_b , соответствующих объекту и фону. Характеристики (4.6) соответствует и случаю, когда объект частично покрывается изображением. В зависимости от расположения объекта на изображении происходит увеличение доли площади k от нуля до максимального значения K , которое при заданных размерах изображения и формы объекта можно считать априорно известным.

Линейность яркостного искажения модели (4.2) дает основание для выбора нормализующего преобразования также из класса линейных поэлементных преобразований вида

$$g(n_1, n_2) = af(n_1, n_2) + b, \quad (n_1, n_2) \in D, \quad (4.7)$$

где a, b - параметры преобразования.

Наиболее распространенным критерием сходства (4.3) изображений является критерий минимума среднеквадратичной погрешности:

$$\varepsilon^2 = E \left\{ \sum_{(n_1, n_2) \in D} (af(n_1, n_2) + b - \bar{f}(n_1, n_2))^2 \right\} \rightarrow \min_{a, b} \quad (4.8)$$

при минимизации которого определяются оптимальные параметры преобразования (4.7):

$$a = \frac{A_o}{A_o^2 + \frac{\sigma_b^2}{k} + \frac{\sigma_o^2}{1-k}}, \quad b = \frac{-A_o A + \sigma_b^2 + \frac{\sigma_o^2}{1-k}}{A_o^2 + \frac{\sigma_b^2}{k} + \frac{\sigma_o^2}{1-k}}. \quad (4.9)$$

Другой способ выбора параметров преобразования (4.7) заключается в приведении среднего и дисперсии в окне к значениям, соответствующим эталонному изображению. Нетрудно показать, что и в этом случае коэффициенты преобразования (4.7) выражаются через характеристики A , A_o яркостного искажения и дисперсии случайных полей:

$$a = \sqrt{\frac{K(1-K)}{A_o^2 k(1-k) + \sigma_b^2(1-k) + \sigma_o^2 k}}, \quad b = -(A + A_o k) a + K, \quad (4.10)$$

где K - доля площади объекта на изображении при его центральном расположении (на эталонном изображении объекта).

При использовании параметров (4.9) и (4.10) преобразования (4.7) реализуются два различных алгоритма нормализации, которые назовем соответственно:

- *алгоритм МНК* (с определением параметров по методу наименьших квадратов) и
- *алгоритм НСД* (с нормализацией среднего и дисперсии).

Каждый из двух приведенных алгоритмов требует априорного знания (или оценки) параметров искажения A , A_o , дисперсий случайных полей $\bar{\xi}_b$, $\bar{\xi}_o$ и величины k и, следовательно, является весьма трудоемким. Другими их недостатками являются невыполнение условия несмещенности (4.5) и зависимость контраста (разности между средними яркости на объекте и фоне) на нормализованном изображении от уровня шума: при увеличении дисперсий случайных составляющих на изображении контраст уменьшается. Кроме того, для определения

оптимальных параметров (4.9) должно быть известно точное положение объекта на изображении.

Тем не менее, алгоритм НСД широко используется для нормализации изображений [44]. При этом параметры (4.10) вычисляются через оценки среднего $\hat{\mu}_f$ и дисперсии $\hat{\sigma}_f^2$ в окне:

$$a = \sqrt{\frac{\sigma_f^2}{\hat{\sigma}_f^2}}, \quad b = \sqrt{\frac{\sigma_f^2}{\hat{\sigma}_f^2}} + \mu_{\bar{f}}, \quad (4.11)$$

где $\mu_{\bar{f}} = K$, $\sigma_{\bar{f}}^2 = K(1-K)$ - среднее и дисперсия на эталонном изображении.

Более простой алгоритм нормализации можно построить, если пренебречь случайной составляющей на изображении. Тогда параметры преобразования (4.7) примут вид:

$$a = \frac{1}{A_o}, \quad b = -\frac{A}{A_o}, \quad (4.12)$$

в которых параметры искажения можно заменить оценками $\hat{\mu}_b$ и $\hat{\mu}_o$ средних на фоне и объекте:

$$a = \frac{1}{\hat{\mu}_o - \hat{\mu}_b}, \quad b = -\frac{\hat{\mu}_b}{\hat{\mu}_o - \hat{\mu}_b}. \quad (4.13)$$

Данные параметры при неограниченном росте отношения «сигнал/шум» на изображении стремятся к оптимальным. Очевидно, что преобразование (4.7) с параметрами (4.13) удовлетворяет требованиям, изложенным выше.

4.2.1.1.1. Особенности реализации алгоритма расчета параметров при локальной нормализации изображения

Использование полученного преобразования возможно как для глобальной, так и для локальной нормализации изображения.

В первом случае формула (4.7) применяется с постоянными параметрами a и b для всего изображения целиком. При локальной нормализации требование вычислительной эффективности вынуждает в ряде задач использовать дополнительные элементы ускорения процессов расчета параметров преобразований (4.9)-(4.13).

В частности, для практического применения преобразования (4.7) с параметрами (4.13) применяется алгоритм (назовем его НЛК-алгоритмом нормализации локального контраста), основанный на одновременном использовании двух скользящих окон [22], в которых реализуется оценка средних значений яркости на объекте и фоне. Одно из окон является основным, и в нем реализуется обработка изображения (нормализация, обнаружение объекта). Во втором (дополнительном) окне вычисляется только сумма значений элементов изображения. Форма этого окна выбирается таким образом, чтобы при центральном расположении объекта в окне, «разность» окон (элементы изображения, принадлежащие только большему из двух окон) соответствовала фону и использовалась для оценки среднего на фоне μ_b . Примером такой пары окон могут служить два прямоугольных окна с общим центром, «разность» которых является прямоугольной рамкой, как показано на рис. 4.1.

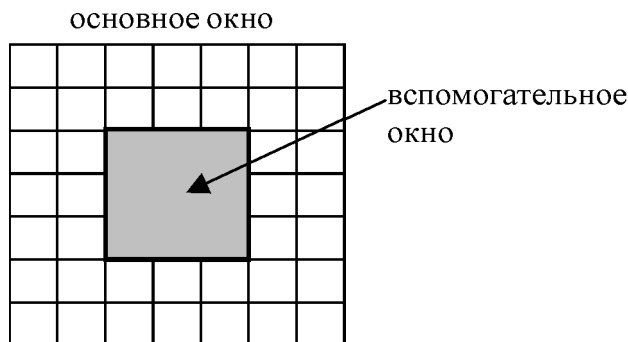


Рис. 4.1. Реализация нормализации при обработке в скользящем окне

Среднее значение яркости на «разности» окон определяется аналогично формуле (4.6):

$$\mu_g = \mu_b + (\mu_o - \mu_b) \tilde{k}^o \quad (4.14)$$

где \tilde{k} - доля площади, занимаемая объектом на «разности» окон.

При обработке изображения «двойным» скользящим окном средние значения на основном окне D и на «разности» окон вычисляются через суммы элементов в окнах:

$$\mu_f = \frac{\sum_{(n_1, n_2) \in D} f(n_1, n_2)}{|D|}, \quad (4.15)$$

$$\mu_g = \left| \frac{\sum_{(n_1, n_2) \in D} f(n_1, n_2) - \sum_{(n_1, n_2) \in \tilde{D}^o} f(n_1, n_2)}{|D| - |\tilde{D}^o|} \right|,$$

где \tilde{D} - область дополнительного окна; $|D|$, $|\tilde{D}|$ - количества элементов в окнах D , \tilde{D} .

С использованием оценок (4.15) строится нормализующее преобразование вида (4.7), параметры которого определяются по формулам:

$$a = \frac{K}{\mu_f - \mu_g}, \quad b = \frac{K \mu_g}{\mu_f - \mu_g}. \quad (4.16)$$

Подставляя параметры (4.16) в (4.7) с учетом (4.6) и (4.14), получим:

$$g(n_1, n_2) = \left. \frac{Kf(n_1, n_2) - K \mu_g}{\mu_f - \mu_g} = \frac{f(n_1, n_2) - \mu_b - (\mu_o - \mu_b) \tilde{k}^o}{(\mu_o - \mu_b) \left(\frac{k}{K} - \frac{\tilde{k}^o}{K} \right)} \right|_{(n_1, n_2) \in D} \quad (4.17)$$

При центральном расположении объекта в окне $\tilde{k} = 0$, $k=K$ и нормализующее преобразование (4.17) эквивалентно преобразованию (4.7) с параметрами (4.13).

Для повышения устойчивости работы алгоритма нормализации (4.17) на участках фона, где разность оценок μ_f и $\tilde{\mu}_f$ близка к нулю, может быть применена регуляризация Тихонова [2], в соответствии с которой вместо параметров (4.16) используются:

$$a = \frac{K(\mu_f - \tilde{\mu}_f)}{(\mu_f - \tilde{\mu}_f)^2 + \gamma^2}, \quad b = -\frac{K \tilde{\mu}_f (\mu_f - \tilde{\mu}_f)}{(\mu_f - \tilde{\mu}_f)^2 + \gamma^2}, \quad (4.18)$$

где γ - параметр регуляризации.

Следует отметить, что регуляризация необходима и при использовании алгоритма НСД в случае наличия на изображении участков фона с малой дисперсией яркости.

4.2.1.1.2. Рекурсивная реализация алгоритма расчета параметров локальной яркостной нормализации

Определение оценок средних значений (4.15) локального нормализующего преобразования (4.17) сводится к вычислению локальных сумм элементов изображения на основном и дополнительном окнах, которое нетрудно организовать рекурсивным образом, если окна являются прямоугольными.

Вычисление суммы в скользящем окне можно рассматривать как обработку изображения фильтром с конечной импульсной характеристикой вида

$$h(m_1, m_2) = \begin{cases} 1 & \text{if } (m_1, m_2) \in D, \\ 0 & \text{if } (m_1, m_2) \notin D. \end{cases}$$

Для прямоугольного окна D

$$-N_1^- \leq n_1 \leq N_1^+, \quad -N_2^- \leq n_2 \leq N_2^+,$$

фильтр является факторизуемым и допускает параллельно-рекурсивную реализацию [41]:

$$\begin{aligned}
S(n_1, n_2) &= S(n_1 - 1, n_2) + f(n_1 + N_1^-, n_2) - f(n_1 - N_1^+ - 1, n_2), \\
s(n_1, n_2) &= s(n_1, n_2 - 1) + S(n_1, n_2 + N_2^-) - S(n_1, n_2 - N_2^+ - 1),
\end{aligned} \tag{4.19}$$

где $S(n_1, n_2)$ - локальные суммы элементов изображения в столбцах окна.

Данные рекуррентные соотношения могут быть также реализованы, если рассматривать локальные суммы как моменты нулевого порядка, при применении параллельно-рекурсивного алгоритма вычисления моментов изображения [10-12, 22].

Основным достоинством алгоритма является независимость его вычислительной сложности от размеров окна (всего 4 арифметической операции для каждого положения скользящего окна) при незначительном увеличении памяти (для одномерного массива сумм элементов в столбцах).

Приведенный алгоритм и оценка его сложности полностью распространяются для второго прямоугольного окна (дополнительного):

$$\tilde{D}: -\tilde{N}_1^- \leq n_1 \leq \tilde{N}_1^+, -\tilde{N}_2^- \leq n_2 \leq \tilde{N}_2^+. \tag{4.20}$$

Однако существуют возможности для упрощения реализации алгоритма. Так, задавая окна равными по высоте ($\tilde{N}_1^\sigma = \tilde{N}_1^\sigma$, $\tilde{N}_2^\sigma = \tilde{N}_2^\sigma$), получим для дополнительного окна, что рекурсивное соотношение для столбцов в (4.19) остается в силе, а вместо второго выражения необходимо использовать рекурсию:

$$\tilde{s}(n_1, n_2) = \tilde{s}(n_1, n_2 - 1) + \tilde{S}(n_1, n_2 + \tilde{N}_2^-) - \tilde{S}(n_1, n_2 - \tilde{N}_2^+ - 1). \tag{4.21}$$

В этом случае «разность окон», используемая для оценки среднего фона (4.15) представляет собой пару вертикальных полос. Для вычисления суммы на втором окне требуется только две операции сложения без дополнительного увеличения памяти.

При реализации преобразования (4.7) с параметрами (4.11) кроме вычисления среднего определяется дисперсия в окне. Для расчета

дисперсии организуется рекурсивное вычисление сумм квадратов отсчетов изображения в окне по формулам, аналогичным (4.19).

4.2.1.1.3. Особенности применения нормализующего преобразования в скользящем окне

При обработке изображения для каждого положения скользящего окна вычисляются свои параметры нормализующего преобразования. Следовательно, при скользящей нормализации невозможно сформировать единое нормализованное изображение, а можно построить серию нормализованных (с различными параметрами a и b) фрагментов изображения, которые соответствуют всем возможным положениям скользящего окна на изображении.

Однако нормализация является, как правило, промежуточным этапом обработки изображения, после которого на нормализованном изображении фрагмента вычисляются признаки, используемые для обнаружения (распознавания) объектов [11,27]. Если процедура расчета признаков реализуется в виде линейной системы с постоянными параметрами, можно вычислять признаки y (например, моментные характеристики) непосредственно на исходном изображении и далее преобразовывать их с помощью рассчитанных параметров нормализации a и b для текущего положения окна:

$$y(g(n_1, n_2)) = a \cdot y(f(n_1, n_2)) + y(b)(n_1, n_2) \in D. \quad (4.22)$$

Следует отметить, что предлагаемое преобразование (4.7) с параметрами (4.16) обеспечивает нормализацию яркостных искажений только в случае полного захвата объекта основным окном D . При частичном расположении объекта в окне нормализация осуществляется неточно, поскольку на «разности» окон присутствует часть объекта. Тем не менее, для обнаружения (распознавания) это обстоятельство не является существенным, поскольку в пространстве признаков область, соответствующую объектам, определяют только векторы признаков, вычисленные для положений окна, полностью покрывающих объект [11].

4.2.1.1.4. Анализ эффективности алгоритмов яркостной нормализации в скользящем окне

Для практического применения любого из рассматриваемых алгоритмов яркостной нормализации изображения необходимо проведение теоретического анализа и экспериментальных исследований алгоритма с целью:

- оценки качества нормализации;
- определения оптимальных параметров нормализующего преобразования;
- определения допустимого диапазона параметров искажения, для которого алгоритм сохраняет работоспособность.

В общем случае оценить качество яркостной нормализации изображения можно только по конечным результатам решения конкретной задачи обработки изображения (например, распознавания объектов в условиях неопределенности яркостных искажений). Однако такой подход к оценке качества требует чрезвычайно трудоемкого моделирования всего многозвенного процесса решения прикладной задачи. Поэтому целесообразно использовать некоторые частные критерии, характеризующие качество решения отдельных этапов задачи и позволяющие провести исследования реализуемых алгоритмов с целью выработки рекомендаций по их практическому использованию. Здесь для оценки качества алгоритма нормализации будем использовать критерий сходства нормализованного и эталонного изображений или признаков, вычисляемых на этих изображениях:

- относительную среднеквадратичную погрешность отклонения нормализованного изображения от эталонного:

$$\bar{\varepsilon}^2 = \frac{\varepsilon^2}{|D|} = E \left\{ \sum_{(n_1, n_2) \in D} (a f(n_1, n_2) + b - \bar{f}(n_1, n_2))^2 \right\} \frac{1}{|D|}, \quad (4.23)$$

где $|D|$ - число элементов изображения в окне;

- расстояние Махаланобиса [40] между векторами признаков, определяемых на эталонном и нормализованном изображениях:

$$R = \left(\mathbf{Y}(g) - \mathbf{y}(\bar{f}) \right)^T \mathbf{B}^{-1} \left(\mathbf{Y}(g) - \mathbf{y}(\bar{f}) \right),$$

где \mathbf{B} – ковариационная матрица случайного вектора $\mathbf{Y}(g)$ (компоненты вектора $\mathbf{y}(\bar{f})$ не являются случайными).

При исследовании алгоритмов будем рассматривать положение объекта в центре изображения, поскольку, как отмечалось выше, основное значение для распознавания объектов имеет качество нормализации именно в такой ситуации.

Пусть заданное на области D и содержащее локальный объект изображения удовлетворяет следующей упрощенной математической модели:

$$f(n_1, n_2) = A_o \bar{f}(n_1, n_2) + A + v(n_1, n_2), \quad (4.24)$$

где

$$\bar{f}(n_1, n_2) = \begin{cases} 0 & \text{при } (n_1, n_2) \in D_b, \\ 1 & \text{при } (n_1, n_2) \in D_o, \end{cases}$$

где $v(n_1, n_2)$ – белый шум, то есть случайное поле с нулевым математическим ожиданием и автоковариационной функцией

$$B_v(n_1, n_2; m_1, m_2) = \sigma_v^2 \delta(n_1 - m_1) \delta(n_2 - m_2),$$

σ_v^2 – дисперсия значений отсчетов изображения, $\delta(n)$ – единичный импульс.

Для яркостной нормализации изображения применим поэлементное линейное преобразование

$$g(n_1, n_2) = a f(n_1, n_2) + b, \quad (n_1, n_2) \in D, \quad (4.25)$$

где a, b – параметры преобразования, определяемые согласно одному из рассмотренных алгоритмов. Очевидно, что автоковариационная функция нормализованного изображения имеет вид:

$$B_g(n_1, n_2; m_1, m_2) = \sigma_g^2 \delta(n_1 - m_1) \delta(n_2 - m_2) = a^2 B_f(n_1, n_2; m_1, m_2),$$

причем, в соответствии с моделью (4.24), выполняется равенство:

$$B_f(n_1, n_2; m_1, m_2) = B_v(n_1, n_2; m_1, m_2). \quad (4.26)$$

Если признаками являются все элементы изображения в окне, то из выражений (4.24)-(4.26) выводится соотношение:

$$\begin{aligned} R &= \sum_{(n_1, n_2) \in D} \sum_{(m_1, m_2) \in D} (g(n_1, n_2) - \bar{f}(n_1, n_2)) B_g^{-1}(n_1, n_2; m_1, m_2) \times \\ &\times (g(m_1, m_2) - \bar{f}(m_1, m_2)) = \\ &= \sum_{(n_1, n_2) \in D} (g(n_1, n_2) - \bar{f}(n_1, n_2))^2 \frac{1}{\sigma_g^2} = \frac{\bar{\varepsilon}^2 |D|}{a^2 \sigma_v^2}. \end{aligned} \quad (4.27)$$

Для каждого из рассмотренных выше алгоритмов определения параметров a , b преобразования (4.25) выведем оценки показателей качества, соответствующие упрощенной модели (4.24).

4.2.1.1.5. Алгоритм МНК

Определение параметров a , b путем минимизации погрешности (4.23) (с помощью метода наименьших квадратов) сводится к решению системы линейных уравнений:

$$\left\{ \begin{aligned} a E \left\{ \sum_{(n_1, n_2) \in D} f^2(n_1, n_2) \right\} + b E \left\{ \sum_{(n_1, n_2) \in D} f(n_1, n_2) \right\} &= \\ &= E \left\{ \sum_{(n_1, n_2) \in D} f(n_1, n_2) \bar{f}(n_1, n_2) \right\}, \\ a E \left\{ \sum_{(n_1, n_2) \in D} f(n_1, n_2) \right\} + b \sum_{(n_1, n_2) \in D} 1 &= \sum_{(n_1, n_2) \in D} \bar{f}(n_1, n_2) \end{aligned} \right.$$

или с учетом (4.24)

$$\begin{cases} a(A_o^2|D_o| + A^2|D| + 2AA_o|D_o| + \sigma_v^2|D|) + \\ \quad + b(A_o|D_o| + A|D|) = (A + A_o)|D_o|, \\ a(A_o|D_o| + A|D|) + b|D| = |D_o|, \end{cases}$$

Из данной системы находим:

$$a = \frac{A_o |D_o| |D_b|}{A_o^2 |D_o| |D_b| + \sigma_v^2 |D|^2} = \frac{A_o}{A_o^2 + \frac{\sigma_v^2}{K(1-K)}},$$

$$b = \frac{-A_o A |D_o| |D_b| + \sigma_v^2 |D_o| |D|}{A_o^2 |D_o| |D_b| + \sigma_v^2 |D|^2} = \frac{-A_o A + \frac{\sigma_v^2}{1-K}}{A_o^2 + \frac{\sigma_v^2}{K(1-K)}},$$

где $K = \frac{|D_o|}{|D|}$ - доля окна, занимаемая объектом. При этих значениях a

и b погрешность $\bar{\varepsilon}^2$ достигает минимума и равна:

$$\begin{aligned} \bar{\varepsilon}^2 &= M \left\{ \sum_{(n_1, n_2) \in D} \bar{f}^2(n_1, n_2) - a \sum_{(n_1, n_2) \in D} f(n_1, n_2) \bar{f}(n_1, n_2) - \right. \\ &\quad \left. - b \sum_{(n_1, n_2) \in D} \bar{f}(n_1, n_2) \right\} \frac{1}{|D|} = \\ &= \left(|D_o| - a(A_o|D_o| + A|D_o|) - b|D_o| \right) \frac{1}{|D|} = \\ &= \left(1 - \frac{A}{A_o^2 + \frac{\sigma_v^2}{K(1-K)}} (A + A_o) - \frac{-A_o A + \frac{\sigma_v^2}{1-K}}{\frac{\sigma_v^2}{K(1-K)}} \right) K = \\ &= \frac{\sigma_v^2}{A_o^2 + \frac{\sigma_v^2}{K(1-K)}} = \frac{1}{d^2 + \frac{1}{K(1-K)}}, \end{aligned}$$

где $d = (A_o/\sigma_v)$ - отношение «сигнал/шум» на искаженном изображении. Расстояние Махаланобиса (4.27) в этом случае определяется:

$$R = \frac{\bar{\varepsilon}^2 |D|}{a^2 \sigma_f^2} = \frac{\sigma_f^2}{A^2 + \frac{\sigma_f^2}{K(1-K)}} \left(\frac{A^2 + \frac{\sigma_f^2}{K(1-K)}}{A} \right)^2 \frac{|D|}{\sigma_f^2} =$$

$$= \left(1 + \frac{1}{K(1-K) \sigma_f^2} \right) |D|.$$

4.2.1.1.6. Алгоритм НСД

Параметры a и b преобразования (4.10), обеспечивающие нормализацию среднего и дисперсии в окне, для упрощенной модели изображения определяются в виде:

$$a = \sqrt{\frac{K(1-K)}{A_o^2 K(1-K) + \sigma_v^2}}, \quad b = -(A + A_o K)a + K. \quad (4.28)$$

Подставляя (4.28) в (4.23) и (4.27), получим:

$$\bar{\varepsilon}^2 = \frac{1}{|D|} E \left\{ a^2 \sum_{(n_1, n_2) \in D} f^2(n_1, n_2) + 2ab \sum_{(n_1, n_2) \in D} f(n_1, n_2) + \right.$$

$$+ b^2 \sum_{(n_1, n_2) \in D} 1 - 2a \sum_{(n_1, n_2) \in D} f(n_1, n_2) \bar{f}(n_1, n_2) -$$

$$\left. - 2b \sum_{(n_1, n_2) \in D} \bar{f}(n_1, n_2) + \sum_{(n_1, n_2) \in D} \bar{f}^2(n_1, n_2) \right\} =$$

$$= a^2 (A_o^2 K + A^2 + 2A A_o K + \sigma_v^2) + 2ab(A_o K + A) + b^2 -$$

$$- 2aK(A + A_o) - 2bK + K = 2(K - K^2) (1 - A_o a) =$$

$$= 2K(1-K) \left(1 - \frac{A_o}{\sqrt{A_o^2 + \frac{\sigma_v^2}{K(1-K)}}} \right) = 2K(1-K) \left(1 - \frac{1}{\sqrt{1 + \frac{1}{K(1-K)d^2}}} \right);$$

$$\begin{aligned}
R &= \frac{2(K-K^2)(1-A_o a) |D|}{\alpha^2 \sigma_v^2} = \\
&= 2(K-K^2) \left(\frac{A_o^2 K(1-K) + \sigma_v^2}{K(1-K)} - A_o \sqrt{A_o^2 + \frac{\sigma_v^2}{K(1-K)}} \right) \frac{|D|}{\sigma_v^2} = \\
&= 2(K-K^2) \left(d^2 + \frac{1}{K(1-K)} - d \sqrt{d^2 + \frac{1}{K(1-K)}} \right) |D|.
\end{aligned}$$

4.2.1.1.7. Алгоритм НЛК

При использовании алгоритма нормализации локального контраста оцениваются параметры (4.12):

$$a = \frac{1}{A_o}, \quad b = -\frac{A}{A_o}.$$

Путем подстановки этих величин и модели (4.24) в выражения (4.23) и (4.27) определим показатели эффективности:

$$\begin{aligned}
\bar{\varepsilon}^2 &= E \left\{ \sum_{(n_1, n_2) \in D} \left(\frac{1}{A_o} (A_o \bar{f}(n_1, n_2) + A + v(n_1, n_2)) - \right. \right. \\
&\quad \left. \left. - \frac{A}{A_o} - \bar{f}(n_1, n_2) \right)^2 \right\} \frac{1}{|D|} = \\
&= E \left\{ \sum_{(n_1, n_2) \in D} \frac{v^2(n_1, n_2)}{A_o^2} \right\} \frac{1}{|D|} = \frac{\sigma_v^2}{A_o^2} = \frac{1}{\sigma_v^2},
\end{aligned}$$

$$R = \frac{\bar{\varepsilon}^2 |D|}{\left(\frac{1}{A_o} \right)^2 \sigma_v^2} = |D|.$$

Таким образом, получены формулы (4.29)-(4.31), используемые для оценки качества алгоритмов яркостной нормализации:

- алгоритм МНК:

$$\bar{\varepsilon}^2 = \frac{\varepsilon^2}{|D|} = \frac{1}{d^2 + \frac{1}{K(1-K)}}, \quad R = \left(1 + \frac{1}{K(1-K) d^2}\right) |D|. \quad (4.29)$$

- алгоритм НСД:

$$\bar{\varepsilon}^2 = 2K(1-K) \left(1 - \frac{1}{\sqrt{1 + \frac{1}{K(1-K) d^2}}}\right),$$

$$R = 2K(1-K) \left(d^2 + \frac{1}{K(1-K)} - d \sqrt{d^2 + \frac{1}{K(1-K)}}\right) |D|: \quad (4.30)$$

- алгоритм НЛК:

$$\bar{\varepsilon}^2 = 1/d^2, \quad R = |D|. \quad (4.31)$$

- Зависимости $\bar{\varepsilon}^2$ и $R/|D|$ от d (при $K=0,5$) и от K (при $d=5$) для рассмотренных алгоритмов показаны на рис. 4.2, на которых кривые соответствуют алгоритмам МНК, НСД, НЛК.
- Анализируя теоретические оценки алгоритмов, можно сделать следующие выводы:
- алгоритм НЛК является наилучшим среди рассматриваемых по критерию расстояния Махаланобиса в пространстве признаков, но худшим по критерию СКО;
- при больших значениях сигнал/шум $d \geq 4$ показатели качества (ε^2 и R) для различных алгоритмов нормализации достаточно близки между собой;

- показатели качества первых двух алгоритмов зависят от доли площади объекта в окне и достигают экстремальных значений при $K=0,5$.

Кроме этих выводов из общих соображений и ранее приведенных данных можно дать следующие рекомендации по применению процедур яркостной нормализации:

- наибольший практический интерес (с точки зрения вычислительной сложности) представляют два алгоритма нормализации - НСД и НЛК; при больших d эффективность применения алгоритмов практически одинакова, поэтому рекомендуется применять алгоритм НЛК, поскольку его вычислительная сложность значительно меньше;
- размеры окон обработки должны обеспечивать необходимую устойчивость статистических оценок и полностью покрывать объект.

4.2.2. Нормализация масштаба объекта

Для построения алгоритма воспользуемся введенной ранее математической моделью наблюдения изображения (4.1) или (4.2). Изменение масштаба объекта приводит к тому, что доля относящихся к объекту отсчетов будет увеличена или уменьшена пропорционально квадрату величины изменения линейных размеров. Пусть K – доля площади изображения, занятая объектом на эталонном изображении, $\mu_o = A_o + A$ и $\mu_b = A$ – средние объекта и фона. При изменении масштаба средние изображения изменяются в соответствии с изменением доли k отсчетов, относящихся к объекту:

$$\mu_f = k\mu_o + (1-k)\mu_b,$$

откуда

$$k = \frac{\mu_f - \mu_b}{\mu_o - \mu_b}.$$

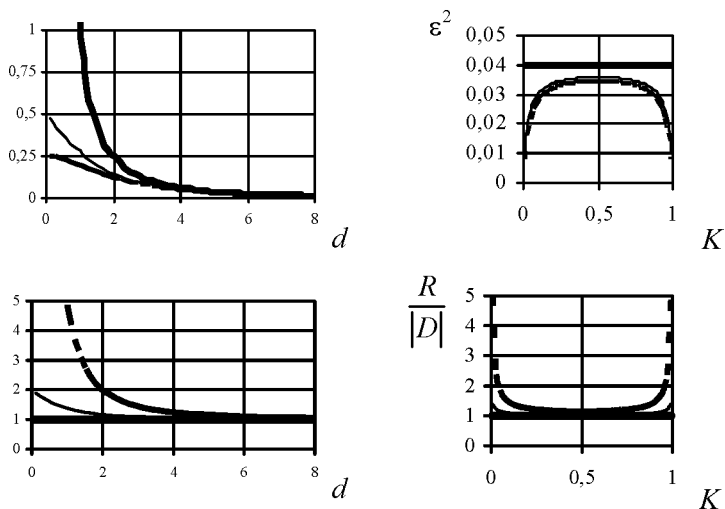


Рис. 4.2. Качество яркостной нормализации изображения с центральным расположением объекта (“- - -” - МНК, “—” - НСД, “—” - НЛК)

В результате линейное изменение масштаба $\alpha = \sqrt{k/K}$. При реализации данного алгоритма для локальной нормализации масштаба, величина среднего с точностью до постоянного множителя может быть эффективно вычислена с использованием параллельно-рекурсивного алгоритма (4.19).

