

Макет аппаратно-программного комплекса обработки данных гиперспектрального аэрокосмического зондирования

Т.В. Кондранин¹, В.В. Козодеров², Е.В. Дмитриев³, В.Д. Егоров³, В.В. Борзяк²

¹ *Московский физико-технический институт (государственный университет)
141700, Московская область, г. Долгопрудный, Институтский переулок, 9.
E-mail: kondr@kondr.rector.mipt.ru*

² *Московский государственный университет им. М.В. Ломоносова
119991, Москва, ГСП-1, Воробьевы горы, МГУ, д. 1.
E-mail: vkozod@mes.msu.ru*

³ *Институт вычислительной математики РАН
119333, г. Москва, ул. Губкина, д. 8.
E-mail: yegor@inm.ras.ru*

Основу создаваемого макета аппаратно-программного комплекса составляют вычислительные процедуры распознавания природно-техногенных объектов с обучением по тестовой выборке. Показаны характерные особенности одного из наиболее перспективных подходов к решению задачи распознавания таких объектов по гиперспектральным данным при использовании статистических методов дискриминантного анализа с обеспечением минимальной среднеквадратической ошибки классификации (байесовский подход). Итогом реализации данного подхода является применение пошагового метода «последовательного дополнения», который позволяет проводить отбор наиболее информативных признаков с оптимизацией каналов гиперспектрального зондирования.

Ключевые слова: гиперспектральное аэрокосмическое зондирование, обработка данных, аппаратно-программный комплекс, распознавание природно-техногенных объектов.

Введение

В существующих зарубежных публикациях по использованию данных гиперспектрального аэрокосмического зондирования рассмотрены следующие основные направления: обоснование необходимости использования гиперспектральных данных для повышения точности распознавания образов и оценки состояния почвенно-растительного покрова; возможности улучшения точности решаемых прикладных задач по данным гиперспектрального зондирования; построение модельных описаний связи биофизических процессов вегетирующей растительности и взаимодействия с ней солнечного излучения при разных уровнях представления рассматриваемой модельной среды. Основные направления соответствующих приложений данных дистанционного гиперспектрального зондирования: использование индексов NDVI для общего понимания разделяемых классов состояния объектов, индексов LAI и их модификаций для общего представления о пологе лесной и другой растительности, других индексов.

Значительно более развиты проблемы построения классификаторов с учетом соотношений между объемом обучающих выборок и размерности признакового пространства. Создаваемый макет аппаратно-программного комплекса обработки данных гиперспектрального аэрозондирования (Кондранин и др., 2011; Козодеров и др., 2011) основан на ре-

лизации наиболее перспективного подхода построения оптимальных классификаторов для распознавания образов разных типов природных объектов.

При решении задачи распознавания нами разделяется два принципиально разных подхода: распознавание на основе спектральной информации каждого отдельного пикселя и распознавание на основе анализа пространственных градаций яркости изображений в фиксированном канале радиометра. На данном этапе классификация по спектральным признакам представляется более перспективной. Действительно, при обеспечении стабильных условий съемки, можно осуществить попиксельную координатную привязку изображения, и на основе имеющихся наземных данных «обучить» алгоритмы классификации и провести подробную независимую валидацию.

Один из наиболее перспективных подходов к решению задачи распознавания природных и антропогенных объектов по гиперспектральным данным основан на статистических методах дискриминантного анализа. Смысл данного подхода заключается в том, чтобы построить функцию в N -мерном пространстве признаков, в данном случае нормированных спектральных яркостей, такую, чтобы на конечном обучающем множестве обеспечить минимальную среднеквадратическую ошибку классификации, либо минимальную полную вероятность ошибки (байесовский подход). Данную функцию называют дискриминантной.

Рассмотрим более подробно основы решения рассматриваемых прикладных задач распознавания природно-техногенных объектов как составной части макета аппаратно-программного комплекса обработки гиперспектральных изображений.

