РАКЕТНО-КОСМИЧЕСКОЕ ПРИБОРОСТРОЕНИЕ И ИНФОРМАЦИОННЫЕ СИСТЕМЫ
2018, том 5, выпуск 4, с. 88–93

СИСТЕМНЫЙ АНАЛИЗ, УПРАВЛЕНИЕ КОСМИЧЕСКИМИ АППАРАТАМИ, ОБРАБОТКА ИНФОРМАЦИИ И СИСТЕМЫ ТЕЛЕМЕТРИИ

УДК 528.854 DOI 10.30894/issn2409-0239.2018.5.4.88.93

Особенности применения нечетких множеств в задачах идентификации растительности по данным дистанционного зондирования

А. И. Стрыков, *к. т. н., contact@spacecorp.ru*
АО «Российские космические системы», Москва, Российская Федерация
Н. П. Морозов, *к. т. н., contact@spacecorp.ru*
АО «Российские космические системы», Москва, Российская Федерация

Аннотация. Рассмотрена методика идентификации растительности с использованием дешифровочных признаков объектов, представленных нечеткими множествами случайных реализаций спектральных характеристик объектов и функций от этих характеристик в различных спектральных диапазонах. Тематическое дешифрирование с использованием нечетких множеств осуществляется методом, названным «голосование по набору признаков». В отличие от классических методов статистической обработки (Махаланобиса и др.), предполагающих распределение значений спектральных характеристик объектов по нормальному закону, «голосование по набору признаков» позволяет идентифицировать объекты с произвольным законом распределения значений спектральных характеристик, представленных нечеткими множествами, и повысить достоверность результатов тематического дешифрирования. Алгоритм дешифрирования не только классифицирует изображение объекта, но и определяет вероятность принадлежности к выбранному классу. Это позволяет управлять порогом для классификации.

Методика реализована и апробирована в виде программного комплекса с использованием данных российской группировки космических аппаратов (КА) дистанционного зондирования «Ресурс-П» и «Канопус-В». Методика может быть использована для проведения космического мониторинга сельскохозяйственной растительности, решения экологических задач, повышения достоверности обнаружения растительности, подлежащей контролю правоохранительными органами.

Ключевые слова: идентификация объектов, дешифровочные признаки, нечеткое множество, спектральные характеристики, набор признаков, метод «голосование по набору признаков»

Application of Fuzzy Sets in Problems of Identification of Vegetation Based on Remote Sensing Data

A. I. Strykov, *Cand. Sci. (Engineering), contact@spacecorp.ru*
Joint Stock Company "Russian Space Systems", Moscow, Russian Federation
N. P. Morozov, *Cand. Sci. (Engineering), contact@spacecorp.ru*
Joint Stock Company "Russian Space Systems", Moscow, Russian Federation

Abstract. The article presents a method of thematic processing of high-resolution satellite images. The method of identification of vegetation with the use of decoding features of objects represented by fuzzy is considered. Thematic interpretation using fuzzy sets is carried out by a method called "voiting on a set of criteria". In contrast to the classical methods of statistical processing (Mahalanobis, etc.), involving the normal distribution of the values of the spectral characteristics of objects, "voiting on a set of criteria" allows to identify objects with an arbitrary law of distribution of the values of the spectral characteristics represented by fuzzy sets, and to increase the reliability of the thematic interpretation results.

The method is implemented and tested in the form of a software package using data from the Russian remote sensing spacecraft "Resurs-P" and "Kanopus-V". The technique can be used for space monitoring of agricultural vegetation, solving environmental problems, etc.

Keywords: identification of objects, decoding characteristics, fuzzy set, spectral characteristics, feature set, "voting on a set of criteria" method

Идентификация объектов растительности, по данным бортовой съемочной аппаратуры, является актуальной задачей тематической обработки данных КА дистанционного зондирования Земли (ДЗЗ). Характерный признак растительности и ее состояния — спектральная отражательная способность, характеризующаяся большими различиями в отражении излучения разных длин волн. Знания о связи структуры и состояния растительности с ее отражательными способностями позволяют использовать космические снимки для идентификации типов растительности и оценки их состояния.