Заметим, что до использования алгоритма нормализации масштаба необходимо воспользоваться алгоритмом яркостной нормализации. В противном случае средняя величина яркости, зависящая от яркости отсчетов изображения, не позволит получить корректное значение k и, следовательно, правильно определить изменение масштаба.

4.2.3 Нормализация положения объекта

При нормализации положения объекта задачей является нахождение такого положения (x_1^c, x_2^c) нового начала координат в плоскости изображения, при котором размещение этого объекта окажется в определенном смысле стандартным (например, центральным). Заметим,

что данное преобразование изображения характерно лишь для глобальной нормализации, поскольку при локальном подходе окно обработки занимает все возможные положения на изображении и обязательно найдется такое, при котором объект будет иметь требуемое расположение в анализируемом фрагменте.

Существует целый ряд подходов к нормализации положения объекта. Из них наиболее типичным является подход к определению начала координат по *центру «тяжести» изображения* [3].

Как известно [3], координаты центра тяжести (x_1^c, x_2^c) плоской фигуры S определяются следующим образом:

$$x_1^c = \frac{\iint_{D_S} x_1 \rho(x_1, x_2) dx_1 dx_2}{\iint_{D_S} \rho(x_1, x_2) dx_1 dx_2}, \quad x_2^c = \frac{\iint_{D_S} x_2 \rho(x_1, x_2) dx_1 dx_2}{\iint_{D_S} \rho(x_1, x_2) dx_1 dx_2}, \quad (4.32)$$

здесь $\rho(x_1, x_2)$ - функция плотности фигуры в точке (x_1, x_2) , D_S - область, занимаемая фигурой. Интерпретируя значения функции яркости как значения «тяжести» в отсчетах наблюдаемого изображения, получаем следующие выражения для вычисления положения нового начала координат в плоскости дискретного изображения:

$$\bar{x}_1^c = \frac{\sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} n_1 f(n_1, n_2)}{\sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} f(n_1, n_2)}, \quad \bar{x}_2^c = \frac{\sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} n_2 f(n_1, n_2)}{\sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} f(n_1, n_2)}.$$

В случае, когда функция яркости имеет всего две градации, то есть наблюдаемое изображение бинарное, приведенные выражения преобразуются к виду:

$$\bar{x}_1^c = \frac{1}{|D_1|} \sum_{(n_1, n_2) \in D_1} n_1, \quad \bar{x}_2^c = \frac{1}{|D_1|} \sum_{(n_1, n_2) \in D_1} n_2,$$

где $D_1 = \{ (n_1, n_2): (n_1, n_2) \neq 1 \}$ - область отсчетов с ненулевыми значениями функции яркости. Полученные выражения отражают уже статистический подход к нахождению нового начала координат - оно определяется как средняя величина координат присутствующих на изображении точек.

В дополнение к изложенному возможны также подходы к определению начала координат, основанные на геометрической модели объекта (центр описывающего многогранника, окружности и т.п.), использующие степенные моменты более высокого порядка [3,6], и др.

4.2.4 Нормализация ориентации объекта

Изложенный ниже *геометрический подход* к определению ориентации объекта предложен в монографии [3]. Он основан на использовании степенных моментов (см. п.3.4.5).

Выберем в качестве нормализованной ориентации такое положение центрированного изображения, при котором момент второго порядка $\mu_{11} = 0$. Для осесимметричного объекта данное условие соответствует ситуации, когда ось симметрии объекта расположена вдоль оси абсцисс или ординат.

Пусть изображение некоторым образом повернуто относительно своего нормализованного положения. Тогда соответствие между моментами второго порядка запишется таким образом:

$$\mu_{11}(0) = \frac{1}{2}(\mu_{20} - \mu_{02}) \cdot \sin 2\varphi + \mu_{11} \cos \varphi,$$

здесь $\mu_{11}(0) = 0$ - момент нормализованного изображения, φ - угол поворота изображения против часовой стрелки, относительно начального положения. Тогда

$$\varphi = \frac{1}{2} \arctg \frac{2\mu_{11}}{\mu_{02} - \mu_{20}} + t \frac{\pi}{2}, \quad t = 0, 1, 2, 3 \quad (4.33)$$

Данное выражение определяет ориентацию не однозначно, а с точностью до поворота на $t(\pi/2)$. Для однозначного ориентирования

изображения (с не более двумя осями симметрии) можно воспользоваться значениями моментов μ_{02} , μ_{01} , μ_{30} следующим образом [3]: в выражении (4.33) коэффициент t выбирается исходя из условий

$$t = \begin{cases} 0, & \text{при } \mu_{20} > \mu_{02}, \quad \mu_{30} > 0, \\ 1, & \text{при } \mu_{20} < \mu_{02}, \quad \mu_{30} > 0, \\ 2, & \text{при } \mu_{20} > \mu_{02}, \quad \mu_{30} < 0, \\ 3, & \text{при } \mu_{20} < \mu_{02}, \quad \mu_{30} < 0. \end{cases}$$

При наличии на объекте (изображении) более двух осей симметрии необходимо использовать более сложные алгоритмы нормализации ориентации [3].

В случае нормализации бинарного изображения, можно использовать *статистический подход*. В этом случае в качестве ориентации объекта выбирают направление собственного вектора x_λ матрицы ковариации \mathbf{V} координат отсчетов ненулевой яркости. При этом собственный вектор должен соответствовать максимальному собственному числу λ_{\max} матрицы ковариации, как показано на рис. 4.3.

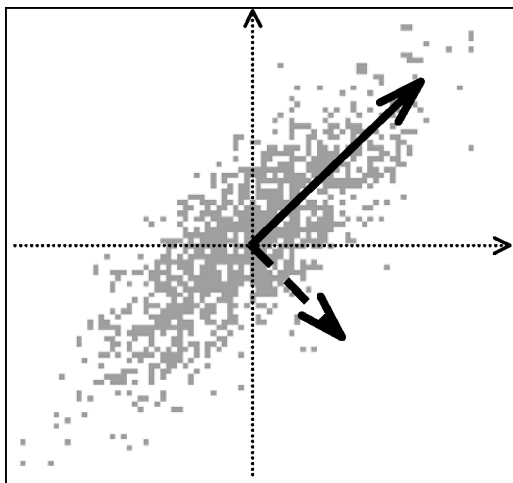


Рис. 4.3. К определению ориентации по собственному вектору матрицы ковариации: «—» — собственный вектор, соответствующий максимальному собственному числу, «- -» — собственный вектор, соответствующий минимальному собственному числу, «.....» — оси координат изображения

Пусть матрица ковариации имеет вид:

$$\mathbf{B} = \begin{pmatrix} a & c \\ c & b \end{pmatrix},$$

возможные собственные значения λ находятся из уравнения:

$$(\mathbf{B} - \lambda \mathbf{E}) \mathbf{x}_\lambda = 0, \quad (4.34)$$

где \mathbf{E} - единичная матрица, \mathbf{x}_λ - собственный вектор, соответствующий числу λ . Из уравнения (4.34) получаем

$$\lambda_{\max} = \frac{a + b + \sqrt{(a - b)^2 + 4c^2}}{2}.$$

Тогда собственный вектор, соответствующий λ_{\max} :

$$\mathbf{x}_\lambda = \begin{bmatrix} \frac{c}{b - a + \sqrt{(b - a)^2 + 4c^2}} \\ \frac{c}{2} \end{bmatrix},$$

что соответствует углу наклона φ такому, что

$$\operatorname{tg} \varphi = \frac{2c}{b - a + \sqrt{(b - a)^2 + 4c^2}}. \quad (4.35)$$

Для центрированного изображения $b = \mu_{20}$, $a = \mu_{02}$, $c = \mu_{11}$:

$$\varphi = \operatorname{arctg} \frac{2\mu_{11}}{\mu_{20} - \mu_{02} + \sqrt{(\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2}}, \quad |\alpha| < \frac{\pi}{2}.$$

Перепишем полученное выражение в виде

$$\operatorname{tg} \alpha = \frac{\frac{2\mu_{11}}{\mu_{20} - \mu_{02}}}{1 + \sqrt{1 + \left[\frac{2\mu_{11}}{\mu_{20} - \mu_{02}} \right]^2}}. \quad (4.36)$$

Учитывая, что для произвольного угла α выполняется тригонометрическое тождество:

$$\operatorname{tg} \alpha = \frac{\operatorname{tg}(2\alpha)}{1 + \sqrt{1 + [\operatorname{tg}(2\alpha)]^2}}, \quad |\alpha| < \frac{\pi}{4},$$

получаем окончательно

$$\alpha = \frac{1}{2} \operatorname{arctg} \left(\frac{2\mu_{11}}{\mu_{20} - \mu_{02}} \right), \quad |\alpha| < \frac{\pi}{4}. \quad (4.37)$$

Легко заметить, что данное выражение для ориентации совпадает с главным значением угла (4.33). То есть результаты определения ориентации объекта на изображении на основе геометрического и статистического подходов в принципе совпадают. В то же время достоинство первого подхода состоит в однозначности определения ориентации, в том числе, при нескольких осях симметрии изображенного объекта.

В заключении заметим, что необходимость нормализации зависит от того, какие именно признаки используются при распознавании. Если используемая система признаков инвариантна к некоторым преобразованиям (яркостным, геометрическим), то необходимость предварительной нормализации изображения отпадает. И наоборот, использование неинвариантных к искажениям признаков приводит к необходимости применения к изображениям нормализующих преобразований.

В п. 3 выделены основные группы признаков, используемых при распознавании на изображениях, а также указано, какие из них требуют нормализующих преобразований, а какие являются инвариантными.

4.3. Моментные инварианты как признаки изображения

Во многих задачах обработки цифровых изображений нашли широкое применение моментные характеристики изображений и рассчитываемые на их основе моментные инварианты [3, 6, 18, 19, 21, 45, 46, 52-57, 71, 73-75]. Основным достоинством моментных инвариантов является нечувствительность к поворотам изображения, что делает эффективным

их применение в качестве признаков в задаче обнаружения и распознавания на изображении объектов неизвестной ориентации. Более того, путем несложных преобразований на базе моментных инвариантов формируются признаки, устойчивые к преобразованиям подобия. Наконец, при поиске и распознавании объектов на изображении в рамках технологии «скользящего окна» применение моментных инвариантов особенно уместно в связи с возможностью параллельно-рекурсивного вычисления локальных моментных характеристик [10-12, 22].

В рамках настоящего раздела:

- произведена формализация общего алгоритма синтеза функционально независимых моментных инвариантов произвольного порядка;
- приведен быстрый алгоритм расчета моментных инвариантов в режиме скользящего окна (локальной нормализации) и приведены оценки вычислительной сложности;
- рассмотрен способ формирования на базе степенных моментов аффинных инвариантов – признаков изображения, инвариантных к аффинным преобразованиям – по материалам работ [52, 53].

4.3.1. Моментные инварианты в обработке изображений

Рассмотрим сначала изображение как функцию двух непрерывных аргументов $f(x_1, x_2)$. Степенные моменты порядка (k, l) определяются согласно формуле

$$\mu_{kl} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1^k x_2^l f(x_1, x_2) dx_1 dx_2, \quad (4.38)$$

$k, l = 0, 1, \dots$

Обычно в задачах распознавания используются центральные моменты, обладающие инвариантностью к сдвигу изображения:

$$\bar{\mu}_{kl} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x_1 - x_1^c)^k (x_2 - x_2^c)^l f(x_1, x_2) dx_1 dx_2, \quad (4.39)$$

$k, l = 0, 1, \dots$

где $x_1^c = \frac{\mu_{10}}{\mu_{00}}$, $x_2^c = \frac{\mu_{01}}{\mu_{00}}$ - координаты центра тяжести изображения

(4.37). Очевидно, что центральные моменты (4.39) выражаются через моменты (4.38):

$$\bar{\mu}_{kl} = \sum_{i=0}^k \sum_{j=0}^l (-1)^{i+j} C_k^i C_l^j \bar{t}_1^{k-i} \bar{t}_2^{l-j} \mu_{ij}, \quad (4.40)$$

где C_k^i, C_l^j - биномиальные коэффициенты. Для центрированного изображения (при $x_1^c = x_2^c = 0$), значения моментов (4.38) и (4.39) совпадают. Для определенности в дальнейшем будем рассматривать именно эту ситуацию.

С помощью степенных моментов определяются характеристики, инвариантные к повороту изображения (моментные инварианты). Так, набор характеристик, приведенный в [18, 21, 71], включает семь инвариантов:

$$\begin{aligned} \Phi_1 &= \mu_{20} + \mu_{02}, \\ \Phi_2 &= (\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2, \\ \Phi_3 &= (\mu_{30} - 3\mu_{12})^2 + (3\mu_{21} - \mu_{03})^2, \\ \Phi_4 &= (\mu_{30} + \mu_{12})^2 + (\mu_{21} + \mu_{03})^2, \\ \Phi_5 &= (\mu_{30} - 3\mu_{12})(\mu_{30} + \mu_{12}) \left[(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2 \right] + \\ &\quad + (3\mu_{21} - \mu_{03})(\mu_{21} + \mu_{03}) \left[3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2 \right], \\ \Phi_6 &= (\mu_{20} - \mu_{02}) \left[(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2 \right] + \\ &\quad + 4\mu_{11}(\mu_{30} + \mu_{12})(\mu_{21} + \mu_{03}), \\ \Phi_7 &= (\mu_{21} - \mu_{03})(\mu_{30} + \mu_{12}) \left[(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2 \right] + \\ &\quad + (\mu_{30} - 3\mu_{12})(\mu_{21} + \mu_{03}) \left[3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2 \right]. \end{aligned} \quad (4.41)$$

В данном наборе используются только моменты до порядка (k, l) при $k+l \leq 3$. В работах [3, 45, 76] используется более узкий набор инвариантов, входящих в указанное множество (4.41). Но при любом выборе набора возникает два вопроса:

- о функциональной избыточности наборов, то есть о возможности выразить один из инвариантов как функцию других [7, 54];
- о полноте набора, то есть о возможности с помощью моментов до заданного порядка построения других функционально независимых инвариантов [7, 54].

4.3.2. Метод построения моментных инвариантов произвольного порядка

Для ответа на поставленные выше вопросы кратко рассмотрим предложенный в [45, 46] метод построения моментных инвариантов с использованием комплексных моментов:

$$M_{kl} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x_1 + i x_2)^k (x_1 - i x_2)^l (x_1, x_2) dx_1 dx_2, \quad (4.42)$$

$k, l = 0, 1.$

Комплексные моменты являются линейной комбинацией обычных моментов:

$$M_{kl} = \sum_{i,j} a_{ij} \mu_{ij} + i \sum_{i,j} b_{ij} \mu_{ij}, \quad \text{при } i + j = k + l, \quad (4.43)$$

где a_{ij} , b_{ij} - некоторые целочисленные коэффициенты.

В полярных координатах комплексные моменты могут быть представлены в виде

$$M_{kl} = \int_0^{2\pi} \int_0^{\infty} \rho^{k+l+1} e^{i(k-l)\vartheta} x(\rho \cos \vartheta, \rho \sin \vartheta) d\rho d\vartheta, \quad (4.44)$$

$k, l = 0, 1.$

С другой стороны $M_{kl} = |M_{kl}| e^{i\psi}$, где $|M_{kl}|$, ψ_0 - модуль и аргумент комплексного числа. Путем несложных преобразований выражения (4.44) показывается, что при повороте изображения на угол $\Delta\vartheta$ значение комплексного момента примет вид

$$M_{kl}(\Delta\vartheta) = |M_{kl}| e^{i\psi_0} e^{-i\Delta\vartheta(k-l)}. \quad (4.45)$$

Из (4.45) следует [46], что при повороте изображения не меняются значения $M_{kk}(\Delta\vartheta) = M_{kk}$ и $M_{kl}(\Delta\vartheta)M_{lk}(\Delta\vartheta) = |M_{kl}|^2$.

Обобщая эти результаты, введем в рассмотрение произведение

$$\begin{aligned} & M_{k_1 l_1}^{r_1} M_{k_2 l_2}^{r_2} \dots K M_{k_q l_q}^{r_q} = \\ & = \left| M_{k_1 l_1}^{r_1} M_{k_2 l_2}^{r_2} \dots K M_{k_q l_q}^{r_q} \right| e^{i(\psi_{01} r_1 + K + \psi_{0q} r_q)} \times \\ & \times e^{-i\Delta\vartheta[(k_1 - l_1)r_1 + (k_2 - l_2)r_2 + K + (k_q - l_q)r_q]}. \end{aligned} \quad (4.46)$$

Оно является инвариантным к повороту, если выполняется условие:

$$\sum_{i=1}^q (k_i - l_i) r_i = 0. \quad (4.47)$$

При этом условии в качестве инвариантов можно использовать модуль произведения комплексных моментов $\left| M_{k_1 l_1}^{r_1} M_{k_2 l_2}^{r_2} \dots K M_{k_q l_q}^{r_q} \right|$, вещественную $\text{Re}\left(M_{k_1 l_1}^{r_1} M_{k_2 l_2}^{r_2} \dots K M_{k_q l_q}^{r_q} \right)$ и мнимую $\text{Im}\left(M_{k_1 l_1}^{r_1} M_{k_2 l_2}^{r_2} \dots K M_{k_q l_q}^{r_q} \right)$ части произведения.

Легко показать, что набор моментных инвариантов (4.41) для центрированного изображения выражается через комплексные моменты следующим образом:

$$\begin{aligned}
\Phi_1 &= M_{11}, \\
\Phi_2 &= M_{20}M_{02}, \\
\Phi_3 &= M_{21}M_{12}, \\
\Phi_4 &= M_{30}M_{03}, \\
\Phi_5 &= \operatorname{Re}\left(M_{21}^3 M_{03}\right), \\
\Phi_6 &= \operatorname{Re}\left(M_{21}^3 M_{02}\right), \\
\Phi_7 &= -\operatorname{Im}\left(M_{21}^3 M_{03}\right).
\end{aligned} \tag{4.48}$$

Отметим, что данный набор является функционально избыточным, так как инварианты $\Phi_3, \Phi_4, \Phi_5, \Phi_7$ связаны соотношением:

$$\Phi_5^2 + \Phi_7^2 = \Phi_3^3 - \Phi_4.$$

Очевидно, что по заданному набору комплексных моментов можно построить любое число инвариантов, так как любая функциональная комбинация инвариантов также является инвариантом. Однако количество функционально независимых инвариантов ограничено и зависит от числа комплексных моментов, используемых при построении.

Алгоритм синтеза набора функционально независимых инвариантов на основе множества Q комплексных моментов $M_{k_1 l_1}, M_{k_2 l_2}, \dots, M_{k_q l_q}$ состоит в следующем:

- из множества Q берутся моменты $M_{k_i l_i}$ при $k_i = l_i$, которые являются инвариантами;
- задается базовый элемент $M_{k_0 l_0}$ ($k_0 = l_0$) множества Q , и для остальных моментов $M_{k_i l_i}$ при $k_i \neq l_i$ строятся инвариантные комбинации, удовлетворяющие требованию (4.47)

$$\Phi_i = M_{k_0 l_0}^{l_i - k_i} M_{k_i l_i}^{k_0 - l_0}, \text{ при } (k_0 = l_0) \neq (k_i = l_i). \tag{4.49}$$

Очевидно, любое другое инвариантное произведение вида (4.46) при условии (4.47), построенное из моментов множества Q , представляется через инварианты (4.49).

Таким образом, число функционально независимых инвариантов, которое можно построить из q комплексных моментов, составляет:

$$I = \begin{cases} q, & \text{если } k_i = l_i \text{ при } 1 \leq i \leq q, \\ q - 1, & \text{в противном случае.} \end{cases}$$

Так, например, из треугольной матрицы комплексных моментов (4.42) при $1 \leq k + l \leq K$, включающей $(K+1)(K+2)/2$ элементов, можно сформировать (неединственным образом) $(K+1)(K+2)/2 - 1$ независимых инвариантов вида (4.46) при условии (4.47).

Для центрированного изображения комплексные моменты M_{01} , M_{10} равны нулю, вследствие чего при построении инвариантов они не используются (например, в (4.41) и (4.48)). Однако, как будет показано ниже, при обработке цифровых изображений в режиме «скользящего окна» оказывается возможным и даже желательным использование моментов M_{01} и M_{10} .

4.3.3. Особенности построения моментных инвариантов для обработки цифровых изображений в режиме скользящего окна

Для обработки цифровых изображений используются дискретные аналоги моментных характеристик (4.38), (4.39), (4.42), в которых операция интегрирования заменена на суммирование дискретных значений произведений изображения на двумерную степенную функцию. При этом моменты обычно вычисляются не по всему изображению, а только по ограниченной области - окну обработки D :

$$\mu_{kl}(n_1, n_2) = \sum_{(m_1, m_2) \in D} m_1^k m_2^l x(n_1 - m_1, n_2 - m_2). \quad (4.50)$$

Вычисление моментных инвариантов в режиме скользящего окна и их использование в качестве признаков для обнаружения и распознавания объектов имеют следующие особенности, обусловленные спецификой этого режима [7]:

- исчезает необходимость центрирования моментов (путем вычисления по дискретному аналогу формулы (4.39) или преобразованием

моментов (4.50) в центральные по формуле (4.40)), поскольку при обработке всегда найдется положение окна, для которого геометрический центр окна будет близок к центру тяжести объекта;

- возникают ситуации, когда окно захватывает только часть объекта, и эта часть постепенно увеличивается при «наполнении» окна на объект, при этом центр тяжести изображения в окне смещается от края окна к его центру.

Для точной локализации (определения координат) объекта на изображении необходимо использовать признаки, характеризующие «захват» объекта окном, например расстояние от центра тяжести изображения в окне до центра окна. Расстояние является инвариантным к повороту изображения и выражается через моменты [3,46]:

$$\Phi = \frac{\sqrt{M_{10} M_{01}}}{M_{00}} = \frac{\sqrt{\mu_{10}^2 + \mu_{01}^2}}{\mu_{00}}. \quad (4.51)$$

Вместо инварианта (4.51) в качестве признаков можно использовать инварианты $M_{10}M_{01}$, M_{00} по отдельности [3,21,46], а также другие инварианты, построенные согласно (4.46), (4.47) с использованием комплексных моментов M_{10} , M_{01} ($M_{20}M_{01}^2$, $M_{30}M_{01}^3$ и т.д.).

Таким образом, набор признаков для распознавания объектов должен состоять из инвариантов:

- построенных с использованием комплексных моментов младших порядков ($k+l \leq 1$) и обеспечивающих локализацию объекта;
- построенных с помощью комплексных моментов старших порядков ($k+l > 1$) и обеспечивающих распознавание (различение) объекта.

4.3.4. Алгоритм расчета моментных инвариантов в скользящем окне

Алгоритм вычисления заданного набора моментных инвариантов для каждого положения скользящего окна на изображении включает следующие этапы [7]:

- вычисление треугольной матрицы моментов μ_{kl} ($k+l \leq K$) до заданного порядка K ;

- вычисление линейных комбинаций моментов, задающих вещественные и мнимые части требуемых комплексных моментов;
- вычисление значений моментных инвариантов.

Рассмотрим указанные этапы и приведем выражения для вычислительной сложности реализуемых алгоритмов.

При использовании окна обработки прямоугольной формы применим параллельно-рекурсивный алгоритм расчета моментных характеристик [10-12,22] с учетом того, что формируется не прямоугольная, а треугольная матрица моментов μ_{kl} ($0 \leq k \leq K$, $0 \leq l \leq K - k$). Для вычисления треугольной матрицы моментов порядка K требуется U_a^{μ} операций сложения и U_m^{μ} операций умножения на каждый отсчет обрабатываемого изображения [7], где

$$U_a^{\mu} = \frac{(K+1)(K+3)(K+8)}{6}, \quad U_m^{\mu} = \frac{K(K+2)(K+7)}{6}. \quad (4.52)$$

На этапе расчета линейных комбинаций моментов (4.43) число операций зависит от набора комплексных моментов, необходимых при вычислении инвариантов. Для пары комплексно-сопряженных моментов достаточно определения только двух комбинаций, поскольку

$$\operatorname{Re}(M_{kl}) = \operatorname{Re}(M_{lk}), \quad \operatorname{Im}(M_{kl}) = -\operatorname{Im}(M_{lk}).$$

Следовательно, при использовании треугольной матрицы комплексных моментов M_{kl} ($0 \leq k \leq K$, $0 \leq l \leq K - k$) порядка K необходимо определить только c элементов матрицы M_{kl} ($0 \leq k \leq K$, $0 \leq l \leq \min(k, K - k)$):

$$c = \left\lceil \frac{K+2}{2} \right\rceil \left\lceil \frac{K+3}{2} \right\rceil,$$

где [...] - целая часть числа n . При этом количество арифметических операций сложения U_a и умножения U_m составляет

$$U_a^M = \frac{1}{3} \sum_{k=0}^K \left\lceil \frac{k}{2} \right\rceil \left(\left\lceil \frac{k}{2} \right\rceil + 1 \right) \left(\left\lceil \frac{k}{2} \right\rceil + 5 \right) - \frac{1}{6} \left\lceil \frac{k}{2} \right\rceil \left(\left\lceil \frac{k}{2} \right\rceil + 1 \right) \left(\left\lceil \frac{k}{2} \right\rceil + 5 \right),$$

$$U_m^M = \frac{1}{3} \sum_{k=2}^K \left(\left[\frac{k}{2} \right] + 1 \right) \left(\left[\frac{k}{2} \right] + 2 \right) \left(\left[\frac{k}{2} \right] + 3 \right) - \frac{1}{6} \left(\left[\frac{k}{2} \right] + 1 \right) \left(\left[\frac{k}{2} \right] + 2 \right) \left(\left[\frac{k}{2} \right] + 3 \right) + 1. \quad (4.53)$$

На практике количество операций умножения может быть ниже приведенной оценки за счет оптимизации вычислительного процесса (например, некоторые коэффициенты в линейных комбинациях оказываются равными единице).

При вычислении моментных инвариантов согласно (4.46), (4.47) число операций определяется количеством комплексных умножений, используемых при построении набора инвариантов. В частном случае, при использовании представленного выше алгоритма с выбором базового элемента (например, M_{01}) число операций для вычисления инвариантов составляет

$$U_a^\Phi = 2 \left[\frac{K+1}{2} \right] \left[\frac{K+2}{2} \right] + K - 1, \quad U_m^\Phi = 2U_a^\Phi. \quad (4.54)$$

Вычислительная сложность отдельных этапов алгоритма вычисления моментных инвариантов, рассчитанных по формулам (4.52)-(4.54), а также суммарные количества операций (U_a и U_m) приведены в следующей табл. 4.1.

Таблица 4.1. Вычислительная сложность алгоритма расчета моментных инвариантов

K	0	1	2	3	4	5	6	7	8
Q_0	4	12	25	44	70	104	147	200	264
U_m^μ	0	4	12	25	44	70	104	147	200
U_a^c	0	0	2	6	15	29	52	84	130
Q	0	0	5	13	27	47	77	117	172
U_a^Φ	0	2	6	12	18	26	34	44	54
Q_0	0	4	12	24	36	52	68	88	108
U_a	4	14	33	62	103	159	233	328	448
Θ_I	0	8	29	62	107	169	249	352	480

Таким образом, проведенный анализ показывает, что число выполняемых операций на всех этапах расчета инвариантов не зависит от размеров окна обработки, что делает рассмотренный алгоритм эффективным при больших размерах окна обработки.

4.3.5. Аффинные инварианты

Введенные авторами (Jan Flusser, Tomáš Suk) *аффинные моментные инварианты* [52, 53, 55] - это признаки изображения, построенные на основе степенных моментов и описывающие силуэт некоторого объекта. В соответствии со своим названием данные признаки являются инвариантными к аффинным преобразованиям изображения:

$$\begin{cases} u_1 = a_0 + a_1x_1 + a_2x_2, \\ u_2 = b_0 + b_1x_1 + b_2x_2. \end{cases}$$

Авторы предложили рассматривать аффинное преобразование как декомпозицию следующих шести элементарных трансформаций:

$$\begin{array}{lll} \text{а)} \begin{cases} u_1 = x_1 + \beta_1 \\ u_2 = x_2 \end{cases} & \text{б)} \begin{cases} u_1 = x_1 \\ u_2 = x_2 + \beta_2 \end{cases} & \text{в)} \begin{cases} u_1 = \alpha x_1 \\ u_2 = \alpha x_2 \end{cases} \\ \text{г)} \begin{cases} u_1 = \delta x_1 \\ u_2 = x_2 \end{cases} & \text{д)} \begin{cases} u_1 = x_1 + t_2 x_2 \\ u_2 = x_2 \end{cases} & \text{е)} \begin{cases} u_1 = x_1 \\ u_2 = t_1 x_1 + x_2. \end{cases} \end{array}$$

Трансформации (а) и (б) представляют преобразование изображения типа сдвига, (в) - пропорциональное масштабирование. Преобразование (г) задает искажение масштаба вдоль одной из осей координат. Последние два преобразования характеризуют поворот и деформацию изображения, не описываемые преобразованием подобия.

Трансформации, связанные с изменением положения объекта, легко могут быть устранены при использовании центрированных моментов (4.39) или, что эквивалентно, введением нового начала координат. Поэтому далее полагается, что вычисляемые моменты являются центрированными.

