Мониторинг ночных судовых огней по данным сенсора VIIRS

М. Н. Жижин^{1,2,3}, А. А. Пойда⁴, Е. О. Тютляева⁵, В. В. Коноплев³, К. Элвидж²

¹ Университет Колорадо, Брумфилд, 80020, США E-mail: Mikhail.zhizhin@colorado.edu

 ² Национальный информационный центр по окружающей среде НОАА, Болдер, 80305, США E-mail: chris.elvidge@noaa.gov
³ Институт космических исследований РАН, Москва, 117997, Россия E-mail: v.konoplev@cosmos.ru

> ⁴ Национальный исследовательский центр «Курчатовский институт», Москва, 123098, Россия E-mail: poyda_aa@nrcki.ru ⁵ РСК Технологии, Москва, 125373, Россия E-mail: xgl@rsc-tech.ru

С 70-х гг. прошлого века известно, что на спутниковых изображениях, полученных в ночное время суток, можно выделять яркие электрические бортовые огни на рыболовных судах, но, несмотря на это, до сих пор не было алгоритма автоматического детектирования судов по спутниковым данным без участия оператора. В статье представлена автоматическая система детектирования ночных судовых огней по изображениям мультиспектрального радиометра VIIRS на спутнике Suomi NPP. Алгоритм детектирования выделяет изолированные яркие точки, резко видимые на ночной поверхности моря. В основу алгоритма положен высокочастотный пространственный фильтр на основе медианного и винеровского фильтров. Результаты детектирования уточняются дополнительными фильтрами грозовых молний, паразитной засветки в оптике прибора, локальной резкости, лунных бликов, корреляции видимого и инфракрасного каналов для удаления помех от лунной засветки и облаков. Описана архитектура хранилища исходных данных, программная реализация алгоритма и способы её многоядерной оптимизации для различных архитектур, включая KNL, Haswell, Broadwell. В качестве экологического приложения проведён ВАСІ-анализ последствий массового отравления рыбы у берегов Вьетнама в апреле 2016 г. Разработанная система может быть использована как для оперативного мониторинга морского рыболовства, так и для анализа долгосрочных экологических последствий ограничений на вылов рыбы и морских экологических катастроф техногенного и природного происхождения.

Ключевые слова: ДЗ ночной поверхности Земли, мультиспектральное ДЗЗ, VIIRS, ночные огни, рыболовные судовые огни, высокопроизводительные вычисления, многоядерная архитектура

Одобрена к печати: 21.12.2017 DOI:10.21046/2070-7401-2018-15-1-101-119

Введение

С 70-х гг. прошлого века известно, что при помощи высокочувствительных сенсоров, размещённых на спутниках, можно обнаруживать яркие судовые огни в ночное время суток (Croft, 1978; Croft, Colvocoresses, 1979). В большинстве случаев речь идёт о рыболовных судах, использующих свет для привлечения рыбы. Недавнее исследование (Straka et al., 2015) показало, что использование спутниковых данных позволяет обнаруживать бортовые огни и других судов. Это имеет большое значение для рыболовной промышленности (Cho et al., 1999; Kiyofuji, Saitoh, 2004; Maxwell et al., 2004; Rodhouse et al., 2001; Waluda et al., 2002, 2004; Zhang et al., 2013), но, несмотря на это, до сих пор не было алгоритма автоматического детектирования судов по спутниковым данным без участия оператора.

Первой системой, которая могла быть использована для глобального мониторинга судовых огней, стал сенсор OLS, установленный на метеорологических спутниках DMSP. Национальный центр геофизических данных США (NGDC), будучи главным центром сбора и сервировки данных DMSP, создал ряд сервисов передачи данных OLS с трёхчасовой задержкой рыболовным агентствам Японии, Кореи, Таиланда и Перу.

Запуск спутника Suomi NPP в 2011 г. ознаменовал новую эру в области спутниковой съёмки в ночное время суток (Elvidge et al., 2013; Miller et al., 2012, 2013; Schueler et al., 2013). Основным сенсором на спутнике NPP является VIIRS — мультиспектральная камера видимого и инфракрасных (ИК) диапазонов с шириной охвата 3000 км. Сенсор VIIRS имеет существенные технические преимущества перед OLS: учетверённое разрешение 750 м на пиксель, динамический диапазон, увеличенный с 6 до 14 бит, постоянную радиометрическую калибровку, а также наличие дополнительных спектральных диапазонов, полезных для определения характеристик облаков, океана и источников горения.

Сенсор VIIRS способен обнаружить гораздо больше судов по их бортовым огням по сравнению с DMSP, при этом объём данных, собираемых VIIRS за один пролёт над регионом, превышает объём данных DMSP, собранных за один пролёт над тем же регионом, более чем в 200 раз. Лавинообразное увеличение объёма данных стало основным препятствием к использованию данных VIIRS для обнаружения рыболовных судов промысловыми агентствами и другими организациями. Для решения этой проблемы была разработана автоматическая система детектирования ночных судовых огней по данным VIIRS. В статье описываются ключевые моменты данной системы: алгоритмы, хранилище данных, программная реализация и её оптимизация, экологические приложения.

1. Метод

1.1. Входные данные

В качестве исходных изображений для детектирования судовых огней используются данные мультиспектральной камеры видимого и ИК-диапазонов VIIRS с шириной охвата 3000 км, установленной на метеорологическом спутнике Suomi NPP и позволяющей вести высокочувствительные наблюдения ночной поверхности Земли.

Для детектирования судовых огней используется в основном видимый панхроматический диапазон (DNB) с полосой чувствительности 0,5–0,9 мкм, но для разделения типов источников излучения и для фильтрации облаков дополнительно используются ИК-каналы, в том числе узкополосный канал I04 с центральной длиной волны 3,7 мкм и шириной полосы пропускания 0,38 мкм. Более подробные характеристики камеры VIIRS можно найти в технической документации*.

Камера VIIRS сканирует данные непрерывно, а в центре приёма на Земле полученные данные объединяются в минутные файлы (так называемые гранулы), которые могут объединяться в ещё более ёмкие структуры по четыре гранулы, называемые агрегатами и содержащие данные за 5 мин полёта спутника.

