

МАЛЫЦЕВ Альберт Владимирович

**РАЗРАБОТКА ПАРАМЕТРИЧЕСКИХ МОДЕЛЕЙ И
ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ РАСПОЗНАВАНИЯ ОБРАЗОВ
ПО ВЫБОРКЕ ФИКСИРОВАННОГО ОБЪЕМА С УЧЕТОМ
ПОГРЕШНОСТЕЙ ПРИЗНАКОВ**

05.13.18 – Теоретические основы математического моделирования, численные методы и комплексы программ

АВТОРЕФЕРАТ

диссертации на соискание ученой
степени кандидата технических наук

Москва 2000

Работа выполнена в научно-производственном закрытом акционерном обществе «Валко-Электроникс» г. Смоленск и на кафедре вычислительной математики и математической физики Московского государственного технического университета им. Н. Э. Баумана

Научный руководитель:

д.т.н., профессор А.А. Грешилов

Официальные оппоненты:

д.ф.-м.н., профессор А.Ф. Кушнир

к.т.н., доцент В.А. Суханов

Ведущая организация: 22 центральный научно-исследовательский
испытательный институт.

Защита состоится " 10 " октября 2000 г. в 11 час 30 мин на заседании
Диссертационного Совета Д 053.15.12 при Московском государственном
техническом университете им. Н. Э. Баумана по адресу: 107005, Москва, 2-я
Бауманская ул., д. 5.

С диссертацией можно ознакомиться в библиотеке МГТУ.

Отзывы в двух экземплярах, заверенные печатью учреждения, просим выслать
по адресу: 107005, Москва, 2-я Бауманская ул., д.5, МГТУ им. Н.Э.Баумана.
Ученому секретарю диссертационного Совета.

Автореферат разослан " " 2000 г.

Ученый секретарь
диссертационного Совета
д.ф.-м.н., профессор

 Волков И.К.

Типография МГТУ им. Н.Э.Баумана

Зак. № 111. Тираж 100 экз. Подп. к печати 10.10 2000г. Объем 1 печ. лист.

НТБ МГТУ им. Н. Э. Баумана



1568602

Мальцев А. В. Разработка па

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы. Статистические задачи распознавания образов характерны для различных областей знаний, в частности при распознавании землетрясений и маломощных взрывов. Актуальным является исследование процедур распознавания образов в условиях, характерных для их практической реализации.

В практической реализации необходимо учитывать погрешности измерений признаков, а также тот факт, что мы имеем небольшое число измерений в обучающей выборке. Таким образом, наряду с усложнением процедуры обучения системы распознавания из-за неопределенности исходной информации необходимо еще и сократить число параметров, подлежащих определению. Поэтому, при достаточном числе точек (измерений в обучающей выборке) будем считать, что нам известен вид функций условных плотностей вероятности, а требуется определить точечные и интервальные оценки их параметров, а затем точечные и интервальные оценки решающих (разделяющих) функций и ошибок I и II рода.

При небольшом числе точек будем предполагать, что известен вид решающих (разделяющих) функций, а необходимо найти точечные и интервальные оценки параметров этих функций, а затем точечные и интервальные оценки самих функций. В этом случае ошибки I и II рода получаются численным методом в процессе обучения (по числу наблюдений, приводящих к ложным решениям). Интервальные оценки решающих (разделяющих) функций определяют нулевую зону – зону невозможности принятия решения.

Учет погрешностей измерений признаков объектов приводит к нелинейным системам и к сложным алгоритмам, выбор которых определяется условиями конкретной задачи. В задачах распознавания образов возникает дополнительная проблема: по единичному наблюдению $X_{набл}$ найти оценку его истинного значения $\xi_{набл}$, по которой впоследствии и проводить идентификацию объекта. Для нахождения $\xi_{набл}$ приходится привлекать всю исходную информацию, используемую в процессе обучения. Эта процедура становится принципиально важной, когда область (эллипсоид) неопределенности $X_{набл}$ пересекается с нулевой зоной.

Неучет погрешностей в наблюдаемых значениях признаков приводит к следующему:

1) оценки функций условных плотностей вероятности признаков образов (если даже их вид известен априори) будут смещенными, а их интервальные оценки – неверными;

2) в процессе идентификации образов, имеющих близкие значения координат векторов признаков, из-за влияния на результаты наблюдений погрешностей может быть принято неверное решение;

3) в традиционных методах функции условных плотностей вероятности признаков выступают в процедурах идентификации образов как

детерминированные (не учитываются их интервальные оценки), и по этой причине в принятии решений не рассматриваются имеющие место зоны неопределенности (нулевые зоны).

В отличие от традиционных методов распознавания образов в работе предложены и используются методы, которые позволяют:

1) учесть погрешности наблюдаемых значений координат векторов признаков объектов;

2) получить несмещенные точечные и интервальные оценки функций условных плотностей вероятности признаков;

3) оценить "истинные" координаты вектора признаков, по которому будет проводиться идентификация объектов, и включить эти данные в процедуру принятия решений.

Решаемая задача представляет интерес для различных областей знаний и практических приложений и является актуальной.

Цель и задачи исследования. Целью диссертационной работы является разработка теоретических и практических методов распознавания образов по выборке фиксированного объема с учетом погрешностей результатов наблюдений признаков, когда алгоритм распознавания включает в себя стохастические параметрические модели (априори известен вид функций условных плотностей вероятности признаков или разделяющих функций). Разрабатываемые методы предназначены для получения точечных и интервальных оценок искомых параметров моделей и оценок "истинных" значений векторов признаков и учета этих оценок в процедурах распознавания образов.

