

Процедура отбора алгоритма идентификации в реальном времени облачных образований из космоса.

Втюрин С.А.¹, Князев Н.А. ¹

¹Федеральное государственное бюджетное учреждение науки Институт космических исследований Российской академии наук (ИКИ РАН), г. Москва

Восемнадцатая Всероссийская Открытая конференция «СОВРЕМЕННЫЕ ПРОБЛЕМЫ
ДИСТАНЦИОННОГО ЗОНДИРОВАНИЯ ЗЕМЛИ ИЗ КОСМОСА (Физические основы, методы и
технологии мониторинга окружающей среды, потенциально опасных явлений и объектов)»

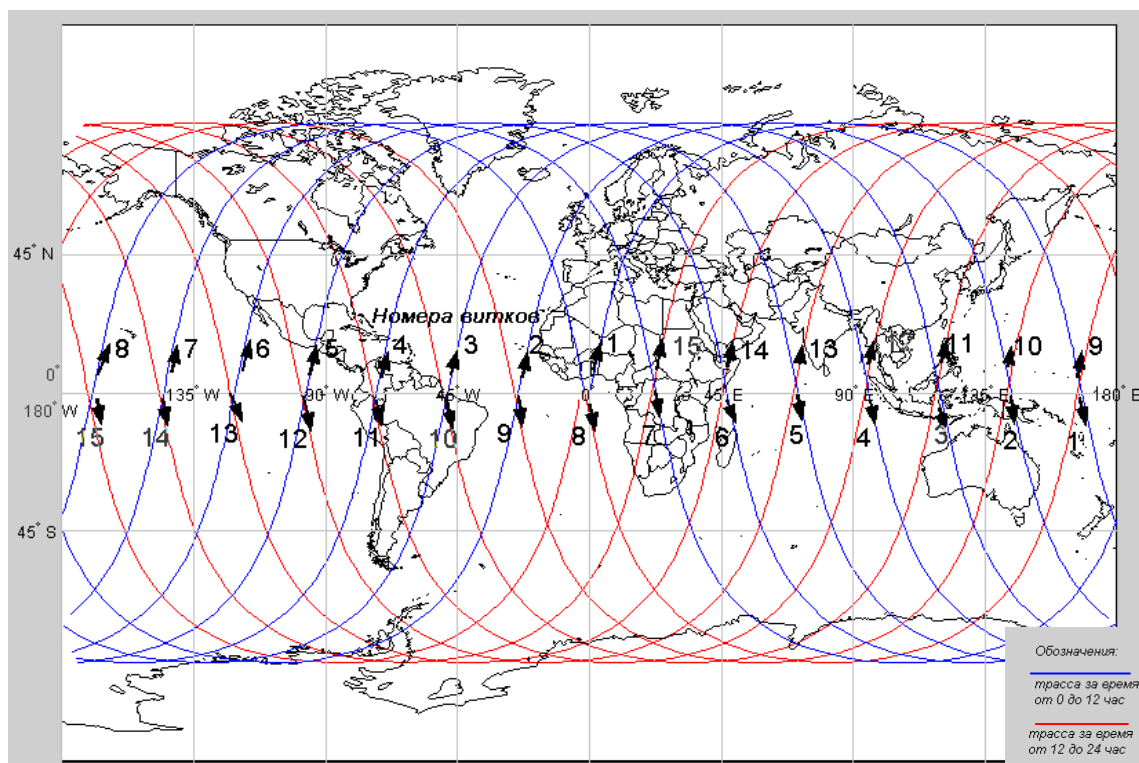
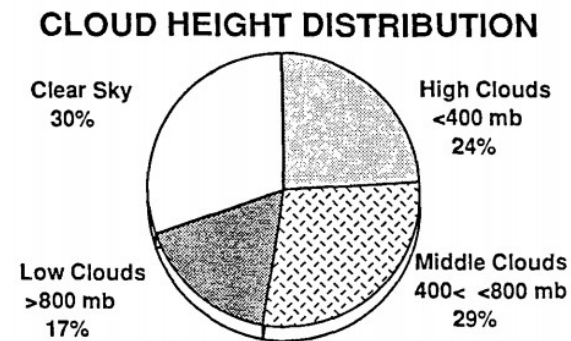
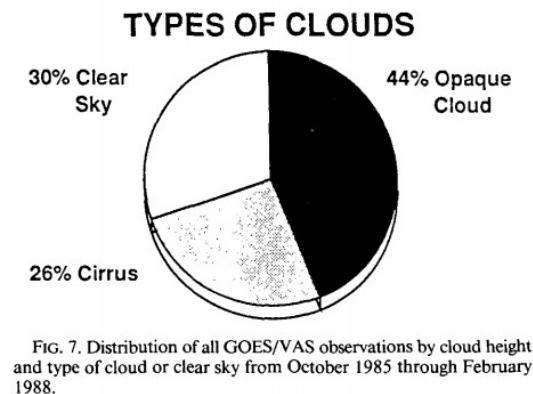
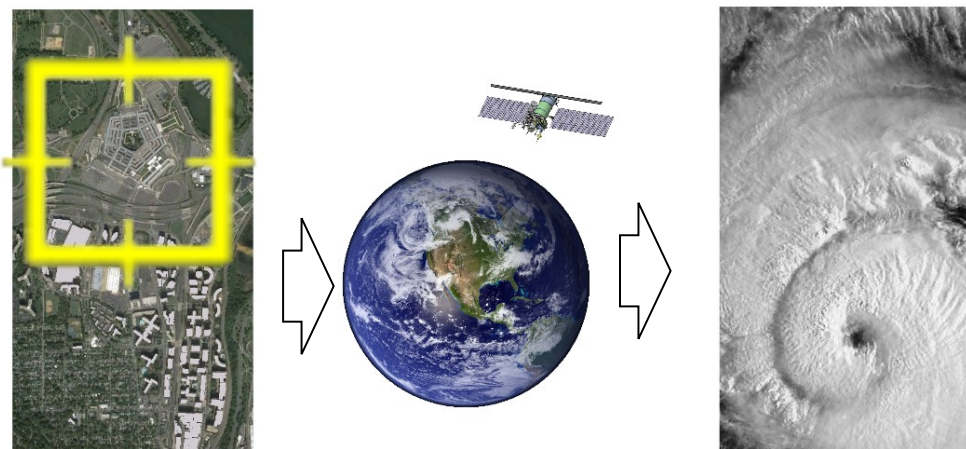
секция Методы и алгоритмы обработки спутниковых данных