1.2. Алгоритм детектирования ночных судовых огней

Первоначально алгоритм детектирования судовых огней создавался на случай «безлунной» ночи (Elvidge et al., 2015), что составляет половину лунного цикла, когда на изображении в видимом канале DNB отсутствуют значимые помехи от лунной подсветки верхних облаков и бликов на воде (lunar glint). В дальнейшем алгоритм был расширен фильтрами облаков и бликов, что понижает его чувствительность при полной луне, но позволяет использовать на протяжении всего лунного цикла. На *рис. 1* (см. с. 103) представлена блок-схема алгоритм а детектирования судовых огней по мультиспектральным спутниковым данным, получаемым с ночной стороны Земли.

^{*} Техническая документация мультиспектральной камеры VIIRS. https://www.star.nesdis.noaa.gov/ jpss/VIIRS.php.



Рис. 1. Блок-схема алгоритма детектирования судовых огней

Для работы с DNB-изображением осуществляется логарифмическое преобразование гистограммы яркости, чтобы уменьшить разницу между сильными и относительно слабыми источниками света, к которым в том числе относятся и судовые огни, и отделить неяркие источники света от шума. Изменённое таким образом изображение назовём «растянутым» (stretched).

На следующем шаге предварительной обработки данных к растянутому изображению применяется адаптивный фильтр Винера (Lim, 1990), который при низком уровне сигнала подавляет высокие частоты, рассматриваемые как тепловой шум сенсора и алгоритма оцифровки. Этот шум распределяется вдоль строки изображения (скана сенсора) неравномерно: по краям он выше. Параметром фильтра является вариация уровня сигнала на изображении вдоль скана (по строке), представленная в виде полинома 6-й степени. Вариация оценивалась для безлунных ночных изображений водной поверхности в центре Тихого океана.

Для поиска полезного сигнала от судовых огней в виде изолированных ярких точек («спайков», *англ.* spike) на изображении используется высокочастотный пространственный фильтр в виде разности исходного изображения и результата применения к нему медианного фильтра (Gonzales, Woods, 2002) в скользящем 3×3 окне (*puc. 2*). Результат применения этого фильтра называется Spike Median Index (SMI).

В качестве кандидатов в судовые огни выбираются точки морской поверхности, для которых значение SMI превышает рассчитанный порог и которые в то же время являются локальными максимумами яркости в видимом канале. Как уже упоминалось ранее, шум неравномерно распределён по изображению, поэтому и порог также неодинаковый на разных участках скана. На *рис. 3* (см. с. 105) приведена зависимость порога SMI вдоль скана.

Оценка локальной резкости изображений необходима в связи с тем, что в безлунные ночи облака не видны, а существующие методы детектирования облаков по инфракрасным снимкам плохо работают в ночное время в областях, покрытых тонкими или относительно тёплыми облаками и аэрозолями (туманом или дымом).

В качестве меры локальной резкости ночного космоснимка, называемой Sharpness Index (SI), используется метод спектральной резкости (Vu et al., 2012). Метод основан на следующем наблюдении: спектр Фурье в области вокруг точечного источника на ночном изображении со спутника являет собой смесь спектров от гало от облаков и аэрозолей (рассеянного излучения) и прямого (не рассеянного) излучения с учётом частичного поглощения в толще атмосферы. Локально спектр Фурье в области изображения, не подверженной рассеянию и потому более резкой, будет медленнее затухать с ростом пространственной частоты по сравнению со спектром в области, покрытой облаками и поэтому менее резкой.



Рис. 2. Вычисление SMI: DNB-изображение скопления рыболовных судов (слева); результат применения к DNB-изображению медианного фильтра в скользящем окне 3×3 (по центру); разность первого и второго изображений (SMI) с применением отсечки по порогу для устранения шумов (справа). Все фрагменты имеют размер 336×320 пикселей



Рис. 3. Изменение порога детектирования SMI вдоль скана (при удалении от точки надира порог SMI повышается). Сплошной красной линией показан порог SMI, пунктирные линии обозначают границы различных зон пиксельной интеграции в канале DNB сенсора VIIRS. Точками показаны значения SMI для пикселей на поверхности воды за 5 мин полёта спутника над Индонезией 27.09.2014 с 18:34:58 по 18:40:16 UTC



Рис. 4. Пример изменения резкости изображения и SI-индекса. Размер фрагментов изображений составляет 336×320 пикселей

На *рис. 4* показано изображение ночных огней от газовых факелов на морских платформах и рыболовных судах севернее Джакарты, полученное 27.09.2014 в 18:34 UTC в канале DNB сенсора VIIRS, и соответствующая ему карта показателя локальной резкости SI, вычисленного в скользящем окне 32×32 пикселя. На нём видно, что в области изображения, размытой из-за дымовых шлейфов и облачности, спектр быстрее убывает с частотой. Для дальнейшего анализа выбираются точки, значение резкости для которых превышает заданный порог (SI > 0,4). Для выделения изолированных ярких точек вычисляется отношение яркости (Spike Height Index, SHI) между центральным пикселем и средней яркостью его ближайших соседей по строке и столбцу в изображении по формуле:

$$SHI = max(SHI_H, SHI_V),$$

где

$$\begin{split} \mathrm{SHI}_{H} = & \begin{cases} I_{0} - \frac{I_{H,PREV} + I_{H,NEXT}}{2} \\ I_{0} \\ 0, \end{cases}, & \text{если } I_{0} > 0, \\ 0, & \text{в противнном случае;} \end{cases} \\ \mathrm{SHI}_{V} = & \begin{cases} I_{0} - \frac{I_{V,PREV} + I_{V,NEXT}}{2} \\ I_{0} \\ 0, & \text{в противнном случае;} \end{cases}, & \text{если } I_{0} > 0, \\ 0, & \text{в противнном случае;} \end{cases} \end{split}$$

 I_0 — центральный (исследуемый) пиксель; $I_{V, PREV}$, $I_{V, NEXT}$, $I_{H, PREV}$, $I_{H, NEXT}$ — ближайшие соседи исследуемого пикселя по строке и столбцу.

Спайк считается изолированным, если SHI больше заданного порога (0,75).

Ещё одна важная фильтрация: обнаружение молний. Пример DNB-изображения с молнией и результат его обработки представлены на *puc. 5*. На спутниковом изображении молния представляется полосой фиксированной ширины. Съёмка осуществляется путём последовательного сканирования перпендикулярно направлению полёта спутника. Каждый скан включает фиксированное количество линий (например, для сенсора VIIRS их число равно 16). Если во время сканирования сенсор «сталкивается» с молнией, то «засвечиваются» сразу все линии скана. Длина засветки объясняется движением камеры во время сканирования вдоль направления скана: за время свечения, вызванного молнией, сенсор успевает отсканировать некоторый участок. Этот участок существенно короче скана, поэтому когда сенсор снимает следующий скан, то свечения от молнии уже нет.

