

Ю.В. Гриднев, К.Т. Протасов

СЕГМЕНТАЦИЯ МНОГОМЕРНЫХ ИЗОБРАЖЕНИЙ АЛГОРИТМОМ КЛАСТЕРИЗАЦИИ, ОСНОВАННЫМ НА ДВУСТОРОННЕМ КРИТЕРИИ ОДНОРОДНОСТИ

Для решения задач тематического дешифрирования аэрокосмических видеоданных, полученных <наблюдением> подстилающей поверхности Земли и облачности, предложен алгоритм автоматической классификации, основанный на двустороннем критерии однородности многомерных выборок с использованием <расстояния> Колмогорова–Смирнова. Особенность построенного алгоритма заключается в его возможности работать с набором <согласованных> изображений многоспектральных каналов видимого диапазона, ИК-, СВЧ-диапазонов электромагнитного спектра.

При решении задач природопользования и климато-экологического мониторинга основным источником оперативной информации являются аэрокосмические изображения подстилающей поверхности Земли и облачности, получаемые одновременной регистрацией данных многоспектральных каналов видимого участка спектра, ИК-, СВЧ-диапазонов электромагнитного излучения. Последующая обработка многомерных наблюдений включает этап тематического дешифрирования, заключающийся в выделении текстурно однородных участков изображений алгоритмом сегментации. Для построения процедуры сегментации естественно воспользоваться алгоритмами автоматической классификации. При этом необходимо, во-первых, сформулировать текстурные признаки фрагментов изображений, во-вторых, ввести меру адекватности анализируемых текстур [1, 4].

Особенность рассматриваемого далее алгоритма автоматической классификации заключается в том, что в данном случае анализируется совокупность согласованных по масштабам многомерных изображений. Несмотря на простоту выявления фрагментов изображения с однотипными текстурными свойствами, осуществляемую дешифровальщиком-оператором, до настоящего времени не существует универсального формального определения текстуры. Известные частные определения отражают лишь эвристические подходы формирования текстурно-признаковых пространств и методов их автоматической классификации [2, 3]. Текстура проявляется как пространственная характеристика изображения, и исчерпывающим статистическим описанием текстуры является соответствующий функционал плотности вероятностей (для непрерывных наблюдений), априорно задать который не представляется возможным. Поэтому оправдано использование более простых моделей для описания статистических свойств оцифрованных полей радиояркостей, например, с помощью гистограмм распределения яркостей, матриц межэлементной связности, регрессионных и автокорреляционных функций [2, 3, 4].

Вначале рассмотрим задачу выделения контурных линий на изображениях, с тем чтобы на этапе текстурного анализа исключить эти участки изображений как <мешающие>. Будем полагать, что подлежащие анализу изображения оцифрованы, согласованы по масштабам, нормированы по яркости, так что цифровое представление имеет вид трехмерной пространственной матрицы чисел $\{z^{ijk}\}$ размера $M \times N \times K$, где z^{ijk} – оцифрованное значение яркости для точки (пиксела) с координатами (i, j) плоскости наблюдаемого изображения формата $M \times N$, а k – номер спектрального канала в наборе изображений видимого, ИК-, СВЧ-диапазонов, $k = 1, \dots, K$. Совокупность элементов $\{z^{ij}\}$ какого-либо из рассматриваемых изображений набора $1, \dots, K$ с координатами (i, j) , принадлежащими квадрату $(2l+1) \times (2l+1)$ пикселей, будем называть фрагментом с центральным элементом $(i=0, j=0)$ и введенной на нем локальной системой координат $i = -l, \dots, -1, 0, +1, \dots, +l$; $j = -l, \dots, -1, 0, +1, \dots, +l$, где l – параметр размера окна. Соответственно фрагмент на квадрате окна $(2l+1) \times (2l+1)$ многокомпонентного изображения представляет собой совокупность элементов $\{z^{ijk}\}$, $(i, j, k) \in (2l+1) \times (2l+1) \times K$.

