

COMBINED CLASSIFICATION OF SIMILAR LOOKING VEGETATIVE COVER TYPES USING HYPERSPECTRAL IMAGERY

Sergey M. Borzov¹, Oleg I. Potaturkin^{1,2}, Mark A. Guryanov^{1,2}

¹Institute of Automation and Electrometry, Siberian Branch,
Russian Academy of Sciences, Novosibirsk, Russia

²Novosibirsk State University, Novosibirsk, Russia

Abstract

In this paper we present an efficiency analysis of a number of methods for spectral and spectral-spatial classification of similar looking vegetative cover types using hyperspectral images for different approaches to construction of the training set.

Keywords: Remote sensing, hyperspectral images, surface type classification, spectral and spatial features

КОМПЛЕКСНАЯ КЛАССИФИКАЦИЯ ТРУДНОРАЗЛИЧИМЫХ ТИПОВ РАСТИТЕЛЬНОСТИ ПО ГИПЕРСПЕКТРАЛЬНЫМ ИЗОБРАЖЕНИЯМ

Борзов С.М. ⁽¹⁾, Потатуркин О.И. ⁽¹⁾⁽²⁾, Гурьянов М.А. ⁽¹⁾⁽²⁾

¹ Институт автоматизации и электрометрии СО РАН, Новосибирск,

² Новосибирский государственный университет, Новосибирск

Работа посвящена исследованию эффективности методов контролируемой спектральной и спектрально-пространственной классификации трудноразличимых типов растительности по данным гиперспектральной съемки при различных способах формирования обучающих выборок (случайным образом из всего изображения и из его части). Рассмотрены зависимости точности классификации от количества спектральных признаков для случаев разделения основных классов и подклассов подстилающей поверхности.

Ключевые слова: дистанционное зондирование Земли, гиперспектральные изображения, классификация типов поверхностей, спектральные и пространственные признаки.

В работе на примере тестового изображения, полученного в рамках программы AVIRIS [1], выполнено сравнение ряда подходов к повышению точности классификации за счет учета окрестности пикселей на различных этапах обработки данных. Рассматриваются алгоритмы, использующие как предварительное пространственное усреднение исходных данных, так и постобработку картосхем попиксельной спектральной классификации. Для спектральной классификации используются методы опорных векторов (SVM) и максимального правдоподобия (ML) (как показано в [2], они обеспечивают лучшую точность классификации данного изображения).

Размер изображения – 614x2677 пикселей, разрешение – 20 м/пикс. число каналов – 220 в диапазоне 0.4-2.5 мкм, при этом 20 каналов, содержащих высокий уровень шума, при проведении исследований удалены из рассмотрения. По результатам наземных наблюдений выполнено разбиение изображения на 58 классов, в том числе 15 классов кукурузы и 18 сои, отличающиеся способами их производства. В частности, имеются участки с различными методами возделывания почвы: no-till – вспашка не производится, min-till – вспашка с применением мало-разрушающих технологий, clear – обычная вспашка, а отдельные из них отличаются лишь направлением посева (север-юг или запад-восток). Общее количество пикселей в размеченной части изображения 334245, из них 128719 (39%) различные классы кукурузы, 91770 (27%) – сои, 63706 (19%) – лес. 15% пикселей относятся к прочим 23 классам. Часть территории не размечена и в работе не рассматривается.

Исследование эффективности алгоритмов классификации выполняется при трех различных способах формирования обучающей выборки из крупноформатного ГС изображения с учетом информации о принадлежности пикселей к классам:

- случайным образом выбирается по 25% пикселей из всего изображения;
- случайным образом выбирается 50% пикселей из левой половины изображения;
- случайным образом выбираются 50% пикселей шахматных клеток одного цвета.

Оценка эффективности классификации выполняется путем сопоставления сформированных картосхем классов и подспутниковых данных. В качестве критерия используется так называемая результирующая точность классификации (overall accuracy), определяемая как выраженный в процентах отношение количества верно классифицированных пикселей к их общему количеству в тестовой выборке. При этом точность классификации рассчитывается:

- по всем 58 классам;
- по основным 27 классам (с объединением 15 подклассов кукурузы и 18 сои в класс кукуруза и класс соя в результирующей картосхеме).

