

АЛГОРИТМЫ ИДЕНТИФИКАЦИИ ГИПЕРСПЕКТРАЛЬНЫХ ХАРАКТЕРИСТИК ОБЪЕКТОВ В ЗАДАЧАХ ДИСТАНЦИОННОГО ЗОНДИРОВАНИЯ ЗЕМЛИ

Демидова Л.А., профессор РГРТУ, д.т.н., e-mail: Liliya.demidova@rambler.ru;

Тишкин Р.В., начальник отдела филиала ФГУП «ГНПРКЦ «ЦСКБ-Прогресс» – ОКБ «Спектр», к.т.н., e-mail: roman.tishkin@mail.ru;

Труханов С.В., главный специалист отдела филиала ФГУП «ГНПРКЦ «ЦСКБ-Прогресс» – ОКБ «Спектр», e-mail: serge_tsv@mail.ru

Ключевые слова: алгоритм идентификации, гиперспектральная характеристика объекта, мера подобия евклидова расстояния, нечеткая мера подобия, угловая мера подобия, нечеткая линейная регрессия, консолидация.

Введение

В последние годы одной из важных задач, решаемых системами дистанционного зондирования Земли (ДЗЗ), является задача обработки и анализа гиперспектральной информации, которая формируется на основе полученных с борта космического аппарата гиперспектральных изображений (ГСИ) земной поверхности, представленных большим набором снимков одной и той же сцены в узких соприкасающихся спектральных диапазонах. Так, например, отечественный космический аппарат «Ресурс-П» с гиперспектральной аппаратурой (ГСА) на борту, запущенный 25 июня 2013 года, выдает наборы снимков в 130 узких соприкасающихся диапазонах видимой области спектра, которые и формируют гиперспектральное изображение (ГСИ), называемое также «гиперкубом».

Большое внимание при обработке ГСИ уделяется проблеме идентификации объектов земной поверхности по их гиперспектральным характеристикам (ГСИ). ГСХ объекта в графическом виде представляет собой отображение связи между длиной волны и значениями коэффициента спектральной яркости (КСЯ) анализируемого объекта. Идентификация объекта по его ГСХ может быть осуществлена посредством сравнения с применением тех или иных мер подобия (сходства) анализируемой ГСХ объекта с некоторыми эталонными ГСХ (эталонами), объединенными в специальные спектральные библиотеки [1–3]. При этом необходимо учитывать, что у анализируемой и эталонных ГСХ должны быть одинаковыми единицы измерения, спектральный диапазон и разрешение данных [3].

Цель работы заключается в создании и исследовании алгоритмов идентификации ГСХ объектов земной поверхности с применением различных обоснованно выбранных мер подобия, а также – в последующей консолидации частных результатов идентификации ГСХ объектов, полученных при использовании предлагаемых алгоритмов.

Рассматривается подход к решению задачи идентификации объектов земной поверхности, основанный на анализе гиперспектральных характеристик объектов, получаемых с обработанных космических изображений, посредством применения алгоритмов идентификации на основе различных мер подобия. Представлены результаты обработки гиперспектральной информации с использованием предлагаемых алгоритмов идентификации.

Алгоритмы идентификации: теоретические исследования

Анализ работ отечественных и зарубежных авторов [4–8] свидетельствует о том, что в настоящее время для решения задач идентификации объектов земной поверхности по их ГСХ наиболее часто используются подходы, основанные на применении: мер подобия (сходства) [5, 9]; метода спектрального угла [4, 5, 10]; искусственных нейронных сетей [5, 12]; алгоритма k -средних [4–7, 10]; метода максимального правдоподобия (к сожалению, оказавшегося неэффективным при работе с 200-полосными ГСИ) [4, 5].

Результаты экспериментальных исследований по оценке качества идентификации объектов земной поверхности на космических снимках с применением перечисленных выше подходов, свидетельствуют о том, что ни один из них не обладает явным преимуществом перед другими, так как не обеспечивает требуемое качество идентификации объектов в преобладающем большинстве случаев.

Для решения задачи идентификации объектов земной поверхности по их ГСХ предлагается одновременно использовать четыре алгоритма идентификации ГСХ, основанные на мере подобия (сходства) евклидова расстояния, угловой мере подобия и двух нечетких мерах подобия, с целью последующей консолидации частных результатов идентификации.

Выбор меры подобия евклидова расстояния обусловлен высокой эффективностью применения данной меры при решении широкого спектра задач идентификации (классификации), в том числе и задач идентификации при обработке космических изображений.

Использование угловой меры подобия, предполагающей реализацию метода спектрального угла (Spectral Angle Mapper – SAM), может быть обосновано тем, что

данная мера обеспечивает получение хороших результатов идентификации ГСХ объектов, имеющих схожие значения яркости точек изображения во всех спектральных диапазонах даже в том случае, когда на исходных изображениях наблюдаются эффекты засветки, поскольку метод спектрального угла не учитывает значения яркости точек изображения.

Целесообразность использования нечетких мер подобия [6, 11, 13] может быть обоснована тем, что некоторая анализируемая ГСХ объекта может одновременно принадлежать к разным классам (категориям) объектов, в то время как алгоритмы «жесткой» идентификации [5], которыми являются алгоритм на основе меры подобия евклидова расстояния и алгоритм на основе угловой меры подобия, однозначно отнесут анализируемую ГСХ объекта к тем или иным однозначно определенным классам, что может быть не всегда верно. Применение алгоритмов «мягкой» идентификации, основанных на применении нечетких мер подобия, позволит более гибко решать задачу идентификации объекта по его ГСХ.

Программная реализация предлагаемых алгоритмов идентификации объекта по его ГСХ (алгоритмов сравнения ГСХ объекта с эталонами на основе меры подобия евклидова расстояния, угловой меры подобия и двух нечетких мер подобия) была выполнена при разработке системы интеллектуальной обработки данных гиперспектральной съемки (СИОДГС), структурно состоящей из базы данных хранения и обработки данных (БД) и программы интеллектуальной обработки данных гиперспектральной съемки [14, 15].

