

## Статистическая модель текстурных признаков перистой облачности по спутниковым снимкам MODIS

В.Г. Астафуров<sup>1,2</sup>, Т.В. Евсюткин<sup>1</sup>, К.В. Курьянович<sup>1</sup>,  
А.В. Скороходов<sup>1\*</sup>

<sup>1</sup>Институт оптики атмосферы им. В.Е. Зуева СО РАН  
634021, г. Томск, пл. Академика Зуева, 1

<sup>2</sup>Томский государственный университет систем управления и радиоэлектроники  
634050, г. Томск, пр. Ленина, 40

Поступила в редакцию 2.02.2014 г.

Сформирована система информативных текстурных признаков для изображений различных типов перистой облачности на спутниковых снимках MODIS с разрешением 250 м. Для описания текстуры применяются методы: Gray-Level Co-occurrences Matrix, Gray-Level Difference Vector и Sum and Difference Histograms. Определены законы распределения, описывающие флуктуации текстурных признаков, и найдены оценки их параметров. Приводятся и обсуждаются результаты классификации подтипов перистой облачности с использованием нейросетевых технологий.

**Ключевые слова:** перистые облака, текстурные признаки, статистическая модель, классификация, спутниковые данные; cirrus clouds, textural features, statistical model, classification, satellite data.

### Введение

В настоящее время существует значительное число публикаций по автоматической классификации облаков на основе данных дистанционного зондирования Земли из космоса, среди которых можно выделить работы [1–4], позволяющие получать наиболее полную информацию о типах облачности согласно действующему метеорологическому стандарту [5]. Большой интерес представляет перистая облачность. Современные климатические модели иллюстрируют ее значительное влияние на перенос солнечной и земной радиации и распределение температуры в системе «атмосфера – подстилающая поверхность» [6]. Важной является проблема удаления полупрозрачной перистой облачности при атмосферной коррекции спутниковых данных [7]. В стандарте [5] выделено 4 типа перистой облачности: когтевидные (*Ci unc*), хребтовидные и перепутанные (*Ci vert & Ci int*), образовавшиеся из наковален кучево-дождевых облаков (*Ci ing*), и хлопьевидные (*Ci floc*). Информацию о наличии перистой облачности (в том числе и полупрозрачной) и некоторых ее характеристиках (температура и высота верхней границы, коэффициент отражения, водность, эффективный радиус частиц, оптическая толщина и др.) можно получить

из тематических продуктов MODIS: MOD06\_L2 и MOD35\_L2. При этом вопросы ее классификации по подтипам в известных нам работах не рассмотрены.

Для описания изображений разновидностей перистой облачности по спутниковым снимкам целесообразно привлекать информацию о текстуре изображений [8], что позволяет представить ее как единую целостную структуру. Под текстурой понимается некоторым образом организованный локальный участок изображения, обладающий однородными статистическими характеристиками. В настоящее время существуют различные подходы к описанию текстуры, каждый из которых оперирует собственным набором текстурных признаков (ТП). При этом ТП не являются одинаково информативными для описания различных изображений. Для описания текстуры разновидностей перистой облачности необходимо сформировать набор информативных ТП.

Целью данной работы является исследование флуктуаций текстурных признаков и построение их статистической модели для различных подтипов перистой облачности по данным спектрорадиометра MODIS с пространственным разрешением 250 м.

### Наборы эталонных изображений разновидностей перистой облачности

Поиск характерных изображений текстур различных подтипов перистых облаков осуществляется

\* Владимир Глебович Астафуров (astafurov@iao.ru); Тимофей Викторович Евсюткин (oomphi.iao@gmail.com); Ксения Викторовна Курьянович (ksuyain@mail.ru); Алексей Викторович Скороходов (vazime@yandex.ru).