«Идентификация» объектов, как правило, предполагает выделение дешифровочных признаков изображений объектов в процессе анализа участков местности на космических снимках и формирование наборов признаков, описывающих определенный класс изображений объектов для данного вида съемочной аппаратуры (эталонов). Весьма эффективным методом получения эталонов является сбор данных о спектральных отражательных свойствах интересующей растительности с проведением космической съемки заданной территории и проведением наземных измерений объектов на выделенных участках (полигонах) этой территории в небольшом временном интервале (наиболее предпочтительно — синхронно).

Для представления дешифровочных признаков объектов используются как спектральные характеристики объектов, так и функции этих характеристик (например, нормализованные дифференцированные индексы — NDVI).

Тематическое дешифрирование таких объектов использует алгоритмы, основанные на методе наименьших квадратов (метод Махаланобиса и др.). Для идентификации объектов, спектры которых существенно отличаются от других объектов и закон распределения которых достаточно точно можно аппроксимировать с помощью нормального закона во всех спектральных каналах, такие алгоритмы обеспечивают приемлемый результат. Использование алгоритмов в тематическом дешифрировании, основанных на методе наименьших квадратов, требует оценки только двух тестовых параметров значения спектральных характеристик искомых объектов в каждом канале (математического ожидания и дисперсии).

Анализ космической информации показывает, что при применении съемочной аппаратуры высокого разрешения российской группировки КА для большинства природных объектов закон распределения случайных значений спектральных признаков существенно отличается от нормального закона. На рис. 1 приведен фрагмент космического изображения, полученного в июне 2015 г. с КА «Ресурс-П» № 2 с помощью аппаратуры «Геотон-Л1».

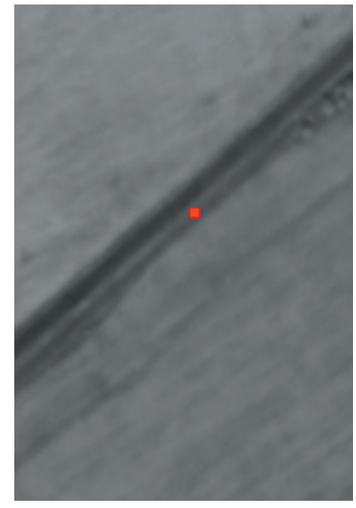


Рис. 1. Фрагмент снимка с КА «Ресурс-П»

На изображении канала 0,72–0,80 красной точкой отмечен пункт наземного наблюдения. Верхнюю часть снимка занимает изображение объекта «Гречиха», а нижнюю часть снимка — объект «Травы».

Кроме того, для большинства объектов растительности имеются маскирующие объекты с похожими спектральными характеристиками. Применение стандартных методов статистической обработки таких данных может привести к недопустимым погрешностям получения оценок анализируемых параметров объектов идентификации.

Одним из возможных способов повышения достоверности таких оценок предлагается использование нечетких множеств [1–3] в тематическом дешифрировании материалов космической съемки высокого разрешения.

Сформируем нечеткое множество, состоящее из упорядоченных пар $Yi = \{fi(x), p(fi(x)\}$, где $fi(x)$ — значение характеристической функции (логической), $p(fi(x))$ — степень истинности (вероятность) $fi(x)$.