Задача классификации гиперспектральных изображений по спектральному признаку

Формальная постановка задачи построения алгоритма классификации с обучением состоит в следующем. Пусть X – множество объектов, а Y – множество ответов (имен или меток классов). Пусть также имеется априорная информация в виде конечного множества пар элементов множеств X и Y : $X^N = \{x(i), y(i)\}_{i=1}^N$, где N – число пар. Очевидно $X^N \subset X \times Y$. Необходимо построить отображение $a: X \rightarrow Y$ (или $s = a(x)$, где $x \in X, s \in Y$) для произвольного элемента $x \in X$, таким образом, чтобы оно приводило к минимальным ошибкам в заданном смысле на множестве X^N . Таким образом, множество X^N называют обучающим множеством, а процесс построения оптимального алгоритма (как правило, оптимизация параметров) называют обучением.

При решении задачи распознавания природно-техногенных объектов по данным гиперспектрального зондирования, классифицируемыми объектами $x \in X$ являются пиксели гиперспектрального изображения. Множество Y содержит названия распознаваемых объектов (например, растительность, почва, водная поверхность, асфальт, бетон и т.п.). В практических задачах вместо непосредственно классифицируемых объектов $x \in X$ рассматривают некоторые их характеристики (признаки), при этом множество X называют признаковым пространством.

В качестве признаков (характеристик пикселей) будем рассматривать значения спектральных плотностей энергетической яркости отраженного излучения видимого и ближнего инфракрасного диапазона, измеряемые с помощью гиперспектрометра. По данным измерений, для каждого пикселя изображения можно построить соответствующий спектр. На рис. 1 (слева) представлен пример спектра пикселя с координатами [629,399], полученный с помощью гиперспектрометра НПО «Лептон», который имеет 290 спектральных каналов.

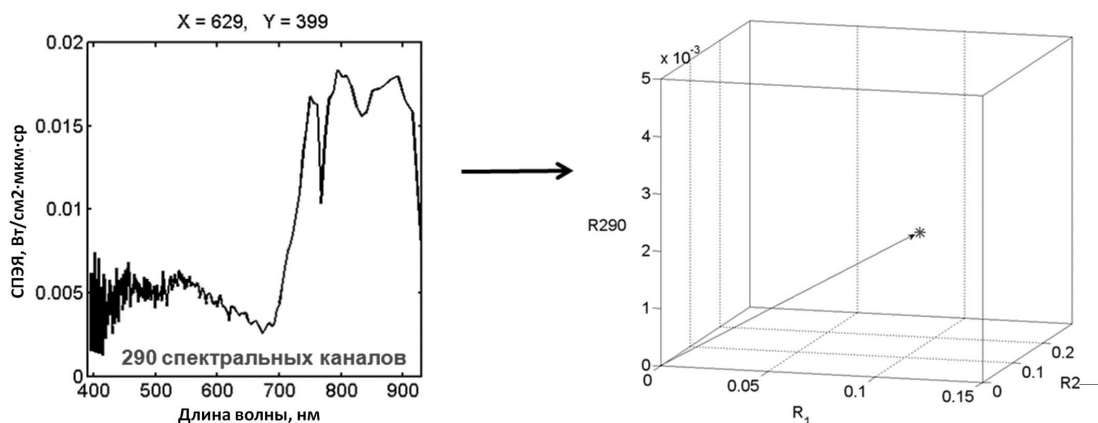


Рис. 1. Пример формирования признакового пространства с размерностью $N=290$ (R_1, R_2, \dots, R_{290} – каналы гиперспектрометра)

Один из способов формирования признакового пространства заключается в том, что каждый спектральный канал рассматривается в качестве отдельного измерения пространства признаков. Таким образом, каждый спектр пикселя гиперспектрального изображения представляется точкой в N -мерном пространстве, где N – число спектральных каналов.

Также следует отметить, что при решении данной задачи не обязательно использовать непосредственно яркости отраженного излучения. Часто более удобно осуществить переход в подпространство, элементы которого получены из исходного пространства с помощью заданного нетривиального и ненулевого преобразования (линейного или нелинейного). Типичным примером таких преобразований является отбор наиболее информативных спектральных каналов, переход в пространство спектральных плотностей яркости нормированных на интегральную яркость, проецирование в пространство коэффициентов разложения по эмпирическим ортогональным функциям, использование ядерных преобразований скалярного произведения.

