

Экология

УДК 57.036:504.064.37

Количество классов при проведении кластерного анализа космических снимков при мониторинге состояния лесов

Ю.Ф. Рожков*, М.Ю. Кондакова**

*Государственный природный заповедник «Олекминский», 678100, Олекминск, ул. Логовая, 31, Россия

**Гидрохимический институт, 344090, г. Ростов-на-Дону, пр. Стачки, 198, Россия

e-mail: olekmazap-nauka@yandex.ru, vesna-dm@mail.ru

Аннотация. Проведена оценка возможности совместного использования двух инструментов кластерного анализа в мониторинге состояния лесов. Мультиспектральные космические снимки высокого разрешения Landsat TM/ETM+ и RGB-покрытие DigitalGlobe сверхвысокого разрешения и их фрагменты подвергались двухэтапной обработке. Вначале проводилась неуправляемая классификация методом ISODATA (Iterative Self-Organizing Data Analysis Technique). Затем проводился расчет тематической разности результатов классификации. Показана связь между количеством классов и количеством объектов, определяемых при классификации на два, четыре, шесть, десять классов. Проведено определение оптимального количества классов для выделения различных уровней структурной организации лесных экосистем. При классификации на два класса в случае снимков высокого разрешения оценивается лесистость выделенных фрагментов снимков. При классификации снимков сверхвысокого разрешения на два класса определяется соотношение между площадью кроны и площадью между кронами деревьев. При классификации снимков высокого разрешения на четыре класса выделяются подклассы более густого и редкого древостоя, подклассы с открытыми пространствами и участки, покрытые кустарниками и редколесьями.

Ключевые слова: дешифрирование космических снимков, классификация ISODATA, тематическая разность, количество классов при проведении кластеризации.

The Number of Classes in the Cluster Analysis of Space Images for Forest State Monitorin

Yu.F. Rozhkov*, M.Yu. Kondakova**

*State Nature Reserve «Olekminsky», 31, Logovaya Str., Olekminsk, 678100, Russia

**Hydrochemical Institute, 198, Stachki Ave., Rostov-on-Don, 344090, Russia

e-mail: olekmazap-nauka@yandex.ru, vesna-dm@mail.ru

Abstract. The possibility of combination of two cluster analysis tools for forests state monitoring was assessed. Multispectral space images of high resolution Landsat TM / ETM + and ultra-high resolution RGB-coating DigitalGlobe and their fragments were subjected to two-stage processing. At first unsupervised classification was performed using the ISODATA (Iterative Self-Organizing Data Analysis Technique) method. Then the thematic difference in the classification results was calculated. A relationship between the number of classes and the number of objects defined in the classification into two, four, six, ten classes is shown. The optimal number of classes for distinguishing different levels of structural organization of forest ecosystems are determined. When classifying into two classes in case of high-resolution images forest cover of selected fragments of the images is estimated. When classifying ultra-high resolution images into two classes, the ratio between the area of crowns and the area between the tree crowns is determined. When classifying high-resolution images into four classes, subclasses of more dense and sparse stand, subclasses with open spaces and areas covered with shrubs and woodlands are distinguished.

Key words: interpretation of space images, ISODATA classification, thematic difference, number of classes in clustering.

Введение

Для решения вопросов оперативного мониторинга состояния лесных экосистем и оценки

биоразнообразия широко используется дешифрирование космических снимков [1, 2]. В качестве инструментов дешифрирования распространён кластерный анализ [3]. При выборе инструментов кластерного анализа исследователи выбирают либо классификацию «с обучением» (с методами: минимального расстояния, макси-

РОЖКОВ Юрий Филиппович – к.х.н., зам. директора; КОНДАКОВА Мария Юрьевна – к.б.н., с.н.с.

мального правдоподобия, дистанции Махаланобиса, параллелепипедов), либо неуправляемую классификацию с методами: К-средних и ISODATA. Метод классификации ISODATA оперирует либо с пикселями, ранжируя их по значениям спектральной яркости отраженного света (в случае мультиспектральных снимков, композитных снимков, монохроматических снимков) [4], либо проводит классификацию индексов – значений отношений между спектральными каналами. В этом случае обработке подвергается распределение цифровых значений отношений, характеризующих каждый пиксел (кластер). При обработке кластерным анализом по методу ISODATA не имеет значения разрешающая способность снимков. Исследователи успешно проводят анализ снимков как сверхвысокого разрешения [5], так и разного разрешения (среднего и сверхвысокого) [6] не только в оптическом диапазоне, но и в инфракрасном спектральном [7].

Не определен только вопрос об оптимальном количестве классов при проведении кластеризации по методу ISODATA. Предполагается, что количество классов задается исследователем в соответствии с целями дешифрирования. Целью представленного ниже исследования является определение оптимального количества классов для выделения различных уровней структурной организации лесных экосистем.

Материалы и методы

Для осуществления непрерывного мониторинга за состоянием бореальных лесов использовалось дешифрирование мультиспектральных космических снимков высокого разрешения Landsat TM/ETM+ (разрешение 30 м/пикс.) и RGB-покрытия DigitalGlobe сверхвысокого разрешения (разрешение 0,6 м/пикс.). Исследования проводились на территории Юго-Западной Якутии площадью 1 млн. га. Для обработки космических снимков использовался пакет программ ENVI-4.0, ArcGis 10.3, ArcView-3.3 с модулями Image Analyst, Spatial Analyst. В качестве показателей, с помощью которых осуществлялся мониторинг состояния лесов, были выбраны два инструмента кластерного анализа: классификация ISODATA и тематическая разность пикселей [8].

В основу инструмента неуправляемой классификации мультиспектральных снимков заложен метод кластерного анализа ISODATA, который использует установленное число итераций (перегруппировка пикселей по классам) и порог сходимости для выбранных классов. Выбранный метод неуправляемой классификации является самоорганизующимся, так как

исследователь указывает только количество классов, на которые нужно разбить весь массив данных [9]. В качестве объектов для анализа выступают пиксели снимков высокого разрешения (которые соответствуют по шкале биоразнообразия урочищам. При разрешающей способности снимка в 30 м размеры пикселя соответствуют площади 900 м²) и пиксели снимков сверхвысокого разрешения (которые по шкале биоразнообразия соответствуют отдельным организмам. Размеры пикселя соответствуют площади 0,36 м²).