Для автоматического обнаружения молнии алгоритм последовательно проходит вдоль границ между сканами (используются две линии: по одной от каждого из смежных сканов) и вычисляет относительное превышение значений одной линии над значениями другой в каждой точке (*puc. 6*, см. с. 107) по формуле:

$$LI = \log 10 \left[\frac{sDNB(i) - sDNB(i-1)}{sDNB(i-1)} \right]$$

где sDNB(i) и sDNB(i-1) — значение «растянутого» DNB-изображения в смежных пикселях скана i и i - 1.



Рис. 5. Пример спутникового изображения с молнией (слева) и результат автоматического детектирования молнии (справа)



Рис. 6. Детектирование молний

Если полученные значения не входят в заданный диапазон (т.е. превышение яркости в одном скане над другим значительное), то анализируемая точка и все 16 точек данного скана, сделанные в тот же момент, помечаются (достаточно, чтобы сигнал поступил только с одной стороны скана, так как с другой стороны он может попасть на источник света — в этом случае большой разницы значений не будет). После этого алгоритм проходит вдоль скана и ищет непрерывные помеченные цепочки. Если находятся цепочки длиной, превышающей некоторую заданную величину L, то все пиксели скана, входящие в эту цепочку, а также отстоящие на длину не более L от начала и конца цепочки, не рассматриваются как кандидаты в детекции судовых огней. Длина L подбирается экспериментально, но для суши значение этой переменной следует брать большим, так как на суше больше ночных огней, которые могут случайно оказаться на границе сканов. Мы используем значения 24 и 16 для суши и моря соответственно.

При лунном свете для детектирования огней от рыболовных судов необходима дополнительная фильтрация лунных бликов на поверхности волн и рассеянного отражения от кучевых облаков. При анализе изображения оценивается минимальная и максимальная величина лунной засветки по всем пикселям обрабатываемого в данный момент изображения. Для прогноза интенсивности лунной засветки (обозначим её как LI от *англ.* Lunar Illuminance) в люксах используется модель (Janiczek, De Young, 1987). Если значения LI для всех пикселей не превышают порогового значения (0,0001 лк), то считается, что лунной засветки нет, и обработка не производится.

Если лунная засветка есть, то все пиксели изображения, в зависимости от значения лунной засветки в них, делятся на шесть зон (*puc.* 7, см. с. 108). Зоны определяются делением диапазона значений логарифма LI на всём изображении на шесть равных интервалов. Зональное деление требуется, чтобы получить статистически значимую оценку для пороговых значений алгоритма с изменением засветки. Если в зоне недостаточно пикселей (например, большая часть пикселей относится к суше), то такая зона объединяется с соседней.

Пиксели, в которых есть лунная засветка, должны быть дополнительно проверены на наличие облачности (так как в лунном свете облака становятся существенно ярче, что может приводить к ложным детекциям). Способ одинаков для всех пикселей с лунной засветкой, но в зависимости от зоны, в которую попал пиксель, используются различные пороговые значения.

Для проверки наличия облаков в обрабатываемом пикселе используется корреляция между значениями DNB-канала и инфракрасного канала I04 в некоторой окрестности данного пикселя. Если в рассматриваемом районе присутствует облачность, то корреляция данных видимого и ИК-каналов стремится к значению -1. Это связано с тем, что облака в свете луны становятся ярче окружающего безоблачного фона, а их яркость в инфракрасном канале, наоборот, уменьшается, так как облака холоднее поверхности воды.



Рис. 7. Деление изображения на шесть зон и зависимость порога индекса SMI от зоны

Более детально алгоритм рассмотрен на *рис.* 8, на котором представлен график зависимости значений индекса SMI в точках локального максимума от величины корреляции между DNB и I04-каналом в окрестности этих же точек. На рисунке представлен случай, когда многие точки попали в зону облачности. Это видно по тому, что центр тяжести облака смещён в сторону отрицательных корреляций. Синяя вертикальная черта на графике — это пороговое значение корреляции (ПЗК), проведённое на уровне -0,6 для всех зон. Если пиксель попадает левее этой отсечки (т. е. значение корреляции в его окрестности меньше -0,6), то он не рассматривается как судовые огни.



Рис. 8. График зависимости значений индекса SMI в точках локального максимума от величины корреляции между DNB и 104-каналом в окрестности этих же точек

Горизонтальная линия (на рисунке обозначена как «Порог SMI для детекций судовых огней», далее — SBD) изображает пороговое значение индекса SMI для возможных детекций. Это пороговое значение зависит от зоны лунной засветки: чем выше значение лунной засветки для обрабатываемой зоны, тем выше порог индекса SMI. Все пиксели, находящиеся выше этого порога (и левее корреляционного порога), являются кандидатами в судовые огни.

Наклонный отрезок (на рисунке обозначен как «Слабые детекции», далее — WBD) — это порог менее достоверных детекций. Отрезок проводится от точки пересечения порогов SBD и ПЗК до точки с координатами (1, SBD/4). Пиксели, находящиеся выше WBD, но ниже SBD, являются кандидатами в детекции с большей вероятностью ошибки. Все пиксели, находящихся ниже порога WBD, не рассматриваются в качестве кандидатов в судовые огни.

Кокс (С. S. Cox) и Манк (W. H. Munk) в 1954 г. создали статистическую модель распределения яркости солнечных бликов на взволнованной поверхности моря (Cox, Munk, 1954). Модель Кокса-Манка была перенесена на случай лунного света (другие модели лунной засветки взволнованной поверхности авторам неизвестны). В модели Кокса-Манка распределение яркости лунных бликов на изображении зависит от скорости и направления ветра, которые создают статистическую упорядоченность волн по положению спутника, направлению съёмки и положению луны. Результатом применения модели является распределение вероятности блика в анализируемой области снимка, а ожидаемая яркость бликов определяется как математическое ожидание, умноженное на яркость лунной засветки с экспериментально выбранным масштабным множителем. В зависимости от ожидаемой яркости выставляется порог индекса SMI (для каждой лунной зоны используется свой порог). Все кандидаты, индекс SMI которых ниже порога, отфильтровываются как недостоверные.

