

# СЕКЦИЯ 13: МЕТОДОЛОГИЯ, МЕТОДЫ И ПРОГРАММНО-АЛГОРИТМИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ОБРАБОТКИ И ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА БОЛЬШИХ МАССИВОВ ИНФОРМАЦИИ

## МЕТОДЫ ОБРАБОТКИ ГИПЕРКУБА ИНФОРМАЦИИ В ПРИЛОЖЕНИЯХ С ИСПОЛЬЗОВАНИЕМ ОБЛАЧНЫХ ВЫЧИСЛЕНИЙ

Алчинов А.И., Жигалов К.Ю.,

*Институт проблем управления им. В.А. Трапезникова,  
Россия, г. Москва, ул. Профсоюзная, д.65  
alchinov46@mail.ru, kshakalov@mail.ru*

*Аннотация: В области дистанционного зондирования Земли активно внедряются гиперспектральные системы. Полученные при этом гиперспектральные изображения представляют большой интерес для решения широкого круга научных и различных практических задач. Однако реализация возможностей гиперкуба информации сдерживается отсутствием необходимых приложений и соответствующих вычислительных систем для обработки изображений, получаемых несколькими сотнями каналов гиперспектрометра. Рассмотрены методы удаленного доступа, анализа, хранения и обработки гиперкуба информации в облачных средах и некоторые подходы использования облачных вычислений.*

Ключевые слова: гиперспектральные изображения, облачные вычисления, гиперкуб информации, квантовые вычисления, дистанционное зондирование земли.

### Введение

Облачные технологии в России находятся на начальном этапе своего развития и связаны с разработкой методов хранения данных на удаленных серверах и методов доступа к данным, разработкой и выбором приложений для их использования различными устройствами доступа и технологиями подключения.

Указанные направления в настоящее время активно развиваются. При этом производится стандартизация и универсализация облачных услуг, приводится в соответствие с новыми решаемыми задачами российское законодательство.

Центрами компетенций по данному направлению являются: МГУ имени М. В. Ломоносова, МФТИ, Mail.ru, Yandex.Cloud - связанные сервисы, облачная платформа Яндекс «Яндекс. Облако», Amazon (Simple Storage) Service, Microsoft (Microsoft Blobs Storage), Google (Google Cloud Storage), Сбербанк и некоторые другие. Эти же компании, являются разработчиками технологий хранения и анализа больших данных с использованием нейронных сетей и глубинного машинного обучения с подкреплением, расположенных в распределенных дата-центрах. Однако в указанных центрах компетенции методы обработки гиперспектральной информации для создания карт в интересах решения задач управления в центрах компетенции, на момент написания данной статьи, не рассматриваются.

Главными преимуществами облачных сервисов являются: быстрдействие и масштабируемость в процессах обработки данных.

Известен проект применения облачных вычислений в виде набора Web-сервисов и систем разработки Web-приложений от таких компаний, как Google, AT&T, Nvidia, Amazon. Система Synaptic Hosting от AT&T, используя которую пользователь получает виртуальную среду с разными вариантами хранения данных, сетевых архитектур, с поддержкой приложений и обеспечением удаленных вычислительных мощностей для них [1]. Система легко масштабируется как в плане новых сервисов, так и в плане нагрузок.

Пользователи могут размещать необходимое облачное оборудование на своих ресурсах и применять инфраструктуру, необходимую для создания собственных локальных облачных сервисов или иметь дистанционный доступ к сервисам с обеспечением необходимого уровня защищенности данных.

По состоянию на сегодняшний день является нерешенная задача по созданию регламента и методик, применяемых для всех наборов данных гиперспектральной информации таким образом, чтобы размещение в облачной среде обеспечивало бы их максимальную защищенность. С

появлением во многих странах гиперспектральной аппаратуры, количественные и качественные характеристики получаемых результатов стали публиковаться мало.

## 1 Облачные технологии

Облачный сервис различного применения является независимым решением для ограниченных условий, позволяющий автономно выполнять задачи с возможностью обмена данными при сохранении защищенной связи с устройствами любого типа. Например, МО США уже практически завершили формирование Единого облака оборонной инфраструктуры JEDI, обеспечивающего новые возможности по сбору, обработке, обмену информацией.

Новейшие технологии позволяют через облако получать программное обеспечение SaaS, инфраструктуру IaaS и платформу PaaS.

Основная проблема использования облачных сервисов – недостаточная скорость доступа к данным и их обработки. Для этой цели исследователи рассматривают возможности использования квантовых компьютеров и фрактальных методов сжатия.