В БД СИОДГС осуществляется хранение и обработка анализируемых ГСХ; в частности, с помощью хранимых процедур проводятся операции нормирования анализируемой ГСХ к эталонным значениям, интерполяция входных значений, расчет меры подобия евклидова расстояния, угловой меры подобия, двух нечетких мер подобия, а также – обработка и хранение результатов. В программе интеллектуальной обработки данных гиперспектральной съемки СИОДГС реализованы четыре алгоритма идентификации ГСХ объекта: алгоритм идентификации ГСХ на основе меры подобия евклидова расстояния, алгоритм идентификации ГСХ на основе угловой меры подобия, два алгоритма идентификации ГСХ на основе нечетких мер подобия, а также алгоритм консолидации частных результатов идентификации ГСХ объекта.

Алгоритм идентификации ГСХ на основе меры подобия евклидова расстояния

Алгоритм идентификации ГСХ на основе меры подобия евклидова расстояния реализует идентификацию ГСХ с использованием известной евклидовой метрики для вычисления расстояния между двумя точками в J -мерном пространстве [9]:

$$E = \sqrt{\sum_{j=1}^J (y_j^A - y_j^S)^2}, \quad (1)$$

где y_j^A – значение коэффициента спектральной яркости (КСЯ) анализируемой (Analyzed) ГСХ для j -го канала гиперспектрометра; y_j^S – значение КСЯ эталонной

(Standard) ГСХ для j -го канала гиперспектрометра; $j = \overline{1, J}$; J – количество каналов гиперспектрометра, равное количеству точек в анализируемой (эталонной) ГСХ (например, $J = 96$).

Следует отметить, что длина волны λ_j для j -го ($j = \overline{1, J}$) канала гиперспектрометра считается известной величиной, которой в соответствие ставятся значения КСЯ анализируемой и эталонной ГСХ.

Алгоритм идентификация ГСХ на основе меры подобия евклидова расстояния предполагает:

- расчет мер подобия евклидова расстояния (1) для анализируемой ГСХ и эталонных ГСХ, хранимых в БД;
- упорядочение по возрастанию вычисленных значений мер подобия евклидова расстояния;
- выбор в качестве искомой той эталонной ГСХ, для которой значение меры подобия евклидова расстояния (1) является минимальным.

Алгоритм идентификации ГСХ на основе угловой меры подобия

Алгоритм идентификации ГСХ на основе угловой меры подобия предполагает использование метода спектрального угла (Spectral Angle Mapper – SAM), реализующего оценку подобия исследуемой и эталонной ГСХ, рассматриваемых как векторы, размерность которых равна количеству каналов гиперспектрометра, посредством вычисления угла между ними [10]:

$$\alpha = \arccos \frac{G \cdot G'}{|G| \cdot |G'|} = \arccos \frac{\sum_{j=1}^J g_j \cdot g'_j}{\sqrt{\sum_{j=1}^J g_j^2} \cdot \sqrt{\sum_{j=1}^J (g'_j)^2}}, \quad (2)$$

где G и G' – спектры анализируемой и эталонной ГСХ соответственно; g_j и g'_j – значения КСЯ анализируемой и эталонной ГСХ соответственно для значения длины волны λ_j ($j = \overline{1, J}$); J – количество каналов гиперспектрометра.

Алгоритм идентификация ГСХ на основе угловой меры подобия предполагает:

- расчет угловых мер подобия (2) для анализируемой ГСХ и эталонных ГСХ, хранимых в БД;
- упорядочение по возрастанию вычисленных значений угловых мер подобия;
- выбор в качестве искомой той эталонной ГСХ, для которой значение угловой меры подобия (2) является минимальным.

Довольно часто из-за погрешностей ГСА, приводящих к неточности информации об анализируемой ГСХ, а также из-за почти полного отсутствия «чистых» анализируемых ГСХ и наличия в большинстве случаев анализируемых ГСХ, представляющих собой смесь из нескольких базовых классов [5], целесообразно рассматривать несколько эталонных ГСХ, имеющих наименьшие значения меры подобия евклидова расстояния (1) и угловой меры подобия (2), в качестве потенциально искомых.

Несмотря на использование при идентификации большого количества эталонных ГСХ, результаты идентификации с применением меры подобия евклидова

расстояния или угловой меры подобия может быть неудовлетворительным. В связи с этим целесообразно осуществлять подтверждение результата идентификации, которое, в частности, может быть получено посредством применения других алгоритмов идентификации ГСХ с последующей консолидацией частных результатов идентификации. Как показывают экспериментальные исследования, хороший консолидирующий результат идентификации обеспечивает совместное использование алгоритма идентификации ГСХ на основе меры подобия евклидова расстояния, алгоритма идентификации ГСХ на основе угловой меры подобия и двух алгоритмов идентификации на основе нечетких мер подобия.

Алгоритмы идентификации ГСХ на основе нечетких мер подобия

Идентификация ГСХ объекта может быть выполнена на основе уравнения классической линейной регрессии (КЛР) посредством решения задач подбора параметров k и b уравнения (например, в соответствии с методом наименьших квадратов [16]) для анализируемой и эталонных ГСХ:

$$y = k \cdot x + b \quad (3)$$

с последующим вычислением среднеквадратичных отклонений (СКО) невязок анализируемой и эталонных ГСХ, которые могут использоваться как характеристики уникальности при идентификации ГСХ.

Если для идентификации ГСХ объекта использовать только одну характеристику уникальности – такую, как СКО, то в ряде случаев ГСХ объектов, имеющие приблизительно равные значения СКО, могут быть отнесены к одному классу, несмотря на то, что форма кривых ГСХ объектов будет различна [17]. В связи с этим для идентификации ГСХ объекта предлагается использовать дополнительные характеристики уникальности, которые могут быть получены на основе уравнения нечеткой линейной регрессии (НЛР) и использованы для расчета нечетких мер подобия.