На *рис.* 9 продемонстрирован результат описанного выше подхода фильтрации лунных бликов на поверхности волн и на кучевых облаках. На графике представлены две кривые, демонстрирующие зависимость числа детекций судовых огней от фазы луны для двух версий алгоритма: базовой и расширенной. Видно, что в случае базового алгоритма количество детекций (очевидно, что большинство из них ложные) при полной фазе возрастало более чем в 10 раз. Расширенная версия алгоритма, наоборот, показывает снижение количества детекций, что связано с загрублением порогов алгоритма в условиях лунной засветки.



Рис. 9. Зависимость количества детекций от фазы луны для двух версий алгоритма: базовой (синяя линия) и расширенной (красная линия)

1.3. Пространственно-временная база данных

За одну ночь алгоритм выявляет порядка 10 тыс. детекций. Для работы с ежедневными данными это не так много, и результаты детектирования могут быть представлены в текстовом формате, например в CSV. Но когда требуется провести анализ данных за более долгосрочный период, то текстовый формат неудобен. Поэтому для ускорения географических поисковых запросов и статистического анализа была разработана пространственно-временная база данных, использующая пространственное расширение PostGIS реляционной СУБД PostgreSQL.

В базу данных вносятся детекции и их параметры (идентификатор, координаты, время наблюдения, качество детекции, значения индексов и т.п.), границы морских экономических зон, заповедные зоны и т.д.

2. Программная реализация алгоритмов для оперативной обработки данных на многоядерном вычислителе

Представленная на *рис.* 1 блок-схема алгоритма детектирования судовых огней выполнена в виде программного продукта с модульной архитектурой. Большинство модулей реализованы на языке программирования MATLAB. Некоторые модули реализованы на языке IDL, Java, Си. В *табл.* 1 представлен результат профилирования созданной программной реализации алгоритма детектирования судовых огней при обработке данных, полученных за 5 мин пролёта спутника. Профилирование проводилось на рабочей станции, имеющей следующие характеристики:

- процессор: intel Xeon E5640@2,67GHz (2 процессора);
- оперативная память: 8 ГБ;
- разрядность операционной системы: 64 бит.

Логические блоки первого уровня	Логические блоки второго уровня	Логические блоки третьего уровня	Количество вызовов	Суммарное время работы, с	Процентное соот- ношение от време- ни родительского блока
Перепроекция из спутниковой проекции в картографическую			1	457	37%
Детектирование судовых огней			1	763	62%
	Вычисление коэффициен- та размытия изображения		1	355	46%
		Двумерное преоб- разование Фурье	782 073	128	37 %
		Кросс- корреляция спек- тральных каналов	782 073	132	39 %
	Интерполяция данных shape-файлов		1	86	11%

Таблица 1. Результат профилирования программной реализации алгоритма детектирования судовых огней (наиболее времяёмкие функции)

В таблице представлены результаты только по наиболее времяёмким функциям: преобразование Фурье, перепроекция, кросс-корреляция спектральных каналов. Как видно из таблицы, время обработки данных на одной станции превышает время их получения. Для обработки потока мультиспектральных спутниковых данных со скоростью, близкой к реальному времени (без накопления задержки), необходим параллельный вычислительный кластер с большим объёмом оперативной памяти и дискового пространства на каждом узле. Но даже при использовании кластера «стандартной» архитектуры из двенадцати 4-процессорных узлов скорость ретроспективного анализа архива данных Suomi NPP лишь в три раза превышает скорость поступления новых данных. Это означает, что обработка всего архива новыми версиями алгоритмов потребует два года вычислений.

Поэтому в 2016 г. в Институте космических исследований РАН при поддержке Минобрнауки России был начат проект, целью которого является разработка технического задела по созданию специализированной многоядерной вычислительной платформы для большого массива данных (изображений), оптимизированной под массивно-параллельные алгоритмы обработки мультиспектральных космических снимков. В первую очередь упор был сделан на наиболее ресурсоёмкие алгоритмы, выявленные в процессе профилирования (см. *табл. 1*).

Оптимизация включала два шага: реализацию наиболее времяёмких алгоритмов на языке C++ и оптимизацию кода C++ для эффективной работы на мультиядерных процессорах Intel Xeon Phi (KNL) и многоядерных процессорах Intel Xeon E5. В рамках оптимизации C++ кода была проведена векторизация кода и распараллеливание в общей памяти:

- 1. Проведена оптимизация программного кода с целью уменьшения количества условных и безусловных переходов, а также увеличения степени векторизации циклов. Для векторизации кода был использован векторизующий компилятор, анализ отчётов о векторизации и вспомогательные директивы (прагмы).
- 2. Было проведено выравнивание данных большинства массивов для ускорения доступа к ним*.
- 3. Для достижения высокой эффективности для ряда вычислений использовались уже готовые примитивы из библиотеки MKL.
- Для оптимизации обработки данных в «скользящем окне» (которая используется при вычислении коэффициента размытия изображения) было использовано распараллеливание в общей памяти при помощи программного интерфейса OpenMP.
- 5. При компиляции кода была задействована максимальная степень оптимизации.

В *табл. 2–4* приведены результаты тестирования, сравнивающие программную реализацию на языке программирования C++ без оптимизации и с оптимизацией. При этом вариант с оптимизацией разделён на два случая в зависимости от количества OpenMP потоков:

- 1. Количество OpenMP потоков равно количеству физических ядер на узле.
- 2. Количество OpenMP потоков равно количеству виртуальных потоков на узле с учётом hyperthreading.

Тестирование проводилось на программном блоке вычисления коэффициента размытия изображения на тех же данных, на которых проводилось профилирование версии, реализованной на МАТLAB.

Таблица 2. Результаты сравнения оптимизированного и неоптимизированного C++ кода для процессора KNL (Intel(R) Xeon Phi(TM) CPU 7250, 68 ядер)

До оптимизаций, С++ версия	С оптимизациями, векторизацией и OpenMP		
4 мин 28 с	68 ОМР потока	272 ОМР потока	
	15,588 c	22,044 c	

^{*} Описание выравнивания данных, размещённое на сайте компании Intel. https://software.intel. com/en-us/articles/data-alignment-to-assist-vectorization.