Москва, Институт космических исследований РАН, 2020

Решаемая проблема

Облачность является серьезным помеховым фактором при проведении спутниковой съёмки в оптическом диапазоне. Современные методы борьбы с помехой, не включающие смену диапазона наблюдения, в основном заключаются в организационных мерах по планированию съёмки с учётом прогноза наличия облачности или последующей сшивки безоблачных участков. Предлагается метод оперативного учёта облачной обстановки в реальном времени за счёт дополнительной камеры и специального программного обеспечения, обеспечивающие автоматическое управление местом и моментом съёмки. Для такого метода требуются достаточно достоверные, но при этом быстрые и способные работать с ограниченным набором данных алгоритмы. Выбору подходящих методов и алгоритмов и посвящён данный доклад.

Задача и актуальность: облачность как помеха при высокодетальном наблюдении в оптическом диапазоне

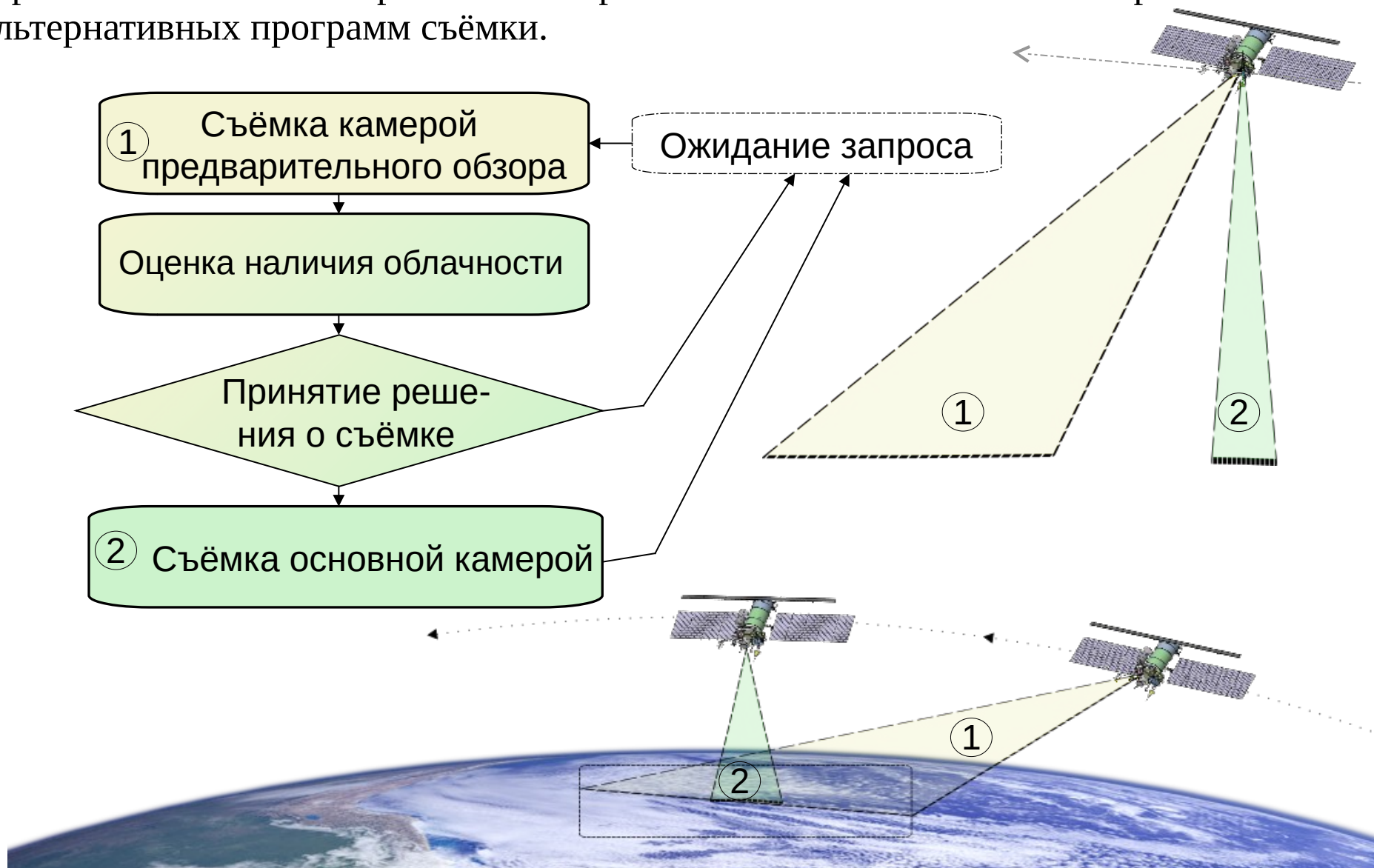


Суточная трасса КА на круговой квазисинхронной орбите с параметрами: Число витков в сутках 15-1/8, период обращения КА – 1,6 час высота орбиты 578 км, наклонение плоскости орбиты 72 градуса