В [18] показано, что наиболее целесообразно использовать уравнение НЛР с асимметричными нечеткими параметрами, так как оно (в отличие от уравнения нечеткой линейной регрессии с симметричными нечеткими параметрами) обеспечивает вычисление значения характеристики уникальности – СКО, которое равно или близко значению, получаемому с помощью уравнения КЛР:

$$Y(x) = A_1 \cdot x + A_0, \quad (4)$$

где $A_1 = (a_1, c_1, d_1)$ и $A_0 = (a_0, c_0, d_0)$ – ТНЧ, соответствующие параметрам k и b уравнения КЛР (3), представляемые с помощью треугольных функций принадлежности (подробно рассмотренных в [17]), и являющиеся асимметричными нечеткими параметрами уравнения НЛР (4).

При разработке алгоритмов идентификации ГСХ объекта на основе нечетких мер могут использоваться различные нечеткие меры подобия [11], при этом сами алгоритмы идентификации будут иметь одинаковые этапы реализации. Укрупненная схема алгоритма идентификации ГСХ объекта на основе нечеткой меры подобия приведена на рис. 1.

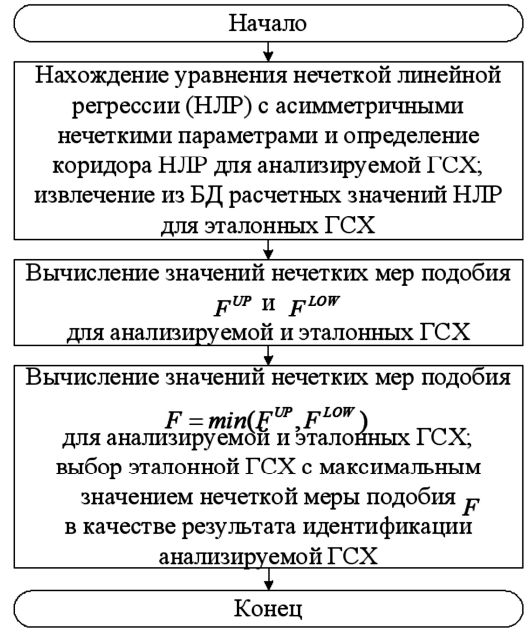


Рис. 1. Укрупненная схема алгоритма идентификации ГСХ объекта на основе нечеткой меры подобия

На первом этапе алгоритма определяются параметры (а именно ТНЧ) уравнения НЛР анализируемой ГСХ, для чего решается задача квадратичного программирования (ЗКП), формулируемая следующим образом [17, 18]:

$$F_{\text{ЗКП}} = k_1 \cdot \sum_{j=1}^J (y_j - \sum_{i=0}^n a_i \cdot x_{ji})^2 + k_2 \cdot (1 - \alpha) \times \quad (5)$$

$$\times \sum_{j=1}^J \sum_{i=0}^n (c_i + d_i) \cdot x_{ji} + \xi \cdot \sum_{i=0}^n (c_i^2 + d_i^2) \rightarrow \min_{a, c, d}$$

при ограничениях:

$$\sum_{i=0}^n a_i \cdot x_{ji} + (1 - \alpha) \cdot \sum_{i=0}^n d_i \cdot x_{ji} \geq y_j; \quad (6)$$

$$\sum_{i=0}^n a_i \cdot x_{ji} - (1 - \alpha) \cdot \sum_{i=0}^n c_i \cdot x_{ji} \leq y_j; \quad (7)$$

$$c_i \geq 0; d_i \geq 0 \quad (j = \overline{1, J}; i = \overline{0, n}; n = 1), \quad (8)$$

где j – порядковый номер канала гиперспектрометра ($j = \overline{1, J}$); J – количество каналов гиперспектрометра;

i – порядковый номер ТНЧ ($i = \overline{0, 1}$ при $n = 1$); $x_{j0} = 1$

($j = \overline{1, J}$), так как соответствуют свободному члену – параметру b – уравнения КЛР (2); x_{j1} – значение длины волны для j -го канала гиперспектрометра; y_j – значение КСЯ ГСХ для длины волны x_{j1} ($x_{j1} = \lambda_j$); α – значение уровня ТНЧ, характеризующее ширину коридора НЛР ($\alpha \in [0, 1]$); k_1, k_2 – весовые коэффициенты, характеризующие вклад первого и второго слагаемых в целевую функцию (4); ξ – малое положительное число такое, что $k_1, k_2 \gg \xi$ (третье слагаемое введено в целевую функцию для того, чтобы она имела квадратичный вид и чтобы при поиске значений ТНЧ можно было сформулировать ЗКП [18]).

При решении ЗКП (5)–(8) предполагается, что $k_1 = k_2 = 1$ (в [17] показано, что выбор значений пара-

метров k_1, k_2 не оказывает сильного влияния на решение задачи); $\xi = 0,001$.

Пусть $a_0 = z_1; c_0 = z_2; d_0 = z_3; a_1 = z_4; c_1 = z_5; d_1 = z_6; x_{j1} = \lambda_j; y_j = g_j$. Тогда ЗКП (5)–(8) может быть записана как:

$$F_{\text{ЗКП}} = \sum_{j=1}^J (g_j - z_1 - z_4 \cdot \lambda_j)^2 + \sum_{j=1}^J (z_2 + z_5 \cdot \lambda_j + z_6 \cdot \lambda_j) + 0,001 \cdot (z_2^2 + z_3^2 + z_5^2 + z_6^2) \rightarrow \min_{z_1, \dots, z_6} \quad (9)$$

при ограничениях:

$$\begin{cases} z_1 + z_3 + z_4 \cdot \lambda_j + z_6 \cdot \lambda_j \geq g_j; \\ z_1 - z_2 + z_4 \cdot \lambda_j - z_5 \cdot \lambda_j \leq g_j; \\ z_2 \geq 0; \\ z_3 \geq 0; \\ z_5 \geq 0; \\ z_6 \geq 0. \end{cases} \quad (10)$$

Для решения ЗКП в виде (8) при ограничениях (9) может быть использован метод неопределенных множителей Лагранжа [16]. Ввиду того, что решение ЗКП методом множителей Лагранжа является весьма трудоемкой задачей, целесообразно использование существующих математических пакетов, позволяющих решать задачи минимизации с ограничениями. В частности, ЗКП может быть решена в системе инженерных и научных расчетов MATLAB с применением встроенной функции «quadprog» [19].