Квантовые компьютеры создают такие крупные организации, как: Google, IBM, Intel и ряд небольших ИТ-компаний и стартапов: D-Wave, Rigetti. Они открывают доступ к своим решениям для университетов и исследовательских организаций с помощью облачной инфраструктуры. Для работы с установками используются облачные платформы IBM Q и Alibaba Cloud. Канадская компания D-Wave Systems тоже запустила свой облачный сервис. Её клиенты получают доступ к квантовым вычислениям на минуту в месяц. Для выполнения ряда задач квантовой машине достаточно всего нескольких миллисекунд. Речь идет о факторизации больших чисел и построении маршрутов. На подобные вычисления классические компьютеры тратят дни и недели. Несколько компаний уже применяют возможности системы D-Wave. Компания Airbus использовала компьютеры на основе кубитов для создания новых самолётов — необходимые расчёты удалось закончить в четыре раза быстрее, чем на обычных вычислительных машинах. Volkswagen использует квантовые компьютеры для проработки алгоритмов управления беспилотными автомобилями.

Сфера квантовых вычислений развивается и в Китае. В 2016 году команда исследователей запустила на орбиту спутник, в котором каналы связи защищены при помощи квантового шифрования. Эту технологию применяют и в наземных сетях: между Пекином и Шанхаем проложили квантовый канал связи длиной около 2000 километров. Квантовыми вычислениями занимаются и в России. В феврале 2018 года несколько организаций: Фонд перспективных исследований, Внешэкономбанк, МГУ, АНО «Цифровая экономика» — создали квантовый консорциум для разработки 50-кубитного компьютера уже в этом году[3]. Кроме того, квантовые технологии входят в список приоритетных сквозных технологий Национальной технологической инициативы (НТИ). Предполагается, что к 2024 году на базе ГК «Росатом» и других участников программы будет создан российский квантовый компьютер, способный решать задачи анализа широкого спектра.

По ряду направлений в облачных технологиях уже предложены решения, связанные с созданием новой элементной базы и ее архитектуры, обеспечивающей формирование «распределенных» квантовых вычислений в облачных средах с помощью квантовых компьютеров в виде экспериментальных образцов, которые необходимо ускоренными темпами применять в России.

В России созданы специализированные программно-технические центры для обработки больших данных в некоторых организациях с использованием высокопроизводительных вычислений в облачных средах [4]. В этих центрах необходимо создать и центры приема гиперкуба информации и его обработки для создания тематических карт.

При этом высокая скорость обмена информацией и высокая степень кодирования обеспечат успешную реализацию применения информационных технологий. Необходимо обеспечить наличие высокоскоростного интернета в пунктах обработки больших объемов данных для реализации облачных технологий. Проблема наличия устойчивого и высокоскоростного интернета в некоторых регионах огромной России есть. При этом исчезнет проблема поиска нужного компьютера или установки переключателей или концентраторов сетей для организации связи между сетями.

В России одновременно развиваются классические методы обработки гиперспектральных изображений[5]. Эти методы в дальнейшем необходимо размещать на удаленных серверах, оснащенных высокопроизводительными компьютерами. Значительное количество каналов, используемых при получении гиперспектральных изображений, порождает проблему "проклятия

размерности", которая обуславливает непригодность традиционных алгоритмов создания тематических карт.

## 2 Гиперспектральные системы зондирования Земли

Существующая в настоящее время в России система информационного обеспечения с использованием гиперкуба информации находится на начальной стадии развития.

В 2019 году разработана гиперспектральная аппаратура СОМ-ГС, с характеристиками, существенно превышающими зарубежные образцы (по открытым публикациям). Например, спектральных каналов 550, пространственное разрешение на местности около 18 м.

Однако слабо еще разработаны теория и методы обработки гиперспектральных изображений. Имеющееся программное обеспечение, практически, не справляется с большими объемами данных, получаемых 550 каналами спектрометра. В настоящее время общедоступных наборов гиперспектральных данных мало, поэтому необходимо одновременно наращивать использование гиперспектральной аппаратуры для производства съемок и создания актуальной базы данных об объектах местности, включая общепринятую классификацию объектов местности, отображаемых на картах и необходимых для решения задач управления в соответствии с нормативно-техническими документами по созданию карт и планов.

Ниже приведена таблица 1 со сравнительными характеристиками гиперспектральной аппаратуры (российской и зарубежными) и рисунок 1 с характеристиками каналов формирования изображения [4].