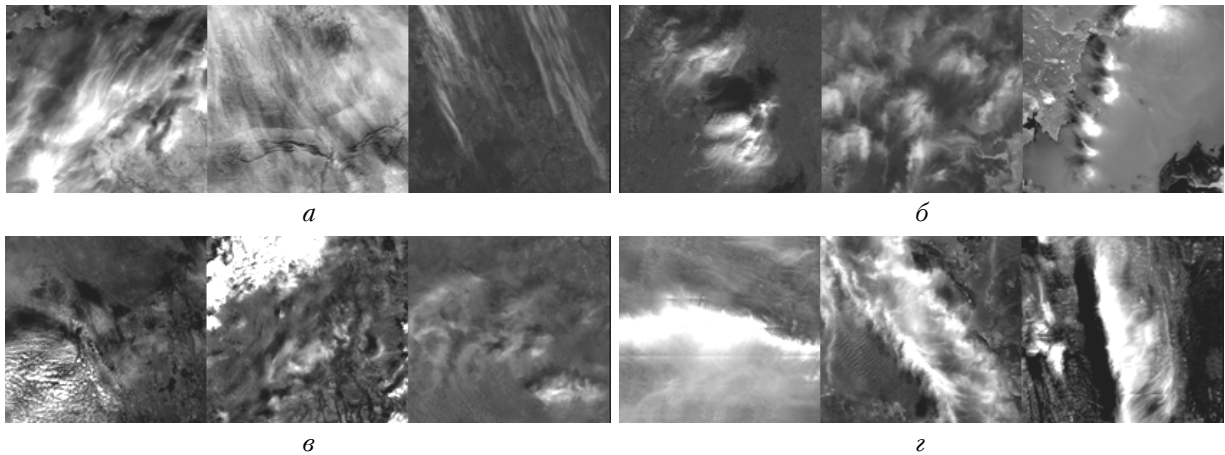


Рис. 1. Фрагменты изображений четырех подтипов перистой облачности: *Ci unc* (а); *Ci floc* (б); *Ci ing* (в); *Ci vert & Ci int* (г)

путем сопоставления архивных данных сети наземных метеостанций со спутниковой съемкой MODIS (<http://modis.gsfc.nasa.gov>). Использовались данные 1-го спектрального канала MODIS (0,62–0,67 мкм) с пространственным разрешением 250 м в дневное время и при отсутствии снежного покрова, наличие которого существенно затрудняет не только классификацию облачности (в том числе и перистой), но и ее идентификацию из-за сходства отражательной способности снега и облаков. Информация о типах облачности извлекалась с сервера «Погода России» (<http://meteo.infospace.ru>). Значительная часть данных получена со следующих метеостанций: Аделаида (Австралия), Алматы (Казахстан), Антофагаста (Чили), Бетлехем (ЮАР), Долорес (Аргентина), Исламабад (Пакистан), Кроун Поинт (Тринидад и Тобаго), острова Санта Круз (Соломоновы острова), Томск (Россия). По результатам исследований сформирован набор характерных изображений четырех разновидностей перистой облачности, примеры фрагментов которых размером 250×250 пикселей показаны на рис. 1.

## Информативные текстурные признаки перистой облачности

Рассматриваются следующие методы описания и анализа текстуры.

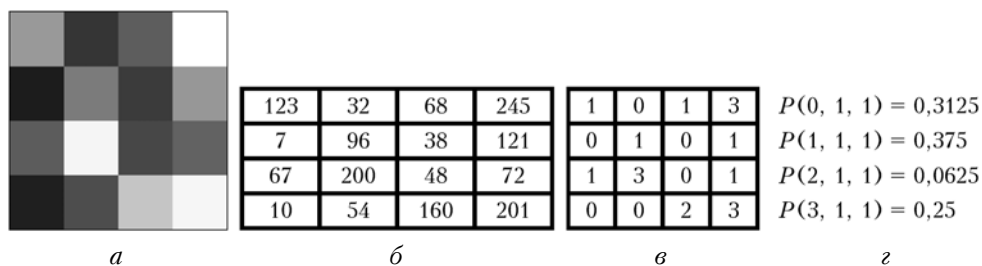


Рис. 2. Пример вычислений оценок функции плотности вероятности: исходный фрагмент изображения (а); яркость пикселей (б); квантованные значения яркости (в); оценки вероятностей для  $\Delta x = 1$  и  $\Delta y = 1$  (г)

1. Gray-Level Difference Vector (GLDV), или метод вектора разности уровней яркости, основанный на вычислении оценок функции плотности вероятности  $P_{\delta}(i, \Delta x, \Delta y)$  значений абсолютных разностей уровней яркости пар элементов изображения (пикселей), отстоящих друг от друга на расстоянии  $\delta(\Delta x, \Delta y)$ :

$$P_{\delta}(i, \Delta x, \Delta y) = \frac{\#\{|I_{x,y} - I_{x+\Delta x, y+\Delta y}| = i\}}{N_x N_y},$$

где символ # — обозначает число элементов множества, удовлетворяющих условию в фигурных скобках;  $x$  и  $y$  — координаты элементов изображения;  $0 < x < N_x$ ,  $0 < y < N_y$ ;  $\Delta x$  и  $\Delta y$  — величины смещений;  $i = 0, \dots, m - 1$ ,  $m$  — число уровней квантования яркости изображения [9]. Последовательность вычислений оценок вероятностей при  $m = 4$  приведена на рис. 2.