В соответствии с указанной целью в работе поставлены следующие задачи:

1) разработать метод идентификации объектов по выборке фиксированного объема, который позволяет оценивать и учитывать "истинные" значения вектора признаков объектов и точечные и интервальные оценки параметров заданных функций плотностей вероятности признаков, участвующих в процессе идентификации объектов, то есть необходимо разработать общий подход и единый алгоритм для задачи обучения системы распознавания и задачи идентификации объектов при априори известном виде функций условных плотностей вероятности признаков;

2) разработать метод распознавания образов, позволяющий учесть погрешности результатов наблюдений векторов признаков объектов при фиксированном объеме выборки при неизвестных условных плотностях вероятности признаков, но известных с точностью до значений параметров разделяющих функций;

3) разработать программное обеспечение, реализующее алгоритмы распознавания объектов по измеренным с погрешностью значениям признаков объектов как в случае заданных с точностью до параметров функций условных плотностей вероятности признаков, так и без учета условных плотностей вероятности признаков при известном виде разделяющих функций;

4) исследовать аналитически и в модельных экспериментах особенности предлагаемых алгоритмов и влияние погрешности в наблюдаемых значениях признаков на сходимость алгоритмов, на статистические свойства решений;

5) исследовать предлагаемые методы на реальных примерах распознавания химических взрывов и слабых землетрясений по сейсмическим наблюдениям, дать сравнительную оценку результатов, полученных традиционным и предлагаемым методами.

Объект исследования. В качестве объекта исследования методами, опирающимися на параметрические модели в виде обобщенных условных плотностей вероятности признаков или (и) разделяющих функций, выступают классы распознаваемых образов, в которых отсутствуют детерминированные исходные данные. В качестве конкретных примеров рассмотрены задачи распознавания химических взрывов и слабых землетрясений по признакам, измеренным с погрешностью.

Методы исследования. В основу исследования положен разработанный в диссертации метод распознавания образов, позволяющий учесть погрешность результатов наблюдений и точечные и интервальные оценки всех величин и функций, участвующих в процедуре распознавания. Метод базируется на конъюнктном анализе.

Разработанный метод основан на работах Р.Дуды, П.Харта, К.Фу, А.Л.Горелика, В.Г.Небабина и других – по теории распознавания образов, работах А.А.Грешилова, В.В.Федорова, W.Fuller, L.Gleser, D.Shafer, M.Kendall и других – по теории вероятностей и математической статистике, работах российских и зарубежных авторов по вычислительной математике и программированию на ЭВМ. При сравнительном анализе использовались результаты других исследователей.

Научная новизна работы состоит в следующем:

1) показано, что в присутствии помехи в результатах наблюдений вектора признаков объектов при фиксированном объеме выборки задача распознавания образов является одновременно и неразделимо задачей обучения системы распознавания и задачей идентификации образов; предложен унифицированный функционал и алгоритм идентификации образов, включающей в себя задачу обучения системы распознавания;

2) разработан алгоритм распознавания образов при фиксированном объеме выборки в присутствии помехи в результатах наблюдений векторов признаков через получение точечных и интервальных оценок "истинных" значений векторов признаков и функций условных плотностей вероятности признаков;

3) разработан алгоритм распознавания образов в присутствии помехи в результатах наблюдений векторов признаков при фиксированном объеме выборки, неизвестных условных плотностях вероятности признаков, но при известном виде разделяющих функций;

4) разработаны программы для ПК, реализующие алгоритмы получения точечных и интервальных оценок "истинных" значений векторов признаков,

условных плотностей вероятности признаков, решающих функций и процедур принятия решений при наличии помехи в результатах наблюдений векторов признаков и фиксированном объеме выборки;

5) в модельных и реальных экспериментах показано отличие результатов полученных традиционными методами и методами, разработанными в диссертации, которые позволяют учесть влияние помехи в наблюдаемых значениях векторов признаков.

Практическая значимость работы состоит в том, что ее результаты позволяют:

1) повысить достоверность принятия решения при распознавании образов за счет учета погрешности наблюдений векторов признаков, правильного расчета точечных и интервальных оценок всех исходных величин и функций;

2) принять адекватное решение при распознавании образов, когда различия в значениях векторов признаков образов соизмеримы с погрешностями наблюдаемых значений признаков;

3) принять адекватное решение по результатам наблюдений признаков с помехами, когда не известны условные плотности вероятности признаков.

Результаты проведенных в диссертации исследований использовались в хоздоговорных темах "Образ-8" (1995–96 г.г.), в госбюджетной тематике, проводимой НИИПИМ МГТУ им. Н.Э.Баумана, в учебных курсах "Принятие решений и распознавание образов" и "Математические методы прогнозирования", читаемых студентам МГТУ им. Н.Э.Баумана и слушателям МИПК при МГТУ им. Н.Э.Баумана по экономическим и управленческим специальностям. Теоретическая часть работы поддерживалась SAIC Global Technology Corporation (SAIC GT) USA, субконтракт №29-990026-142.

На защиту выносятся следующие основные результаты:

1) метод идентификации объектов по выборке фиксированного объема, учитывающий погрешности измерений векторов признаков объектов, включающий в себя процедуры определения точечных и интервальных оценок функций условных плотностей вероятности признаков, решающих функций и оценок "истинных" значений векторов признаков;

2) метод распознавания образов при фиксированном объеме выборки по измерениям признаков объектов, наблюдаемым с погрешностью, при неизвестных функциях условных плотностей вероятности признаков, но известном виде разделяющих функций;

3) теоремы и утверждения, определяющие свойства предлагаемого алгоритма и влияние погрешностей измерений вектора признаков на результаты идентификации объектов;

4) комплекс ПК-программ реализации разработанных методов распознавания образов по измерениям признаков в присутствии помехи при фиксированном объеме выборки.

5) результаты идентификации химических взрывов и слабых землетрясений по сейсмическим измерениям векторов признаков в присутствии помехи.