Вероятностная постановка задачи

Формирование спектров излучения отраженного от распознаваемых объектов определяется множеством факторов – оптическими свойствами элементарных однородных компонент объектов. Так, например, если элементу разрешения гиперспектрального прибора соответствует участок лесного полога, то даже при однородном породном и возрастном составе измеряемый спектр будет значимо зависеть от объемного распределения листвы и ветвей, вариаций отражающей способности листьев, ветвей и подстилающей поверхности, а также от соотношения указанных элементов. Поскольку на практике нет точной информации о процентном соотношении элементарных компонент и их взаимном пространственном расположении, то формирование спектра пикселя гиперспектрального прибора имеет случайный характер. Таким образом, для построения алгоритмов классификации вполне естественно использовать результаты теории вероятностей.

Общая вероятностная постановка задачи может быть сделана следующим образом. В качестве вероятностного пространства рассматривается множество $X \times Y$, которое содержит все возможные пары элементов множеств X и Y . Элементы множества $X \times Y$ будем также называть прецедентами. Учитывая конечность множества Y , плотность распределения прецедентов можно представить в виде $p(x, y) = P(y)p(x | y)$, где $P(y) \equiv P_y$ – вероятность класса

у рассматриваемой сцене (априорная вероятность класса), $p(x | y) \equiv p_y(x)$ – плотность вероятности распределения признаков внутри класса (функция правдоподобия класса).

В рамках рассматриваемой физической задачи, включение в вероятностное пространство всех возможных пар обосновано тем, что при определенных условиях, касающихся распределения элементарных компонент распознаваемых объектов, спектры различных классов могут совпадать. Т.е. одной и той же точке признакового пространства может с разной ненулевой вероятностью соответствовать несколько ответов. В рамках вероятностного подхода требуется построить алгоритм, который приводит к минимальной вероятности ошибки классификации (байесовский алгоритм).

Построение алгоритма «четкой» классификации сводится к разбиению признакового пространства на непересекающиеся области, каждая из которых ставится в соответствие одному из ответов. Разбиение проводится оптимальным в заданном смысле образом с помощью поверхности (кривой) называемой дискриминантной поверхностью (рис. 2). Функция, определяющая данную поверхность, называется дискриминантной функцией, а алгоритм классификации иногда называют дискриминантом.

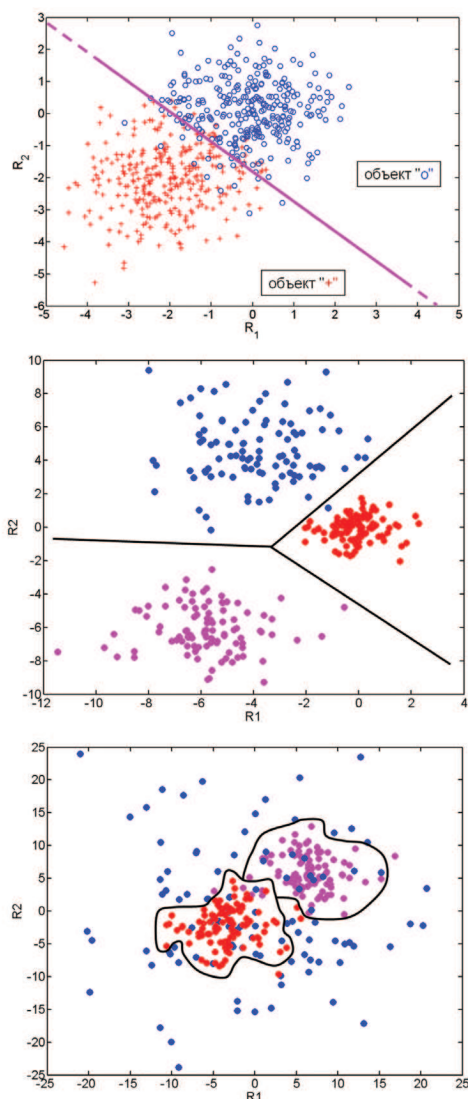


Рис. 2. Построение дискриминантной поверхности (кривой): верхний – линейный дискриминант для бинарной (двуклассовой) классификации, средний – линейный дискриминант для случая многих классов, нижний – нелинейный дискриминант для случая многих классов

Если бы были известны точные P_y и $P_y(x)$, то оптимальный в вероятностном смысле алгоритм строился бы однозначно на основе принципа наибольшего правдоподобия $s = a(x) = \arg \max_{y \in Y} P(y)p(x|y)$. Попадание точно на дискриминантную поверхность является событием с нулевой вероятностью. На практике ввиду дискретности ряда машинных чисел, вероятность попадания на дискриминантную поверхность отличается от нуля. В этом случае классификация производится случайным образом между классами, для которых данный участок дискриминантной поверхности является граничным.