Рассмотрим, как последующие преобразования изменяют значения степенных (центрированных) моментов. Пропорциональное масштабирование изображения

$$u_1 = \alpha x_1, \quad u_2 = \alpha x_2 \quad (4.55)$$

изменяет значение степенного момента $\mu_{k\ell}$ следующим образом:

$$\begin{aligned} \mu_{k1}^\alpha &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} u_1^k u_2^1 f(u_1, u_2) du_1 du_2 = \\ &= \alpha^{k+1+2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1^k x_2^1 f(\alpha x_1, \alpha x_2) dx_1 dx_2 = \alpha^{k+1+2} \mu_{k1}. \end{aligned}$$

Из полученного соотношения следует, что величина

$$\frac{\mu_{k1}}{\left(\mu_{00}\right)^{\frac{k+1}{2}+1}} \quad (4.56)$$

является инвариантом относительно преобразования (4.55). В более общем случае функционал

$$\Phi = \sum_{i \in I} \left[r_i \frac{\mu_{k_1(i)l_1(i)} \prod \mu_{k_{c(i)}l_{c(i)}}}{\mu_{00}^{z(i)}} \right], \quad (4.57)$$

где r_i - некоторые вещественные коэффициенты, является инвариантом к этому виду преобразований, если выполняются равенства:

$$z(i) = \frac{\sum_{j=1}^{c(i)} (k_j(i) + 1) l_j(i)}{2} + c(i). \quad (4.58)$$

Искажение масштаба вдоль оси x_1 , задаваемое преобразованием

$$u_1 = \delta x_1, \quad u_2 = x_2, \quad (4.59)$$

приводит к подобным изменениям значений степенных моментов:

$$\mu_{kl}^\delta = \delta^{p+1} \cdot \mu_{kl}.$$

Подставив данное равенство в соотношение (4.57), получим выражение для функционала Φ в виде

$$\Phi = \sum_{i \in I} \left[r_i \frac{\mu_{k_1(i) 1_1(i) \dots k_{c(i)}(i) 1_{c(i)}(i)} \cdot \delta^{\sum_{j=1}^{c(i)} k_j(i) + c(i)}}{\mu_{00}^{z(i)}} \cdot \frac{\delta^{z(i)}}{\delta^{z(i)}} \right].$$

Очевидно, для инвариантности к преобразованию (4.59) достаточно выполнения соотношений:

$$\forall i \in I: \sum_{j=1}^{c(i)} k_j(i) + c(i) = z(i).$$

Учитывая выражение (4.58), получим следующее условие, при выполнении которого обеспечивается инвариантность к искажению масштаба вдоль одной из осей координат:

$$\forall i \in I: \sum_{j=1}^{c(i)} k_j(i) = \sum_{j=1}^{c(i)} \ell_j(i).$$

Преобразования (д) и (е) однотипны. Поэтому рассмотрим способ построения инвариантов к первому из них.

Из теории алгебраических инвариантов [55] известно, что условием инвариантности функционала Φ к преобразованию (д) является выполнение равенства:

$$\sum_k \sum_l k \mu_{k-1, l+1} \frac{\partial \Phi}{\partial \mu_{kl}} = 0. \quad (4.60)$$

Подставляя выражение (4.57) для функционала Φ в уравнение (4.60) можно найти выражение для коэффициентов r_i ($i \in I$) при заданном конкретном множестве слагаемых I .

В заключение приведем шесть аффинных инвариантов, полученных авторами подхода [52].

$$I_1 = \frac{1}{\mu_{00}^4} (\mu_{20}\mu_{02} - \mu_{11}^2),$$

$$I_2 = \frac{1}{\mu_{00}^{10}} (\mu_{30}^2 \mu_{03}^2 - 6\mu_{30}\mu_{21}\mu_{12}\mu_{03} + 4\mu_{30}\mu_{12}^3 + 4\mu_{03}\mu_{21}^3 - 3\mu_{21}^2 \mu_{12}^2),$$

$$I_3 = \frac{1}{\mu_{00}^7} (\mu_{20} (\mu_{21}\mu_{03} - \mu_{12}^2) - \mu_{11} (\mu_{30}\mu_{03} - \mu_{21}\mu_{12}) + \mu_{02} (\mu_{30}\mu_{12} - \mu_{21}^2)),$$

$$I_4 = \frac{1}{\mu_{00}^{11}} (\mu_{20}^3 \mu_{03}^2 - 6\mu_{20}^2 \mu_{11} \mu_{12} \mu_{03} - 6\mu_{20}^2 \mu_{02} \mu_{21} \mu_{03} + 9\mu_{20}^2 \mu_{02} \mu_{12}^2 + \\ + 12\mu_{20} \mu_{11}^2 \mu_{21} \mu_{03} + 6\mu_{20} \mu_{11} \mu_{02} \mu_{30} \mu_{03} - 18\mu_{20} \mu_{11} \mu_{02} \mu_{21} \mu_{12} - \\ - 8\mu_{11}^3 \mu_{30} \mu_{03} - 6\mu_{20} \mu_{02}^2 \mu_{30} \mu_{12} + 9\mu_{20} \mu_{02}^2 \mu_{21}^2 + \\ + 12\mu_{11}^2 \mu_{02} \mu_{30} \mu_{12} - 6\mu_{11} \mu_{02}^2 \mu_{30} \mu_{21} + \mu_{02}^3 \mu_{30}^2),$$

$$I_5 = \frac{1}{\mu_{00}^6} (\mu_{40}\mu_{04} - 4\mu_{31}\mu_{13} + 3\mu_{22}^2),$$

$$I_6 = \frac{1}{\mu_{00}^9} (\mu_{40}\mu_{04}\mu_{22} + 2\mu_{31}\mu_{22}\mu_{13} - \mu_{40}\mu_{13}^2 - \mu_{04}\mu_{22}^3).$$

5. ОБНАРУЖЕНИЕ И ЛОКАЛИЗАЦИЯ ОБЪЕКТОВ НА ИЗОБРАЖЕНИИ

5.1 Постановка задачи и анализ современного состояния

Обнаружение и локализация объектов на изображении – один из основных этапов при решении задачи поиска и распознавания объектов на изображениях. Основными причинами этого являются следующие:

- любую многоальтернативную задачу классификации можно заменить несколькими простыми задачами двухальтернативного распознавания - обнаружения;
- при классификации на изображении решающее правило работает в условиях с пространственной протяженностью обрабатываемых данных, что приводит к необходимости согласования критериев и/или алгоритмов обнаружения и локализации для качественного решения задачи [79];
- высокая вычислительная сложность «хорошего» алгоритма распознавания и «хороших» признаков не позволяет практически использовать их при поиске объекта по всему изображению, что приводит к необходимости применения достаточно простых решающих правил предварительной обработки, локализующей области возможного интереса, где может появиться соответствующий объект.

В настоящее время известно огромное число алгоритмов построения классификаторов для решения задачи обнаружения [13, 14, 17, 33, 38, 39, 40]. Каждый из них имеет свои недостатки и свои преимущества. В настоящей работе рассмотрена линейная модель. Это обусловлено следующими причинами:

- простотой применения (реализации) классификатора;
- возможностью приведения классификаторов с полиномиальной дискриминантной функцией к линейному виду путем повышения размерности пространства признаков;

- существованием для решающего правила явного выражения вычислительной сложности, которая, к тому же, является самой низкой среди всех групп классификаторов (с тем же числом признаков);
- возможностью проведения аналитических расчетов.

Заметим, что при поиске объекта на изображении задача не исчерпывается построением классификатора. Дополнительно к классификации возникает необходимость *локализации объекта* – то есть определения координат его местоположения. Это обусловлено следующими причинами:

- область, выделенная классификатором-обнаружителем и относящаяся к одному объекту, как правило, является размытой и не дает конкретных координат объекта;
- возможно существование несвязных областей, относящихся к одному и тому же объекту, что может привести к появлению ложного объекта;
- области, соответствующие различным объектам одного и того же класса могут быть связными - это может привести к потере объекта;
- области, относящиеся к объектам разных классов, могут находиться вблизи друг друга, что может привести к потере объекта или его неверной классификации.

Локализация позволяет обойти обозначенные недостатки процесса обнаружения и распознавания. Однако для локализации должен быть выбран критерий, по которому одно из возможных положений считается предпочтительнее другого положения. Хотя количество работ в области обнаружения и распознавания достаточно велико, вопросу локализации объекта на изображении практически уделяется недостаточно внимания, хотя одно из основных свойств изображения состоит в передаче человеку информации о взаимном расположении объектов. Именно конкретные координаты играют для многих реальных задач ключевую роль. Так, например, все алгоритмы, использующие описание одной и той же реальной картины (сцены) в виде нескольких изображений, применяют так называемую процедуру *согласования изображений*. Похожие процедуры используются и в задачах калибровки

видеокамер, задачах стереозрения, построения рельефа местности, привязки изображений к реальным или существующим картам и т.д. [26, 44, 47, 85]. Все они, по существу, сводятся к вычислению корреляции изображения заданного объекта с наблюдаемым изображением. Небольшие вариации в подходах и решениях этой сложной задачи, предпринимаемые с целью увеличения качества обнаружения и локализации, связаны, как правило, с эвристическими соображениями. Это может быть и квантование сигнала, и специальное препарирование и т.д. Подобные пробные подходы к локализации объекта все чаще используются в связи с переходом от полутоновых к цветным изображениям - локализацию предлагается проводить только по характерному цвету искомого объекта [62, 68]. При этом подавляющее большинство алгоритмов определения местоположения не учитывает, что одновременно с локализацией происходит и обнаружение объекта.

Комплексный подход к рассмотрению задачи обнаружения и локализации предложен в работах [43, 44]. Он определяет качество измерения координат объектов ошибками двух видов. Ошибки первого рода возникают вследствие неверных отождествлений искомого объекта с отдельными деталями на наблюдаемом изображении. Они дают большие отклонения результата измерения координат от истинного значения, превышающие некоторые наперед заданные предельные размеры (или размеры объекта). Такие ошибки называют *аномальными*. Аномальные ошибки характеризуются вероятностями ложного обнаружения и пропуска объекта. Ошибки второго рода или *нормальные* ошибки имеют величину меньше некоторого наперед заданного значения (или меньше линейных размеров объекта). Можно считать, что нормальные ошибки характеризуются своими среднеквадратичными значениями отклонения в определении координат. Для построения локализатора с минимальным числом аномальных ошибок используется подход [41], принятый в обработке изображений (использование частотной характеристики фильтра, гистограммы и спектра изображения и т.д.). Однако этот подход невозможно применить для задачи локали-

зации, где в качестве признаков используются не отсчеты входного изображения, и классификатор, в свою очередь, не является линейным.

Все сказанное выше приводит к выводу об актуальности проблемы разработки и реализации критериев и алгоритмов локализации, согласованных с показателями качества процесса обнаружения (нормальные ошибки, таким образом, в данном случае не рассматриваются).

5.2. Критерии локализации объектов

5.2.1. Локализация, согласованная с критерием качества обнаружения и распознавания

Качество процедуры обнаружения и распознавания может быть охарактеризовано величиной (1.5) общего риска R . При этом, оптимальный в смысле минимума общего риска классификатор - байесовский - строится таким образом, чтобы при решении о каждом конкретном векторе y выбирать тот класс Ω_l ($l = \overline{0, L-1}$), для которого значение условного среднего риска (1.3) по этому классу:

$$R_l(y) = \sum_{i=0}^{L-1} r_i P(\Omega_i/y), \quad l = \overline{0, L-1}.$$

минимально [13, 17, 33, 38, 40]. Здесь $P(\Omega_i/y)$ - апостериорная вероятность отнесения данного конкретного вектора y к классу Ω_i . Для определенности в дальнейшем будем интерпретировать класс Ω_0 как класс, соответствующий фоновой составляющей изображения.

Поскольку процедуру локализации объектов на изображении можно рассматривать как частный случай процедуры обнаружения или распознавания, то представляется целесообразным производить локализацию объекта по тому же критерию, по которому производится и классификация, то есть по минимуму общего риска. Пусть n_1 и n_2 - отсчеты в плоскости изображения. Пусть область локализации (то есть область предполагаемого размещения объекта j -ого класса на изобра-

жении) обозначена \tilde{D}_j . Пусть $\mathbf{y} = \mathbf{y}(n_1, n_2)$ - вектор признаков для фрагмента изображения с координатами n_1 и n_2 . Тогда критерий локализации выглядит следующим образом:

$$(n_1, n_2) : R_j(\mathbf{y}(n_1, n_2)) = \min_{(m_1, m_2) \in \tilde{D}_j^0} R_j(\mathbf{y}(n_1 - m_1, n_2 - m_2)). \quad (5.1)$$

В силу реализационных ограничений области \tilde{D}_j , как правило, принимаются одинаковыми и равными некоторой наперед заданной области \tilde{D} , размеры которой согласуют с размером искомого объекта. Тогда существует потенциальная возможность упрощения процесса обнаружения и локализации, если классификация также производится на основе анализа значения среднего риска. Используем тот факт, что выбор класса l для анализируемого фрагмента происходит по критерию (2.1), то есть:

$$l_{(n_1, n_2)} : R_l(\mathbf{y}(n_1, n_2)) = \min_{j=0, L-1} R_j(\mathbf{y}(n_1, n_2)). \quad (5.2)$$

В этом случае общий критерий обнаружения (распознавания) и локализации запишется в следующем виде:

$$(n_1, n_2) : R_l(\mathbf{y}(n_1, n_2)) = \min_{i=0, K-1} R_j(\mathbf{y}(n_1 - m_1, n_2 - m_2)). \quad (5.3)$$

$$(m_1, m_2) \in \tilde{D}^0$$

Процесс обнаружения и локализации, основанный на приведенном критерии реализуется следующим образом: область \tilde{D} (окно локализации) занимает все возможные положения на плоскости изображения, и если при очередном положении (n_1, n_2) координаты-аргументы оптимума критерия (5.3) совпадают с текущими, а соответствующий номер класса отличен от нуля (не фоновый класс), то текущее положение рассматривается как искомое положение объекта соответствующего класса.

В реализационном плане использование выражения (5.3) напрямую оказывается не всегда целесообразно, поскольку вычислять значения рисков R_j неудобно. Более того, известные правила байесовской классификации [33,40] существенно проще (5.2). Возникает вопрос, нельзя ли упростить и правило локализации. Покажем, что в общем случае сделать это невозможно.

Пусть сравниваются два значения риска для двух различных местоположений предполагаемого объекта j -го и l -го классов с координатами (n_1, n_2) и $(\tilde{n}_1, \tilde{n}_2)$:

$$R_j(\mathbf{y}(n_1, n_2)) = \frac{\sum_{i=0}^{K-1} r_{ji} P(\Omega_i) p(\mathbf{y}(n_1, n_2)/\Omega_i)}{p(\mathbf{y}(n_1, n_2))}, \quad j = \overline{0, L-1},$$

$$R_l(\mathbf{y}(\tilde{n}_1, \tilde{n}_2)) = \frac{\sum_{i=0}^{K-1} r_{li} P(\Omega_i) p(\mathbf{y}(\tilde{n}_1, \tilde{n}_2)/\Omega_i)}{p(\mathbf{y}(\tilde{n}_1, \tilde{n}_2))}, \quad l = \overline{0, L-1}.$$

Здесь $p\left(\frac{\mathbf{y}}{\Omega_i}\right)$ - условная плотность вероятностей случайного векторного признака \mathbf{Y} в классе Ω_i , $p(\mathbf{y})$ - безусловная плотность вероятностей случайного вектора \mathbf{Y} . В случае, если $(n_1, n_2) = (\tilde{n}_1, \tilde{n}_2)$, значения безусловной плотности вероятности в этих двух формулах одинаковы. Именно это обстоятельство позволяет производить дальнейшие упрощения байесовского правила классификации при конкретном векторе признаков. Но при локализации координаты различны, следовательно, различны векторы и, наконец, значения безусловной плотности вероятности. Значит, в общем случае локализация по минимуму риска должна производиться по самим значениям риска. Однако для задачи обнаружения объектов можно получить существенные упрощения.

5.2.2. Локализация при обнаружении объекта

Пусть класс объектов - Ω_1 , класс фона (шума) - Ω_0 . Тогда значение риска для вектора \mathbf{y} имеет вид:

$$R_l(\mathbf{y}(n_1, n_2)) = \frac{r_{l0}P(\Omega_0)p(\mathbf{y}(n_1, n_2)/\Omega_0) + r_{l1}P(\Omega_1)p(\mathbf{y}(n_1, n_2)/\Omega_1)}{P(\Omega_0)p(\mathbf{y}(n_1, n_2)/\Omega_0) + P(\Omega_1)p(\mathbf{y}(n_1, n_2)/\Omega_1)},$$

$$l = \{0, 1\}.$$

Его можно переписать следующим образом:

$$R_l(\mathbf{y}(n_1, n_2)) = \frac{r_{l0} + r_{l1} \cdot \frac{P(\Omega_1)p(\mathbf{y}(n_1, n_2)/\Omega_1)}{P(\Omega_0)p(\mathbf{y}(n_1, n_2)/\Omega_0)}}{1 + \frac{P(\Omega_1)p(\mathbf{y}(n_1, n_2)/\Omega_1)}{P(\Omega_0)p(\mathbf{y}(n_1, n_2)/\Omega_0)}}, \quad l = \{0, 1\}.$$

Введем отношение правдоподобия:

$$\Lambda(n_1, n_2) = \frac{p(\mathbf{y}(n_1, n_2)/\Omega_1)}{p(\mathbf{y}(n_1, n_2)/\Omega_0)}. \quad (5.4)$$

Тогда риск за принятие решения об отнесении текущего фрагмента к каждому из классов будет равен:

$$R_l(n_1, n_2) = \frac{r_{l0} + r_{l1} \cdot \frac{P(\Omega_1)}{P(\Omega_0)} \Lambda(n_1, n_2)}{1 + \frac{P(\Omega_1)}{P(\Omega_0)} \Lambda(n_1, n_2)}, \quad l = \{0, 1\}.$$

Учитывая, что за неверные решения штраф больше, чем за верные, график зависимости рисков от величины $\Lambda(n_1, n_2)$ примет вид, приведенный на рис. 5.1.

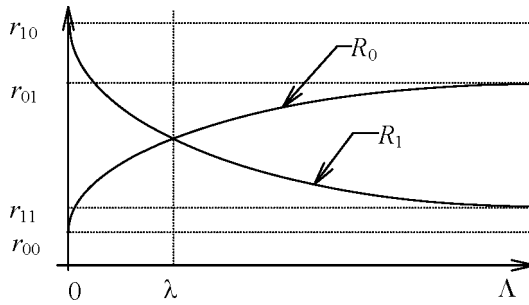


Рис. 5.1. Изменение риска в задаче обнаружения

Классификация производится путем сравнения величины Λ с порогом:

$$\lambda = \frac{r_{10} - r_{00}}{r_{01} - r_{11}} \cdot \frac{P(\Omega_0)}{P(\Omega_1)}.$$

При $\Lambda > \lambda$ принимается решение о наличии объекта и производится его локализация. Очевидно, что значение риска R тем меньше, чем больше отношение правдоподобия (5.4). Поэтому критерий локализации (5.1) может быть приведен к виду

$$(n_1, n_2) : \Lambda(\mathbf{y}(n_1, n_2)) = \max_{(m_1, m_2) \in \mathcal{B}^0} \Lambda(\mathbf{y}(n_1 - m_1, n_2 - m_2)). \quad (5.5)$$

Полученный результат тем более важен, поскольку выражение (5.4) используется для построения байесовского классификатора в виде (см. п.2.1.1):

$$\Lambda(\mathbf{y}(n_1, n_2)) \begin{matrix} > \\ < \end{matrix} \lambda \Rightarrow \begin{cases} \Omega_1 \\ \Omega_0 \end{cases}.$$

А значит, критерий обнаружения и локализации будет единым:

$$\begin{aligned} &(n_1, n_2) : \Lambda(\mathbf{y}(n_1, n_2)) = \\ &= \max_{\{(m_1, m_2) \in D^L \ \& \ \Lambda(\mathbf{y}(n_1 - m_1, n_2 - m_2)) > \lambda\}} \Lambda(\mathbf{y}(n_1 - m_1, n_2 - m_2)), \end{aligned} \quad (5.6)$$

здесь «&» – операция логической конъюнкции (логическое «и»).

Процесс обнаружения и локализации, основанный на приведенном критерии реализуется следующим образом: область \tilde{D} занимает все возможные положения на плоскости изображения, и если при очередном положении (n_1, n_2) координаты-аргументы оптимума критерия (5.6) совпадают с текущими, то оно и рассматривается как искомое положение объекта.

5.2.3. Локализация объекта по максимуму правдоподобия

Пусть функция правдоподобия фонового класса такова, что для всех отличающихся от нуля значений функции правдоподобия класса объектов она приблизительно постоянна, то есть:

$$p(\mathbf{y}/\Omega_0) \approx \text{const} \quad \text{для} \quad \{\mathbf{y} : p(\mathbf{y}/\Omega_1) > \varepsilon\}_{\varepsilon \rightarrow 0}. \quad (5.7)$$

Тогда критерий локализации (5.5) преобразуется к следующему виду:

$$(n_1, n_2) : p\left(\frac{\mathbf{y}(n_1, n_2)}{\Omega_1}\right) = \max_{(m_1, m_2) \in D^L} p\left(\frac{\mathbf{y}(n_1 - m_1, n_2 - m_2)}{\Omega_1}\right). \quad (5.8)$$

Выражение (5.8) означает, что локализация производится по максимуму правдоподобия класса объектов. Гипотезу (5.7) целесообразно использовать в случае, если разброс значений векторов признаков в классах объектов и фона существенно различается. С точки зрения статистических свойств вектора признаков это означает существенное различие в матрицах ковариации векторов \mathbf{Y} в каждом из классов:

$$|\lambda_{\min}(\mathbf{B}_0)| \gg |\lambda_{\max}(\mathbf{B}_1)|,$$

где

$$\mathbf{B}_l = E\left\{\left(\mathbf{Y} - E\{\mathbf{Y}\}\right)\left(\mathbf{Y} - E\{\mathbf{Y}\}\right)^T / \Omega_l\right\}$$

или, в одномерном случае, существенное различие в дисперсиях случайных величин:

$$\sigma_0^2 \gg \sigma_1^2,$$

где

$$\sigma_i^2 = E \left\{ (Y - E\{Y\})^2 / \Omega_i \right\}.$$

Здесь $\lambda_{\min}(\cdot)$ и $\lambda_{\max}(\cdot)$ - минимальное и максимальное собственные числа соответствующих ковариационных матриц. Для одномерного случая выполняется: $\lambda_{\max} = \lambda_{\min} = \sigma$. Данный критерий целесообразно использовать также в ситуации, если функция правдоподобия фонового класса неизвестна или предполагается произвольной.

Локализация по максимуму функции правдоподобия имеет понятный смысл: производится выбор того положения фрагмента, для которого предположение о схожести на объект (его описание в признаковом пространстве) наиболее правдоподобно.

Процесс локализации идентичен приведенному ранее.

5.2.4. Локализация объекта по минимуму значения альтернативной функции правдоподобия

Пусть функция правдоподобия класса объектов такова, что для всех отличающихся от нуля значений функции правдоподобия фонового класса она приблизительно постоянна, то есть:

$$p(\mathbf{y}/\Omega_1) \approx \text{const} \quad \forall \mathbf{y} \in \{ \mathbf{y} : p(\mathbf{y}/\Omega_0) > \varepsilon \}_{\varepsilon \rightarrow 0}. \quad (5.9)$$

Тогда критерий локализации (5.5) преобразуется к следующему виду:

$$(n_1, n_2) : p(\mathbf{y}(n_1, n_2)/\Omega_0) = \min_{(m_1, m_2) \in \tilde{D}} p(\mathbf{y}(n_1 - m_1, n_2 - m_2)/\Omega_0),$$

что означает, что локализация производится по минимальному значению функции правдоподобия фонового класса, то есть альтернативной к функции правдоподобия класса объектов. Аналогично предшествующему случаю, гипотезу (5.9) целесообразно использовать в случае, если разброс значений векторов признаков в классах объектов и фона существенно различается. С точки зрения статистических свойств вектора признаков это означает существенное различие в матрицах ковариации:

$$|\lambda_{\min}(\mathbf{B}_1)| ? |\lambda_{\max}(\mathbf{B}_0)|$$

или, в одномерном случае,

$$\sigma_1^2 ? \sigma_0^2.$$

Приведенный критерий целесообразно использовать также в ситуации, если функция правдоподобия класса объектов неизвестна или предполагается произвольной. Такое часто происходит, когда производится поиск на изображении каких-либо локальных аномалий, характеристики которых, разумеется, заранее неизвестны. При этом, как правило, большая часть изображения (фон) заполнена некоторой однородной структурой (например, текстура) с совершенно понятными для идентификации характеристиками. В такой ситуации действительно целесообразно выбирать как аномальные именно те точки на изображении, которые плохо согласуются с описанием фона. А поскольку фон занимает основное пространство изображения, следовательно, вероятности появления аномальных фрагментов малы, то есть малы и значения функции правдоподобия класса фона.

Локализация по минимуму альтернативной функции правдоподобия имеет понятный смысл: производится выбор того положения фрагмента, для которого предположение о похожести на «необъект» (описание в признаковом пространстве) наименее правдоподобно (альтернативная функция правдоподобия имеет наименьшее значение). Процесс локализации по данному критерию аналогичен изложенному ранее.

Предложенный здесь критерий локализации объектов достаточно интенсивно используется, но без постановки задачи именно таким образом. Примером может служить работы [33,77,83] по распознаванию отпечатков пальцев. В качестве признаков для распознавания принимались те фрагменты изображения отпечатка, которые выглядели аномальными по сравнению с общей структурой. А именно, предполагалось, что изображение при локальном рассмотрении состоит из параллельных и достаточно регулярных перепадов функции яркости, таких как на рис. 5.2а. Все, что значительно отличается от подобной модели принимается за аномальный фрагмент. Примеры аномалий приведены на

рис. 5.2б-г. Именно величина отличия, с точки зрения теории распознавания, может рассматриваться как значение признака, по которому определяется значение функции правдоподобия. Очевидно, что для фонового класса более правдоподобными оказываются более низкие значения признака (слабое отличие), а менее правдоподобными - более высокие, как изображено на рис. 5.3. В результате локализация производится по максимуму величины отличия. Пример обнаружения и локализации аномальных фрагментов приведен на рис. 5.4. В качестве значения, характеризующего отличие фрагмента изображения от изображения регулярной структуры, была принята ошибка аппроксимации фрагмента двумерной цилиндрической полиномиальной функцией.

В заключение отметим, что представленные здесь критерии могут быть использованы безотносительно к тому, какой именно вид области использует процедура локализации. Однако можно предположить, что если все же адаптироваться на вид конкретной области локализации, то результаты могут быть лучше. Более того, не всегда построение классификатора связано с построением функции правдоподобия. И тогда оказывается невозможным воспользоваться данными критериями. Поэтому целесообразно произвести рассмотрение возможности построения совместного алгоритма обнаружения-локализации, который одновременно (а не последовательно, как это предполагалось ранее) выполнял бы и обнаружение, и локализацию по выбранному критерию качества.

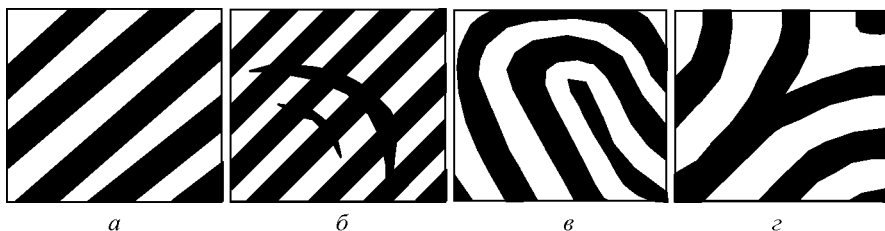


Рис. 5.2. Схематичные изображения отдельных фрагментов дактилограмм:
 фрагмент изображения с регулярной структурой (а);
 фрагмент изображения с искусственным нарушением регулярной структуры (б);
 фрагмент изображения с естественным нарушением регулярной структуры (в);
 фрагмент изображения с естественным нарушением регулярной структуры (г)

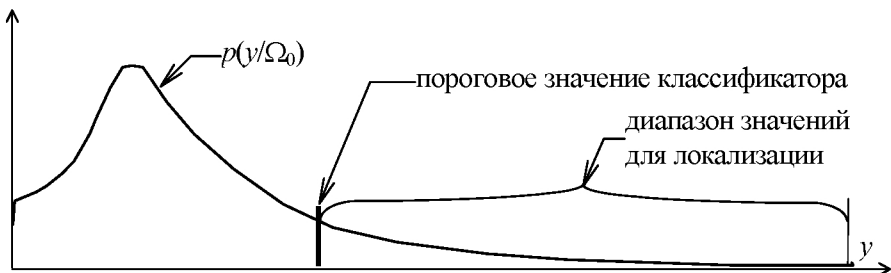


Рис. 5.3. Функция правдоподобия фонового класса в задаче обнаружения и локализации особенностей на дактилограмме

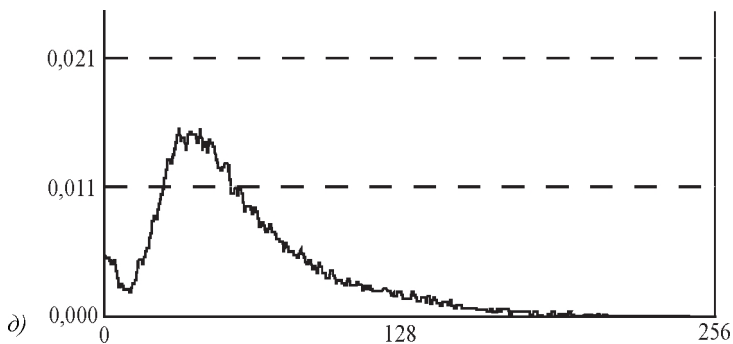
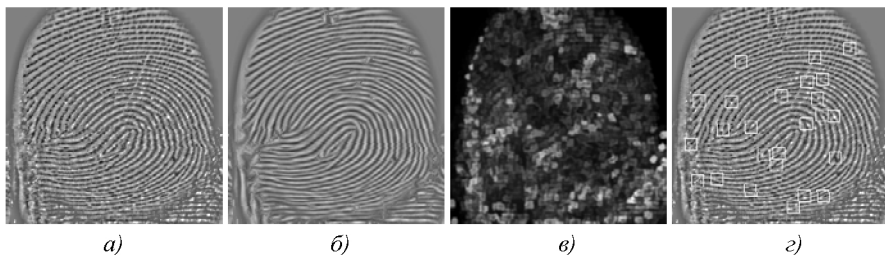


Рис. 5.4. Процесс обнаружения и локализации аномальных фрагментов на дактилограмме: исходная дактилограмма (а); изображение – аппроксимация (б); поле локальных ошибок аппроксимации (в); результаты обнаружения и локализации (г); функция правдоподобия (эмпирическая) для класса регулярной структуры изображения (фоновый класс)

5.3. Совместное обнаружение и локализация объектов

5.3.1. Общее описание процедуры локализации

Вначале определим понятие «локализатора-обнаружителя» объектов. Очевидно, что он должен наследовать свойства классификатора для обнаружения объектов, а следовательно, содержать некоторую функцию (подобно дискриминантной функции) $d(\cdot)$, задаваемую на пространстве признаков и имеющую некоторое заранее выбранное пороговое значение, на основе которого определяется, нужно ли отнести данную точку на изображении к области локализации. Без ограничения общности можно считать порог равным нулю. В случае, если пороговое значение не равно нулю, всегда можно сделать линейное преобразование функции $d(\cdot)$ таким образом, чтобы пороговое значение обратилось в ноль. С помощью функции $d(\cdot)$ для каждого положения (n_1, n_2) окна обработки производится вычисление одного отсчета изображения дискриминантной функции $\eta(n_1, n_2)$. Далее, в отличие от обычного классификатора, «локализатор-обнаружитель» должен содержать некоторое правило, по которому среди нескольких точек области локализации выбирается какая-либо одна. Для этого целесообразно использовать известный в обработке изображений *пиковый фильтр* [29]. Алгоритм его работы следующий:

$$g(n_1, n_2) = \begin{cases} \eta(n_1, n_2), & \text{если } \eta(n_1, n_2) = \max_{(m_1, m_2) \in B^0} \eta(n_1 - m_1, n_2 - m_2), \\ \overline{\min}, & \text{если } \eta(n_1, n_2) \neq \max_{(m_1, m_2) \in B^0} \eta(n_1 - m_1, n_2 - m_2). \end{cases}$$

Величина B^0 - заранее выбранная область локализации, $\overline{\min}$ - некоторое постоянное значение, не превышающее ни одно из значений яркости изображения $\eta(n_1, n_2)$, $g(n_1, n_2)$ - изображение, полученное в ре-

зультате пиковой фильтрации изображения $\eta(n_1, n_2)$. Как правило, область локализации выбирается симметричной относительно нулевого отсчета, который и принимается за выходной отсчет фильтра. А именно:

$$B^0 = [-M_1^0, M_1^0] \times [-M_2^0, M_2^0].$$

Здесь $(2M_1^0 + 1)$, $(2M_2^0 + 1)$ - размеры пикового фильтра по вертикали и горизонтали, соответственно.

