

УДК 528.8

О технологии создания новых технологий в области дистанционного зондирования Земли

М. А. Попов

ГУ “Научный центр аэрокосмических исследований Земли ИГН НАН Украины”, Киев, Украина

Изложен методический подход к поэтапному созданию новых технологий в области дистанционного зондирования Земли. Приводится схема поэтапного создания новых технологий в ДЗЗ, обсуждаются содержание и наиболее важные аспекты работ на каждом из этапов. Изложенный подход рекомендуется для использования в качестве методической основы при формировании технических заданий и планов выполнения соответствующих исследований и разработок в области ДЗЗ.

Ключевые слова: дистанционное зондирование Земли, аэрокосмическое изображение, требования к ДЗЗ-технологии, классификация изображений, сбор данных, объем выборки

© М. А. Попов. 2018

1. Вступление

Непрерывное совершенствование технических средств дистанционного зондирования Земли (ДЗЗ), развитие глобальной инфраструктуры доставки данных создали в настоящее время благоприятные условия для обеспечения пользователей практически в любой точке земного шара необходимыми им материалами аэрокосмической съемки. Более того, сегодня пользователю предоставляется большой выбор разных программных средств для обработки аэрокосмических изображений, причем работа с ними не требует высокой квалификации. Такая ситуация является благоприятной для разработки новых эффективных методов решения самых разных тематических задач и, безусловно, способствует дальнейшему расширению сфер использования методов ДЗЗ. Среди таких сфер — сельское и городское хозяйства, природные ресурсы, экология, транспорт и дорожные сети, водная среда и т. д., и в каждой из них учеными и специалистами постоянно ведутся новые разработки [1–3].

Однако проведенный нами анализ показывает, что создатели подобных разработок, с учетом их научного профиля и специфики предметной области, значительно различаются своими представлениями относительно уровня завершенности и технологичности разработок, глубины способов проверки их на эффективность, желаемой степени автоматизации и т. д. Данное обстоятельство затрудняет возможности как объективного сравнения ДЗЗ-разработок, так и их использование в смежных сферах с минимальными доработками.

Представляется, что указанных недостатков можно во-многом избежать, если при создании новых ДЗЗ-разработок, независимо от сферы их последующего применения, придерживаться некоего общего подхода. Такой технологический подход предлагается в данной статье. Приводится схема поэтапного создания новых технологий в ДЗЗ, обсуждаются содержание и наиболее важные аспекты работ на каждом из этапов.

2. О понятии “технология” в ДЗЗ

Кембриджский словарь делового английского языка [4] дает следующее определение понятия “технология”:

Technology = knowledge, equipment, and methods that are used in science and industry (знания, оборудование и методы, используемые в науке и промышленности).

Большой словарь современного украинского языка [5] трактует данное понятие таким образом:

Технология — совокупность способов обработки или переработки материалов, изготовления изделий, выполнения разных производственных операций и т. д.

Приведенные определения отражают сущность понятия “технология”, однако носят весьма общий характер. Заметим, что когда рассматривается какая-то конкретная область человеческой деятельности, то для достижения здесь новых конструктивных результатов необходимо учитывать присущий

* E-mail: mpopov@casre.kiev.ua
Тел.: +38 (044) 482-01-66,

ей методологический инструментарий, существующие в этой области специфические взгляды и представления, наработанные способы, методы, понятийный аппарат и т. д.

С учетом этих соображений для области ДЗЗ понятие “технология” можно выразить следующим образом:

Технология ДЗЗ — это набор процедур получения, обработки и комплексного анализа материалов аэрокосмической съемки, наземных данных и другой необходимой информации, в совокупности обеспечивающих решение конкретной тематической задачи или группы подобных задач и реализованных аппаратно и программно.

Под тематической задачей понимается задача в соответствующей отрасли человеческой деятельности — промышленности, науке, сельском хозяйстве и т. д., для решения которой предполагается использовать дистанционный подход.

К технологиям ДЗЗ предъявляются следующие требования:

- четко поставленная цель относительно конечной реализации технологического процесса — информационного продукта, обеспечивающего решение конкретной тематической задачи или группы задач;
- высокая степень разделения процесса на стадии (фазы);
- системная полнота (целостность) процесса, который должен включать весь набор элементов, обеспечивающих завершенность действий в достижении требуемого конечного результата;
- возможность интерактивного (диалогового) режима работы;
- способность взаимодействовать и функционировать с другими технологиями и системами.