В случае, когда априорные вероятности и функции правдоподобия классов неизвестны, но имеется обучающее множество $X^N = \{x(i), y(i)\}_{i=1}^N$ – случайная выборка объема N из $X \times Y$, выполненная независимым образом, можно вместо истинной функции $p(x,y)$ использовать ее статистическую оценку. В этом случае вероятностная оптимизация ведется в конечном, выборочном пространстве. Оптимизация во всем вероятностном пространстве является некорректной задачей, поскольку многим $p(x,y)$ может соответствовать одна и та же выборка X^N .

Формализуем определение ошибок классификации. Рассмотрим множество всех признаков, которые будут отнесены алгоритмом к классу $s \in Y$, как событие: $A_s = \{x \in X | a(x) = s\}$. Вероятность появления объекта «истинного» класса y , который будет отнесен алгоритмом $a(x)$ к классу s , определяется как $P_y P(A_s | y) = P_y \int p(x|y) dx$. Если $y=s$, то классификация считается правильной, если $y \neq s$ – ошибочной. Для каждой пары (y,s) вводится характеристика λ_{ys} , определяющая величину потери при отнесении объекта класса y к классу s . Способ определения λ_{ys} зависит от конкретной задачи, однако при правильной классификации ($y=s$) обычно полагают $\lambda_{ys} = \lambda_{yy} = 0$, а при ошибочной ($y \neq s$) – $\lambda_{ys} > 0$. Совокупность элементов λ_{ys} может быть записана в виде матрицы потерь

$$\Lambda = \begin{pmatrix} 0 & \lambda_{12} & \cdots & \lambda_{1M} \\ \lambda_{21} & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \lambda_{(M-1)M} \\ \lambda_{M1} & \cdots & \lambda_{M(M-1)} & 0 \end{pmatrix},$$

M – число всех классов, имеющей на диагонали нулевые элементы и, вообще говоря, не симметричной, поскольку потери от ошибок в различных классах могут существенно различаться.

В качестве глобальной характеристики ошибки алгоритма классификации $a(x)$ вводится «функционал среднего риска»

$$R(a) = \sum_{y \in Y} \sum_{s \in Y} \lambda_{ys} P_y P(A_s | y), \quad (1)$$

который является математическим ожиданием потери λ_{ys} – случайной величины с дискретным распределением (поскольку распределения s и y дискретные). Заметим, что если $\lambda_{ys} = 1, \forall (y, s : y \neq s)$, то $R(a)$ представляет собой полную вероятность ошибки алгоритма $a(x)$.

Поиск оптимального в байесовском смысле алгоритма классификации $a(x)$ сводится к минимизации функционала среднего риска $R(a)$.

Байесовские классификаторы: от общего к частным алгоритмам

Приведем общий алгоритм байесовской классификации в наиболее компактной форме:

$$a(x) = \arg \min_{s \in Y} \sum_{y \in Y} \lambda_{ys} P_y p_y(x). \quad (2)$$

Таким образом, для решения задачи распознавания необходимо для каждого класса задать соответствующие функции правдоподобия, априорные вероятности и определить потери от ошибок классификации. Победителем считается тот класс, который имеет минимальное математическое ожидание потери.

Покажем, что данный алгоритм доставляет минимум функционалу среднего риска. Для этого зафиксируем произвольный класс t из множества Y и перепишем этот функционал в виде

$$R(a) = \sum_{y \in Y} \lambda_{yt} P_y P(A_t | y) + \sum_{s \in \{Y \setminus \{t\}\}} \sum_{y \in Y} \lambda_{ys} P_y P(A_s | y).$$

