

**ОБРАБОТКА СПЕКТРАЛЬНОГО ИНДЕКСОВ GNDVI И CVI
НА ОСНОВЕ ДАННЫХ ДИСТАНЦИОННОГО ЗОНДИРОВАНИЯ**

К.К.Сеитназаров – доктор технических наук, профессор
Нукусский государственный технический университет

С.М.Мадиримова – соискатель
Ташкентский университет информационных технологий

Б.К.Туремуратова – магистрант

Б.Б.Айтмуратов – студент
Нукусский государственный технический университет

**GNDVI VA CVI СПЕКТРАЛ ИНДЕКСЛАРИНИ МАСОФАВИЙ ЗОНДЛАШ
МАЪЛУМОТЛАРИ АСОСИДА ИШЛАБ ЧИҚИШ**

К.К.Сеитназаров – техника фанлар доктори, профессор
Нукус давлат техника университети

С.М.Мадиримова – тадқиқотчи
Тошкент ахборот технологиялари университети

Б.К.Туремуратова – магистрант

Б.Б.Айтмуратов – студент
Нукус давлат техника университети

PROCESSING GNDVI AND CVI SPECTRAL INDICES BASED ON DISTANCE SCOMING DATA

K.K.Seitnazarov – Doctor of Technical Sciences, Professor
Nukus State Technical University

S.M.Madirimova – applicant
Tashkent University of Information Technologies

B.K.Turemuratova – master's student

B.B. Aitmuratov – student
Nukus State Technical University

Таянч сўзлар: ернинг масофадан зондланиши, мультиспектрал тасвирлар, маълумотларни даслабки қайта ишлаш, вегетацион индекслар, GNDVI, CVI, атмосфера коррекцияси, эксперт тизимлари, геоинформацион технологиялар.

Резюме. Мазкур мақолада ўсимлик қопламининг ҳолатига, айниқса хлорофилл микдорига юқори сезгирликка эга бўлган GNDVI ва CVI вегетацион индексларини ҳисоблаш масалалари кўриб чиқилади. Шунингдек, модуль архитектураси конвеер (pipeline) ёндашуви асосида ишлаб чиқилган бўлиб, бу катта ҳажмдаги сунъий йўлдош тасвирларини қайта ишлаш жараёнида модуллилик, масштабланувчанлик ва барқарорликни таъминлайди. Амалга ошириш жараёнида ҳисоблашларни опти-маллаштириш, шунингдек PostgreSQL/PostGIS геофазовий маълумотлар базаси билан интеграция масалаларига алоҳида эътибор қаратилган.

Ключевые слова: дистанционное зондирование Земли, мультиспектральные изображения, предварительная обработка данных, вегетационные индексы, GNDVI, CVI, атмосферная коррекция, экспертные системы, геоинформационные технологии.

Резюме. В данной статье проводится вычислению индексов GNDVI, CVI, которые обладают высокой чувствительностью к состоянию растительного покрова, в особенности содержанию хлорофилла. Также стоит отметить, что архитектура модуля построена на основе конвейерного подхода, которая обеспечивает модульность, масштабируемость и устойчивость обработки больших объёмов спутниковых изображений. Реализация включает оптимизацию вычислений, а также интеграцию с геопространственной базой данных PostgreSQL/PostGIS.

Key words: earth remote sensing, multispectral imagery, data preprocessing, vegetation indices, GNDVI, CVI, atmospheric correction, expert systems, geoinformation technologies

Summary. This article focuses on the computation of the GNDVI and CVI vegetation indices, which exhibit high sensitivity to vegetation condition, particularly to chlorophyll content. It is also noted that the module architecture is based on a pipeline approach, which ensures modularity, scalability, and robustness in the processing of large volumes of satellite imagery. The implementation includes computational optimization as well as integration with a geospatial PostgreSQL/PostGIS database.

Введение. В этой связи особую актуальность приобретает разработка специализированных программных модулей, обеспечивающих преобразование «сырых» спутниковых данных в калиброванные и нормализованные показатели, пригодные для применения механизмов логического вывода и экспертных систем. Вегетационные индексы, такие как GNDVI, CVI, широко используются для оценки состояния растительности, продуктивности биомассы и уровня хлорофилла, однако точность их расчёта напрямую зависит от корректности предшествующих этапов обработки.

Метаданные сцены, включающие дату съемки, углы наблюдения Солнца, параметры сенсора и коэффициенты радиометрической калибровки, извлекаются из сопутствующих файлов и сохраняются в объекте SceneMetadata. Эта информация используется на последующих этапах обработки для применения корректных преобразований.

