

Алгоритм обнаружения облачных объектов на многоспектральных снимках

О.В. Николаева¹

¹ Институт прикладной математики РАН, Москва, Россия

Аннотация

Представлен многошаговый алгоритм обнаружения облачных объектов на многоспектральных снимках земной поверхности. На каждом шаге выполняется кластеризация пространственных пикселей многоспектрального снимка методом *k*-средних и *k* фрагментам полученных кластеров применяются спектральные критерии облачности/чистого неба. Эта процедура повторяется до тех пор, пока находятся «облачные» пиксели. «Облачные» пиксели, найденные на одном шаге, объединяются в один облачный объект.

Представлены результаты тестирования этого алгоритма на снимках прибора HYPERION (199 ненулевых спектральных каналов в спектральном диапазоне 426 нм–2400 нм) с высоким пространственным разрешением (30 м). Выбраны снимки с частичной облачностью над поверхностями пяти типов: океан, зеленая растительность, пустынная местность, городская застройка и снег.

Выполнено сравнение результатов обработки этих снимков представленным алгоритмом и альтернативным алгоритмом, в котором те же спектральные критерии применяются независимо к каждому пространственному пикселю. Представлены средние для каждого облачного объекта спектры. Показано, что представленный алгоритм на каждом снимке находит от 1 до 3 облачных объектов, отвечающих распределению яркостей на соответствующих RGB-изображениях. Применение альтернативного алгоритма (без предварительной кластеризации) приводит к ошибкам обнаружения на краях облаков.

Предложено три параметра для оценки качества получаемых облачных масок. Из них выбран наиболее информативный: отношение разброса спектров в «облачных» пикселях к разбросу спектров в «чистых» пикселях (отношение должно быть много меньше 1).

Ключевые слова: обнаружение облаков, многоспектральные снимки, спектральные критерии, показатели качества.

Цитирование: Николаева, О.В. Алгоритм обнаружения облачных объектов на многоспектральных снимках / О.В. Николаева // Компьютерная оптика. – 2022. – Т. 46, № 5. – С. 808-817. – DOI: 10.18287/2412-6179-CO-1076.

Citation: Nikolaeva OV. Technique of detecting cloudy objects in multispectral images. Computer Optics 2022; 46(5): 808-817. DOI: 10.18287/2412-6179-CO-1076.

Введение

Обнаружение облаков на многоспектральных снимках является актуальной задачей. Обнаруженные облачные объекты либо исследуются (если выполняется мониторинг облаков), либо удаляются со снимков (если нужны изображения только земной поверхности).

Также облачные пиксели исключаются из рассмотрения, если решается задача определения содержания поглощающих газов. В этом случае измерения выполняются в узких спектральных диапазонах с очень высоким спектральным разрешением. Пространственный пиксель считается «чистым», если измеренный спектр яркости может быть получен при моделировании взаимодействия излучения с «чистой» атмосферой (при построении модели «чистой» атмосферы используются актуальные данные метеонаблюдений) [1].

Съемка земной поверхности выполняется с меньшим спектральным разрешением. В этом случае основной алгоритм нахождения облачных объектов

является набор спектральных пороговых критериев, применяемых отдельно к каждому пространственному пикселю. В ранних алгоритмах «облачными» назначались пиксели, чья яркость в видимом диапазоне больше порогового значения [2]. В широко используемых в настоящее время алгоритмах дополнительно применяются условия на значения спектральных индексов NDVI (Normalized Difference Vegetation Index) и NDSI (Normalized Difference Snow Index), а также радиационной температуры, определяемой по измерениям в спектральном диапазоне 11–13 мкм [3–7]. При использовании данных радиометра используются только значения радиационной температуры [8]. Поскольку водяной пар располагается главным образом на высотах 0–3 км над поверхностью земли, а облака экранируют приземные газы, то для обнаружения стратосферных облаков используется условие на яркость пикселя в полосе поглощения водяного пара 1,38 мкм [9]. Для обнаружения плотных облаков используется условие на яркость пикселя в полосе поглощения кислорода [10]. Также для обна-

ружения облаков используется условие на белизну пикселя (отклонение спектра в видимом диапазоне от константы) [10].

Важным является вопрос о значении порогов в критериях. Для снижения ошибок при обнаружении облаков используются значения порогов, зависящие от типа подстилающей поверхности и условий освещения и съемки [3], высоты и температуры земной поверхности [8]. Отдельные значения порогов для каждого снимка могут быть найдены предварительно на земле и использованы на орбите [11].

Из-за трудностей, связанных с выбором значений порогов, развивается другой подход – обнаружение облаков с помощью сравнения нескольких снимков одного участка земной поверхности, но сделанных в разные моменты времени. Снимки делаются либо двумя приборами на одном космическом аппарате, либо одним прибором, но на разных витках [12]. Во втором случае создается архив безоблачных изображений различных участков земной поверхности; облако определяется применением спектральных критериев к измененной части снимка [7, 13]. Также облако может выявляться как искажение сцены на одном из снимков в последовательной серии [14].

Еще один подход к обнаружению облачных объектов опирается на предварительную кластеризацию снимка (например, методом *k*-средних). Однако применение такого алгоритма к монохромным снимкам приводит к большим ошибкам [15].

В настоящей работе предлагается алгоритм, в котором процедура кластеризации применяется к многоспектральному снимку. Далее определяется, какой из кластеров может содержать облако, и облачные пиксели выделяются из этого кластера с помощью набора спектральных критериев.

Важным также является вопрос об оценке качества получаемых облачных масок. Часто выполняется сравнение двух облачных масок, полученных двумя алгоритмами на близких по времени и месту съемки данных разных приборов [1]. Также можно сравнить две облачные маски, построенные двумя разными алгоритмами на одних и тех же данных. Однако представляется необходимым иметь объективные критерии для оценки качества алгоритмов.

В настоящей работе предлагается три показателя качества, каждый из которых представляет собой отношение некоторого функционала для спектров в «облачных» пикселях к аналогичному функционалу для спектров в «чистых» пикселях.