Поскольку A_s – непересекающиеся события, то по формуле полной вероятности имеем

$$P(A_t | y) = 1 - \sum_{s \in \{Y \setminus \{t\}\}} P(A_s | y).$$

Таким образом, исходный функционал может быть представлен в виде:

$$\begin{aligned} R(a) &= \sum_{y \in Y} \lambda_{yt} P_y + \sum_{s \in \{Y \setminus \{t\}\}} \sum_{y \in Y} (\lambda_{ys} - \lambda_{yt}) P_y P(A_s | y) = \text{const}(a) + \sum_{s \in \{Y \setminus \{t\}\}} \int_{A_s} \sum_{y \in Y} (\lambda_{ys} - \lambda_{yt}) P_y p_y(x) dx = \\ &= \text{const}(a) + \sum_{s \in \{Y \setminus \{t\}\}} \int_{A_s} (g_s(x) - g_t(x)) dx, \end{aligned}$$

$$\text{где } g_s(x) = \sum_{y \in Y} \lambda_{ys} P_y p_y(x).$$

В итоге алгоритм оптимальный байесовский укажет на класс s (т.е. выполнится $a(x)=s$) тогда и только тогда, когда класс имеет минимальную ожидаемую потерю

$$s = \arg \min_{t \in Y} g_t(x) = \arg \min_{t \in Y} \sum_{y \in Y} \lambda_{yt} P_y p_y(x),$$

что и требовалось доказать. Если таких классов несколько, то можно взять любой из них, поскольку в этом случае подынтегральное выражение равно нулю.

Рассмотрим частный случай, который соответствует *классической формулировке байесовского классификатора*. Предположим, что функция потери для данного истинного класса y не зависит от класса s , который выбирается алгоритмом распознавания, т.е. мы можем представить функцию потерь в виде

$$\lambda_{ys} = \begin{cases} 0, & s = y, \\ \lambda_y, & s \neq y. \end{cases}$$

Тогда при заданных априорных вероятностях и функциях правдоподобия минимум функционала среднего риска достигается на алгоритме

$$a(x) = \arg \max_{y \in Y} \lambda_y P_y p_y(x). \quad (3)$$

Вероятность того, что измеренному набору признаков x соответствует класс y , является апостериорной вероятностью и может быть вычислена по формуле Байеса

$$P(y | x) = \frac{p(x, y)}{p(x)} = \text{const} \cdot P_y p_y(x).$$

Тогда алгоритм (3) может быть представлен в виде

$$a(x) = \arg \max_{y \in Y} \lambda_y P(y | x),$$

что является классической формулировкой байесовского решающего правила. Если положить потери $\lambda_y \equiv 1$, то имеем алгоритм, известный как *принцип максимума апостериорной вероятности*

$$a(x) = \arg \max_{y \in Y} P(y | x). \quad (4)$$

Если при этом все классы равновероятны, т.е. $P_y = 1/M$, $\forall y \in Y$, то имеем классическую формулировку алгоритма, известного как *принцип максимального правдоподобия*

$$a(x) = \arg \max_{y \in Y} p_y(x).$$

Оценивание плотностей вероятностей распределения признаков внутри классов требует рассмотрения двух случаев.

1. Семейство распределений не определено – непараметрическая оценка.

Хорошо известным приближением функции плотности вероятности является нормализованная гистограмма. Будем сразу рассматривать многомерный случай. Пусть m – число элементов обучающего множества, соответствующих заданному классу, n – размерность признакового пространства. Тогда задав некоторый вектор, состоящий из положительных чисел $h = [h_1, \dots, h_p, \dots, h_n]T$, $h_j > 0$ (например, это могут быть ширины $n+1$ -мерного столбика гистограммы по каждой размерности признакового пространства), имеем оценку

$$\hat{p}_h(x) = \frac{1}{m} \sum_{i=1}^m \prod_{j=1}^n \frac{1}{h_j} \delta \left(|x^j - x_i^j| < \frac{h_j}{2} \right),$$

где x^j – компонента вектора признаков x , δ – индикаторная функция, действующая на пространстве событий и определяемая выражением

$$\delta(A) = \begin{cases} 1, & \text{событие } A \text{ произошло;} \\ 0, & \text{иначе.} \end{cases}$$

Часто используют более общую (настраиваемую) оценку Парзена-Розенблатта

$$\hat{p}_h(x) = \frac{1}{m} \sum_{i=1}^m \prod_{j=1}^n \frac{1}{h_j} K \left(\frac{x^j - x_i^j}{h_j} \right),$$

где K – произвольная непрерывная, ограниченная, четная функция, называемая ядром.

2. Параметрический подход предполагает, что функция плотности вероятности известна с точностью до параметров – $p(x) = \varphi(x, \theta)$, т.е. задано параметрическое семейство распределений Φ^θ . В этом случае, если выполнены условия регулярности, оптимальное значение вектора параметров θ можно оценить исходя из принципа максимального правдоподобия.