Для описания поведения значений яркостей $\{z^{ij}\}$ в пределах фрагмента $(2l+1) \times (2l+1)$ воспользуемся простой моделью, введенной в [5], при этом локальные характеристики изображения описываются отрезками плоскости – фасетами.

Формализованное описание окрестности точки (x, y) определяется следующим уравнением плоскости, записанным для непрерывных координат x, y, z трехмерного пространства

$$z = \alpha x + \beta y + \mu, \quad (1)$$

где z – результирующее значение яркости фасеточной модели; α, β, μ – параметры плоскости. Обозначив через i и j дискретные значения координат x, y в (1) соответственно, согласуем модель (1) с фактическими наблюдениями z^{ij} фрагмента изображения $(2l+1) \times (2l+1)$ наилучшим образом в смысле метода наименьших квадратов, а именно

$$I(\alpha, \beta, \gamma) = \sum_{j=-l}^{+l} \sum_{i=-l}^{+l} [\alpha i + \beta j + \mu - z^{ij}]^2 = \min_{\{\alpha, \beta, \mu\}}.$$

Оценки неизвестных параметров, обеспечивающие минимум $I(\alpha, \beta, \gamma)$, будут иметь следующий вид:

$$\begin{aligned} \hat{\alpha} &= 3 \sum_{i=-l}^{+l} i \sum_{j=-l}^{+l} z^{ij} / l(l+1)(2l+1)^2; \\ \hat{\beta} &= 3 \sum_{j=-l}^{+l} j \sum_{i=-l}^{+l} z^{ij} / l(l+1)(2l+1)^2; \\ \hat{\mu} &= \sum_{j=-l}^{+l} \sum_{i=-l}^{+l} z^{ij} / (2l+1)^2. \end{aligned} \quad (2)$$

Описание фрагмента с помощью фасеточной модели (1) будем использовать для решения задачи выделения градиентных участков изображения. Величину градиента изображения в пределах фрагмента будем оценивать пространственной производной, определяемой как отношение площади dS наклонной плоскости (1), ограниченной квадратом фрагмента $(2l+1) \times (2l+1)$, к площади основания $d\Delta$ этого фрагмента. Если уравнение (1) представить в виде уравнения плоскости в направляющих косинусах, то

$$\frac{dS}{d\Delta} \cong \left| \left(\frac{3 \sum_{i=-l}^{+l} i \sum_{j=-l}^{+l} z^{ij}}{l(l+1)(2l+1)^2} \right)^2 + \left(\frac{3 \sum_{j=-l}^{+l} j \sum_{i=-l}^{+l} z^{ij}}{l(l+1)(2l+1)^2} \right)^2 + 1 \right|^{1/2}. \quad (3)$$

Полученное оценочное значение градиента отнесем к центральной точке фрагмента с локальными координатами пиксела $i=0, j=0$. Если теперь аналогичным образом <продифференцировать> каждое из совокупности анализируемых изображений, относя к центральным элементам <скользящего> окна с параметром l значения соответствующих градиентов, то от исходных изображений рельефа радиояркостей $\{z^{ijk}\}$ аэрокосмических снимков можно перейти к изображениям с <подчеркнутым> градиентом со значениями $\{w^{ijk}\}$, определяемыми по формуле (3).

Построим гистограмму распределения значений градиентов одного из градиентных изображений и попытаемся выделить на этой гистограмме две области: левую, начинающуюся от нуля, – область, <не градиентов>, и правую, – область <градиентов>.

Формализованная постановка этой задачи приводит к задаче декомпозиции совместного распределения градиентов $p(w)$, которым является упомянутая гистограмма, на два составляющих распределения [1, 4]:

$$p(w) = P f(w) + Q g(w), \quad (4)$$

где $g(w)$ – функция плотности распределения вероятностей при условии гипотезы H_1 , заключающейся в том, что наблюдение w относится к классу градиентов; $f(w)$ – функция плотности при условии альтернативы H_0 , заключающейся в том, что наблюдение w относится к классу не градиентов; P, Q – параметры смеси; $P + Q = 1$.