Принимая описанную модель локализатора-обнаружителя, схема процесса обнаружения и локализации становится такой, как изображена на рис.5.5. Существенным отличием работы локализатора-обнаружителя от процесса локализации, описанного в п.5.2, является то, что порядок следования локализации и классификации (процедур пиковой фильтрации и порогового отбора, соответственно) в данной ситуации не влияет на результат работы. Это связано с тем, что одни и те же значения изображения $\eta(n_1, n_2)$ используются и для локализации, и для порогового отбора. А значит, при реализации алгоритма возможно построение процедуры, которая осуществляет локализацию и обнаружение за один проход по изображению: выделенное процедурой локализации экстремальное значение $g(n_1, n_2)$ просто сравнивается с некоторым порогом.

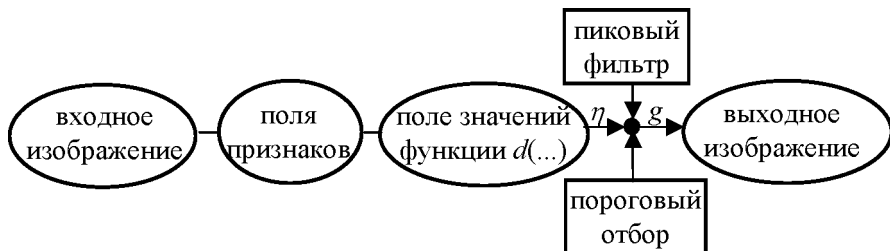


Рис.5.5. Схема процесса обнаружения и локализации

С целью упрощения дальнейшего изложения примем следующие соглашения. Считаем, что отсчеты в окне пикового фильтра упорядочены по мере удаления от выбранного выходного отсчета фильтра

и их нумерация производится в диапазоне $[0, N-1]$, где $N = |\tilde{D}|$ - число отсчетов в области локализации. Для определения аномальных ошибок процедуры совместного обнаружения и локализации (вероятностей ложной тревоги и пропуска объекта) зададим событие Q , которое состоит в принятии процедурой решения о том, что некоторый (случайный) фрагмент изображения содержит объект. В соответствии с выбранной схемой обнаружения и локализации это событие запишется следующим образом:

$$Q = \{\eta_0(\mathbf{Y}) \geq 0\} \cdot \{\eta_0(\mathbf{Y}) \geq \eta_1(\mathbf{Y})\} \cdot K \cdot \{\eta_0(\mathbf{Y}) \geq \eta_{N-1}(\mathbf{Y})\}.$$

Здесь первое из событий $\{\eta_0(\mathbf{Y}) \geq 0\}$ отвечает за процедуру обнаружения, а все остальные события - за пиковую фильтрацию. Тогда вероятности ошибок классификации для процедуры совместного обнаружения и локализации запишутся в обычном виде:

$$p_0 = P(Q/\Omega_0), \quad p_1 = P(\bar{Q}/\Omega_1). \quad (5.10)$$

Критерием качества процесса «обнаружения-локализации» является критерий минимума суммарной ошибки обнаружения в виде

$$R = r_0 p_0 + r_1 p_1 \rightarrow \min_{d(\dots)}, \quad (5.11)$$

где в качестве параметров оптимизации выступают параметры функции $d(K)$. При таком критерии формируемая процедура обнаружения и локализации будет оптимальна только с точки зрения величины аномальных ошибок: объект найден или не найден. Следовательно, возле конкретного местоположения каждого из объектов однозначно должна быть задана некоторая область, для которой принимается допущение, что для любой ее точки объект присутствует. Вне такой области объект считается отсутствующим. Размеры области целесообразно задавать, учитывая размер используемого при локализации пикового фильтра. А именно, они должны быть такими, чтобы в пределах области каждого конкретного объекта процедура совместного обнаружения и локализации выделяла не более одного объекта. Такое возможно,

только если эта область имеет линейные размеры не больше половины соответствующих линейных размеров пикового фильтра (области локализации), как показано на рис. 5.6.

Один из приводимых ниже алгоритмов построения «локализатора-обнаружителя» основывается на моделях функции правдоподобия в признаковом пространстве, а также на предположении линейности функции $d(\cdot)$ и фиксированном окне локализации. Процесс построения функции $d(\cdot)$ в такой ситуации возможен в предположении независимости значений отсчетов изображения дискриминантной функции в окне локализации. Другой алгоритм не использует никаких априорных предположений, но требует для настройки «локализатора-обнаружителя» (то есть нахождения параметров функции $d(\cdot)$ и размера окна пикового фильтра) исходные данные в виде обучающей выборки. Этот алгоритм является итерационным.

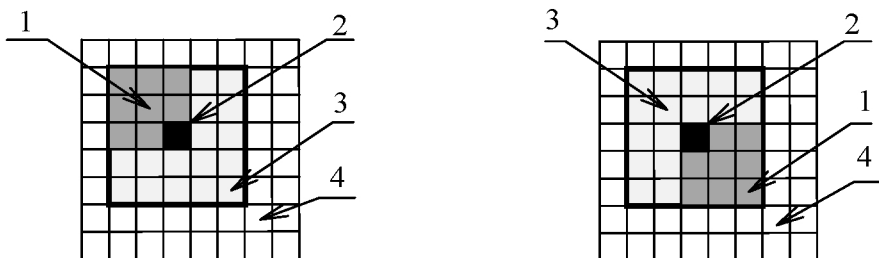


Рис. 5.6. Обнаружение и локализация объекта: область расположения объекта на изображении (1), область локализации (3), выходной отсчет пикового фильтра (2), плоскость изображения (4)

5.3.2 Настройка процедуры обнаружения и локализации при независимых значениях отсчетов изображения дискриминантной функции в пределах окна локализации

Пусть плотность вероятности отсчетов изображения в окне пикового фильтра задается следующим образом:

$$p(\eta_0, \eta_l, K, \eta_{N-1}/\Omega_l), l = \{0, 1\}. \quad (5.12)$$

Тогда значение вероятностей классификации (5.10) равно:

$$P(Q/\Omega_k) = \int_0^{+\infty} \int_{-\infty}^{\eta_0} K \int_{-\infty}^{\eta_0} p(\eta_0, \eta_l, K, \eta_{N-1}/\Omega_l) d\eta_{N-1} K d\eta_l d\eta_0, \\ l = \{0, 1\}. \quad (5.13)$$

Предположим независимость значений пикселей изображения дискриминантной функции в окне пикового фильтра. Заметим, что такое предположение, как правило, не отражает реальную ситуацию. Поскольку значения изображения функции в окне локализации рассчитываются по признакам сильно перекрывающихся фрагментов изображения, то и зависимость значений, несомненно, присутствует. Но предположение позволяет сделать существенные упрощения и, как будет показано, получить алгоритм настройки процедуры совместного обнаружения и локализации. Этот же алгоритм можно использовать и в ситуации, когда предположение не выполняется, понимая, однако, что решение в этом случае не является оптимальным. Итак, пусть выполняется равенство

$$p(\eta_0, \eta_l, K, \eta_{N-1}/\Omega_l) = \prod_{i=0}^{N-1} p_i(\eta_i/\Omega_l), \quad l = \{0, 1\},$$

здесь $p_i(\cdot)$ - плотность вероятности i -го отсчета изображения в окне пикового фильтра. В таком случае выражение (5.13) преобразуется к виду

$$P(Q/\Omega_l) = \\ = \int_0^{+\infty} p_0(\eta_0/\Omega_l) \int_{-\infty}^{\eta_0} p_1(\eta_l/\Omega_l) K \int_{-\infty}^{\eta_0} p(\eta_{N-1}/\Omega_l) d\eta_{N-1} K d\eta_l d\eta_0 = \\ = \int_0^{+\infty} p_0(\eta_0/\Omega_l) \prod_{i=1}^{N-1} P_i(\eta_0/\Omega_l) d\eta_0, \quad l = \{0, 1\}.$$

Здесь $P(\eta_0/\Omega_l)$ - функция распределения значений пикового фильтра в классе Ω_l .

В процессе обработки окно пикового фильтра последовательно занимает все возможные положения в плоскости дискретных аргументов изображения. Поэтому практически все пиксели изображения попадает в каждый из отсчетов окна пикового фильтра. А значит, разумно предположить, что законы распределения значений отсчетов изображения дискриминантной функции, попадающих в различные отсчеты окна пикового фильтра одинаковы, то есть $p_i(\eta/\Omega_l) = p(\eta/\Omega_l)$, $P_i(\eta/\Omega_l) = P(\eta/\Omega_l)$, $i = \overline{0, N-1}$. Тогда [23]:

$$P(Q/\Omega_l) = \int_0^{+\infty} p(\eta_0/\Omega_l) P^{N-1}(\eta_0/\Omega_l) d\eta_0 = \frac{1 - P^N(0/\Omega_l)}{N},$$

$$l = \{0, 1\}$$

и вероятности ошибок запишутся в виде

$$p_0 = \frac{1 - P^N(0/\Omega_0)}{N}, \quad p_1 = \frac{N-1}{N} + \frac{P^N(0/\Omega_1)}{N}. \quad (5.14)$$

Для полученных значений вероятностей (5.14) критерий качества (5.11) преобразуется к виду (отбрасываем несущественные постоянные величины):

$$r_1 P^N(0/\Omega_1) - r_0 P^N(0/\Omega_0) \rightarrow \min_{d(\dots)}, \quad (5.15)$$

а необходимое условие экстремума показателя качества запишется в виде

$$\frac{P^N(0/\Omega_1)}{P^N(0/\Omega_0)} = \frac{r_0}{r_1} \cdot \frac{P^{N-1}(0/\Omega_0)}{P^{N-1}(0/\Omega_1)}.$$

Производная вычисляется по неизвестным параметрам функции $d(\cdot)$, вид которой оказывает влияние на функции распределения отсчетов изображения в окне пикового фильтра в каждом из классов.

Для окончательного построения алгоритма обнаружения-локализации необходимо знание функции $d(\cdot)$ и параметров распределения вектора признаков в признаковом пространстве. Аналитический результат можно получить, если предположить линейность функции $d(\cdot)$ и специальный вид ее функции распределения в каждом из классов. В этом случае процесс построения классификатора можно получить в виде алгоритма Петерсона-Матсона [23,82].

*Случай линейной функции $d(\cdot)$
и функции распределения специального вида*

Пусть $P_d(\eta/\Omega_l)|_{l=\{0,1\}}$ - функция распределения значений функции $d(\cdot)$ с непрерывной плотностью вероятности $p_d(\eta/\Omega_l)|_{l=\{0,1\}}$. И пусть при изменении параметров функции $d(\cdot)$ изменения в законе распределения касаются только параметров среднего и дисперсии μ_l, σ_l в каждом из классов.

$$\begin{aligned}\mu_l &= \mathbf{W}^T \mathbf{M}_l + W_N, \\ \sigma_l^2 &= \mathbf{W}^T \mathbf{B}_l \mathbf{W},\end{aligned}$$

где \mathbf{W}, W_N - коэффициенты линейной формы и свободный член - неизвестные параметры функции $d(\cdot)$. $\mathbf{M}_l, \mathbf{B}_l$ - соответственно математическое ожидание и ковариационная матрица вектора признаков в каждом из классов:

$$\mathbf{M}_l = E\{\mathbf{Y}/\Omega_l\}, \mathbf{B}_l = E\{(\mathbf{Y} - \mathbf{M}_l)(\mathbf{Y} - \mathbf{M}_l)^T / \Omega_l\}, \quad l = \overline{0, L-1}.$$

Тогда (см. п.5.4) поиск значений линейной функции можно записать в виде (алгоритм Петерсона-Матсона):

$$\mathbf{W} = [t\mathbf{B}_0 + (1-t)\mathbf{B}_1]^{-1} (\mathbf{M}_0 - \mathbf{M}_1),$$

$$W_N = -\frac{t\sigma_0^2 \mathbf{W}^T \mathbf{M}_1 + (1-t)\sigma_1^2 \mathbf{W}^T \mathbf{M}_0}{t\sigma_0^2 + (1-t)\sigma_1^2}.$$

Здесь параметр $t \in \mathcal{R}$. По этому параметру производится оптимизация критерия (5.15) в следующем виде:

$$r_1 P_d^N \left(-\frac{\mu_1}{\sigma_1} \right) - r_0 P_d^N \left(-\frac{\mu_0}{\sigma_0} \right) \rightarrow \min_t.$$

Заметим, что при достаточной разделимости классов и близких априорных вероятностях, значение параметра находится в диапазоне $t \in [0, 1]$.

5.3.3 Настройка процедуры обнаружения и локализации при параметрическом задании дискриминантной функции

В ситуации, когда предположение о независимости значений отсчетов изображения дискриминантной функции в окне пикового фильтра не выполняется, для настройки локализатора-обнаружителя можно предложить итерационную процедуру, которая в качестве априорной информации использует выборочные данные в виде наборов векторов признаков.

Решение о классе объекта в текущей точке изображения принимается в ситуации, когда текущее значение $d(\cdot)$ оказывается больше, чем окружающие значения в области локализации, и выше порога, встроенного в качестве параметра в выражение функции $d(\cdot)$. Тогда событие Q принятия решения о том, что текущий фрагмент является искомым, представляется следующим образом:

$$Q = \{Tsd\}, \tag{5.16}$$

здесь d задает процесс расчета значений функции $d(\cdot)$, T - пороговый отбор (значений функции d), S - процедура локализации (по значениям функции d). Независимо от порядка следования процедур локализации и порогового отбора и та, и другая работают со значениями функции $d(\cdot)$. Поэтому результирующее решение в действительности получается как совместное решение процедур T и S об отнесении текущего фрагмента к классу объектов. Таким образом, событие (5.16) есть произведение двух событий - события локализации значения функции $d(\cdot)$ и события превышения заданного порога значением этой функции, то есть:

$$Q = \{Sd\} \cdot \{Td\}.$$

Первое из событий происходит тогда, когда решение об отнесении текущего фрагмента к классу объектов принимает процедура локализации. Второе событие - когда подобное решение принимает процедура порогового отбора. На процесс локализации-обнаружения, таким образом, можно смотреть как на процедуру принятия коллегиального решения двумя простейшими классификаторами - пороговым и экстремальным. Однако полностью отождествлять представление о локализаторе-обнаружителе, как о двух независимых классификаторах, некорректно, поскольку величина порогового значения связана с видом функции $d(\cdot)$, которая и формирует новое признаковое пространство для вышеназванных простейших классификаторов.

Поскольку порядок следования событий несущественен, вероятности классификации, определяющие качество классификации, могут быть записаны следующим образом:

$$\begin{aligned} P(Q/\Omega_i) &= P(\{Sd\} \cdot \{Td\}/\Omega_i) = \\ &= P(\{Sd\}/\Omega_i)P(\{Td\}/\Omega_i). \end{aligned} \tag{5.17}$$

На основе (5.17) строится следующая итерационная процедура формирования локализатора-обнаружителя (для определенности \mathbf{a}^k - вектор параметров функции $d(\cdot)$ на k -й итерации):

- задается некоторый начальный вид функции $d(\cdot)$, то есть задается событие $\{Td(\mathbf{a}^0)\}$ путем решения, например, задачи обнаружения:

$$\sum_{l=0}^1 (-1)^l r_l P(\{Td(\mathbf{a}^0)\}/\Omega_l) \rightarrow \min_{\mathbf{a}^0}, \quad \mathbf{a}^0 = \frac{\mathbf{a}^0}{\|\mathbf{a}^0\|};$$

- на очередной k -й итерации определяются оптимальные параметры и соответствующее событие $\{Td(\mathbf{a}^k)\}$ как решение задачи:

$$\sum_{l=0}^1 (-1)^k r_l P(\{Sd(\mathbf{a}^{k-1})\}/\Omega_l) P(\{Td(\mathbf{a}^n)\}/\{Sd(\mathbf{a}^{n-1})\}\Omega_l) \rightarrow \min_{\mathbf{a}^n};$$

- найденный вектор \mathbf{a}^n рассматривается как направление на оптимум, поэтому коррекция параметров осуществляется следующим образом:

$$\mathbf{a}^k = \mathbf{a}^{k-1} + \alpha_k \left(\frac{\mathbf{a}^k}{\|\mathbf{a}^k\|} - \mathbf{a}^{k-1} \right), \quad \mathbf{a}^k = \frac{\mathbf{a}^k}{\|\mathbf{a}^k\|};$$

- итерации заканчиваются, когда решение слабо меняется.

Для обеспечения сходимости итерационной процедуры используем известный из метода стохастической аппроксимации результат и зададим последовательность α_k величины коррекции таким образом, чтобы выполнялись условия [40]:

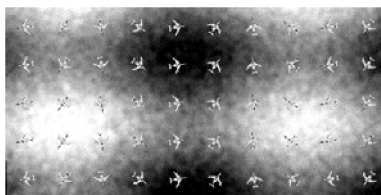
$$\alpha_k \xrightarrow[k \rightarrow \infty]{} 0, \quad \sum_{k=0}^{\infty} \alpha_k = \infty, \quad \sum_{n=0}^{\infty} \alpha_k^2 < \infty.$$

Для иллюстрации работоспособности предложенного алгоритма рассмотрим две задачи обнаружения и локализации объектов на изображении, приведенных на рис. 5.7а и рис. 5.7д.

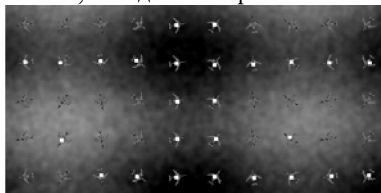
Функцию $d(\cdot)$ выберем линейной. В качестве признаков для первой задачи используем нормализованные показатели выборочных среднего и среднеквадратического отклонений в окне обработки. Для второй задачи используем единственный признак - локальное выборочное среднее.

На рис. 5.7б-5.7г и рис. 5.7е-5.7з приведены результаты решения соответствующих задач. Из иллюстрации к первой задаче, характеризующей высокой априорной вероятностью появления объекта на изображении, видно, что суммарная ошибка обнаружения, пропорциональная в данном случае паре чисел (число пропущенных объектов, число ложно обнаруженных), оказалась самой большой для ситуации, когда процедура обнаружения не учитывала наличие последующей локализации (б) – (26, 0). Самый низкий уровень ошибок оказался у процедуры совместного обнаружения и локализации, настраиваемой в предположении независимости отсчетов в окне локализации (в) – (0, 2).

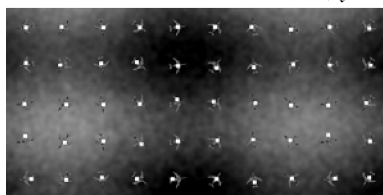
Из иллюстрации ко второй задаче, характеризующей относительно низкой априорной вероятностью появления объекта на изображении, видно, что без учета локализации при обнаружении ни один из объектов не был на изображении обнаружен и ошибка обнаружения в (е) составила (72,0). Следует заметить, что подобный результат типичен при обнаружении в ситуации с низкими априорными вероятностями появления объектов. В отличие от классической схемы решения при использовании совместных процедур обнаружения и локализации и соответствующих алгоритмов их настройки ошибка обнаружения оказалась достаточно низкой. Для представленных изображений (ж) и (з) она пропорциональна величинам (5, 45) и (5, 21). Таким образом, очевидно преимущество процедур совместного обнаружения и локализации по сравнению с последовательным использованием алгоритма обнаружения объекта и его последующей локализацией.



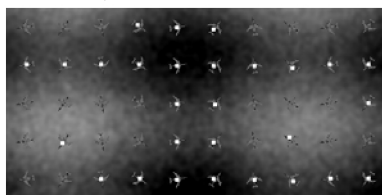
а) исходное изображение



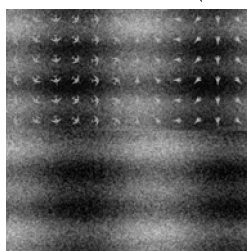
б) результат применения процедуры обнаружения с критерием (5.17) и последующей локализацией



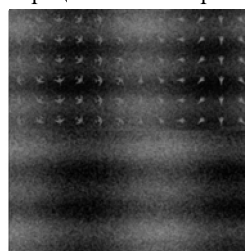
в) результат локализации-обнаружения в предположении независимости отсчетов в окне локализации



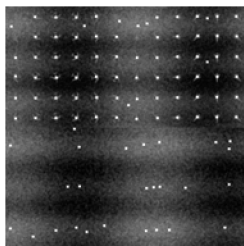
г) результат локализации-обнаружения при настройке процедуры с использованием итерационного алгоритма



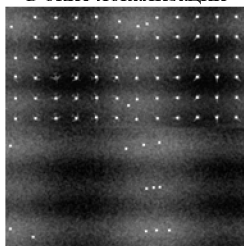
д) Исходное изображение



е) Результат применения процедуры обнаружения с критерием (5.17) и последующей локализацией



ж) Результат локализации-обнаружения в предположении независимости отсчетов в окне локализации



з) Результат локализации-обнаружения при настройке процедуры с использованием итерационного алгоритма

Рис. 5.7. Результаты обнаружения и локализации: задача 1 (а-з); задача 2 (д-з)

5.4. Распознавание объектов двух классов

В настоящем разделе рассмотрен один из известных алгоритмов, называемый ниже алгоритмом Петерсона-Матсона [40,82], который был разработан указанными авторами для нахождения параметров линейной разделяющей функции $d(\mathbf{y})$, минимизирующей величину среднего риска в предположении, что значения этой функции имеют нормальный закон распределения в каждом из классов. Ниже будет показано, что без каких-либо изменений данный алгоритм может быть использован и без предположения о нормальности закона распределения. Также будет предложен алгоритм нахождения параметров для случая, когда математические ожидания вектора признаков для раз-

личных классов совпадают - для такой ситуации известный алгоритм не позволяет получить решение. Кроме того, будут сделаны некоторые замечания о диапазоне параметра, используемого в алгоритме Петерсона-Матсона. И, наконец, будет показано, что тот же самый алгоритм можно использовать и при решении задачи Неймана-Пирсона. Для начала рассмотрим известный алгоритм.

5.4.1. Модификация алгоритма Петерсона-Матсона

Пусть линейная дискриминантная функция $d(\mathbf{y}) = \mathbf{W}^T \mathbf{y} + W_N$ имеет нормальный закон распределения в каждом из классов. Для этого предположения есть две предпосылки:

- вектор признаков \mathbf{Y} имеет нормальный закон распределения в каждом из классов, тогда и линейная комбинация элементов вектора имеет нормальный закон распределения;
- вектор признаков \mathbf{Y} не имеет нормального закона распределения, но количество признаков достаточно большое. В силу центральной предельной теоремы линейная комбинация элементов вектора имеет нормальный закон распределения.

Поведение значений функции $d(\mathbf{y})$ в каждом из классов полностью характеризуется математическим ожиданием и дисперсией, а именно:

$$\begin{aligned} \mu_l &= E\{d(\mathbf{Y})/\Omega_l\} = \mathbf{W}^T \mathbf{M}_l + W_N, \\ \sigma_l^2 &= E\left\{\left(d(\mathbf{Y}) - \mu_l\right)^2 / \Omega_l\right\} = \mathbf{W}^T \mathbf{B}_l \mathbf{W}, \quad l = \{0, 1\}. \end{aligned}$$

Здесь \mathbf{M}_l , \mathbf{B}_l - вектор математических ожиданий и ковариационная матрица вектора признаков в классе Ω_l .

Пусть, далее, необходимо минимизировать величину среднего риска (по параметрам дискриминантной функции):

$$R_{\min} = r_0 P(Q/\Omega_0) + r_1 P(\bar{Q}/\Omega_1) \rightarrow \min_{\mathbf{W}, w_n}, \quad (5.18)$$

где событие $Q = \{d(\bar{X}) > 0\}$. Тогда этот критерий эквивалентен следующему:

$$r_1 \Phi\left(-\frac{\mu_1}{\sigma_1}\right) - r_0 \Phi\left(-\frac{\mu_0}{\sigma_0}\right) \rightarrow \min_{\mathbf{W}, W_N}.$$

Здесь $\Phi(\cdot)$ - функция Лапласа. Необходимым условием минимума, как известно, является следующее:

$$r_1 \Phi'\left(-\frac{\mu_1}{\sigma_1}\right) - r_0 \Phi'\left(-\frac{\mu_0}{\sigma_0}\right) = 0.$$

Тогда, учитывая, что [40, 82]:

$$\begin{aligned} \frac{\partial}{\partial \mathbf{W}} \left(-\frac{\mu_l}{\sigma_l}\right) &= -\frac{\mathbf{M}_l}{\sigma_l} + \frac{\mathbf{W}^T \mathbf{B}_l \mu_l}{\sigma_l^3}, \\ \frac{\partial}{\partial W_N} \left(-\frac{\mu_l}{\sigma_l}\right) &= -\frac{1}{\sigma_l}, \quad l = \{0, 1\}, \end{aligned} \tag{5.19}$$

получим систему уравнений:

$$\begin{cases} \frac{r_0}{\sqrt{2\pi\sigma_0}} \exp\left(-\frac{\mu_0^2}{2\sigma_0^2}\right) \left(\mathbf{M}_0 + \frac{\mu_0}{\sigma_0^2} \mathbf{W}^T \mathbf{B}_0\right) = \\ = \frac{r_1}{\sqrt{2\pi\sigma_1}} \exp\left(-\frac{\mu_1^2}{2\sigma_1^2}\right) \left(\mathbf{M}_1 + \frac{\mu_1}{\sigma_1^2} \mathbf{W}^T \mathbf{B}_1\right), \\ \frac{r_0}{\sqrt{2\pi\sigma_0}} \exp\left(-\frac{\mu_0^2}{2\sigma_0^2}\right) = \frac{r_1}{\sqrt{2\pi\sigma_1}} \exp\left(-\frac{\mu_1^2}{2\sigma_1^2}\right). \end{cases} \tag{5.20}$$

Из этой системы очевидно следующее равенство:

$$\mathbf{M}_0 + \frac{\mu_0}{\sigma_0^2} \mathbf{W}^T \mathbf{B}_0 = \mathbf{M}_1 + \frac{\mu_1}{\sigma_1^2} \mathbf{W}^T \mathbf{B}_1. \tag{5.21}$$

Поскольку точного решения для выражения (5.21) получить невозможно, Петерсон и Матсон предложили следующую итерационную

процедуру. Вектор \mathbf{W} определяется как функция некоторого числового параметра t :

$$\mathbf{W} = a \left[t \mathbf{B}_0 + (1-t) \mathbf{B}_1 \right]^{-1} (\mathbf{M}_0 - \mathbf{M}_1), \quad (5.22)$$

где

$$t = \frac{-\mu_0/\sigma_0^2}{\mu_1/\sigma_1^2 - \mu_0/\sigma_0^2}, \quad a = \mu_1/\sigma_1^2 - \mu_0/\sigma_0^2. \quad (5.23)$$

Анализируя второе из равенств в (5.20), можно предположить, что оптимальное решение приведет к тому, что μ_0 и μ_1 будут иметь разные знаки, то есть $\mu_0 < 0$, $\mu_1 > 0$. В этом случае величина a - всего лишь масштабный множитель (всегда одного знака) и ее можно не учитывать, а диапазон $t \in [0, 1]$. Теперь значение W_N находится по формуле

$$W_N = - \frac{t \sigma_0^2 \mathbf{W}^T \mathbf{M}_1 + (1-t) \sigma_1^2 \mathbf{W}^T \mathbf{M}_0}{t \sigma_0^2 + (1-t) \sigma_1^2}. \quad (5.24)$$

Существует обобщение изложенного подхода для произвольного вида функционала качества [40], которое, однако, приводит только к системе уравнений и не дает окончательного решения. Одним из частных решений, полученных таким образом, является классификатор Фишера.