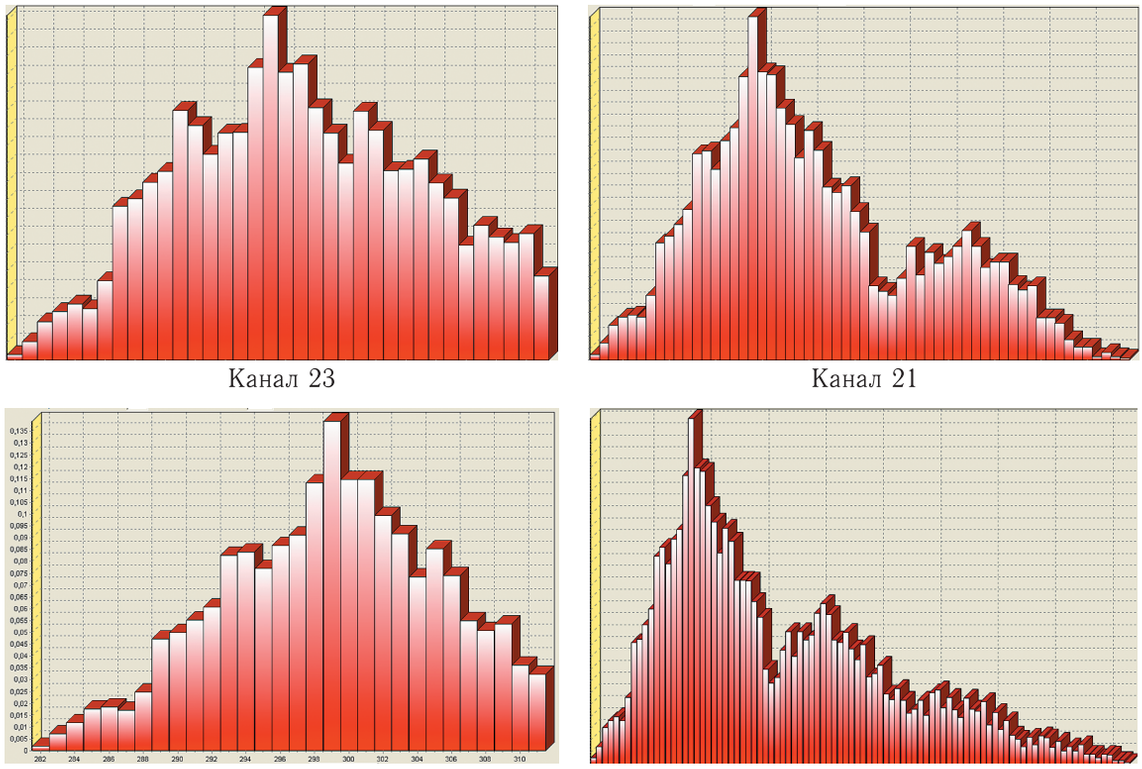


Рис. 2. Гистограммы спектральных характеристик объекта «Гречиха»

Для обработки данных множеств применим методику, названную «голосованием по набору признаков» [4]. Методика реализуется в два этапа. На первом этапе выделяется, оценивается и записывается в базу данных (БД) совокупность дешифровочных признаков, представленных в виде массивов нечетких множеств.

На втором этапе осуществляется анализ дешифровочных признаков в соответствии с методикой «голосование по набору признаков». Алгоритм анализа основан на оптимизации целевой функции разделения множества объектов на классы тестовых объектов.

Этап 1, в свою очередь, состоит из этапов 1.1 и 1.2. На этапе 1.1 формируются нечеткие множества значений спектральных характеристик объектов, а на этапе 1.2 — нечеткие множества, являющиеся функциями спектральных характеристик.

Для формирования набора дешифровочных признаков объектов должны быть проведены периодические (3–4 раза за вегетационный период) наземные наблюдения на тестовых полигонах. В результате должно быть определено географическое расположение и дано геоботаническое описание объектов в данный период вегетации. В этот же период должна быть проведена космическая съемка объектов тестовых полигонов и оценены спектральные характеристики изображений этих объектов.

Этап 1.1

Пусть имеются No объектов, которые подлежат идентификации (классификации) с использованием Nk спектральных каналов мультиспектральной или гиперспектральной космической информации (космического снимка). Значение Nk может быть меньше или равно общему числу всех каналов, предусмотренных для данного вида аппаратуры.

Рассмотрим прямоугольный фрагмент изображения объекта с номером $n \leqslant No$ на космическом снимке, состоящего из Ni строк и Nj столбцов в каждом спектральном канале. Обозначим через P_{nijk} спектральное значение пикселя (по возможности после атмосферной коррекции)

Таблица 1. Запись в БД о спектральных характеристиках объекта «Гречиха»

KDob	Col	45-52
6	13488883	416_419$1!419_422$6!422_425$10!425_428$19!428_430$14!430_433$19!433_436$10!436_439$12!439_442$7!442_444$2!430#5,68^
6	13488883	416_421$9!421_425$4!425_430$11!430_435$13!435_440$10!440_444$3!444_449$16!449_454$17!454_458$9!458_462$9!441#12,74^

в i-й строке ($i \leqslant Ni$), j — столбце ($j \leqslant Nj$) и k-м спектральном канале ($k \leqslant Nk$). Для каждого объекта с номером n получим Nk множеств случайных величин Ω_{nk} ($k \leqslant Nk$).