Кроме того, для проведения анализа использовались серии фрагментов космических снимков с разной площадью по возрастанию. Сохранялись фрагменты по размеру экрана дисплея с разным масштабом [8]. Ниже представлены результаты совместного применения двух инструментов кластерного анализа пикселей.

Результаты и обсуждение

1. Анализ результатов классификации на два класса снимков высокого разрешения LandsatETM+. Для определения изменения структурной организации экосистем для снимков высокого разрешения были подготовлены серии фрагментов космических снимков, каждый из которых подвергался обработке (рис. 1–3).

При проведении неуправляемой классификации по методу ISODATA на два класса мультиспектральных снимков программа разделяет все пиксели по значениям спектральной яркости отраженного света на два диапазона: 0–50 % и 50–100 %. В 1-й класс попадают пиксели с наименьшими и малыми значениями спектральной яркости (в случае покрытых лесом территорий – это площади с густым древостоем). Во 2-м классе сосредоточены пиксели с высокими значениями оптической плотности (это площади с открытыми пространствами, редколесья и пустоши). Классификация на два класса позволяет определить индекс лесистости [1] как отношение площади, покрытой лесной растительностью, к общей площади: $D = df/S$, где D – лесистость; df – площадь, покрытая лесной растительностью, м²; S – общая площадь территории, м². В случае индекса NDVI получают два класса – низкопродуктивные и высокопродуктивные. Распределение пикселей на два класса представлено в табл. 1–3.

Результаты классификации фрагментов снимков высокого разрешения на два класса показывают соотношение между древостоем и пустошами. Угол наклона кривых при изменении масштаба связан с изменением лесистости территории. Чем круче угол наклона кривой 1-го