5.4.2 Обобщение процедуры Петерсона-Матсона на законы распределения специального вида

Требование нормальности распределений для получения процедуры решения задачи построения классификатора является слишком жестким. Для того, чтобы получить ту же самую итерационную процедуру вполне достаточно, чтобы выполнялись следующие требования:

- плотности вероятности являлись дифференцируемыми функциями по аргументу;

- изменение параметров функции $d(\cdot)$ приводило бы только к изменению математического ожидания и дисперсии в функциях распределения $P_d(\cdot/\Omega_0)$ и $P_d(\cdot/\Omega_1)$.

Тогда критерий (5.18) преобразуется к виду

$$r_1 P_d\left(-\frac{\mu_1}{\sigma_1}/\Omega_1\right) - r_0 P_d\left(-\frac{\mu_0}{\sigma_0}/\Omega_0\right) \rightarrow \min_{\mathbf{w}, W_N}.$$

Определим плотности вероятности:

$$p_l(u/\Omega_l) = \frac{dP_d(u/\Omega_l)}{du}, \quad l = \{0, 1\}.$$

А система уравнений примет следующий вид:

$$\begin{cases} \frac{r_0}{\sigma_0} p_0\left(-\frac{\mu_0}{\sigma_0}\right) \left(\mathbf{M}_0 + \frac{\mu_0}{\sigma_0^2} \mathbf{W}^T \mathbf{B}_0\right) = \\ = \frac{r_1}{\sigma_1} p_1\left(-\frac{\mu_1}{\sigma_1}\right) \left(\mathbf{M}_1 + \frac{\mu_1}{\sigma_1^2} \mathbf{W}^T \mathbf{B}_1\right), \\ \frac{r_0}{\sigma_0} p_0\left(-\frac{\mu_0}{\sigma_0}\right) = \frac{r_1}{\sigma_1} p_1\left(-\frac{\mu_1}{\sigma_1}\right). \end{cases} \quad (5.25)$$

Откуда снова приходим к выражению (5.21) и итеративной процедуре Петерсона-Матсона (5.22)-(5.24).

При этом нет необходимости адаптировать алгоритм Петерсона-Матсона к какому-либо конкретному распределению - вполне достаточно лишь предположения о выполнении указанных требований. Нахождение же оптимального значения параметра t , вообще говоря, можно производить и не зная конкретный вид закона, то есть проверяя оптимальность решения по выборочным данным при условии, что выборка репрезентативна.

5.4.3. Построение линейного классификатора при совпадающих средних вектора признаков в классах

При построении классификатора возможна ситуация, когда математические ожидания \mathbf{M}_0 и \mathbf{M}_1 вектора признаков оказываются одинаковыми, то есть

$$\mathbf{M}_0 = \mathbf{M}_1 = \mathbf{M}. \quad (5.26)$$

Подобная ситуация может возникнуть, например, если набор признаков не обладает достаточными дискриминантными свойствами.

Совершенно очевидно, что итерационная процедура (5.22)-(5.24) не даст решения в такой ситуации. Поэтому требуется так модифицировать алгоритм, чтобы получить решение. Обратимся к выражению (5.21). Учитывая равенство (5.26), (5.21) можно переписать следующим образом:

$$\frac{\mu}{\sigma_0^2} \mathbf{W}^T \mathbf{B}_0 = \frac{\mu}{\sigma_1^2} \mathbf{W}^T \mathbf{B}_1,$$

Здесь $\mu = \mu_0 = \mu_1$. Перепишем последнее соотношение в виде

$$\left(\frac{1}{\sigma_0^2} \mathbf{B}_0 - \frac{1}{\sigma_1^2} \mathbf{B}_1 \right) \mathbf{W} = 0. \quad (5.27)$$

Решить такое уравнение, также как и (5.21), аналитически невозможно. Поэтому введем переменную t :

$$t = \frac{-\frac{1}{\sigma_1^2}}{\frac{1}{\sigma_0^2} - \frac{1}{\sigma_1^2}} = \frac{\sigma_0^2}{\sigma_0^2 - \sigma_1^2}.$$

Тогда, с точностью до ненулевого масштабного множителя, условие (5.27) переписывается в виде

$$(t\mathbf{B}_1 + (1-t)\mathbf{B}_0) \mathbf{W} = 0. \quad (5.28)$$

Заметим, что диапазон изменения переменной t , в отличие от указанного в [40,82], не ограничен отрезком $[0, 1]$, а определяется всей

числовой прямой $t \in \mathcal{R}$. Кроме того, для конкретного значения t вектор \mathbf{W} может и не существовать. Действительно, выражение (5.28) относительно вектора \mathbf{W} есть определение ортогонального векторного пространства к набору векторов-строк матрицы

$$t\mathbf{B}_1 + (1-t)\mathbf{B}_0. \quad (5.29)$$

Для существования такого ортогонального пространства необходимо (и достаточно), чтобы среди строк матрицы (5.29) была хотя бы одна линейно зависимая. В этом случае определитель матрицы оказывается нулевым. Если же определитель ненулевой, значит, все строки независимы и ортогональное пространство построить невозможно - оно оказывается пустым. Таким образом, бессмысленно решать уравнение (5.28) в ситуации, когда не выполняется условие:

$$\det(t\mathbf{B}_1 + (1-t)\mathbf{B}_0) = 0. \quad (5.30)$$

Учитывая, что матрицы \mathbf{B}_1 и \mathbf{B}_0 невырожденные, последнее условие можно преобразовать к виду

$$\det\left(\mathbf{B}_1\mathbf{B}_0^{-1} + \frac{1-t}{t}\mathbf{I}\right) = 0, \quad (5.31)$$

здесь \mathbf{I} - единичная матрица. Значения величины $\frac{1-t}{t}$ лежат на всей числовой оси, за исключением точки $t = 0$. Но при этом значении определитель (5.30) ненулевой, так как матрица (5.29) оказывается равной ковариационной матрице \mathbf{B}_0 . Поэтому можно переписать (5.31) в виде:

$$\det\left(\mathbf{B}_1\mathbf{B}_0^{-1} + \frac{1}{t}\mathbf{I}\right) = 0, \quad (5.32)$$

Если при распознавании используется N признаков, размер матрицы $\mathbf{B}_1\mathbf{B}_0^{-1}$ составляет $N \times N$. Тогда, следуя определению определителя, (5.32) можно переписать в виде уравнения N -ого порядка относительно величины \tilde{t} . А, значит, можно решать это уравнение относительно переменной \tilde{t} теми же методами, что и обычные уравнения

такого рода [2, 22]. Максимальное количество линейно независимых векторов-решений \mathbf{W} исходного уравнения (5.28) не превышает размера матрицы N , причем при одном значении переменной \tilde{t} (корне) их число равно кратности корня. Сам же вектор \mathbf{W} может легко быть найден путем процесса ортогонализации Грама-Шмидта набора векторов: строк матрицы (5.32) и некоторого произвольного вектора \mathbf{W} .

Пусть теперь вектор \mathbf{W} найден. Для получения окончательного ответа необходимо определить значение свободного члена W_N . Это можно сделать на основе выражения (5.25). Покажем это на примере гауссовского распределения. Учитываем, что выполняется равенство (5.25) в виде:

$$\frac{r_0}{\sqrt{2\pi\sigma_0}} \exp\left(-\frac{\mu^2}{2\sigma_0^2}\right) = \frac{r_1}{\sqrt{2\pi\sigma_1}} \exp\left(-\frac{\mu^2}{2\sigma_1^2}\right), \quad (5.33)$$

тогда

$$\mu = \pm \sigma_0 \sigma_1 \sqrt{\frac{2}{\sigma_1^2 - \sigma_0^2} \ln\left(\frac{r_0 \sigma_1}{r_1 \sigma_0}\right)}.$$

Учитывая, что $\mu = \mathbf{W}^T \mathbf{M} + W_N$, получаем выражение для искомой величины:

$$W_N = \pm \sigma_0 \sigma_1 \sqrt{\frac{2}{\sigma_1^2 - \sigma_0^2} \ln\left(\frac{r_0 \sigma_1}{r_1 \sigma_0}\right)} - \mathbf{W}^T \mathbf{M}. \quad (5.34)$$

Задача неразрешима, если выражение под корнем отрицательно. Это происходит, если выполняется одно из следующих соотношений:

$$\sigma_1 > \sigma_0 \text{ и } r_0 \sigma_1 < r_1 \sigma_0,$$

либо

$$\sigma_1 < \sigma_0 \text{ и } r_0 \sigma_1 > r_1 \sigma_0.$$

Графически эта ситуация выглядит так, как показано на рис.5.8. То есть штрафы за решения выбраны таким образом, что минимум риска

достигается только в ситуации, если все пространство классифицируется в какой-то один класс.

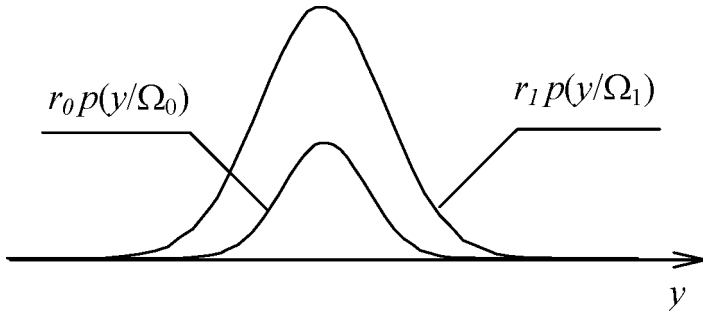


Рис. 5.8. Пример ситуации с отсутствием решения

Такая ситуация невозможна, если штрафы одинаковы, и решение принимается на основе анализа только плотностей. Единственная аномалия возникает, когда плотности совпадают, но тогда абсолютно бессмысленно выбирать какое-либо конкретное решение, так как все они дают одинаковую величину риска.

5.4.3.1. Двумерное пространство признаков ($N=2$)

Рассмотрим простую ситуацию, когда пространство признаков двумерно. Тогда возможно получить окончательное аналитическое решение для классификатора. Итак, пусть матрицы \mathbf{V}_1 и \mathbf{V}_0 имеют следующий вид:

$$\mathbf{V}_0 = \begin{pmatrix} a & c \\ c & b \end{pmatrix}, \quad \mathbf{V}_1 = \begin{pmatrix} d & q \\ q & e \end{pmatrix},$$

причем выполняются равенства $ab > c^2$ и $de > q^2$. Тогда матрица (5.29) запишется следующим образом:

$$\mathbf{V} = \begin{pmatrix} ta + (1-t)d & tc + (1-t)q \\ tc + (1-t)q & tb + (1-t)e \end{pmatrix}. \quad (5.35)$$

Условие существования ортогонального пространства (5.30) преобразуется к виду:

$$(ta + (1-t)d)(tb + (1-t)e) - (tc + (1-t)q)^2 = 0,$$

либо, в виде квадратного уравнения:

$$t^2 \left((a-d)(b-e) - (c-q)^2 \right) + \\ + t(ae + bd + 2q^2 - 2cq - 2de) + (de - q^2) = 0.$$

Обозначим

$$A = (a-d)(b-e) - (c-q)^2, \quad B = ae + bd + 2q^2 - 2cq - 2de, \\ C = de - q^2.$$

Тогда сразу можно записать решение:

$$t_{1,2} = \frac{-B \pm \sqrt{B^2 - 4AC}}{2A}.$$

Параметр W_N определяется из равенства (5.25).

Наиболее простое решение получается в ситуации, когда дисперсии по координатам одинаковы, то есть $a = b$ и $d = e$. Тогда

$$B|_{t_1} = \gamma_1 \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad B|_{t_2} = \gamma_2 \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}.$$

А соответствующие векторы линейной дискриминантной функции, определяющие ортогональное векторное пространство, таковы:

$$\mathbf{W}|_{t_1} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \quad \mathbf{W}|_{t_2} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

На рис. 5.9 приведен пример-иллюстрация двумерного случая.

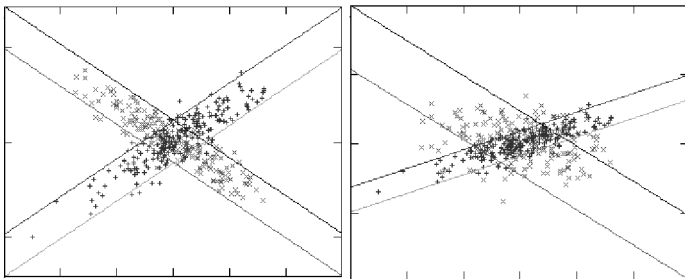


Рис. 5.9. Линейные разделяющие функции при равных средних вектора признаков

5.4.3.2. *Замечание о диапазоне переменной в итерационной процедуре алгоритма Петерсона-Матсона*

В работах [40,82] диапазон величины t , определенной выражением (5.23), принимается ограниченным: $t \in [0,1]$. Такое ограничение следовало из выполнения в системе уравнений (5.25) второго равенства, которое предполагает ясную геометрическую трактовку, изображенную на рис. 5.10. В такой ситуации $\mu_0 < 0 < \mu_1$, и значение t действительно принадлежит вышеуказанному диапазону. Но это возможно только в ситуации достаточно качественного разделения гиперплоскостью классов в пространстве признаков.

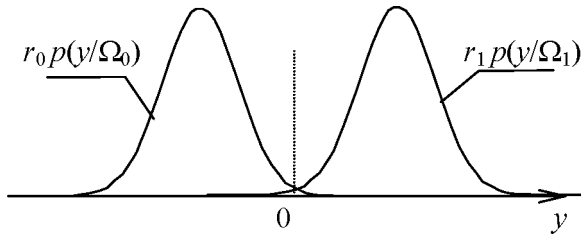


Рис. 5.10. Геометрическая интерпретация необходимого условия экстремума: ситуация высокой разделимости классов

В действительности вполне возможна ситуация, изображенная на рис. 5.11. Такое возможно, например, если одно из весовых значений в целевой функции существенно больше другого, или если один из классов существенно больше «размыт» в признаковом пространстве. А поскольку это достаточно характерные ситуации, то необходимо, оставаясь в рамках линейного классификатора, все же найти наилучшее решение. При этом заметим, что возможно получение не одного, а нескольких классификаторов, среди которых и надо отобрать наилучший.

Для указанного обобщения необходимо расширить область параметра t на всю вещественную прямую: $t \in \mathcal{R}$. На рис. 5.12 приведен пример для одной из ситуаций, когда оптимальное значение параметра t лежит вне интервала $[0, 1]$.

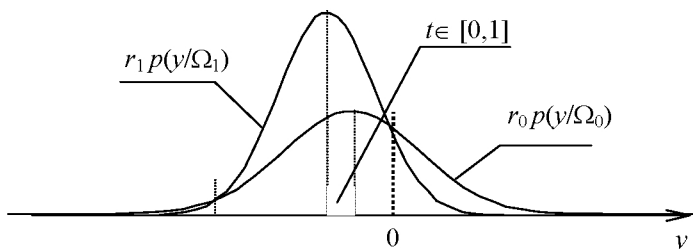


Рис. 5.11. Геометрическая интерпретация необходимого условия экстремума: ситуация низкой разделимости классов

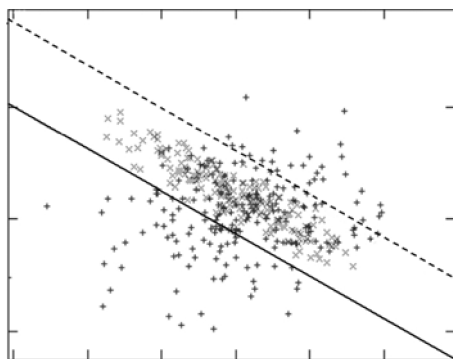


Рис. 5.12. Множество локальных решений линейного классификатора: “- - -” $t = -0,094$, “—” $t = -0,024$

5.4.4 Линейный классификатор, оптимальный по критерию Неймана-Пирсона

Во многих задачах вместо линейного критерия оптимизации используется критерий Неймана-Пирсона (4.3.3):

$$\begin{cases} p_1 \rightarrow \min, \\ p_0 = p_0^*. \end{cases} \quad (5.36)$$

Заметим, что в ситуации, когда множество значений вероятностей ошибок (p_0, p_1) имеет выпуклый характер, критерии Неймана-Пирсона и минимума риска эквивалентны при соответствующем выборе параметров оптимизации, то есть величин r_1, r_0 и p_0^* [40]. В альтернативной ситуации линейная задача оптимизации не позволяет

получить все решения на границе множества. В то же время критерий Неймана-Пирсона позволяет это сделать. Поэтому важно уметь решать задачу Неймана-Пирсона для рассматриваемого линейного классификатора.

Итак, пусть решается задача (5.36) для линейного классификатора со всеми описанными выше предположениями на законы распределения значений дискриминантной функции. Перепишем ее, используя метод множителей Лагранжа, в виде безусловной целевой функции вида:

$$p_1 + \alpha (p_0 - p_0^*)^2 \rightarrow \min,$$

здесь α - множитель Лагранжа ($\alpha \gg 0$). Тогда необходимым условием минимума является:

$$p_1' = 2\alpha (p_0^* - p_0) p_0'.$$

Учитывая, что

$$p_1 = P_d \left(-\frac{\mu_1}{\sigma_1} / \Omega_1 \right), \quad p_0 = 1 - P_d \left(-\frac{\mu_0}{\sigma_0} / \Omega_0 \right),$$

получим следующее уравнение:

$$p_1 \left(-\frac{\mu_1}{\sigma_1} \right) \left(-\frac{\mu_1}{\sigma_1} \right)' = 2\alpha \left(1 - p_0^* - P_d \left(-\frac{\mu_0}{\sigma_0} \right) \right) p_0 \left(-\frac{\mu_0}{\sigma_0} \right) \left(-\frac{\mu_0}{\sigma_0} \right)'.$$

Подставляя в последнюю формулу выражение для производной по вектору \mathbf{W} и порогу W_N , снова приходим к аналогичной (5.25) системе уравнений и, далее, получаем такое же уравнение (5.21) относительно вектора \mathbf{W} . Понятно, что результатом рассуждений снова окажется итерационная процедура Петерсона-Матсона (5.22)-(5.24).

Тот факт, что, используя рассматриваемую процедуру, для любого значения вероятности ложной тревоги можно найти соответствующее ей наименьшее значение вероятности пропуска объекта означает, что процедура Петерсона-Матсона позволяет получить всю границу множества значений вероятностей ошибок. А следовательно, любую зада-

чу оптимизации с критерием, требующим совместного понижения значений вероятностей p_0 и p_1 , и линейным классификатором можно решить, используя алгоритм Петерсона-Матсона - то есть простым перебором значений величины t . Этот факт очень удобен при построении линейного классификатора на базе других критериев - нет необходимости разрабатывать какую-то другую процедуру поиска решающего правила.

6. МЕТОДЫ СОВМЕСТНОЙ КЛАССИФИКАЦИЯ

Основной целью проектирования любой системы классификации является построение такого решающего правила, которое обеспечивало бы максимально возможное качество распознавания. Именно поэтому в последнее время начали интенсивно развиваться математические методы и алгоритмы построения решающих правил с повышенной структурной сложностью - *мультиклассификаторов* [49, 58, 66, 67, 69, 70, 72, 84, 87]¹.

Идея мультиклассификатора достаточно проста - использовать результаты распознавания сразу нескольких классификаторов для улучшения качественных показателей всей системы в целом. Подобная *совместная классификация*, очевидно, в целом должна оказаться не хуже индивидуальных решений, которые принимаются каждым из используемых классификаторов, называемых в ряде работ *экспертами* (experts) [80].

В то же время, ориентация на схему принятия решения со многими классификаторами породила целый ряд теоретических и практических проблем: как наилучшим образом комбинировать классификаторы, как настраивать параметры получаемого мультиклассификатора, можно ли улучшить качество совместной классификации при коррекции параметров классификаторов-экспертов и как это делать, какова граница качественных показателей при совместной классификации и т.д. На многие из этих вопросов до сих пор нет исчерпывающих и полных ответов. Это еще раз показывает, что теория и методы совместной классификации - новое, развивающееся направление в распознавании образов. Несмотря на это, существует целый ряд теоретически интересных и полезных с практической точки зрения результатов.

¹ В настоящее время в рамках данного направления нет устоявшейся терминологии. В связи с этим некоторые используемые далее термины не являются общепринятыми.

6.1. Стратегии совместной классификации

Поиски путей нахождения наилучшей схемы совместной классификации привели к появлению целого ряда различных мультиклассификаторов. В то же время стратегий совместной классификации, как показывает анализ этих решающих функций, достаточно немного. Выделим их.

С точки зрения реализации процесса совместной классификации очевидно разделение стратегий совместной классификации на *последовательные* и *параллельные*.

В рамках *последовательной стратегии классификации* вектор(ы) признаков подвергаются последовательному анализу каждым из входящих в мультиклассификатор экспертом. В случае, если данный эксперт достаточно "эрудирован", ответ о принадлежности к конкретному классу выносится немедленно. В противном случае анализ данных осуществляется следующим экспертом. Задачей настройки последовательной схемы классификации является нахождение параметров всех

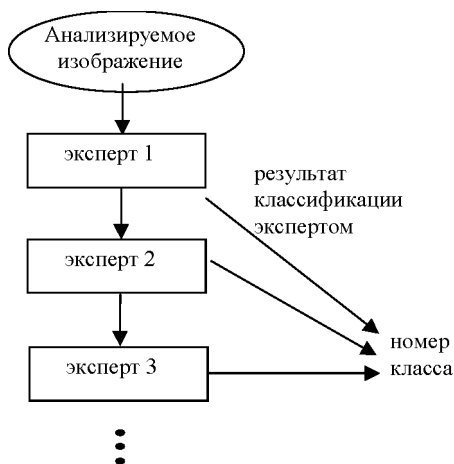


Рис. 6.1. Схема мультиклассификатора для последовательной стратегии классификации

классификаторов и правила их взаимодействия. Иллюстрация последовательной стратегии классификации представлена на рис. 6.1.

Число экспертов в такой схеме принятия решения может быть различным: от двух классификаторов до неограниченного их количества.

Надо отметить, что появление и развитие последовательной стратегии классификации во многом было обусловлено существованием хорошо развитой

статистической теорией последовательного анализа гипотез [4, 5, 20, 30, 39].

Классической работой в области последовательного анализа считается работа Вальда [5]. Суть метода при проверке двух гипотез заключается в следующем. Все выборочные данные разбиваются, искусственным или естественным путем, на группы *наблюдения*. Начиная с первого наблюдения производится расчет вектора признаков, а затем и *отношения правдоподобия*:

$$\Lambda(\mathbf{y}^i) = \frac{p(\mathbf{y}^k / \Omega_1)}{p(\mathbf{y}^k / \Omega_0)}, \quad k = 1, 2, \dots$$

Здесь \mathbf{y}^k - объединенный вектор признаков до k -го наблюдения. Значение отношения правдоподобия при очередном наблюдении используется для вынесения одного из трех решений:

- если $\Lambda(\mathbf{y}^k) < A$, то принимается гипотеза Ω_0 ,
- если $\Lambda(\mathbf{y}^k) > B$, то принимается гипотеза Ω_1 ,
- если $A < \Lambda(\mathbf{y}^k) < B$, то производится следующее наблюдение.

Пороговые значения A и B выбираются таким образом, чтобы обеспечить требуемые значения ошибок классификации. Метод проверки двух гипотез, называемый *последовательным критерием отношения правдоподобия Вальда*, является оптимальным в том смысле, что минимизирует одновременно средний риск и среднюю длительность наблюдений [20]. Дальнейшее его развитие привело к появлению методов многоальтернативного последовательного анализа гипотез, методов усечения (остановки) последовательных процедур, различных обобщений и дополнений [4, 30, 31, 32, 86]. В одной из последних отечественных работ в области последовательного статистического оценивания [32] вместо отношения правдоподобия используется некоторая наперед заданная статистика, а пороги варьируются на каждом шаге.

Однако, несмотря на достаточно широкий круг работ в этой области, использование методологии последовательного анализа для построения и оптимизации последовательной схемы совместной классификации в общем случае не достаточно эффективно. Это обусловлено несколькими

причинами. Во-первых, решение на каждом этапе анализа в последовательной схеме классификации в общем случае принимается не на основании статистики, а классификатором, параметры которого заранее могут быть неизвестными. Во-вторых, эксперты различных уровней могут быть «компетентны» лишь на ограниченном числе классов. Например, в двухэтапной процедуре распознавания [24,50] первый эксперт предназначен для решения задачи обнаружения, а второй - задачи распознавания. В последовательных же процедурах классификации все элементы структурно эквивалентны.

Выделенные основные недостатки существующих решений в области последовательной организации классификации делают целесообразным разработку новых алгоритмов параметрической и структурной оптимизации последовательных мультиклассификаторов. Этим обусловлено появление большого числа работ в этой области [30, 31, 32].

В рамках *параллельной стратегии классификации* вектор(ы) признаков подвергаются одновременному анализу несколькими классификаторами-экспертами. Каждый из классификаторов выносит свое решение о принадлежности анализируемых данных к тому или иному классу. В этом случае задачей настройки параллельной схемы классификации является задача поиска функции «агрегирования» решений различных экспертов, а также поиск оптимальных параметров классификации каждого эксперта в составе общей схемы. Структура параллельного мультиклассификатора приведена на рис. 6.2.

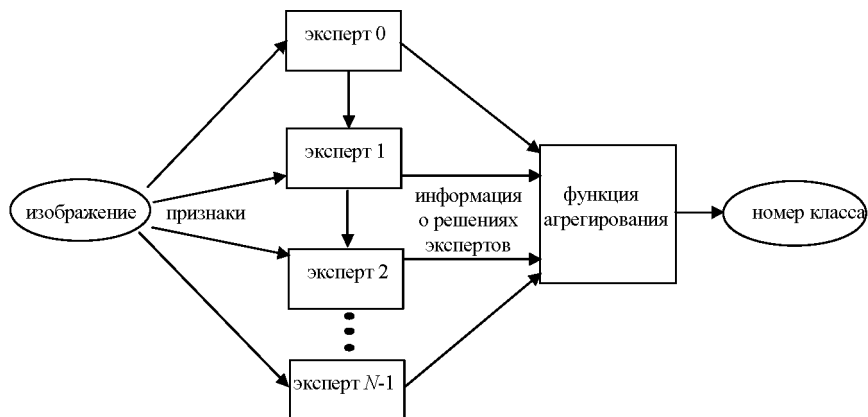


Рис.6.2. Схема мультиклассификатора для параллельной стратегии классификации

Заметим, что при параллельной комбинации экспертов возможно различное представление информации об их решениях. Наиболее типичными формами представления решений являются:

- номер класса, к которому принадлежат классифицируемые данные по мнению эксперта (минимальный объем информации);
- апостериорные вероятности принадлежности классифицируемых данных к каждому классу (максимальный объем информации).

Появление параллельных стратегий классификации вначале было обусловлено желанием разработчиков повысить качество систем распознавания. В последнее время параллельная стратегия классификации бурным развитием обязана в первую очередь возросшим потенциалом вычислительной техники, в особенности компьютерам с параллельной архитектурой. Использование многопроцессорных вычислительных комплексов для решения задачи распознавания позволяет естественным образом разделить процедуру классификации на ряд параллельно выполняемых процессов, каждый из которых реализует процесс принятия решения некоторого конкретного эксперта. Подобная технология в настоящее время эффективно используется, например, в нейронных сетях [66, 67, 84].

Заметим также, что поиск функции агрегирования в первое время проводился эвристическими методами, часто использовались аналогии с коллективными методами принятия решений, которые характерны для людей. Наиболее ярким примером такой функции агрегирования является классификация по правилу «максимального голосования». В соответствии с этим правилом решение о принадлежности к некоторому классу принималось, если за этот класс «проголосовало» наибольшее число экспертов. Однако до недавнего времени для большинства подобных схем классификации не были получены теоретические обоснования эффективности. Поэтому их использование во многом отражало детерминистический подход к построению мультиклассификаторов.

В рамках этого параграфа представлены некоторые методы и алгоритмы для параллельной и последовательной стратегий совместной классификации. В п.6.2. приведены алгоритмы настройки правила

агрегирования параллельного мультиклассификатора с минимальным объемом информации о решениях экспертов. Рассмотрены особенности настройки и реализации мультиклассификатора в частных случаях: при независимой работе экспертов, при малом числе классов. Даны выражения для оценки ошибок мультиклассификации.

Информацию о методе построения мультиклассификаторов, использующих максимальную информацию о решениях экспертов, и математическое обоснование соответствующих алгоритмов можно найти, например, в работе авторов J.Kittler, M.Hatef и R.P.W.Duin [72]. Полученные ими результаты направлены на теоретическое обоснование ряда эвристических методов совместной классификации, широко используемых в распознавании образов.

В п.6.3. приведен метод построения итерационных алгоритмов параметрической оптимизации классификаторов-экспертов в составе двухэтапной последовательной процедуры классификации. Рассмотрены особенности реализации такой процедуры, алгоритмов ее настройки и эффективности на примере решения задачи поиска и распознавания локальных объектов на изображении.

6.2. Параллельная схема совместной классификации с минимальной информацией о решениях экспертов

Как отмечалось в п.6.1, схемы параллельной классификации могут различаться способом представления информации о решениях экспертов. Рассмотрим ситуацию, когда информация от эксперта поступает в виде номера класса. В этом случае не требуется никаких дополнительных предположений о способе представления образа и виде решающей функции. Вначале дадим формальную постановку задачи мультиклассификации, а затем рассмотрим возможные методы ее решения.