2) Gray-Level Co-occurrence Matrix (GLCM), или метод матрицы смежности, элементами которой являются относительные частоты  $P_{ij}$  наличия на изображении соседних точек с яркостями  $I_i$  и  $I_j$ , расположенных на расстоянии  $d$  друг от друга в одном из четырех угловых направлений  $\varphi = 0, 45, 90, 135^\circ$  [10]. Последовательность вычислений матриц смежности для четырех угловых направлений при  $m = 4$  и  $d = 1$  приведена на рис. 3.

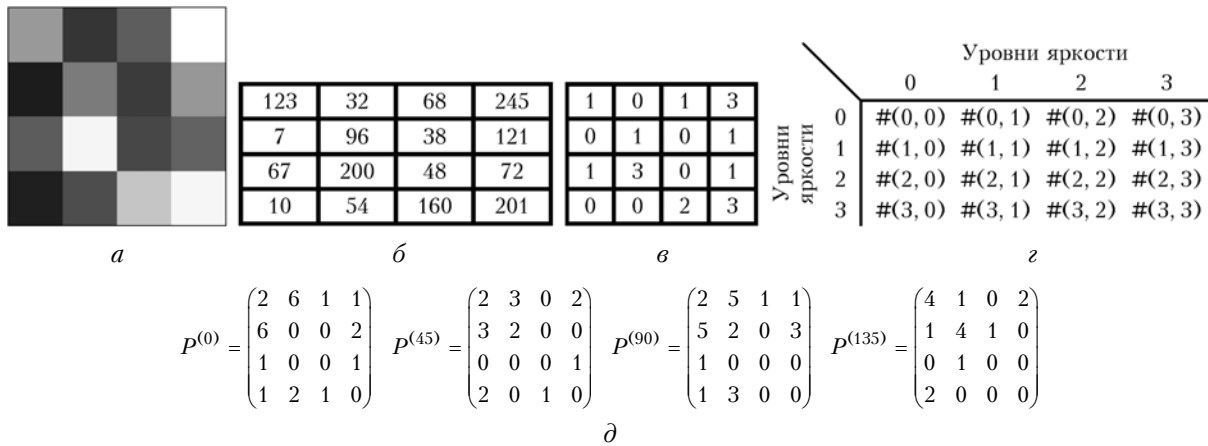


Рис. 3. Пример вычислений матриц смежности: исходный фрагмент изображения (а); яркость пикселей (б); квантованные значения яркости (в); общий вид матриц смежности (г) (здесь # – число смежных точек с яркостями  $I_i$  и  $I_j$  на квантованном изображении); матрицы смежности для угловых направлений 0, 45, 90 и 135° (д)

3) Sum and Difference Histograms (SADH), или метод, основанный на вычислении локальных свойств гистограмм суммы  $P_s(i)$  и разности  $P_d(j)$  уровней яркости пар элементов изображения:

$$P_s(i) = \frac{\#\{(x, y) \in D, s_{x,y} = i\}}{N_x N_y},$$

$$P_d(j) = \frac{\#\{(x, y) \in D, d_{x,y} = j\}}{N_x N_y},$$

где  $i = 2, \dots, 2m$ ;  $j = -m + 1, \dots, m - 1$ ,  $m$  – число уровней квантования яркости изображения. При этом сумма  $s_{x,y}(\Delta x, \Delta y) = I_{x,y} + I_{x+\Delta x, y+\Delta y}$  и разность  $d_{x,y}(\Delta x, \Delta y) = I_{x,y} - I_{x+\Delta x, y+\Delta y}$  элементов изображения зависят от величины относительных смещений  $\delta_s(\Delta x, \Delta y)$  и  $\delta_d(\Delta x, \Delta y)$  соответственно [11]. Схема вычислений оценок функции плотности вероятности при  $m = 4$  приведена на рис. 4.

Для формирования системы информативных ТП использовался метод Add [12], который состоит в циклическом переборе всех исследуемых ТП и последовательном добавлении наиболее эффективных из них к набору информативных признаков. В качестве классификатора в методе Add использовалась самоорганизующаяся нейронная сеть Кохонена с числом нейронов, равным четырем по числу подтипов перистых

облаков. Для обучения сети использовался алгоритм «победитель забирает все с механизмом утомления», который является наиболее эффективным среди известных алгоритмов самоорганизации с точки зрения быстрейшего действия и качества обучения сетей Кохонена [13]. Обучающая выборка содержит по 100 образцов текстуры для каждого типа облачности размером  $21 \times 21$  пиксель, а тестовая выборка – 200 фрагментов для *Ci ing* и по 500 для остальных трех подтипов.