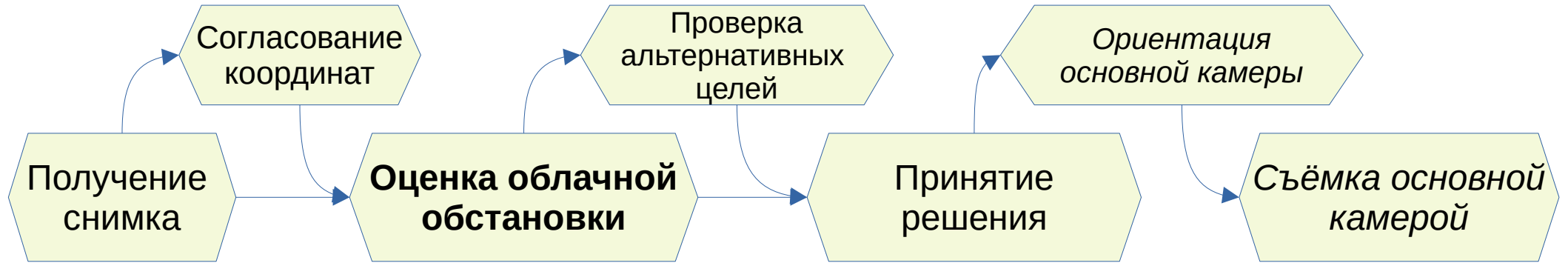
Облачность, как правило, является помехой для съёмок в оптическом диапазоне. При этом, например, согласно (Чернокульский А.В., Мохов И.И.), по данным спутниковых и наземных наблюдений доля покрытия облаками Земли составляет около двух третей, достигая по отдельным данным 0.75, при этом над сушей эта доля находится в диапазоне от 0.49 до 0.58 (по отдельным данным до 0.69) что в целом означает вероятность подобного перекрытия достаточно значительна. Сходные данные приводятся у Wylie D.P., Menzel W. ** где приводится диаграмма распределения облачности различной высоты и типа, процитированная на этом слайде (период наблюдений окт.1985-фев.1988).

Основная идея решения: алгоритм съёмки

В качестве решения предлагается система, включающая камеру предварительного обзора (КПО), установленную под углом в направлении движения КА, на основе анализа данных с которой средствами бортовой ЭВМ возможно построение актуальной карты облачности в районе интереса и автоматический выбор из нескольких альтернативных программ съёмки.



Требования к алгоритму



← Время от съёмки КПО до полёта к месту цели →

Необходимость выполнения всех операций по оценке облачности и принятия решения между снимком КПО и основным прибором.

Возможные технические характеристики камер предварительного обзора (КПО) **

Многоэлементный фотоприемник (МФП)	InGaAs
Формат МФП	2048×1
Размеры пикселя МФП, мкм	12,5×12,5
Спектральный диапазон, мкм	0,9÷1,7
Размер изображения на МФП, мм	25,6
Фокусное расстояние объектива КПО, мм	51,2
Энергопотребление КПО, Вт	~ 6
Размеры КПО, мм	~ 80×80×90
Масса КПО, кг	~ 0,9
Размеры пикселя объекта, км	0,25×0,25
Размер изображения объекта, км	~ 500×0,25
Высота орбиты КА, км	~ 600

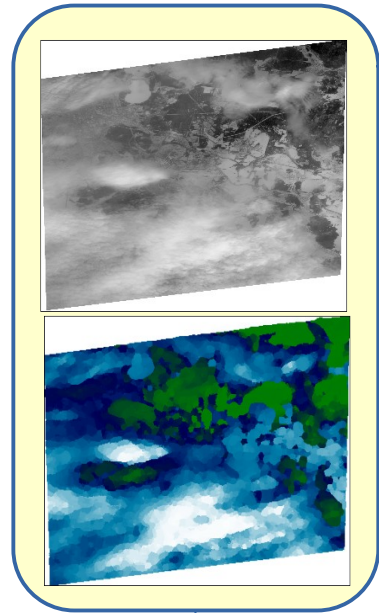
**Конфигурация камеры предложена: Бойко В.А. НПП "Геофизика-космос"; (для случая стереоскопической схемы – используются 2 камеры)

Автоматическая и автономная работа

Низкая требовательность к характеристикам и подробности входных данных

Быстродействие, обеспечивающее обработку за полётное время

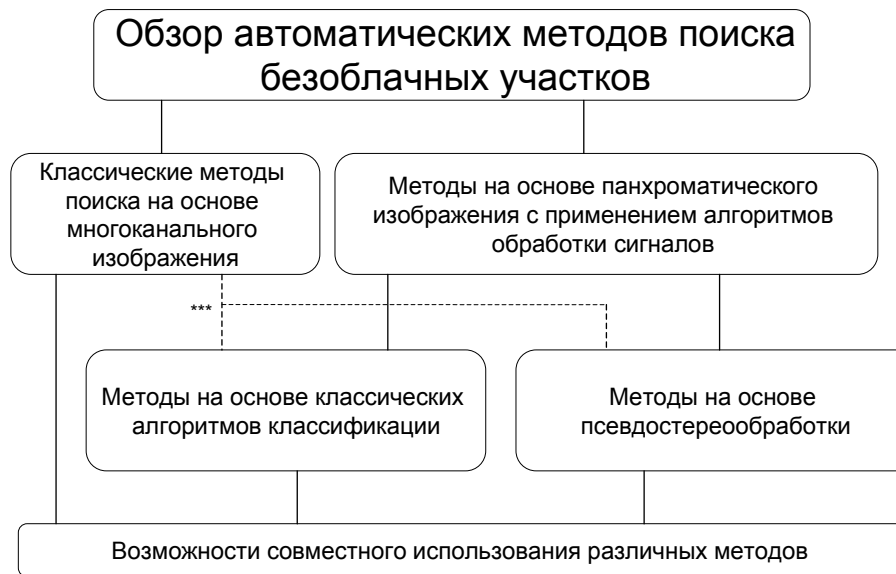
Обеспечение достаточно надёжного и достоверного распознавания



