

МНОГОМЕРНАЯ СТАТИСТИЧЕСКАЯ МОДЕЛЬ ДЛЯ ОБНАРУЖЕНИЯ МЕСТ НЕФТЕЗАГРЯЗНЕНИЙ ПО МАТЕРИАЛАМ КОСМИЧЕСКИХ СЪЕМОК

Аловсат Шура оглы Гулиев

Государственная Нефтяная Компания Азербайджанской Республики (SOCAR), AZ1000, Азербайджан, г. Баку, пр. Гейдара Алиева 121, старший геодезист, тел. (99450)492-93-18, e-mail: alov_soc@yahoo.com

Татьяна Александровна Хлебникова

Сибирский государственный университет геосистем и технологий, 630108, Россия, г. Новосибирск, ул. Плеханова, 10, доктор технических наук, профессор кафедры инженерной геодезии и маркшейдерского дела, тел. (913)474-19-70, e-mail: t.a.hlebnikova@ssga.ru

В статье рассматривается алгоритм определения статистической модели по нескольким неоднородным изображениям земной поверхности, полученными разными датчиками (оптико-электронное сканирующее устройство, радиолокатор с синтезированной апертурой (РСА)) над морскими районами. Объектом исследования являются методы дистанционного зондирования Земли, используемые для обнаружения и картирования разливов нефти. Целью исследований ставилось выполнение тестирования для возможной вариации статистической модели внутри неоднородного скользящего окна основанной на полуавтоматическом подходе. Предлагаемый алгоритм позволяет определять пространственную протяженность мест нефтеразработок и нефтяных загрязнений в шельфовых акваториях с использованием многовременных данных РСА и многозонального совмещенного изображения с пространственным разрешением 10 м. Сначала на изображении анализируются однородные области, а затем модель зоны анализа расширяется до более общего случая неоднородных областей, которые наблюдаются в окнах анализа.

Ключевые слова: Sentinel-1A, Sentinel-2A, радиолокационные снимки, многозональные космические снимки, мониторинг нефтяных загрязнений

MULTIDIMENSIONAL STATISTICAL MODEL FOR DETECTING OIL POLLUTION SITES BASED ON SATELLITE IMAGERY

Alovsat Shura oglu Guliyev

State Oil Company of Azerbaijan Republic (SOCAR), 121 Heydar Aliyev avenue, Baku, Azerbaijan, AZ1000, Senior Surveyor, phone: (99450)492-93-18, e-mail: alov_soc@yahoo.com

Tatiana A. Khlebnikova

Siberian State University of Geosystems and Technologies, 10 Plakhotnogo St., Novosibirsk, 630108, Russia, D.Sc., Professor, Department of Engineering Geodesy and Mine Surveying, phone: (913)474-19-70, e-mail: t.a.hlebnikova@ssga.ru

The article considers an algorithm for determining the statistical model from several inhomogeneous images of the Earth's surface obtained by different sensors (optoelectronic scanning device, synthetic aperture radar (SAR)) over the sea areas. The object of the study are the methods of remote sensing of the Earth used for detection and mapping of oil spills. The aim of the research was to perform testing for a possible variation of the statistical model inside a non-uniform sliding window based on a semi-automatic approach. The proposed algorithm makes it possible to determine the spa-

tial extent of oil production sites and oil pollution in offshore waters using multi-time RSA data and a multi-zone combined image with a spatial resolution of 10 m. First, homogeneous regions are analyzed in the image, and then the model of the analysis zone is expanded to the more general case of inhomogeneous regions that are observed in the analysis windows.

Keywords: Sentinel-1A, Sentinel-2A, radar imagery, multi-zone satellite imagery, oil pollution monitoring

Введение

В тех случаях, когда процесс получения полевых материалов невозможен (большое расстояние до исследуемого объекта, прошедшее время события), для обнаружения и картирования разливов нефти в шельфовых акваториях используются материалы дистанционного зондирования Земли – космические снимки. Для дешифрирования объектов на цифровом изображении выполняется компьютерное дешифрирование, основанное на использовании универсального подхода – классификации.

Под классификацией понимается процедура определения принадлежности объекта к одному из выбранных классов. Результатом классификации являются объекты, разделенные по классам. Метод контролируемой классификации основан на использовании и учете априорной информации о типах объектов и эталонных значениях спектральных характеристик этих объектов. В процессе классификации выполняется сравнение значения яркости текущего пикселя с эталонными признаками. При отсутствии априорной информации используется классификация без обучения (неконтролируемая классификация).