Апробация работы. Материалы диссертации докладывались на кафедре вычислительной математики и математической физики МГТУ им. Н. Э. Баумана и ученом совете МИПК при МГТУ им. Н. Э. Баумана, научно-методической конференции, посвященной 35-летию образования факультета "Фундаментальные науки" МГТУ им. Н. Э. Баумана в 1999 году.

Публикации. Основные результаты диссертации опубликованы в печати: книг – 2, тезисов докладов – 3.

Объем работы. Диссертационная работа состоит из введения, трех глав, заключения и списка литературы. Основная часть работы изложена на 145 страницах машинописного текста, содержит 37 рисунков, 2 таблицы, 78 наименований цитируемой литературы.

СОДЕРЖАНИЕ РАБОТЫ

Во введении формулируются проблемы, которые возникают при решении статистических задач распознавания образов при фиксированном объеме выборки: неучет погрешностей наблюдений признаков ведет к смещенным точечным оценкам функций условных плотностей вероятности признаков, а так же разделяющих (решающих) функций и к неверным интервальным оценкам этих функций. На этапе идентификации образов по наблюдаемому вектору признаков присутствие помехи в наблюдениях может привести к ложным выводам. Показывается, что неучет в традиционных методах интервальных оценок исключает из рассмотрения нулевую зону – зону неопределенности в принятии решения. Делается вывод, что для адекватного принятия решений алгоритмы обучения системы распознавания и идентификации образов должны основываться на конфлюентном анализе – совокупности методов, позволяющих проводить статистические исследования с учетом неопределенности всех исходных факторов.

В первой главе дана математическая постановка задачи, показано, что при учете погрешностей наблюдений признаков процедуры обучения системы распознавания и идентификации образов по измеренному вектору признаков сливаются в единую процедуру (рис. 1). Здесь же дан анализ традиционных методов, применяемых для идентификации объектов.

В традиционном подходе задача распознавания образов содержит два практически независимых этапа. На первом этапе проводится обучение системы распознавания по результатам наблюдений над объектами, которые в будущем подлежат распознаванию, и вырабатывается решающее правило.

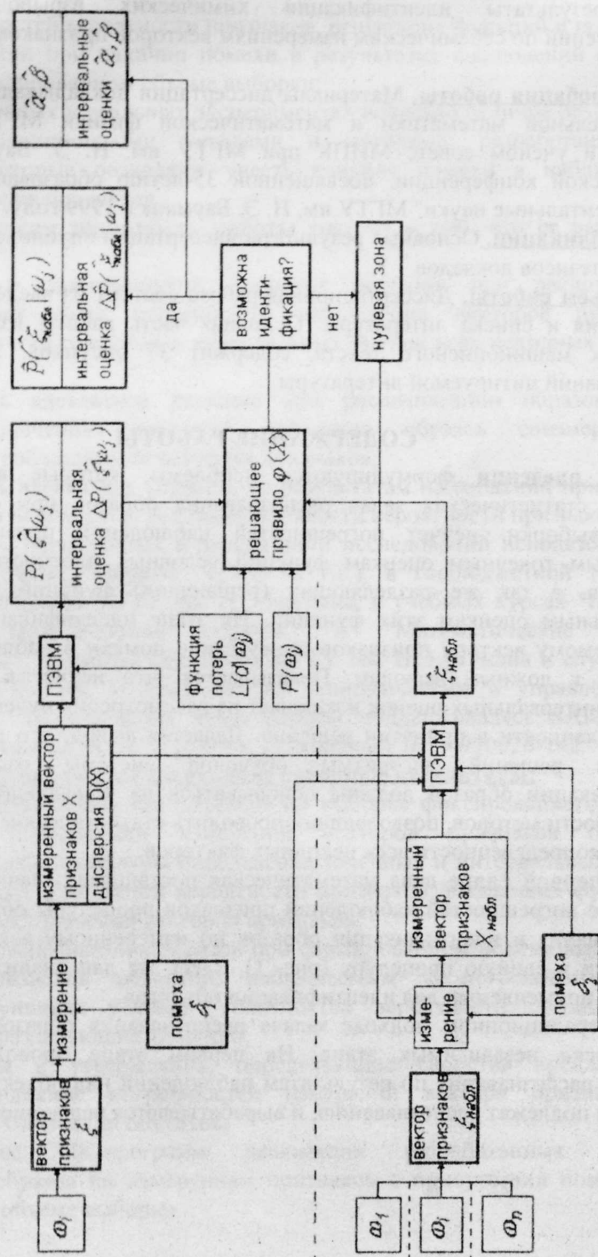


Рис. 1. Блок-схема единой процедуры обучения системы и распознавания образов

На втором этапе по результатам наблюдений вектора признаков неизвестного объекта, с помощью выработанного на первом этапе решающего правила, наблюдаемый объект идентифицируют – относят к одному из элементов допустимого множества объектов.

В действительности в процессе наблюдений измеренные значения вектора признаков имеют две компоненты: истинное значение ξ и помеху δ т. е. $X = \xi + \delta$. Неучет помехи δ приводит к смещенным оценкам всех характеристик, в расчете которых участвует значение $X_{\text{набл}}$. Кроме этого, необходимо учесть и интервальные оценки всех последующих рассчитываемых функций и величин: $p(X|\omega_j)$, ошибок I и II рода. Это не делается в традиционных методах, что на практике приводит к появлению зон неопределенности в принятии решений.

Идентификацию объектов следует проводить по оценке истинного значения $\xi_{\text{набл}}$, а не по наблюдаемому значению $X_{\text{набл}}$. В диссертации показано, что получить оценку $\xi_{\text{набл}}$ можно, если использовать всю информацию, полученную на первом этапе. В итоге, задача идентификации практически не разделяется на независимые этапы, а имеет общую схему (рис.1).