Эти требования являются неизменными условиями достижения результатов в соответствии с поставленной целью, однако, в зависимости от сложности задачи, результат — конечный информационный продукт — может достигаться при различной степени “жесткости” осуществления технологического процесса.

В случае достаточно простых тематических задач этот процесс может быть полностью автоматизирован, и тогда говорят о программируемой технологии. Однако решение большинства тематических задач ДЗЗ требует непосредственного и активного участия специалиста в управлении данными и весьма сложных алгоритмов интеллектуальной обработки, результаты применения которых не всегда дают однозначный результат. Если при этом в технологии последовательность выполнения процессов алгоритмической обработки данных и инфор-

мации остается неизменной, то такую технологию называют научно-технической. Если же в технологии предусмотрена возможность изменять последовательность обработки данных, включать в контур обработки данных для получения желаемого результата некоторые дополнительные логико-вычислительные процедуры, то такую технологию характеризуют как научно-исследовательскую.

В зависимости от сложности тематической задачи ДЗЗ можно ориентироваться на тот или иной тип технологии. Безусловно, при выборе типа технологии важное значение имеют и другие факторы (уровень квалификации потенциальных пользователей, стоимость и др.). Однако при создании новой технологии ДЗЗ можно, независимо от выбранного типа и сферы ее применения, придерживаться единого общего подхода. Такой технологический подход рассматривается ниже.

3. Схема создания новых технологий в ДЗЗ

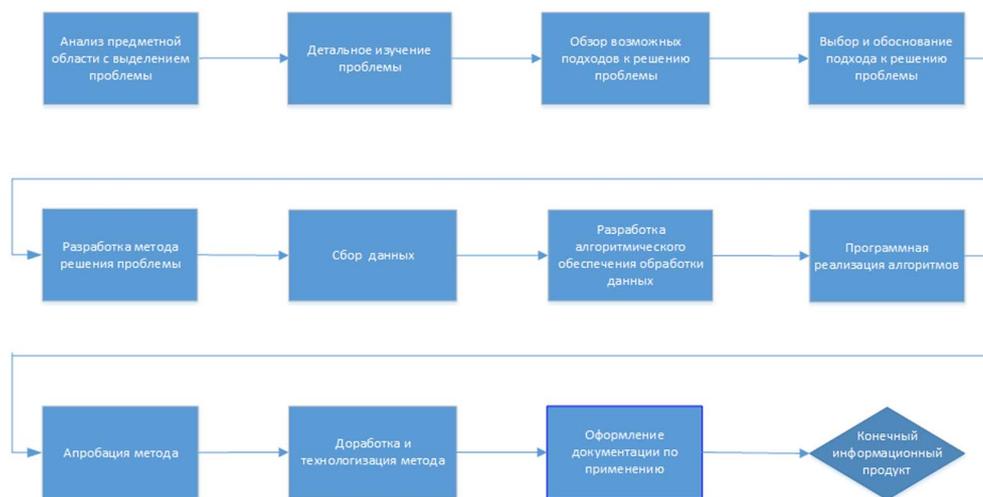
Схема, которая отображает содержание процесса создания новых технологий в различных сферах ДЗЗ, приведена на рисунке (см. ниже).

Новые технологии создаются либо по заказу, либо инициативно, при наличии мотивации со стороны разработчика [6].

В первом случае формулировка проблемы, которая требует нового технологического решения, принадлежит заказчику. Он же, как инвестор разработки и будущий ее пользователь, имеет право задать основные требования к качеству конечного информационного продукта и устанавливает ограничения по срокам создания, стоимости и др. Всю эту информацию заказчик передает исполнителю — разработчику технологии.

Во втором случае инициатором разработки проводится анализ выбранной предметной области и в ней выделяется проблема, требующая решения. Далее эта проблема всесторонне изучается: производится оценка ее актуальности, выявляются причины, по которым она остается нерешенной, оцениваются востребованность создаваемой технологии для ее решения и возможные доходы от ее коммерческого использования и т. д. При положительных результатах такого изучения инициатором разработки решается вопрос с источниками финансирования предстоящей работы и переходят к анализу известных попыток ее решения, для чего привлекаются патенты, научные и научно-технические публикации, отчеты по выполненным в этой области научно-исследовательским работам, другие доступные информационные материалы, в том числе архивные.