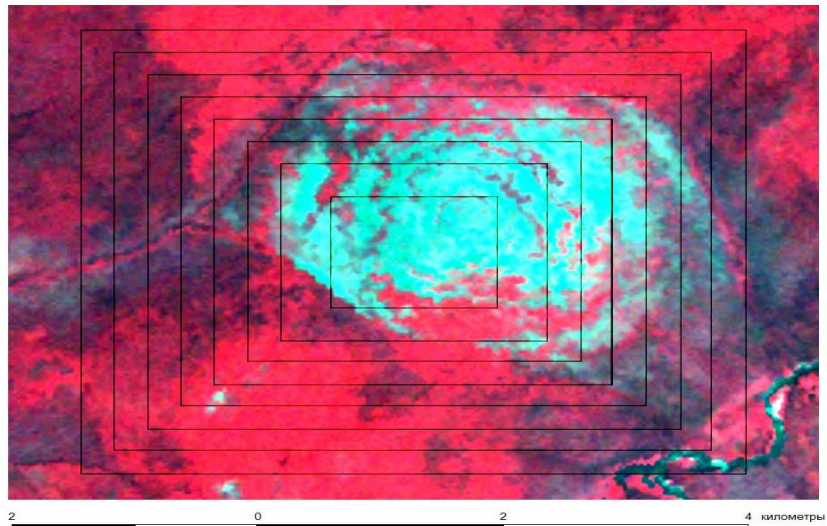


Рис. 1. Голец Тулая. Изменение масштабов на снимке от 1:5000 до 1:20000

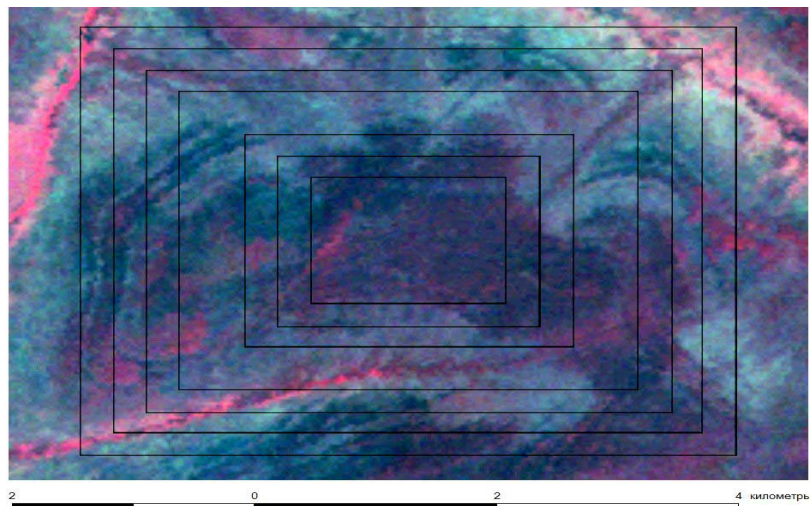


Рис. 2. Исток ручья. Изменение масштабов на снимке от 1:4000 до 1:20000

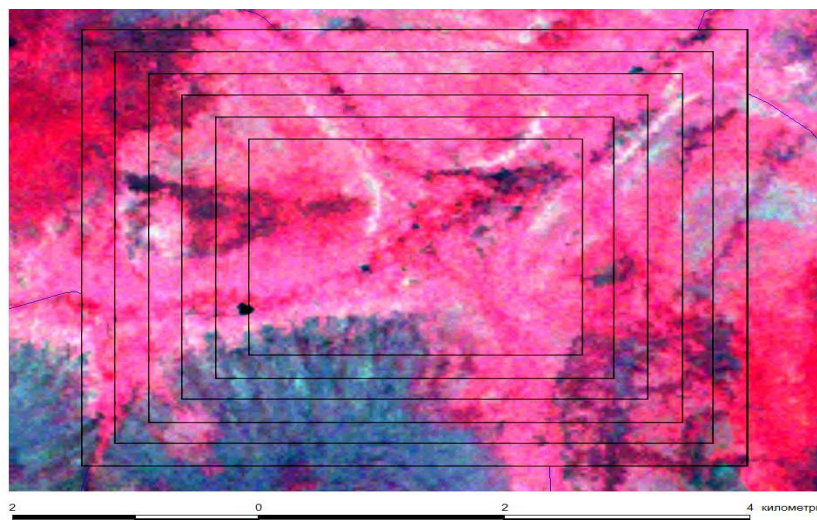
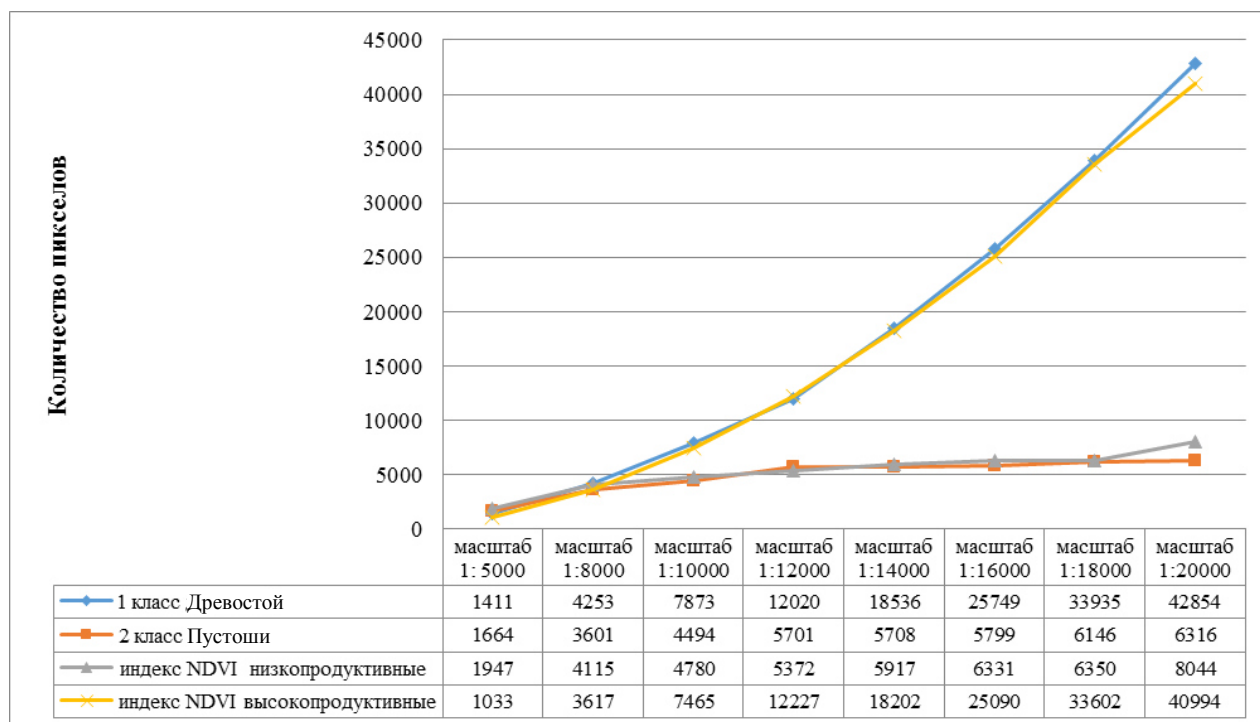


Рис.3. Ерники в долинах ручьев. Изменение масштабов на снимке от 1:4000 до 1:20000

КОЛИЧЕСТВО КЛАССОВ ПРИ ПРОВЕДЕНИИ КЛАСТЕРНОГО АНАЛИЗА

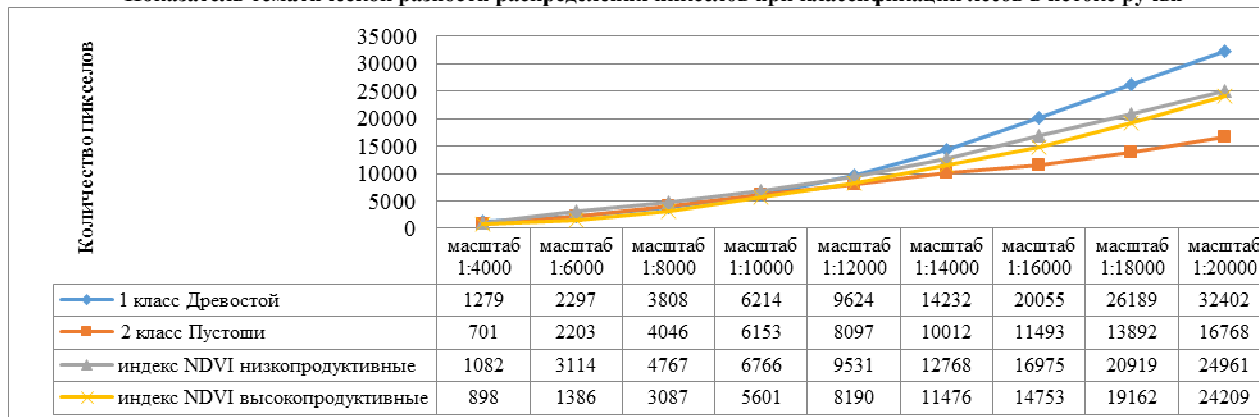
Т а б л и ц а 1

Показатель тематической разности распределения пикселей при классификации лесов вокруг гольца Тулаях



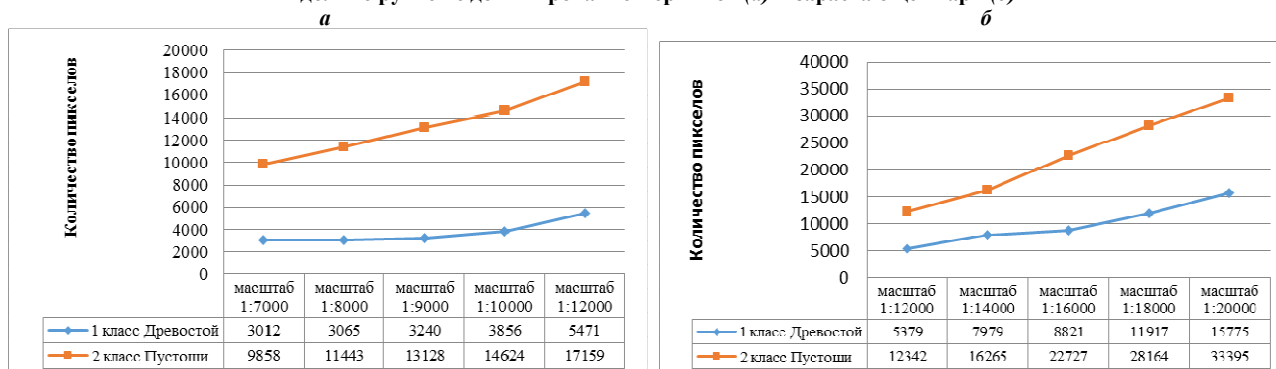
Т а б л и ц а 2

Показатель тематической разности распределения пикселей при классификации лесов в истоке ручья



Т а б л и ц а 3

Показатель тематической разности распределения пикселей при классификации лесов в долине ручьев с доминированием ерников (а) и зарастающей гари (б)



класса (древостой) с увеличением площади (масштаба), тем больше лесистость территории.

