

М. А. Сквазников¹, Д. Л. Колыгин¹*

Подход к разработке математической модели процесса интерпретации данных дистанционного зондирования Земли с использованием статистических методов регрессионного анализа

¹ Военно-космическая академия имени А. Ф. Можайского,
г. Санкт-Петербург, Российская Федерация
* e-mail: vka@mil.ru

Аннотация. Широкое использование средств дистанционного зондирования Земли в интересах информационного обеспечения решения социально-экономических задач диктует необходимость оценивания качества процесса интерпретации аэрокосмической информации. Существующие в настоящее время подходы к определению достоверности результатов анализа визуальной информации применимы только для простых объектов искусственного происхождения. Поэтому целью проведенного исследования явилась разработка математической модели процесса интерпретации разнородных данных дистанционного зондирования Земли. Для определения параметров математической модели процесса интерпретации предложен статистический метод регрессионного анализа с использованием теории планирования экспериментов. Введено новое понятие - «форм-фактор» объекта, равное соотношению значений коэффициента распознавания его формы и линейного размера. Для вычисления коэффициентов уравнения регрессии был проведен дробный факторный эксперимент первого порядка с двухуровневым варьированием факторов. Проверка значимости полученной модели осуществлялась с использованием критерия Фишера. Разработанная модель позволяет оценивать качество интерпретации объектов в зависимости от сложности их формы и качества полученных данных ДЗЗ, характеризуемого линейным разрешением на местности.

Ключевые слова: интерпретация аэрокосмической информации, данные дистанционного зондирования Земли, форм-фактор объекта, вероятность распознавания объекта, объект искусственного происхождения

М. А. Skvaznikov¹, D. L. Kolygin¹*

An approach to the development of a mathematical model for the interpretation of earth remote sensing data using statistical methods of regression analysis

¹ Mozhaisky Military Space Academy, Saint Petersburg, Russian Federation
* e-mail: vka@mil.ru

Abstract. The widespread use of Earth remote sensing tools in the interests of information support for solving socio-economic problems dictates the need to assess the quality of the process of interpreting aerospace information. The current approaches to determining the reliability of the results of the analysis of visual information are applicable only to simple objects of artificial origin. Therefore, the purpose of the study was to develop a mathematical model for the process of interpreting heterogeneous Earth remote sensing data. To determine the parameters of the mathematical model of the interpretation process, a statistical method of regression analysis is proposed using the theory of planning experiments. A new concept is introduced - the "form factor" of an object, which is equal to the ratio of the values of the coefficient of recognition of its shape and linear size. To calculate the coef-

ficients of the regression equation, a fractional factorial experiment of the first order was carried out with a two-level variation of the factors. The significance of the obtained model was checked using the Fisher criterion. The developed model makes it possible to evaluate the quality of interpretation of objects depending on the complexity of their shape and the quality of the obtained remote sensing data, characterized by a linear resolution on the ground.

Keywords: interpretation of aerospace information, earth remote sensing data, object form factor, object recognition probability, object of artificial origin

Введение

Процесс интерпретации данных дистанционного зондирования Земли (ДЗЗ) занимает более четверти всего объема работ на этапе планирования различных социально-экономических задач. Визуальный анализ аэрокосмических снимков заключается в обнаружении, распознавании и определении характеристик объектов по их изображению. При этом распознавание является основным содержательным этапом данного процесса, обуславливающим его основную цель [1-4].

Вследствие воздействия априорно неизвестных факторов, характеризующих как условия функционирования, так и применения системы ДЗЗ, интерпретация является процессом стохастическим. Поэтому для оценивания ее эффективности необходимо использовать вероятностные показатели качества [5-8].

Конечным результатом процесса интерпретации данных ДЗЗ является информация, необходимая потребителю. Следовательно, результативность, как одно из основных операционных свойств данного процесса, характеризуется качеством получаемой информации. Показателем её выбрана достоверность информации, которая имеет смысл вероятности правильного распознавания объектов наблюдения [9-14].

Информация об основных свойствах (признаках) изображенных на снимках объектов сравнивается с априорной информацией о классах (эталонах, образах). При этом анализируются совпадения наблюдаемых свойств реализаций с аналогичными свойствами установленных классов. По результатам сравнения принимается решение о принадлежности изучаемого изображения объекта к определенному классу [15-19]. Решение принимается с учетом заданных критериев распознавания. Полученные сведения о выбранном классе являются результатом распознавания наблюдаемых объектов.

Материалы и методы

Существующий в настоящее время подход [20] к оценке вероятности распознавания применим только для простых объектов искусственного происхождения и предполагает использование следующего математического выражения:

$$P = \exp \left[- \left(B \frac{L}{A} \right)^2 \right], \quad (1)$$

где P – вероятность распознавания простого объекта,

B – коэффициент распознавания формы простого объекта;

L – линейное разрешение на местности,

A – максимальный геометрический размер простого объекта.