1. Постановка задачи

Рассмотрим многоспектральный снимок, состоящий из M и N пространственных пикселей поперек и вдоль полосы съемки соответственно. Каждому пространственному пикселю снимка сопоставим пару индексов (m, n) . Пусть измеренная в этом пикселе спектральная зависимость интенсивности излучения

задана вектором $\mathbf{I}_{m,n}$ длины L . Преобразуем вектор интенсивности в вектор яркости по формуле [16]

$$\mathbf{R}_{m,n} = \mathbf{I}_{m,n} \pi / (\mathbf{I}_0 \cos \theta_0). \quad (1)$$

Здесь θ_0 – зенитный угол солнца, \mathbf{I}_0 – вектор значений интенсивности внеатмосферного излучения. Операции с векторами в равенстве (1) выполняются покомпонентно.

Все векторы $\mathbf{R}_{m,n}$ образуют множество точек в пространстве R^L . Элемент вектора $\mathbf{R}_{m,n}$, отвечающий длине волны λ_{ℓ} , обозначим как $R_{m,n,\ell}$.

Рассмотрим задачу: разделить все пространственные пиксели (векторы $\mathbf{R}_{m,n}$) на «облачные» и «чистые». При решении таких задач обычно используются следующие спектральные величины.

1. $R_{m,n,\ell Br}$ – коэффициент яркости для выбранной длины волны $\lambda_{\ell Br}$, где облако должно быть ярче земной поверхности.

2. Индекс NDVI [3]

$$NDVI_{m,n} = \frac{R_{m,n,\ell NIR} - R_{m,n,\ell RED}}{R_{m,n,\ell NIR} + R_{m,n,\ell RED}}, \quad (2)$$

где $R_{m,n,\ell NIR}$ и $R_{m,n,\ell RED}$ – коэффициенты яркости для длин волн из ближнего инфракрасного $\lambda_{\ell NIR}$ и видимого красного $\lambda_{\ell RED}$ диапазонов.

3. Индекс NDSI [3]

$$NDSI_{m,n} = \frac{R_{m,n,\ell VIS} - R_{m,n,\ell SNIR}}{R_{m,n,\ell VIS} + R_{m,n,\ell SNIR}}, \quad (3)$$

где $R_{m,n,\ell SNIR}$ и $R_{m,n,\ell VIS}$ – коэффициенты яркости для длин волн из коротковолнового инфракрасного $\lambda_{\ell SNIR}$ и видимого зеленого $\lambda_{\ell VIS}$ диапазонов.

4. $R_{m,n,\ell O_2}$ – коэффициент яркости для длины волны $\lambda_{\ell O_2}$ в полосе поглощения кислорода.

Используются следующие пороговые критерии «облачности» пикселей:

$$R_{m,n,\ell Br} > R^{Br}, \quad (4)$$

$$|NDVI_{m,n}| < N_0, \quad (5)$$

$$R_{m,n,\ell O_2} > R^{O_2}. \quad (6)$$

В условии (4) предполагается, что облако ярче, чем поверхность, для длины волны $\lambda_{\ell Br}$ [3]. Условие (5) позволяет отделить облако от растительности, однако может быть выполнено и для снега, и для обнаженной породы [3]. В условии (6) используется тот факт, что облака экранируют приземные газы, в том числе кислород, и увеличивают в полосе поглощения кислорода коэффициент яркости [10]. Пороговые значения R^{Br} , N_0 , R^{O_2} и длина волны $\lambda_{\ell Br}$ зависят от конкретного алгоритма, от альбедо подстилающей поверхности и условий съемки.

Кроме того, чтобы отличить облако от снега, применяется критерий [4]

$$NDSI_{m,n} < 0,4. \quad (7)$$

Здесь используется тот факт, что в инфракрасной области снег много темнее, чем в видимой.

От обнаженной породы облако можно отличить с помощью условия

$$NDSI_{m,n} > -0,1. \quad (8)$$

Здесь используется тот факт, что в инфракрасной области порода много ярче, чем в видимой. Часто используется условие на значение радиационной температуры (облако холоднее, чем земная поверхность). Однако измерения в нужном спектральном диапазоне (11 – 13 мкм) не всегда имеются.

В нашем алгоритме, помимо критериев (4–8), будем учитывать величину относительного отклонения (разброса) спектров группы пикселей. Эту величину определим формулой

$$D(G) = \frac{1}{P(G)} \sqrt{\frac{1}{L \cdot |G|} \sum_{(n,m) \in G} \sum_{\ell=1}^L (R_{m,n,\ell} - S_\ell(G))^2}, \quad (9)$$

где

$$S_\ell(G) = \frac{1}{|G|} \sum_{(n,m) \in G} R_{m,n,\ell}, \quad (10)$$

– средний спектр в группе G , $|G|$ – число пикселей в группе, $P(G)$ – среднее значение спектра $S_\ell(G)$

$$P(G) = \frac{1}{|G|} \sum_{\ell=1}^L |S_\ell(G)|.$$

2. Алгоритм обнаружения облачных объектов

Каждый шаг алгоритма состоит из трех этапов.

Пусть снимок сделан не при сплошной облачности, то есть на снимке может присутствовать облако. Можно предположить, что среднее расстояние между двумя «облачными» векторами много меньше среднего расстояния между «облачным» и «чистым» вектором. Поэтому на первом этапе выполняется разбиение всего множества векторов $R_{m,n}$ на два кластера методом к-средних.

На втором этапе определяется, какой из двух кластеров с большей вероятностью содержит облако. Для каждого кластера G_i , $i=1,2$, находятся разброс $D(G_i)$ по формуле (9), а также средние по кластеру значения спектральных индексов

$$R^{Br}(G_i) = \frac{1}{|G_i|} \sum_{(m,n) \in G_i} R_{m,n,\ell^{Br}},$$

$$|NDVI(G_i)| = \frac{1}{|G_i|} \sum_{(m,n) \in G_i} |NDVI_{m,n}|,$$

$$R^{O_2}(G_i) = \frac{1}{|G_i|} \sum_{(m,n) \in G_i} R_{m,n,\ell^{O_2}}.$$

Пусть значения спектральных индексов двух кластеров отличаются менее, чем отличаются величины разбросов, то есть пусть выполнено условие

$$\max \{u, v, w\} < |1 - D(G_1)/D(G_2)|, \quad (11)$$

где

$$u = |1 - R^{Br}(G_1)/R^{Br}(G_2)|, \quad v = |1 - R^{O_2}(G_1)/R^{O_2}(G_2)|,$$

$$w = |1 - |NDVI(G_1)|/|NDVI(G_2)||.$$

Тогда «облачный» кластер определяется только по величинам разброса. «Облачным» назначается кластер с наименьшим значением разброса (здесь предполагается, что «облачный» кластер является более однородным, чем «необлачный»).