Автоматические методы поиска безоблачных участков

Многоканальное изображение

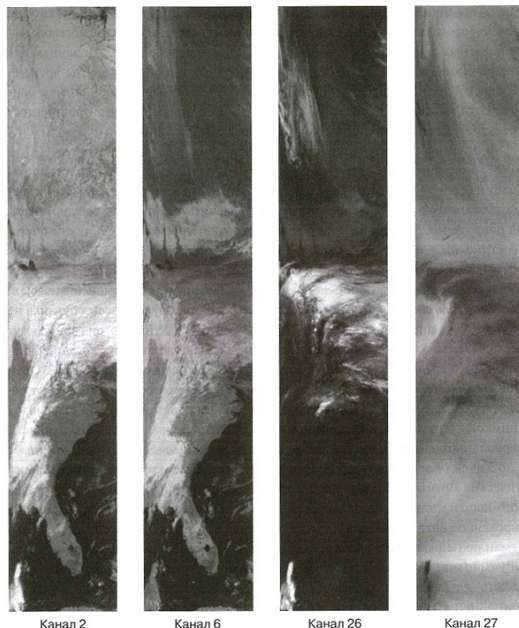
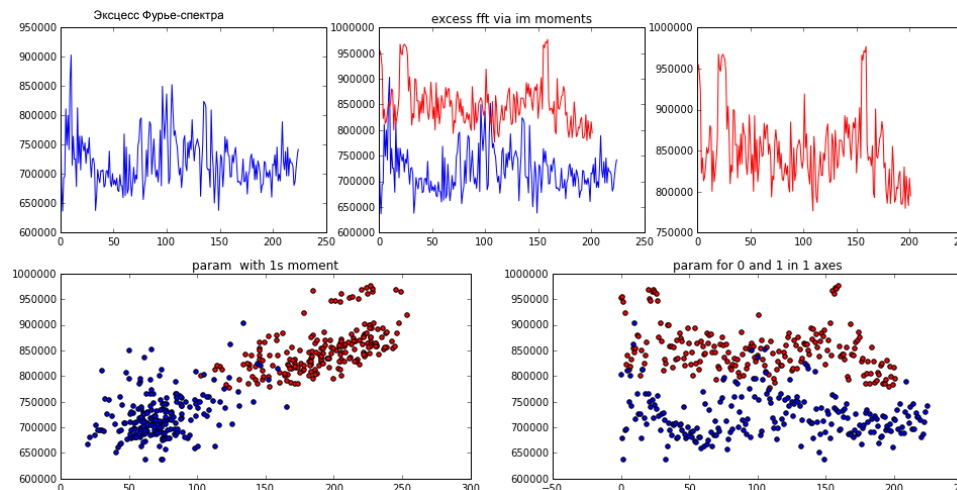
В качестве примера в (Шовенгердт Р.А.) приводится различие облачности от снежно-ледового покрова на основе многоканальных данных аппаратуры MODIS при помощи канала 6 т.к. в этом диапазоне снег и лёд имеют более низкую чем облака отражательную способность. Кроме того на приведённых снимках, легко заметить что в задаче селекции полезен и канал № 26, выделяющий перьевые облака.



*** Могут использовать многоканальные данные как дополнительную информацию или стереобазу.

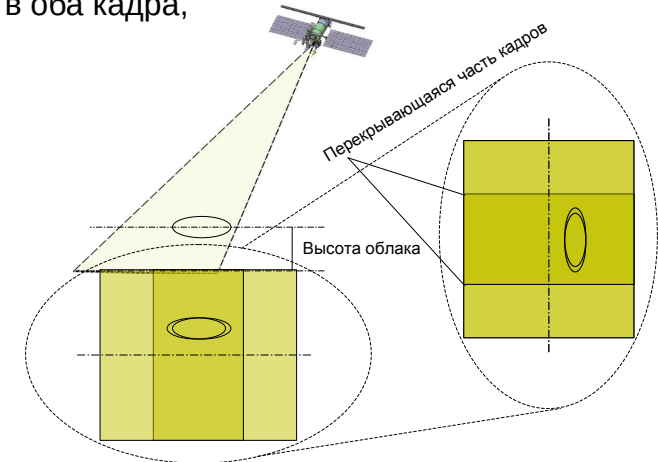
Классификация на основе текстурных признаков

Обзор параметров возможных для использования в пространстве признаков показывает что вполне возможно достичь линейной разделимости или свести к ней.



Псевдостереообработка

Идея метода заключается в том, что если камера предварительного обзора снимает два кадра так, что область интереса входит в оба кадра,



снятых на разных участках маршрута, то за счёт появления стереобазы можно получить информацию о наличии и высоте облачности.

Так в работе (Королев Е.Е. и др) это используется для дополнительной селекции предварительно найденных облачных объектов.

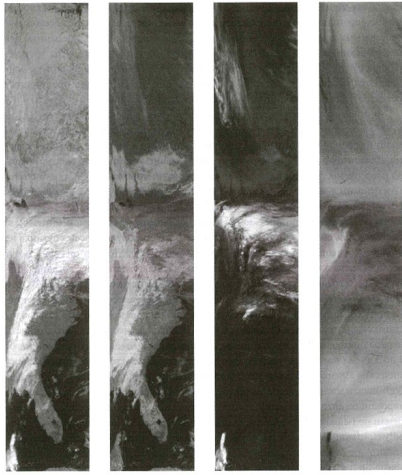
Там же, со ссылкой на (Пошехонов В.И.), приводится пример использования

алгоритма стереообработки, в соответствии с которым определяются трёхмерные координаты каждого пикселя.

В данной работе предлагается иной подход - на основе современных алгоритмов анализа движения

Классические методы на основе многоканального изображения

на примере Landsat 7 ETM+ ACCA algorithm



© рисунок из [1] (Шовенгердт Р.А.)

