

# Технология построения информативного признака области натурального гиперспектрального изображения для задачи классификации

М.И. Хотилин<sup>1</sup>

<sup>1</sup>Самарский национальный исследовательский университет им. академика С.П. Королева, Московское шоссе 34а, Самара, Россия, 443086

## Аннотация

В данной статье описывает процесс определения признаков области гиперспектрального изображения, характеризующий данную область. Описаны методы и алгоритмы поиска признаков принадлежности к определенной области. Указаны дальнейшие перспективы развития алгоритма на больших данных.

## Ключевые слова

Гиперспектральные изображения, MaZda, дискриминантный анализ, классификация, отбор признаков

## 1. Введение

Гиперспектральные изображения – это трёхмерный массив данных, включающий в себя пространственную информацию об объекте, дополненную спектральной информацией по каждой пространственной координате [1]. Обработка и анализ гиперспектральных изображений является популярной тематикой исследований в области обработки изображений и компьютерного зрения. В рамках данной работы рассматривается технология построения признаков гиперспектрального изображения для задачи классификации, на примере отдельно взятой его области.

## 2. Постановка задачи

Весь ход работы можно условно разделить на последовательно выполняемые этапы. На первом этапе, посредством использования модуля предобработки выделяется небольшая область, используемая для исследования и обработки (6х6 пикселей, количество слоев 242).

Далее, из гиперкуба исследуемой области выделяется набор всех двумерных сечений плоскостями. Этот этап необходим, поскольку заранее не представляется возможным сказать, какие именно слои являются значимыми, а также потому, что применение существующих методов и средств вычисления признаков, является затруднительным к рассматриваемым гиперспектральным изображениям.

Для исследования и обработки полученных выше сечений было принято решение использовать ПО «MaZda» [4], позволяющее произвести расчет различных групп признаков.

Дискриминантный анализ позволяет определить какие признаки, слои, совокупности слоев и совокупности пикселей (гиперпиксели) являются значимыми для каждого изображения.

## 3. Практическое исследование и программная реализация

В качестве исходного изображения рассматривалось изображение, размером 1311x3411 пикселей с количеством спектральных каналов равным 242. Далее из исходного изображения, посредством написанного модуля предобработки, вырезалась небольшая область, после чего полученный массив яркостей разделялся на двумерные по каждой паре координат.

Следующим этапом являлось исследование полученных сечений и извлечение характеризующих их признаков с помощью «MaZda». Входными данными являлся массив

изображений-сечений, а в качестве выходных представлялся отчет, содержащий различные характеристики и признаки. Объединяя все данные отчетов, был получен массив признаков, характеризующих указанную область изображения.

Далее применялся дискриминантный анализ, позволяющий на основании множества признаков, сформировать показатель для прогнозирования принадлежности признаков к группам. Также линейный дискриминантный анализ можно интерпретировать как построение разделяющей прямой (плоскости), позволяющей разделить два или более классов.

В ходе работы рассматривались области двух классов, количество рассмотренных областей было 10 по каждому классу, количество измерений по каждому экземпляру области 260, количество признаков - 260. Размерность итогового массива входных данных – 5200x260. Данный массив был разбит на два – обучающую и тестовую выборки.

Для ускорения вычислений было сокращено количество используемых признаков, путем удаления линейно-зависимых столбцов и столбцов с корреляцией выше 0.999 (при меньшем значении корреляции возрастала ошибка классификации). В результате произошло сокращение количества столбцов-характеристик с 260 до 58.

Был произведен расчет необходимых параметров, в том числе средних признаков и коэффициентов линейных дискриминантов. В результате расчетов было выявлено, что наибольший вклад (наибольшее значение коэффициента) вносят функции углового второго момента (ASM), которые являются мерой однородности изображения. Они позволяют отличать изображения с разной текстурой.

Далее происходила классификация тестовой выборки на основе результатов дискриминантного анализа. В результате классификации, получилось, что общее число ошибок классификации было 130, что, в общем, составило 5%.

Рассмотренные области являлись небольшими, и все вычисления производились на персональном компьютере. Однако при возрастании размеров исследуемых областей и их числакратно возрастает сложность вычислений и время их выполнения.

## 4. Заключение

Определение признаков, определяющих однозначно принадлежности объектов (областей изображения) к определенному классу является одной из значимых задач классификации и обработки изображений. В данной работе описан алгоритм, выделяющий из изображения информацию о его признаках, благодаря которому можно проводить распознавание образов

В настоящее время ведется работа над реализацией алгоритма, повышающего эффективность выполнения задачи поиска и извлечения признаков для различных классов областей гиперспектральных изображений с помощью методов работы с данными больших объемов.

## 5. Литература

- [1] Зимичев, Е.А. Пространственная классификация гиперспектральных изображений с использованием метода кластеризации K-MEANS++ / Е.А. Зимичев, Н.Л. Казанский, П.Г. Серафимович // Компьютерная оптика. – 2014. – Т. 38, № 2. – С. 281-286. DOI: 10.18287/0134-2452-2014-38-2-281-286.
- [2] Сергеев, В.В. Применение методологии распознавания образов в задачах цифровой обработки изображений / В.В. Сергеев // Автометрия. – 1998. – Т. 2. – С. 63-76.
- [3] Strzelecki, M. A software tool for automatic classification and segmentation of 2D/3D medical images / M. Strzelecki, P. Szczypinski, A. Materka, A. Klepaczko // Nuclear Instruments and Methods In Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment.– 2013. – Vol. 702. – P. 137-140.
- [4] Хотилин, М.И. Классификация объектов натуральных гиперспектральных изображений / М.И. Хотилин, Н.С. Кравцова, И.А. Рыцарев, А.В. Куприянов // Информационные технологии и нанотехнологии (ИТНТ): сборник трудов. – Самара: Изд-во Самарского университета. – 2020. – С. 86-90.