

**Л. Руис** (L. Ruiz),  
**Т. Эрмосилла** (T. Hermosilla),  
**Г. Сериса** (G. Serisa) и др.

Научно-исследовательская группа по геоэкологическому картографированию и дистанционному зондированию Земли. Политехнический университет, г. Валенсия (Испания).

## Многовариантная и объектно-ориентированная стратегия обновления базы данных земельных ресурсов и растительного покрова по снимкам WorldView-2\*

### ВВЕДЕНИЕ

Одна из стратегических задач, возникающих при разработке геоинформационных систем, заключается в разработке и применении новых методов, позволяющих учитывать создание и обновление баз географических данных по космическим снимкам и другим источникам информации. Создание и поддержка баз геоданных земельных ресурсов и растительного покрова в различных масштабах являются обязательными для эффективного управления на региональном и общегосударственном уровне. Доступность, точность и актуальность таких баз данных являются основой для решения задач в различных областях, таких, как использование природных ресурсов, защита окружающей среды, контроль городской застройки и т.п. Тем не менее существующие способы обновления баз данных являются довольно медленными и дорогостоящими из-за трудоемкого процесса дешифрирования фотоснимков и большого объема полевых работ, требующихся в большинстве

случаев. В последние годы особенно остро ощущается необходимость в улучшении качества данных дистанционного зондирования Земли, повышении их пространственного, спектрального и радиометрического разрешения, а также степени их доступности. Новые условия требуют развития новых методов и алгоритмов для работы с такой информацией и для ее эффективного использования при решении конкретных задач, связанных с картографированием и обновлением баз геоданных.

Космические снимки используются для картографирования земельных ресурсов и растительного покрова уже много лет из-за высокой периодичности получения данных, а зачастую также из-за более низкой стоимости по сравнению с аэрофотоснимками. Традиционные методы дешифрирования снимков предусматривают присваивание каждому пикселю тематической характеристики, основанной на различных цифровых значениях пикселя в разных спектральных каналах. Этот метод подходит для однородных классов или категорий, но не под-

\* Сокращенный перевод с английского языка. Статья была представлена на конкурс «8-Band Research Challenge», проведенный компанией DigitalGlobe, и отмечена жюри как одна из лучших. Оригинал статьи опубликован на сайте компании DigitalGlobe ([http://www.digitalglobe.com/downloads/8bc/Paper8bandWW2Challenge\\_LARuiz\\_et\\_al\(CGAT\)def.pdf](http://www.digitalglobe.com/downloads/8bc/Paper8bandWW2Challenge_LARuiz_et_al(CGAT)def.pdf)). Перевод подготовлен к публикации Б.А. Дворкиным (Компания «Совзонд»)

Рис. 1.  
Космический снимок WorldView-2  
(ближний инфракрасный канал)  
на исследуемый район



ходит для классов, характеризующихся высокой степенью спектрального смешения или неоднородным распределением элементов ландшафта. Использование текстурного анализа и текстурных характеристик (Haralick et al., 1973; Laws, 1985; Curran, 1988; etc.) улучшает результаты в таких областях, но также имеет ряд ограничений.

Благодаря новым спутниковым сенсорам, обладающим высоким разрешением, за последнее десятилетие увеличилось количество проектов в области землепользования, инициированных различными министерствами и ведомствами, нацеленных на создание и обновление баз геоданных на региональном, национальном и международном уровнях. Однако в случае использования космических снимков высокого разрешения для этого класса задач возникают две основные практические и методологические проблемы. Во-первых, чем выше детализация изображений, тем выше внутренняя вариативность соответствующих областей или участков. Во-вторых, спектральные характеристики большинства космических снимков высокого разрешения обычно ограничиваются тремя видимыми и одним инфракрасным каналами в дополнение к одному панхроматическому каналу, которые во многих случаях используются только для получения цветного изображения. Эта про-

блема особенно актуальна для картографирования растительности, а также для проектов, охватывающих крупные географические объекты с разнообразной растительностью.

Проблему внутренней вариативности можно решить, используя объектно-ориентированные методы классификации. Они позволяют учитывать и, более того, эффективно использовать пространственное распределение элементов, составляющих конкретный ландшафт или административную единицу, с целью повышения точности классификаций земельных ресурсов и растительного покрова. При использовании этих методов каждая пространственная единица (например, многоугольник) анализируется в целом, что позволяет связывать значения пикселей внутри границ и создавать новые дескриптивные характеристики, дающие информацию о внутренней вариативности и распределении элементов в объекте. При этом, с одной стороны, могут быть рассмотрены текстурные характеристики в виде уникального значения для объекта и могут быть выделены новые структурные характеристики, что дает информацию о пространственном расположении элементов внутри объектов. С другой стороны, проблема ограничения спектральной информации, типичная для спутниковых сенсоров высокого разрешения, может

быть решена путем использования новых космических снимков, полученных со спутника WorldView-2, в которых имеется 8 спектральных каналов в видимой и инфракрасной областях электромагнитного спектра, а разрешение в панхроматическом режиме составляет примерно 0,5 м.