Если условие (11) не выполнено, то «облачным» назначается кластер G_i , для которого выполнена большая часть неравенств

$$R^{Br}(G_i) > R^{Br}(G_j), \quad R^{O_2}(G_i) > R^{O_2}(G_j),$$

$$|NDVI(G_i)| < |NDVI(G_j)|.$$

Иначе говоря, в соответствии с критериями (4–6) «облачным» назначается кластер, более яркий для выбранных длин волн $\lambda_{\ell^{Br}}$, $\lambda_{\ell^{O_2}}$ и с меньшим средним значением модуля индекса $|NDVI|$.

На третьем этапе из выбранного на втором этапе «облачного» кластера отбираются «облачные» пиксели. Для этого сначала значения спектральных критериев в каждом пикселе из «облачного» кластера усредняются по всем соседним пикселям, принадлежащим тому же кластеру. Далее критерии (4–8) применяются отдельно к каждому пикселю из «облачного» кластера.

Если пикселей, удовлетворяющих этим критериям, не находится, поиск облачных объектов заканчивается. Иначе удовлетворяющие критериям (4–8) пиксели объединяются в облачный объект и исключаются из дальнейшего рассмотрения. Далее с оставшимися пикселями выполняются все этапы следующего шага.

Отметим, что на каждом шаге находится один облачный объект. Иначе говоря, число найденных облачных объектов M определяется числом шагов, а не задается заранее. Если на первом же шаге никакого облачного объекта не найдено, то считается, что снимок сделан при чистом небе.

Если на снимке был найден хотя бы один облачный объект, то выполняется оценка качества построенной облачной маски. Обозначим через $Cloud_i$ группу пикселей, отнесенных к i -му облачному объекту. Обозначим через $Clear$ пиксели, отнесенные к чистому небу. Найдем следующие величины:

- 1) величины разброса $D(Cloud_i)$ и $D(Clear)$;
- 2) средние значения коэффициента яркости в полосе поглощения кислорода $R^{O_2}(Cloud_i)$ и $R^{O_2}(Clear)$;
- 3) величины пространственной вариации коэффициента яркости в видимом зеленом диапазоне $V(Cloud_i)$ и $V(Clear)$ по формуле:

$$V(G) = \frac{1}{|G|} \sum_{(m,n) \in G} |1 - R_{m,n,\ell^{VIS}} / \tilde{R}_{m,n,\ell^{VIS}}|,$$

$$\tilde{R}_{m,n,\ell^{VIS}} = \sum_{m'=m-1}^{m+1} \sum_{n'=n-1}^{n+1} R_{m',n',\ell^{VIS}} / 8 \Big|_{\text{при } |m-m'|+|n-n'|>0}.$$

Показатели качества определим как отношения этих величин

$$d = \sum_i K_i D(Cloud_i) / D(Clear) K, \tag{12}$$

$$r = \sum_i K_i R^{O_2}(Cloud_i) / R^{O_2}(Clear) K, \tag{13}$$

$$v = \sum_i K_i V(Cloud_i) / V(Clear) K. \tag{14}$$

Здесь K_i – число пикселей в i -м облачном объекте, $K = \sum_i K_i$ – общее число облачных пикселей.

Показатель (12) должен быть много меньше 1, поскольку разброс в спектрах «облачных» пикселей меньше, чем разброс в спектрах пикселей чистого неба над неоднородной подстилающей поверхностью. Показатель (13) должен быть много больше 1, поскольку облака экранируют приземный кислород и коэффициент яркости над облаком больше коэффициента яркости над безоблачной неяркой земной поверхностью. Показатель (14) также должен быть больше 1, так как в видимом спектральном диапазоне (0,559 мкм) яркость облаков отличается большей пространственной вариабельностью по сравнению с яркостью чистого неба [17].

3. Численные результаты

Тестирование алгоритма выполним на снимках прибора HYPERION (199 ненулевых спектральных каналов в диапазоне (426 нм, 2400 нм), пространственное разрешение 30 м). Будем рассматривать снимки, содержащие одинаковое число пикселей вдоль и поперек полосы съемки $M=N=256$. Выберем снимки, сделанные над разными подстилающими поверхностями, см. табл. 1.

В снимке номер 1 используем два фрагмента.

При обработке снимков в критериях (4), (6) используются те же значения параметров, что и в алгоритме MODIS [3, 4]:

$\lambda_{\ell^{Br}} = 0,86$ мкм над водой или пустынной местностью,

$\lambda_{\ell^{Br}} = 0,65$ мкм над растительностью, $N_0 = 0,15$,

$R^{Br} = 0,055$ над водой,

$R^{Br} = 0,034$ над пустынной местностью,

$R^{Br} = 0,18$ над растительностью.

В полосе поглощения кислорода используется длина волны $\lambda_{\ell^{O_2}} = 0,762$ мкм. При определении индексов NDVI и NDSI используются длины волн $\lambda_{\ell^{RED}} = 0,65$ мкм, $\lambda_{\ell^{NIR}} = 0,84$ мкм, $\lambda_{\ell^{SNIR}} = 1,638$ мкм, $\lambda_{\ell^{VIS}} = 0,559$ мкм.

Обработка выполняется как представленным алгоритмом (будем называть его алгоритм 1), так и

применением критериев (4–8) отдельно к каждому пикселю (алгоритм 2).

Табл. 1. Используемые снимки

№	Поверхность	ID	Дата съемки	Место съемки
1	Океан	EO1H05207 02011329110T 2 PF2 01	25.11.2011	Тихий океан
2	Растительность	EO1H128031 2016330110K 7 АК3 01	09.10.2012	Таиланд
3	Пустынная местность	EO1H128031 2016330110 K7 АК3 01	25.11.2016	Внутренняя Монголия, КНР
4	Город	EO1H04403420021 96110PF_AGS_01	15.07.2002	Сан-Франциско, США
5	Снег	EO1H04202420090 46110KX_SGS_01	15.02.2009	Близ Калгари, Канада

Процесс построения облачной маски на примере фрагмента 1 снимка 1 иллюстрирует рис. 1. Представлены RGB-изображения, на которые на панелях (б), (в), (г) нанесены кластеры, построенные на первом, втором и третьем шагах алгоритма соответственно. «Облачные» кластеры отмечены белым цветом, «чистые» – черным. На первом и втором шагах из «облачных» кластеров с помощью критериев (4–8) образуются первый и второй облачные объекты, см. рис. 2б. На третьем шаге в соответствии с этими критериями «облачный» кластер оказывается чистым небом.