На рис. 1 обозначено $L(d|\omega_j)$ – ущерб от принятия решения d , когда выбран объект ω_j ; $P(\omega_j)$ – вероятность появления ω_j -го объекта. Прямоугольником с надписью ПЭВМ, обозначена процедура обработки результатов наблюдений и ее алгоритм; жирно обведенными прямоугольниками обозначены блоки, которых в традиционных методах либо вообще нет, либо эти блоки претерпевают существенные изменения.

В первой главе формулируется математическая постановка задачи.

Пусть до принятия решения d , принадлежащего множеству решений D ($d \in D$), задано параметрическое пространство образов Ω с параметром $\omega \in \Omega$, а для всех $\omega \in \Omega$ задана обобщенная плотность вероятности $P(\omega)$. Задана на произведении $\Omega \times D$ вещественная функция полезности $U(\omega, d)$ или функция потерь $L(\omega, d) = -U(\omega, d)$. Пусть S выборочное пространство возможных значений наблюдений признака ξ образа ω , когда значение признака не наблюдается, а наблюдается случайная величина X , связанная со значением признака ξ соотношением $X = \xi + \delta$, где δ – аддитивная помеха, закон распределения которой $f(X)$ и его числовые характеристики считаются известными, в частности, это может быть гауссовский закон распределения с математическим ожиданием $M(\delta) = 0$, и дисперсионной матрицей $D(\delta) = \sigma^2 * I$ (I – единичная матрица).

Пусть задано с точностью до значений параметров параметрическое семейство условных обобщенных вероятностных плотностей (о.в.п.) $p(\xi, \theta|\omega)$, $\omega \in \Omega$; при неизвестных о.в.п. $p(\xi, \theta|\omega)$ задано параметрическое семейство разделяющих функций $h(\xi, \theta)$.

Из класса решающих функций Δ выбрана решающая функция $\varphi(\xi) \in \Delta$, для любого возможного значения $\xi \in S$ определяющая решение $d(\xi) \in D$, минимизирующее средний риск:

$$p[\omega, d(\xi)] = M\{L[\omega, d(\xi)]\} = \int_{\Omega} \int_S L[\omega, d(\xi)] p(\xi, \theta | \omega) P(\omega) d\mu(\xi) d\nu(\omega).$$

Для всех $\omega \in \Omega$ функция $L[\omega, d(\xi)]$ измерима и интегрируема на множестве S .

Требуется по фиксированной выборке результатов наблюдений X :

1) найти точечные и интервальные оценки параметров θ условных о.в.п. $p(\xi, \theta | \omega)$, точечные и интервальные оценки условных о.в.п. $p(\xi, \theta | \omega)$; при неизвестных о.в.п. $p(\xi, \theta | \omega)$ – точечные и интервальные оценки параметров θ и самих разделяющих функций $h(\xi, \theta)$;

2) найти оценки "истинных" значений $\hat{\xi}$ вектора признаков ξ , по которым впоследствии принимается решение $d \in D$;

3) произвести идентификацию наблюдаемого объекта, т. е. поставить в соответствие наблюдаемому значению вектора признаков X некоторый образ ω , и определить при этом точечные и интервальные оценки суммарной ошибки и ошибок I и II рода, возникающих при принятии решения $d \in D$.

В традиционной постановке задачи распознавания помехи δ не учитываются, а величины ξ и X отождествляются; не определяются интервальные оценки функций $p(\xi, \theta | \omega)$, $h(\xi, \theta)$ и $\varphi(\xi)$, не определяются оценки $\hat{\xi}$ и интервальные оценки ошибок I и II рода (α и β).

Неучет помех δ и оценок $\hat{\xi}$ приводит к смещенным точечным оценкам $p(\xi, \theta | \omega)$, $h(\xi, \theta)$ и $\varphi(\xi)$, к неверным интервальным оценкам и ложным практическим выводам.

Во второй главе рассмотрены способы учета погрешностей наблюдений признаков в задачах распознавания образов при фиксированном объеме выборки, когда вид функций условных плотностей вероятности признаков $p(\xi, \theta | \omega_j)$ известен априори, но неизвестны значения их параметров θ .

Предполагается, что l -мерный вектор признаков ξ связан с результатами n наблюдений вектора X условием $X = \xi + \delta$, а вектор относительной частоты наблюдений n_0 связан с условной плотностью $p(\xi, \theta | \omega_j)$ выражением:

$$n_0 = p(\xi, \theta | \omega_j) + \varepsilon,$$

где δ , ε – аддитивные помехи, подчиняющиеся нормальному закону распределения с нулевыми математическими ожиданиями и с известными ковариационными матрицами.

Тогда оценки параметров θ по методу максимума правдоподобия находятся из условия минимума функционала:

$$F(X, n_0 | \theta) = [n_0 - p(\xi, \theta | \omega_j)]^T K^{-1} [n_0 - p(\xi, \theta | \omega_j)] + \sum_{i=1}^n \sum_{k=1}^l \frac{(x_{ik} - \xi_{ik})^2}{\sigma^2(x_{ik})}, \quad (1)$$

где K – дисперсионная матрица наблюдений относительных частот n_0 , $\sigma^2(x_{ik})$ – дисперсия x_{ik} . Далее показывается, что этот же функционал может быть получен как функционал ортогональной регрессии, а ограничения, связанные с нормальным законом распределения, могут быть сняты.