Уравнение НЛР (3) для анализируемой ГСХ объекта строится на основе решения ЗКП (9)–(10). При этом на основе вычисленных ТНЧ $A_1 = (a_1, c_1, d_1)$ и $A_0 = (a_0, c_0, d_0)$ уравнения НЛР (4) для значений длин волн $x = \lambda_j (j = \overline{1, J})$ определяются характерные точки уравнения КЛР $Y_{\text{НЛР}}^{\text{КЛР}}(\lambda_j)$:

$$Y_{\text{НЛР}}^{\text{КЛР}}(\lambda_j) = a_0 + a_1 \cdot \lambda_j, \quad (11)$$

а также уравнений верхней $Y_{\text{НЛР}}^{\text{UP}}(\lambda_j)$ и нижней $Y_{\text{НЛР}}^{\text{LOW}}(\lambda_j)$ границ коридора НЛР анализируемой характеристики:

$$Y_{\text{НЛР}}^{\text{UP}}(\lambda_j) = a_0 + d_0 + (a_1 + d_1) \cdot \lambda_j, \quad (12)$$

$$Y_{\text{НЛР}}^{\text{LOW}}(\lambda_j) = a_0 - c_0 + (a_1 - c_1) \cdot \lambda_j. \quad (13)$$

Для эталонных ГСХ уравнения НЛР (3), а также уравнения КЛР $Y_{\text{НЛР}}^{\text{КЛР}}(\lambda_j)$, верхней $Y_{\text{НЛР}}^{\text{UP}}(\lambda_j)$ и нижней $Y_{\text{НЛР}}^{\text{LOW}}(\lambda_j)$ границ коридора НЛР, определяемые в соответствии с (11), (12) и (13), строятся аналогичным образом, при этом все расчетные значения НЛР для эталонных ГСХ хранятся в БД СИОДГС.

Из-за асимметричности ТНЧ $A_1 = (a_1, c_1, d_1)$ и $A_0 = (a_0, c_0, d_0)$ коридоры НЛР анализируемой и эталонных ГСХ также асимметричны (рис. 2). В связи с этим точки ГСХ (как анализируемой, так и каждой эталонной) могут быть разбиты на 2 подмножества: подмножество точек, лежащих в верхней части *UP* коридора НЛР (между линией верхней границы коридора НЛР и линией КЛР, определяемыми соответственно уравнениями (12) и (11)), и подмножество точек, лежащих в нижней части *LOW* коридора НЛР (между линией нижней границы

коридора НЛР и линией КЛР, определяемыми соответственно уравнениями (13) и (11)).

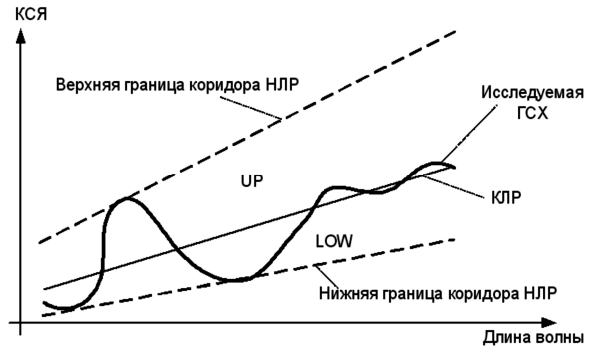


Рис. 2. Представление ГСХ в коридоре НЛР

На втором этапе алгоритма для точек анализируемой и каждой эталонной ГСХ, лежащих в верхней *UP* и нижней *LOW* частях коридоров своих НЛР, находятся значения нечетких мер подобия F^{UP} и F^{LOW} , для вычисления которых используется одна из двух нечетких мер подобия вида [11, 17]:

$$f_1 = 1 - \frac{\sum_{j=1}^J |u_A(\lambda_j, g'_j) - u_S(\lambda_j, g''_j)|}{\sum_{j=1}^J (u_A(\lambda_j, g'_j) + u_S(\lambda_j, g''_j))}, \quad (14)$$

$$f_2 = \frac{1}{J} \cdot \sum_{j=1}^J \frac{\min(u_A(\lambda_j, g'_j), u_S(\lambda_j, g''_j))}{\max(u_A(\lambda_j, g'_j), u_S(\lambda_j, g''_j))}, \quad (15)$$

где $u_A(\lambda_j, g'_j)$ – значение функции принадлежности нечеткого множества *A* (Analyzed) анализируемой ГСХ уравнению НЛР этой ГСХ для значения КСЯ g'_j , соответствующего длине волны $\lambda_j (j = \overline{1, J})$; $u_S(\lambda_j, g''_j)$ – значение функции принадлежности нечеткого множества *S* (Standard) эталонной ГСХ уравнению НЛР этой ГСХ для значения КСЯ g''_j , соответствующего длине волны $\lambda_j (j = \overline{1, J})$; *J* – количество каналов гиперспектрометра, равное количеству точек в анализируемой (эталонной) ГСХ.

Данные меры подобия были выбраны из множества известных нечетких мер подобия как показавшие наивысшее качество результатов идентификации при выполнении оценки подобия тестовых наборов данных, описываемых с помощью уравнений НЛР [17].