В данной работе задачей анализа космических изображений является определение границ областей (нефтяных пятен) на изображениях с использованием неконтролируемых методов классификации. В большинстве программных приложений выполнение классификации требует эмпирической интерпретации, визуального распознавания объектов, что занимает много времени.

Предлагается стратегия обнаружения изменений на изображениях, полученных средствами однородных и неоднородных датчиков (оптико-электронные и RSA-изображения). Для уменьшения влияния шума и учета пространственной корреляции, присутствующих в большинстве изображений дистанционного зондирования, при анализе изображений следует учитывать обобщенный подход при анализе группы пикселей, содержащиеся в скользящем окне.

Распределение этих групп пикселей имеет место во многих приложениях обработки изображений. Приложения обработки изображений включают: обнаружение изменений, сегментацию изображений, регистрацию изображений и обновление базы данных [1–5].

Распределения, рассматриваемые в литературе, включают двумерные распределения гаммы для двух изображений радиолокатора с синтезированной апертурой RSA [1] и двумерные распределения Пирсона для неоднородных оптических и RSA изображений [2]. Однако вопросам обнаружения изменений между изображениями, полученными с помощью гетерогенных датчиков, уделя-

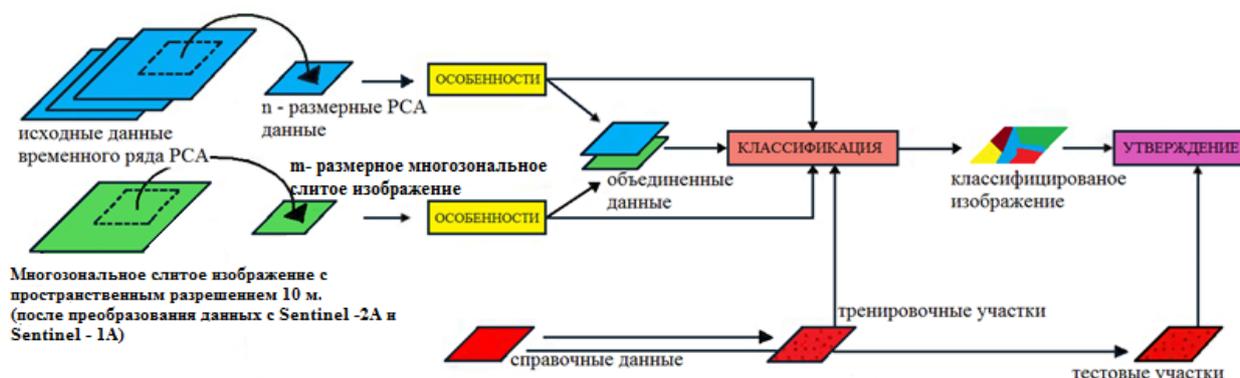
лось меньше внимания в литературе, чем оптическим или радиолокационным / радарным. Известен недавно разработанный подход, при котором одно из двух изображений преобразуется с целью получения характеристик, аналогичных другому изображению, с использованием теории связей. Однако этот метод требует изучения соответствующей связки с использованием обучающих выборок, и его вряд ли можно обобщить на ситуации дешифрирования, когда доступно более двух изображений.

Методы и материалы

Для исследований многозональные совмещенные изображения и спутниковые данные радиолокатора с синтезированной апертурой были взяты на район акватории Нефтяных Камней (Азербайджанский экономический сектор Каспийского моря). Использовались одна сцена Sentinel-2A и временной ряд данных Sentinel-1A, охватывающий период с 2017 по 2018гг. Европейского Космического Агентства (ЕСА). Протяженность исследуемой территории с запада на восток – 52,08 км, с севера на юг – 45,20 км, площадь 2354 кв. км. Географические координаты 40°26' N, 50°30' E – 40°01' N, 51°07' E.

Обработка выбранных снимков выполнялась средствами системы анализа данных дистанционного зондирования Земли ERDAS Imagine версии 6.7 и SNAP Desktop версии 8.0.0. Некоторые статистические показатели использовались в качестве функций для данных временных рядов Sentinel-1A. Для анализа данных временных рядов использовался набор инструментов пакета прикладных программ Matlab. Рабочий процесс интегрированного мониторинга состоит из нескольких этапов (рисунок). В начале на изображениях была определена область достоверно определенных объектов.