Таблица 3. Результаты сравнения оптимизированного и неоптимизированного C++ кода для процессора Haswell (Intel(R) Xeon(R) CPU E5-2697 v3, 14 ядер на процессор, 2 процессора на узле)

До оптимизаций, С++ версия	С оптимизациями, векторизацией и OpenMP		
51,389 c	28 OMP	56 OMP	
	5,003 c	4,418 c	

Таблица 4. Результаты сравнения оптимизированного и неоптимизированного C++ кода для процессора Broadwell (Intel(R) Xeon(R) CPU E5-2698 v4, 20 ядер на процессор, 2 процессора на узле)

До оптимизаций, С++ версия	С оптимизациями, векторизацией и OpenMP		
48,832 c	40 OMP	80 OMP	
	4,563 c	4,310 c	

Результаты тестирования показывают, что проведённая оптимизация позволила радикально уменьшить время работы вычислительного модуля на многоядерных процессорах за счёт эффективного использования всех доступных вычислительных ядер.

На многоядерных процессорах Haswell и Broadwell дополнительный выигрыш даёт технология Hyperthreading*, так что для получения наилучшего результата стоит использовать количество потоков, равное количеству доступных на системе виртуальных потоков.

На процессоре с архитектурой Intel Xeon Phi использование технологии Hyperthreading уже не позволяет получить дополнительное ускорение в связи с тем, что вычислительная задача в каждом потоке имеет достаточную вычислительную сложность, чтобы полностью загрузить «легковесное» ядро. Поэтому для мультиядерных процессоров Intel Xeon Phi наилучшего результата можно добиться при использовании количества потоков, равного количеству физических ядер на процессоре.

3. Программная СХД для хранения и обработки данных проекта

Для успешной реализации проекта требуется система хранения данных (СХД), которая бы сочетала высокую отказоустойчивость, производительность и экономическую эффективность. До недавнего времени решения, сочетающие высокую производительность и отказоустойчивость, были доступны только в сегменте специализированных аппаратных СХД. Однако стоимость хранения терабайта данных в таких системах в несколько раз выше, чем та, которую можно было бы получить на типовом общедоступном оборудовании. В нашем случае, когда объём хранимых данных оценивается сотнями терабайт, использование специализированной аппаратной СХД не представляется экономически оправданным.

В последнее время появилась возможность создавать высокопроизводительные и отказоустойчивые СХД на базе связки свободного программного обеспечения и недорогих широко распространённых аппаратных компонент. Одним из ключевых факторов здесь стало быстрое развитие технологий твердотельных дисковых накопителей. Это позволило отказаться от энергонезависимой памяти с каналом синхронизации между контроллерами СХД и вынести «КЭШ» непосредственно на твердотельные накопители в дисковые корзины. На *рис. 10* (см. с. 113) в общих чертах показан механизм трансформации от специализированной аппаратной СХД к программной.

^{*} Описание технологии Hyperthreading: https://www.intel.ru/content/www/ru/ru/architecture-and-technology/hyper-threading/hyper-threading-technology.html.



Рис. 10. Трансформация архитектуры СХД

Архитектура программного хранилища, как правило, предполагает два уровня представления данных — оконечное хранилище (backend) и уровень клиентского доступа (frontend). Система хранения данных проекта, реализованная на базе архитектуры программно-аппаратного комплекса SCARP (Коноплев и др., 2015), допускает вариации, показанные в *табл. 5*.

Таблица 5. Уровни представления данных в SCARP

Вариант	Frontend	Backend
1	Lustre/RDMA	ZFS
2	NFS/RDMA	ZFS



Рис. 11. Архитектура программно-аппаратного комплекса

В качестве оконечного хранилища выбрана файловая система ZFS. На данный момент это наиболее функциональная свободная файловая система, готовая для штатной эксплуатации. Поддержка кэширования на дисках SSD позволяет создавать недорогие отказоустойчивые решения, поддерживающие КЭШ на запись. Архитектура «копирование при записи», изначально заложенная в ZFS, даёт возможность создавать снимки без ущерба производительности. Это позволяет получить дополнительную защиту пользовательских данных.

Параллельная файловая система Lustre позволяет автоматически распределять данные и балансировать нагрузки по множеству оконечных файловых хранилищ. К сожалению, начиная с 2014 г. стабильные сборки стали доступны только в коммерческих версиях этого продукта, в связи с чем использование Lustre перестало быть экономически оправданным для больших научных хранилищ.

Текущая реализация программного хранилища ориентирована на использование файловой системы NFS поверх RDMA на уровне представления. Общая архитектура программноаппаратного вычислительного комплекса показана на *рис. 11* (см. с. 113). Пользовательские приложения могут работать как непосредственно на физических серверах, так и внутри контейнеров (лёгкая виртуализация). В последнем случае становится возможной быстрая миграция пользовательского «рабочего окружения» между вычислительными узлами в случае отказа оборудования или при необходимости балансировки нагрузки. Лёгкая виртуализация также повышает уровень безопасности, поскольку пользователи не имеют прямого сетевого доступа к серверам хранения. Ветки файловых систем программного хранилища реэкспортируются с физических вычислительных узлов внутрь контейнеров посредством *связанного монтичования* (bind mount). Такой подход даёт хороший уровень изоляции пользователей и практически не требует вычислительных затрат. В том случае, когда использование виртуализации для запуска пользовательских задач оказывается затруднительным (например, при использовании спецвычислителей Xeon Phi), остаётся возможность запуска задач непосредственно на физических серверах вычислительного кластера.

4. Результаты. BACI-анализ динамики рыбного лова при экологическом воздействии

Помимо оперативного применения детектора судовых огней для предупреждения несанкционированного лова в зонах ограничения и запрета рыболовства (за один проход спутника), пространственно-временная база данных всех детекций судовых огней за продолжительный период времени (с апреля 2012 г. по настоящее время) позволяет анализировать долгосрочные тренды и пространственное распределение вероятного лова рыбы.

Метод BACI (Before – After Control – Impact, что переводится как «до – после; контроль – влияние») представляет собой двухфакторный дисперсионный анализ (two-factor analysis of variance, ANOVA), который позволяет установить одновременное влияние на экосистему двух факторов — места и времени, а также взаимодействие между этими факторами (Stewart-Oaten 2001). Для этого наблюдения за системой ведутся до и после внешнего воздействия одновременно в двух местах, про одно из которых известно, что оно было подвержено влиянию (пострадавшая зона, ПЗ), а другое — нет (контрольная зона, КЗ). Задача BACI состоит в оценке статистической значимости изменения средней разницы (разности или отношения) наблюдаемого параметра между ПЗ и КЗ в периоды до и после внешнего воздействия, дата которого известна.