При этом аргументы B и A характеризуют свойства дешифрируемого объекта, а аргумент L – качество изображения (данных ДЗЗ). Введем новое понятие – «форм-фактор» простого объекта искусственного происхождения F , равное соотношению значений коэффициента распознавания его формы и линейного размера:

$$F = \frac{B}{A}. \quad (2)$$

Простые объекты представляют собой элементарные геометрические фигуры: круг, квадрат, прямоугольник, треугольник и т.д. Реальные (сложные) объекты можно декомпозировать на совокупность простых. Например, на плановом изображении коттеджной застройки (рис. 1) можно выделить прямоугольные жилые и вспомогательные постройки, круглую детскую площадку; на изображении водоочистных сооружений – круглые отстойники, прямоугольные аэротенки и вспомогательные здания; на изображении буровой платформы – треугольные подъемные опоры и главную палубу, а также прямоугольную вертолетную площадку.

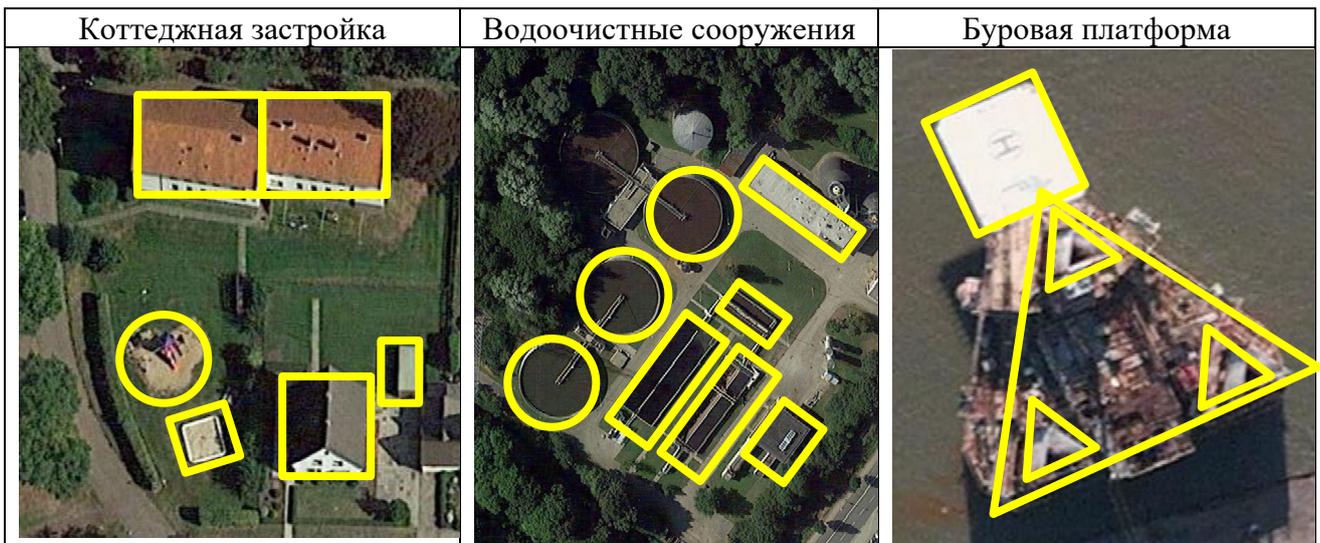


Рис. 1 Декомпозиция изображений сложных объектов

Для расчета форм-фактора сложного объекта F_c предлагается использовать среднее значение форм-факторов его простых элементов F_i , которые выступают в качестве признаков при распознавании:

$$F_c = \frac{1}{n} \sum_{i=1}^n F_i, \quad (3)$$

где n – количество простых элементов, используемых при распознавании объекта.

В табл. 1 приведены результаты расчета значений форм-фактора некоторых простых и сложных искусственных объектов.

Таблица 1

Значения форм-факторов простых и сложных объектов

Типы объектов	Резервуар ГСМ	Контейнер	Участок трубопровода	Конвейерная галерея	Коттеджная застройка	Водоочистные сооружения	Буровая платформа
F_c	0,0315	0,0416	0,0470	0,0018	0,1291	0,0467	0,1103

Экспериментальные исследования показывают, что математическое выражение (1) применимо только для расчета вероятности распознавания простых объектов.

Результаты и обсуждение результатов

Математическая модель распознавания сложных объектов искусственного происхождения разработана на основе использования теории планирования эксперимента [21-25].

При планировании эксперимента были выявлены следующие значимые факторы, оказывающие наиболее существенное влияние на величину вероятности распознавания объекта P : форм-фактор сложного объекта F_c и линейное разрешение на местности L .