Полученные для всех снимков результаты представлены в табл. 2, 3 и на рис. 2–13.

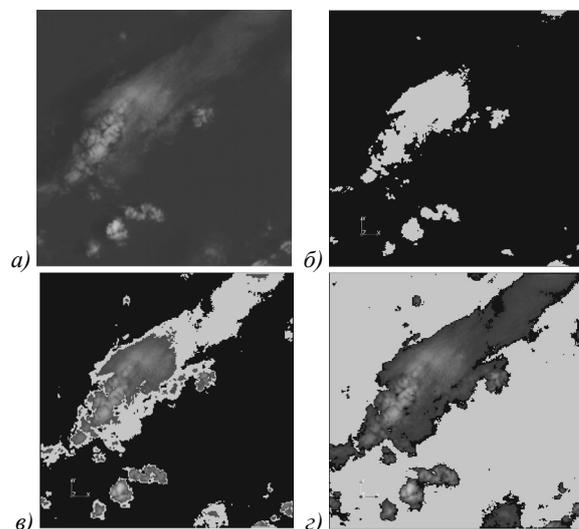


Рис. 1. Кластеры, получаемые в процессе построения облачной маски для снимка 1, фрагмента 1

На рис. 2, 4, 6, 8, 10, 12 представлены:

- а) – RGB-изображение,
- б) – облачная маска, полученная алгоритмом 1,
- в) – облачная маска, полученная алгоритмом 2,
- г) – средние спектры в каждом облачном объекте и в чистом небе.

На облачных масках черным цветом обозначено «чистое небо».

На рис. 3, 5, 7, 9, 11, 13 представлены коэффициенты яркости в полосе поглощения кислорода λ_{O_2} . В табл. 2 приведены значения показателей качества (12–14). В табл. 3 приведены доли (%) пропущенных «облачных» пикселей:

S_{12} – доля «облачных» пикселей алгоритма 1, отнесенных алгоритмом 2 к чистому небу.

S_{21} – доля «облачных» пикселей алгоритма 2, отнесенных алгоритмом 1 к чистому небу.

Из таблиц и рисунков можно сделать следующие выводы.

Повышенные значения коэффициента яркости в полосе поглощения кислорода отвечают облачным объектам для всех подстилающих поверхностей, кроме снега (ср. рис. 2а и рис. 3, рис. 4а и рис. 5, рис. 6а и рис. 7, рис. 8а и рис. 9, рис. 10а и рис. 11, рис. 12а и рис. 13). Это связано с тем, что повышение коэффициента яркости за счет отражения от снежной поверхности компенсирует его уменьшение за счет поглощения в кислороде.

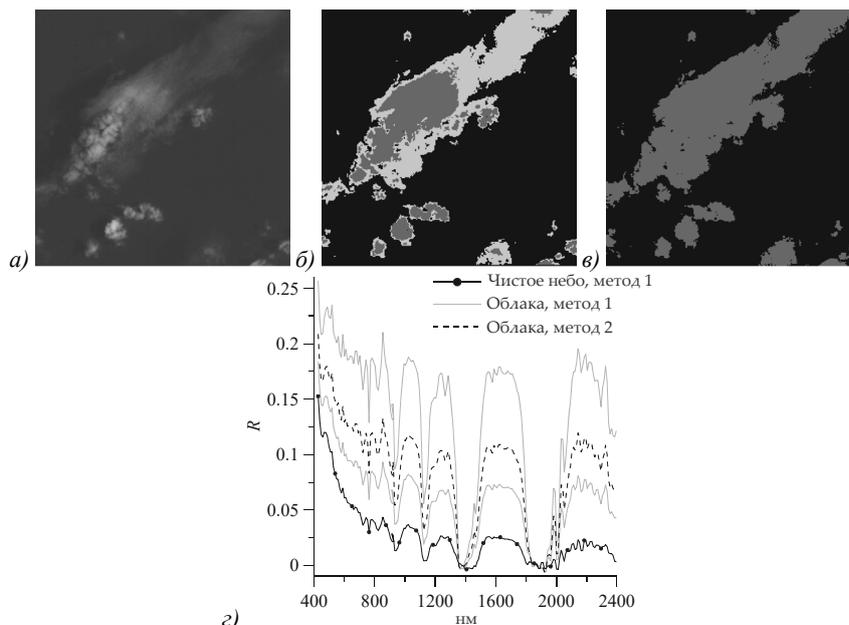


Рис. 2. Результаты обработки снимка 1, фрагмент 1

Алгоритм 1 выделяет от 1 до 3 компактных облачных объектов на каждом снимке. Расположение найденных облачных объектов соответствует расположению облаков на RGB-изображениях (ср. рис. 2а с 2б, рис. 4а с 4б, рис. 6а и 6б, рис. 8а с 8б, рис. 10а с 10б, рис. 12а с 12б).

Алгоритм 2 по построению всегда находит одно облако. Это облако является излишне обширным на фрагменте 2 снимка 1 (океан), ср. рис. 4а, 5 и рис. 4б, в, и на снимке 5 (снег), ср. рис. 12а, 13 и рис. 12б, в. Здесь алгоритм 2 включает в облако приоблачный аэрозоль. В результате средние «облачные» спектры в алгоритме 2 оказываются ближе к «чистым», а не к «облачным» спектрам алгоритма 1 (см. рис. 4д и рис. 12д).

Для фрагмента 1 снимка 1 (океан), снимка 2 (растительность) и снимка 4 (город) доля пикселей, отнесенных одним алгоритмом к облаку, а другим алгоритмом – к чистому небу, не превышает 10%.

Для снимка 3 (пустынная местность) облако, найденное алгоритмом 2, практически совпадает с облаком, найденным алгоритмом 1, см. табл. 3.

Для облачных масок, полученных алгоритмом 1, показатель качества $d \ll 1$ (разброс спектров в «об-

лачных» пикселях всегда много меньше, чем в «чистых»). При переходе к алгоритму 2 показатель d возрастает и может оказаться много больше 1. Это происходит на фрагменте 2 снимка 1 и на снимке 5, когда алгоритм 2 включает в облако часть приоблачного аэрозоля, см. табл. 2.