В случае территорий с доминированием пустошей (табл. 3) наблюдается превышение количества пикселей 2-го класса (пустоши) над пикселями 1-го класса (древостой). При этом отмечается сходный характер изменений кривых при классификации на 2 класса с увеличением площадей как для территорий с доминированием ерников, так и для зарастающих гарей.

2. Анализ результатов классификации на два класса снимков сверхвысокого разрешения (RGB покрытие DigitalGlobe). Были подготовлены серии фрагментов снимков сверхвысокого разрешения с возрастающим масштабом (рис. 4–6).

При проведении классификации фрагментов снимков сверхвысокого разрешения на два класса формируются классы пикселей, отражающие площади крон (2-й класс) и площади пространства между кронами (1-й класс). Результаты изменения распределения пикселей с изменением масштаба на выделенных фрагментах представлены в табл. 4,5.

3. Анализ результатов классификации на четыре класса снимков высокого разрешения LandsatETM+. При классификации на четыре класса происходит разделение всех пикселей на четыре диапазона значений спектральной яркости с шагом в 25 % (0–25; 25–50; 50–75; 75–100). Причем при ранжировании учитываются не количество пикселей, а значения спектральной яркости у каждого пикселя. Поэтому распределение пикселей в каждом диапазоне неравномерное, зависящее от свойств объекта классификации. В случае гольца Тулаях из класса 1 (древостой) формируются три подкласса, а из класса 2 (пустоши) – два (табл. 6).

Анализ этих двух кривых показывает, что подкласс 3 соответствует открытым пространствам (каменистым россыпям вокруг вершины гольца) и, начиная с масштаба 1:12000, доля этих россыпей резко уменьшается. Тогда как подкласс 4 соответствует редколесьям и доля его возрастает с увеличением площади фрагментов снимка (рис.1).

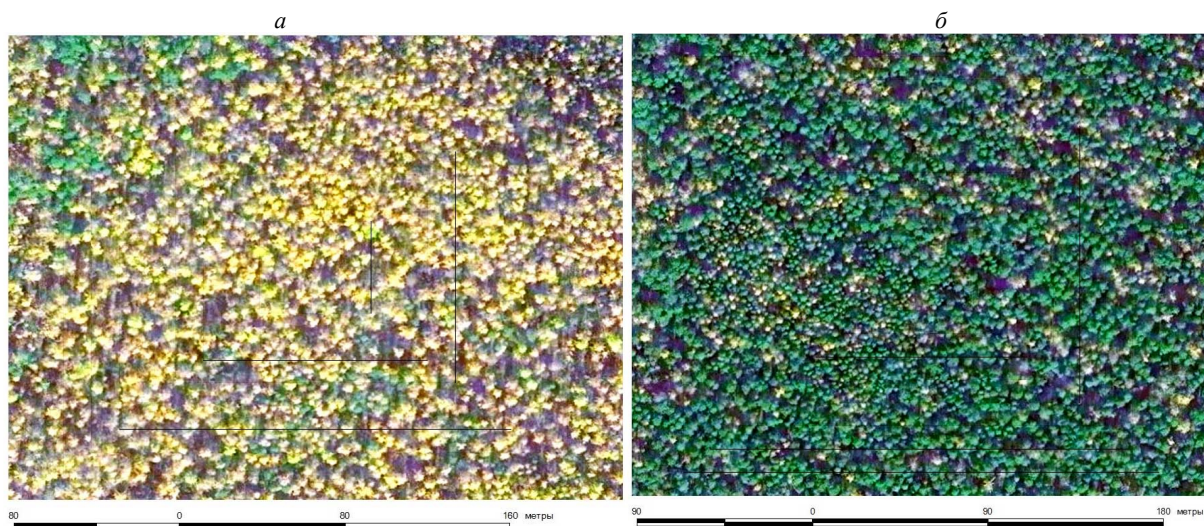


Рис.4. Фрагменты леса с доминированием лиственницы(а) и сосны обыкновенной (б). Изменение масштабов на снимке от 1:200 до 1:900

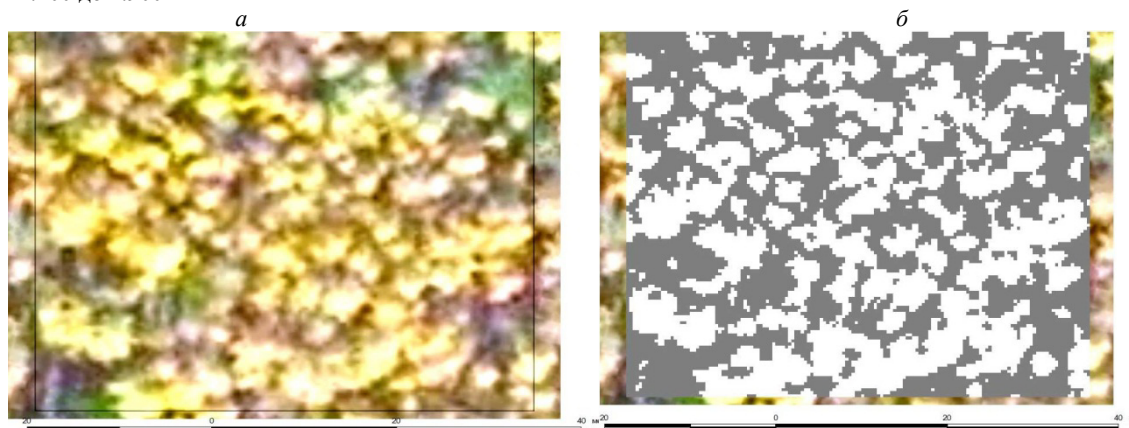


Рис.5. Фрагмент леса с доминированием лиственницы (а), результаты классификации фрагмента на 2 класса (б). Масштаб 1:200