Методы. Класс RadiometricCalibrator выполняет преобразование значений цифровых чисел, хранящихся в исходных файлах, в физически значимые величины отражательной способности. Для данных Landsat применяется двухэтапное преобразование: сначала DN конвертируются в значения радиации на уровне сенсора с использованием коэффициентов масштабирования и смещения, специфичных для каждого спектрального канала [1, 2]:

$$L_{\lambda} = M_L \times DN + A_L$$

где L_{λ} – спектральная радиация, M_L – мультипликативный коэффициент масштабирования, A_L – аддитивный коэффициент смещения. Затем радиация преобразуется в отражательную способность на уровне сенсора (TOA reflectance):

$$\rho_L = \frac{\pi \times L_{\lambda} \times d^2}{ESUN_{\lambda} \times \cos(\theta_s)}$$

где d – расстояние Земля-Солнце в астрономических единицах, $ESUN_{\lambda}$ – среднее значение солнечной экзотмосферной освещенности для канала λ , θ_s – зенитный угол Солнца. Для Sentinel-2 используются предоставляемые в метаданных коэффициенты QUANTIFICATION_VALUE для прямого преобразования DN в отражательную способность.

Класс CloudMaskGenerator формирует бинарные маски облачности и теней облаков, необходимые для исключения недостоверных пикселей из анализа. Для данных Landsat-8/9 используется алгоритм Quality Assessment band, основанный на анализе битовых флагов в специальном канале качества BQA, поставляемом вместе с данными. Алгоритм извлекает

информацию о высокой и средней уверенности наличия облаков, облачных теней, снега и насыщения сенсора, формируя комплексную маску качества.

Для сцен без готовых масок качества или для их уточнения применяется алгоритм на основе анализа спектральных характеристик. Облака идентифицируются по высоким значениям отражательной способности в видимом диапазоне и низкой температуре яркости в тепловых каналах. Используется пороговый классификатор с адаптивными порогами, вычисляемыми на основе статистики распределения значений в сцене:

$$mask_{cloud} = (p_{blue} > \mu_{blue} + 2\sigma_{blue}) \wedge (p_{green} > \mu_{green} + 2\sigma_{green}) \wedge (p_{blue} > \mu_{blue} + 2\sigma_{blue})$$

где μ и σ – среднее значение и стандартное отклонение соответствующих каналов. Тени облаков выявляются по аномально низким значениям в ближнем инфракрасном диапазоне в сочетании с нормальными значениями в видимом диапазоне. Для уточнения границ облаков и теней применяются морфологические операции расширения, что гарантирует исключение переходных зон с частичной облачностью.

Особое внимание уделяется обработке граничных случаев: деление на ноль, недопустимые значения, выход за допустимый диапазон. Использование контекстного менеджера `np.errstate` предотвращает генерацию предупреждений при делении на ноль, а функция `np.where` обеспечивает замену недопустимых результатов нулевыми значениями. Функция `np.clip` ограничивает итоговые значения диапазоном $[-1, 1]$, что соответствует теоретически возможному диапазону индекса [3, 4].

Расчёт индекса GNDVI (с англ., Green Normalized Difference Vegetation Index) выполняется классом GNDVICALculator по формуле:

$$GNDVI = (NIR - GREEN) / (NIR + GREEN)$$

На рис.1. показан расчёт индекса GNDVI на примере мультиспектральных данных Нукусского района. Реализация аналогична расчёту NDVI, но использует зеленый канал вместо красного, что обеспечивает более высокую чувствительность к содержанию хлорофилла:

```
def calculate_gndvi(nir, green):
    numerator = nir - green
    denominator = nir + green
    with np.errstate(divide='ignore', invalid='ignore'):
        gndvi = np.where(denominator != 0, numerator / denominator, 0)
    gndvi = np.clip(gndvi, -1.0, 1.0)
    return gndvi
```

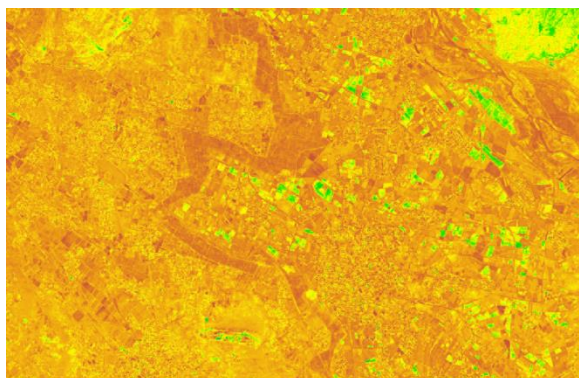


Рис.1. показан расчёт индекса GNDVI на примере мультиспектральных данных Нукусского района