Примером такого экологического воздействия на морскую экосистему может служить несанкционированный выброс токсических отходов в апреле 2016 г. у берегов Вьетнама с металлургического комбината Formosa, принадлежащего тайваньской На Tinh Steel Corporation, вызвавший гибель более 100 т рыбы за месяц, что вынудило вьетнамские власти 4 мая того же года ввести санитарный запрет на рыбную ловлю в 20-мильной прибрежной зоне у четырёх провинций, показанных коричневым цветом на карте на *рис. 12* (см. с. 115), и потребовать у компании компенсации за экологический ущерб в 500 млн дол. (Paddock, 2016).



Рис. 12. Зона санитарного запрета на рыболовство у берегов Вьетнама, введённого 4 мая 2016 г.



Рис. 13. Временные ряды разностей детекций для северной контрольной зоны и каждой из пострадавших зон за период с марта 2013 г. по декабрь 2016 г. Красной вертикальной чертой отмечена дата экологической катастрофы

Входными данными для BACI-анализа последствий этой экологической катастрофы являются суммарные количества детекций судовых огней в каждой зоне за календарный месяц в период с апреля 2012 г. по декабрь 2016 г. В качестве контрольных используются прибрежные зоны соседних провинций на севере и юге от пострадавших зон. Метод сводится к проверке с помощью t-статистики Стьюдента нулевой гипотезы о том, что два набора месячных разностей числа детекций в одной контрольной (например, северной) зоне и одной из четырёх пострадавших зон, полученные до (48 месяцев до апреля 2016 г.) и после катастрофы (9 месяцев, апрель – декабрь 2016 г.), имеют нормальное распределение с одинаковыми средними значениями. Предполагается, что средние квадратические отклонения для наборов разностей до и после катастрофы не известны и равны между собой. Для оценки статистической значимости этой гипотезы можно воспользоваться функцией ttest2 из пакета MATLAB. Графики временных рядов разностей детекций для северной контрольной и каждой из пострадавших зон представлены на *рис. 13* (см. с. 115).

нулевой гипотезы ВАСІ-анализа					
	ПЗ 1, Ha Tinh	ПЗ 2, Quang Binh	ПЗ 3, Quang Tri	ПЗ 4, Thua Thien	
КЗ 1, Север	0,0001	0,0002	0,0005	4,00E-05	
КЗ 2, Юг	0,9747	0,8174	0,983	0,6515	

Таблица 6. Уровни статистической значимости нулевой гипотезы ВАСІ-анализа

Результаты проверки нулевой гипотезы представлены в *табл. 6.* Из неё следует, что с вероятностью 0,999 средние разности числа детекций между любой из ПЗ и северной КЗ до и после катастрофы не равны (первая строка таблицы), т.е. снижение числа детекций во всех ПЗ по сравнению с северной КЗ статистически значимо на протяжении всего периода после катастрофы. Напротив, для южной КЗ статистически значимой разницы в числе детекций до и после катастрофы не наблюдается (вторая строка таблицы). Это можно объяснить не только чувствительностью метода ВАСІ к выбору контрольной зоны, но и большой протяжённостью зоны экологической катастрофы в южном направлении.

Заключение

В статье представлен алгоритм детектирования ночных судовых огней по мультиспектральным спутниковым данным, получаемым с сенсора VIIRS. Алгоритм детектирования выделяет изолированные яркие точки, резко видимые на ночной поверхности моря. При свете луны учитываются помехи от облаков и лунных бликов.

Так как однопотоковая реализация алгоритма достаточно ресурсоёмка, в статье описаны методы её оптимизации и ускорения путём использования современных технологий и подходов к организации параллельных распределённых вычислений на многоядерных процессорах. Также в статье показана возможность многократного ускорения работы системы за счёт организации архива входных/выходных данных и оптимизации хранилища данных, в котором этот архив размещается.

Разработанная система может быть использована как для оперативного мониторинга, так и для анализа долгосрочных последствий морских экологических катастроф техногенного и природного происхождения. Это подтверждается примером анализа последствий массовой гибели рыбы и последующего запрета на вылов во Вьетнаме методом Before – After Control – Impact, который приведён в статье.

Представленный алгоритм может быть улучшен по нескольким направлениям. Вопервых, требуется обеспечить работоспособность алгоритма в условиях слабой облачности или плохой погоды (например, тумана), приводящих не к полной потере сигнала от источника, а к его ослаблению и размытию. Во-вторых, требуется провести калибровку алгоритма на источниках с заранее известными характеристиками. Это позволит не только детектировать местоположение судов, но и оценивать мощность их бортовых огней.

В качестве дальнейшей работы интерес также представляет сравнение результатов детектирования с данными систем мониторинга судов VMS/AIS.

Работа выполнялась при поддержке Минобрнауки России (Государственный контракт 14.607.21.0165, уникальный идентификатор ПНИЭР RFMEFI60716X0165).