Значение функции принадлежности некоторой точки (λ_j, g_j) ГСХ к уравнению НЛР этой ГСХ может быть определено как [17, 18]:

$$u(\lambda_j, g_j) = \begin{cases} 1 - \frac{a_0 + a_1 \cdot \lambda_j - g_j}{c_0 + c_1 \cdot \lambda_j}, & \text{если} \\ a_0 + a_1 \cdot \lambda_j - c_0 - c_1 \cdot \lambda_j \leq g_j \leq a_0 + a_1 \cdot \lambda_j; \\ 1 - \frac{g_j - a_0 - a_1 \cdot \lambda_j}{d_0 + d_1 \cdot \lambda_j}, & \text{если} \\ a_0 + a_1 \cdot \lambda_j \leq g_j \leq a_0 + a_1 \cdot \lambda_j + d_0 + d_1 \cdot \lambda_j; \\ 0, & \text{иначе;} \end{cases} \quad (16)$$

где $a_0, c_0, d_0, a_1, c_1, d_1$ – параметры ТНЧ, вычисленные при решении ЗКП (9) – (10).

На завершающем **третьем этапе** алгоритма для анализируемой и каждой эталонной ГСХ сначала проводится вычисление результирующей нечеткой меры подобия, определяемой как минимум из нечетких мер подобия F^{UP} и F^{LOW} :

$$F = \min(F^{UP}, F^{LOW}), \quad (17)$$

а затем все эталонные ГСХ упорядочиваются по убыванию вычисленных таким образом значений нечетких мер подобия, при этом в качестве искомой выбирается та эталонная ГСХ, для которой значение нечеткой меры подобия (17) является максимальным.

Алгоритм консолидации частных результатов идентификации ГСХ объекта

Как уже было отмечено, для повышения качества идентификации анализируемой ГСХ объекта целесообразно выполнение консолидации частных результатов идентификации (тем или иным образом) [20].

При применении четырех алгоритмов идентификации ГСХ к анализируемой ГСХ и некоторой эталонной ГСХ вычисляются четыре значения: значение меры сходства евклидова расстояния E (1); значение угловой меры подобия α (2); два значения нечетких мер подобия F_1 и F_2 , вычисленных соответственно с применением формул (14), (15) и (17).

При работе с БД СИОДГС, содержащей информацию об K эталонных ГСХ, для анализируемой ГСХ объекта будут получены четыре идентифицирующих набора, каждый из которых является результатом применения одного из четырех алгоритмов идентификации ГСХ и содержит K значений (в соответствии с количеством эталонных ГСХ в БД). Каждый такой идентифицирующий набор может быть упорядочен по убыванию (возрастанию) значений используемой меры подобия, в результате чего эталонным ГСХ из БД будут присвоены некоторые рейтинговые оценки (порядковые номера в рейтинге): чем выше значения меры подобия евклидова расстояния и угловой меры сходства и чем ниже значения нечетких мер подобия, тем меньше номер в рейтинге.

Пусть R_E^k – рейтинговая оценка k -й эталонной ГСХ при использовании алгоритма идентификации ГСХ на основе меры подобия евклидова расстояния E ; R_α^k – рейтинговая оценка k -й эталонной ГСХ при использовании алгоритма идентификации ГСХ на основе угловой меры подобия α ; $R_{F_1}^k$ – рейтинговая оценка k -й эталонной ГСХ при использовании алгоритма идентификации ГСХ на основе нечеткой меры подобия F_1 ; $R_{F_2}^k$ – рейтинговая оценка k -й эталонной ГСХ при использовании алгоритма идентификации ГСХ на основе нечеткой меры подобия F_2 ($k = \overline{1, K}$).

Алгоритм консолидации частных результатов идентификации ГСХ объекта предполагает:

консолидацию частных результатов идентификации, полученных с применением меры сходства евклидова расстояния E , угловой меры подобия α , а также двух

нечетких мер подобия F_1 и F_2 посредством применения формулы:

$$\bar{R}^k = \frac{(R_E^k + R_\alpha^k + R_{F_1}^k + R_{F_2}^k)}{4}; \quad (18)$$

упорядочение эталонных ГСХ из БД по возрастанию усредненных значений рейтинговых оценок \bar{R}^k ($k = \overline{1, K}$).

Пусть консолидирующие рейтинговые оценки \hat{R}^k определены как номера в рейтинге для усредненных оценок \bar{R}^k ($k = \overline{1, K}$). Очевидно, что для дальнейшего анализа целесообразно использовать незначительную часть лучших эталонных ГСХ из БД (от 5 до 10 эталонных ГСХ), имеющих наименьшие значения (порядковые номера) консолидирующих рейтинговых оценок \hat{R}^k .

После визуального сравнительного анализа результатов идентификации анализируемой ГСХ объекта с помощью некоторого количества лучших эталонных ГСХ из БД, полученных вышеуказанным образом и представленных как в текстовом, так и в графическом виде, оператор программы СИОДГС может принять окончательное решение о соответствии анализируемой ГСХ объекта некоторой эталонной ГСХ из БД.

Экспериментальные исследования

Предлагаемая СИОДГС была использована при решении задачи идентификации ГСХ различных объектов, в частности, была проанализирована ГСХ «Сосна, крона, Московская обл., 1 августа, 1970 г.», взятая с сайта GIS-LAB [3]. В БД на момент выполнения данного анализа находилось 281 эталонная ГСХ, включая анализируемую ГСХ.