В качестве параметрических моделей распределений, составляющих смесь (4), возьмем S_B – распределения Джонсона [6], обладающие высокими аппроксимирующими способностями в сочетании с небольшим числом параметров, среди которых есть параметры формы

$$g(w), f(w) = \frac{1}{\sqrt{2\pi} (w - \varepsilon) (\lambda - w + \varepsilon)} \exp \left\{ -\frac{1}{2} \left[\gamma + \eta \ln \left(\frac{w - \varepsilon}{\lambda - w + \varepsilon} \right) \right]^2 \right\}, \quad (5)$$

где $\varepsilon \leq w \leq \varepsilon + \lambda$, $\eta > 0$, $\lambda > 0$, $-\infty < \gamma < \infty$.

Задачу декомпозиции распределений (4) с учетом моделей (5) будем решать численными методами, выбрав в качестве критерия оптимальности минимум следующего квадратичного критерия:

$$[\hat{p}(w) - P\hat{f}(w) - Q\hat{g}(w)]^2 = \min_{\{\Theta\}}, \quad (6)$$

где $\hat{p}(w)$ – гистограмма градиентов изображения; $\Theta = \{\varepsilon_0, \lambda_0, \gamma_0, \eta_0; \varepsilon_1, \lambda_1, \gamma_1, \eta_1; P\}^T$ – набор неизвестных параметров; t – знак транспонирования, причем $\varepsilon_0, \lambda_0, \gamma_0, \eta_0$ – параметры распределения $f(w)$ в формуле (5) при условии альтернативы H_0 ; $\varepsilon_1, \lambda_1, \gamma_1, \eta_1$ – параметры распределения $g(w)$ в форме (5) при условии гипотезы H_1 ; P – вес распределения $f(w)$ в (4), а $Q = 1 - P$. Задачу минимизации (6) многоэкстремального функционала по набору из 9 параметров можно решить численно одним из стандартных методов, использующих случайный поиск исходных точек с последующим градиентным спуском.

После того, как оценки параметров распределений $f()$, $g()$ будут получены, решающее правило проверки гипотез о наличии контура (градиента) или отсутствии градиента запишется в виде байесова решающего правила

$$u(w^{ij}) = \arg \max \{ \hat{P}\hat{f}(w^{ij}), \hat{Q}\hat{g}(w^{ij}) \}, \quad (7)$$

где $u()$ – решение; w^{ij} – значение анализируемого градиентного изображения в точке (i, j) . В том случае, когда мы имеем дело с совокупностью градиентных изображений, обобщенный градиент определяется вектором $\frac{d\mathbf{S}}{d\Delta} = \left(\frac{dS_1}{d\Delta_1}, \dots, \frac{dS_K}{d\Delta_K} \right)^T$.

Компоненты этого вектора есть значения отдельных градиентных изображений.

Для решения задачи восстановления многомерных вероятностных распределений компонент вектора $\frac{d\mathbf{S}}{d\Delta}$ в модели (4) естественно воспользоваться многомерным аналогом S_B -распределения Джонсона [6], что существенно усложнит задачу оценивая неизвестных параметров. Упрощения можно достигнуть ценой некоторой потери качества решения, вводя понятие нормы вектора градиентов

$$\left\| \frac{d\mathbf{S}}{d\Delta} \right\| = \sqrt{\left(\frac{dS_1}{d\Delta_1} \right)^2 + \dots + \left(\frac{dS_K}{d\Delta_K} \right)^2},$$

что позволяет полностью использовать описанный выше вариант с отдельным градиентным изображением.

Наконец, перейдем к рассмотрению основной задачи текстурного анализа многомерных видеоданных, в которых исключены контурные линии высокого градиента.

Так как текстура проявляется как пространственная характеристика изображения, необходимо исходные изображения представить в виде фрагментов и исследовать статистические характеристики каждого из фрагментов.

Воспользуемся одним из критериев проверки гипотез на однородность двух выборок [7], использующим в качестве теста <расстояние> Колмогорова–Смирнова между многомерными эмпирическими распределениями, когда последние восстанавливаются на парах сравниваемых фрагментов.