Проанализирована исходная система из 181 ТП, рассчитываемых методами GLCM, GLDV и SADH [9–11] для различных угловых направлений  $\varphi = 0, 45, 90, 135^\circ$ , расстояния  $d = 1$ , числа уровней квантования  $m = 20$  и различных значений смещений  $\Delta x = 0, 1$  и  $\Delta y = 0, 1$ . Выбор числа уровней квантования для задачи классификации облачности по типам был основан на результатах наших исследований [14]. Итоговая система информативных ТП перистой облачности включает в себя:

- дисперсию (GLCM) при  $\varphi = 45^\circ$ ,
- информационную меру корреляции – 1 (GLCM) при  $\varphi = 0^\circ$ ,
- среднее (SADH) при  $\Delta x = 0$  и  $\Delta y = 1$ ,
- дисперсию (SADH) при  $\Delta x = 0$  и  $\Delta y = 1$ .

Следует отметить, что метод Add может использоваться совместно с любым другим классификатором. При этом полученные наборы информативных признаков могут отличаться друг от друга.

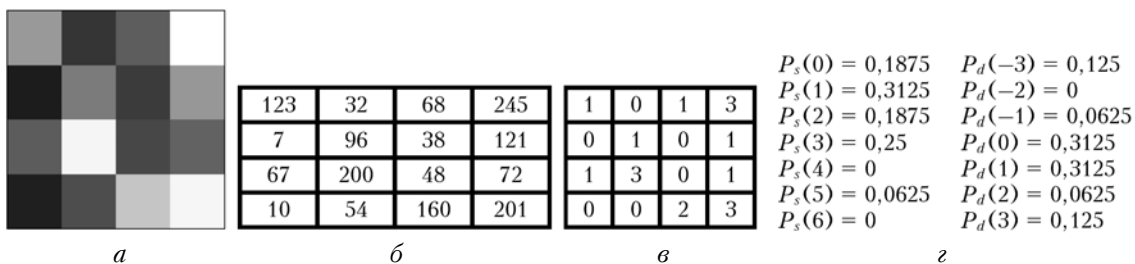


Рис. 4. Пример вычислений оценок функции плотности вероятности: исходный фрагмент (а) изображения; яркость пикселей (б); квантованные значения яркости (в); оценки функции вероятности для  $\Delta x = 1$  и  $\Delta y = 1$  (г)

## Статистическая модель текстуры изображений перистой облачности

Для сформированного набора информативных ТП построена статистическая модель текстуры. Для каждого подтипа перистой облачности по наборам фрагментов их характерных изображений составлены выборки значений всех информативных ТП и построены гистограммы. С помощью критерия Колмогорова–Смирнова [15] для каждой гистограммы подбирались плотность вероятности и вычислялись оценки ее параметров. Критерий Колмогорова–Смирнова основан на мере расхождения  $D_n = \sup_{|x| < \infty} |F_n(x) - \Phi(x)|$  между эмпирической  $F_n(x)$  и теоретической  $F_n(x)$  функциями распределения. Здесь  $\sup(\cdot)$  обозначает точную верхнюю границу. Вид распределения определялся исходя из минимума величины  $D_n$ .

В результате исследований для описания гистограмм информативных ТП подтипов перистой облачности подобрано 6 распределений, плотности вероятности которых приведены в табл. 1, а их подробное описание можно найти в [15, 16]. В табл. 2 приведены параметры распределений для различных подтипов перистой облачности.

На рис. 5 показаны графики подобранных плотностей вероятностей и гистограммы масштабированных значений ТП:

$$x = \frac{T - T_{\min}}{T_{\max} - T_{\min}},$$

где  $T$  – текущее значение ТП,  $T_{\max}$  и  $T_{\min}$  – его максимальное и минимальное значения на исследуемых фрагментах изображений.