Затем в справочных данных были определены тренировочные и тестовые участки. Далее были выбраны объекты из наборов данных, используемых для классификации. С помощью данных обучения эти признаки были использованы для классификации признаков нефтепроявлений. Вторая классификация была проведена путем объединения двух наборов данных. Наконец, все классификации были проверены с помощью тестовых участков.



Рабочий процесс экспериментов, проведенных для этого исследования

Меры сходства, основанные на статистической зависимости между значениями интенсивностями пикселей, не всегда подходят для обнаружения изменений между изображениями, полученными с помощью разнородных датчиков. Чтобы восполнить этот пробел, в исследовании предлагается новая мера подобия, основанная на статистической модели. Мы предполагаем, что некоторые свойства параметров статистической модели сохраняются только в неизменных областях. Это предложение приводит к определению новой меры подобия, основанной на параметрах этой статистической модели.

Поскольку объект является одним и тем же для всех изображений, связанных с неизменной областью, физические свойства P (свойства исследуемого объекта) для данного пикселя не изменяются для всех изображений. Такой объект создаст компоненту в распределении статистических данных, следуя результату. Следовательно, вектор параметров, описывающий эту компоненту, принадлежит многообразию M , характеризующего отношения между задействованными датчиками, определенному в рамках гипотезы «без изменений».

Результаты и обсуждения

В данных исследованиях использовались два алгоритма классификации: правило ближайшего соседа (k -NN) и классификатор максимального правдоподобия. Использовались три различных варианта правила ближайшего соседа. Все наборы данных были классифицированы этими четырьмя методами.

Алгоритм k -ближайшего соседа (k -NN) – это метод непараметрической классификации, который классифицирует объекты на основе ближайших обучающих примеров в пространстве признаков. k -NN – это тип обучения на основе экземпляров, при котором функция аппроксимируется только локально, а все вычисления откладываются до классификации. Алгоритм k -ближайшего соседа является одним из самых простых из всех алгоритмов машинного обучения: объект классифицируется большинством голосов его соседей, причем объект присваивается классу, наиболее распространенному среди его k -ближайших соседей (k – положительное целое число).

В исследовании использовались три версии правила ближайшего соседа. Первые два были k NN с $k = 1$ и $k = 9$. Это означает, что первый использовал для классификации одного соседа (1-NN), а второй – девять соседей (9-NN). В третьем методе использовались средние значения областей обучения (далее MDM). В этом методе вычислялось среднее значение всех обучающих пикселей в классе, и эти значения сравнивались с пикселями с неизвестной классификацией для присвоения класса. Средние значения рассчитывались отдельно для каждого изображения. Наиболее близкое среднее значение по сравнению с пикселем с неизвестным классом определяет присвоенный класс. Этот классификатор также называется минимальным расстоянием до средних значений.

Визуальная интерпретация дает хорошее представление о результате классификации, но следует также проверить окончательную точность. В таблице

приведены результаты всех выполненных экспериментов. Вариант 9-NN в основном работает так же хорошо, как лучший метод для определенного набора данных, тогда как максимальная вероятность варьируется от лучшего до худшего. Можно отметить, что уменьшение размерности в основном приводит к худшей точности (за исключением PCA). Однако даже в этих случаях по некоторым классификациям были достигнуты хорошие результаты.

Результаты классификации данных Sentinel-1

Параметры классификации	MDM	1-NN	9-NN	ML
	%			
Поляризация вместе(vv+vh)	68,4	71,2	67,2	44,4
3 изображения PCA, рассчитанные на основе всех поляризованных изображений (vv и vh)	73,2	71,2	63,4	64,4
Выбор сцены (1 дата, одна поляризация):				
– использование коэффициента Фишера; (17.08.2018, vv поляризовано)	5,6	33,2	40,1	42,9
– использование ступенчатой регрессии; (23.09.2018, vv поляризовано)	54,8	53,2	54,1	51,6
Выбор сцены (2 даты, обе поляризации):				
– использование коэффициента Фишера (среднее, минимальное, среднее квадратическое, энтропия)	40,0	55,2	57,6	62,8
– использование пошаговой регрессии (std, max, min, эксцесс)	47,6	59,2	58,8	61,6

В таблице приведены значения достоверности автоматизированной классификации участков нефтезагрязнений на исследуемой территории на основе априорных данных о дешифровочных признаках объекта.