Точечные оценки N параметров θ находятся из условия

$$\left. \frac{\partial F}{\partial \theta_\nu} \right|_{\theta=\hat{\theta}} = 0, \quad \nu = 1, \dots, N;$$

при дополнительном ограничении

$$\left. \frac{\partial F}{\partial \xi_{ik}} \right|_{\xi_{ik}=\hat{\xi}_{ik}} = 0, \quad i = 1, \dots, n, \quad k = 1, \dots, l. \quad (2)$$

Запишем (2) в более явном виде

$$\frac{n_{0i} - p(\hat{\xi}_{ik}, \theta | \omega_j)}{\sigma^2(n_{0i})} p'(\hat{\xi}_{ik}, \theta | \omega_j) + \frac{x_{ik} - \xi_{ik}}{\sigma^2(x_{ik})} = 0, \quad i = 1, \dots, n, \quad k = 1, \dots, l. \quad (3)$$

Выражение (3) определяет оценки $\hat{\xi}$ "истинных" значений наблюдаемых признаков ξ , которые необходимы процедуре идентификации образов по наблюдаемому вектору признаков и выработанному решающему правилу.

Дисперсионная матрица оценок параметров $\hat{\theta}$ имеет вид:

$$D(\hat{\theta}) = M^{-1}; \quad M = \left\| \frac{\partial^2 F}{\partial \theta_{\nu_1} \partial \theta_{\nu_2}} \right\|_{\theta=\hat{\theta}}; \quad \nu_1, \nu_2 = 1, \dots, N.$$

Получив значения оценок $\hat{\theta}$ и $D(\hat{\theta})$, можно в интересующей области значений ξ определить интервальную оценку функции условной плотности вероятности признаков:

$$p \left\{ p(\hat{\xi}, \hat{\theta} | \omega_j) - t_\gamma \sqrt{D[p(\hat{\xi}, \hat{\theta} | \omega_j)]} \leq p(\xi, \theta | \omega_j) \leq p(\hat{\xi}, \hat{\theta} | \omega_j) + t_\gamma \sqrt{D[p(\hat{\xi}, \hat{\theta} | \omega_j)]} \right\} = \gamma,$$

где γ - доверительная вероятность; t_γ - квантиль распределения Стьюдента; а

$$D[p(\hat{\xi}, \hat{\theta} | \omega_j)] = \sum_{\nu=1}^N \left[\frac{\partial p(\xi, \theta | \omega_j)}{\partial \theta_\nu} \right]_{\theta=\hat{\theta}}^2 D(\hat{\theta}_\nu) + 2 \sum_{\nu_1 > \nu_2}^N \frac{\partial p(\xi, \theta | \omega_j)}{\partial \theta_{\nu_1}} \cdot \frac{\partial p(\xi, \theta | \omega_j)}{\partial \theta_{\nu_2}} \cdot D(\hat{\theta}_{\nu_1}, \hat{\theta}_{\nu_2}) \quad (4)$$

Здесь $D(\hat{\theta}_{\nu_1}, \hat{\theta}_{\nu_2})$ - корреляционный момент. Под интервальной оценкой

$p(\xi, \theta | \omega_j)$ понимается огибающая семейства всевозможных значений $p(\hat{\xi}, \hat{\theta} | \omega_j)$, полученных при различных значениях $\hat{\theta}$ в интервале $\hat{\theta} \pm t_\gamma \sqrt{D(\hat{\theta})}$.

Зная точечные и интервальные оценки $p(\xi, \theta | \omega_j)$, $j=1, \dots, m$, можно получить решающее правило (решающую функцию) для различных объектов ω_j ($j=1, \dots, m$) (например, воспользовавшись критерием отношения правдоподобия), интервальную оценку найденной решающей функции (т.е. возможную нулевую зону), точечные и интервальные оценки ошибок I и II рода (и суммарной ошибки). На этом процедура обучения заканчивается.

В процедуре идентификации объекта наблюдаемый вектор признаков $X_{набл}$ надо соотнести с ω_j ($j=1, \dots, m$). Решение принимается не по $X_{набл}$, а по $\xi_{набл}$. Поэтому для различных гипотез H_j ($j=1, \dots, m$) по формуле (3), в случае достаточного числа точек, при нахождении оценки $p(\hat{\xi}, \hat{\theta}|\omega_j)$ находятся m

оценок $\hat{\xi}_{набл}$, по которым и принимается решение. При недостаточном числе точек, оценки $\hat{\xi}_{набл}$ и $\hat{\theta}$ надо получить снова, присовокупив для каждой гипотезы значение $X_{набл}$ к набору значений векторов признаков, полученных в процессе обучения системы. Когда получена решающая функция $\varphi(\xi)$ и ее интервальные оценки $\varphi_*(\xi)$ и $\varphi_*(\xi)$, то правило распознавания примет вид:

$$\begin{cases} \text{если } \varphi(\xi) > \varphi_*(\xi), \text{ то } \xi \approx \omega_1; \\ \text{если } \varphi(\xi) < \varphi_*(\xi), \text{ то } \xi \approx \omega_2; \\ \text{если } \varphi_*(\xi) < \varphi(\xi) < \varphi_*(\xi), \text{ то продолжить измерения.} \end{cases}$$

Далее во второй главе исследуются статистические свойства оценок $\hat{\theta}$, влияние погрешности признаков на изменение погрешности оценок $\hat{\theta}$; показывается, что результаты модельных экспериментов подтверждают аналитические выкладки.

Один из вариантов распознавания образов при наблюдении двух нормально распределенных признаков X_1 и X_2 показан на рис. 2, где принято: — без учета погрешностей признаков; — — с учетом погрешностей признаков; - - - - интервальная оценка решающей функции. Традиционный подход в этом случае приводит к неверному решению.