6.2.1. Формальная постановка задачи мультиклассификации

Пусть все множество объектов разделено на L классов Ω_l , $l = \overline{0, L-1}$. Пусть для классификации используются N классификаторов,

каждый из которых выносит свое решение Q_n ($n = \overline{0, N-1}$) относительно принадлежности конкретного объекта ω некоторому классу. Для определенности будем считать, что n -й классификатор относит анализируемый объект к l -му классу, если $Q_n = l$. Таким образом множество возможных решений каждого из классификаторов $\{l: l = \overline{0, L-1}\}$, и на решение Q_n можно смотреть как на дискретную случайную величину с некоторым законом распределения. В этом случае на вход мультиклассификатора (функции агрегирования) подается случайный вектор $\mathbf{Q} = (Q_0, K, Q_{N-1})^T$, в качестве компонент содержащий мнения экспертов. Тогда решением задачи настройки мультиклассификатора является нахождение функции или отображения, которое по заданному конкретному вектору решений $\mathbf{q} = (q_0, K, q_{N-1})^T$ определяет, к какому именно классу Ω_l принадлежит соответствующий этому вектору объект ω .

Очевидно, постановка задачи таким образом приводит нас к стандартной байесовской процедуре классификации дискретного случайного вектора (см. п.4).

6.2.2 Байесовская процедура совместной классификации

Пусть на всем множестве возможных векторов (возможных решений экспертов) $q = (q_0, K, q_{N-1})^T$ для каждого класса Ω_l ($l = \overline{0, L-1}$) определена его функция правдоподобия $p(\mathbf{Q} = \mathbf{q} / \Omega_l)$. Тогда оптимальным решением задачи классификации, как известно, является правило максимума апостериорной вероятности (см. п.2.1)², объект относят к тому классу Ω_l , для которого

$$P(\Omega_l / \mathbf{Q} = \mathbf{q}) = \max_{j=0, L-1} P(\Omega_j / \mathbf{Q} = \mathbf{q}). \quad (6.1)$$

² Здесь и далее в разделе матрица штрафов предполагается простейшей.

Здесь апостериорные вероятности $P(\Omega_j/\mathbf{Q}=\mathbf{q})$ определяется по формуле Байеса:

$$P(\Omega_j/\mathbf{Q}=\mathbf{q}) = \frac{P(\Omega_j) P(\mathbf{Q}=\mathbf{q}/\Omega_j)}{\sum_{i=0}^{L-1} P(\Omega_i) P(\mathbf{Q}=\mathbf{q}/\Omega_i)}$$

На практике процедура построения байесовского мультиклассификатора сводится к получению для каждого класса Ω_l ($l=\overline{0, L-1}$) таблицы размера L^N , которая задает распределение вероятностей случайного вектора $\mathbf{Q}: P(\mathbf{Q}=\mathbf{q}/\Omega_l)$. Таким образом, необходимо произвести оценку L^{N+1} вероятностей $P(\mathbf{Q}=\mathbf{q}/\Omega_l)$ ($l=\overline{0, L-1}$). При реализации процедуры распознавания достаточно одной таблицы размером L^N , которая содержит для каждого возможного вектора \mathbf{q} соответствующий номер класса, получаемый заранее в соответствии с формулой (6.1). В табл.6.1 приведены размеры такой таблицы.

Таблица 6.1. Размер таблицы байесовского мультиклассификатора

		число классов L							
		2	4	8	16	32	64	128	256
число экспертов N	2	2 ²	2 ⁴	2 ⁶	2 ⁸	2 ¹⁰	2 ¹²	2 ¹⁴	2 ¹⁶
	4	2 ⁴	2 ⁸	2 ¹²	2 ¹⁶	2 ²⁰	2 ²⁴	2 ²⁸	2 ³²
	8	2 ⁸	2 ¹⁶	2 ²⁴	2 ³²	2 ⁴⁰	2 ⁴⁸	2 ⁵⁶	2 ⁶⁴
	16	2 ¹⁶	2 ³²	2 ⁴⁸	2 ⁶⁴
	32	2 ³²	2 ⁶⁴

Для реализации байесовской процедуры мультиклассификации в режиме реального времени (например, при автоматическом чтении текста) занимаемый соответствующей таблицей объем данных должен помещаться в оперативную память компьютера, что означает практическую невозможность использования мультиклассификатора для тех параметров (N, L) таблицы, которые выделены серым цветом.

Кроме того, при настройке такого мультиклассификатора возникают дополнительные ограничения, связанные с устойчивостью получаемых оценок вероятностей соответствующих законов распределения. Действительно, получение «хорошей» в статистическом смысле оценки одного значения вероятности требует десятков, а иногда сотен и тысяч экспериментов. Учитывая при этом, что число оцениваемых значений вероятностей составляет L^{N+1} , то объем выборочных данных становится просто катастрофически большим.

Таким образом, объективно существуют причины, которые в ряде случаев делают невозможным использование байесовской процедуры мультиклассификации. Это привело к появлению различных алгоритмов, основанных на определенных упрощениях. Одним из типичных упрощений является допущение независимости решений классификаторов-экспертов и/или их признаков. Ниже приведен подход построения мультиклассификатора для такой ситуации. Однако надо заметить, что в ситуации небольшого числа классов и малого количества классификаторов целесообразно использовать приведенную выше байесовскую процедуру в силу ее оптимальности.

6.2.3. Мультиклассификаторы при независимых экспертах

Предположим, что решения Q_n , выносимые различными решающими правилами в мультиклассификаторе, являются независимыми в каждом из классов. С вероятностной точки зрения это предположение означает выполнение равенств:

$$P(\mathbf{Q} = \mathbf{q}/\Omega_l) = \prod_{n=0}^{N-1} P(Q_n = q_n/\Omega_l), \quad l = \overline{0, L-1}.$$

Тогда байесовская дискриминантная функция l -го класса

$$d_l(\mathbf{q}) = P(\Omega_l)P(\mathbf{Q} = \mathbf{q}/\Omega_l), \quad l = \overline{0, L-1},$$

представима в виде:

$$d_l(\mathbf{q}) = P(\Omega_l) \prod_{n=0}^{N-1} P(Q_n = q_n/\Omega_l), \quad l = \overline{0, L-1}.$$

Введем в рассмотрение функцию:

$$G_s(q) = \begin{cases} 1, & \text{если } q = s, \\ 0, & \text{иначе.} \end{cases}$$

Тогда значение вероятности $P(Q_n = q_n/\Omega_l)$ может быть представлено следующим образом:

$$P(Q_n = q_n/\Omega_l) = \sum_{s=0}^{L-1} P(Q_n = s/\Omega_l) G_s(q_n).$$

В этом случае выражение для дискриминантной функции мультиклассификатора переписывается в виде:

$$d_l(\mathbf{q}) = P(\Omega_l) \prod_{n=0}^{N-1} \sum_{s=0}^{L-1} P(Q_n = s/\Omega_l) G_s(q_n), \quad l = \overline{0, L-1}. \quad (6.2)$$

Применяя к полученному выражению логарифмическое преобразование, а также учитывая, что для конкретного n функция $G_s(q_n) = 0$ при всех $s \neq q_n$, получим

$$d_l(q) = \ln P(\Omega_l) + \sum_{n=0}^{N-1} \sum_{s=0}^{L-1} G_s(q_n) \cdot \ln P(Q_n = s/\Omega_l), \quad l = \overline{0, L-1}. \quad (6.3)$$

Обозначим:

$$z_j = G_s(q_n), \quad \text{где } j = L \cdot n + s, \quad j = \overline{0, L \cdot N - 1}, \\ s = \overline{0, L-1}, \quad z = (z_0, \dots, z_{L \cdot N - 1}). \quad (6.4)$$

Введем $z = (z_0, \dots, z_{L \cdot N - 1})^T$ - бинарный вектор длины $L \cdot N$. Введем коэффициенты

$$W_j^l = \ln P(Q_n = s/\Omega_l), \quad j = Ln + s, \quad n = \overline{0, N-1}, s = \overline{0, L-1}. \quad (6.5)$$

Тогда окончательным решением задачи является мультиклассификатор с дискриминантной функцией вида:

$$d_l(\mathbf{y}) = [\mathbf{W}^l]^T \mathbf{z} + \ln P(\Omega_l),$$

здесь \mathbf{W}^l - вектор-столбец, составленный из коэффициентов W_j^l ($j = \overline{0, L \cdot N - 1}$). Таким образом, оптимальный мультиклассификатор

оказался линейным относительно введенного бинарного вектора признаков (6.4) и преобразованного признакового пространства.

Замечание 1

В качестве функции $G_s(q)$ может выступать, например, полином $(L-1)$ -го порядка:

$$G_s(q) = \frac{1}{\prod_{l=0}^{L-1} (s-l)} \prod_{l=0}^{L-1} (q-l). \quad (6.6)$$

Из выражений (6.3) и (6.6) видно, что в признаковом пространстве решений экспертов мультиклассификатор является *полиномиальным*, и степень полинома, описывающего решающую функцию мультиклассификатора, равна $(L-1)$. То есть при построении совместного решения правила распознавания для двух классов мнения экспертов комбинируются в линейной форме, для трех классов они учитываются в полиноме второй степени и т.д.

Замечание 2

Представленный алгоритм построения мультиклассификатора при независимых решениях экспертов естественным образом переносится на ситуацию распознавания векторов-признаков со значениями в конечном поле.

В дополнение к изложенному общему алгоритму мультиклассификации рассмотрим один важный частный случай, когда решается задача двухальтернативного распознавания (обнаружения) и мультиклассификатор, в соответствии с первым замечанием, комбинирует мнения экспертов в линейной решающей функции. Данная задача известна как задача распознавания бинарных векторов признаков и ее решение дано, например, в монографии [33]. В данном случае ее решение вытекает как частный случай (6.2) и (6.6).

Итак, пусть $L=2$. Тогда функция $G_S(q)$ определяется следующим образом:

$$G_0(q) = 1 - q, \quad G_1(q) = q.$$

В результате дискриминантная функция (6.2) может быть переписана следующим образом:

$$d_l(\mathbf{q}) = P(\Omega_l) \prod_{n=0}^{N-1} \left((1 - P(Q_n = 1/\Omega_l))(1 - q_n) + P(Q_n = 1/\Omega_l)q_n \right),$$

$$l = \overline{0, L-1}$$

Окончательным решением задачи является мультиклассификатор с дискриминантной функцией вида:

$$d_l(\mathbf{q}) = \left[\ln(P(\Omega_l)) + \sum_{n=0}^{N-1} \ln(1 - P(Q_n = 1/\Omega_l)) \right] +$$

$$+ \sum_{n=0}^{N-1} q_n \cdot \ln \left(\frac{P(Q_n = 1/\Omega_l)}{1 - P(Q_n = 1/\Omega_l)} \right),$$

который, как ожидалось, является линейным относительно решений экспертов.

Аналогичным образом можно получить выражение для мультиклассификатора в терминах отношения правдоподобия (см. п.1.2.):

$$\forall j \neq l \quad \chi_{jl}^0(\bar{q}) \geq \chi_{jl}^0 \Rightarrow \mathbf{q} \in \Omega_l,$$

где

$$\chi_{jl}^0(\mathbf{q}) = \sum_{n=0}^{N-1} W_{jl}^n \cdot q_n, \quad W_{jl}^n = \ln \left(\frac{P(Q_n = 1/\Omega_l)}{1 - P(Q_n = 1/\Omega_l)} \cdot \frac{1 - P(Q_n = 1/\Omega_j)}{P(Q_n = 1/\Omega_j)} \right),$$

$$\chi_{jl}^0 = \ln \left(\frac{P(\Omega_j)}{P(\Omega_l)} \right) + \sum_{n=0}^{N-1} \ln \left(\frac{1 - P(Q_n = 1/\Omega_j)}{1 - P(Q_n = 1/\Omega_l)} \right).$$

Очевидно, отношение правдоподобия также является линейной функцией решений экспертов.

Замечание 3

В отличие от общей ситуации построения мультиклассификатора, количество требуемых для оценки параметров при независимых решениях экспертов в соответствии с выражением (6.5) составляет $L \times N$ для каждого класса. В этом случае объем данных, используемых при классификации равен $N \times L^2$, что существенно меньше, чем приведенные в табл. 6.1 значения. Это позволяет использовать представленную здесь процедуру совместной классификации для большого числа классов и классификаторов-экспертов.

6.2.4. Вычисление вероятностей ошибочной классификации при совместной классификации

Аналитическое выражение для вероятностей ошибок совместной классификации в общем случае не получено. Однако при небольшом количестве экспертов значения этих вероятностей можно вычислить на ПЭВМ, используя дискретный аналог формулы (4.26):

$$p_{lj} = \sum_{\bar{q} \in D_j} P(\mathbf{Q} = \bar{q} / \Omega_l), \quad l \neq j, \quad l, j = \overline{0, L-1},$$

поскольку в признаковом пространстве находится всего L^N элементов. Когда число компонентов вектора решений экспертов велико, можно получить приближенные выражения для вероятностей ошибочной классификации в случае $L=2$, если предположить независимость решений экспертов. Результаты расчета вероятностей известны для бинарного вектора признаков и даны, например, в [33]. В подобной ситуации в соответствии с центральной предельной теоремой можно считать закон распределения случайной величины $\tilde{\Lambda}_{lj} \equiv \tilde{\Lambda}_{lj}(\mathbf{Q})$ нормальным. Числовые характеристики закона распределения случайной величины $\tilde{\Lambda}_{10}$ имеют вид:

$$\begin{aligned}
\mu_l &= E\left\{\lambda_{10}^o(\mathbf{Q})/\Omega_l\right\} = \\
&= \sum_{n=0}^{N-1} \ln\left(\frac{P(Q_n=1/\Omega_1)}{1-P(Q_n=1/\Omega_1)} \frac{1-P(Q_n=1/\Omega_0)}{P(Q_n=1/\Omega_0)}\right) P(Q_n=1/\Omega_l), \\
\sigma_l^2 &= E\left\{\left(\lambda_{10}^o(\mathbf{Q}) - \mu_l\right)^2/\Omega_l\right\} = \\
&= \sum_{n=0}^{N-1} \left[\ln\left(\frac{P(Q_n=1/\Omega_1)}{1-P(Q_n=1/\Omega_1)} \frac{1-P(Q_n=1/\Omega_0)}{P(Q_n=1/\Omega_0)}\right)\right]^2 \times \\
&\times P(Q_n=1/\Omega_l)(1-P(Q_n=1/\Omega_l)), \quad l=0,1.
\end{aligned}$$

а для вероятностей ошибочной классификации получаем следующие приближенные выражения:

$$p_0 = 1 - \Phi\left(\frac{\lambda_0^o - \mu_0}{\sigma_0}\right), \quad p_1 = \Phi\left(\frac{\lambda_0^o - \mu_1}{\sigma_1}\right), \quad \text{где } \lambda_0^o = \ln \frac{P(\Omega_0)}{P(\Omega_1)}.$$

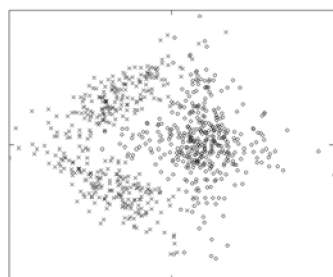
В ситуации, когда условия центральной предельной теоремы не выполняются, а решения экспертов независимы, можно воспользоваться следующими выражениями для верхних границ вероятностей ошибок, вытекающих из неравенства Чебышева:

$$p_0 \leq \frac{\sigma_0^2}{(\mu_0 - \lambda_0^o)^2}, \quad p_1 \leq \frac{\sigma_1^2}{(\mu_1 - \lambda_0^o)^2}.$$

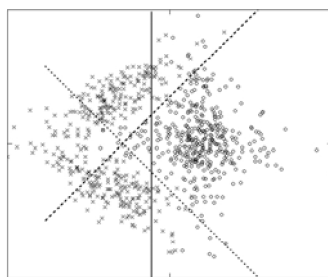
6.2.5. Исследование эффективности правил совместной классификации

Для оценки эффективности использования совместного решающего правила был проведен ряд экспериментальных исследований по классификации. Данные были представлены в виде векторов смеси нормальных законов распределений, как показано на рис. 6.3а. В качестве экспертов выступали линейные классификаторы, параметры которых оказывались локально оптимальными. Результаты экспериментальных исследований представлены в табл. 6.2 и на рис. 6.3б,в. Формируемая разделяющая граница мультиклассификатора, базируясь на линейных

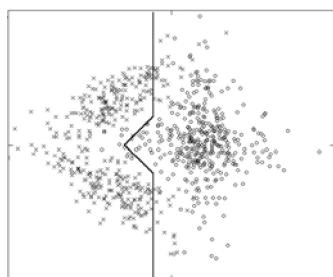
разделяющих границах экспертов, показанных на рис. 6.3б, оказывается кусочно-линейной (см. рис. 6.3в). Как следствие такой «адаптации» к типу распределения данных, качество работы мультиклассификатора оказалось лучше качества любого из экспертов и для вероятностей ошибок p_0 и p_1 , и для величины риска в целом (см. таблицу 6.2).



а) исходные данные



б) разделяющие границы экспертов



в) совместная классификация с независимыми решениями экспертов

“.....” – эксперт 1,

“- - - -” – эксперт 2,

“—” – эксперт 3,

“х” – объекты класса Ω_0 ,

“о” – объекты класса Ω_1 ,

Рис. 6.3. Вектора признаков и результаты классификации с минимальной информацией о решениях экспертов

Таблица 6.2. Эффективность совместной классификации с минимальной информацией о решениях экспертов

Классификатор	p_0	p_1	Риск
эксперт 1	0,785	0,085	0,435
эксперт 2	0,775	0,08	0,425
эксперт 3	0,23	0,12	0,175
Классификация оптимальным мультиклассификатором с независимыми решениями экспертов	0,23	0,055	0,143

6.3. Двухэтапная последовательная процедура классификации

6.3.1. Структура двухэтапной последовательной процедуры классификации

Примем модель данных, в соответствии с которой большая часть анализируемых (классифицированных) данных относится к не интересующему нас «фоновому» классу, а интересующие нас данные, интерпретируемые как «объекты», встречаются достаточно редко. Подобная модель характерна для многих приложений, в частности для задач поиска и распознавания локальных объектов на изображении [1, 3, 9, 10, 11, 50].

Из принятой модели следует, что прямое использование расширенного набора признаков и качественного решающего правила для большей части анализируемых фрагментов данных неэффективно и вычислительно избыточно, так как их занимает не интересующая нас фоновая составляющая. Поэтому представляется целесообразным:

- до процесса распознавания провести предварительный анализ данных с целью выделения из них фрагментов, которые потенциально содержат интересующую информацию - объекты. Объем «полезных» данных по сравнению с первоначальным оказывается сравнительно небольшим;
- по выделенным данным производить по возможности полный анализ с целью достаточно качественной классификации. Большая вычислительная сложность производимого анализа компенсируется малым числом анализируемых фрагментов.

Тогда процесс распознавания данных может быть синтезирован в рамках двухэтапной процедуры, представленной на рис. 6.4.

Для каждого фрагмента данных:

в рамках первого этапа процедуры

- с использованием быстрого алгоритма производится формирование вектора признаков;
- на основе сформированного вектора принимается решение о принадлежности текущего фрагмента либо к классу объектов, либо к классу фона (этап предварительного обнаружения).

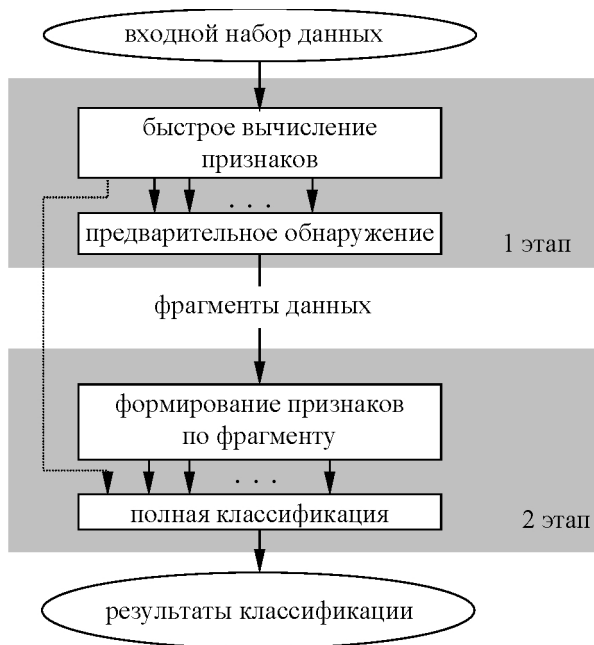


Рис. 6.4. Схема двухэтапной последовательной процедуры классификации

Пусть текущий фрагмент классифицируется как «объект», тогда классификатор второго этапа производит распознавание. А именно:

(в рамках второго этапа процедуры)

- по выделенному фрагменту данных производится формирование признаков, информативность которых должна быть достаточной для принятия решения с необходимым качеством;
- производится классификация фрагмента с использованием достаточно качественного классификатора второго этапа. При этом классификатор может использовать информацию, полученную на первом этапе.

Легко заметить, что первый этап по своему содержанию близок к обычной схеме обнаружения. Однако к нему предъявляются дополнительные требования: низкой вычислительной сложности и малой вероятности пропуска объекта при одновременно небольшом числе выделяемых фрагментов.

Основное преимущество данной двухэтапной процедуры обнаружения и распознавания состоит в возможности, с одной стороны,

снижения вычислительной сложности (первый этап) и, с другой стороны, использования расширенного набора средств формирования признаков и классификации (второй этап).

6.3.2 Соглашения об обозначениях

Введем вначале обозначения для обычного алгоритма распознавания, а затем, используя полученную терминологию, используем ее для описания двухэтапной последовательной процедуры классификации.

Пусть по фрагменту анализируемых данных γ , являющегося реализацией случайного фрагмента \mathbf{Y} , с помощью однозначного (но не взаимнооднозначного) преобразования ϕ формируется вектор признаков $\mathbf{y} = \{y_n\}_{n=0}^{N-1}$ (в частности, вектор \mathbf{y} может совпадать с самим фрагментом данных), который, в свою очередь, рассматривается как реализация случайного вектора \mathbf{Y} . Вектор признаков, как показано на рис.6.5, используется для классификации фрагмента γ в один из L классов $\{\Omega_l\}_{l=0}^{L-1}$, представители которых и формируют исходный набор данных. Следуя принятой выше терминологии, будем интерпретировать класс Ω_0 как класс фона, а оставшиеся $(L-1)$ классов - как классы объектов. Как известно, справедливы следующие соотношения [40]:

$$\Omega = \bigcup_{l=0}^{L-1} \Omega_l, \quad \Omega_i \cap \Omega_j = \emptyset, \quad i \neq j, \quad i, j = \overline{0, L-1}.$$

В то же время, пространство Ω , содержащее все возможные фрагменты данных γ , оказывается разбитым на множество областей $\Theta_l \left(\overline{l=0, L-1} \right)$, определяющих, по сути, работу классификатора. А именно, классификатор принимает решение о принадлежности фрагмента γ классу Ω_l в том случае, если $F(\gamma) = \mathbf{y} \in D_l$, при этом $\Theta_l = \phi^{-1}(D_l)$ является прообразом области D_l N -мерного метрического пространства признаков. Необходимость задания областей D_l и при-

знакового пространства D обусловлена тем фактом, что производить классификацию в исходном пространстве Ω затруднительно, так как области Ω_l не являются компактными, а размерность самого пространства часто оказывается достаточно высокой.

Преобразование ϕ позволяет, с одной стороны, снизить размерность пространства признаков, а с другой стороны, добиться удобного для классификации вида областей $\phi(\Omega_l)$. Заметим, что выполняются следующие соотношения:

$$D = \bigcup_{i=0}^{L-1} D_i, \quad D_i \cap D_j = \emptyset, \quad i \neq j, \quad i, j = \overline{0, L-1},$$

и, соответственно, в силу однозначности преобразования ϕ ,

$$\Omega = \bigcup_{i=0}^{L-1} \Theta_i, \quad \Theta_i \cap \Theta_j = \emptyset, \quad i \neq j, \quad i, j = \overline{0, L-1}.$$

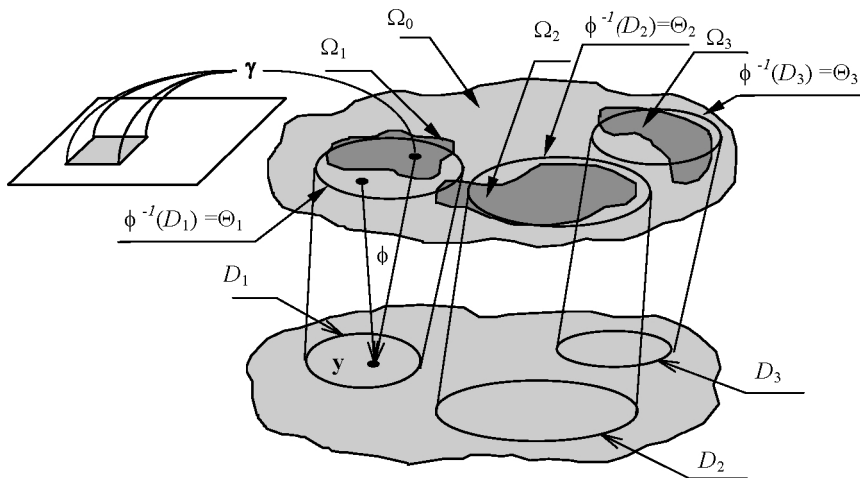


Рис. 6.5. Схема процесса распознавания в признаковом пространстве

В статистической теории распознавания описание классов производится путем задания или определения функций правдоподобия в признаковом пространстве для каждого из классов [17,33,40]. Отойдём от подобного традиционного подхода. Будем производить все

рассуждения в пространстве Ω , но при этом рассматривать его как обычное признаковое пространство. В этом случае для описания любого из классов достаточно знать соответствующую область Ω_l , $l = \overline{0, L-1}$, а также единственную безусловную плотность вероятности:

$$f(\gamma) = \sum_{l=0}^{L-1} P(\Omega_l) p(\gamma/\Omega_l),$$

где $p(\gamma/\Omega_l)$, $l = \overline{0, L-1}$ - функция правдоподобия l -ого класса в признаковом пространстве фрагментов данных, $P(\Omega_l)$, $l = \overline{0, L-1}$ - некоторая числовая величина, задающая априорную вероятность появления объектов l -ого класса. Действительно, функция правдоподобия какого-либо класса легко может быть получена следующим образом:

$$p\left(\frac{\gamma}{\Omega_l}\right) = \begin{cases} f(\gamma)/P(\Omega_l), & \gamma \in \Omega_l, \\ 0, & \gamma \notin \Omega_l, \end{cases} \quad l = \overline{0, L-1}.$$

А значение соответствующей априорной вероятности равно:

$$\begin{aligned} P(\Omega_l) &= P(\Omega_l) \cdot 1 = P(\Omega_l) \cdot \int_{\Omega_l} p(\gamma/\Omega_l) d\gamma = \\ &= \int_{\Omega_l} P(\Omega_l) p(\gamma/\Omega_l) d\gamma = \int_{\Omega_l} p(\gamma) d\gamma = P(\mathbf{Y} \in \Omega_l) \end{aligned}$$

Тогда выражение для вероятностей классификации из класса в класс, задаваемое в обычном признаковом пространстве следующим образом

$$p_{ij} = \int_{D_i} p(\mathbf{y}/\Omega_j) d\mathbf{y}, \quad i, j = \overline{0, L-1}$$

в нашем случае преобразуется к виду:

$$\begin{aligned}
p_{ij} &= \int_{\Theta_i} p(\gamma/\Omega_j) d\gamma = \frac{1}{P(\Omega_j)} \int_{\Theta_i} P(\Omega_j) p(\gamma/\Omega_j) d\gamma = \\
&= \frac{1}{P(\Omega_j)} \int_{\Theta_i \cap \Omega_j} p(\gamma) d\gamma = \frac{P(\mathbf{Y} \in \Theta_i \cap \Omega_j)}{P(\Omega_j)} = \\
&= \frac{P((\mathbf{Y} \in \Theta_i) \cdot (\mathbf{Y} \in \Omega_j))}{P(\mathbf{Y} \in \Omega_j)} = P(\mathbf{Y} \in \Theta_i / \mathbf{Y} \in \Omega_j), \\
i, j &= \overline{0, L-1}.
\end{aligned}$$

Здесь « \cdot » - произведение событий, а окончательное выражение для вероятности задает обычную условную вероятность. Очевидное преимущество предлагаемого описания состоит в том, что и классы, и классификатор описываются в исходном пространстве единообразно, используя области этого пространства Ω_i и Θ_i , соответственно.

Введем теперь следующие обозначения событий:

$$\begin{aligned}
Q_l &= \{\mathbf{Y} \in \Theta_l\} = \{\phi(\mathbf{Y}) = \mathbf{Y} \in D_l\}, & l = \overline{0, L-1}, \\
W_l &= \{\mathbf{Y} \in \Omega_l\},
\end{aligned} \tag{6.7}$$

здесь Q_l - событие, состоящее в принятии классификатором решения о принадлежности случайного вектора признаков \mathbf{Y} области D_l , то есть об отнесении случайного фрагмента \mathbf{Y} к l -ому классу. Тогда записанное выражение для вероятностей классификации переписывается в виде:

$$p_{ij} = P(\mathbf{Y} \in \Theta_i / \mathbf{Y} \in \Omega_j) = P(Q_i / W_j), \quad i, j = \overline{0, L-1}.$$

Будем использовать введенные обозначения в рамках задачи распознавания (второй этап). Для задачи обнаружения (первый этап) примем:

$$L = 2, Q = Q_1, \bar{Q} = Q_0, \Theta = \Theta_1, \bar{\Theta} = \Theta_0, D = D_1, \bar{D} = D_0. \tag{6.8}$$

Заметим, что в ситуации, когда для распознавания используется одно и то же признакововое пространство, предлагаемое описание использовать нецелесообразно. Однако оно очень удобно, если имеется несколько признакововых пространств и несколько классификаторов. Действительно, пусть классификатор первого этапа, подразделяющий пространство признаков на области «объекты» - «фон» и решающий задачу обнаружения, связан с событием классификации Q и соответствующей областью Θ , которые определены в формуле (6.8). Событие Q , очевидно, в подобной ситуации интерпретируется как решение об отнесении вектора признаков к классу объектов, а \bar{Q} - к классу фона.