Индекс CVI (с англ., Chlorophyll Vegetation Index) вычисляется классом CVICalculator согласно формуле:

$$CVI = (NIR \times RED) / (GREEN^2)$$

На рис.2. показан расчёт индекса CVI на примере мультиспектральных данных Нукусского района. Этот индекс особенно чувствителен к концентрации хлорофилла благодаря использованию произведения NIR и RED в числителе и квадрата GREEN в знаменателе:

```
def calculate_cvi(nir, red, green):
    numerator = nir * red
    denominator = green ** 2
    with np.errstate(divide='ignore', invalid='ignore'):
        cvi = np.where(denominator != 0, numerator / denominator, 0)
    # CVI не имеет фиксированного верхнего предела,
    # но значения выше 10 обычно нереалистичны
    cvi = np.clip(cvi, 0.0, 10.0)
    return cvi
```

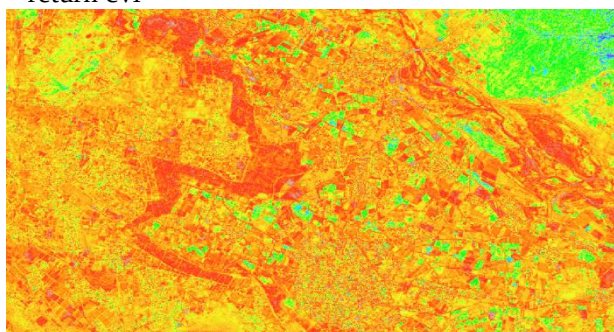


Рис.2. показан расчёт индекса CVI на примере мультиспектральных данных Нукусского района

Для CVI применяется иной диапазон ограничения (0-10), поскольку этот индекс не является нормализованным и теоретически может принимать любые положительные значения, хотя практически значимые значения редко превышают 3-5 [5, 6].

Класс IndexStatisticsCalculator вычисляет статистические характеристики рассчитанных индексов, необходимые для контроля качества и последующего анализа. Для каждого индекса

определяются: минимальное и максимальное значения, среднее арифметическое, медиана, стандартное отклонение, процентиля (5%, 25%, 75%, 95%) и гистограмма распределения значений. Статистика вычисляется только для валидных пикселей, исключая области, замаскированные как облака, тени или NoData, на рис.3. показана блок-схема принципа работы класса.

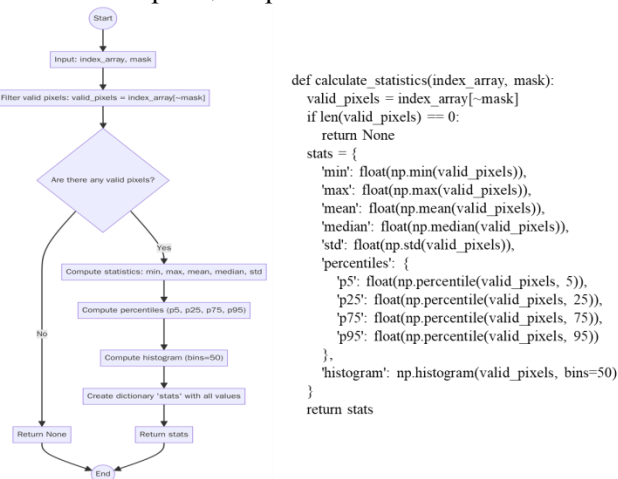


Рис.3. Процесс работы классификатора индексов

Обработка больших сцен и оптимизация производительности. При работе с полноразмерными спутниковыми сценами загрузка всех данных в память может быть неэффективной или невозможной на машинах с ограниченными ресурсами. Для решения этой проблемы реализована блочная обработка через класс TiledProcessor.