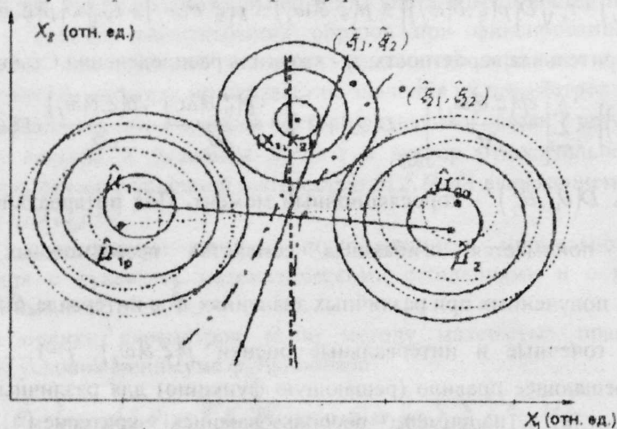


Рис. 2. Распознавание образов с учетом погрешности наблюдений признаков при известном виде функций условных плотностей вероятности признаков

На рис. 3 представлена блок-схема программной реализации предложенного метода.

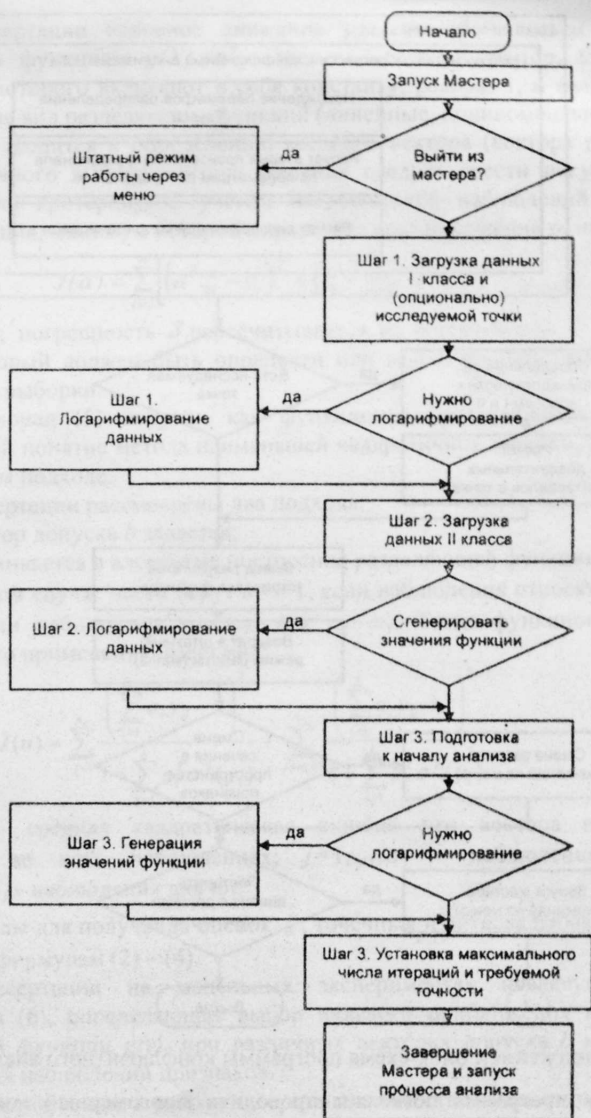


Рис. 3. Блок-схема программы конфлюэнтного анализа данных

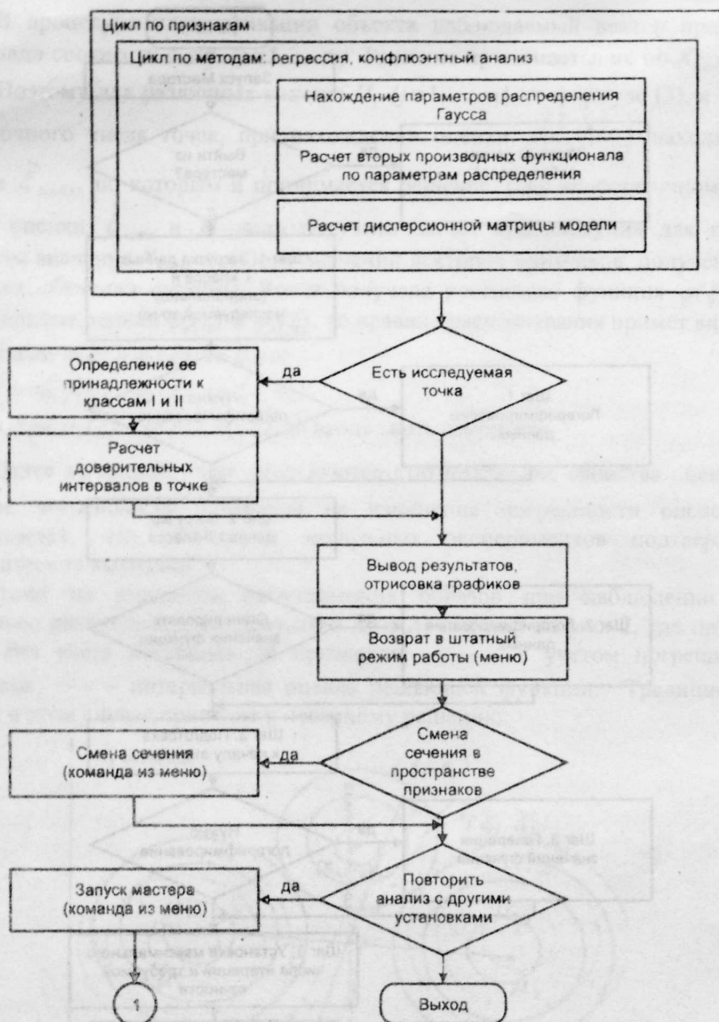


Рис. 3 (продолжение). Блок-схема программы конфлюэнтного анализа данных

Данная программа позволяет проводить многомерный конфлюэнтный анализ данных двух классов объектов, распределенных по нормальному закону.

В третьей главе рассмотрена задача распознавания образов, когда известен вид разделяющих функций. По результатам наблюдений признаков получают точечные и интервальные оценки решающих функций и проводят идентификацию объектов по наблюдаемому вектору признаков.