Таблица 1. Технические характеристики гиперспектрометров

Характеристики Гиперспектрометров	Показатели							
	Российские и зарубежные гиперспектрометры							
	Ресурс-П Россия 2013	HIRIS США 1994	AVIRIS США 1999	HyMap Австралия 1998	CHRIS ЕКА, Европа 2001	MODIS США 1999	HY-PERION США 2000	PRISMA ЕКА Европа 2018
Масса, кг.	-	-	-	-	14	-	49	90
Количество спектральных каналов	96	192	224	126	56	36	220	237
Спектральное разрешение, нм.	5-10	9,5-12	10	14,3	1,5-11	-	10	10-12
Спектральный диапазон, мкм.	0,4-1,1	0,4-2,5	0,4-2,5	0,4-2,5	0,41-1	-	0,4-2,5	450-2500
Разрешение, на Местности, м	30	25	-	-	25-50	250-1000	30	30
Полоса обзора, Км.	25	-	-	-	13	233	7,5	30

Российский гиперспектрометр «Ресурс-П» сейчас практически не используется. Зарубежные аналоги HIRIS, AVIRIS, HyMap, CHRIS-1, MODIS, HYPERION, PRISMA ЕКА имеет высокую стоимость, большую массу гиперспектральной аппаратуры. Наиболее современным техническим уровнем обладает PRISMA ЕКА.

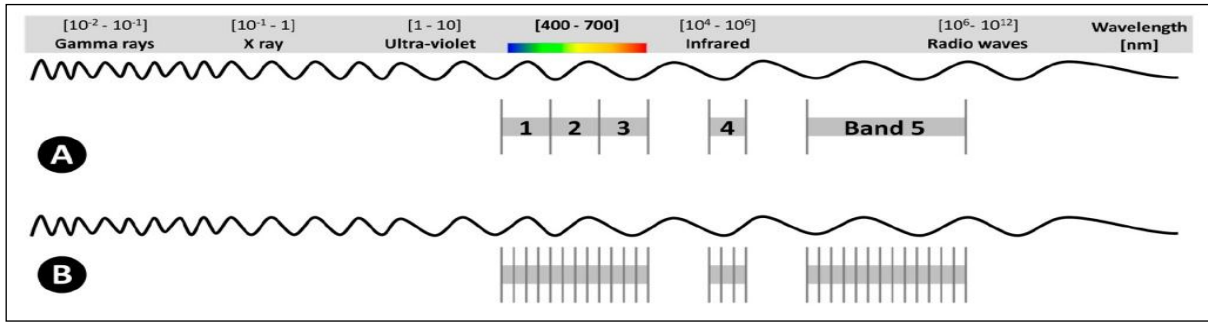


Рис. 1. Сравнительная характеристика каналов формирования изображения

(А) - мультиспектральный диапазон, включающий 5 каналов;

(В) - гиперспектральный диапазон, включающий узкие полосы – каналы. Полос-каналов – может быть несколько сотен (Российский гиперспектрометр – 550 полос, таблица 1).

Гиперспектральные изображения имеют большой потенциал для перехода на совершенно новый уровень дешифрирования в военной области, обеспечивая при этом даже вскрытие замаскированных объектов [5,6].

Давайте рассмотрим пример оптимизации гиперспектрального изображения для маскировки ниже.

Пусть известно распределение яркостей физического поля Земли  $a_{ij}$  в районе маскируемого объекта. Имеется  $N$  средств маскировки, причем  $k$ -е средство при нахождении в начале координат создает известную добавочную яркость физического поля  $\delta_k(i, j)$ . Пусть средства  $k$ -е средство размещено в точке с координатами  $(x_k, y_k)$  ( $k = 1, \dots, N$ ). Необходимо вычислить оптимальные параметры  $A_k$  каждого из  $N$  средств маскировки, обеспечивающие минимум корреляции между текущим и эталонным изображениями. Их взаимная корреляция вычисляется по окну с размерами  $1 \leq i \leq M_x, 1 \leq j \leq M_y$ .