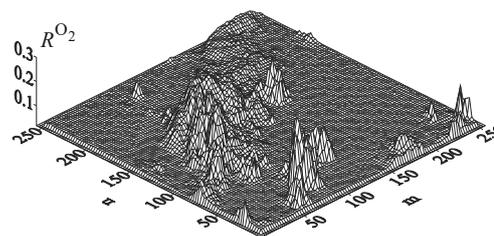


Рис. 3. Коэффициент яркости при $\lambda = \lambda_{O_2}$ для снимка 1, фрагмент 1

Почти для всех построенных облачных масок показатель качества $r \gg 1$ (в полосе поглощения кислорода «облачные» пиксели в среднем ярче, чем «чистые»). Только над снегом (снимок 5) показатель r близок к 1. При включении в облако части аэрозоля алгоритмом 2 (фрагмент 2 снимка 1 и снимок 5) значение показателя r лишь несколько снижается, см. табл. 2.

Табл. 2. Показатели качества

Снимок	Алгоритм 1			Алгоритм 2		
	d	r	v	d	r	v
1, фрагмент 1	0,78	2,92	1,97	1,88	2,82	1,93
1, фрагмент 2	0,58	3,76	2,8	1,96	2,72	2,02
2	0,46	1,52	1,79	0,94	1,43	1,78
3	0,012	2,09	1,83	0,015	1,98	1,82
4	0,46	1,5	2	0,46	1,5	2
5	0,39	1,08	1,08	1,04	0,96	0,97

Это утверждение справедливо и для показателя качества v , см. табл. 2.

Таким образом, из всех рассмотренных показателей качества d, r, v только отношение величин разброса d позволяет обнаружить включение в облако части чистого неба (в этом случае величина d становится заметно больше 1).

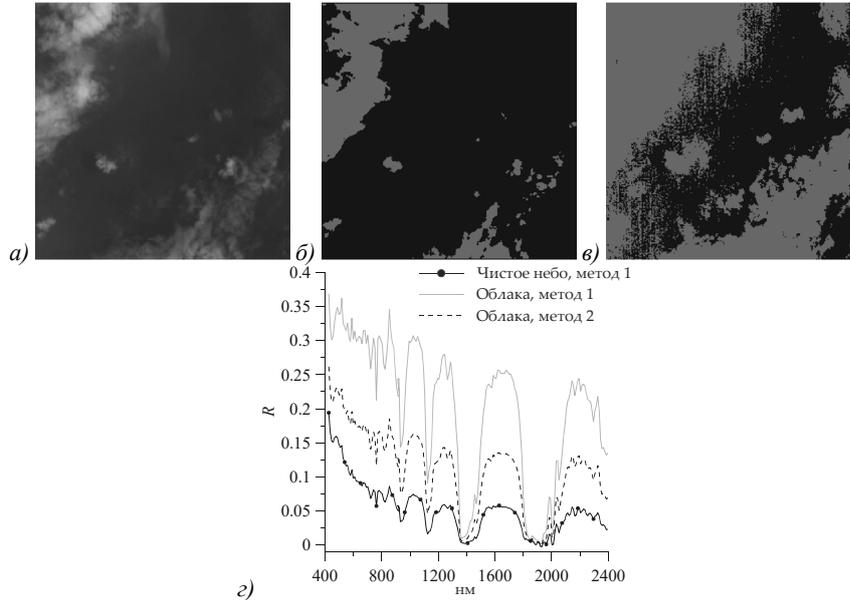


Рис. 4. Результаты обработки снимка 1, фрагмент 2

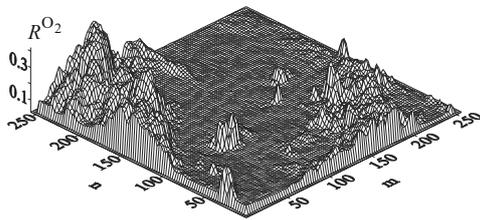


Рис. 5. Коэффициент яркости при $\lambda = \lambda_{O_2}$, для снимка 1, фрагмент 2

Табл. 3. Доли (%) пропущенных облачных пикселей

Снимок	S_{12}	S_{21}
1, фрагмент 1	0,26	8,65
1, фрагмент 2	0	69,7
2	1,9	6,6
3	1,25	0,66
4	6	9,6
5	7	37,5

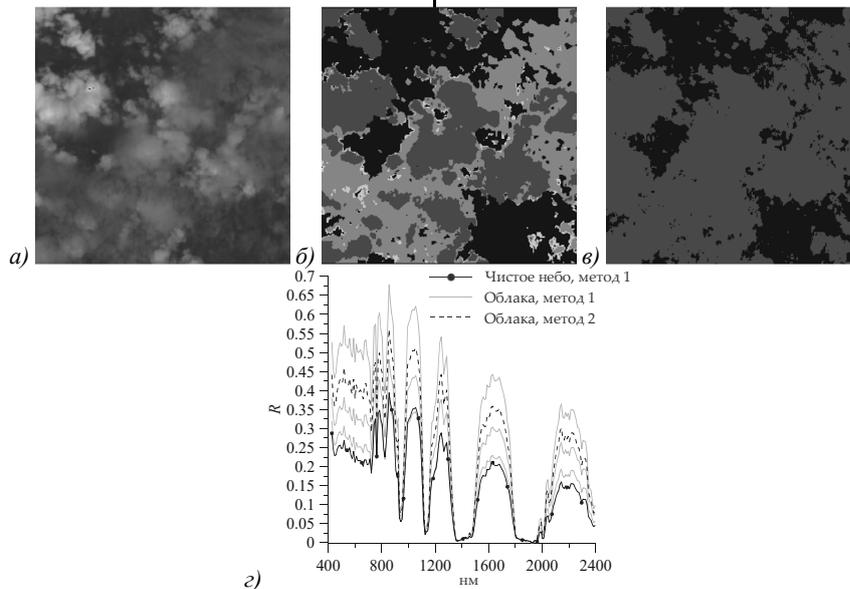


Рис. 6. Результаты обработки снимка 2

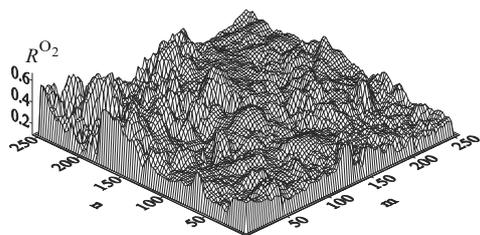


Рис. 7. Коэффициент яркости при $\lambda = \lambda_{rO_2}$ для снимка 2

При использовании алгоритма 1 число облачных объектов M определяется автоматически. В алгорит-

ме 2 все пиксели, определенные как «облачные», объединяются в одно облако. Разобьем это облако на M кластеров методом к-средних.