В диссертации основное внимание уделено обобщенным линейным разделяющим функциям $g(X) = a^T y$, где a – весовой вектор, y – вектор, координаты которого включают в себя константу, равную 1, и функции от X , определяющие вид разделяющих функций (линейные, полиномиальные и т.д.).

Задача сводится к определению весового вектора (вектора решения) a . Для однозначного выбора вектора решения следует ввести некую целевую функцию (или критерий). С учетом погрешностей наблюдений признаков целевая функция, минимум которой определяет вектор решений a , имеет вид:

$$J(a) = \sum_{i=1}^n \left[(a^T \xi_i - b_i)^2 + (\xi_i - y_i)^2 \right], \quad (5)$$

где $y_i = \xi_i + \delta_i$; погрешность δ пересчитывается из погрешности X ; b – вектор допуска, который должен быть определен или задан. Функция критерия $J(a)$ включает все выборки.

Функционал (5) получен как функционал ортогональной регрессии, расширяющий понятие метода наименьшей квадратичной ошибки, принятого в традиционном подходе.

В диссертации рассмотрены два подхода:

- 1) вектор допуска b задается;
- 2) оценивается в алгоритме построения разделяющей функции.

В первом случае часто берут $b = +1$, если наблюдения относятся к $y_1 \sim \omega_1$, и $b = -1$, если наблюдения относятся к $y_2 \sim \omega_2$. Тогда функционал (5) для практического применения имеет вид:

$$J(a) = \sum_{j=1}^{n_1} \frac{\left(\sum_{i=1}^k a_i y_{ij} - 1 \right)^2}{1 + \sum_{i=1}^k a_i^2 \sigma^2(y_i)} + \sum_{j=n_1+1}^{n_2} \frac{\left(\sum_{i=1}^k a_i y_{ij} + 1 \right)^2}{1 + \sum_{i=1}^k a_i^2 \sigma^2(y_i)}, \quad (6)$$

где $\sigma(y_i)$ – средняя квадратическая ошибка i -го вектора наблюдений, одинаковая во всех наблюдениях; $j = 1, \dots, n_1$ – наблюдения для ω_1 , $j = n_1 + 1, \dots, n_2$ – наблюдения для ω_2 .

Формулы для получения оценок $\hat{\xi}$, точечных и интервальных оценок $g(y)$ аналогичны формулам (2) – (4).

В диссертации на модельных экспериментах исследуется форма функционала (6), определяющая выбор нулевого приближения и поведения разделяющей функции $g(y)$, при различных векторах допуска b и различных погрешностях наблюдений признаков.

Все предлагаемые в работе алгоритмы в качестве нулевых приближений используют традиционные подходы, т.е. они улучшают традиционные методы. Таким образом, всегда имеется возможность сопоставить результаты по предлагаемому и традиционному методу.

В третьей главе диссертации описаны возможности и процедура работы с разработанным программным обеспечением, функционирующем в среде

Windows 95/98. Ниже представлена общая блок-схема компьютерной программы, реализующей данный алгоритм (рис.4).

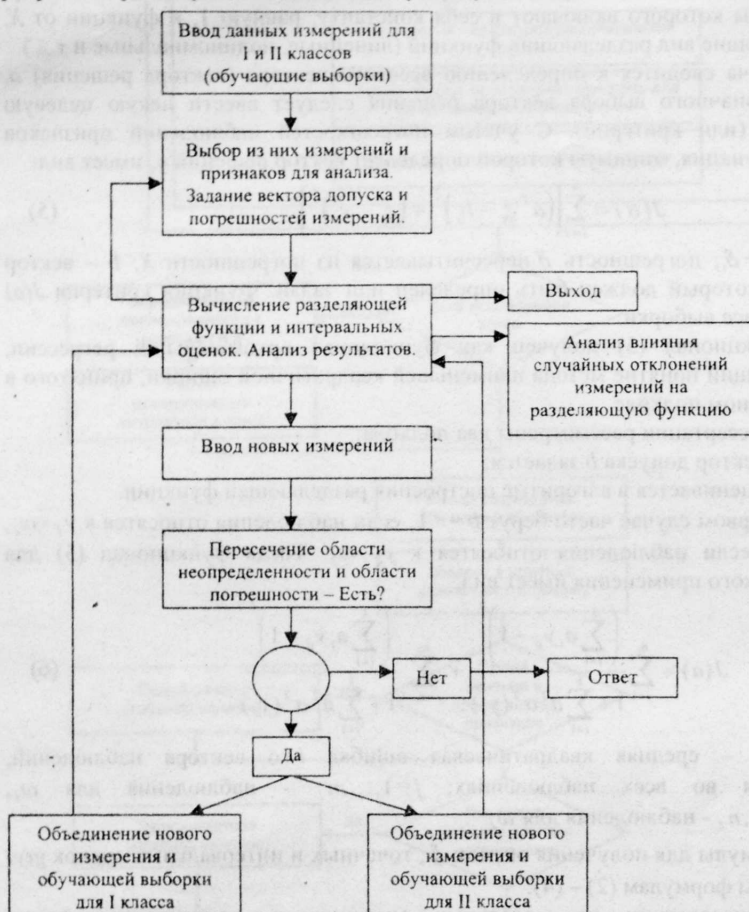


Рис. 4. Общая блок-схема программы

Разработанный метод был применен для идентификации слабых локальных землетрясений и искусственных взрывов по сейсмограммам. Обучение системы проводилось по пяти наиболее информативным признакам из 18 исходных: отношению средних энергий S- и P-фаз в частотном диапазоне 10-15 Гц; доле средней энергии S-фазы, принадлежащей частотному диапазону 6-10 Гц; отношению пиковых энергий S- и P-фазы в частотном диапазоне 1-3 Гц; отношению средних энергий S- и P-фазы в частотном диапазоне 6-10 Гц; доле средней энергии S – фазы, принадлежащей частотному диапазону 1-3 Гц.