Пусть $\{z^{ijk}\}$ и $\{v^{ijk}\}$ – наборы радиояркостей, зафиксированные на двух, анализируемых на однородность, фрагментах изображения с непрерывными интегральными функциями распределения $F(z)$ и $G(z)$ соответственно. Мерой идентичности текстур фрагментов изображения в этом случае будет значение расстояния Колмогорова–Смирнова между многомерными распределениями, восстанавливаемыми на соответствующих фрагментах [7]:

$$D_{mn} = \sup_{\{z\}} |F_m(z) - G_n(z)|, \quad (8)$$

где $F_m(\cdot)$ и $G_n(\cdot)$ – многомерные эмпирические функции распределений следующего вида:

$$F_m(z) = \frac{1}{m} \sum_{\{i,j\}} \prod_{k=1}^K C(z^k - z^{ijk}); \quad (9)$$

$$G_n(z) = \frac{1}{n} \sum_{\{i,j\}} \prod_{k=1}^K C(z^k - v^{ijk});$$

m – число элементов фрагмента $\{z^{ijk}\}$; n – число элементов фрагмента $\{v^{ijk}\}$, причем в этих фрагментах отсутствуют элементы, принадлежащие контуру, если таковой был; $C(\cdot)$ – функция сравнения, $C(t) = \{1, t \geq 0; 0, t < 0\}$.

Необходимо проверить гипотезу H_0 , заключающуюся в том, что интегральные функции распределения яркостей, восстанавливаемые на упомянутых фрагментах, равны между собой в статистическом смысле и две выборки, по которым они восстановлены, принадлежат одному статистическому ансамблю, что свидетельствует о статистической однородности сравниваемой пары фрагментов. Альтернатива H_1 утверждает неравенство интегральных функций плотности, а значит, и отсутствие статистической эквивалентности текстур этих фрагментов.

Другими словами, надо проверить гипотезу $H_0: F(z) = G(z)$, по отношению к альтернативе $H_1: F(z) \neq G(z)$. Однако следует иметь в виду, что для статистики (8) двустороннего теста однородности получены асимптотические распределения D_{mn} лишь для одномерных функций $F(\cdot)$, $G(\cdot)$.

Выход из этой ситуации можно найти, воспользовавшись идеей декомпозиции смешанной статистики D_{mn} на составляющие распределения, которые затем используются для построения байесова решающего правила проверки гипотез H_0 и H_1 , как и в случае модели (4).

Процесс выделения кластеров текстурно-однородных фрагментов в этом случае строится следующим образом. Исходное многомерное изображение декомпозируется координатной сетью с некоторым шагом на отдельные многомерные фрагменты со значениями радиояркостей $\{z^{ijk}\}$, $i = 1, \dots, l_x$; $j = 1, \dots, l_y$; $k = 1, \dots, K$. Размер фрагментов $l_x \times l_y$ выбирается из тех соотношений, чтобы текстурные свойства могли в полной мере проявиться в области фрагмента. При этом следует учитывать противоречивые требования, а именно: выбирая фрагмент больших размеров, можно полнее описать текстурные характеристики фрагмента, повышая устойчивость оценок распределений, однако при этом теряется точность детализации границ областей однородности, кроме того, в рамках большего фрагмента могут вероятнее всего нарушиться свойства внутренней однородности текстуры.

После того как набор фрагментов на изображениях зафиксирован, выберем один из них в качестве опорного, а вторым – последовательно каждый из оставшихся, затем оценим степень однородности пар фрагментов в смысле расстояния (8). Этот процесс порождает статистику значений $\{D_{mnj}\}$, $j = 1, \dots, M$, где M – количество фрагментов, сравниваемое с опорным. По этой статистике строим гистограмму смешанного распределения, описывающую поведение D_{mn} при условии гипотезы H_0 и альтернативы H_1 .