При решении задачи идентификации анализируемой ГСХ с применением четырех рассмотренных выше алгоритмов идентификации были использованы все эталонные ГСХ из БД СИОДГС. В табл. 1 представлены пять лучших результатов идентификации анализируемой ГСХ, при этом указаны порядковые номера \hat{R}^k лучших эталонных ГСХ, определенные посредством упорядочения по возрастанию консолидирующих рейтинговых оценок \bar{R}^k ($k = \overline{1, K}$) всех эталонных ГСХ (при идентификации анализируемой ГСХ четырьмя алгоритмами идентификации), наименования лучших эталонных ГСХ с указанием их класса и типа, а также порядковые номера приведенных в таблице эталонных ГСХ, определенные при индивидуальном применении алгоритмов идентификации ГСХ к ГСХ «Сосна, крона, Московская обл., 1 августа, 1970 г.» на основе меры подобия евклидова расстояния, угловой меры подобия и нечетких мер подобия ($R_E, R_\alpha, R_{F_1}, R_{F_2}$). Следует отметить, что во всех представленных ниже таблицах, в том числе и в таблице 1, жирным шрифтом выделены порядковые номера трех лучших ГСХ, определенные как при индивидуальном применении алгоритмов идентификации ГСХ, так и в результате консолидации частных результатов идентификации ГСХ.

На рис. 3 и 4 приведены фрагменты главного окна программы, содержащие графическую интерпретацию результатов идентификации ГСХ. При этом показаны: анализируемая ГСХ, линия КЛР, границы коридора НЛР

и эталонная ГСХ (выбранная из таблицы результатов идентификации ГСХ). Точки анализируемой ГСХ изображены темными квадратами, точки эталонной ГСХ – светлыми окружностями.

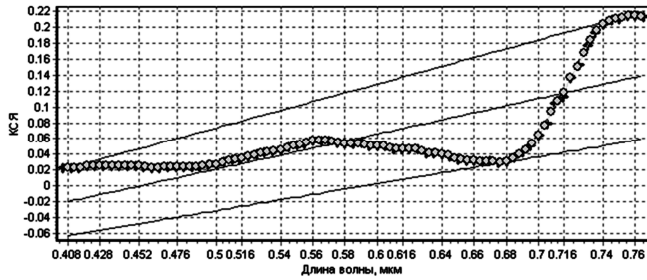


Рис. 3. Фрагмент главного окна программы с результатом идентификации, представленным в графическом виде (для эталона с общим порядковым номером, равным 1)

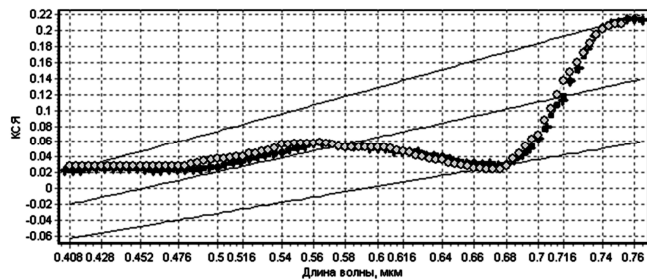


Рис. 4. Фрагмент главного окна программы с результатом идентификации, представленным в графическом виде (для эталона с общим порядковым номером, равным 2)

На рис. 3 представлены анализируемая ГСХ и эталонная ГСХ с порядковым номером 1 (табл. 1), соответствующей объекту «Сосна, крона, Московская обл., 1 августа, 1970 г.». При этом все точки анализируемой ГСХ полностью совпадают с точками эталонной ГСХ. На

рис. 4 представлены анализируемая ГСХ и эталонная ГСХ с порядковым номером 2 (табл. 1), соответствующей объекту «Сосна, крона, Московская обл., 6 августа, 1970 г.». При этом можно увидеть наличие некоторых расхождений между анализируемой и эталонной ГСХ.

Из табл. 1 видно, что все эталонные ГСХ лучших результатов идентификации принадлежат к тому же классу («Деревья») и к тому же типу («Растительность»), что и анализируемая ГСХ «Сосна, крона, Московская обл., 1 августа, 1970 г.» (хотя все эталонные ГСХ, участвовавшие в идентификации, разделены на 4 типа и 21 класс).

Следует отметить, что при использовании только одного алгоритма идентификации ГСХ, результат идентификации может быть менее точным и даже неверным. Так, например, из табл. 2, содержащей 5 лучших результатов идентификации анализируемой ГСХ с помощью только одного алгоритма идентификации ГСХ на основе угловой меры подобия R_α , видно, что, если бы БД не содержала эталонную ГСХ «Сосна, крона, Московская обл., 1 августа, 1970 г.», то результат идентификации – выбранная в качестве искомой эталонная ГСХ – мог бы иметь неправильно определенные класс и тип анализируемой ГСХ. В то же время консолидация частных результатов идентификации ГСХ, полученных с применением четырех алгоритмов идентификации, позволяет

установить более точные рейтинговые оценки \bar{R}^k эталонных ГСХ из БД и избежать принятия ложного решения о результатах идентификации анализируемой ГСХ (при отсутствии одноименного эталона ГСХ в БД) как ГСХ «Чернозем типичный, гор. А1 (0 - 50 см), 1975 г.» (табл. 2, вторая строка). При этом анализируемая ГСХ была бы идентифицирована как ГСХ «Сосна, крона, Московская обл., 6 августа, 1970 г.» (табл. 1, вторая строка) с правильно определенными классом и типом.

Таблица 1. Лучшие результаты идентификации ГСХ «Сосна, крона, Московская обл., 1 августа, 1970 г.» (консолидация)

Порядк. номер \hat{R}^k	Наименование объекта с эталонной ГСХ	Класс	Тип	Среднее арифм. \bar{R}^k	R_E	R_α	R_{F_1}	R_{F_2}
1	Сосна, крона, Московская обл., 1 августа, 1970 г.	Деревья	Растительность	1	1	1	1	1
2	Сосна, крона, Московская обл., 6 августа, 1970 г.	Деревья	Растительность	6,25	4	8	10	3
3	Сосна, крона, Московская обл., 23 августа, 1970 г.	Деревья	Растительность	7	10	11	5	2
4	Береза, крона, Московская обл., 6 августа, 1970 г.	Деревья	Растительность	14	3	10	24	19
5	Береза, крона, Московская обл., 14 июля, 1970 г.	Деревья	Растительность	14,25	18	12	15	12