Т а б л и ц а 1

Статистические распределения, используемые в модели перистой облачности	
Распределение	Плотность вероятности
Дагума (Dagum)	$f(x) = \frac{\alpha k (x/\beta)^{\alpha k - 1}}{\beta (1 + (x/\beta)^{-\alpha})^{-k}}$
Обобщенное распределение экстремальных значений	$f(x) = \begin{cases} \frac{1}{\sigma} \exp(-(1+kz)^{-1/k}) (1+kz)^{-1-1/k}, & k \neq 0 \\ \frac{1}{\sigma} \exp(-z - \exp(-z)), & k = 0 \end{cases}, \text{ где } z = \frac{x - \mu}{\sigma}$
Обобщенное логистическое	$f(x) = \begin{cases} \frac{(1+kz)^{-1-1/k}}{\sigma (1 + (1+kz)^{-1/k})^2}, & k \neq 0 \\ \frac{\exp(-z)}{\sigma (1 + \exp(-z))^2}, & k = 0 \end{cases}, \text{ где } z = \frac{x - \mu}{\sigma}$
SB-распределение Джонсона	$f(x) = \frac{\delta}{\sigma \sqrt{2\pi z(1-z)}} \exp\left(-\frac{1}{2}\left(\gamma + \delta \ln\left(\frac{z}{1-z}\right)\right)^2\right), \text{ где } z = \frac{x - \mu}{\sigma}$
Перта (Pert)	$f(x) = \frac{1}{B(\alpha_1, \alpha_2)} \frac{(x-a)^{\alpha_1-1} (b-x)^{\alpha_2-1}}{(b-a)^{\alpha_1+\alpha_2-1}}, \text{ где } B - \text{ бета-функция};$ $\alpha_1 = \frac{4m+b-5a}{b-a}, \quad \alpha_2 = \frac{5b-a-4m}{b-a}$
Вакеби (Wakeby)	$f(x) = \frac{(1-F(x))^{\delta+1}}{\alpha(1-F(x))^{\beta+\delta} + \gamma},$ $x = \xi + \frac{\alpha}{\beta} \left(1 - (1-F(x))^\beta\right) - \frac{\gamma}{\delta} \left(1 - (1-F(x))^{-\delta}\right),$ где $F(x) = \int_{-\infty}^x f(t) dt$

Т а б л и ц а 2

Статистическая модель текстуры подтипов перистой облачности			
Тип	Текстурный признак	Распределение	Оценки параметров распределения
1	2	3	4
Ci unc	Дисперсия (GLCM)	Вакеби	$\alpha = -0,26; \beta = 1,9; \gamma = 0,336;$ $\delta = -0,17; \xi = 0$
	Информационная мера корреляции – 1 (GLCM)	Вакеби	$\alpha = 1,3; \beta = 320; \gamma = 0,18;$ $\delta = 0,050; \xi = 0$

1	2	3	4
<i>Ci unc</i>	Среднее (SADH)	Вакеби	$\alpha = 0,64; \beta = 7,5; \gamma = 0,37;$ $\delta = -0,31; \xi = -0,017$
	Дисперсия (SADH)	Обобщенное логистическое	$k = -0,14; \sigma = 0,082;$ $\mu = 0,73$
<i>Ci floc</i>	Дисперсия (GLCM)	Вакеби	$\alpha = -0,14; \beta = 3,6; \gamma = 0,34;$ $\delta = -0,25; \xi = 0,0077$
	Информационная мера корреляции - 1 (GLCM)	Вакеби	$\alpha = 0,51; \beta = 13; \gamma = 0,10;$ $\delta = 0,12; \xi = -0,002$
	Среднее (SADH)	SB-распределение Джонсона	$\gamma = 0,56; \delta = 0,96; \sigma = 1,1;$ $\mu = -0,018$
<i>Ci ing</i>	Дисперсия (SADH)	Дагума	$k = 0,32; \alpha = 9,5; \beta = 0,59$
	Дисперсия (GLCM)	Дагума	$k = 0,29; \alpha = 3,4; \beta = 0,41$
	Информационная мера корреляции - 1 (GLCM)	Вакеби	$\alpha = 0,28; \beta = 0,66; \gamma = 0,006;$ $\delta = 0,72; \xi = -0,003$
	Среднее (SADH)	Перга	$m = 0,32; a = -0,012; b = 1,2$
	Дисперсия (SADH)	Обобщенное распределение экстремальных значений	$k = -0,39; \sigma = 0,23;$ $\mu = 0,46$
<i>Ci vert &amp; Ci int</i>	Дисперсия (GLCM)	SB-распределение Джонсона	$\gamma = 1,0; \delta = 0,69; \sigma = 1,1;$ $\mu = -0,019$
	Информационная мера корреляции - 1 (GLCM)	Дагума	$k = 0,15; \alpha = 4,0; \beta = 0,45$
	Среднее (SADH)	Вакеби	$\alpha = 1,8; \beta = 14; \gamma = 0,51;$ $\delta = -0,56; \xi = -0,014$
	Дисперсия (SADH)	SB-распределение Джонсона	$\gamma = 0,13; \delta = 1,4; \sigma = 1,34;$ $\mu = -0,20$