Для алгоритма классификации (распознавания), используемого на втором этапе, в качестве событий и соответствующих областей примем Q_i и Θ_i , определенные формулой (6.7). Событие Q_0 , таким образом, интерпретируется как решение об отнесении вектора признаков к классу фона классификатором второго этапа технологии. Заметим, что в общем случае события Q и Q_0 не являются несовместными (может выполняться неравенство $P(QQ_0) \neq 0$), то есть заведомо допускается попадание «фоновых» фрагментов на второй этап технологии.

В рамках вышеуказанных соглашений вероятности классификации при обнаружении и распознавании в двухэтапной процедуре задаются следующим образом:

$$\begin{aligned} p_{ij} &= P(QQ_i/W_j), \quad i = \overline{1, L-1}, \quad j = \overline{0, L-1} \\ p_{0j} &= P(\bar{Q}/W_j) + P(QQ_0/W_j), \end{aligned} \quad (6.9)$$

Здесь произведение событий QQ_i означает, что случайный фрагмент данных Y классификатор первого этапа отнес к классу объектов, и одновременно с этим классификатор второго этапа определил его к i -му классу. При отнесении случайного фрагмента к классу фона возможно, с одной стороны, сделать это на этапе предварительного обнаружения (событие \bar{Q}), с другой стороны, проклассифицировать

его в нулевой класс на втором этапе (событие QQ_0). Поскольку события QQ_0 и \bar{Q} несовместны, общая вероятность отнесения случайного фрагмента к классу «фон» может быть записана как сумма вероятностей первого и второго случаев.

В дополнение к качественным показателям процесса распознавания введем в рассмотрение характеристику, характеризующую вычислительную сложность процедуры. Заметим, что традиционно в обработке изображений [9, 10, 12, 28] оценка сложности алгоритмических и программных средств производится по формуле

$$U_a + \eta U_m, \quad (6.10)$$

где U_a , U_m - количество, соответственно, сложений и умножений, используемых для обработки или классификации одного фрагмента изображения, то есть сложность на отсчет изображения; η - коэффициент относительной сложности операции умножения по отношению к операции сложения, типичное значение которого для наиболее массовых компьютеров лежит в диапазоне (1...3).

В единицах (6.10) средняя вычислительная сложность одного акта обнаружения и распознавания составляет:

$$U = U_I + P(Q)U_{II}, \quad (6.11)$$

где U_I и U_{II} , соответственно, сложность первого и второго этапов технологии.

Разумный компромисс между качеством решения задачи и затраченными вычислительными ресурсами достигается, если рассматривать величины сложности и качества либо как составляющие некоторой целевой функции, которая отражает приемлемость выбора каждой конкретной пары, либо как критерии оптимальности и ограничения в задаче оптимизации технологии. Наиболее часто встречающейся задачей *параметрической оптимизации*, используемой при построении алгоритма распознавания,

является задача улучшения его качества при одновременном ограничении на сложность процесса принятия решения [1, 3, 10, 11, 50]:

$$\begin{cases} R(\mathbf{a}) \rightarrow \min, \\ \mathbf{a} \\ U(\mathbf{a}) \leq U_{\text{lim}}. \end{cases} \quad (6.12)$$

Здесь U_{lim} - верхняя граница сложности, \mathbf{a} - вектор параметров алгоритма, по которому производится оптимизация.

6.3.3 Параметрическая оптимизация двухэтапной последовательной процедуры классификации

Пусть решается задача параметрической оптимизации двухэтапной процедуры в виде (6.12) по вектору параметров

$$\mathbf{a} = (\mathbf{a}_I, \mathbf{a}_{II}) ,$$

где \mathbf{a}_I и \mathbf{a}_{II} задают, соответственно, параметры первого и второго этапов. Тогда, учитывая выражения (6.9), (6.11) и (6.12) задача оптимизации переписывается в виде:

$$\begin{cases} \sum_{j=0}^{L-1} r_{0j} + \sum_{i,j=0}^{L-1} (r_{ij} - r_{0j}) P(W_j) P(Q(\mathbf{a}_I) Q_i(\mathbf{a}_{II}) / W_j) \rightarrow \min_{(\mathbf{a}_I, \mathbf{a}_{II})}, \\ P(Q(\mathbf{a}_I)) \leq \frac{U_{\text{lim}} - U_I}{U_{II}}. \end{cases}$$

Очевидно, что показатель сложности двухэтапной процедуры не содержит зависимости от совместного вектора параметров \mathbf{a} . В то же время показатель качества функционально с ним связан, а, следовательно, не может выступать в качестве критерия настройки ни одного из двух классификаторов-экспертов. Получаем противоречие - мы не можем настроить процедуру без настройки каждого из ее этапов, и не можем настроить ни один из этапов по критерию качества процедуры в целом. Для того чтобы изменить эту ситуацию, воспользуемся

возможностью представления вероятностей классификации из класса в класс в виде двух альтернативных выражений:

$$\begin{aligned} P(QQ_i/W_j) &= P(Q/W_j)P(Q_i/QW_j), \\ P(QQ_i/W_j) &= P(Q_i/W_j)P(Q/Q_iW_j), \end{aligned} \quad i, j = \overline{0, L-1} \quad (6.13)$$

Пусть теперь в процессе оптимизации получено некоторое решение $(\mathbf{a}_I^{fix}, \mathbf{a}_{II}^{fix})$. Зафиксируем какой-либо один из этапов, например второй, и будем изменять параметры другого. Тогда получим:

$$\begin{aligned} P(Q(\mathbf{a}_I)Q_i(\mathbf{a}_{II}^{fix})/W_j) &= P(Q_i(\mathbf{a}_{II}^{fix})/W_j)P(Q(\mathbf{a}_I)/Q_i(\mathbf{a}_{II}^{fix})W_j), \\ i, j &= \overline{0, L-1}. \end{aligned}$$

То есть изменение вероятности классификации из класса в класс пропорционально изменению вероятности $P(Q(\mathbf{a}_I)/Q_i(\mathbf{a}_{II}^{fix})W_j)$, причем событие, стоящее как условие, является фиксированным в процессе изменения параметров первого этапа. Аналогично, фиксированным оказывается и коэффициент пропорциональности изменения $P(Q_i(\mathbf{a}_{II}^{fix})/W_j)$. Точно также, в соответствии с первым выражением в (6.13), можно получить величину изменения искомой вероятности при фиксированном первом этапе.

Таким образом, используя представление классификатора как события, можно, фиксируя то первое, то второе из них, находить вектор оптимальных изменений и производить соответствующую коррекцию, что в целом должно привести к решению задачи настройки последовательной процедуры классификации. Заметим, что подобный подход требует итерационного представления алгоритма оптимизации. При этом указанные возможности позволяют построить итерационные алгоритмы различных типов. В настоящем разделе рассматриваются два из них.

6.3.4. Итерационный алгоритм параметрической оптимизации на основе градиентного метода

Для построения итерационного алгоритма оптимизации на основе градиентного метода необходимым является наличие возможности вычисления действительного значения вектора изменения параметров. Для этого требуется вычисление производных вида

$$\frac{\partial P(Q(\mathbf{a}_I)Q_i(\mathbf{a}_{II})/W_j)}{\partial \mathbf{a}}$$

Учитывая выражение (6.13), получаем следующий способ их вычисления в некоторой точке $(\mathbf{a}_I^{fix}, \mathbf{a}_{II}^{fix})$:

$$\begin{aligned} & \left. \frac{\partial P(Q(\mathbf{a}_I)Q_i(\mathbf{a}_{II})/W_j)}{\partial \mathbf{a}_I} \right|_{(\mathbf{a}_I, \mathbf{a}_{II})=(\mathbf{a}_I^{fix}, \mathbf{a}_{II}^{fix})} = \\ & = P(Q_i(\mathbf{a}_{II})/W_j) \left. \frac{\partial P(Q(\mathbf{a}_I)/Q_i(\mathbf{a}_{II})W_j)}{\partial \mathbf{a}_I} \right|_{\mathbf{a}_I=\mathbf{a}_I^{fix}}, \\ & \left. \frac{\partial P(Q(\mathbf{a}_I)Q_i(\mathbf{a}_{II})/W_j)}{\partial \mathbf{a}_{II}} \right|_{(\mathbf{a}_I, \mathbf{a}_{II})=(\mathbf{a}_I^{fix}, \mathbf{a}_{II}^{fix})} = \\ & = P(Q(\mathbf{a}_I)/W_j) \left. \frac{\partial P(Q_i(\mathbf{a}_{II})/Q(\mathbf{a}_I)W_j)}{\partial \mathbf{a}_{II}} \right|_{\mathbf{a}_{II}=\mathbf{a}_{II}^{fix}}. \end{aligned}$$

Зная теперь значения производных, можно воспользоваться градиентным или любым другим методом оптимизации первого порядка [2, 15, 22].

Использовать тот же подход для получения матрицы вторых производных не представляется возможным, поскольку невозможно записать выражение для смешанных производных. Как следствие, метод Ньютона и другие методы оптимизации второго порядка не могут быть использованы [2].

К сожалению, воспользоваться алгоритмом оптимизации на основе градиентного метода достаточно сложно. Основной проблемой является то, что на практике вид зависимости вероятностей классификации от вектора параметров алгоритма классификации не может быть представлен в аналитическом виде, а значит и вычисление указанных производных невозможно. В тех редких случаях, когда получение градиента все же допускается, существуют дополнительные трудности применения градиентного алгоритма: учет ограничений в задаче оптимизации; выбор шага при коррекции параметров, высокая вычислительная сложность такого алгоритма применительно к задаче оптимизации процедуры распознавания. Все это делает использование итерационного алгоритма оптимизации на основе градиентного подхода малоэффективным или невозможным в большинстве реальных задач. Поэтому необходима его модификация. Эта модификация, исходя из перечисленных недостатков градиентного алгоритма, должна касаться процесса модификации параметров. А именно, если на очередной итерации происходит полная коррекция вектора параметров каждого из этапов, то можно без каких бы то ни было затруднений использовать известные алгоритмы оптимизации. Учет ограничений по сложности в таком случае также не является проблематичным.

Таким образом, итерационный алгоритм параметрической оптимизации двухэтапной последовательной процедуры классификации может быть представлен следующим образом.

6.3.5 Итерационный алгоритм поэтапной параметрической оптимизации

- Решается задача оптимизации алгоритма распознавания второго этапа (второго эксперта) в виде

$$\sum_{j=0}^{L-1} r_{0j} + \sum_{i,j=0}^{L-1} (r_{ij} - r_{0j}) P(W_j) P(Q(\mathbf{a}_I)/W_j) P(Q_i(\mathbf{a}_I)/Q(\mathbf{a}_I)W_j) \rightarrow \min_{\mathbf{a}_I}. \quad (6.14)$$

- Решается задача оптимизации алгоритма распознавания первого этапа (первого эксперта) в виде:

$$\begin{cases} \sum_{j=0}^{L-1} r_{0j} + \sum_{i,j=0}^{L-1} (r_{ij} - r_{0j}) P(W_j) P(Q_i(\mathbf{a}_I)/W_j) P(Q(\mathbf{a}_I)/Q_i(\mathbf{a}_I)W_j) \rightarrow \min_{\mathbf{a}_I}, \\ P(Q(\mathbf{a}_I)) \leq \frac{U_{\text{lim}} - U_I}{U_{II}}. \end{cases} \quad (6.15)$$

- Если точность решения задачи не удовлетворяет, то возвращаемся к решению задачи оптимизации второго этапа.

Для определенности будем считать, что алгоритм прекращает свою работу в ситуации, если новые изменения величины риска R и сложности $P(Q)$ не превышают некоторых изначально заданных величин. В качестве начального состояния в данном алгоритме выбирается такое состояние классификатора первого этапа, когда $Q=W_0$, то есть на первой итерации в качестве условий будут выступать только события W_j .

Сделаем некоторые замечания относительно задач (6.14) и (6.15). При решении задачи распознавания число сформированных классов не превышает их количество в первоначальной задаче. В задаче же обнаружения (6.15) число всех условных событий составляет L^2 , то есть квадратично зависит от первоначального числа классов. Такое увеличение их числа может привести к неэффективному решению задачи - известным является факт повышения значения среднего риска при увеличении числа классов и неизменном пространстве признаков [40]. Поэтому при решении каждой конкретной задачи необходимо, учитывая конкретные значения штрафов r_{ij} , производить модификацию выражения (6.15) для снижения числа условных событий. Это может быть эффективно сделано, если матрица штрафов является простейшей.

6.3.6. Итерационный алгоритм поэтапной параметрической оптимизации в случае простейшей матрицы штрафов.
Итерационные схемы алгоритма

Пусть матрица штрафов является простейшей. В этом случае задачи оптимизации (6.14) и (6.15) меняются. А именно, задача (6.14) может быть записана в виде:

$$P(W_0)P(Q(\mathbf{a}_I)/W_0) + \sum_{j=0}^{K-1} P(W_j)P(Q(\mathbf{a}_I)/W_j)P(Q_i(\mathbf{a}_{II})/Q(\mathbf{a}_I)W_j) \rightarrow \min_{\mathbf{a}_{II}}. \quad (6.16)$$

Для задачи (6.15) показатель качества можно переписать следующим образом:

$$\begin{aligned} R &= \sum_{j=1}^{L-1} P(W_j)P(\bar{Q}/W_j) + \sum_{j=0}^{L-1} \sum_{\substack{i=0, \\ i \neq j}}^{L-1} P(W_j)P(QQ_i/W_j) = \\ &= P\left(\bar{Q} \cdot \bigcup_{j=1}^{L-1} W_j\right) + \sum_{j=0}^{L-1} P\left(QW_j \bigcup_{\substack{i=1, \\ i \neq j}}^{L-1} Q_i\right) = P\left(\bar{Q} \cdot \bigcup_{j=1}^{L-1} W_j\right) + \sum_{j=0}^{L-1} P(QW_j \bar{Q}_j) = \\ &= P\left(\bar{Q} \cdot \bigcup_{j=1}^{L-1} W_j\right) + P\left(Q \bigcup_{j=0}^{L-1} W_j \bar{Q}_j\right). \end{aligned}$$

Окончательно:

$$\left\{ \begin{aligned} &P\left(\bar{Q}(\mathbf{a}_I) \cdot \bigcup_{j=1}^{L-1} W_j\right) + P\left(Q(\mathbf{a}_I) \bigcup_{j=0}^{L-1} W_j \bar{Q}_j(\mathbf{a}_{II})\right) \rightarrow \min_{\mathbf{a}_{II}}, \\ &P(Q(\mathbf{a}_I)) \leq \frac{U_{\text{lim}} - U_I}{U_{II}}. \end{aligned} \right. \quad (6.17)$$

Очевидно, что показатель качества этой задачи может быть представлен в виде

$$P(\bar{Q} \cdot V_1) + P(Q \cdot V_0),$$

где

$$V_1 = \bigcup_{j=1}^{L-1} W_j, \quad V_0 = \bigcup_{j=0}^{L-1} W_j \bar{Q}_j. \quad (6.18)$$

В то же время, используя другие преобразования, возможно получение и других форм представления критерия (6.17). Некоторые из вариантов приведены в табл. 6.3.

Аналогичным образом может быть изменено и выражение (6.16). Подобная неоднозначность представления задачи оптимизации в общем случае может привести к множественности результатов ее решения. Однако если на каждой итерации выносятся оптимальное или байесовское решение, то результат оптимизации не будет зависеть от выбора итерационной схемы.

Таблица 6.3. Формы представления составляющих критериев качества

Итерационная схема	V_0	V_1
1	W_0	$\bigcup_{j=0}^{L-1} W_j Q_j$
2	$W_0 \bar{Q}_0$	$\bigcup_{j=1}^{L-1} W_j Q_j$
3	$\bigcup_{j=0}^{L-1} W_j \bar{Q}_j$	$\bigcup_{j=1}^{L-1} W_j$
4	$\bigcup_{j=0}^{L-1} \bar{W}_j Q_j$	$\bigcup_{j=1}^{L-1} W_j$

Покажем это на примере второй и третьей схем. Для первой из них запишем байесовский классификатор. Как известно, его разделяющая граница определяется из неравенства [17, 33, 40]:

$$P\left(W_0 \bigcup_{l=1}^{L-1} Q_l\right) P\left(\mathbf{y} / W_0 \bigcup_{l=1}^{L-1} Q_l\right) > P\left(\bigcup_{l=1}^{L-1} W_l Q_l\right) P\left(\mathbf{y} / \bigcup_{l=1}^{L-1} W_l Q_l\right).$$

Граница не изменится, если к правой и левой частям неравенства добавить одну и ту же величину $P\left(\bigcup_{l=1}^{L-1} W_l \bar{Q}_l\right) P\left(\mathbf{y} / \bigcup_{l=1}^{L-1} W_l \bar{Q}_l\right)$. Учитывая, что присутствующее в этом выражении событие несовместно ни с одним из событий третьей схемы, получим:

$$P\left(\left(W_0 \bigcup_{l=1}^{L-1} Q_l\right) \bigcup_{l=1}^{L-1} \left(\bigcup_{l=1}^{L-1} W_l \bar{Q}_l\right)\right) f\left(\mathbf{y} / \left(W_0 \bigcup_{l=1}^{L-1} Q_l\right) \bigcup_{l=1}^{L-1} \left(\bigcup_{l=1}^{L-1} W_l \bar{Q}_l\right)\right) > \\ < P\left(\left(\bigcup_{l=1}^{L-1} W_l Q_l\right) \bigcup_{l=1}^{L-1} \left(\bigcup_{l=1}^{L-1} W_l \bar{Q}_l\right)\right) P\left(\mathbf{y} / \left(\bigcup_{l=1}^{L-1} W_l Q_l\right) \bigcup_{l=1}^{L-1} \left(\bigcup_{l=1}^{L-1} W_l \bar{Q}_l\right)\right).$$

Из последнего неравенства следует:

$$P\left(\bigcup_{l=0}^{L-1} W_l \bar{Q}_l\right) P\left(\mathbf{y} / \bigcup_{l=0}^{L-1} W_l \bar{Q}_l\right) > P\left(\bigcup_{l=1}^{L-1} W_l\right) P\left(\mathbf{y} / \bigcup_{l=1}^{L-1} W_l\right),$$

что задает выражение разделяющей границы байесовского классификатора для третьей итерационной схемы. Аналогичным образом можно показать эквивалентность, с точки зрения байесовского решающего правила, остальных итерационных схем. В то же время, при использовании классификатора, отличного от байесовского, результаты оптимизации будут зависеть от вида задачи, а следовательно, и от итерационной схемы. Значит, для нахождения наилучшего решения необходимо также исследовать результаты оптимизации для каждой из схем и выбрать наилучшую для решения задачи.

Вернемся к вопросу размерности задач распознавания и обнаружения - числу классов в каждой из них. Очевидно, что поскольку алгоритм обнаружения может быть сведен к виду (6.18), где V_0 и V_1 выполняют роль событий принадлежности к соответствующему классу,

то и число классов в задаче обнаружения снижено до двух. Число классов в алгоритме распознавания (6.16), как и в (6.14), осталось идентичным числу классов в первоначальной задаче.

6.3.7. Результаты экспериментальных исследований

Для исследования предложенного итерационного алгоритма оптимизации двухэтапной последовательной процедуры классификации рассмотрим задачу поиска и распознавания локальных объектов на изображении.

Характер изображений в этой задаче и специфика процесса их анализа позволяют говорить об удовлетворении принятой ранее модели данных. А именно, для задачи поиска и распознавания локальных объектов на изображении характерно, что [1, 3, 9, 10, 11, 50]:

- подавляющую часть изображения занимает не интересующая нас «фоновая» составляющая;
- интересующие нас на изображении локальные области - объекты - малы по сравнению с размерами самого изображения;
- расположение объектов в плоскости изображения таково, что для каждого из них можно указать такое положение некоторой выпуклой наперед заданной области, что никакой другой объект в ней не будет присутствовать;
- количество содержащихся на изображении объектов мало по сравнению с общим количеством отсчетов изображения.

В рамках данной модели одним из наиболее удобных способов анализа данных является их локальная обработка [10, 11, 63]. В задачах обнаружения и распознавания локальных объектов на изображении при каждом положении окна обработки анализируемый фрагмент рассматривается как отдельное изображение, для которого и производится классификация.

Для экспериментальных исследований были выбраны изображения, содержащие объекты двух классов. Изображение, приведенное на рисунке рис.6.6а использовалось для обучения и оценки результатов классификации. В качестве признаков в работе были выбраны следующие:

- на первом этапе - два локальных средних с различными размерами окна усреднения (15*15, 25*25), вычислительная сложность их расчета составляет 8 операций на отсчет [22];
- на втором этапе - пять моментных инвариантов (см. п.4.3), вычисленных по градиенту выделенного фрагмента (25*25) изображения с использованием алгоритма прямой свертки; вычислительная сложность расчета инвариантов составляет приблизительно 6300 операций на отсчет изображения.

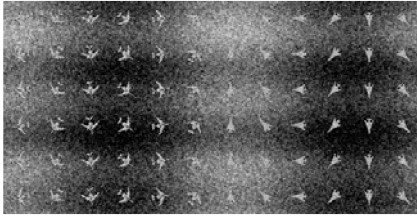
В качестве классификаторов в работе были использованы:

- на первом этапе - линейный классификатор, рассчитываемый с использованием процедуры Петерсона-Матсона; вычислительная сложность его использования для классификации составляет 4 операции на отсчет (см. п. 5.4);
- на втором этапе - квадратичный классификатор, вычисляемый в предположении нормальности закона распределения вектора признаков на втором этапе, вычислительная сложность его использования для классификации изображения при двух классах объектов и одном классе фона составляет приблизительно 200 операций на отсчет (см. п. 2.2).

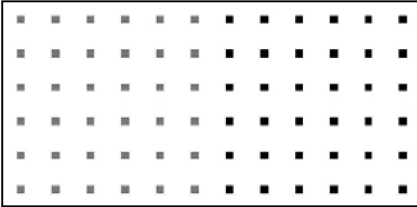
Таким образом, в проведенных исследованиях:

$$U_I=12, U_{II}=6500.$$

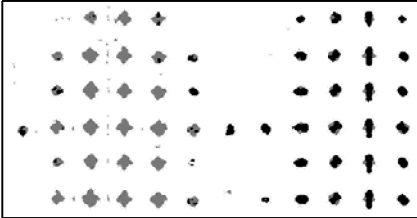
Сравнение качества работы оптимизированной двухэтапной процедуры производилось с наилучшим достижимым качеством классификатора второго этапа. Ограничения на сложность процедуры не задавались. Процесс оптимизации приведен на рис.6.6в –рис. 6.6з. Сравнение показателей работы приведено на диаграммах рис. 6.6ж - рис. 6.6з. Оно позволяет утверждать о приблизительно двукратном выигрыше в качестве распознавания при использовании двухэтапной процедуры. При этом до 15% выигрыша достигалось за счет использования итерационного алгоритма оптимизации. Заметим, что также произошло снижение сложности процесса принятия решения: по сравнению со сложностью алгоритма классификации второго этапа значение вычислительной сложности снизилось приблизительно в 10 раз (см. рис. 6.6з).



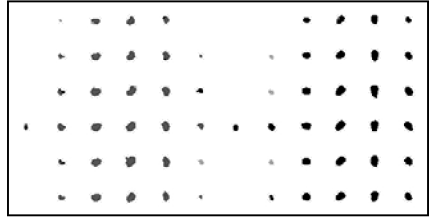
а) входное изображение



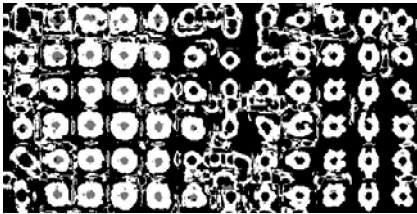
б) требуемая классификация



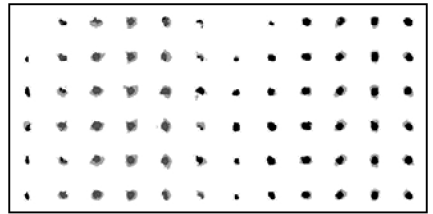
в) наилучший результат распознавания классификатора второго этапа (риск равен 0,05634) и его состояние на первой итерации алгоритма оптимизации



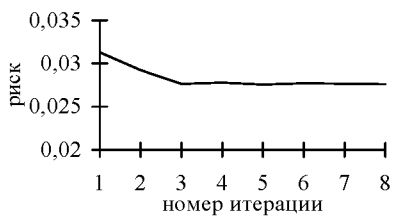
г) результаты работы двухэтапной процедуры классификации на первой итерации алгоритма оптимизации



д) результат работы классификатора второго этапа на последней итерации алгоритма оптимизации



е) результаты работы двухэтапной процедуры классификации на последней итерации алгоритма оптимизации



ж) величина среднего риска
при распознавании



з) вычислительная сложность
при распознавании

Рис. 6.6. Экспериментальные результаты оптимизации двухэтапной последовательной процедуры классификации

7. ЗАДАЧИ, УПРАЖНЕНИЯ, КОНТРОЛЬНЫЕ ВОПРОСЫ

7.1. Задачи

Задача 1.

Рассматривается задача разделения образов на два класса Ω_0 и Ω_1 , каждый из которых описывается нормальным законом распределения в одномерном пространстве

$$p(x/\Omega_0) = N(m_i, \sigma_i^2): \begin{matrix} m_0 = 0, & \sigma_0 = 2, \\ m_1 = 1, & \sigma_1 = 2. \end{matrix}$$

Матрица штрафов - простейшая.

1. Построить классификатор Байеса для следующих случаев задания априорных вероятностей классов:

а) $P(\Omega_0) = P(\Omega_1)$,

б) $P(\Omega_0) = 2P(\Omega_1)$,

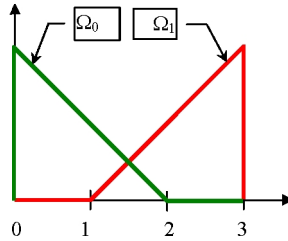
в) $P(\Omega_0) = \frac{1}{2}P(\Omega_1)$.

2. Построить минимаксный классификатор,

3. Построить классификатор Неймана-Пирсона для заданной вероятности ошибки первого рода $p_0^* = 0,05$ (при расчетах принять $\Phi^{-1}(1 - p_0^*) = 1,645$).

Задача 2.

Рассматривается задача разделения образов на два класса Ω_0 и Ω_1 с использованием одномерного вектора признаков. Законы распределения случайного вектора признаков в каждом из классов задаются графически.



1. Построить классификатор Байеса для априорных вероятностей $P(\Omega_0) = P(\Omega_1)$,
2. Построить минимаксный классификатор,
3. Построить классификатор Неймана-Пирсона для заданной вероятности ошибки первого рода $p_0^* = 0,25$.

Задача 3.

Рассматривается задача разделения образов на L классов Ω_l ($l = \overline{0, L-1}$), каждый из которых описывается распределением Рэлея:

$$p(x/\Omega_l) = \begin{cases} \frac{x^2}{\sigma_l^2} \exp\left(-\frac{x^2}{2\sigma_l^2}\right), & x \geq 0, \\ 0, & x < 0. \end{cases}$$

Параметры распределений в каждом из классов: $\sigma_0 = 1$, $\sigma_l = l\sigma_0$, ($l \geq 1$). Для равных априорных вероятностей и простейшей матрицы штрафов найти и изобразить графически байесовскую разделяющую границу для трех классов.

Задача 4.

Рассматривается задача разделения образов на два класса Ω_0 и Ω_1 , каждый из которых описывается нормальным законом распределения $p(\bar{x}/\Omega_0) = N(\bar{M}_l, B_l)$ в двумерном пространстве признаков, математические ожидания классов:

$$\bar{M}_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \bar{M}_1 = \begin{pmatrix} -1 \\ 0 \end{pmatrix}.$$

Рассматриваются два варианта постановки задачи с различными параметрами распределений:

$$\text{a) } B_0 = B_1 = \begin{pmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{pmatrix}; \text{ b) } B_0 = B_1 = \begin{pmatrix} 1 & 0,5 \\ 0,5 & 1 \end{pmatrix}.$$

1. Построить классификатор Байеса для равных априорных вероятностей $P(\Omega_0) = P(\Omega_1)$ и простейшей матрицы штрафов. Найти общую вероятность ошибочной классификации.

2. Построить минимаксный классификатор для простейшей матрицы штрафов, найти общую вероятность ошибочной классификации.

3. Построить классификатор Неймана-Пирсона для заданной вероятности ошибки первого рода $p_0^* = 0,05$ (при расчетах принять $\Phi^{-1}(1 - p_0^*) = 1,645$), найти общую вероятность ошибочной классификации.

Задача 5.

Рассматривается задача разделения образов на два класса Ω_0 и Ω_1 , каждый из которых описывается нормальным законом распределения $p(\bar{x}/\Omega_0) = N(\bar{M}_0, B_0)$ в двумерном пространстве признаков, математические ожидания классов:

$$\bar{M}_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \bar{M}_1 = \begin{pmatrix} -1 \\ 0 \end{pmatrix}.$$

$$B_0 = \begin{pmatrix} 1 & 0,5 \\ 0,5 & 1 \end{pmatrix}, \quad B_1 = \begin{pmatrix} 1 & -0,5 \\ -0,5 & 1 \end{pmatrix}.$$

Построить классификатор Байеса для равных априорных вероятностей $P(\Omega_0) = P(\Omega_1)$ и простейшей матрицы штрафов. Изобразить разделяющую границу графически.

Задача 6.