Если $p(x) \in N(\mu, \Sigma)$, т.е. принадлежит семейству нормальных распределений, то совокупность получаемых алгоритмов составляет нормальный дискриминантный анализ. Напомним, что плотность многомерного нормального распределения определяется выражением

$$p_{N(\mu, \Sigma)}(x) = \frac{1}{\sqrt{(2\pi)^n \det(\Sigma)}} \exp\left(-\frac{(x - \mu)^T \Sigma^{-1} (x - \mu)}{2}\right),$$

где $\mu = [Ex^1, Ex^2, \dots, Ex^n]^T$ – вектор математических ожиданий, $\Sigma = E(x - \mu)(x - \mu)^T$ – ковариационная матрица.

Если имеется m выборочных реализаций случайного вектора $x \sim N(\mu, \sigma)$ размерности n , то несмещенные оценки неизвестных статистических моментов имеют вид

$$\hat{\mu} = \frac{1}{m} \sum_{i=1}^m x_i, \quad \hat{\Sigma} = \frac{1}{m-1} \sum_{i=1}^m (x_i - \hat{\mu})(x_i - \hat{\mu})^T.$$

Получение оценок статистических моментов собственно и есть процесс обучения.

При поиске максимума апостериорной вероятности можно использовать монотонные преобразования. В рассматриваемом случае удобно использовать логарифмирование. Алгоритм (4) преобразуется к виду $a(x) = \arg \max_{y \in Y} (\ln P_y + \ln p_y(x))$. Подставляя функцию плотности многомерного нормального распределения, имеем алгоритм

$$a(x) = \arg \max_{y \in Y} \left(\ln(P_y) - \frac{1}{2} (x - \hat{\mu}_y)^T \hat{\Sigma}_y^{-1} (x - \hat{\mu}_y) - \frac{1}{2} \ln(\det(\hat{\Sigma}_y)) \right), \quad (5)$$

известный как *квадратичный дискриминантный анализ*. Название связано с тем, что дискриминантные поверхности, разделяющие каждую пару классов, задаются полиномами 2 степени (рис. 3)

$$F = x^T Q x + L x + K,$$

где $Q = \hat{\Sigma}_2^{-1} - \hat{\Sigma}_1^{-1}$, $L = 2(\hat{\mu}_1^T \hat{\Sigma}_1^{-1} - \hat{\mu}_2^T \hat{\Sigma}_2^{-1})$ и $K = 2 \ln \left(\frac{P_1}{P_2} \right) + \ln \left(\frac{\det \hat{\Sigma}_2}{\det \hat{\Sigma}_1} \right) + \hat{\mu}_2^T \hat{\Sigma}_2^{-1} \hat{\mu}_2 - \hat{\mu}_1^T \hat{\Sigma}_1^{-1} \hat{\mu}_1$.

Выбор наиболее информативных признаков

Пошаговый метод «последовательного дополнения» (в англоязычной литературе – «step up method») позволяет проводить отбор наиболее информативных признаков. Суть метода заключается в следующем (рис. 4). Выбирается метод распознавания и априорная вероятность для распознаваемых классов. Пространство всех признаков делится на две группы – принятых в модель и остальных признаков (для которых производится оценка возможности принятия в модель). Для каждого признака из множества «остальных признаков» делается оценка ошибки распознавания при условии его добавления в модель. Из полученного набора ошибок выбирается минимальная и производится ее сравнение с ошибкой предыдущей модели. Если произошло значимое уменьшение ошибки, то соответствующий признак принимается в модель, если нет, то процесс останавливается.