КОЛИЧЕСТВО КЛАССОВ ПРИ ПРОВЕДЕНИИ КЛАСТЕРНОГО АНАЛИЗА

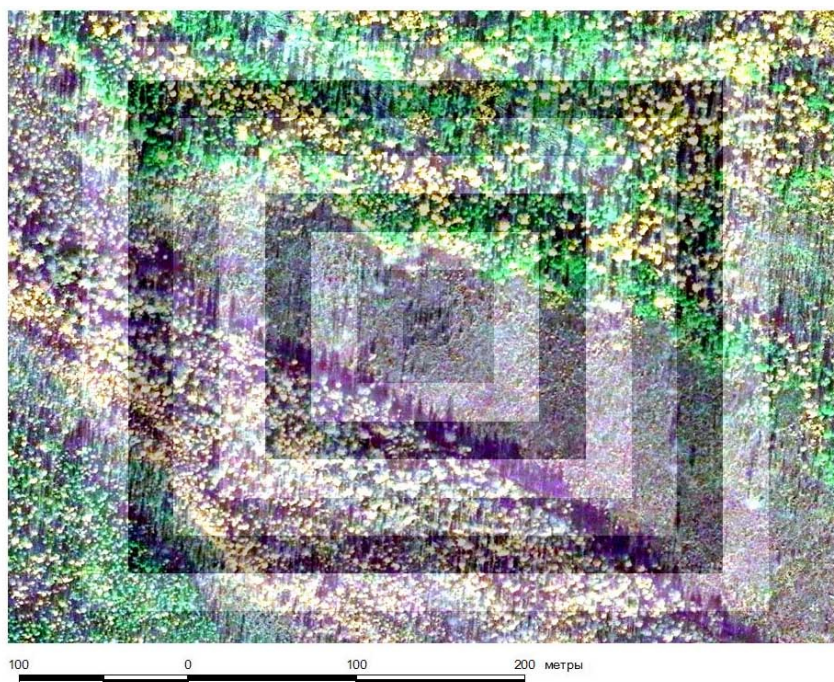
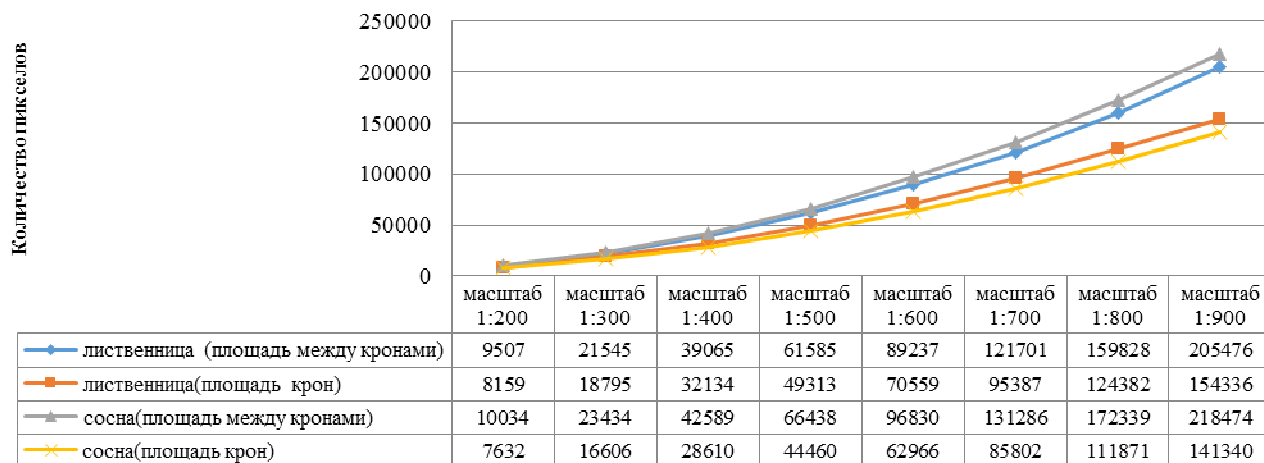


Рис.6. Фрагмент поймы ручья. Изменение масштабов от 1:100 до 1:1500

Т а б л и ц а 4

Результаты классификации на два класса фрагментов снимков с доминированием сосны и лиственницы



Т а б л и ц а 5

Результаты классификации на два класса фрагмента снимка поймы ручья

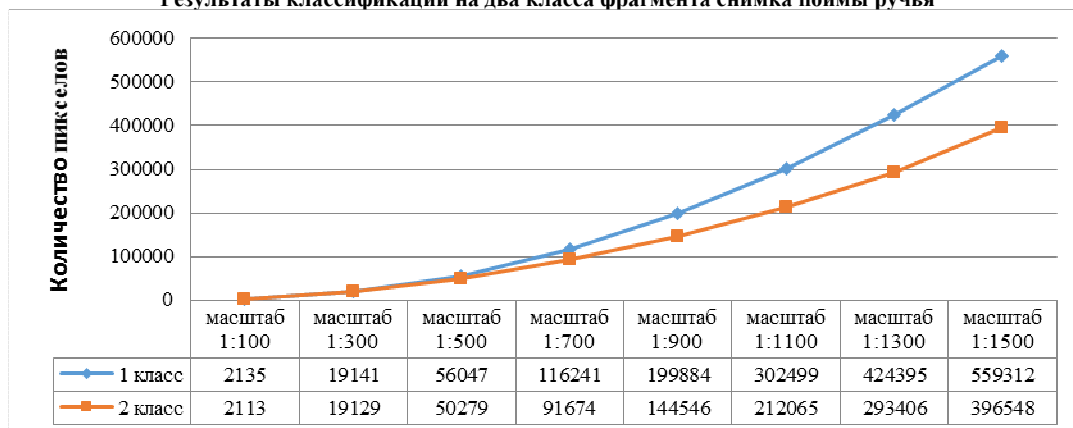


Таблица 6

Показатель тематической разности распределения пикселей при классификации лесов на четыре класса вокруг гольца Тулаях (пустоши)