Цель данной работы состояла в изучении объектно-ориентированных методов классификации по космическим снимкам с высоким спектральным и пространственным разрешением, полученным со спутника WorldView-2, и в оценке эффективности этих методов в классификации земельных ресурсов

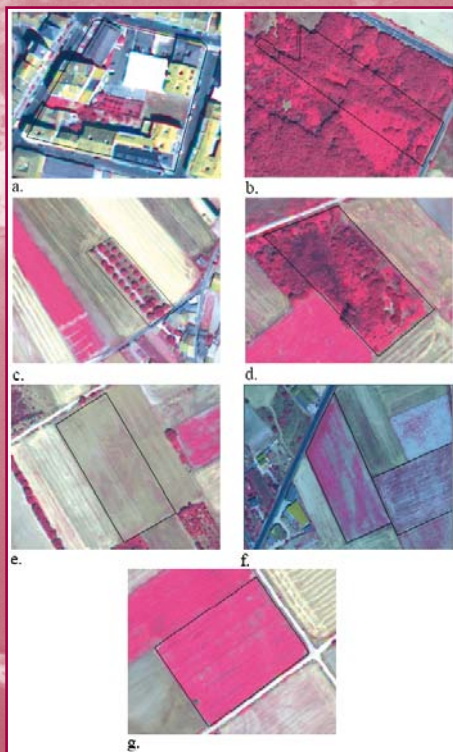


Рис. 2.  
Цветные инфракрасные композиты эталонных участков:  
а) городская застройка; б) леса; в) фруктовые сады;  
д) кустарники; е) пашни; ф) редкая растительность;  
г) орошаемые поля

и растительного покрова для автоматического создания и обновления базы геоданных. Использовался снимок, покрывающий площадь в 100 кв. км. Исследуемая область расположена в Галисии (северо-запад Испании). Эта область выбрана из-за разнообразия ландшафта (рис. 1) и доступности данных полевых наблюдений.

Сенсор WorldView-2 обеспечивает получение панхроматических (разрешение 0,46 м) и мультиспектральных изображений в восьми каналах видимой и ближней инфракрасной областях спектра (разрешение 1,84 м): фиолетовый или coastal (0,400–0,450 нм), синий (0,450–0,510 нм), зеленый (0,510–0,580 нм), желтый (0,585–0,625 нм), красный (0,630–0,690 нм), крайний красный или red-edge (0,630–0,690 нм), ближний ИК-1 (0,770–0,895 нм), ближний ИК-2 (0,860–1,040 нм).

Границы исследуемых участков были получены из испанского земельного кадастра SIGPAC. Площадки составляют сплошную область, входящую в участок сельскохозяйственного назначения. Кроме того, использовались эталоны, полученные при проведении полевых исследований. Эти эталоны имеют форму прямоугольников со сторонами от 350 до 500 м.

## МЕТОДИКА ИССЛЕДОВАНИЯ

### Выделение классов и отбор эталонов

Было выделено в общей сложности семь классов с учетом классификаций, используемых местными и национальным картографическими агентствами, разрабатывающими базы геоданных в исследуемой области (рис. 2). Большая часть проверочных эталонов была выбрана по результатам полевых исследований. Кроме того, были использованы дополнительные данные, полученные путем дешифрирования снимков, для обеспечения полноты представления некоторых классов. В общей сложности было использовано 1172 пробных участка.

Для обеспечения точности классификации участки площадью менее 60 кв. м были исключены.

### Выделение характеристик по объектам

Использование эффективных дескриптивных характеристик очень важно для точной классификации. На этом этапе каждый участок был обработан