Результаты для фрагмента 1 снимка 1 и для снимка 2 представлены на рис. 14 и в табл. 4. Можно видеть, что распределение кластеров не отвечает распределению яркостей на RGB-изображении (ср. рис. 14 с рис. 2а, б, 6а, б). Разбросы полученных таким способом облачных объектов много больше, чем разбросы облачных объектов, найденных алгоритмом 1, см. табл. 4.

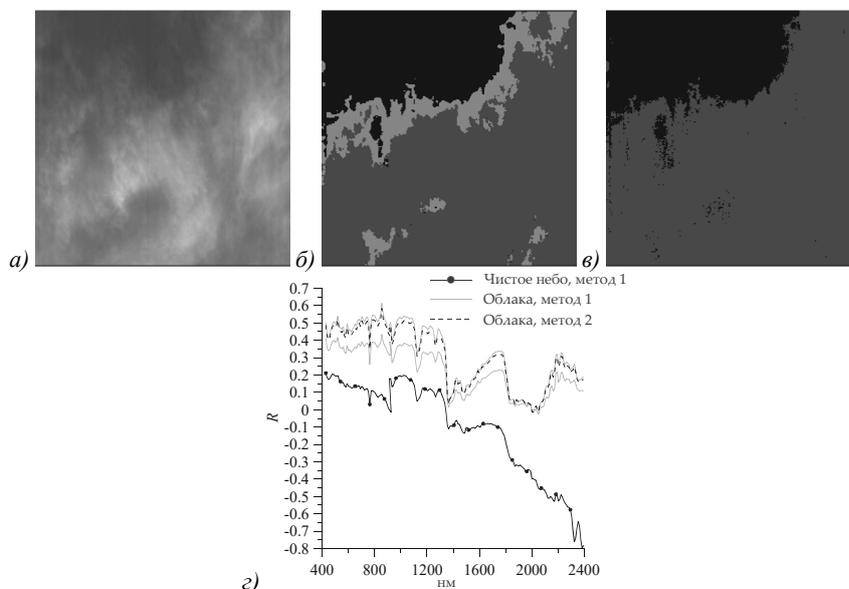


Рис. 8. Результаты обработки снимка 3

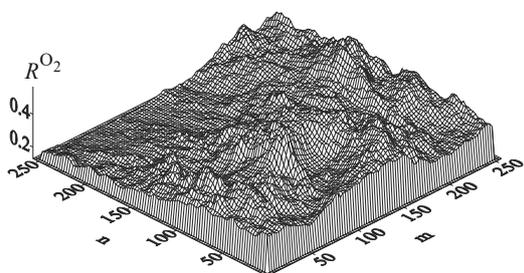


Рис. 9. Коэффициент яркости при $\lambda = \lambda_{rO_2}$ для снимка 3

Табл. 4. Разбросы облачных объектов

Снимок	Алгоритм 1	Алгоритм 2
1, фрагмент 1	0,34/0,23	0,63/0,58
2	0,21/0,11/0,06	0,28/0,34/0,3

Заключение

Предложен алгоритм обнаружения облачных объектов на многоспектральных снимках. В алгоритме используется тот факт, что среднее расстояние между двумя «облачными» спектрами снимка много меньше среднего расстояния между «облачным» и «чистым» спектрами того же снимка. Поэтому выделение каждого облачного объекта начинается с разделения пространственных пикселей снимка на два кластера методом к-средних. Далее из двух кластеров выбирается один, потенциально содержащий облако. Облачный

объект формируется из тех пикселей этого кластера, которые удовлетворяют пороговым спектральным критериям. Далее метод к-средних применяется к оставшимся пикселям и формируется следующий облачный объект. Процесс останавливается, если ни один из пикселей «облачного» кластера не удовлетворяет спектральным критериям. Таким образом, число облачных объектов определяется в процессе работы алгоритма. При этом фактически в алгоритме выполняется не только обнаружение, но и сегментация облаков (каждому сегменту отвечает один облачный объект).

Приведены результаты тестирования представленного алгоритма на снимках прибора HYPERION, сделанных в условиях частичной облачности над поверхностями разных типов (вода, растительность, пустынная местность, город, снег). Алгоритм определяет от 1 до 3 облачных объектов, расположение которых отвечает расположению облаков на соответствующих RGB-изображениях.

Помимо облачных масок, найденных представленным в работе алгоритмом 1, приведены маски, полученные попиксельным применением тех же пороговых критериев без предварительной кластеризации (алгоритм 2). Тогда в двух случаях (над океаном и над снегом) находятся излишне обширные

облака (в облако включается часть приоблачного аэрозоля). Таким образом, предварительная кластеризация позволяет более точно провести границу облака. Этот же результат мог быть достигнут более

тонкой настройкой значений порогов в спектральных критериях. Использование предварительной кластеризации позволяет обойтись без этой настройки.