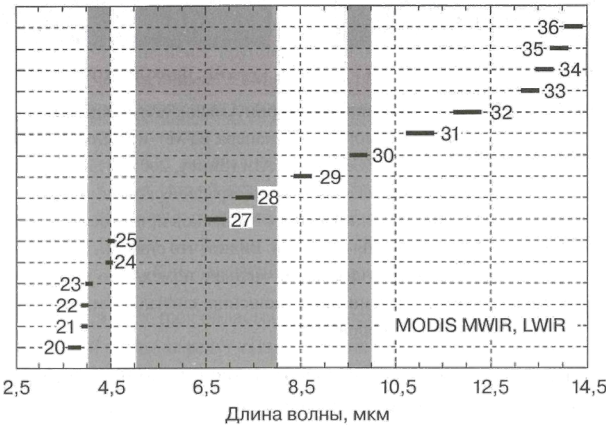
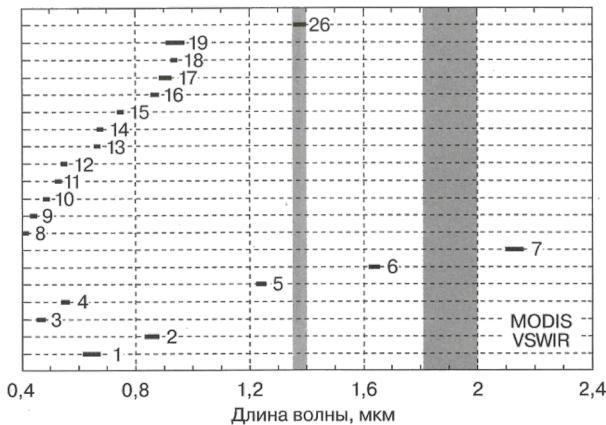
Основной принцип данных методов - сравнение изображения в каналах, соответствующих частотам, на которых искомый объект контрастно отличается на возможном помеховом фоне.

Существуют методы, хорошо зарекомендовавшие себя в постобработке данных дистанционного зондирования, обеспечивающие достаточно высокую надёжность, но требующие применения множества спектральных каналов.

Основной принцип многих из таких методов - поэтапная многоканальная пороговая фильтрация самих данных, либо метрик, на них основанных. Это обеспечивает хорошее качество, но, как уже отмечалось выше, требует значительных вычислительных ресурсов и реализуется, как правило, уже в наземной обработке.

Так, например, алгоритм классификации, используемый для данных Landsat7, «Landsat 7 ETM+ ACCA algorithm», согласно [1](El-Araby и др.) основывается на простом принципе: облачность контрастно выделяется высокой отражающей способностью и более низкой температурой. Однако для повышения надёжности, привлекается несколько спектральных каналов.

Вычислительная сложность. В приведённом примере, а также возможных его модификациях, учитывающих особенности входных данных, большинство операций сводятся к пороговой обработке, что составляет $O(n)$ от количества пикселей.



Распределение каналов MODIS по частотам. Закрашенные области представляют собой основные окна поглощения атмосферы. [1] (Шовенгердт Р.А.)

В данном алгоритме используется два прохода и в первом применяется 8 пороговых фильтров, принцип которых показан на рис. Слева (цитируется по [1](El-Araby и др.)). Второй проход предназначенный для разрешения неопределённостей, возникших на первом этапе, уже включает использование вычисленных более сложных метрик: (среднее (η), стд. Отклонение (σ^2), асимметрия (Skewness) и эксцесс (Kurtosis)), а также уточнённое адаптивное значение температурного порога, вычисленное по результатам 1го прохода.

Фильтры первого этапа

	Фильтр	функция
1	Яркостный порог $B_1 > 0.08$	удаляет тёмные изображения
2	Нормализованный дифференциальный снежный индекс $NDSI = \frac{B_2 - B_5}{B_2 + B_5} < 0.07$	удаляет множество видов снега
3	Температурный порог $B_6 < 300K$	удаляет горячие участки
4	Комбинация каналов 5/6 $(1 - B_5)B_6 < 255$	удаляет множество категорий в том числе лёд
5	отношение каналов 4/3 $\frac{B_4}{B_3} < 2$	удаляет яркую растительность и грунт
6	отношение каналов 4/2 $\frac{B_4}{B_2} < 2$	удаляет неоднозначные признаки
7	отношение каналов 4/5 $\frac{B_4}{B_5} > 1$	удаляет камни и пустыни
8	Комбинация каналов 5/6 $(1 - B_5)B_6 > 210^{(1)}$ или $< 210^{(2)}$	различение тёплых (1) и в холодных (2) облаков

B_k - обозначает k-й канал MODIS

$$\eta = \frac{1}{n} \sum_{i=1}^n x_i \quad \text{Skewness} = \frac{1}{n} \sum_{i=1}^n \left(\frac{x_i - \eta}{\sigma} \right)^3$$

$$\sigma^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \eta)^2 \quad \text{Kurtosis} = \frac{1}{n-3} \sum_{i=1}^n \left(\frac{x_i - \eta}{\sigma} \right)^4$$

Обобщённые правила классификации для первого этапа

Вывод	Правило
Снег	$\left(\frac{B_2 - B_5}{B_2 + B_5} > 0.7 \right) AND (B_4 > 0.1)^{*A}$
Пустыня	$\frac{B_4}{B_5} < 0.83^{*B}$
НеОблако	$B_3 < 0.08 OR (b_6 > 300) OR (Снег)$
Неопределённость	$\left(((1 - B_5)B_6 > 255) OR \left(\frac{B_4}{B_3} > 2 \right) OR \left(\frac{B_4}{B_2} > 2 \right) OR (Пустыня) \right) AND (\sim \text{НеОблако})$
Холодное облако	$((1 - B_5)B_6 \geq 210) AND (\sim \text{Неопределённость}) AND (\sim \text{НеОблако})$
Тёплое облако	$((1 - B_5)B_6 < 210) AND (\sim \text{Неопределённость}) AND (\sim \text{НеОблако})$

Приводится по статье [El-Araby и др.] с предложенными модификациями:
^{*A} Тест по каналу 4 добавлен после множественной некорректной классификации пикселей как снега на основе NDSI при применении к данным MODIS
^{*B} Порог понижен по сравнению с исходной версией (1.0) для исправления множественной неверной классификации облаков как пустыни (величина определена экспериментально)

Классификация на основе текстурных признаков

Этапы применения классификационных методов:

- Выделение и вычисление характерных признаков, дающих хорошую разделимость для искомого объекта и формирующих затем пространство признаков (Текстурные, частотные, статистические).
- Применение методов и алгоритмов классификации. Как то: метод К-ближайших соседей, метод опорных векторов, метод К-средних, различные версии Байесовского классификатора, различные более нестандартные версии совместного учёта признаков как, например в [6] где они используются для формирования псевдоцветов.
- Ещё одним этапом может быть последующая селекция результатов с применением дополнительной информации.