По результатам проведенного критического анализа разработчиком выбирается подход к решению проблемы, делаются предварительные оценки его



Технологическая схема создания новых технологий в различных сферах ДЗЗ

перспективности и целесообразности в соответствии с принятыми критериями.

Выбранный подход требует конкретизации в виде структурно-логической схемы метода, которая содержит все стадии (фазы) процесса решения проблемы, обеспечивая системную полноту (целостность) процесса и завершенность действий в достижении конечного результата в виде обусловленного информационного продукта ДЗЗ. Составляется перечень спутниковых изображений, картографических, геологических и других материалов, которых требует метод, формулируются требования к их качеству, объемам, полноте. Ставятся задания на разработку совокупности алгоритмов для реализации логико-вычислительных процедур, предусмотренных телом метода.

Изучаются возможные источники данных, необходимых для применения метода, и производится их сбор. Значительную помощь в получении необходимых материалов (текстовых, графических) могут оказать существующие открытые базы данных и информационно-поисковые системы, в частности, семантические веб-сервисы (Semantic Web Services).

Опыт показывает, что поиск в глобальных сетях оказывается быстрым и результативным, если данные представлены в одном из общепринятых или стандартных форматов и снабжены метаданными, как, например, большинство спутниковых изображений или топокарты. Однако на данное время большая часть имеющихся данных о физических полях Земли не структурирована, а метаданные о них еще ждут своей стандартизации. Поэтому исследователю приходится задавать обобщенные, размытые критерии поиска, вследствие чего собираемые массивы данных оказываются избыточными и далеко не всегда адекватными потребностям задачи и требуют дополнительного рассмотрения и отбора. Существует и еще одна серьезная проблема, с которой приходится сталкиваться при привлече-

нии материалов из глобальных сетей, — это проблема доверия к их точности и достоверности.

Наряду со сбором необходимых исходных данных ведется разработка алгоритмического обеспечения их обработки и анализа. Состав алгоритмического обеспечения определяется содержанием логико-вычислительных процедур, которые должны быть реализованы в разрабатываемом методе. Как правило, при создании алгоритмического обеспечения в значительной мере удастся опираться на уже известные разработки, однако часть алгоритмов всегда является оригинальными.

Одной из логико-вычислительных процедур, которая присутствует практически в любом методе ДЗЗ, является контролируемая классификация. На сегодня известно достаточно много самых различных алгоритмов классификации, что упрощает задачу разработчиков.

Разработанные алгоритмы требуют программной реализации, т. е. разработки специализированного программного обеспечения. Перед тем, как решать эту задачу, необходимо определиться с рядом ключевых вопросов: языком программирования, требованиями к процессору и памяти, выбором операционной системы, протоколами взаимодействия, дизайном интерфейса пользователя и др. После разработки программного обеспечения каждый программный модуль в отдельности и система в целом подвергаются тестированию, в процессе которого выявляются и устраняются ошибки.

Описанная выше и выполненная авторами новой технологии совокупность работ позволяет перейти к апробации разработанного метода. На этом этапе определяется работоспособность метода и оцениваются его возможности по получению конечного информационного продукта, отвечающего принятым критериям.

Один из наиболее важных критериев — это достоверность получаемой дистанционной информа-

ции [7, 8]. Она определяется точностью классификации аэрокосмических изображений. Известно, что при выбранном методе классификации статистическая корректность оценок точности зависит от объема экзаменационной выборки, которая должна включать в себя репрезенты всех учитываемых в задаче классов объектов. Наиболее благоприятный случай — это когда удается сформировать выборку, в которой примерно в одинаковых пропорциях представлены все классы. В таком случае для определения статистически достаточного объема экзаменационной выборки используется подход, основанный на биномиальной модели [9]. Коротко опишем его суть.

Предположим, экзаменационная выборка состоит из N объектов, принадлежащих соответствующим классам (число классов не менее двух), а по результатам классификации получают \mathbf{s} ошибочных результатов. Обозначим P_0 — точность классификации (оценка вероятности правильной классификации), q — оценка вероятности ошибки (неправильной классификации); $q = \mathbf{s} / N$. Очевидно, что $P_0 + q = 1$.

Величина \mathbf{s} является дискретною случайною величиной, вероятность определяется биномиальным распределением

$$\Pr \{ \mathbf{s} = s / q \} = \frac{N!}{(N-s)!s!} q^s (1-q)^{N-s}. \quad (1)$$

Согласно [6], объем N^* экзаменационной выборки, обеспечивающий корректность статистических оценок точности классификации, рассчитывается по формуле:

$$N^* = \frac{Z_a^2 P_0 (1 - P_0)}{b^2}, \quad (2)$$

где Z_a — стандартная табулированная оценка для принятой вероятности α двухсторонней ошибки 1-го рода; b — допустимая (приемлемая) ошибка.