Выводы

Выполненные исследования позволяют сделать следующие выводы.

Полученные результаты позволили прийти к выводу, что за счет совместной обработки мезозональных и временных рядов радиолокационных изображений одной и той же территории представляется возможным обработать множество идентифицируемых объектов водной поверхности и распознать зоны нефтеразливов на ранних стадиях.

Предложенная статистическая модель анализа значений яркостей пикселей на нескольких изображениях, полученных различными датчиками, позволяет определить изменения границ нефтяных пятен на водной поверхности и распознать зоны нефтеразливов на ранней стадии.

Выполнен анализ нескольких алгоритмов классификации на основе предложенной статистической модели. Выявлено, что уменьшение размерности окон анализа приводит к понижению точности. Подход к получению этой статистиче-

ской модели отличается от известных, поскольку он исследует физические процессы, которые генерируют неоднородные изображения.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. M. Chabert and J.-Y. Tournet, Bivariate Pearson distributions for remote sensing images, in Proc. IEEE Int. Geosci. Remote Sens. Symp., Vancouver, Canada, July 2011, pp. 4038–4041 [Electronic resource]. – Mode of access: <https://ieeexplore.ieee.org/document/6050118> (дата обращения 21.01.2021).

2. G. Mercier, G. Moser, and S. B. Serpico, Conditional copulas for change detection in heterogeneous remote sensing images, IEEE Trans. Geosci. Remote Sens., vol. 46, no. 5, pp. 1428–1441, May 2008 [Electronic resource]. – Mode of access: <https://ieeexplore.ieee.org/abstract/document/4481231> (дата обращения 21.01.2021).

3. M. S. Allili, D. Ziou, N. Bouguila, and S. Boutemedjet, Image and video segmentation by combining unsupervised generalized Gaussian mixture modeling and feature selection, IEEE Trans. Circuits Syst. Video Technol., vol. 20, no. 10, pp. 1373–1377, Oct. 2010 [Electronic resource]. – Mode of access: <https://ieeexplore.ieee.org/document/5580019> (дата обращения 21.01.2021).

4. F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, Multimodality image registration by maximization of mutual information, IEEE Trans. Med. Imag., vol. 16, no. 2, pp. 187–198, 1997 [Electronic resource]. – Mode of access: <https://ieeexplore.ieee.org/document/563664> (дата обращения 21.02.2021).

5. M. Chabert, J.-Y. Tournet, V. Poulain, and J. Inglada, Logistic regression for detecting changes between databases and remote sensing images, in Proc. IEEE Int. Geosci. Remote Sens. Symp., Honolulu, USA, 2010, pp. 3198–3201 [Electronic resource]. – Mode of access: <https://ieeexplore.ieee.org/document/5649669> (дата обращения 21.02.2021).

6. C. Tison, J.-M. Nicolas, F. Tupin, and H. Maitre, A new statistical model for Markovian classification of urban areas in high-resolution SAR images”, IEEE Trans. Geosci. Remote Sens., vol. 42, no. 10, pp. 2046–2057, Oct. 2004 [Electronic resource]. – Mode of access: <https://ieeexplore.ieee.org/document/1344157> (дата обращения 21.01.2021).

7. L. Bruzzone and D. F. Prieto, Automatic analysis of the difference image for unsupervised change detection, IEEE Trans. Geosci. Remote Sens., vol. 38, no. 3, pp. 1171–1182, May 2000 [Electronic resource]. – Mode of access: <https://ieeexplore.ieee.org/document/843009> (дата обращения 21.02.2021).

8. Аловсат Ш. Г., Хлебникова Т. А. Выявление мест нефтезагрязнений шельфовой зоны по материалам космических съемок (на примере акватории Нефтяных Камней (Каспий)) // Вестник СГУГиТ. – 2019. – Т. 24, № 3. – С. 52–64.

9. L. Wojnar, Image Analysis: Applications in Materials Engineering, Taylor & Francis, 1998. ISBN-13: 978-0849382260, ISBN-10: 0849382262, Mode of access: <https://www.amazon.com/Image-Analysis-Applications-Engineering-Technology/dp/0849382262> (дата обращения 21.01.2021).

© Аловсат Шура оглы Гулиев, Т. А. Хлебникова, 2021