Литература

- Коноплев В. В., Боярский М. Н., Назиров Р. Р. SCARP: масштабируемая и отказоустойчивая платформа легкой виртуализации для интенсивной работы с большими данными и облачных приложений // Вычислительные технологии в естественных науках: Методы суперкомпьютерного моделирования: сб. тр. Ч. 3. ИКИ РАН, 2015. С. 132–139.
- 2. *Cho K., Ito R., Shimoda H., Sakata T.* Fishing fleet lights and sea surface temperature distribution observed by DMSP/OLS sensor // Intern. J. Remote Sensing. 1999. V. 20. P. 3–9.
- 3. *Cox C., Munk W.* Measurement of the Roughness of the Sea Surface from Photographs of the Sun's Glitter // J. Optical Society of America. 1954. V. 44. Iss. 11. P. 838–850.
- 4. Croft T.A. Nighttime Images of the Earth from Space // Scientific American. 1978. V. 239. P 86–98.
- 5. *Croft T.A., Colvocoresses A. P.* The Brightness of Lights on Earth at Night, Digitally Recorded by DMSP Satellite: USGS Open File Report. Reston, VA: U.S. Geological Survey, 1979. 57 p.
- 6. *Elvidge C. D., Baugh K. E., Zhizhin M., Hsu F. C.* Why VIIRS data are superior to DMSP for mapping night-time lights // Proc. Asia-Pacific Advanced Network. 2013. V. 35. P. 62–69.
- 7. *Elvidge C. D., Baugh K. E., Zhizhin M., Hsu F. C.* Automatic Boat Identification System for VIIRS Low Light Imaging Data // Remote Sensing. 2015. V. 7. No. 3. P. 3020–3036. DOI:10.3390/rs70303020.
- 8. *Gonzales R. C., Woods R. E.* Digital Image Processing. 2nd ed. Englewood Cliffs. New Jersey: Prentice-Hall, 2002. 122 p.
- 9. *Janiczek P. M., De Young J. A.* Computer programs for sun and moon illuminance with contingent tables and diagrams: U.S. Naval Observatory Circulars. 1987.
- Kiyofuji H., Saitoh S. Use of nighttime visible images to detect Japanese common squid Todarodes pacificus fishing areas and potential migration routes in the Sea of Japan // Marine Ecology Progress Series. 2004. V. 276. P. 173–186.
- 11. *Lim J. S.* Two-Dimensional Signal and Image Processing. Englewood Cliffs. New Jersey: Prentice Hall, 1990. 548 p.
- 12. Maxwell M. R., Henry A., Elvidge C. D., Safran J., Hobson V. R., Nelson I., Tuttle B. T., Dietz J. B., Hunter J. R. Fishery Dynamics of the California market squid (Loligo opalescens), as measured by satellite remote sensing // Fishery Bulletin. 2004. V. 102. P. 661–670.
- Miller S. D., Mills S. P., Elvidge C. D., Lindsey D. T., Lee T. F., Hawkins J. D. Suomi satellite brings to light a unique frontier of environmental imaging capabilities // Proc. National Academy of Sciences USA. 2012. V. 109. P. 15 706–15 711.
- 14. Miller S. D., Straka W. III, Mills S. P., Elvidge C. D., Lee T. F., Solbrig J., Walther A., Heidinger A. K., Weiss S. C. Illuminating the Capabilities of the Suomi National Polar-Orbiting Partnership (NPP) Visible Infrared Imaging Radiometer Suite (VIIRS) Day/Night Band // Remote Sensing. 2013. V. 5. P. 6717–6766.
- 15. *Paddock R. C.* Taiwan-Owned Steel Factory Caused Toxic Spill, Vietnam Says // New York Times. June 30, 2016.
- Rodhouse P. G., Elvidge C. D., Trathan P. N. Remote sensing of the global light-fishing fleet: An analysis of interactions with oceanography, other fisheries and predators // Advances in Marine Biology. 2001. V. 39. P. 261–303.
- 17. Schueler C. F., Lee T. F., Miller S. D. VIIRS constant spatial-resolution advantages // Intern. J. Remote Sensing. 2013. V. 34. P. 5761–5777.
- Stewart-Oaten A., Bence J. R. Temporal and Spatial Variation in Environmental Impact Assessment // Ecological Monographs. 2001. V. 71. No. 2. P. 305–339.
- Straka W. C. III, Seaman C. J., Baugh K., Cole K., Stevens E., Miller S. D. Utilization of the Suomi National Polar-Orbiting Partnership (NPP) Visible Infrared Imaging Radiometer Suite (VIIRS) day/night band for Arctic ship tracking and fisheries management // Remote Sensing. 2015. V. 7. P. 971–989.

- 20. *Vu C. T., Phan T. D., Chandler D. M.* S3: A spectral and spatial measure of local perceived sharpness in natural images // IEEE Trans. Image Process. 2012. V. 21. P. 934–945.
- 21. Waluda C. M., Trathan P. N., Elvidge C. D., Hobson V. R., Rodhouse P. G. Throwing light on straddling stocks of Ilex argentinus: Assessing fishing intensity with satellite imagery // Canadian J. Fisheries and Aquatic Sciences. 2002. V. 59. P. 592–596.
- Waluda C. M., Yamashiro C., Elvidge C. D., Hobson V. R., Rodhouse P. G. Quantifying light-fishing for Dosidicus gigas in the eastern Pacific using satellite remote sensing // Remote Sensing Environment. 2004. V. 91. P. 129–133.
- 23. Zhang X., Saitoh S.-I., Hirawake T., Nakada S., Koyamada K., Awaji T., Ishikawa Y., Igarashi H. An attempt of dissemination of potential fishing zones prediction map of Japanese common squid in the coastal water, southwestern Hokkaido, Japan // Proc. Asia-Pacific Advanced Network. 2013. V. 36. P. 132–141.

Monitoring of night fishing boat lights with VIIRS

M. N. Zhizhin^{1,2,3}, A. A. Poyda⁴, E. O. Tyutlyaeva⁵, V. V. Konoplev³, C. D. Elvidge²

 ¹ University of Colorado, Broomfield 80020, USA E-mail: mikhail.zhizhin@colorado.edu
² NOAA National Centers for Environmental Information, Boulder 80305, USA E-mail: chris.elvidge@noaa.gov
³ Space Research Institute RAS, Moscow 117997, Russia E-mail: v.konoplev@cosmos.ru
⁴ National Research Center "Kurchatov Institute", Moscow 123098, Russia E-mail: poyda_aa@nrcki.ru
⁵ RSC Technologies, Moscow 125373, Russia E-mail: xgl@rsc-tech.ru

Since the 70s of the last century, it was known that the satellite images obtained at night time can be used to detect bright electric lights from fishing boats. However, there was still no algorithm for automatic detection of ships using satellite data without operator participation. The paper presents an automatic system for detecting fishing boat lights at nighttime using images from the VIIRS multispectral radiometer on the Suomi NPP satellite. The new algorithm detects isolated bright spots that are sharply visible on the night sea's surface. The algorithm is based on a high-frequency spatial filter based on the median and Wiener filters. The detection results are supported by additional filters of thunderstorm lightning, straylight in the instrument optics, local sharpness, lunar glint, local correlation in visible and infrared channels to remove interference from moonlight and clouds. The architecture of the data source storage, software implementation of the algorithm and methods of its multi-core optimization for various architectures are described, including KNL, Haswell, Broadwell. As an environmental application, a BACI analysis of the effects of mass fish poisoning off the coast of Vietnam in April 2016 was conducted. The boat detection system can be used both for operational monitoring of marine fisheries and for analyzing the long-term environmental consequences of fishing bans and marine environmental disasters of man-made and natural origin.