Значение P_0 может быть первоначально выбрано исходя из предыдущих экспериментов или других источников, а потом уточняться. Ошибка 1-го рода задает вероятность α того, что истинное значение вероятности правильной классификации P выйдет за границы доверительного интервала: $\alpha = \Pr \{ |P_0 - P| \geq Z_a \cdot \sqrt{P(1-P)/N} \}$.

Заметим, что из формулы (2) следует, что с ростом вероятности правильной классификации (при увеличении качества классификации) пропорционально понижаются требования к необходимому объему выборки.

Однако во многих практических задачах экзаменационную выборку, в которой примерно одинаково представлены все классы объектов, сформировать не удастся. В подобных случаях применима для определения статистически достаточного объема экзаменационной выборки используется подход, основанный на мультиномиальной модели [10], которая позволяет

учитывать требования к точности классификации на уровне класса. Суть подхода состоит в следующем. Пусть выборка состоит из объектов — репрезентов K классов, причем известны доли Π_i объектов произвольного i -го класса в выборке, $i = 1, 2, \dots, K$. Кроме того, установлена некоторая приемлемая ошибка классификации объектов i -го класса b_i , а также задана вероятность α ошибки 1-го рода. Тогда, согласно [9], необходимая численность n_i объектов i -го класса в выборке может быть рассчитана по формуле:

$$n_i = \frac{B \Pi_i (1 - \Pi_i)}{b_i^2}, \quad (3)$$

где число $B = \chi^2_{(1-\alpha/K)}$ берется из табулированного кумулятивного распределения χ^2 с одной степенью свободы для значения α/K (см., например, [11]).

Для определения потребного объема N^* выборки сначала формулу (3) применяют последовательно для каждого класса, после чего используют правило:

$$N^* = \max_{i=1,2,\dots,K} \{ n_i \} \quad (4)$$

Отметим, что что мультиномиальная модель дает наиболее хорошие результаты в условиях достаточно больших выборок, особенно когда численность объектов какого-то класса доминирует над остальными.

Примеры расчетов по определению статистически достаточного объема экзаменационной выборки при оценке точности результатов контролируемой классификации аэрокосмических изображений для обоих рассмотренных выше моделей можно найти в [12].

Если разработанный метод в результате проведенных испытаний (апробации) продемонстрировал свою эффективность, то переходят к его технологизации. Окончательно определяются с аппаратурой (компьютер, интерфейс, другие необходимые средства) и загружают в нее весь инструментарий разработанного метода. На этом этапе также разрабатывается сопроводительная документация (описание технологии, руководство по применению, указания по эксплуатации и др.).

Описание разработанной технологии должно включать в себя:

1. Название технологии.
2. Область применения.
3. Целевое предназначение.
4. Описание сущности разработанного метода (теоретические предпосылки, модели, расчетные формулы, схема выполнения операций преобразования данных и др.).
5. Требования к исходным данным и другой необходимой информации.
6. Требования к аппаратному обеспечению.
7. Пример (учебный или из практики) использования технологии.

Руководство по применению технологии должно

содержать требования к квалификации пользователей, перечень необходимых исходных данных и пути их получения, последовательность действий пользователя, особенности применения в разных задачах и др.

Указания по эксплуатации технологии должны содержать описание условий, предусмотренных Техническим заданием на проект и необходимых для штатного функционирования аппаратного обеспечения технологии, информацию по проверке работоспособности инструментария технологии, рекомендации по устранению возможных ошибок и неисправностей в системе и т. д.

4. Заключительные замечания

После завершения процесса создания новой технологии ДЗЗ и установления факта ее эффективности встает вопрос о ее дальнейшей судьбе. Если технология создавалась по чьему-то заказу, то она передается заказчику на обусловленных договором или другим документом условиях. При этом для создателя новой технологии очень важным моментом является обеспечить сохранность авторских имущественных прав на нее как объект интеллектуальной собственности. Одним из шагов, способствующих решению данной проблемы, является своевременное оформление патентной заявки на изобретение.