Один из вариантов расчетов представлен на рис. 5. По оси абсцисс отложены номера точек наблюдения. По оси ординат – расстояние от точек до разделяющей поверхности. Светлые и серые эллипсы соответствуют точкам первого и второго классов. Их вертикальный размер определяется ошибками измерений, которые задаются. Прямоугольники, примыкающие одной стороной к оси абсцисс, определяют интервальные оценки для соответствующих точек. Как видно из рис. 5, неправильно идентифицируются точки первого класса с номерами 9 и 23, а также точка второго класса с номером 24.

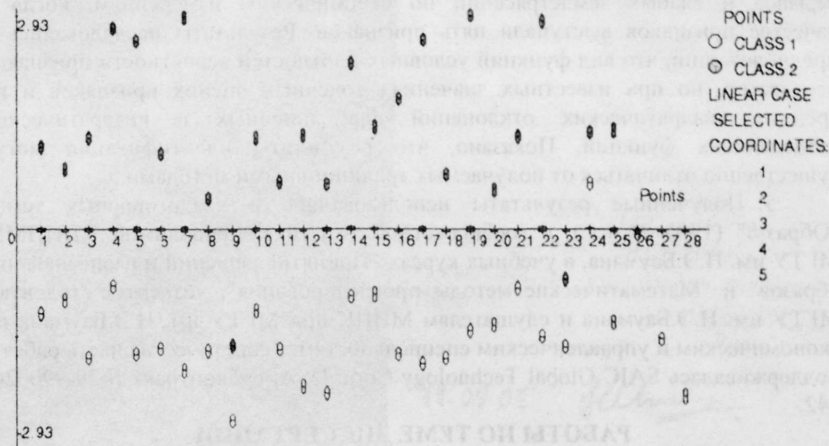


Рис. 5. Пример идентификации по сейсмическим признакам

ОСНОВНЫЕ РЕЗУЛЬТАТЫ ДИССЕРТАЦИИ

1. Разработан метод распознавания образов по фиксированной выборке при заданных параметрических моделях условных плотностей вероятности признаков или разделяющих функций, позволяющий учитывать погрешности наблюдений значений признаков, получать точечные и интервальные оценки функций условных плотностей вероятности признаков или разделяющих функций, точечные и интервальные оценки наблюдаемого вектора признаков, по которому ведется идентификация объектов. Метод объединяет процедуру обучения системы распознавания и процедуру идентификации образов. Разработана процедура принятия решений (идентификации объекта) при пересечении нулевой зоны и области неопределенности наблюдаемых значений признаков. Показано, что метод работоспособен как при известном виде функций условных плотностей вероятности признаков, так и в случае, когда вид функций условных плотностей вероятности признаков не задан, а заданы только средние квадратические погрешности наблюдений и общий вид разделяющих функций.

2. Исследованы основные особенности предложенного алгоритма распознавания образов и форма получаемых функционалов, показана

несмещенность и единственность получаемых оценок, определен выбор нулевого приближения в процедурах оценивания параметров моделей и исследованы статистические свойства полученных оценок, показана согласованность результатов моделирования и аналитических расчетов.

3. Разработано программное обеспечение, реализующее на ПК предложенный метод распознавания образов и обеспечившее проведение статистических испытаний получаемых результатов.

4. Предложенный метод применен для распознавания химических взрывов и слабых землетрясений по сейсмическим измерениям, когда в качестве признаков выступали пять признаков. Результаты исследовались в предположении, что вид функций условных плотностей вероятности признаков неизвестен, но при известных значениях точечных оценок признаков и их средних квадратических отклонений для линейных и квадратических разделяющих функций. Показано, что результаты идентификации могут существенно отличаться от получаемых традиционными методами.

5. Полученные результаты использовались в хоздоговорных темах "Образ-8" (1995–96 г.г.), в госбюджетной тематике, проводимой НИИПММ МГТУ им. Н.Э.Баумана, в учебных курсах "Принятие решений и распознавание образов" и "Математические методы прогнозирования", читаемых студентам МГТУ им. Н.Э.Баумана и слушателям МИПК при МГТУ им. Н.Э.Баумана по экономическим и управленческим специальностям. Теоретическая часть работы поддерживалась SAIC Global Technology Corp. USA, субконтракт №29-990026-142.

РАБОТЫ ПО ТЕМЕ ДИССЕРТАЦИИ

1. Мальцев А.В. Параметрическое распознавание образов по выборке фиксированного объема с погрешностями в признаках / Под ред. А.А.Грешилова. –М.: Радио и связь, 1999. – 73 с.

2. Анализ стохастических систем при учете погрешностей исходных данных / А.А.Грешилов, М.П.Гордин, А.В.Мальцев и др. // Научно-методическая конференция, посвященная 35-летию образования факультета "Фундаментальные науки" МГТУ им. Н.Э.Баумана: Тезисы докладов. –М., 1999. – С. 49.

3. Грешилов А.А., Мальцев А.В. Математические проблемы в задачах распознавания образов // Научно-методическая конференция, посвященная 35-летию образования факультета "Фундаментальные науки" МГТУ им. Н.Э.Баумана: Тезисы докладов. –М., 1999. – С. 50.

4. Грешилов А.А., Мальцев А.В., Пархоменко В.П. Принятие решений с помощью обобщенных линейных разделяющих функций. –М.: Радио и связь, 2000. – 49 с.

5. Грешилов А.А., Мальцев А.В., Пархоменко В.П. Математические проблемы учета погрешностей и программная реализация алгоритмов распознавания образов // Системы компьютерной математики и лингвистики: Материалы междунар. конф. – Смоленск, 2000. – С 17-19.