Критерии отбора признаков

- Линейная разделимость т.е. задача заключается в поиске таких признаков, которые бы давали разное значение для картинки с облаком и без, таким образом чтобы в пространстве векторов признаков нашлась бы такая (n-1 мерная) гиперплоскость чтобы значения для разных классов находились по разные её стороны.
- Предпочтение независимым наборам. Некоторые параметры, формально являясь разными, выражают схожие характеристики изображения и являются зависимыми. Одновременное использование таких характеристик мало повышает информативность набора, повышая при этом затраты на вычисление.
- Низкая вычислительная сложность.

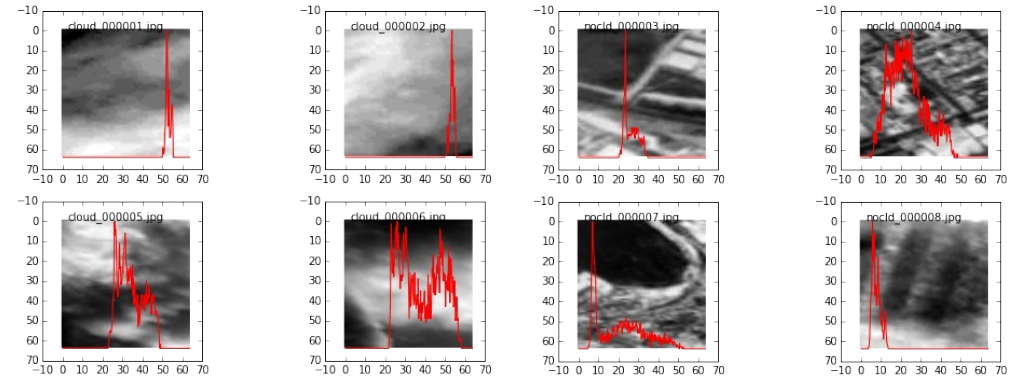
Вычислительная сложность

...алгоритмов классификации складывается из двухосновных составляющих: этапа формирования признаков, составляющего, для большинства из них $O(n)$ где n- количество пикселей окна обработки, и классификации, для приведённого примера k-средних, согласно (А.Часовских), $O(nkl)$, где k – число кластеров (в нашем случае 2), l – число итераций (также константа). Для других методов в том же источнике приводится

метод	выч. сложность
Иерархический	$O(n^2)$
k-средних	$O(nkl)$, где k – число кластеров,,
c-средних	l – число итераций
Выделение связанных компонент	зависит от алгоритма
Минимальное покрывающее дерево	$O(n^2 \log n)$
Послойная кластеризация	$O(\max(n, m))$, где $m < n(n-1)/2$

Пример библиотеки эталонных образцов

Для обучения классификатора и тестирования алгоритма сформирована библиотека эталонных образцов. На рисунке представлен пример нескольких изображений из такой библиотеки**. На каждом фрагменте нанесены гистограммы, соответствующие ранее приведённым образцам. Параметры фрагментов: 1 каналное 8 битовые размером 16x16 ***. Всего в библиотеке 575 фрагмента с облаком и 629 без облаков.



Слева – изображения содержащие облака, справа- не содержащие.

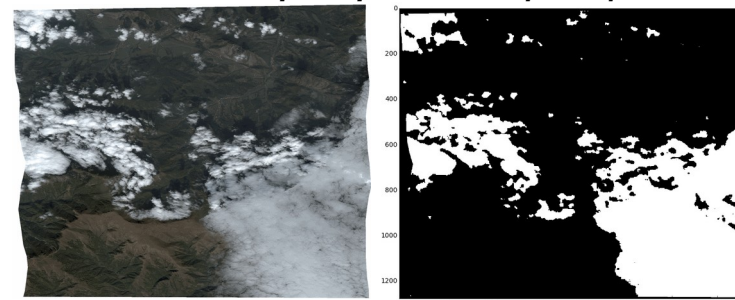
** для тестов использованы демонстрационные снимки quickbird и worldview-1 от catalog.sovzond.ru

***для наглядности на слайде показаны фрагменты 64x64, использованные в одной из предыдущих версий расчёта

На основе тестовых изображений и, сформированной по ним библиотеки образцов было проведено исследование по отбору наиболее информативных признаков для выявления облачности. В исходном наборе участвовали различные усреднённые меры яркости, статистические признаки, меры контраста, текстуры, а также моменты изображения. Всего на начальном этапе анализировалось 816 векторов по 3 признака и 120 по 2.

По приведённым критериям при помощи автоматического и ручного отбора были отвергнуты малоинформативные признаки и зависимые пары. Фрагмент полученного набора приведён в таблице справа.

Тестовые примеры классификации



Вектор признаков ('enegrg1IMstd', 'histStd', 'histmaxind')

Признаки
Лучшие по количеству ошибок
('histStd', 'semvarDeriv2', 'histmaxind')
('ent', 'contrast3', 'histmaxind')
('enegrg1IMstd', 'contrast3', 'histmaxind')
('histStd', 'contrast3', 'histmaxind')
('ent', 'histmaxind', 'imoment selK nu02')
('semvarDeriv2', 'contrast3', 'histmaxind')
('enegrg1IMstd', 'ent', 'histmaxind')
('contrast3', 'histmaxind', 'imoment selK nu02')
('histStd', 'ent', 'histmaxind')
('enegrg1IMstd', 'histStd', 'histmaxind')
('semvarDeriv2', 'ent', 'histmaxind')
('enegrg1IMstd', 'semvarDeriv2', 'histmaxind')
Выбранные вручную
('semvarDeriv2', 'ent', 'histmaxind')
('histStd', 'semvarDeriv2', 'histmaxind')
('histStd', 'ent', 'histmaxind')
('ent', 'histmaxind', 'imoment selK nu02')
('ent', 'contrast3', 'histmaxind')
('enegrg1IMstd', 'histStd', 'contrast3')
('enegrg1IMstd', 'ent', 'histmaxind')