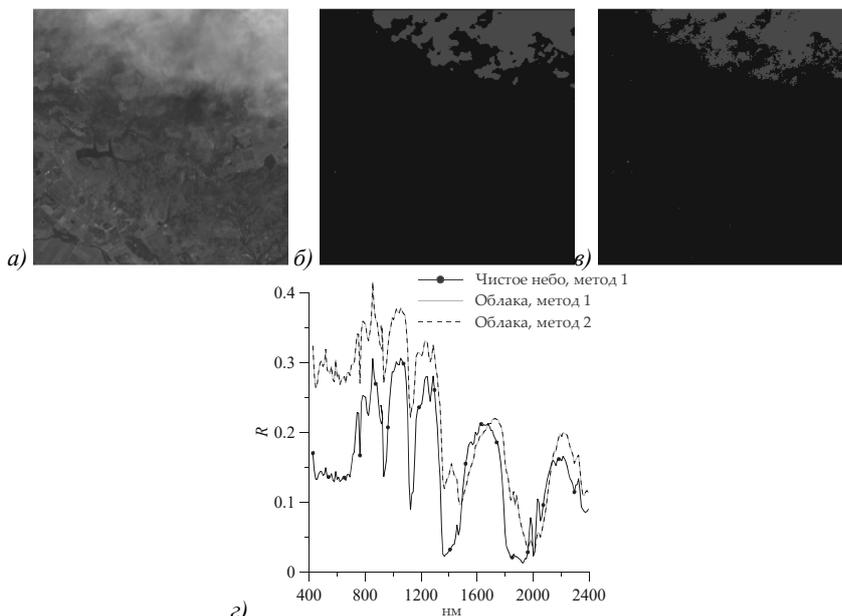


Рис. 10. Результаты обработки снимка 4

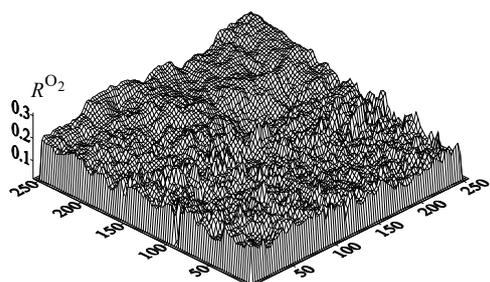


Рис. 11. Коэффициент яркости при $\lambda = \lambda_{r^O_2}$ для снимка 4

Найденное алгоритмом 2 облако может быть разделено методом к-средних на столько же кластеров, сколько облачных объектов найдено алгоритмом 1. Однако положение этих кластеров на снимке не отвечает распределению яркостей на RGB-изображении. Таким образом, кластеризация должна применяться ко всему снимку и разделять две сильно отличающиеся группы пикселей («облачные» и «чистые»), а не несколько групп близких друг к другу «облачных» пикселей.

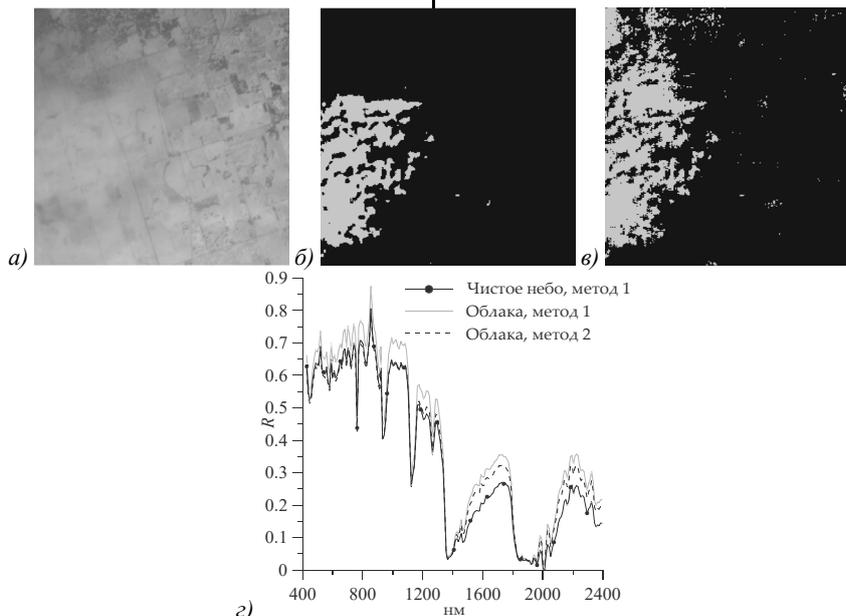


Рис. 12. Результаты обработки снимка 5

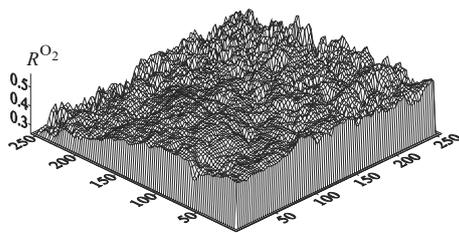


Рис. 13. Коэффициент яркости при $\lambda = \lambda_{O_2}$ для снимка 5

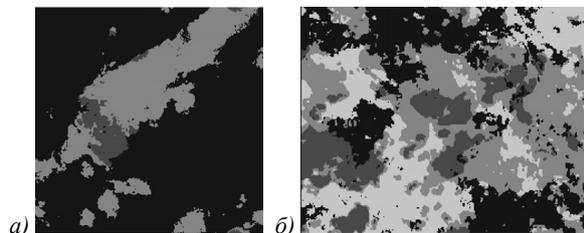


Рис. 14. Результаты кластеризации облака, построенного алгоритмом 2: (а) снимок 1, фрагмент 1, (б) снимок 2

Рассмотрены три показателя качества построенной облачной маски:

- отношение разброса спектров в «облачных» пикселях к разбросу спектров в «чистых» пикселях;
- отношение среднего коэффициента яркости в полосе поглощения кислорода в «облачных» пикселях к аналогичной величине в «чистых» пикселях;
- отношение вариации коэффициента яркости в видимом диапазоне в «облачных» пикселях к аналогичной величине в «чистых» пикселях.

По результатам тестирования предлагается использовать отношение разбросов.

Отметим, что представленный алгоритм применим только к многоспектральным снимкам, поскольку на монохроматическом снимке метод к-средних выделяет облачный кластер с большими ошибками [15]. Перед применением представленного алгоритма к снимкам иных многоспектральных приборов, помимо HYPERION, необходимо убедиться, что на снимке имеется достаточно спектральных каналов для выделения облачного кластера методом к-средних.

Очевидно, что качество облачной маски зависит от того, насколько хорошо выбраны спектральные критерии и значения порогов в них. Предметом дальнейших исследований является поиск новых спектральных критериев, особенно для обработки снимков, где облака расположены над снежным покровом.

Также предметом дальнейших исследований должен стать детальный статистический анализ рассмотренных показателей качества облачной маски с использованием других снимков как HYPERION, так и иных приборов. Актуален также поиск новых показателей качества.

References

[1] Taylor TE, O'Dell CW, O'Brien DM, Kikuchi N, Yokota T, Nakajima TY, Ishida H, Crisp D, Nakajima T. Compari-

son of cloud-screening methods applied to GOSAT near-infrared spectra. *IEEE Trans Geosci Remote Sens* 2012; 50(1): 295-309. DOI: 10.1109/TGRS.2011.2160270.

[2] Richter R. Atmospheric correction satellite data with haze removal including a haze/clear transition region. *Comput Geosci* 1996; 22(6): 675-681. DOI: 10.1016/0098-3004(96)00010-6.