В случае описанной расстановки средств маскировки яркость искаженного на текущей карте (изображении) в точке  $(i, j)$  равна

$$a_{ij} + \sum_{k=1}^N A_k \delta_k(i - x_k, j - y_k). \quad (1)$$

Корреляция между текущим и эталонным изображениями равна:

$$C = \frac{\sum_{i=1}^{M_x} \sum_{j=1}^{M_y} a_{ij} \left( a_{ij} + \sum_{k=1}^N A_k \delta_k(i - x_k, j - y_k) \right)}{\sqrt{\sum_{i=1}^{M_x} \sum_{j=1}^{M_y} a_{ij}^2} \cdot \sqrt{\sum_{i=1}^{M_x} \sum_{j=1}^{M_y} \left( a_{ij} + \sum_{k=1}^N A_k \delta_k(i - x_k, j - y_k) \right)^2}}. \quad (2)$$

Необходимо найти значения вектора амплитуд  $(A_1, \dots, A_N)$ , при котором значение  $C$  минимально. Условием минимальности является равенство нулю всех частных производных  $C$  по  $A_k$ . Вычислим частные производные и получим систему из  $N$  квадратных уравнений с  $N$  неизвестными.

Коэффициент при  $A_k^2$  равен

$$\sum_{i=1}^{M_x} \sum_{j=1}^{M_y} a_{ij} \delta_s(i - x_s, j - y_s) \cdot \sum_{i=1}^{M_x} \sum_{j=1}^{M_y} \delta_k(i - x_k, j - y_k)^2 - \sum_{i=1}^{M_x} \sum_{j=1}^{M_y} a_{ij} \delta_k(i - x_k, j - y_k) \cdot \sum_{i=1}^{M_x} \sum_{j=1}^{M_y} (\delta_s(i - x_s, j - y_s) \cdot \delta_k(i - x_k, j - y_k)). \quad (3)$$

При  $s = k$  получается 0,  $\frac{\partial C}{\partial A_s}$  - зависит от  $A_s$  линейно, при фиксированных  $A_k, k \neq s$ , оптимальное значение  $A_s$  получается однозначно.

Приведем алгоритм нахождения оптимальных параметров средств маскировки:

Вычисляем квадратичные функции, являющиеся числителями полученных выражений для частных производных.

Цикл до сходимости. В каждом цикле:

Цикл по  $s$  от 1 до  $N$ .

В каждом цикле: Вычисляем при фиксированных  $A_k$ ,  $k \neq s$ , оптимальное  $A_s$ , получаемое однозначно.

Каждая итерация обеспечивает минимизацию корреляции за счет выбора одной амплитуды. В результате каждой итерации значений корреляции уменьшается, при этом монотонность изменения целевой функции обеспечивает сходимость приведенного алгоритма. Для реализации приведенного алгоритма нужны значительные вычислительные ресурсы.

### **3 Использование облачных вычислений при обработке гиперспектральных изображений**

В настоящее время широкое распространение получает концепция «облачных вычислений», суть которой состоит в том, что вычислительные ресурсы и среды предоставляются пользователям удаленно через сеть Internet. Это освобождает пользователей от необходимости обслуживать задействованное в вычислениях аппаратное и программное обеспечение («облако») и позволяет им сосредоточиться непосредственно на решении прикладных задач.

Компания Hewlett-Packard, предоставляет свои вычислительные мощности через набор виртуальных машин, работающих через один web-портал от Hewlett-Packard. Альянс Intel, Yahoo! и HP под названием Cloud Computing Test Bed, представляющий собой тестовый полигон для облачных вычислений, который состоит из шести дата-центров, в каждом из которых находится «облако» из нескольких тысяч процессоров.

Основным недостатком и ограничением применения концепции «облачных вычислений» в ГИС является узкополостность информационного канала, связывающего пользователя системы с ее сервисами. При этом стандартных алгоритмов сжатия информации, передаваемой по информационному каналу, уже недостаточно для обеспечения нормальной работы пользователя в системе в процессе решения большинства геоинформационных задач, особенно при обработке гиперспектральных и радиофотонных изображений местности. Для решения этой проблемы необходимо использование высокоуровневых стратегий, например, уменьшение количества обращений пользователя к сервисам приложений.

В общем случае нужна строгая формулировка задачи для отправки на обработку в web-сервис. При этом можно получить решение этой задачи. Здесь необходима декомпозиция общей задачи обработки изображений [7]. Частных задач может быть много. Например, обработка изображений с целью дешифрирования объектов местности может включать задачи по распознаванию объектов местности, сначала по типам, затем конкретно, по каждому объекту. Некоторые позиции при этом сложно формализовать.

Пока решение картографической задачи может быть представлено в виде запроса, аналогичного по форме запросу SQL. GIS Webservice Special Edition (SE) – web-сервис для публикации пространственных данных по протоколам: OGC WMS, OGC WMTS, OGC WFS, OGC WFS - T, OGC WCS на виртуальных платформах.