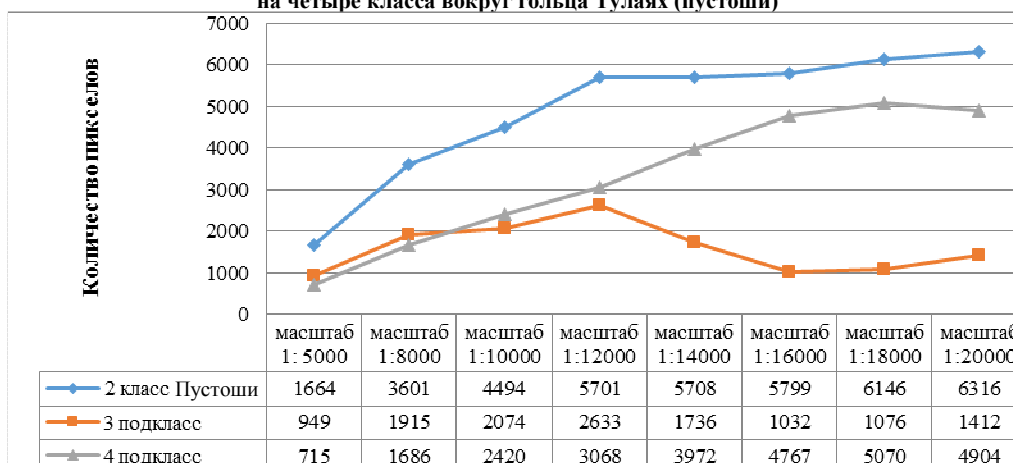
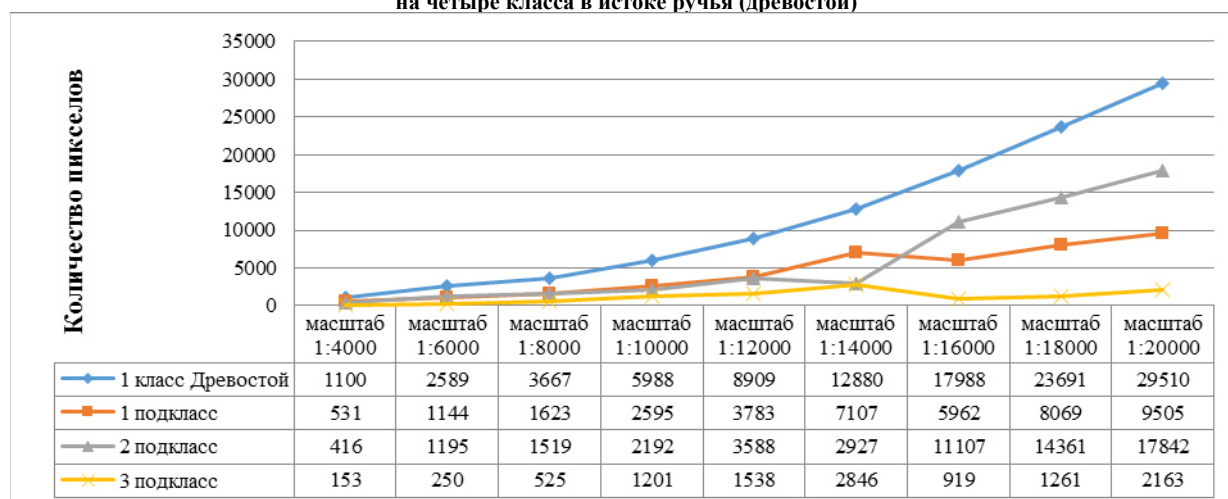


Таблица 7

Показатель тематической разности распределения пикселей при классификации лесов на четыре класса в истоке ручья (древостой)



В случае серии фрагментов в истоке ручья (рис. 2) из класса 1 (древостой) формируются три подкласса (табл. 7) и из класса 2 (пустоши) – три.

Сравнение кривых изменения подклассов древостоя при увеличении площади фрагмента показывает, что, начиная с фрагмента масштаба 1:14000, резко увеличивается доля подкласса 2 в общем количестве пикселей. На рис. 2 видно, что более густой древостой в центре окружен более редким лесом по периметру. Это соответствует и градации пикселей по спектральной яркости. Первый подкласс соответствует наиболее густым лесам и наименьшей спектральной яркости, второй – более редким древостоям, которые начинают преобладать, начиная с фрагмента с масштабом 1:14000. И наконец, при классификации на четыре класса выявляется различие в характере распределения пикселей у ерников и зарастающей гари, которые при классификации на два класса были похожи.

В сериях фрагментов из ерников и зарастающей гари из класса 1 (древостой) формируются два подкласса, а из класса 2 (пустоши) – три (табл. 8).

Видны два явных различия в характере распределения пикселей по спектральной яркости. Во-первых, в долине ручьев с доминированием ерников все три подкласса пустошей возрастают пропорционально увеличению площадей фрагментов. Тогда как доли подклассов пустошей зарастающей гари изменяются непропорционально при увеличении площадей фрагментов. Во-вторых, в серии фрагментов снимка зарастающей гари преобладает доля пикселей 4-го подкласса с наибольшей спектральной яркостью отраженного света. Это участки послепожарных пустошей с выгоревшей растительностью. А в серии фрагментов с доминированием ерников преобладает доля пикселей 3-го подкласса с меньшей спектральной яркостью, который и соответствуют ерникам.