Для каждого множества Ω_{nk} определяются минимальное и максимальное значения пикселей в каждом канале, которые обозначим $Pmin_{nk}$ и $Pmax_{nk}$. Для каждого множества Ω_{nk} построим гистограмму значений пикселей объекта. Последовательность значений пикселей, находящихся на отрезке $[Pmin_{nk}, Pmax_{nk}]$, обозначим через P_{ntk} ($Pmin_{nk} \leqslant P_{ntk} \leqslant Pmax_{nk}$), ($1 \leqslant t \leqslant (Pmax_{nk} - Pmin_{nk} + 1)$. При построении гистограммы множества Ω_{nk} получим количество реализаций NP_{ntk} спектра P_{nijk} для выбранного объекта. Вероятность реализации значений спектра P_{ntk} можно определить как $V(P_{ntk}) = NP_{ntk}/(Ni \cdot Nj)$. ($1 \leqslant t \leqslant (Pmax_{nk} - Pmin_{nk} + 1)$; $k \leqslant Nk$). Для выбранного объекта рассчитываем нечеткую переменную, характеризующую значение спектра P_{ntk} для k-го канала. Совокупность P_{ntk} для ($1 \leqslant t \leqslant (Pmax_{nk} - Pmin_{nk} + 1)$; $k \leqslant Nk$) можно записать в БД в виде нечеткого множества $(P_{ntk}, V(P_{ntk})$. Это нечеткое множество представляет собой набор значений, зависящих от спектральных характеристик объектов (рис. 2).

В табл. 1 представлен пример записи в БД спектральных характеристик объекта «Гречиха».

KDob — код объекта (6 — код гречихи). Col — код цвета, который будет назначен пикселю изображения, идентифицированного как гречиха. В строковом поле БД, который соответствует спектральному каналу №33 КА «Геотон-Л1», записаны 10 отрезков:

– через _ записаны начало и конец отрезка;

– после $ до ! записана вероятность попадания в данный отрезок (значение, умноженное на 100, чтобы оперировать с целыми числами);

– между последним ! и # находится величина математического ожидания случайной величины спектра в канале и далее находится величина среднеквадратического отклонения случайной величины спектра в канале.

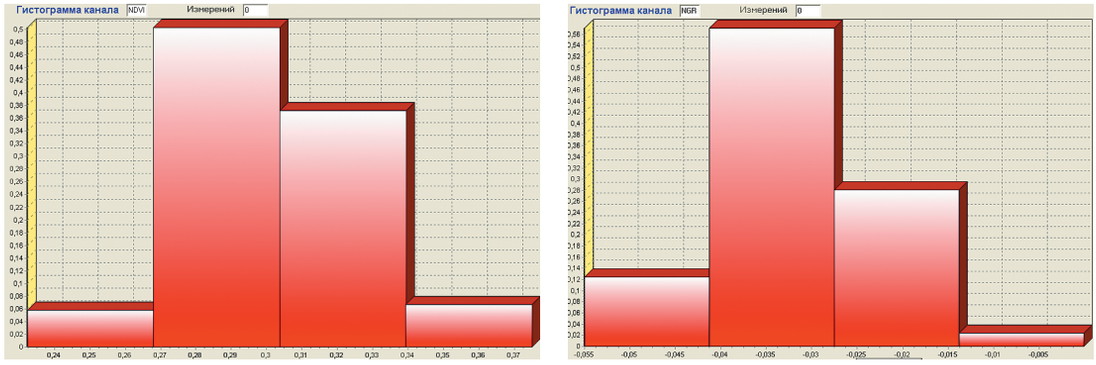
Этап 1.2

Формируется нечеткое множество, представляющее собой набор значений, зависящих от функций спектральных характеристик объектов.

Обозначим набором спектральных признаков объекта совокупность функций, аргументами которых являются P_{nijk}. Число спектральных признаков обозначим через Npr. Значение спектрального признака с номером m объекта с номером n (PR_m) в точке фрагмента изображения с номерами строки и столбца (i, j) имеет вид: $PR_{nijm} = f_m(P_{nij1}, P_{nij2}, P_{nijNk})$; ($i \leqslant Ni$; $j \leqslant Nj$; $m \leqslant Nm$). Получилось множество реализаций спектральных признаков объекта. Вероятность реализации значений спектра PR_{ntk} можно определить как $V(PR_{ntk}) = N\,PR_{ntk}/(Ni \cdot Nj)$; ($1 \leqslant t \leqslant (PRmax_{nk} - PRmin_{nk} + 1)$; $k \leqslant Nk$). Для выбранного объекта нечеткую переменную, характеризующую значение реализаций спектральных признаков объекта PR_{ntk} для «k» канала, можно записать в виде $(PR_{ntk}, V(PR_{ntk})$. Множество реализаций значений спектрального признака с номером m объекта с номером n обозначим через Ωpr_{nm}.