Показано, что одновременный учет пространственных и спектральных признаков позволяет достичь качественной классификации трудноразличимых подклассов растительности

только при использовании обучающих выборок с максимальной степенью распределения пикселей по изображению. При этом, наибольшей эффективностью обладает метод, сочетающий пространственную обработку исходных изображений и постобработку формируемых картограмм.

Установлено, что результирующая точность при рассматриваемых способах обучения в значительной степени различается. Так при формировании ОВ из всего изображения точность классификации основных типов поверхности методом опорных векторов (SVM) составила 67,3%, а методом максимального правдоподобия (ML) 72,5%, пространственная обработка позволила ее повысить до 84,0% и 93,1%, соответственно. Точность классификации при обучении по части изображения оказалась ниже (64,0% и 56,8%), а применение предварительной пространственной обработки улучшило результаты приблизительно на 3%. В случае разделения подклассов растительности при формировании ОВ из всего изображения точность составила 52,8%(SVM) и 53,8%(ML), а пространственная обработка позволила ее повысить до 63,4% и 72,1%. При формировании ОВ из части изображения точность составила лишь 43,1% и 42,8%, и при применении предварительной пространственной обработки – 48,0% и 49,1%.

При использовании большего количества признаков расхождение в результатах еще более значительное. Так в случае 200 признаков для ОВ, сформированной из всего изображения, точность классификации методом максимального правдоподобия составила 59,0%, а после предварительной пространственной обработки достигла 95,2%, тогда как для ОВ близкого объема, сформированной из части изображения, точность не превысила 43,1% и практически не менялась при предварительной пространственной обработке,

Основными причинами невысокой точности классификации при обучении по части изображения (помимо схожести спектров отражения различных классов), по всей видимости, является неполное устранение влияния неоднородности параметров регистрирующей системы по полю зрения, а также определенная внутриклассовая изменчивость характеристик растительности, связанная с локальными условиями ее прорастания (состав почвы, влажность, высота, направление склона и т.п.). Отметим, что из двух используемых способов формирования ОВ из части изображения более эффективным оказывается способ, основанный на клетчатой структуре. Он, очевидно, более точно характеризует параметры каждого класса на всем изображении, поскольку в определенной степени исключает влияние неоднородности характеристик регистрирующей системы.

Из этого следует, что для разделения трудноразличимых классов растительности необходимо использовать обучающие выборки с максимальной степенью распределения пикселей по всему изображению. Однако такой подход на практике трудно реализуем, поскольку в большинстве реальных задач невозможно обеспечить подспутниковые данные по всей наблюдаемой территории. Как правило, достоверные наземные данные могут быть получены лишь на отдельных ее участках и, следовательно, только по ним формировать ОВ. Поэтому исследования, ориентированные на оценку эффективности традиционных и оригинальных алгоритмов классификации ГС изображений при обучении по всему полю наблюдения, следует признать недостаточно значимыми с точки зрения практических приложений.

Дальнейшие исследования в этой области целесообразно направить на разработку алгоритмов классификации, обеспечивающих повышение эффективности разделения трудноразличимых типов растительности, при обучении по отдельным фрагментам анализируемого ГС изображения. В этом плане могут быть рассмотрены алгоритмы с адаптивным формированием обучающей выборки из определенного количества пространственно наиболее близких пикселей каждого класса.

Работа выполнена при финансовой поддержке РФФИ (грант № 13-07-12202-офи_м).

ЛИТЕРАТУРА

- [1] Baumgardner M.F., Biehl L.L., Landgrebe D.A. 220 Band AVIRIS Hyperspectral Image Data Set: June 12, 1992 Indian Pine Test Site 3. Purdue University Research Repository. 2015. doi:10.4231/R7RX991C.
- [2] Борзов С.М., Потатуркин А.О., Потатуркин О.И., Федотов А.М. Исследование эффективности классификации гиперспектральных спутниковых изображений природных и антропогенных территорий // Автометрия. 2016. № 1. С. 3-14.