Вновь воспользовавшись моделью (4) и задав параметрические аппроксимации неизвестных условных функций плотности при условии гипотезы H_0 и альтернативы H_1 , эти функции

можно восстановить вместе с априорными вероятностями кластеров, решая задачу идентификации смеси распределений (6). После чего решающее правило проверки гипотез H_0 и H_1 имеет стандартный байесов вид (7). Среди возможных вариантов выбора опорного фрагмента берется тот вариант, устанавливаемый перебором всех возможных вариантов, при котором гистограмма распределения статистики D_{mn} в большей степени <сосредоточена> у нуля, что свидетельствует о компактности кластера.

Этот вариант со своим опорным фрагментом и выбирается в качестве рабочего для агрегирования фрагментов в кластер. Решающее правило разделяет все множество фрагментов на множество, образующее класс текстурно-однородных фрагментов, и множество фрагментов, не вошедших в этот класс; последние вновь подвергаются описанной процедуре выбора опорного фрагмента с последующей кластеризацией.

После некоторого числа итераций остается некоторое количество фрагментов, которые можно охарактеризовать как разнородные или выбросы наблюдений.

Рассмотрим иллюстративный пример сегментации конкретного космического снимка, любезно предоставленного В.И. Хамариным (Институт экологии природных комплексов СО РАН). Снимок представляет собой совокупность трех оцифрованных значений радиояркостей в диапазонах 0,5–0,6 мкм (первый канал), 0,6–0,7 мкм (второй канал), 0,8–0,9 мкм (третий канал).

Анализируемый фрагмент зафиксировал радиояркости участка лесной зоны сибирской тайги с просеками, вырубками и имеет размер 640×448 точек. На рис. 1 представлен вид исходного изображения третьего канала. Для выделения контурных участков видеоданных было задано окно фасеточной модели (1) с параметром $l = 1$ и получено изображение градиентов, определяемых нормой пространственного градиента. Гистограмма распределения градиентов, представленная на рис. 2, а, явилась входной информацией для решения задачи декомпозиции смешанного распределения на две составляющие функции плотности (4), заданные в виде параметрических распределений S_B Джонсона (5). Восстановленные функции плотности с весовыми коэффициентами изображены на рис. 2, а. Они определяют байесово решающее правило (7) для выявления контурных линий на градиентном изображении. Выделенные контуры показаны светлым тоном на рис. 3. Затем область определения видеоданных – прямоугольник 640×448 отсчетов – была декомпозирована координатной сетью на отдельные фрагменты размером 16×16 отсчетов и проведена автоматическая классификация полученного набора фрагментов, исключая те из них, которые содержат элементы контурных линий.

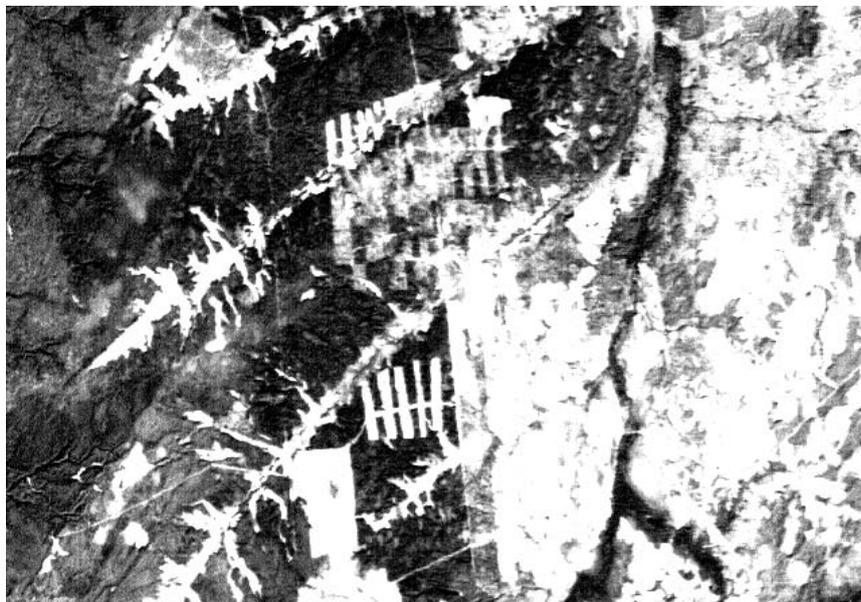


Рис. 1. Исходное изображение третьего канала

С использованием трехмерного варианта расстояния Колмогорова–Смирнова (8) на парах анализируемых фрагментов, один из которых выбирался опорным, а вторым являлся последо-