Рассматривается задача разделения образов на два класса Ω_0 и Ω_1 , каждый из которых описывается нормальным законом распределения

$p(\bar{x}/\Omega_0) = N(\bar{M}_1, B_1)$ в двумерном пространстве признаков с параметрами:

$$\bar{M}_0 = \bar{M}_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad B_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad B_1 = \begin{pmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{pmatrix} \quad \sigma \neq 1.$$

1. Построить классификатор Байеса для случая равных априорных вероятностей $P(\Omega_0) = P(\Omega_1)$ и простейшей матрицы штрафов. Изобразить разделяющую границу графически.

2. Найти выражения для вероятностей ошибочной классификации первого и второго рода.

Задача 7.

Рассматривается задача разделения образов на три класса Ω_0 , Ω_1 и Ω_2 , каждый из которых описывается нормальным законом распределения $p(\bar{x}/\Omega_0) = N(\bar{M}_1, B_1)$ в двумерном пространстве признаков с параметрами:

$$\bar{M}_0 = \bar{M}_1 = \bar{M}_2 = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

$$B_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad B_1 = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}, \quad B_2 = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}.$$

1. Построить классификатор Байеса для случая равных априорных вероятностей $P(\Omega_0) = P(\Omega_1) = P(\Omega_2)$ и простейшей матрицы штрафов. Изобразить разделяющую границу графически.

2. Найти выражения для вероятностей ошибочной классификации без вычисления интегралов.

Задача 8.

Рассматривается задача обнаружения объектов на фоне помех, то есть модель наблюдаемого вектора «сигнал+шум». Наблюдаемый «шум» соответствует классу Ω_0 , наблюдаемый «сигнал+шум» - классу

Ω_1 . Закон распределения наблюдаемого случайного вектора является нормальным с параметрами:

$$\bar{M}_0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \bar{M}_1 = \begin{pmatrix} a \\ a \end{pmatrix}, \quad B_0 = B_1 = \begin{pmatrix} \sigma^2 & \eta \\ \eta & \sigma^2 \end{pmatrix}.$$

Для случая равных априорных вероятностей $P(\Omega_0) = P(\Omega_1)$ и простейшей матрицы штрафов исследовать поведение суммарной вероятности ошибочной классификации (риска) от величин σ^2, η, a для классификатора Байеса.

Задача 9.

Показать, что увеличение числа классов, повышающее степень детализации распознаваемых объектов, при неизменном векторе (словаре) признаков увеличивает ошибку распознавания.

Примечание. Платежную матрицу принять в виде

$$c_{ij} = \begin{cases} c, & i \neq j, \\ 0, & i = j. \end{cases}$$

Рассмотреть выражение для риска в ситуации трех классов, затем объединить два произвольных класса в один. Посмотреть, как изменится выражение для риска при неизменном расположении разделяющих границ.

7.2. Упражнения

Читателю предлагается выполнить приведенные ниже упражнения в одной из сред математического программирования:

- MathCad,
- MatLab,
- Maple.

Для выполнения указанных упражнений, в частности выполнения заданий по синтезу выборочных данных, необходимы сведения по дисциплине «Теория вероятностей и математическая статистика» в объеме второго курса по специальности 010200 - прикладная математика и информатика.

7.2.1. Упражнение 1. Моделирование выборочных данных для решения задач распознавания

1. Разработать алгоритм моделирования нормально распределенного случайного вектора с заданными математическим ожиданием и корреляционной матрицей.
2. Смоделировать и изобразить графически обучающие выборки объема $N=200$ для двух нормально распределенных двумерных случайных векторов с заданными математическими ожиданиями и самостоятельно подобранными равными корреляционными матрицами. Варианты математических ожиданий случайного вектора приведены в п.7.2.4.
3. Смоделировать и изобразить графически обучающие выборки объема $N=200$ для трех нормально распределенных двумерных случайных векторов с заданными математическими ожиданиями и с неравными корреляционными матрицами, которые выбрать самостоятельно. Варианты математических ожиданий случайного вектора приведены в п.7.2.4.
4. На основании полученных выборок найти точечные оценки параметров нормального закона для каждого из распределений.

7.2.2. Упражнение 2. Построение классификаторов, реализующих оптимальные статистические стратегии распознавания для нормальных векторов признаков

1. Построить байесовскую решающую границу между классами Ω_0 и Ω_1 двумерных нормально распределенных векторов признаков для случая равных корреляционных матриц и равных априорных вероятностей и изобразить ее графически. Вычислить вероятности ошибочной классификации и суммарную вероятность ошибочной классификации в этом случае.
2. Построить минимаксный классификатор и классификатор Неймана-Пирсона для вероятности ошибки первого рода $p_0^* = 0,05$ для двух

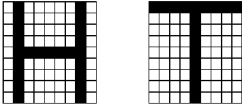
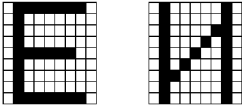
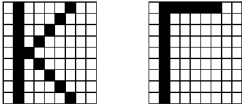
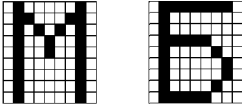
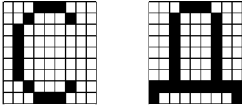
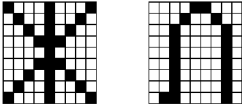
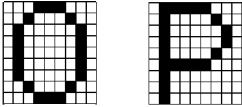
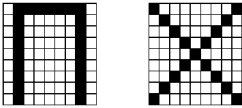
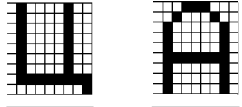
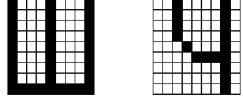
классов Ω_0 и Ω_1 двумерных нормально распределенных векторов признаков в случае равных корреляционных матриц. Изобразить решающие границы полученных классификаторов графически.

3. Построить байесовскую решающую границу между классами Ω_0 , Ω_1 и Ω_2 двумерных нормально распределенных векторов признаков для неравных корреляционных матриц и равных априорных вероятностей. Изобразить полученные решающие границы графически. Для любых двух классов оценить экспериментально вероятность ошибочной классификации в этом случае и определить относительную погрешность полученных оценок для заданного объема обучающей выборки N . Определить объем обучающей выборки, обеспечивающий получение оценок вероятностей ошибочной классификации с погрешностью не более 5%.

7.2.3 Упражнение 3. Построение классификатора для распознавания бинарного вектора признаков

1. Смоделировать обучающие выборки объема $N=200$ двух бинарных случайных векторов с распределениями, которые обеспечивают вероятность изменения указанной в представителе компоненты случайного вектора, равную $p = 0,3$. Формируемый для распознавания бинарный вектор представляет собой и горизонтальную развертку бинарного изображения буквы русского алфавита. Варианты представителей бинарных векторов приведены в п.7.2.4.
2. Произвести оценку распределений $P(\bar{x}/\Omega_l)$ ($l = 0,1$) бинарного случайного вектора на основании смоделированных данных.
3. Построить байесовскую разделяющую границу между классами Ω_0 и Ω_1 двумерных бинарных векторов признаков. Вычислить вероятности ошибочной классификации аналитически и оценить их экспериментально.

7.2.4. Варианты заданий при выполнении упражнений

№	Математические ожидания трех наборов нормально распределенных случайных векторов	Представители бинарных случайных векторов, \square \sim "0", \blacksquare \sim "1"
1.	$\bar{M}_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \bar{M}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \bar{M}_3 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$	
2.	$\bar{M}_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \bar{M}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \bar{M}_3 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$	
3.	$\bar{M}_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \bar{M}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \bar{M}_3 = \begin{pmatrix} -1 \\ -1 \end{pmatrix}.$	
4.	$\bar{M}_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \bar{M}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \bar{M}_3 = \begin{pmatrix} -1 \\ -1 \end{pmatrix}.$	
5.	$\bar{M}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \bar{M}_2 = \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \bar{M}_3 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$	
6.	$\bar{M}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \bar{M}_2 = \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \bar{M}_3 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$	
7.	$\bar{M}_1 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \bar{M}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \bar{M}_3 = \begin{pmatrix} -1 \\ -1 \end{pmatrix}.$	
8.	$\bar{M}_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \bar{M}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \bar{M}_3 = \begin{pmatrix} -1 \\ -1 \end{pmatrix}.$	
9.	$\bar{M}_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \bar{M}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \bar{M}_3 = \begin{pmatrix} -2 \\ 1 \end{pmatrix}.$	
10.	$\bar{M}_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \bar{M}_2 = \begin{pmatrix} -1 \\ -1 \end{pmatrix}, \bar{M}_3 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$	

$$11. \quad \bar{M}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \bar{M}_2 = \begin{pmatrix} -1 \\ -1 \end{pmatrix}, \quad \bar{M}_3 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$



$$12. \quad \bar{M}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \bar{M}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \quad \bar{M}_3 = \begin{pmatrix} 1 \\ -2 \end{pmatrix}.$$



$$13. \quad \bar{M}_1 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \quad \bar{M}_2 = \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \quad \bar{M}_3 = \begin{pmatrix} 2 \\ -1 \end{pmatrix}.$$



$$14. \quad \bar{M}_1 = \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \quad \bar{M}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad \bar{M}_3 = \begin{pmatrix} -1 \\ 2 \end{pmatrix}.$$



$$15. \quad \bar{M}_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \bar{M}_2 = \begin{pmatrix} -1 \\ -1 \end{pmatrix}, \quad \bar{M}_3 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$



$$16. \quad \bar{M}_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \bar{M}_2 = \begin{pmatrix} -1 \\ -1 \end{pmatrix}, \quad \bar{M}_3 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$



$$17. \quad \bar{M}_1 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \quad \bar{M}_2 = \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \quad \bar{M}_3 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$



$$18. \quad \bar{M}_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad \bar{M}_2 = \begin{pmatrix} 2 \\ 2 \end{pmatrix}, \quad \bar{M}_3 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$



$$19. \quad \bar{M}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \bar{M}_2 = \begin{pmatrix} 2 \\ -2 \end{pmatrix}, \quad \bar{M}_3 = \begin{pmatrix} -1 \\ -1 \end{pmatrix}.$$



$$20. \quad \bar{M}_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad \bar{M}_2 = \begin{pmatrix} -2 \\ -2 \end{pmatrix}, \quad \bar{M}_3 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$



$$21. \quad \bar{M}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \bar{M}_2 = \begin{pmatrix} -2 \\ 2 \end{pmatrix}, \quad \bar{M}_3 = \begin{pmatrix} -1 \\ -1 \end{pmatrix}.$$



$$22. \quad \bar{M}_1 = \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \quad \bar{M}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \quad \bar{M}_3 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$



$$23. \quad \bar{M}_1 = \begin{pmatrix} 0 \\ 2 \end{pmatrix}, \quad \bar{M}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad \bar{M}_3 = \begin{pmatrix} -2 \\ 0 \end{pmatrix}.$$



$$24. \quad \bar{M}_1 = \begin{pmatrix} -2 \\ 0 \end{pmatrix}, \quad \bar{M}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \bar{M}_3 = \begin{pmatrix} 0 \\ -2 \end{pmatrix}.$$



$$25. \quad \bar{M}_1 = \begin{pmatrix} 0 \\ -2 \end{pmatrix}, \quad \bar{M}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \quad \bar{M}_3 = \begin{pmatrix} 2 \\ 0 \end{pmatrix}.$$



7.3 Контрольные вопросы

1. Постановка задачи классификации.
2. Определение классификатора.
3. Способы задания классификатора.
4. Качество классификатора.
5. Байесовский классификатор.
6. Минимаксный классификатор.
7. Классификатор Неймана-Пирсона.
8. Отношение правдоподобия.
9. Байесовский классификатор для нормально распределенных векторов признаков.
10. Меры близости нормальных распределений.
11. Инвариантность расстояний к линейным преобразованиям.
12. Характер линейного преобразования, обеспечивающего инвариантность евклидова расстояния.
13. Байесовский классификатор для распознавания бинарных векторов признаков.
14. Вычисление вероятностей ошибочной классификации.
15. Процедура Петерсона-Матсона построения линейной дискриминантной функции, минимизирующей суммарную вероятность ошибочной классификации.
16. Перечислите основные группы признаков, используемых для распознавания.
17. Что такое геометрические признаки? Приведите их примеры.
18. Что такое топологические признаки? Приведите их примеры.
19. Что такое вероятностные признаки? Приведите их примеры.
20. Что такое спектральные признаки? Приведите их примеры.
21. Перечислите основные группы задач распознавания на изображениях. Охарактеризуйте их. Приведите примеры.

22. Перечислите и прокомментируйте основные требования к признакам, вычисляемым по изображениям.
23. Понятие нормализации вектора признаков. Основные типы нормализации при распознавании изображений.
24. Яркостная нормализация и метод определения ее параметров.
25. Рекурсивная реализация нормализующего преобразования.
26. В чем различие алгоритмов яркостной нормализации МНК, НСД, НКЛ.
27. Нормализация масштаба объекта.
28. Нормализация положения объекта.
29. Нормализация ориентации объекта.
30. Моментные инварианты. Метод построения моментных инвариантов.
31. Моментные инварианты. Метод расчета в скользящем окне.
32. Аффинные инварианты.
33. Постановка задачи обнаружения и локализации объектов на изображении.
34. Критерии локализации объектов.
35. Общее описание процедуры локализации. Методы ее настройки.
36. Алгоритм Петерсона-Матсона нахождения параметров линейной дискриминантной функции.
37. Модификации алгоритма Петерсона-Матсона.
38. Стратегии совместной классификации.
39. Параллельная стратегия совместной классификации при минимальной информации о решениях экспертов. Вид классификатора при независимых решениях экспертов.
40. Двухэтапная последовательная процедура классификации. Постановка задачи.
41. Двухэтапная последовательная процедура классификации. Итерационные алгоритмы поэтапной параметрической оптимизации.

СПИСОК ЛИТЕРАТУРЫ

1. Абчук В. А., Суздаль В. Г. *Поиск объектов* (М.: Сов. радио, 1977)
2. Амосов А.А., Дубинский Ю.А., Кошленова Н.В. *Вычислительные методы для инженеров* (М.: Высш. шк., 1994)
3. Анисимов Б.В., Курганов В.Д., Злобин В.К. *Распознавание и цифровая обработка изображений* (М.: Высш. шк., 1983)
4. Бапаринов А. Е., Флейшман Б. С. *Методы статистического последовательного анализа и их радиотехнические приложения* (М.: Сов. радио, 1962)
5. Вальд А. *Последовательный анализ* (М.: Физматгиз, 1960)
6. Буреев В.А. и др. Методы сокращения вычислительных затрат в задачах распознавания изображений. *Зарубежная радиоэлектроника* (4) 52-75 (1980)
7. Глумов Н.И. Построение и применение моментных инвариантов для обработки изображений в скользящем окне. *Деп. в ВИННИТИ, №1880-В94* (Самара: Самарский государственный аэрокосмический университет, 1994)
8. Глумов Н.И. Яркая нормализация цифрового изображения в скользящем окне. *Деп. в ВИННИТИ, №1881-В94* (Самара: Самарский государственный аэрокосмический университет, 1994)
9. Глумов Н.И., Егунев И.П., Коломиец Э.И., Мясников В.В., Сергеев В.В. Оптимизация информационной технологии обнаружения локальных объектов на изображении. *Распознавание образов и анализ изображений: новые информационные технологии: 2-ая Всероссийская с участием стран СНГ конференция* (Ульяновск, Часть 2, 91-93, 1995)
10. Глумов Н.И., Коломиец Э.И., Сергеев В.В. Информационная технология обнаружения объектов на изображении в режиме скользящего окна. *Научное приборостроение* (1) 72-88 (1993)
11. Глумов Н.И., Коломиец Э.И., Сергеев В.В. Информационная технология распознавания объектов на изображении в режиме скользящего окна *Тезисы докл. 6-й науч. конф. «Математические методы распознавания образов»* (Москва, 89-90, 1993)
12. Глумов Н.И., Сергеев В.В. Разложение двумерного сигнала по степенному и полиномиальным базисам в скользящем окне *Тезисы докладов 6-й научной конф. «Математические методы распознавания образов»* (Москва, 90-91, 1993)
13. Горелик А.Л., Скрипкин В.А. *Методы распознавания* (М.: Высш. шк., 1984)
14. Горелик А.Л., Гуревич И.Б., Скрипкин В.А. *Современное состояние проблемы распознавания* (М.: Высш. шк., 1985)
15. Давыдов Э.Г. *Исследование операций* (М.: Высш. шк., 1990)
16. Даджон Д., Мерсеро Р. *Цифровая обработка многомерных сигналов* (М.: Мир, 1988)
17. Дуда Р., Харт П. *Распознавание образов и анализ сцен* (М.: Мир, 1976)
18. Круглов В.Н., Лабунец В.Г. Быстрый алгоритм совмещения изображений, инвариантный к повороту и масштабу. *Деп. в ВИННИТИ, №5104-85* (Свердловск: Уральский политехнический институт, 1985)
19. Круглов В.Н., Лабунец В.Г. Распознавание образов при помощи модульных инвариантов моментов. *Деп. в ВИННИТИ, N 5105-85* (Свердловск: Уральский политехнический институт, 1985)
20. Леман Э. *Проверка статистических гипотез* (М.: Наука, 1979)
21. Майтра С. Моментные инварианты. *ТИИЭР* (4) 297-299 (1979)
22. Методы компьютерной обработки изображений. Под ред. В.А.Сойфера. – (М.: Физматлит, 2003)

23. Мясников В.В. Локализация объектов в задаче их распознавания на изображении *Сборник материалов 2-ой международной конференции «Расознавание 95»* (Курск, 88-89, 1995)
24. Мясников В.В. Оптимизация двухэтапной технологии распознавания объектов на изображении. *3-я Всероссийская с участием стран СНГ конференция «Расознавание образов и анализ изображений: Новые информационные технологии» (РОАИ-3-97)* (Нижний Новгород, Часть 1, 203-207, 1997)
25. Прэтт У.К. *Цифровая обработка изображений 1* (М.: Мир, 1982)
26. Прэтт У.К. *Цифровая обработка изображений 2* (М.: Мир, 1982)
27. Пулятин Е.П., Аверин С.И. *Обработка изображений в робототехнике* (М.: Машиностроение, 1990)
28. Сергеев В.В. Параллельно-рекурсивные КИХ-фильтры в задачах обработки изображений *Радиотехника* (8) 38-41 (1990)
29. Сергеев В.В., Чичева М. А. Некоторые модификации цифрового коррелятора для обнаружения объектов на изображении. *Компьютерная оптика* (5) 78-84 (1989)
30. Синдлер Ю.Б. *Метод двухступенчатого статистического анализа и его приложени в технике* (М.: Наука, 1973)
31. Сосулин Ю.Г., Гаврилов К.Ю. К-этапные процедуры проверки статистических гипотез. *2-ая Всероссийская с участием стран СНГ конференция «Расознавание образов и анализ изображений: новые информационные технологии»* (Ульяновск, Часть 1, 23-25, 1995)
32. Сосулин Ю.Г., Гаврилов К.Ю. Модифицированные К-этапные процедуры проверки статистических гипотез. *3-ая Всероссийская с участием стран СНГ конференция «Расознавание образов и анализ изображений: новые информационные технологии»* (Нижний Новгород, Часть 1, 84-88, 1997)
33. Ту Дж., Гонсалес Р. *Принципы распознавания образов* (М.: Мир, 1978)
34. Федотов Н.Г. *Методы стохастической геометрии в распознавании образов* (М.: Радио и связь, 1990)
35. Федотов Н.Г. Теория признаков распознавания образов, основанная на стохастической геометрии *Тезисы докладов III Конференции Расознавание образов и анализ изображений: новые информационные технологии (РОАИ-3-97)* (Нижний Новгород, 278-282, 1997)
36. Федотов Н.Г., Мельников М.М., Тужилов И.В., Костюпин Р.А., Шульга Л.А. Система распознавания изображений на основе методов стохастической геометрии. *Труды IV Конференции Расознавание образов и анализ изображений: новые информационные технологии (РОАИ-4-98)* (Новосибирск, 187-189, 1998)
37. Федотов Н.Г., Шульга Л.А. Перспективы развития теории автоматической генерации признаков в задачах распознавания образов. *Труды V Международной Конференции Расознавание образов и анализ изображений: новые информационные технологии (РОАИ-5-2000) 1* (Самара, 156-161, 2000)
38. Фор А. *Восприятие и распознавание образов* (М.: Машиностроение, 1989)
39. Фу К. *Последовательные методы в распознавании образов и обучении машин* (М.: Наука, 1971)
40. Фукунага К. *Введение в статистическую теорию распознавания образов* (М.: Наука, 1979)
41. Ярославский Л.П. *Введение в цифровую обработку изображений* (М.: Советское радио, 1979)
42. Хелгасон С. *Преобразование Радона* (М.: Мир, 1983)

43. Ярославский Л.П. Точность и достоверность измерения положения двумерного объекта на плоскости. *Радиотехника и электроника* (5) 17-25 (1972)
44. Ярославский Л.П. *Цифровая обработка сигналов в оптике и голографии: Введение в цифровую оптику* (М.: Радио и связь, 1987)
45. Abu-Mostafa Y., Psaltis D. Image Normalization by Complex Moments. *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-7**(1) 46-55 (1985)
46. Abu-Mostafa Y., Psaltis D. Recognitive Aspects of Moment Invariants. *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-6**(6) 698-706 (1984)
47. Aksak I., Feist Ch., Kijko V., Knoefel R., Matsello V., Oganovskij V., Schlesinger M., Schlesinger D., Stanke G. Detection of the Objects with Given Shape on the Grey-Valued Pictures. *Proceedings of 7-th International Conference on Computer Analysis of Images and Pattern* (Kiel, Germany, 551-558, Springer, 1997)
48. Dana H. Ballard, Christopher M. Brown *Computer vision* (Prentice-hall, Inc., Englewood Cliffs, New Jersey, 1982)
49. J.Cao, M.Ahmadi, & M.Shridhar Recognition of hand-written numerals with multiple features and multistage classifiers. *Pattern Recognition* **28**(2) 153-160 (1995)
50. Casasent D. New Techniques for Object Detection and Recognition. *Proceedings of The 10th Scandinavian Conference on Image Analysis II* (Lappeenranta, Finland, 597-604, 1997)
51. A.J.Elms A connected character recognizer using Level Building of HMMs. *Proceedings 12th IAPR International Conference and Neural Networks, Conference B: Pattern Recognition Methodology and Systems II* 439-441 (1994)
52. Flusser J. and Suk T. Affine moment invariants: a new tool for character recognition. *Pattern Recognition Letters* (15) 433-436 (1994)
53. Flusser J. and Suk T. A moment-base approach to registration of images with affine geometric distortion. *IEEE Transactions on Geoscience and Remote Sensing* **32**(2) 382-387 (1994)
54. Flusser J. On the independence of rotation moment invariants. *Pattern Recognition* **33**(9) 1405-1410 (2000)
55. Flusser J. and Suk T. Pattern recognition by affine moment invariants. *Pattern recognition* **26**(1) 167-174 (1993)
56. Flusser J., Suk T., Saic S. Image features invariant with respect to blur. *Pattern Recognition* **28** 1723-1732 (1995)
57. Flusser J., Suk T., Saic S. Recognition of blurred images by the method of moments. *IEEE Trans. Image Proc.* **5** 533-538 (1996)
58. Franke J., Mandler E. A Comparison of Two Approaches for Combining the Votes of Cooperating Classifiers. *Proceedings 11th IAPR International Conference on Pattern Recognition, Conference B: Pattern Recognition Methodology and Systems II* 611-614 (1992)
59. Freeman H. Computer processing of line drawing images *Computer Surveys* **6**(1) 57-98 (1974)
60. Freeman H. On the encoding of arbitrary geometric configurations. *IRE Trans. Elec. Comp.* **EC-10** 260-268 (1961)
61. Freeman H. Techniques for the digital computer analysis of chain-encoded arbitrary plane curves. *Proc. Natl. Elec. Conf.* (18) 312-324 (1961)
62. Gevers T., Smeulders W. Object Recognition based on Photometric Color Invariants *Proceedings of The 10th Scandinavian Conference on Image Analysis II* (Lappeenranta, Finland, 861-868, 1997)

63. Glumov N.I., Egunov I.P., Kolomiets E.I., Myasnikov V.V., Sergeyev V.V. Optimization of Information Technology for Detection of Local Objects on an Image. *Pattern Recognition and Image Analysis* **6**(1) 120-121 (1996)
64. Glumov N.I., Krainukov N.I., Sergeyev V.V., Khramov A.G. The Fast Algorithm of Image Approximation in a Sliding Window. *Pattern Recognition and Image Analysis* (4) 424-426 (1991)
65. A.Golovan, M.H.Yoo, S.W.Lee Pre-Attentive Detection of Perceptually Important Regions in Facial Images. *Proc. of the 15th Int. Conf. On Pattern Recognition* **1** (Barcelona, Spain, 1092-1095, 2000)
66. L.K.Hansen, P.Salamon Neural network ensembles. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **12**(10) 993-1001 (1990)
67. Hashem, B.Schmeiser Improving model accuracy using optimal linear combinations of trained neural networks. *IEEE Transactions on Neural Networks* **6**(3) 792-794 (1995)
68. Healey G. Recognition Objects in Color Images. *Proceedings of The 10th Scandinavian Conference on Image Analysis* **II** (Lappeenranta, Finland, 823-829, 1997)
69. Ho J., Hull J., Srihari S.N. Decision combination in multiple classifier systems. *IEEE Transaction Pattern Analysis and Machine Intelligence* **16**(1) 66-75 (1994)
70. T.K.Ho, J.J.Hull, S.N.Srihari Decision combination in multiple classifier systems. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **16**(1) 66-75 (1994)
71. Hu M. Visual pattern recognition by moment invariants. *IRE Trans. Information Theory* **IT-8** 179-187 (1962)
72. Kittler J., Hater M., Duin R. Combining Classifiers. *Proceedings of 13th International Conference on Pattern recognition. Track B: Pattern Recognition and Signal Analysis* **II** (Vienna, Austria, 897-901, 1996)
73. Lambert G., Noll J. Discrimination Property of Invariants Using the Line Moments of Vectorized Contours. *Proceedings of 13th International Conference on Pattern recognition. Track B: Pattern Recognition and Signal Analysis* **II** (Vienna, Austria, 735-739, 1996)
74. Li B. High-Order Moment Computation of Grey-Level Images. *IEEE Trans. on Image Processing* **4**(4) 502-505 (1995)
75. Li B., Shen J. Two-dimensional local moment, surface fitting and their fast computation. *Pattern Recognition* **27**(6) 785-790 (1994)
76. Maghsoodi R., Rezaie B. Image Registration Using a Fast Adaptive Algorithm. *Methods of Handling and Processing Imagery. SPIE* **757** 58-63 (1987)
77. D.Maio, D.Maltoni Direct Gray-Scale Minutiae Detection in Fingerprints. *IEEE Trans. Pattern Anal. Machine Intell.* **19**(1) 27-40 (1997)
78. Matuš F., Flusser J. Image Representations via a Finite Radon Transform. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **15**(10) 996-1006 (1993)
79. Myasnikov V.V. On the modified quality criterion for a procedure to detect objects in spatially-extended fields *Proceedings of the 10th Scandinavian Conference on Image Analysis SCIA'96* **1** (Lappeenranta, Finland, 405-410, 1997)
80. D.Nair, J.K.Aggarwal Hierarchical, Modular Architectures for Object Recognition by Parts. *Proc. 13th Int. Confer. On Pattern Recognition* **1** (Vienna, Austria, 601-606, 1996)
81. D.Nandy, J.Ben-Arie Recovery of 3-D Face Structure using Recognition. *Proc. of the 15th Int. Conf. On Pattern Recognition* **1** (Barcelona, Spain, 1104-1108, 2000)

82. Peterson D. Y., Mattson R. T. A method of finding linear discriminant functions for a class of performances criteria. *IEEE Trans. Information Theory* **IT-12**(Chapter 4) 380-387 (1966)
83. S.Prabhaker, A.K.Jain, J.Wang, S.Pankanti, R.Bolle Minutia Verification and Classification for Fingerprint Matching. *Proc. of the 15th Int. Conf. On Pattern Recognition 1* (Barcelona, Spain, 25-29, 2000)
84. G.Rogova Combining the results of several neural network classifiers. *Neural Networks* **7**(5) 777-781 (1994)
85. Suri J., Haralick R., Sheehan F., Jamin V. Effect of Edge Detection, Pixel Classification and Classification-Edge Fusion Over LV Calibration: A Two Stage Automatic System. *Proceedings of The 10th Scandinavian Conference on Image Analysis I* (Lappeenranta, Finland, 197-204, 1997)
86. Tung H., Lee J., Tsai J. Multi-stage pre-candidate selection in handwritten Chinese character recognition systems. *Pattern Recognition* **27**(8) 1093-1102 (1994)
87. Xu A., Krzyzak C., Suen Y. Methods of combining multiple classifiers and their applications to handwriting recognition. *IEEE Trans. SMC* **22**(3) 418-435 (1992)

Учебное издание

*Мясников Владислав Валерьевич
Глумов Николай Иванович
Сергеев Владислав Викторович*

**МЕТОДЫ ОБНАРУЖЕНИЯ И РАСПОЗНАВАНИЯ ОБЪЕКТОВ
НА ЦИФРОВЫХ ИЗОБРАЖЕНИЯХ**

Учебное пособие

Технический редактор *В.А. Фурсов*
Редакторская обработка *Т.К. Кретинина*
Корректорская обработка *О.Ю. Дьяченко*
Доверстка *А.В. Ярославцева*

Подписано в печать 28.12.06. Формат 60x84 1/16.
Бумага офсетная. Печать офсетная.
Усл. печ. л. 9,77. Усл. кр.-отт. 9,89. Печ.л. 10,5.
Тираж 50 экз. Заказ . ИП-54/2006

Самарский государственный
аэрокосмический университет.
443086 Самара, Московское шоссе, 34.

Изд-во Самарского государственного
аэрокосмического университета.
443086 Самара, Московское шоссе, 34.