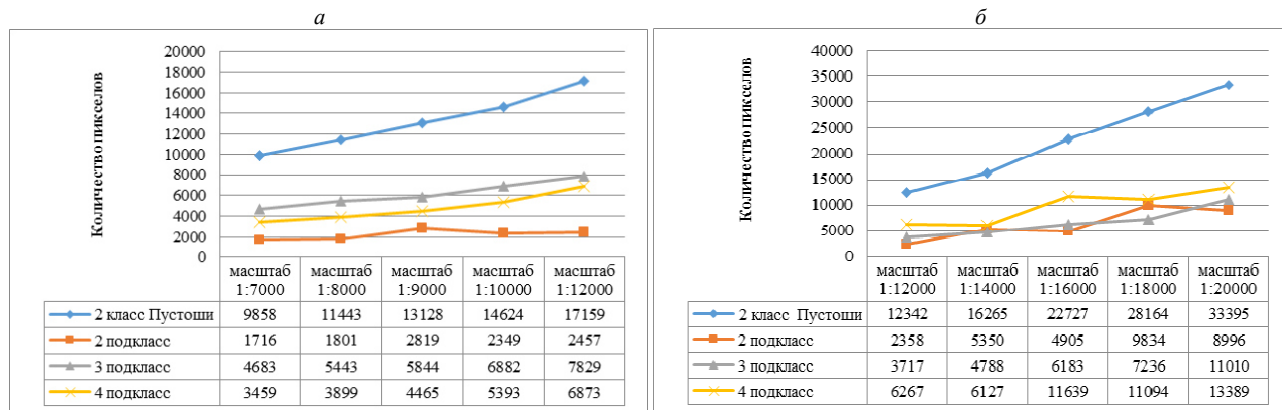
КОЛИЧЕСТВО КЛАССОВ ПРИ ПРОВЕДЕНИИ КЛАСТЕРНОГО АНАЛИЗА

4. Анализ результатов классификации на четыре класса, шесть классов снимков сверхвысокого разрешения (RGB покрытие Digital-Globe). При проведении классификации на 4 класса и 6 классов в случае серии фрагментов с доми-

нированием лиственницы, сосны обыкновенной и в пойме ручья (табл.9,10) не выявлено особенностей в распределении пикселей по спектральной яркости. Доля пикселей каждого класса возрастает пропорционально увеличению площади фрагмента.

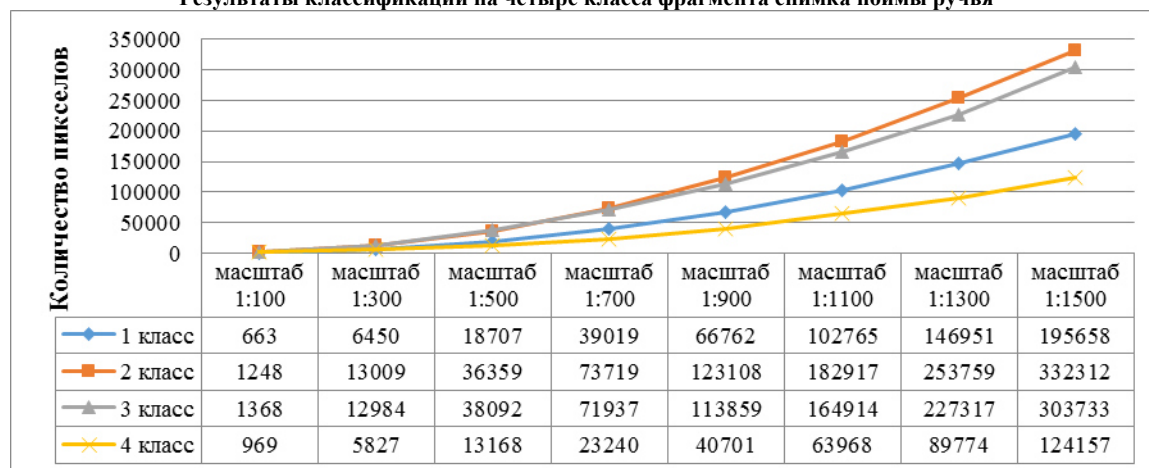
Т а б л и ц а 8

Показатель тематической разности распределения пикселей при классификации лесов на четыре класса в долине ручьев с доминированием ерников (а) и зарастающей гари (б)