вательно каждый из оставшихся, были получены гистограммы распределения значений статистик $\{D_{mn}\}$. Одно из получающихся смешанных распределений статистики D_{mn} , восстановленное с помощью непараметрической оценки неизвестной функции плотности с ядром Епанечникова, приведено на рис. 2, б. Решающее правило формирует очередной класс статистически однородных фрагментов из той части общего набора фрагментов, которая связана с левой модой восстановленной функции плотности. На этапе кластеризации было выделено 5 классов, которые показаны на рис. 3 вариацией серого тона.

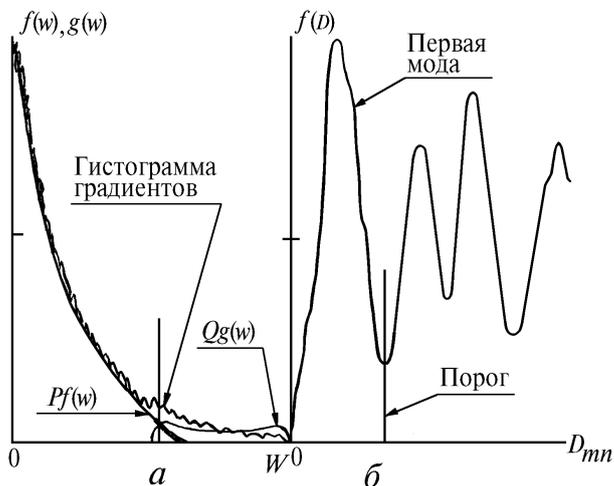


Рис. 2. Гистограмма распределения градиентов изображения и ее аппроксимация (а) и непараметрическое восстановление распределения расстояний Колмогорова–Смирнова (б)



Рис. 3. Изображение с выделенными контуром и классами однородных фрагментов

В заключение отметим, что в описанной схеме построения процедуры автоматической классификации можно использовать и другие двусторонние критерии проверки двух выборок на однородность, например критерий ω^2 [7].

1. Дуда Р., Харт П. Распознавание образов и анализ сцен. М.: Мир, 1976. 512 с.
2. Хомяков Ю.Н., Сушкин В.А. //Зарубежная радиоэлектроника. 1986. N 2. С. 33–46.
3. Денисов Д.А., Низовкин В.А. // Зарубежная радиоэлектроника. 1985. N 10. С. 5–30.
4. Розенфельд А. Распознавание и обработка изображений с помощью вычислительных машин. М.: Мир, 1972. 230 с.

5. Чочиа П. А. // Кодирование и обработка изображений. М.: Наука, 1988. С. 69–87.
6. Протасов К. Т. // Оптика атмосферы и океана. 1994. Т. 7. N 6. С. 840–847.
7. Тарасенко Ф. П. Непараметрическая статистика. Томск: Изд-во ТГУ, 1976. 292 с.

Институт оптики атмосферы СО РАН,
Томск

Поступила в редакцию
29 декабря 1994 г.

Ju. V. Gridnev, K.T. Protasov. Segmentation of Many-dimensional Images by the Clustering Algorithm Based on Bilateral Test of Homogeneity.

An algorithm of automatic classification is proposed for thematic decoding the aerospace videodata obtained from observations of the Earth underlying surface and clouds. The algorithm is based on bilateral test of homogeneity of many-dimensional samples with use of Kolmogorov–Smirnov <space>. The feature of the constructed algorithm is its applicability to a set of <coordinated> images from multispectral channels of visible, IR, and SHF ranges of the electromagnetic spectrum.