В ходе опыта каждый результат распознавания сравнивается с имеющейся достоверной информацией об объекте. Оценка вероятности распознавания k -го объекта производится с использованием выражения:

$$\tilde{P}_k = \frac{m_k}{n_k}, \quad (4)$$

где m_k – количество правильно распознанных k -х объектов;

n_k – общее количество предъявленных для распознавания k -х объектов.

При линейной зависимости вероятности распознавания сложных объектов P от двух независимых факторов уравнение регрессии принимает следующий вид:

$$P = b_0 + b_1 F_c + b_2 L, \quad (5)$$

где b_0, b_1, b_2 – коэффициенты регрессии;

F_c – форм-фактор сложного объекта.

Полный факторный эксперимент при варьировании n факторов на двух уровнях обуславливает постановку 2^n опытов.

Коэффициенты в уравнении регрессии рассчитаны по формулам:

$$b_0 = \frac{\sum_{i=1}^n \bar{P}_j x_{0j}}{n}; \quad (6)$$

$$b_1 = \frac{\sum_{i=1}^n \bar{P}_j F_{cj}}{n}; \quad (7)$$

$$b_2 = \frac{\sum_{i=1}^n \bar{P}_j L_j}{n}, \quad (8)$$

где \bar{P}_j – среднее значение оценок вероятностей распознавания сложных объектов, полученных в результате эксперимента, состоящего из 2^n опытов;

$$j = \overline{1, n}.$$

Для проверки значимости коэффициентов регрессии использовался критерий Стьюдента [26].

В результате расчетов получено следующее уравнение регрессии:

$$\hat{P} = 0,7775 - 0,0875F_c - 0,0475L, \quad (9)$$

где \hat{P} – модельное значение вероятности распознавания объектов.

Адекватность разработанной модели при доверительной вероятности $\beta = 0,95$ доказана на основе использования критерия Фишера [27].

Наибольший интерес представляет получение математического выражения вероятности распознавания определенного класса объектов. Если в качестве основных признаков класса объектов рассматривать их размеры, форму, а также конструкционные особенности, то среднее значение форм-факторов сложных объектов i -го класса $\overline{F_{ci}}$ будет представлять его кардинальную отличительную особенность.

Так, например, для класса объектов «водоочистные сооружения» $\overline{F_{c_{bc}}} = 0,0277$.

При этом математическое выражение вероятности распознавания водоочистных сооружений приобретает следующий вид:

$$\hat{P}_{bc} = 0,7751 - 0,0475L. \quad (10)$$

Задавая значения директивной вероятности распознавания объектов наблюдения, можно предъявлять обоснованные требования к характеристикам бортовой специальной аппаратуры (линейному разрешению на местности) средств ДЗЗ.

Заключение

Разработанная модель позволяет оценивать качество интерпретации объектов искусственного происхождения в зависимости от сложности их формы (количества элементов, определяющихся геометрической формой и значением форм-фактора) и качества полученных данных ДЗЗ, характеризуемого линейным разрешением на местности.

Предлагается использовать среднее значение форм-факторов сложных объектов определенного класса, как его кардинальное отличительное свойство при автоматизации процесса интерпретации данных ДЗЗ.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Вершовский Е.А. Интерпретация данных дистанционного зондирования Земли // Перспективы развития информационных технологий. – 2010. – № 2. – С. 43-45.
2. Ерошевич М.О., Ермакович В.Р., Саечников В.А. Обработка и интерпретация данных дистанционного зондирования, полученных с БКА // Сборник тезисов докладов шестнадцатой Всероссийской открытой конференции "Современные проблемы дистанционного зондирования Земли из космоса". Институт космических исследований Российской академии наук. – 2018. – С. 36.
3. Колесников А.А. Анализ методов и средств искусственного интеллекта для анализа и интерпретации данных активного дистанционного зондирования // Вестник СГУГиТ (Сибирского государственного университета геосистем и технологий). – 2022. – Т. 27, № 3. – С. 74-94.
4. Чехарин Е.Е. Алгоритмы интерпретации данных дистанционного зондирования // Славянский форум. – 2015. – № 3 (9). – С. 301-308.
5. Попова О.А. Повышение точности представления и обработки данных дистанционного зондирования Земли // Региональные проблемы дистанционного зондирования Земли. Материалы III Международной научной конференции. – 2016. – С. 174-177.
6. Андриянов Н.А., Дементьев В.Е., Ташлинский А.Г. Обнаружение объектов на изображении: от критериев Байеса и Неймана-Пирсона к детекторам на базе нейронных сетей // Компьютерная оптика. – 2022. – Т. 46, № 1. – С. 139-159.
7. Евстратова Е. К вопросу точности обработки материалов дистанционного зондирования Земли при мониторинге территорий // Сборник тезисов докладов шестнадцатой Всероссийской открытой конференции "Современные проблемы дистанционного зондирования Земли из космоса". Институт космических исследований Российской академии наук. – 2018. – С. 31.
8. Гарафутдинова Л.В., Каличкин В.К., Хлебникова Е.П. Оценка методов классификации многозональных космических снимков // Вестник Омского государственного аграрного университета. – 2022. – № 4 (48). – С. 19-28. Sonka, M. Image processing, analysis and machine vision / M. Sonka, V. Hlavac, R. Boyle. — Boston : Cengage Learning, 2015. — 920 p.
9. USGS digital spectral library [Electronic resource] / U.S. Geological Survey. — Electronic data. — [USA], 2007. — Mode of access: <http://speclab.cr.usgs.gov/spectral-lib.html>.
10. Zucker, S. W. Finding structure in co-occurrence matrices for texture analysis / S. W. Zucker, D. Terzopoulos // Computer Graphics and Image Processing. — 1980. — Vol. 12, Issue 3. — P. 286-308.
11. Wang, W.N. Image Emotional Semantic Query Based on Color Semantic Description / W.N.Wang, Y.L.Yu // Proc. Of 2005 International Conference on Machine Learning and Cybernetics, Vol.7. — Guangzhou, 2005. — pp. 4571-4576.