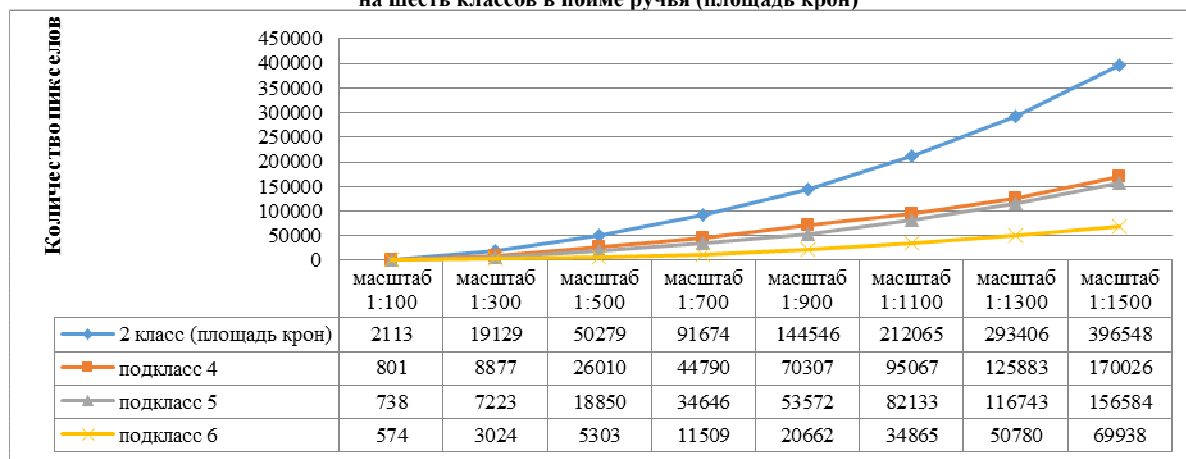
Т а б л и ц а 9

Результаты классификации на четыре класса фрагмента снимка поймы ручья

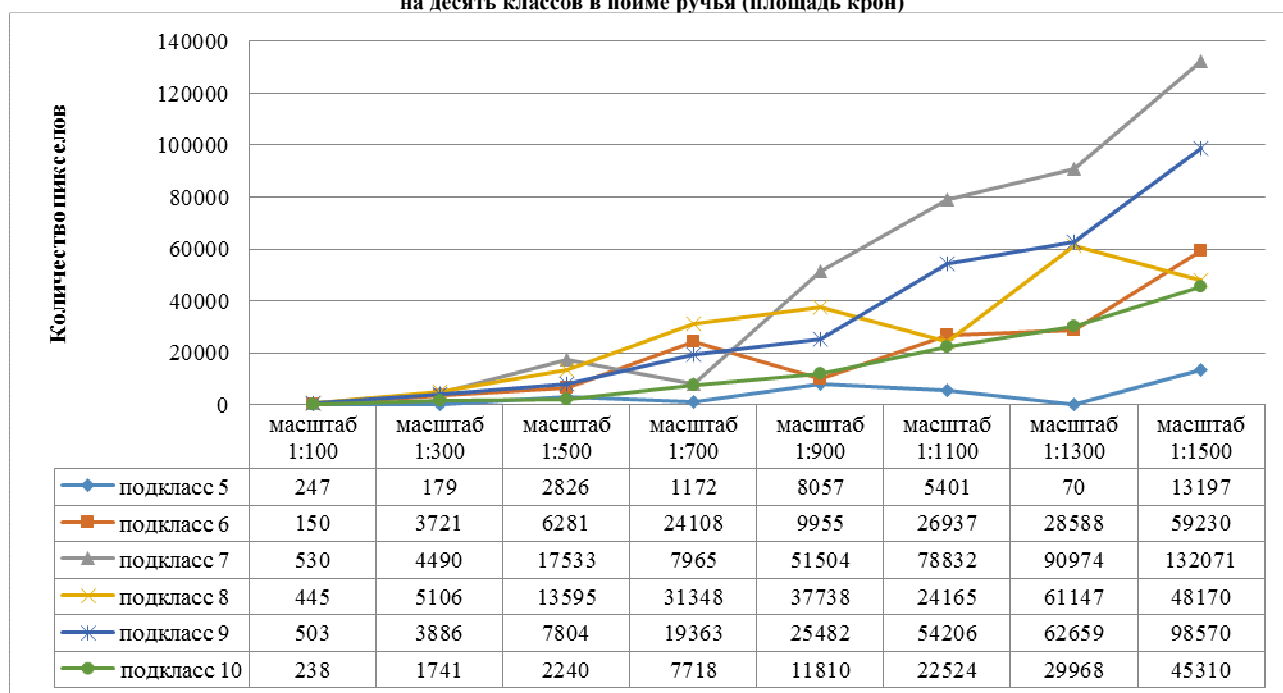


Т а б л и ц а 10

Показатель тематической разности распределения пикселей при классификации лесов на шесть классов в пойме ручья (площадь крон)



Показатель тематической разности распределения пикселей при классификации лесов на десять классов в пойме ручья (площадь крон)



5. Анализ результатов классификации на десять классов снимков сверхвысокого разрешения (RGB покрытие DigitalGlobe). При проведении классификации на десять классов все пиксели по значениям спектральной яркости разделяются на десять диапазонов по возрастанию с шагом в 10% (0–10, 10–20, 20–30, 30–40, 50–60, 60–70, 70–80, 80–90, 90–100). И подобная детализация распределения пикселей позволяет выявить особенности структурной организации лесных экосистем. В случае серии фрагментов в пойме ручья (рис. 6) 1-й класс (площади между кронами) делится на 6 подклассов, а 2-й класс (площади крон) – на 6 подклассов тематической разности (первые четыре подкласса тематической разности имеют нулевые значения) (табл. 11).

Начиная с фрагмента с масштабом 1:700 отмечается резкое увеличение доли подклассов 7 и 9 в общем количестве пикселей. Именно с этого фрагмента доля лесов начинает преобладать над пойменными редколесьями и кустарником.

Заключение

Проведено определение оптимального количества классов для выделения различных уровней структурной организации лесных экосистем. При классификации на два класса в случае снимков высокого разрешения оценивается лесистость выделенных фрагментов снимков, соотношение между древостоем и пустошами.

При классификации снимков сверхвысокого разрешения на два класса определяется соотношение между площадью крон и площадью между кронами деревьев. При классификации снимков высокого разрешения на четыре класса выделяются подклассы более густого и редкого древостоя, подклассы с открытыми пространствами и участки, покрытые кустарниками и редколесьями. При классификации снимков сверхвысокого разрешения на десять классов выделяются подклассы пойменных редколесий, кустарников и древостоя.

Литература

- Исаев А.С., Князева С.В., Пузаченко М.Ю., Черненко Т.В. Использование спутниковых данных для мониторинга биоразнообразия лесов // Исследование Земли из космоса. 2009. № 2. С. 1–12.
- Литинский П.Ю. Трехмерное моделирование структуры и динамики таежных ландшафтов. Петрозаводск: Карельский научный центр, 2007. 107 с.
- Кочуб Е.В. Анализ методов обработки материалов дистанционного зондирования Земли // Вестник Полоцкого государственного университета. Серия Ф. Строительство. Прикладные науки. Геодезия. 2012. № 16. С. 132–140.
- Коросов А.В., Коросов А.А. Техника ведения ГИС. Приложение в экологии. Петрозаводск: Изд-во ПетрГУ, 2002. 188 с.

5. *Фомин В.В., Залесов С.В., Магасумова А.Г.* Методики оценки густоты подроста и древостоев при зарастании сельскохозяйственных земель древесной растительностью с использованием космических снимков высокого пространственного разрешения // Аграрный вестник Урала. 2015. № 1. С. 125–129.

6. *Кравцова В.И.* Методические подходы в аэрокосмических исследованиях северной границы леса // География и природные ресурсы. 2012. № 3. С. 133–139.

7. *Балдина Е.А., Грищенко М.Ю.* Методика дешифрирования разновременных космических

снимков в тепловом инфракрасном диапазоне // Вестник МГУ. Серия 5. География. 2014. № 3. С. 35–42.

8. *Рожков Ю.Ф., Кондакова М.Ю.* Мониторинг состояния лесов с использованием кластерного анализа при дешифрировании космических снимков среднего и высокого разрешения // Наука и образование. 2016. № 3. С. 95–101.

9. *ArcView Image Analysis.* Руководство пользователя. М.: Дата+, 1998. 270 с.

Поступила в редакцию 15.05.2017