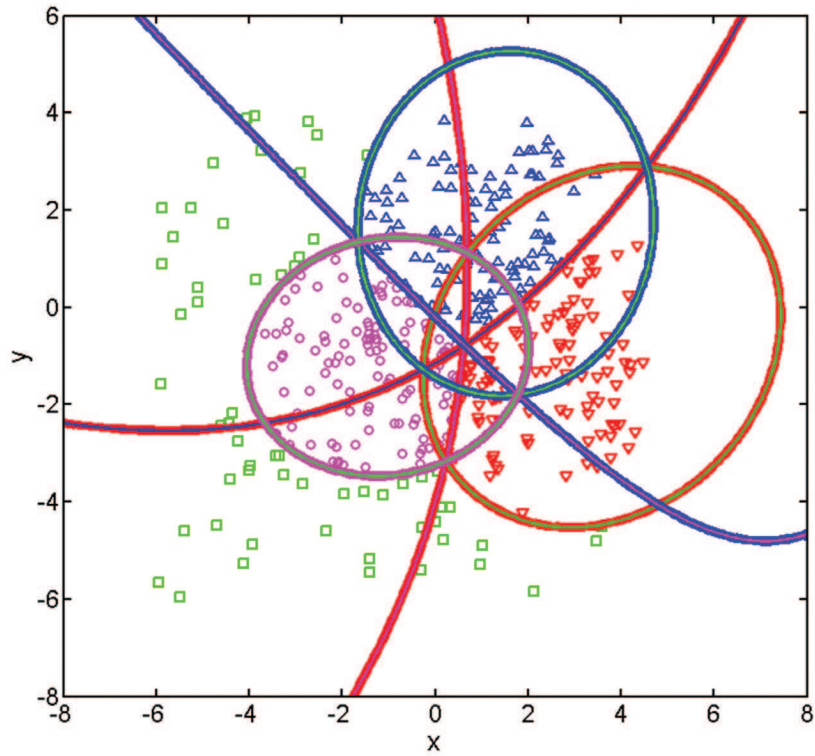


Рис. 3. Квадратичный дискриминантный анализ для случая 4 классов (2-мерное признаковое пространство). Дискриминантные кривые обозначены двумя цветами в соответствии с цветами разделяемых классов

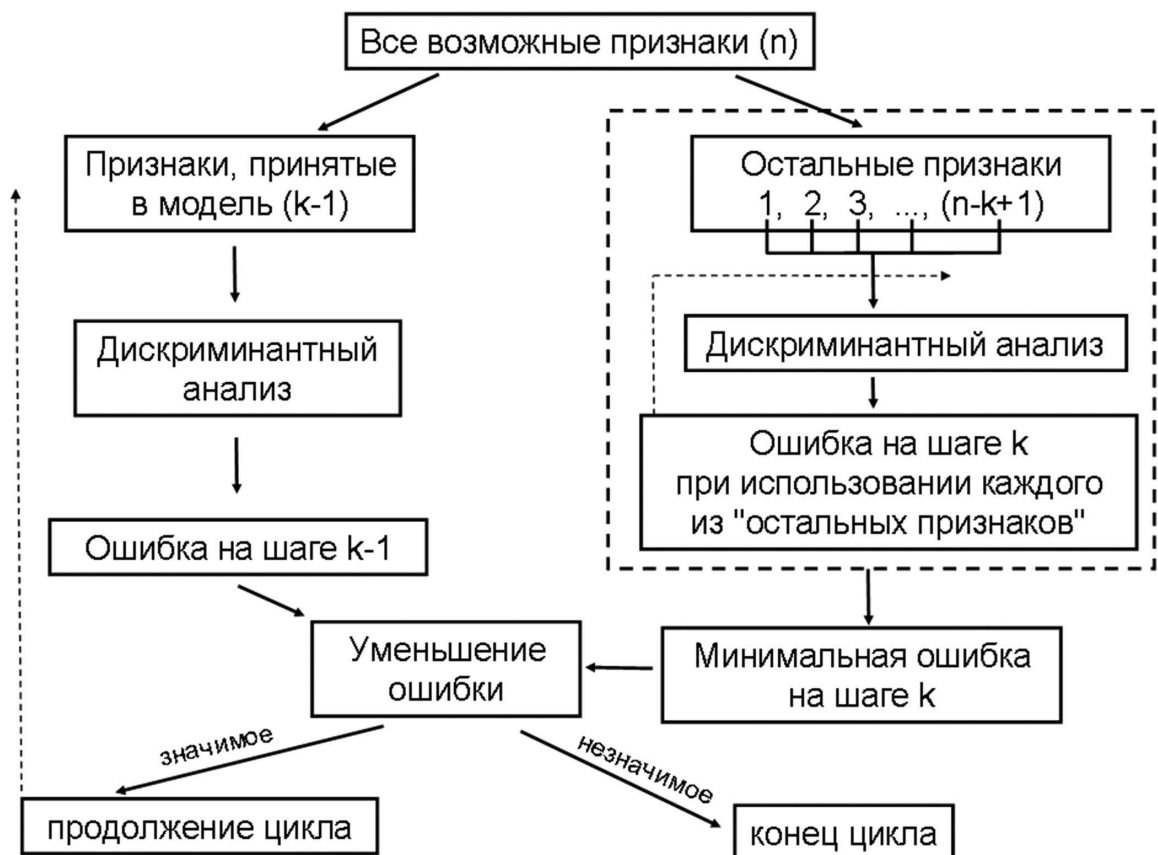


Рис. 4. Схема работы метода «последовательного дополнения», используемого для выбора наиболее информативных признаков