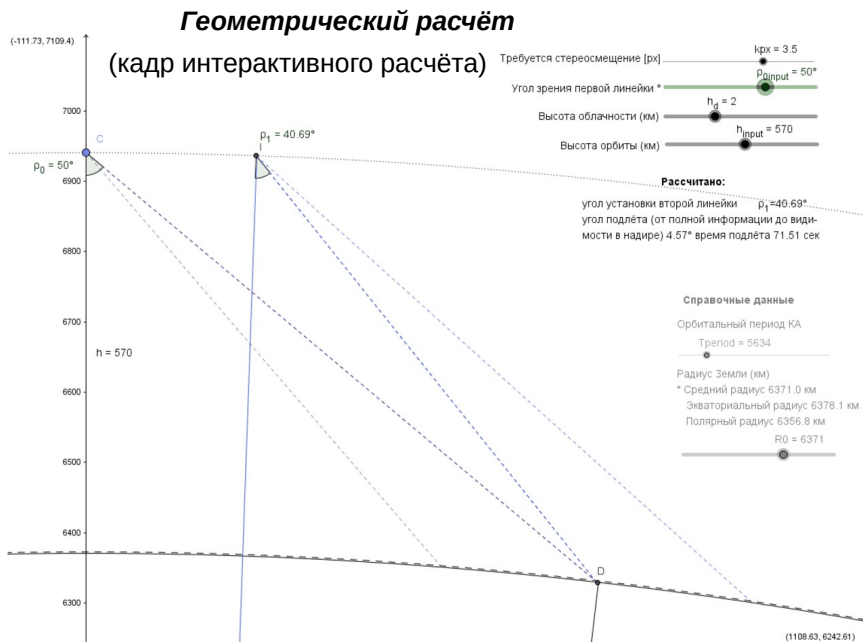
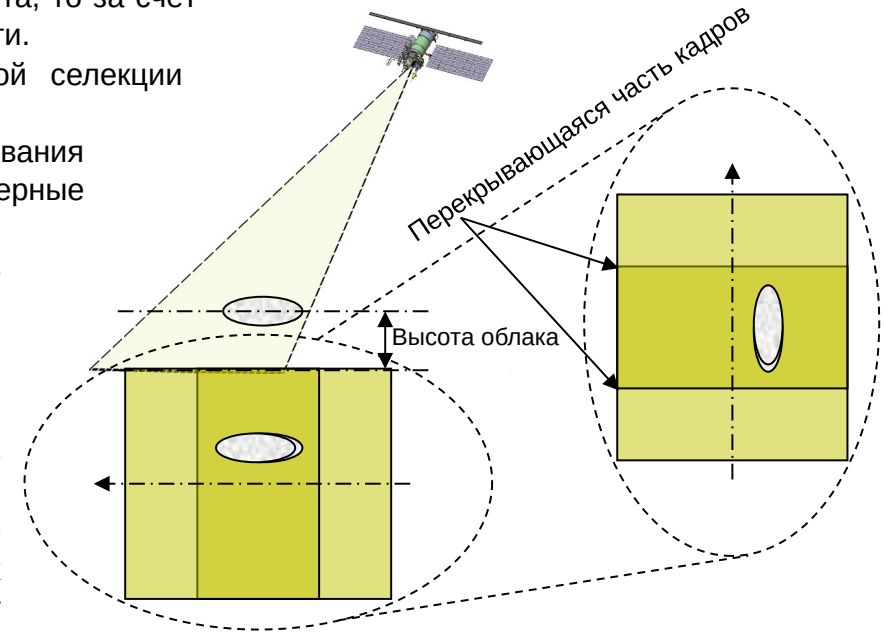
Методы на основе псевдостереообработки

Идея метода заключается в том, что если камера предварительного обзора снимает два кадра так, что область интереса входит в оба кадра, снятых на разных участках маршрута, то за счёт появления стереобазы можно получить информацию о наличии и высоте облачности.

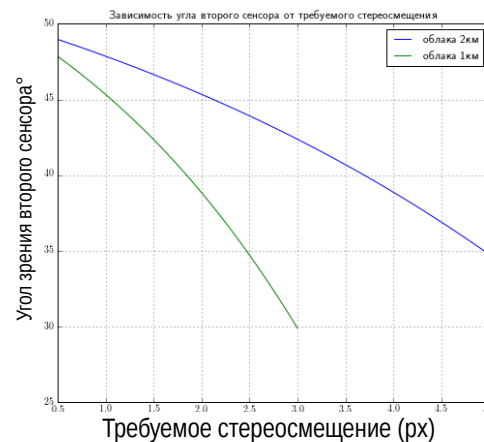
Так в работе [4] (Королев Е.Е. и др) это используется для дополнительной селекции предварительно найденных облачных объектов.

Там же, со ссылкой на [5] (Пошехонов В.И.), приводится пример использования алгоритма стереообработки, в соответствии с которым определяются трёхмерные координаты каждого пикселя.

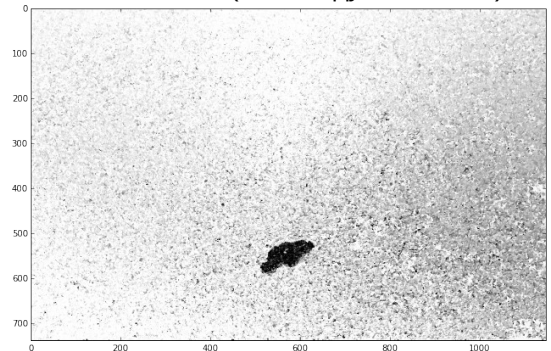
В данной работе предлагается иной подход: на основе современных алгоритмов анализа движения. Суть его состоит в том, что для целей детектирования не важен строгий и точный расчёт высоты облачности. Достаточно зафиксировать видимое смещение облачности, возникающее с другой точки съёмки. Для его фиксации достаточно применить относительно более быстрые алгоритмы анализа движения, подобные тем что используется, например, для анализа видеосигнала что позволит значительно упростить и ускорить расчёт. И, хотя **вычислительная сложность** таких алгоритмов достаточно велика и может достигать, согласно (Baker S.), $O(n^3)$, небольшой размер кадра и использование более быстрых вариантов анализа смещения (в тесте, например (Farneback, G.), позволит достичь приемлемых значений.