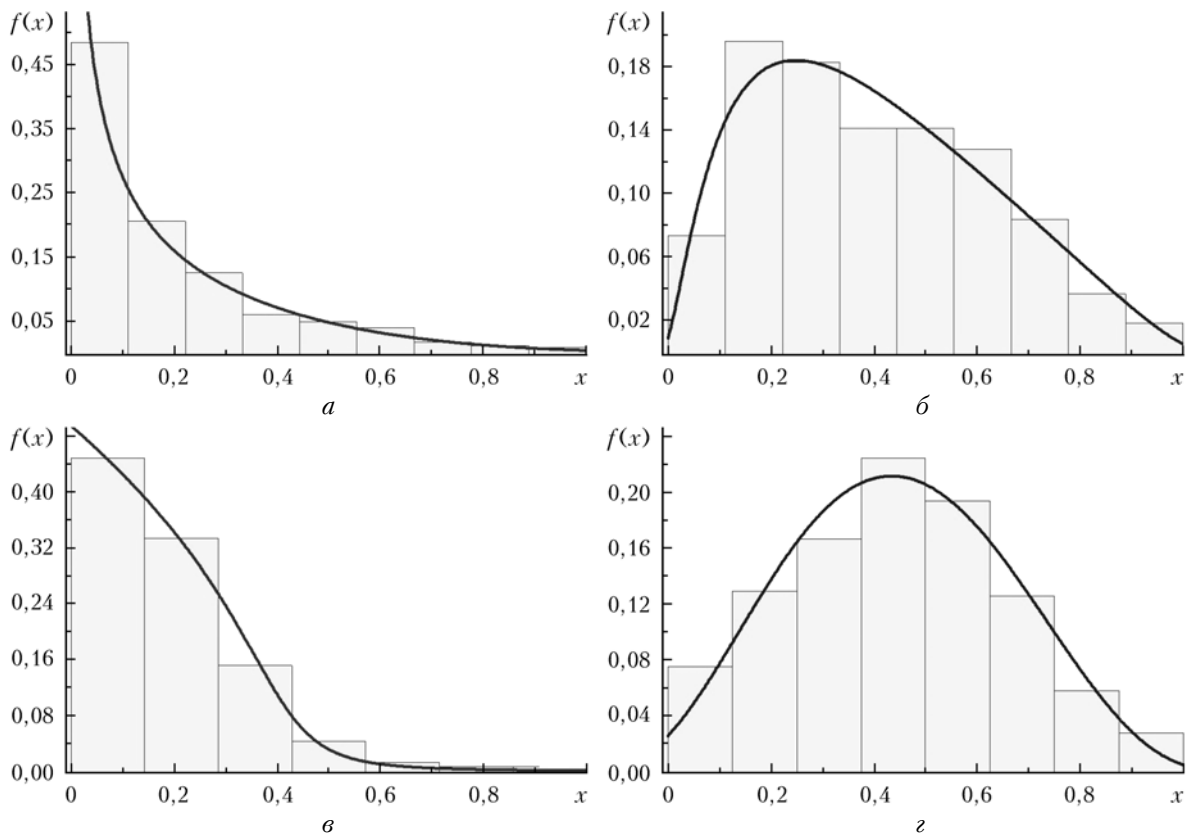


Рис. 5. Распределение Вакеби значений признака «дисперсия» (GLCM) для *Ci unc* (а); SB-распределение Джонсона значений признака «среднее» (SADH) для *Ci floc* (б); обобщенное распределение экстремальных значений признака «дисперсия» (SADH) для *Ci ing* (в) и распределение Дагума значений признака «информационная мера корреляции» - 1 (GLCM) для *Ci vert & Ci int* (г)

## Результаты классификации перистой облачности

Оценка эффективности сформированного набора информативных ТП проводилась путем классификации тестовой выборки. В качестве классификатора использовалась самоорганизующаяся сеть Кохонена, которая была включена в процедуру Add при формировании системы информативных ТП. Оценка достоверности результатов классификации тестовых образцов вычислялась по формуле

$$E_i = N_R^{(i)} / N_T^{(i)},$$

где  $N_R^{(i)}$  – количество правильно проклассифицированных тестовых образцов  $i$ -го подтипа перистой облачности при  $i = \overline{1, 4}$ , а  $N_T^{(i)}$  – число тестовых образцов для  $i$ -го подтипа перистой облачности. Значения  $E_i$  приведены в табл. 3.

Таблица 3  
Достоверность результатов классификации тестовой выборки сетью Кохонена

Тип облачности	$Ci unc$	$Ci flo$	$Ci ing$	$Ci vert$ & $Ci int$
$E_i$	0,85	0,53	0,39	0,12

Низкую эффективность классификации  $Ci vert$  &  $Ci int$  можно объяснить значительным сходством

ее текстуры с другими разновидностями перистой облачности (особенно  $Ci unc$ ) на тестовых фрагментах. Также следует отметить, что уникальность текстуры подтипа  $Ci vert$  &  $Ci int$  (см. рис. 1,  $z$ ) заключается в том, что они организованы в вытянутую структуру с повторяющимся элементом текстуры (хребет), характерный размер которых превышает величину тестовых фрагментов (21 пиксель). Одним из возможных вариантов увеличения эффективности классификации  $Ci vert$  &  $Ci int$  является поиск на спутниковых снимках серий текстур этого типа, ориентированных в одном направлении.