Сцена разбивается на регулярную сетку блоков размером 1024×1024 пикселя с перекрытием по 32 пикселя на каждой стороне для корректной обработки граничных областей. Каждый блок загружается, обрабатывается и сохраняется независимо, что позволяет поддерживать постоянное потребление памяти независимо от размера сцены:

```
def process_scene_tiled(scene_path,
    tile_size=1024, overlap=32):
    with rasterio.open(scene_path) as src:
        width, height = src.width, src.height
        for row_start in range(0, height, tile_size - overlap):
            for col_start in range(0, width, tile_size - overlap):
                # Определение границ текущего тайла
                window = Window(col_start, row_start,
                    min(tile_size, width - col_start),
                    min(tile_size, height - row_start))
                # Загрузка данных тайла
                tile_data = src.read(window=window)
                # Обработка тайла
                processed_tile = process_tile(tile_data)
                # Сохранение результата
                save_tile(processed_tile, row_start, col_start)
```

Для дополнительного ускорения обработки применяется параллелизм на уровне тайлов: независимые блоки обрабатываются параллельно в нескольких процессах с использованием multiprocessing.Pool. Количество параллельных процессов адаптивно выбирается на основе доступных процессорных ядер и оперативной памяти.

Класс MemoryManager контролирует использование памяти в процессе обработки. При приближении к заданному лимиту потребления памяти менеджер инициирует освобождение кэшированных данных и принудительную сборку мусора. Это предотвращает исчерпание памяти и связанные с ним сбои при обработке больших объемов данных [7, 8].

Интеграция с базой данных и экспорт результатов. После завершения расчёта индексов результаты сохраняются в базе данных PostgreSQL/PostGIS через класс IndexRepository. Сохранение выполняется в два этапа: сначала создаются записи метаданных в таблице vegetation_indices с информацией о сцене, территории, типе индекса и статистических характеристиках, затем загружаются растровые данные [11, 12].

Для эффективной загрузки растров используется подход с промежуточными GeoTIFF файлами: рассчитанные NumPy массивы сохраняются во временные GeoTIFF файлы с корректными геопространственными метаданными, которые затем импортируются в PostGIS с

использованием функции raster2pgsql или напрямую через ST_FromGDALRaster. После успешной загрузки в базу данных временные файлы удаляются.

Класс IndexExporter обеспечивает экспорт рассчитанных индексов в стандартные форматы для использования в внешних ГИС-приложениях или программах обработки изображений. Поддерживаются форматы GeoTIFF, GeoJSON и Cloud Optimized GeoTIFF для эффективного веб-доступа к данным. При экспорте сохраняются все пространственные метаданные, цветовые палитры для визуализации и дополнительные атрибуты в тегах файлов.

Заключение. Модульная архитектура, использование оптимизированных вычислительных методов и поддержка обработки больших сцен гарантируют масштабируемость и устойчивость программного комплекса. Полученные результаты подтверждают, что рассчитанные индексы GNDVI и CVI обладают высокой стабильностью и пригодны для последующего применения в механизмах логического вывода экспертных систем.

Разработанный модуль может быть использован для задач экологического мониторинга, анализа агроландшафтов и оценки состояния растительного покрова, а также служит основой для дальнейшего развития интеллектуальных геоинформационных систем и баз знаний в области дистанционного зондирования Земли.

Литература

1. Мельников А.В. Обработка многозональных изображений. – Москва: «Наука», 202
2. Seitnazarov K. et al. Hydrogeological-Mathematical Model of Formation and Management of Resources and Quality of Fresh Underground Water of the Karakalpak Artesian Basin. // International Conference on Information Science and Communications Technologies (ICISCT). – IEEE, 2021. – P. 01-05.
3. Ткаченко В.Н. Основы дистанционного зондирования Земли. – Санкт-Петербург: «Политехника», 2020.
4. Seitnazarov K., Turdışov D., Dosimbetov A. Knowledge base of algorithmic software complex for providing agricultural fields with water resources. //AIP Conference Proceedings. – AIP Publishing, 2024. Т. 3147. № 1.
5. Гектин Ю. М. и др. Температура и методы ее определения с помощью оптико-электронной аппаратуры космических систем ДЗЗ. // Журнал Ракетно-космическое приборостроение и информационные системы. 2024. Т.11. № 2. – С. 11-21.
5. Seitnazarov K.K., Turemuratova B.K., Aytanov A.K. Stages and Methods of Data Collection for Developing an Artificial Intelligence Model for Recognizing Letters of the Karakalpak Sign Language. // IEEE 25th International Conference of Young Professionals in Electron Devices and Materials (EDM). IEEE, 2024. – С. 2530-2534.
6. Murphy Kevin P. Machine learning: a probabilistic perspective. – MIT press, 2012.
7. Kenesbaevich S.K., Muxambetmustapayevich D.A., Arzubaevich N.A. Development of software for calculating the forecast of groundwater regime based on probabilistic and statistical methods. // ACADEMICIA: An International Multidisciplinary Research Journal 10 (7). – P. 526-530.