[3] Ackerman S, Frey R, Strabala K, Liu Y, Gumley L, Baum B, Menzel P. Discriminating clear-sky from cloud with Modis. Algorithm theoretical basis document (MOD35). Madison: University of Wisconsin; 2010.

[4] Hall DK, Riggs GA, Solomonson VA, Barton JS, Casey K, Chien JYL, Digirolamo NE, Klein AG, Powell HW, Tait AB. Algorithm theoretical basis document for the Modis snow and sea ice-mapping algorithms. Madison: University of Wisconsin; 2001.

[5] Irish RR, Barker JL, Goward SN, Arvidson T. Characterization of the Landsat-7 ETM+ Automated Cloud-Cover Assessment (ACCA) algorithm. *Photogramm Eng Remote Sensing* 2006; 72(10): 1179-1188. DOI: 10.14358/PERS.72.10.1179.

[6] Griggin M, Burke H, Mandle D, Miller J. Cloud cover detection algorithm for EO-1 Hyperion imagery. *Proc IEEE International Geoscience and Remote Sensing Symposium (IGARSS)* 2003; 1: 86-89. DOI: 10.1109/IGARSS.2003.1293687.

[7] Lyapustin A, Wang Y, Frey R. An automatic cloud mask algorithm based on time series of MODIS measurements. *J Geophys Res* 2008; 113(D16): 207. DOI: 10.1029/2007JD009641.

[8] Volkova EV. Automatic estimation of cloud cover and precipitation parameters obtained by AVHRR NOAA for day and night conditions [In Russian]. *Current Problems in Remote Sensing of the Earth from Space* 2017; 14(5): 300-320. DOI: 10.21046/2070-7401-2017-14-5-300-320.

[9] Gao BC, Kaufman YJ, Wiscombe W. Removal of thin cirrus path radiances in the 0.4-1.0 μ m spectral region using the 1.375- μ m strong water vapor absorption channel. *Summaries of the Seventh JPL Airborne Earth Science Workshop* 1998; 1: 121-130.

[10] Gómez-Chova L, Camps-Valls G, Calpe-Maravilla J, Guanter L, Moreno J. Cloud-screening algorithm for ENVISAT/MERIS multispectral images. *IEEE Trans Geosci Remote Sens* 2007; 45(12): 4105-4118. DOI: 10.1109/TGRS.2007.905312.

[11] Thompson DR, Green RO, Keymeulen D, Lundeen SK, Mouradi Y, Nunes D, Castano R, Chien SA. Rapid spectral cloud screening onboard aircraft and spacecraft. *IEEE Trans Geosci Remote Sens* 2014; 52(11): 6779-6792. DOI: 10.1109/TGRS.2014.2302587.

[12] Korolev EE, Kochergin AM, Kuznetsov AE, Pobaruiev VI. Automatic segmentation of cloud objects on the high spatial resolution surface image [In Russian]. *Modern Problems of Science and Education* 2014; 5: 32-36.

[13] Trigo IF, Freitas SC, Barroso C, Macedi J. Gio global land component – Lot I "Operation of the global land component". Algorithm Theoretical Basis Document for cloud mask for LST retrieval. Copernicus; 2014.

[14] Belov AM, Denisova AY. Scene distortion detection algorithm using multitemporal remote sensing images. *Computer Optics* 2019; 43(5): 869-885. DOI: 10.18287/2412-6179-2019-43-5-869-885.

[15] Vetrov AA, Kuznetsov AE. Segmentation of cloud objects on panchromatic photos of earth surface [In Russian]. *Digit Signal Process* 2011; 3: 32-36.

[16] Sobolev VV. Light scattering in planetary atmospheres. New York: Pergamon Press; 1975.

-
- [17] Martins JV, Tanre D, Remer L, Kaufman Y, Mattoo S, Levy R. Modis cloud screening for remote sensing of aerosols over oceans using spatial variability. Geophys Res Lett 2002; 29(12): 1619-1622. DOI: [10.1029/2001GL013252](https://doi.org/10.1029/2001GL013252).
-

Сведения об авторе

Николаева Ольга Васильевна, 1971 года рождения, в 1993 году окончила факультет вычислительной математики и кибернетики Московского государственного университета им. М.В. Ломоносова. Старший научный сотрудник Института прикладной математики РАН. Область научных интересов: прямые и обратные задачи для уравнения переноса излучения.

ГРНТИ: 27.35.47

Поступила в редакцию 1 декабря 2021 г. Окончательный вариант – 16 февраля 2022 г.

Technique of detecting cloudy objects in multispectral images

O.V. Nikolaeva¹

¹*Keldysh Institute of Applied Mathematics RAS, Moscow*

Abstract

A multistep algorithm to detect cloudy objects in multispectral images is presented. Clustering spatial pixels by the k-means method and applying spectral criteria of cloudy/clear sky to fragments of obtained clusters are carried out in each step of the algorithm. One cloudy object is found in one step.

Results of testing the algorithm on images from a sensor HYPERION (199 non-zero spectral bands in a 425 nm–2400 nm interval under high spatial resolution of 30 m) are given. Images with discontinuous cloud cover above different surfaces (ocean, vegetation, desert, town, snow) are considered.

An alternative method, in which the same spectral criteria are applied to each pixel, is also used in testing. Cloud masks obtained by both algorithms are compared. Mean spectra of obtained cloudy objects are given. The presented algorithm finds 1-3 cloudy objects corresponding to the brightness distribution in RGB images. Using the alternative algorithm (without preliminary clustering) leads to detection errors on the cloud edges.

Three quality parameters are offered. The ratio of dispersion of “cloudy” spectra to dispersion of “clear” spectra is found to be most informative. This ratio should be much less than 1 when using a good cloudy mask.

Keywords: cloud detection, multispectral images, spectral criteria, quality parameters.

Citation: Nikolaeva OV. Technique of detecting cloudy objects in multispectral images. *Computer Optics* 2022; 46(5): 808-817. DOI: 10.18287/2412-6179-CO-1076.

Author's information

Olga V. Nikolaeva (b. 1971) graduated from the Computational Mathematic and Cybernetics department of the Moscow State University. She works as senior researcher in the Keldish Institute of Applied Mathematics RAS. Research interests are direct and inverse problems for radiation transport equation.

Received December 1, 2021. The final version – February 16, 2022.