На рис. 6 показаны результаты классификации перистой облачности по данным MODIS и зафиксированной метеостанциями Бетлехем (28°15' ю.ш., 28°20' в.д.) над территорией ЮАР 23.05.2011 г. (рис. 6,  $a$ ) и Аделаида (34°56' ю.ш., 138°31' в.д.) над побережьем Австралии 30.09.2008 г. (рис. 6,  $b$ ). Аббревиатура «Nc» обозначает неклассифицированные участки изображения, к которым могут относиться подстилающая поверхность или другие типы облачности. Процедура классификации исходных изображений проводилась путем их сканирования окном размером 21 × 21 пиксель. Центральный пиксель проклассифицированного фрагмента изображения перекрашивался в определенный цвет в соответствии с выбранной легендой. Полученные результаты иллюстрируют значительную неоднородность перистой облачности в пределах отдельно взятых фрагментов спутниковых снимков.

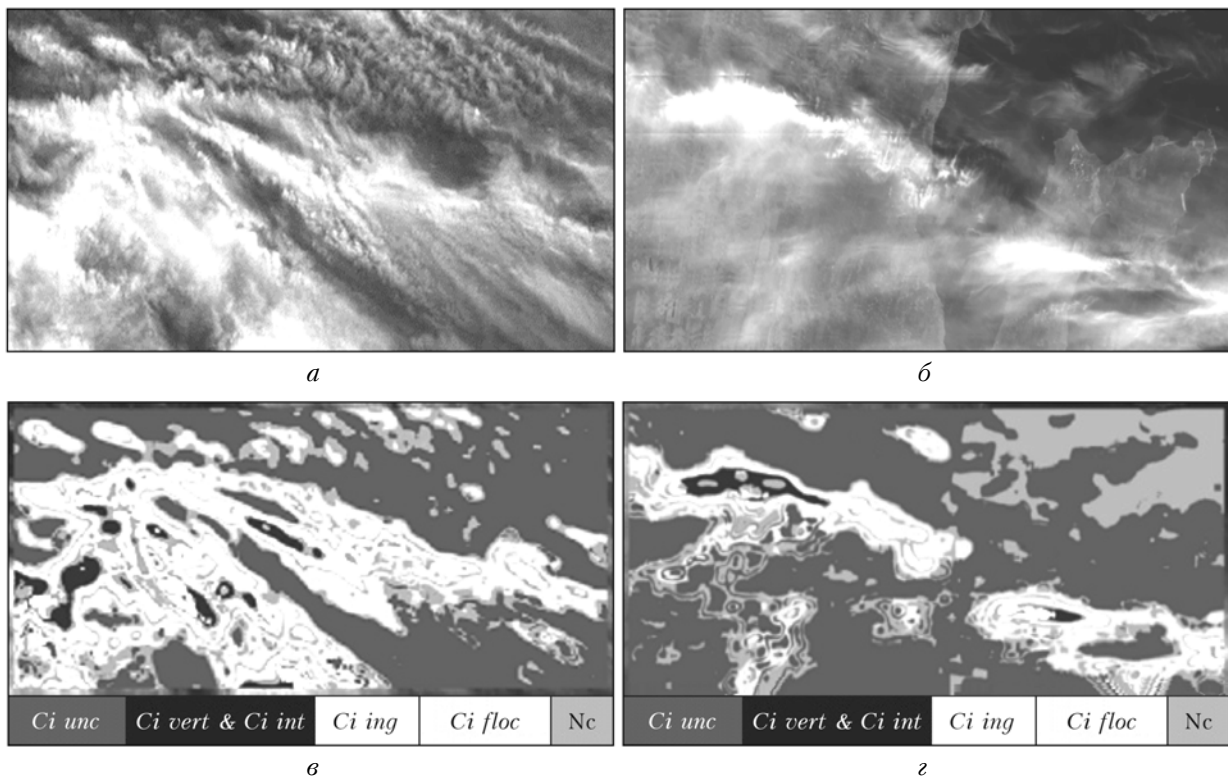


Рис. 6. Результаты классификации фрагментов изображений подтипов перистой облачности ( $a$ ,  $b$ ), самоорганизующейся сетью Кохонена ( $v$ ,  $z$ )

## Обсуждение результатов

Основным результатом работы является статистическая модель флуктуаций текстуры изображений подтипов перистой облачности на основе спутниковых снимков MODIS с пространственным разрешением 250 м, которую можно рассматривать как дополнение к существующим моделям облачных полей. Данную модель можно применять также для разработки более совершенных методов классификации перистой облачности, например для задания функций фазификации (переход к нечеткости) в нечетком классификаторе, использование которого может позволить не только детектировать многослойную облачность, но и определять тип облаков каждого яруса. Полученные результаты по классификации перистой облачности на основе простейшей архитектуры самоорганизующихся сетей Кохонена являются обнадеживающими и позволяют сделать выводы о возможности классификации облачности верхнего яруса по подтипам, а также построения более эффективных систем с привлечением дополнительной информации, получаемой по спутниковым данным.