Алгоритм решения картографических задач необходимо рассматривать в виде последовательности элементарных операций, которые бы выполнил картограф в процессе решения конкретной задачи. Решение задачи будет представлять преобразование или получение картографических данных с помощью имеющегося в его распоряжении web-приложения с учетом опыта картографа.

Задачи по обработке изображений для создания карт и планов должны быть типизированы и классифицированы. Решения должны быть представлены в виде иерархического дерева, в котором произведена декомпозиция основной задачи до набора элементарных процедур. Пользователь будет вынужден загрузить с сервера ветви дерева решения подобной задачи, отобрать для его собственной задачи, заменить неподходящие ветви подходящими из набора предоставляемых системой ветвей решения или создать свои.

Использование концепции «облачных вычислений» в реализации распределенных геоинформационных систем обладает рядом важных достоинств - платформенной независимостью, освобождением конечных пользователей от поддержки вычислительных систем.

Популярность модели «все как услуга» (Everything-as-a-Service, XaaS). Такие корпорации, как Adobe, Workday, LinkedIn и Salesforce, реализовали данную модель, предоставив пользователям возможность использовать приложения в облаках. Для большинства из нас эта модель ассоциируется с подпиской на приложения и развлекательный контент. Подобные модели могут быть реализованы и при создании тематических карт, проводя основные вычислительные операции в облаках.

Коммерческие программные продукты российского рынка тяжело конкурируют с зарубежными, поскольку значительная часть компонентной базы выпускается за рубежом.

## Заключение

Быстрое развитие научного направления, связанного с получением гиперкуба информации о земной поверхности, обуславливают необходимость оперативного внедрения в практику средств получения гиперкуба информации методов обработки этой информации. Разработанные методы создания карт необходимо развивать и адаптировать для работы с большими объемами данных в облачных средах с использованием квантовых компьютеров.

Некоторое отставание России от США и некоторых ведущих западных стран в области разработки средств дистанционного зондирования Земли, конкретно в области гиперспектрометрии немного усложняет проблему обеспечения автоматизации создания картографического материала в России. Однако существующие методы обработки материалов гиперспектральной съемки позволяют решать прикладные задачи с их использованием.

Недостаточная готовность промышленности по созданию современных средств получения гиперкуба информации из-за возможного отсутствия элементной базы разрабатываемого оборудования также сдерживает практическое применение гиперспектрометрии.

Также недостаточно разработаны методы обработки больших объемов информации, получаемых средствами дистанционного зондирования, с использованием облачных вычислений.

Применение облачных вычислений может дать преимущества по сравнению со стандартным протоколом работы обычных ГИС - приложений:

- пользователи могут использовать любые удобные для работы мобильные устройства и операционные системы;
- отсутствует необходимость в наличии большого количества прикладных приложений, ориентированных на работу с определенными web-сервисами или предназначенными для решения конкретных задач - все необходимые приложения пользователь получает, подключившись к «облаку»;
- автоматическая поддержка масштабируемости «облака» представляет удобную среду для используемых приложений.

Для использования облачных вычислений необходимо произвести декомпозицию общей задачи вычислений до отдельных операций по направлениям использования приложений, которая находится на завершающем этапе.

## Литература

1. *Василенко В.С.* Применение концепции облачных вычислений в ГИС. Научно-технический и прикладной журнал «ИЗВЕСТИЯ ЮФУ. ТЕХНИЧЕСКИЕ НАУКИ», с.227-231.
2. <https://russiandrone.ru/publications/vozmozhnosti-oblachnykh-tehnologiy-v-voennoy-sfere>. Теги: 1cloud.ru, IaaS, VPS, VDS, Частное и публичное облако, SSL, 02.02.2019 г.
3. *С.М. Борзов, П.В. Мельников, И.А. Пестунов, О.И. Потатуркин, А.М. Федотов.* Комплексная обработка гиперспектральных изображений на основе спектральной и пространственной информации. ИВТ СО РАН, Вычислительные технологии, том 21, №1, 2016 г., г. Новосибирск, с.25-37.
4. <http://gistechinik.ru/pub/3-publik/37-ws.html>.
5. <https://russiandrone.ru/publications/vozmozhnosti-oblachnykh-tehnologiy-v-voennoy-sfere>.
6. *Telmo Adão, Jonáš Hruška, Luís Pádua, José Bessa, Emanuel Peres, Raul Morais, Joaquim João Sousa.* Гиперспектральные снимки: обзор сенсоров для БПЛА, систем обработки данных и приложений для сельского и лесного хозяйства. 29.01.2019 г. (Совзонд).