12. Соловьева А.Н. Многоуровневое описание космических снимков в задаче автоматизированного дешифрирования / А.В. Кучуганов, А.Н. Соловьева // Интеллектуальные системы в производстве. – 2014. №2 (24). С. 164-166.
13. Blaschke, T. Object based image analysis for remote sensing / T.Blaschke// ISPRS Journal of Photogrammetry and Remote Sensing. – 2010. – Vol.65, Issue 1. – pp. 2-16.
14. Object-based image analysis: spatial concepts for knowledge-driven remote sensing applications / T.Blaschke, S.Lang, G.J.Hay (eds.). – Berlin: Springer, 2008. – 817 p. – (Lecture notes in geoinformation and cartography).
15. Huang, L. Object-oriented classification of high resolution satellite image for better accuracy / L.Huang, L.Ni // Proc. of the 8th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences. —Shanghai, 2008. —P.211–218.
16. Du, F. L., 2004. Object-oriented Image Classification Analysis and Evaluation, remote sensing technology and application, Vol.19, No.1, pp. 20-23.
17. Lee, J.Y., 2004. Image classification with a region based approach in high spatial resolution imagery, ISPRS, Istanbul, 2004. ACK.
18. GEOBIA state of the art, science or technology [Electronic resource] / P. Aplin. — Electronic data (1 file: 2431 Kbytes).— Thessaloniki, 2014. — Mode of access: <http://geobia2014.web.auth.gr/geobia14/sites/default/files/pictures/aplin.pdf>.
19. Sonka, M. Image processing, analysis and machine vision / M. Sonka, V. Hlavac, R. Boyle. — Boston : Cengage Learning, 2015. — 920 p.
20. Живичин А.Н., Соколов В.С. Дешифрирование фотографических выражений. – М.: Недра, 1980. – 253 с.
21. Сидняев Н.И. Теория планирования эксперимента и анализ статистических данных: учеб. пособие для магистров /– 2-ое изд., перераб. и доп. – М.: Издательство – 2015. – 496 с.
22. Спириин Н.А., Лавров В.В. Методы планирования и обработки результатов инженерного эксперимента: Конспект лекций (отдельные главы из учебника для вузов). – Под общ. Ред. Н.А.Спирина. Екатеринбург: ГОУ ВПО УГТУ-УПИ, 2004. – 257 с.
23. Белокуренько Н.С. Корреляционно-регрессионный анализ как статистический метод // Наука и общество в условиях глобализации. – 2020. – № 1 (7). – С. 54-55.
24. Астафьев Е.Р. Методы корреляционного и регрессионного анализа обработки статистических данных / Е.Р.Астафьев, Е.В.Михайленко. – Краснодар: Краснодарская акад. МВД России, 2006. – 67 с., ил.
25. Маслова Е.Ю., Гриневиц Я.А. Регрессионный анализ, как основной метод изучения статистической взаимосвязи // Галактика науки–2021. Материалы Всероссийской (национальной) научно-практической конференции. Кубанский государственный университет. – 2021. – С. 133-138.
26. Приходько А.А., Порошина Е.А., Порошина Л.А. Статистические методы в различных предметных областях // Наука и техника. Мировые исследования. Материалы международной научно-практической конференции. – 2020. – С. 184-187.
27. Статистические методы обработки результатов наблюдений / Под ред. Р.М. Юсупова. – Б.М.: МО СССР,1984. – 563 с.

© М. А. Сквзников, Д. Л. Колыгин, 2023