Таблица 2. Лучшие результаты идентификации ГСХ «Сосна, крона, Московская обл., 1 августа, 1970 г.» на основе угловой меры подобия (2)

R_α	Наименование объекта с эталонной ГСХ	Класс	Тип	Порядк. номер \hat{R}^k	Среднее арифм. \bar{R}^k	R_E	R_{F_1}	R_{F_2}
1	Сосна, крона, Московская обл., 1 августа, 1970 г.	Деревья	Растительность	1	1	1	1	1
2	Чернозем типичный, гор. А1 (0 - 50 см), 1975 г.	Почвы	Грунты	98	24,5	15	219	200
3	Дерново-карбонатная почва. А1 (0 - 12 см), 1975 г.	Почвы	Грунты	153	38,25	74	280	280
4	Дерново-карбонатная почва А1 (>12 см), 1975 г.	Почвы	Грунты	203	50,75	208	281	281
5	Сосна, крона, Московская обл., 24 июня (молодые побеги), 1970 г.	Деревья	Растительность	25	6,25	136	2	7

Таблица 3. Лучшие результаты идентификации ГСХ «Асфальтированное шоссе сухое, 17 июля, 1960 г.» на основе угловой меры подобия (2)

R_{α}	Наименование объекта с эталонной ГСХ	Класс	Тип	Порядк. номер \hat{R}^k	Среднее арифм. \bar{R}^k	R_E	R_{F_1}	R_{F_2}
1	Асфальтированное шоссе сухое, 17 июля, 1960 г.	Сооружения	Антропогенные объекты	1	1	1	1	1
2	Чернозем обыкновенный, почва суббореального пояса	почвы	Грунты	52	87,25	108	85	104
3	Шоссе с асфальтовым покрытием, сухое, 17 июля 1961 г.	сооружения	Антропогенные объекты	2	4,25	2	7	6
4	Дерново-карбонатная почва. А1 (0 - 12 см).	почвы	Грунты	141	173,75	124	231	199
5	Каолин	горные породы	Грунты	61	95	278	20	21

Таблица 4. Лучшие результаты идентификации ГСХ «Асфальтированное шоссе сухое, 17 июля, 1960 г.» на основе нечеткой меры подобия (14)

R_{F_1}	Наименование объекта с эталонной ГСХ	Класс	Тип	Порядк. номер \hat{R}^k	Среднее арифм. \bar{R}^k	R_{α}	R_E	R_{F_2}
1	Асфальтированное шоссе сухое, 17 июля, 1960 г.	Сооружения	Антропогенные объекты	1	1	1	1	1
2	Средние горные породы: трахит	Горные породы	Грунты	46	79,25	18	251	2
3	Береза, крона, Московская обл., 1 октября 1970 г.	Деревья	Растительность	88	152,75	256	264	3
4	Такыр типичный, распаханная почва субтропического пояса	Почвы	Грунты	58	91,5	65	236	7
5	Грубообломочные горные породы: желтовато-серый мелкогалечный конгломерат	Горные породы	Грунты	4	28,5	84	22	4

Таблица 5. Лучшие результаты идентификации ГСХ «Асфальтированное шоссе сухое, 17 июля, 1960 г.» на основе меры подобия евклидова расстояния (1)

R_E	Наименование объекта с эталонной ГСХ	Класс	Тип	Порядк. номер \hat{R}^k	Среднее арифм. \bar{R}^k	R_{α}	R_{F_1}	R_{F_2}
1	Асфальтированное шоссе сухое, 17 июля, 1960 г.	Сооружения	Антропогенные объекты	1	1	1	1	1
2	Шоссе с асфальтовым покрытием, сухое, 17 июля 1961 г.	Сооружения	Антропогенные объекты	2	4,5	3	7	6
3	Почва: солонец мелкий	Почвы	Грунты	15	44	48	61	52
4	Песчаные подзолистые почвы. Содержание влаги в почве - 0-4%	Почвы	Грунты	43	76,25	21	10 8	133
5	Шоссе, мощенное камнем, сухое, 17 июля 1960 г.	Сооружения	Антропогенные объекты	5	30	16	46	53

Алгоритмы идентификации ГСХ на основе угловой меры подобия и нечетких мер подобия также не всегда дают приемлемые по точности результаты идентификации ГСХ [в частности, при идентификации ГСХ сухого асфальтированного шоссе алгоритмы идентификации ГСХ на основе угловой меры подобия и нечетких мер подобия дают менее точные результаты идентификации, чем алгоритм идентификации ГСХ на основе меры сходства евклидова расстояния (табл. 3–5)].

Так, например, из табл. 3 и 4, содержащих 5 лучших результатов идентификации ГСХ «Асфальтированное шоссе сухое, 17 июля, 1960 г.» соответственно с помощью только одного алгоритма идентификации ГСХ на основе угловой меры подобия (2) R_{α} и только одного алгоритма идентификации ГСХ на основе нечеткой меры (14) R_{F_1} видно, что, если бы БД не содержала эталонную ГСХ «Асфальтированное шоссе сухое, 17 июля, 1960 г.», то результаты идентификации – выбранные в качестве искомой эталонные ГСХ «Чернозем обыкновенный, почва суббореального пояса» (табл. 3, строка 2)

и ГСХ «Средние горные породы: трахит» (табл. 4, строка 2) – могли бы иметь неправильно определенные класс и тип анализируемой ГСХ.

В тоже время из табл. 5, содержащей 5 лучших результатов идентификации ГСХ «Асфальтированное шоссе сухое, 17 июля, 1960 г.» с помощью только одного алгоритма идентификации ГСХ на основе меры сходства евклидова расстояния (1) R_E , видно, что, даже если бы БД не содержала эталонную ГСХ «Асфальтированное шоссе сухое, 17 июля, 1960 г.», то результат идентификации – выбранная в качестве искомой эталонная ГСХ «Шоссе с асфальтовым покрытием, сухое, 17 июля 1961 г.» (табл. 5, строка 2) – имел бы, по крайней мере, правильно определенные класс и тип анализируемой ГСХ.