Геометрический расчёт (на примере параметров КА Ресурс-ДК) показывает что достаточно надёжная разделимость в несколько пикселей (3.5 на кадре) смещения может быть получена для высоты облачности 1-2 км при, например, углах просмотра 50 и 40 градусов уже более чем за 70 секунд до съёмки основной камерой, что позволяет провести всю необходимую обработку.



Пример на модельном изображении получен с применением алгоритма выделения оптического потока (по методу Farneback)



Сравнение методов поиска безоблачных участков

Метод, основанный на	Пригодность к автоматической работе	Требовательность к входным данным	Надёжность распознавания	Общая оценка пригодности
Анализе многоканального изображения	Все перечисленные методы и их комбинации пригодны или могут быть адаптированы к работе без участия оператора.	Высокая: многоканальное изображение	Высокая, при обеспечении требуемыми входными данными	Малопригодно в полноценном виде при сохранении требований к стоимости и массогабаритным характеристикам КПО ***
Текстурных признаках панхроматического изображения		Может использовать один канал, но качество повышается при использовании более высокого разрешения	Достаточно высокая, но может давать ложноположительные срабатывания в присутствии схожих целей (например снег и лёд)	Могут быть использованы при объединении в рамках одной съёмочной системы как взаимодополняющие ***
Псевдостереоизображении		Может использовать два кадра относительно низкого разрешения с одного канала	Достаточная (**на основе моделирования)	

*** Также рассматривалась опция использования второго частотного канала при наличии возможности установки второй сенсорной линейки в камеру, что позволит использовать сильные стороны всех рассмотренных методов для повышения надёжности распознавания.

Выводы и предложения

- 1) Предлагается дополнить состав аппаратуры КА высокодетального зондирования в оптическом диапазоне камерой предварительного обзора для оценки на борту облачной обстановки в предполагаемом районе съёмки что обеспечит повышение эффективности.
- 2) Для реализации наиболее качественного определения безоблачных участков, предлагается использовать схему, содержащую два двухканальных линейных сенсора, установленных под разными зенитными углами по курсу КА, что обеспечит необходимые данные для достаточно надёжного автоматического детектирования облачности на борту КА.
- 3) Представляет интерес возможность установки подобной камеры также на платформе мини (нано-) спутника или их группировки..
- 4) Наиболее перспективными для автономного автоматического принятия решений в условиях ограниченного времени представляется использование алгоритмов классификации на основе текстурных признаков, совместно с анализом на основе псевдостереоэффекта.
- 5) Анализ быстродействия достоверности и надёжности существующих методов позволяет сделать вывод о возможности реализации автоматической оценки распределения облачности на борту КА, т.к. скорость обработки обеспечивает получение результатов за время полёта к месту съёмки основной камерой.
- 6) Определение проектного облика системы проводится с учетом факторов низкой стоимости создания прибора, его возможных существующих прообразов высокой заводской готовности при требованиях к минимизации массово-габаритных и по энергопотреблению характеристик. Оценочные расчёты технических параметров системы показывают её практическую реализуемость и низкие требования к технической сложности элементов.
- 7) На предлагаемый способ получен патент

Список литературы

- Шовенгердт Р.А. - Дистанционное зондирование. Методы и модели обработки изображений
- Чернокульский А.ф В., Мохов И. И. Сравнительный анализ характеристик глобальной и зональной облачности по различным спутниковым и наземным наблюдениям. / Исследование Земли из космоса, (3):12–29, 2010.
- Скороходов А. В. Нейросетевой классификатор облачности.
- Оценки качества для анализа цифровых 046-Monich.pdf
- Ю.И. Монич, В.В. Старовойтов Оценки качества для анализа цифровых изображений. «Искусственный интеллект» 4'2008
- Королев Е.Е., Кочергин А.М., Кузнецов А.Е., Побаруев В.И. Автоматическая сегментация облачных объектов на снимках земной поверхности высокого пространственного разрешения ФГБОУ ВПО «Рязанский государственный радиотехнический университет»
- Пошехонов В.И. Алгоритмы и модели стереофотограмметрической обработки данных от систем спектральной съемки Земли
- Родионова Н.В. Текстуриная сегментация одноканальных изображений: примеры применения. "Современные проблемы дистанционного зондирования Земли из космоса" 2012. Т.9.№3. С.65-69
- В.Г. Астафуров и др. Классификация перистой облачности по данным MODIS с помощью нечеткой нейронной сети. Современные проблемы дистанционного зондирования Земли из космоса. 2014. Т. 11. №4. С. 265–275
- П.П.Гейко. Контроль концентрации отравляющих веществ лидаром дифференциального поглощения на основе CO2-лазера. Доклады ТУСУРа, №2 (24), часть 3, декабрь 2011
 - Michael E. Webber, Michael Pushkarsky, C. Kumar N. Patel. Optical detection of chemical warfare agents and toxic industrial chemicals: Simulation Journal of applied physics 97, 113101 (2005)
 - Wylie D.P., Menzel W. Two years of cloud cover statistics using VAS // Journal of Climate. 1989. Vol. 2, № 4. P. 380–392.
 - А. Часовских. Обзор алгоритмов кластеризации данных habr.com/ru/post/101338/, 11 августа 2010
 - Farneback, G. "Two-Frame Motion Estimation Based on Polynomial Expansion." In Proceedings of the 13th Scandinavian Conference on Image Analysis, 363 - 370. Halmstad, Sweden: SCIA, 2003.
 - Simon Baker and Iain Matthews - Lucas-Kanade 20 Years On: A Unifying Framework: Part 1, The Robotics Institute, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213, USA



Сведения об авторах:

Вторин Сергей Александрович, Институт космических исследований РАН, (117997, г. Москва, ул. Профсоюзная 84/32)
т. (495) 333-51-66, e-mail: vtiurin@iki.rssi.ru
Князев Николай Александрович, Институт космических исследований РАН, (117997, г. Москва, ул. Профсоюзная 84/32)
т. (495) 333-51-66, e-mail: nknyazev@iki.rssi.ru