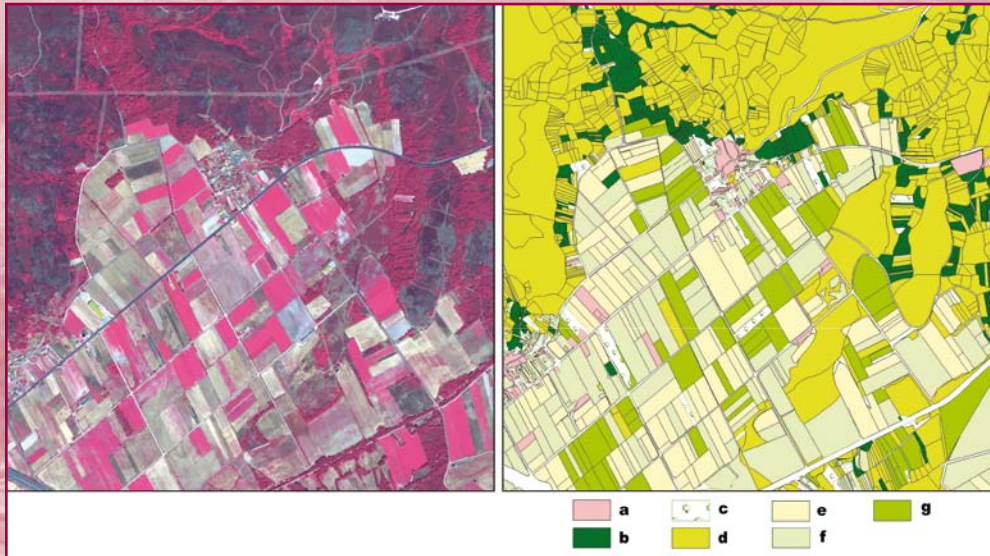


Рис. 3. Пример результата классификации. Слева – цветной композит (ближний инфракрасный канал), справа – тематическая карта (условные обозначения: а) городская застройка; б) леса; с) фруктовые сады; д) кустарники; е) пашни; ф) редкая растительность; г) орошаемые поля)

отдельно для выделения дескриптивных характеристик, определяющих текущий тип его землепользования. Подход, использованный в данном исследовании при расчете характеристик, можно считать многосторонним, т.к. было выделено три категории характеристик: спектральные, текстурные и структурные.

**Спектральные характеристики**, использованные в исследовании, дают информацию об отражательных свойствах объектов в видимой и ближней инфракрасной областях, которые зависят от типа земной поверхности, состояния растительности, состава почвы, строительных материалов и т.д. Эти характеристики особенно полезны для определения спектрально однородных классов, таких, как зеленые культуры или земли под паром; при этом снимки WorldView-2 предоставляют более конкретную спектральную информацию по сравнению с другими доступными сенсорами. Среднее отклонение, стандартное отклонение, минимальное и максимальное значения были рассчитаны на основе восьми спектральных

каналов WorldView-2, а также на основе приведенной разности вегетативного индекса (NDVI), вычисленной для красного и ближнего ИК-1 каналов.

**Текстурные характеристики** дают информацию о пространственном распределении параметров интенсивности в изображении, они полезны для количественного определения таких свойств, как однородность, контрастность, однородность, для каждого объекта. Эти свойства дают дополнительную информацию об объекте. Текстурные характеристики были рассчитаны для каждого объекта по панхроматическому каналу для оптимального использования его высокого пространственного разрешения.

**Структурные характеристики** дают информацию о пространственном расположении различных элементов внутри объекта, выраженную в хаотичности или равномерности распределения элементов. Например, упорядоченность, или правильный рисунок в схемах искусственных ландшафтов, таких, как схемы посадки зерновых культур и деревьев на сельскохозяйственных площадях.

### Классификация с использованием дерева решений

Объекты были классифицированы с использованием дерева решений, представляющего собой набор условий, организованных в иерархическую структуру таким образом, что класс, присвоенный объекту, может быть определен с помощью прослеживания условий, которые выполняются от корней дерева (начальный набор данных) к любому из его листьев (присвоенный класс).

Процесс построения дерева решений начинается с разделения набора эталонов с использованием взаимоисключающих условий. Каждая из образованных подгрупп итерационно разделяется до тех пор, пока все новые сформированные подгруппы не станут однородными, т. е. когда все элементы в подгруппе будут принадлежать одному классу. Этот процесс повторяется до тех пор, пока не будет выполнено разделение исходных данных на однородные подгруппы с использованием соотношения прироста информации в качестве критерия разбиения. Этот критерий использует информационную теорию для оценки размера поддеревьев для каждого возможного атрибута и выбирает атрибут с наибольшим ожидаемым приростом информации, т.е. атрибут, приводящий к наименьшему ожидаемому размеру поддеревьев.

В общей сложности для классификации были сформированы десять деревьев решений с использованием метода возрастания с несколькими классификаторами, позволяющего повысить точность классификаторов. Метод возрастания с несколькими классификаторами основан на соотношении весов с проверочными пробами. Чем выше вес пробы, тем сильнее его влияние в классификаторе. После каждого построения дерева вектор весов корректируется с целью определения эффективности модели. Таким образом, проба, классифицированная ошибочно, увеличивает свой вес, в то время как вес правильно классифицированной пробы уменьшается. Соответственно модель, полученная при следующей итерации, учитывает пробы, ошибочно классифицированные в предыдущем шаге. После построения набора деревьев решений каждому объекту назначается класс с учетом расчетной погрешности, определенной при построении каждого дерева.

