

Применение классификаторов на основе самокорректирующихся кодов в задаче дистанционной инвентаризации лесных территорий

А.О. Дементьев¹, Е.В. Дмитриев^{1,2}, В.В. Козодеров В.В.^{3,1}

¹ *Институт вычислительной математики РАН*

² *Московский физико-технический институт (государственный университет)*

³ *Московский государственный университет имени М.В.Ломоносова*

Тезисы.

Создание и обновление карт лесотаксации и решение задач управления с использованием этих данных, особенно в регионах, где затруднено создание аналогичных карт наземным способом, является одной из важных задач современного управления лесным хозяйством. Основным и наиболее точным источником информации являются данные наземных обследований. Получение такого рода информации сопряжено с существенными временными затратами и финансовыми расходами, поэтому в последние годы активно развиваются дистанционные методы инвентаризации лесов.

Используемые на сегодняшний день обучаемые алгоритмы распознавания объектов по мультиспектральным и гиперспектральным изображениям обычно основываются на стандартных подходах различной сложности: метрические классификаторы, методы ближайших соседей, байесовские классификаторы, нелинейный метод опорных векторов. Указанные классификаторы отличаются точностью, устойчивостью к переобучению и вычислительной эффективностью. Следует отметить, что из указанных подходов нельзя выбрать оптимальный во всех трех смыслах. Выбор подхода определяется конкретной задачей классификации, таким образом, усовершенствование методики определения параметров древостоев по данным дистанционного зондирования лежит в комплексировании указанных выше методов.

В рамках данной работы рассматривается задача построения многоклассовых методов распознавания на основе комплексирования бинарных классификаторов. Для этого применяется методика построения алгоритмов обучаемой классификации, основанная на теории самокорректирующихся кодов (Error-Correcting Output Codes - ECOC) [1]. Проводится численный анализ указанных алгоритмов на основе тестов с модельными данными. Рассматриваются возможности использования бинарных классификаторов различного типа. Приводятся результаты расчетов связанных с решением задачи оптимизации кодовой матрицы.

Указанная методика была отработана на тестовом участке, расположенном на территории Савватеевского лесничества Тверского лесхоза, для которой имеются стандартные данные наземной лесотаксации. Породный состав леса на тестовой территории представлен древостоями сосны, березы, осины и ели. Более подробное описание указанной измерительной кампании приводится в работе [2]. Представлена методика эффективного использования участков со смешанным породным составом древостоев для построения обучающего набора данных. Приводятся примеры распознавания породного и возрастного состава древостоев по данным авиационных измерений с помощью отечественной гиперспектральной аппаратуры производства НПО Лептон (г. Зеленоград, г. Москва) с соответствующими численными оценками точности классификации.

Исследования проводятся в рамках проекта РНФ №16-11-00007 «Автоматизация распознавания природно-техногенных объектов на аэрокосмических изображениях высокого спектрального и пространственного разрешения», а также проектов РФФИ №14-07-00141 «Распознавание и оценка параметров состояния лесной растительности по данным гиперспектральной аэрокосмической съемки», №14-05-00598 «Дистанционные методы определения таксационных, биологических и продукционных параметров древостоев на основе гиперспектральных изображений», №16-01-00107 «Повышение информационного содержания вычислительных процедур распознавания природно-техногенных объектов по данным аэрокосмического гиперспектрального зондирования».

Литература.

1. Dietterich T.G., Bakiri G. Solving Multiclass Learning Problems via Error-Correcting Output Codes // Journal of Artificial Intelligence Research. 1995. V. 2. P. 263-286.

2. Дмитриев Е.В. Классификация лесного покрова тверской области на основе гиперспектральных аэроизображений // Исследование Земли из космоса. 2013. №3. С. 22-32.