Работа выполнена при частичной финансовой поддержке РФФИ в рамках научного проекта № 14-07-31018 мол\_а.

1. *Bankert R.L., Mitrescu C., Miller S.W., Wade R.H.* Comparison of GOES cloud classification algorithms employing explicit and implicit physics // *J. Appl. Meteorol. Climatol.* 2009. V. 48. P. 1411–1421.
2. *Lafont D., Jourdan O., Guillement B.* Mesoscale cloud pattern classification over ocean with a neural network using a new index of cloud variability // *Int. J. Remote Sens.* 2006. V. 27. P. 3533–3552.
3. *Астафуров В.Г., Скороходов А.В.* Нейросетевой классификатор облачности по спутниковым данным // *Информ. технол.* 2012. № 7. С. 32–37.
4. *Волкова Е.В., Успенский А.Б.* Оценки параметров облачного покрова по данным геостационарного МИСЗ METEOSAT-9 круглосуточно в автоматическом режиме // *Современные проблемы дистанционного зондирования Земли из космоса.* 2010. Т. 7, № 3. С. 65–73.
5. *Код* для оперативной передачи данных приземных гидрометеорологических наблюдений с сети станций Госкомгидромета СССР, расположенных на суше (включая береговые станции) КН-01, национальный вариант международного кода FM-12-IX SYNOP. Л.: Гидрометеоздат, 1989. 64 с.
6. *Baran A.* A review of the light scattering properties of cirrus // *J. Quant. Spectrosc. Radiat. Transfer.* 2009. V. 110. P. 1239–1260.
7. *Соломатов Д.В., Афонин С.В., Белов В.В.* Построение облачной маски и удаление полупрозрачной облачности на спутниковых снимках ETM+/Landsat-7 // *Оптика атмосф. и океана.* 2013. Т. 26, № 12. С. 798–803.
8. *Гонсалес Р., Вудс Р.* Цифровая обработка изображений. М.: Техносфера, 2005. 1072 с.
9. *Weszka J.S., Dyer C.R., Rosenfeld A.* A comparative study of texture measures for terrain classification // *IEEE Transaction on Systems, Man and Cybernetics.* April 1976. V. SMC-6, N 4. P. 269–285.
10. *Haralick R.M., Shanmugam K., Dinstein I.* Textural features for image classification // *IEEE Transactions on Systems, Man and Cybernetics.* November 1973. V. SMC-3, N 6. P. 610–621.
11. *Unser M.* Sum and difference histograms for texture classification // *IEEE Transaction on Systems, Pattern Analysis and Machine Intelligence.* January 1986. V. PAMI-8, N 1. P. 118–125.
12. *Загоруйко Н.Г.* Прикладные методы анализа данных и знаний. Новосибирск: ИМ СО РАН, 1999. 270 с.
13. *Осовский С.* Нейронные сети для обработки информации / Пер. с польск. И.Д. Рудинского. М.: Финансы и статистика, 2002. 344 с.
14. *Астафуров В.Г., Скороходов А.В.* Сегментация спутниковых снимков облачности по текстурным признакам на основе нейросетевых технологий // *Исслед. Земли из космоса.* 2011. № 6. С. 10–20.
15. *Кобзарь А.И.* Прикладная математическая статистика. Для инженеров и научных работников. М.: Физматлит, 2006. 816 с.
16. *MathWave* [Electronic resource]: EasyFit – Easily Fit Distributions to Your Data! Electronic data. – Dnepropetrovsk, 2004–2014. – URL: <http://www.mathwave.com/help/easyfit/index.html> (Accessed 12.02.2014).

### *V.G. Astafurov, T.V. Evsyutkin, K.V. Kuriyanovich, A.V. Skorokhodov. Statistical model of cirrus cloud textural features based on MODIS satellite images.*

A set of informative textural features is formed for various cirrus cloud patterns presented in MODIS satellite images with 250-m resolution. To describe the texture, the following methods are applied: Gray-Level Co-occurrences Matrix, Gray-Level Difference Vector, and Sum and Difference Histograms. Laws are determined that describe textural features' fluctuations, and assessments of their parameters are presented as well. Results of cirrus cloud subtype classification using neural network technologies are presented and discussed.