Это нечеткое множество представляет собой набор значений, зависящих от функций спектральных характеристик объектов, и записывается в БД. Результаты формирования нечетких множеств значений, зависящих от функций спектральных характеристик (NDVI и NGRDI) объекта «Гречиха», приведены на рис. 3 в виде гистограмм. В табл. 2 представлен пример записи значений, зависящих от функций спектральных характеристик объекта «Гречиха». Структура записи аналогична табл. 1.

После завершения этапа 1 проводится анализ информации о дешифровочных характеристиках объектов и делается заключение о принадлежности рассматриваемых объектов к одному из имеющихся



NDVI NGRDI

Рис. 3. Гистограммы распределения множества значений индексов NDVI и NGRDI

Таблица 2. Запись в БД значений, зависящих от функций спектральных характеристик объекта «Гречиха»

NDVI	NGRDI
0,174_0,219$4!0,219_0,264$36!0,264_0,309$50!0,309_0,354$10!0,27#0^	-0,083_-0,063$6!-0,063_-0,044$60!-0,044_-0,024$33!-0,024_-0,004$1!-0,047#0^
0,216_0,247$1!0,247_0,278$39!0,278_0,309$44!0,309_0,34$17!0,286#0^	-0,075_-0,061$1!-0,061_-0,047$28!-0,047_-0,033$56!-0,033_-0,019$16!-0,042#0^
0,22_0,273$1!0,273_0,326$2!0,326_0,379$35!0,379_0,432$63!0,382#0^	-0,071_-0,044$1!-0,044_-0,016$2!-0,016_0,012$43!0,012_0,04$55!0,012#0^

в БД объектов. Этап 2 состоит из этапов 2.1 и 2.2. На этапе 2.1 проводится анализ нечетких множеств значений спектральных характеристик объектов, а на этапе 2.2 — нечетких множеств, являющихся функциями спектральных характеристик.

Этап 2.1

Пусть pix_{ijk} — значение пикселя в точке (i, j) изображения в k-м спектральном канале. В каждой строке БД содержатся нечеткие значения спектральных характеристик разделенных на «n» отрезков для объекта с индексом m. Для каждого k-го канала определяется попадание pix_{ijk} в один из отрезков строки БД, т. е. определяется $Vm(pix_{ijk})$ — вероятность реализации pix_{ijk} в k-м канале для объекта m. Признаком будет являться выполнение выражения $(Vm(pix_{ijk}) > V1km)$, где pix_{ijk} — значение пикселя в точке (i, j) изображения в k-м спектральном канале, $Vm(pix_{ijk})$ — вероятность реализации вектора pix_{ijk} в k-м спектральном канале, $V1km$ — некоторая константа, зависящая от номера канала и свойств объекта с индексом m в канале k.

Если $(Vm(pix_{ijk}) > V1km)$ истинно, то для объекта с индексом m признак выполнения условия засчитывается, в противном случае — не засчитывается.

Этап 2.2

Алгоритм расчета аналогичен алгоритму этапа 2.1. Основное отличие заключается в том, что вместо pix_{ijk} и $Vm(pix_{ijk})$ для каждого канала используется $f(pix_{ij1}, \ldots, pix_{ijr})$ — функция некоторых каналов, например NDVI. Для данного вида съемочной аппаратуры определяются правила (критерии, весовые коэффициенты каждого признака) голосования по набору признаков с учетом нечетких значений признаков. Признак засчитывается в случае выполнения условия $(Vm(f(pix_{ij1}, \ldots, pix_{ijr}) > V2m)$, где $V2m$ — некоторая константа, зависящая от функции спектральных характеристик и свойств объекта с индексом m.

Для каждого пикселя изображения проводится процесс голосования путем подсчета количества засчитанных признаков. Пиксель считается принадлежащим искомому объекту, если количество признаков превышает минимальное заданное значение.