### Оценка

Оценка классификации основывалась на анализе матрицы неточностей путем сравнения класса, назначенного каждому эталону, с информацией, содержащейся в эталонной базе данных. Была рассчитана общая точность классификаций, а также точность производителя и точность пользователя для каждого класса.

Поскольку область исследования была небольшой и для корректной оценки требовался представительный набор испытательных данных, использовался метод перекрестной проверки достоверности с исключением объектов по одному (leave-one-out). Этот метод основан на единственном наблюдении исходной пробы в качестве данных для проверки достоверности и оставшихся наблюдений в качестве проверочных данных. Процесс повторяется таким образом, что каждое наблюдение в пробе используется один раз в качестве данных для проверки достоверности.

Из-за большого количества определенных параметров некоторые характеристики могут оказаться избыточными для эффективного описания объектов, используемых в нашем исследовании. Включение этих параметров может внести помехи в классификацию. Поэтому было выполнено исследование связей между характеристиками и их вкладом в точность классификации. Был использован статистический линейный дискриминационный анализ для определения важности характеристик для конкретной задачи классификации и для каждого конкретного случая применения.

### РЕЗУЛЬТАТЫ И КОММЕНТАРИИ

Анализ точности для каждого независимого класса при использовании спектральных, текстурных и структурных характеристик показывает, что классы «городская застройка», «кустарники» и «фруктовые сады» выделяются с высокой точностью (около или более 80). Начальная точность для классов «леса» и «орошаемые поля» составляет 50%, но после добавления характеристик в модель их точность повышается. Класс «редкая растительность» по расчетам характеризуется очень низкой точностью при использовании двух дополнительных дискриминационных характеристик, но при использовании среднего значения характеристики для крайнего красного кана-

Таблица 1

**Общая точность классификации при использовании различных сочетаний дескриптивных характеристик**

Характеристики	Общая точность
Спектральные	84,0
Спектральные и текстурные	87,2
Спектральные, текстурные и структурные	89,0

ла расчетная точность значительно возрастает (до 90%). Для достижения максимальной расчетной точности класса «пашни» требуется наибольшее число дескриптивных характеристик.

Общая точность классификации при использовании различных сочетаний дескриптивных характеристик указана в табл. 1.

При использовании только спектральных характеристик достигается общая точность 84%, но при этом для класса «фруктовые сады» достигается очень низкая точность. Этот класс характеризуется высокой степенью смешиваемости с классами «леса», «кустарники» и «редкая растительность». Это объясняется тем, что участки фруктовых садов часто имеют маленькие площади и характеризуются смешанным типом землепользования. Классы «леса» и «кустарники» часто смешиваются при классификации из-за сходных спектральных свойств. По этой же причине при классификации «редкая растительность» и «орошаемые поля» также характеризуются заметной степенью смешиваемости.

Сочетание спектральных и текстурных характеристик позволяет достичь повышения общей точности классификации до 87,2%, главным образом за счет значительного улучшения результатов классификации класса «фруктовые сады». В результате наблюдается некоторый прирост точности для классов «леса» и «кустарники».

При последующем включении структурных характеристик в процесс классификации наблюдается умеренный прирост точности для всех классов. В результате общая точность классификации достигает 89% (рис. 3).

## ВЫВОДЫ

Использование крайнего красного спектрального канала сенсора WorldView-2 однозначно повышает точность классификации участков сельскохозяйственных культур и является одним из самых эффек-

тивных средств дифференциации этих классов земельных участков. Учитывая тот факт, что область исследования представляет собой сочетание участков с растительностью и без растительности, можно утверждать, что полный набор спектральных каналов дает разнообразную спектральную информацию, повышая общую точность классификации для представленной задачи, в которой были выделены семь классов землепользования.

Тектурные и структурные характеристики, выделенные по панхроматическому каналу для каждого отдельного участка, обеспечивают очень интересный синергетический эффект со спектральной информацией при классификации объектов. Информация, которую они добавляют, является дополнительной к спектральной, что позволяет извлечь преимущество из внутренней вариативности при характеристике участков. Это объясняется не только очень высоким пространственным разрешением панхроматического изображения, но также радиометрическим разрешением 11 бит, которое повышает чувствительность расчетных переменных и контрастность захваченной информации.

Проведенное исследование показало, что используемая методика является репрезентативной для многих текущих проектов, связанных с созданием и обновлением базы данных земельных ресурсов и растительного покрова. Полученные хорошие результаты (общая точность близка к 90%) позволяют с оптимизмом рассматривать перспективу использования данных со спутника WorldView-2. Сочетание высокого спектрального, пространственного и радиометрического разрешения данных и исчерпывающего извлечения информации при использовании объектно-ориентированных методов и многовариантного подхода, основанного на спектральных, текстурных и структурных характеристиках, обеспечивает многообещающие перспективы для их практического применения.