В нашей задаче контроль остановки цикла может осуществляться двумя способами.

1. Проверка значимости различия ошибок на текущем и предшествующем шагах на основе параметрического метода проверки гипотезы о равенстве эмпирических частот. Уровень значимости 0.1.

2. Оценка максимально возможного числа используемых признаков – цикл продолжается до тех пор, пока оценка полной вероятности ошибки распознавания уменьшается. Метод основан на том, что при наличии достаточного количества незначимых признаков неизбежно возникнут биения оценки ошибки распознавания. Этот метод позволяет оценить верхнюю границу числа используемых признаков, потенциально полезных для распознавания.

Некоторые примеры реализации этого подхода для оптимизации спектральных каналов гиперспектрального аэрозондирования в заданной предметной области показаны в работе (Козодеров, Дмитриев, 2012). Здесь отметим новизну предлагаемого подхода, который позволяет уменьшить избыточность конкретных наборов гиперспектрального зондирования с обоснованием точности решения задачи распознавания объектов.

Заключение

Макет аппаратно-программного комплекса обработки данных гиперспектрального зондирования строится в соответствии с требованиями использования оптимальных классификаторов для распознавания объектов с заданной точностью. Изложены основы вероятностной трактовки разрабатываемых вычислительных процедур. Показаны особенности непараметрических и параметрических методов оценки распределения вероятности при минимизации функции потерь неправильной классификации.

Исследования выполняются при финансовой поддержке ФЦП «Научные и научно-педагогические кадры инновационной России» на 2009-2013 годы, г/к № ПЗ49 и № 14.740.11.1091, проектов РФФИ № 11-07-00382, 11-07-12006-офи_м.

Литература

1. *Кондранин Т.В., Козодеров В.В., Казанцев О.Ю., Бобылев В.И., Дмитриев Е.В., Егоров В.Д., Каменцев В.П., Борзяк В.В.* Проблемы классификации гиперспектральных авиакосмических изображений // Современные проблемы дистанционного зондирования Земли из космоса, 2011. Вып.8. Т.1. С.90-98.
2. *Козодеров В.В., Дмитриев Е.В., Егоров В.Д., Борзяк В.В.* Вычислительные аспекты построения классификаторов разной сложности при обработке гиперспектральных аэрокосмических изображений // Современные проблемы дистанционного зондирования Земли из космоса, 2011. Вып.8. Т.3. С.55-63.
3. *Козодеров В.В., Дмитриев Е.В.* Дистанционное зондирование лесного покрова: инновационный подход // Вестник Московского государственного университета леса, 2012. №1 (в печати).

A lay-out of the apparatus and programmatic complex of hyperspectral airborne remote sensing data processing

T.V. Kondranin¹, V.V. Kozoderov², E.V. Dmitriev³, V.D. Egorov³, V.V. Borzyak²

¹ *Moscow Institute for Physics and Technology (State University)*

E-mail: kondr@kondr.rector.mipt.ru

² *M.V. Lomonosov Moscow State University*

³ *Institute of Numerical Mathematics of Russian Academy of Sciences*

Computational procedures of natural and anthropogenic objects recognition with supervising using test samples give a basis of the created lay-out of the apparatus and programmatic complex. Characteristic features of one of the most perspective approaches are outlined to solve the recognition problem for such objects while applying hyperspectral data with the utilization of statistical techniques of discriminant analysis that ensures the minimal mean square error of classification (the Bayesian approach). The usage of a step up method that enables to conduct the selection of the most informative features with the optimization of the applied remote sensing channels is realized as a result of the proposed approach.

Keywords: hyperspectral airborne remote sensing, data processing, apparatus and programmatic complex, recognition of natural and anthropogenic objects.