Keywords: nighttime remote sensing, multispectral remote sensing, VIIRS, fishing boat lights

Accepted: 21.12.2017 DOI:10.21046/2070-7401-2018-15-1-101-119

Reference

1. Konoplev V.V., Boyarskiy M.N., Nazirov R.R., SCARP: masshtabiruemaya i otkazoustoichivaya platforma legkoi virtualizatsii dlya intensivnoi raboty s bol'shimi dannymi i oblachnykh prilozhenii (SCARP: scalable and fault tolerant light-weight virtualization platform for data intensive and cloud applications), Vychislitel'nye tekhnologii v estestvennykh naukakh: Metody superkomp'yuternogo modelirovaniya, chast' 3, Moscow: IKI RAN, 2015, pp. 132–139.

- 2. Cho K., Ito R., Shimoda H., Sakata T., Fishing fleet lights and sea surface temperature distribution observed by DMSP/OLS sensor, *Int. J. Remote Sens.*, 1999, Vol. 20, pp. 3–9.
- 3. Cox C., Munk W., Measurement of the Roughness of the Sea Surface from Photographs of the Sun's Glitter, *J. Opt. Soc. Am.*, 1954, Vol. 44, Issue 11, pp. 838–850.
- 4. Croft T.A., Nighttime Images of the Earth from Space., Sci. Am., 1978, Vol. 239, pp. 86–98.
- 5. Croft T.A., Colvocoresses A.P., *The Brightness of Lights on Earth at Night, Digitally Recorded by DMSP Satellite*, USGS Open File Report, Reston, VA: U.S. Geological Survey, 1979, 57 p.
- 6. Elvidge C. D., Baugh K. E., Zhizhin M., Hsu F. C., Why VIIRS data are superior to DMSP for mapping nighttime lights, *Proc. Asia-Pac. Adv. Netw.*, 2013, Vol. 35, pp. 62–69.
- Elvidge C. D., Baugh K. E., Zhizhin M., Hsu F. C., Automatic Boat Identification System for VIIRS Low Light Imaging Data, *Remote Sens.*, 2015, Vol. 7, No. 3, pp. 3020–3036, DOI:10.3390/rs70303020.
- 8. Gonzales R. C., Woods R. E., Digital Image Processing, 2nd ed., Englewood Cliffs, NJ: Prentice-Hall, 2002.
- 9. Janiczek P. M., De Young J. A., Computer programs for sun and moon illuminance with contingent tables and diagrams, U.S. Naval Observatory Circulars, 1987.
- Kiyofuji H., Saitoh S., Use of nighttime visible images to detect Japanese common squid Todarodes pacificus fishing areas and potential migration routes in the Sea of Japan, *Mar. Ecol. Prog. Ser.*, 2004, Vol. 276, pp. 173–186.
- 11. Lim J. S., Two-Dimensional Signal and Image Processing, Englewood Cliffs, NJ: Prentice Hall, 1990, 548 p.
- 12. Maxwell M. R., Henry A., Elvidge C. D., Safran J., Hobson V. R., Nelson I., Tuttle B. T., Dietz J. B., Hunter J. R., Fishery Dynamics of the California market squid (Loligo opalescens), as measured by satellite remote sensing, *Fish. Bull.*, 2004, Vol. 102, pp. 661–670.
- Miller S. D., Mills S. P., Elvidge C. D., Lindsey D. T., Lee T. F., Hawkins J. D., Suomi satellite brings to light a unique frontier of environmental imaging capabilities, *Proc. Natl. Acad. Sci. USA*, 2012, Vol. 109, pp. 15706–15711.
- Miller S. D., Straka W. III, Mills S. P., Elvidge C. D., Lee T. F., Solbrig J., Walther A., Heidinger A. K., Weiss S. C., Illuminating the Capabilities of the Suomi National Polar-Orbiting Partnership (NPP) Visible Infrared Imaging Radiometer Suite (VIIRS) Day/Night Band, *Remote Sens.*, 2013, Vol. 5, pp. 6717–6766.
- 15. Paddock R.C., Taiwan-Owned Steel Factory Caused Toxic Spill, Vietnam Says, *New York Times*, June 30, 2016.
- Rodhouse P. G., Elvidge C. D., Trathan P. N., Remote sensing of the global light-fishing fleet: An analysis of interactions with oceanography, other fisheries and predators, *Adv. Mar. Biol.*, 2001, Vol. 39, pp. 261–303.
- 17. Schueler C. F., Lee T. F., Miller S. D., VIIRS constant spatial-resolution advantages, *Int. J. Remote Sens.*, 2013, Vol. 34, pp. 5761–5777.
- 18. Stewart-Oaten A., Bence J. R., Temporal and Spatial Variation in Environmental Impact Assessment, *Ecological Monographs*, 2001, Vol. 71, pp. 305–339.
- 19. Straka W.C. III, Seaman C.J., Baugh K., Cole K., Stevens E., Miller S.D., Utilization of the Suomi National Polar-Orbiting Partnership (NPP) Visible Infrared Imaging Radiometer Suite (VIIRS) day/night band for Arctic ship tracking and fisheries management, *Remote Sens.*, 2015, Vol. 7, pp. 971–989.
- 20. Vu C. T., Phan T. D., Chandler D. M., S3: A spectral and spatial measure of local perceived sharpness in natural images, *IEEE Trans. Image Process*, 2012, Vol. 21, pp. 934–945.
- Waluda C. M., Trathan P. N., Elvidge C. D., Hobson V. R., Rodhouse P. G., Throwing light on straddling stocks of Ilex argentinus: Assessing fishing intensity with satellite imagery, *Can. J. Fish. Aquat. Sci.*, 2002, Vol. 59, pp. 592–596.
- Waluda C. M., Yamashiro C., Elvidge C. D., Hobson V. R., Rodhouse P. G., Quantifying light-fishing for Dosidicus gigas in the eastern Pacific using satellite remote sensing, *Remote Sens. Environ.*, 2004, Vol. 91, pp. 129–133.
- 23. Zhang X., Saitoh S.-I., Hirawake T., Nakada S., Koyamada K., Awaji T., Ishikawa Y., Igarashi H., An attempt of dissemination of potential fishing zones prediction map of Japanese common squid in the coastal water, southwestern Hokkaido, Japan, *Proc. Asia-Pac. Adv. Netw.*, 2013, Vol. 36, pp. 132–141.