При этом консолидация частных результатов идентификации ГСХ «Асфальтированное шоссе сухое, 17 июля, 1960 г.», полученных с применением четырёх предлагаемых алгоритмов идентификации, позволяет принять правильное решение по идентификации анализируемой ГСХ.

Приведенные примеры свидетельствуют о целесообразности совместного применения нескольких алгоритмов идентификации ГСХ с целью последующей консолидации частных результатов идентификации.

Заключение

Результаты экспериментальных исследований подтверждают целесообразность дальнейшего развития предлагаемого подхода к решению задачи идентификации ГСХ объектов, основанного на консолидации частных результатов идентификации ГСХ объектов, полученных с применением различных обоснованно выбранных алгоритмов идентификации ГСХ объектов, с целью повышения надежности классификационного решения. Использование такого подхода позволит решить задачу идентификации объектов земной поверхности по выделенным с обработанных космических снимков КА «Ресурс-П» гиперспектральным характеристикам с последующим накоплением эталонных характеристик в базе данных, что, в свою очередь, обеспечит создание актуальной отечественной спектральной библиотеки эталонов, которую можно будет применять при мониторинге состояния сельскохозяйственных угодий, лесных массивов, водных ресурсов, экологического состояния почв и т.п.

Литература

1. USGS Spectroscopy Lab. <http://speclab.cr.us-gs.gov>.
2. Jet Propulsion Laboratory. ASTER Spectral Library. NASA. <http://speclab.jpl.nasa.gov>.
3. Спектральные библиотеки – источник данных по спектрам. GIS-LAB. <http://www.gis-lab.info>.
4. Программный комплекс ENVI. Уч. пособие. Компания «Совзонд». – 2009. <http://www.sovzond.ru>.
5. Шовенгердт Р.А. Дистанционное зондирование. Модели и методы обработки изображений. – М.: Техносфера, 2010. – 560 с.
6. Чандра А.М., Кош С.К. Дистанционное зондирование и географические информационные системы. – М.: Техносфера, 2008. – 312 с.
7. Рис У.Г. Основы дистанционного зондирования. – М.: Техносфера, 2006. – 336 с.
8. Пылькин А.Н., Тишкин Р.В., Труханов С.В. Задачи DATA MINING и их решение в современных реляционных СУБД // Вестник Рязанского государственного радиотехнического университета. – 2011. – №4 (выпуск 38). – С. 60-65.
9. Чубукова И.А. Data Mining. Основы информационных технологий. Специальные курсы. Издательство «Бинном». Лаборатория знаний, 2006. – 384 с.
10. Yang C., Everitt J H., Bradford J.M. Yield estimation from hyperspectral imagery using spectral angle mapper (SAM). – American Society of Agricultural and Biological Engineers. Vol. 51(2): 729-737.
11. Van der Weken D., Nachtegaele M., Kerre E.E. An overview of similarity measures for images // Proceedings of

ICASSP 2002 (IEEE Int. Conf. Acoustics, Speech and Signal Processing). – Orlando, USA. – 2002. – pp. 3317-3320.

12. Пылькин А.Н., Тишкин Р.В. Методы и алгоритмы сегментации изображений. – М.: Горячая линия – Телеком, 2010. – 92 с.

13. Мятлов Г.Н., Тишкин Р.В., Ушенкин В.А., Юдаков А.А. Применение нечетких мер подобия в задаче совмещения изображений поверхности Земли // Вестник Рязанского государственного радиотехнического университета. – 2013. – №2 (выпуск 44). – С. 18-26.

14. Труханов С.В., Юдаков А.А. Создание структуры базы данных системы интеллектуальной обработки данных гиперспектральной съемки. Свидетельство о государственной регистрации программы для ЭВМ в Федеральной службе по интеллектуальной собственности №2013611036 от 09.01.2013.

15. Труханов С.В., Юдаков А.А. Программа интеллектуальной обработки данных гиперспектральной съемки. Свидетельство о государственной регистрации программы для ЭВМ в Федеральной службе по интеллектуальной собственности №2013610619 от 09.01.2013.

16. Кремер Н.Ш., Путко Б.А., Тришин И.М., Фридман М.Н. Высшая математика для экономистов: Учебник для вузов, М.: ЮНИТИ, 2002. – 471 с.

17. Демидова Л.А., Мятлов Г.Н. Подход к оценке уникальности кусочно-линейных объектов с использованием нечеткой линейной регрессии // Системы управления и информационные технологии. 2013. Т. 51. № 1. С. 85-89.

18. Haekwan Lee, Hideo Tanaka Fuzzy approximations with non-symmetric fuzzy parameters in fuzzy regression analysis. Osaka prefecture University. Journal of the Operations Research Society of Japan. Vol.42, №. 1, March 1999.

19. Ануфриев Е.И., Смирнов А.Б., Смирнова Е.Н. MATLAB 7. – СПб.:БХВ-Петербург, 2005. – 1104 с.

20. Демидова Л.А., Тишкин Р.В., Юдаков А.А. Разработка ансамбля алгоритмов кластеризации на основе матриц подобия меток кластеров и алгоритма спектральной факторизации // Вестник Рязанского государственного радиотехнического университета. 2013. № 4-1 (46). С. 9-17.

THE OBJECTS HYPERSPECTRAL FEATURE IDENTIFICATION ALGORITHMS IN THE EARTH REMOTE SENSING TASKS

Demidova L.A., Tishkin R.V., Trukhanov S.V.

The approach to resolve the task of Earth objects' identification founded on hyperspectral objects' feature analyze, which are got from processed space images, by means of identification algorithms, based on the different similarity measures, is considered. The results of hyperspectral information processing by means of the offered identification algorithms, are represented.