Иная ситуация, если у технологии конкретный заказчик отсутствовал; например, она создавалась в инициативном порядке. Тогда, после получения информационного продукта — новой технологии ДЗЗ, перед разработчиком стоит задача найти и заинтересовать ее возможностями потенциальных пользователей. Достижению этой цели способствует правило: “не забывать, что понятия “технология” и “пользователь” должны рассматриваться в одной связке на всех этапах разработки”.

Итак, в статье изложен технологический подход к поэтапному созданию новых технологий в ДЗЗ, рассмотрено содержание и наиболее важные аспекты выполнения работ на каждом из этапов. Изложенный подход может быть использован как основа при формировании технических заданий и планов выполнения соответствующих исследований и разработок в области ДЗЗ.

Литература

1. Schowengerdt R. A. Remote Sensing: Models and Methods for Image Processing. 3rd Ed. / R. A. Schowengerdt // Amsterdam: Elsevier 2007. — 558 p.
2. Спутниковые методы поиска полезных ископаемых / Под ред. акад. НАН Украины В. И. Лялька и д. т. н. М.А. Попова // Киев: Карбон-Лтд, 2012. — 436 с.
3. Сучасні методи дистанційного пошуку корисних копалин / За ред. В. І. Лялька і М. О. Попова // Київ: НАН України, ЦАКДЗ, 2017. — 221 с. <http://ujrs.org.ua/ujrs/>
4. Cambridge Business English Dictionary <https://dictionary.cambridge.org/dictionary/english/technology>.
5. Великий тлумачний словник сучасної української мови / Уклад. і голов. ред. В. Т. Буссел. — Київ: Ірпінськ: ВТФ “Перун”, 2003. — 1440 с.
6. Hong Z. The Development and Industrialization Recommendation of Current Aerial Remote Sensing Technology / Z. Hong // International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XL-7/W1, 3rd ISPRS IWIDF 2013, 20–22 August 2013, Antu, Jilin Province, PR China. P. 37–39.
7. Попов М. О. Правила та процедура сертифікації методик використання даних дистанційного зондування Землі при вирішенні тематичних задач // М. О. Попов, Т. В. Цимбал // Космічна наука і технологія. — 2004. — Т. 10. — № 4. — С. 104–112.
8. Дистанційне зондування Землі з космосу. Методика оброблення даних: Порядок розроблення / Національний стандарт України ДСТУ 7894:2015. — 8 с.
9. Congalton R. Assessing the Accuracy of Remotely Sensed Data: Principles and Practices. 2nd Ed. / R. Congalton, K. Green // Boca Raton: Taylor & Francis, 2009. — 183 p.
10. Tortora R. A note on sample size estimation for multinomial populations / R. Tortora // American Statistician, 1978, 32 (3). P. 100–102.
11. Вентцель Е. С. Теория вероятностей / Е. С. Вентцель // М.: Наука, 1969. — 576 с.
12. Попов М. А. Методология оценки точности классификации объектов на космических изображениях / М. А. Попов // Проблемы управления и информатики. — 2007. — № 1. — С. 97–103.

ПРО ТЕХНОЛОГІЮ СТВОРЕННЯ НОВИХ ТЕХНОЛОГІЙ В ГАЛУЗІ ДИСТАНЦІЙНОГО ЗОНДУВАННЯ ЗЕМЛІ

М. О. Попов

Розглянуто методичний підхід до поетапного створення нових технологій в галузі дистанційного зондування Землі. Наводиться схема поетапного створення нових технологій в ДЗЗ, обговорюються зміст і найбільш важливі аспекти робіт на кожному з етапів. Викладений підхід рекомендується для використання в якості методичної основи при формуванні технічних завдань і планів виконання відповідних досліджень і розробок в області ДЗЗ.

Ключові слова: дистанційне зондування Землі, аерокосмічне зображення, вимоги до ДЗЗ-технології, класифікування зображень, збирання даних, обсяг вибірки

ON THE TECHNOLOGY OF CREATION OF NEW TECHNOLOGIES IN THE AREA OF REMOTE SENSING OF THE EARTH

M. O. Popov

The methodical approach to the phased elaboration of new technologies in the area of remote sensing of the Earth is outlined. A scheme for the phased elaboration of new technologies in remote sensing is suggested, the content and the most important aspects of the work at each stage are discussed. This approach is offered to use as a methodological basis at the forming specifications and the implementation plans of relevant research and development in the area of remote sensing.

Keywords: remote sensing of the Earth, aerospace image, requirements for RS-technology, image classification, data collection, sample size