Результат проведения голосования представляет собой набор следующих величин:

– нечеткое значение функции голосования и величина вероятности события;

– код объекта и код цвета его легенды на карте.

При величине вероятности максимального значения функции голосования меньше определенного значения объект считается не определенным.

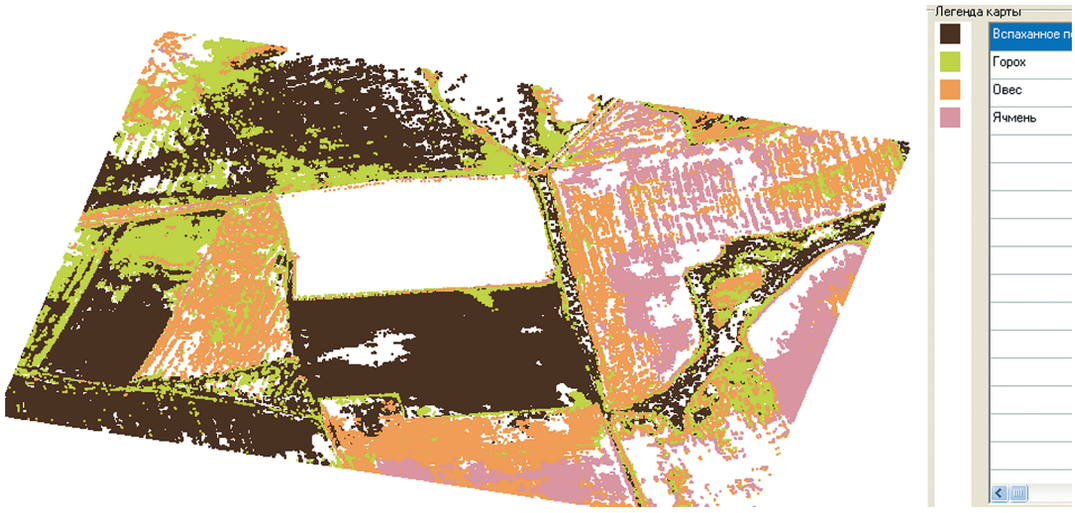


Рис. 4. Результаты тематической обработки информации о растительности участка территории Болховского района Орловской области

Результат обработки представляется в виде набора пикселей, которым присвоены условные цвета объектов.

На рис. 4 представлены результаты тематической обработки фрагмента космического снимка полей Орловской области (КА «Ресурс-П» №2, прибор «Геотон-Л1»). На рисунке различными цветами отображены участки с произрастанием различной сельскохозяйственной растительности (горох, овес, ячмень), вспаханные поля. Белым цветом изображены участки с неидентифицированной растительностью (из-за недостаточности наземных наблюдений для формирования эталонных признаков дешифрации).

Заключение

1. Идентификация объектов по данным КА высокого разрешения вызывает необходимость разработки методов обработки информации об объектах со спектральными характеристиками, отличающимися от нормального закона.

Предложена методика тематической обработки данных космического наблюдения, включающая представление дешифровочных признаков нечеткими множествами с произвольным законом распределения спектральных характеристик и способ обработки таких данных — «голосование по набору признаков».

Методика позволяет повысить достоверность результатов тематического дешифрирования информации КА детального и высокого разрешения.

2. Методика реализована и апробирована в виде программного комплекса с использованием данных российской группировки КА дистанционного зондирования «Ресурс-П» и «Канопус-В». Методика может быть использована для проведения космического мониторинга сельскохозяйственной растительности, решения экологических задач, повышения достоверности обнаружения растительности, подлежащей контролю правоохранительных органов.

Список литературы

1. *Заде Л.* Понятие лингвистической переменной и его применение к принятию приближенных решений. М.: Мир, 1976. 167 с.
2. *Круглов В.В., Дли М.И.* Интеллектуальные информационные системы: компьютерная поддержка систем нечеткой логики и нечеткого вывода. М.: ФИЗМАТЛИТ, 2002. 256 с.
3. Прикладные нечеткие системы / Под ред. Т. Тэрано, К. Асаи, М. Сугэно. М.: Мир, 1993. 368 с.
4. *Стрыков А.И.* Особенности тематической обработки гиперспектральной информации с КА «Ресурс-П» в задачах мониторинга и распознавания объектов природной среды // Ракетно-космическое приборостроение и информационные системы, 2017, т. 4, вып. 2. С. 11–19.