

УДК 629.7

## МЯГКИЕ ВЫЧИСЛЕНИЯ В ЗАДАЧАХ СЕГМЕНТАЦИИ КОСМИЧЕСКИХ ИЗОБРАЖЕНИЙ

Тишкин Р.В., начальник отдела, филиал ФГУП «ГНПРКЦ «ЦСКБ-Прогресс» - ОКБ «Спектр»,  
4374@spectr.ryazan.ru

**Ключевые слова:** дистанционное зондирование Земли, искусственные нейронные сети, нечеткая логика, сегментация изображений.

### Введение

В настоящее время значительно увеличился поток информации, получаемой от космических аппаратов дистанционного зондирования Земли (КА ДЗЗ). Растет и число потребителей. Если раньше работать с высокодетальной информацией могли себе позволить лишь ряд организаций, то в настоящее время практически любой человек способен получить такую информацию. В связи с этим возрастает и номенклатура задач, решаемых с использованием данных ДЗЗ. При этом возникает необходимость в автоматизации не только предварительной обработки данных ДЗЗ (до уровня 4), но и обработки тематической, т.е. обработки, нацеленной на конечного потребителя. Здесь ведущую роль начинает играть сегментация изображений, которая представляется собой первый шаг к переходу от количественного представления изображения к качественному. На сегодняшний день, в большинстве случаев, тематическая обработка изображений производится в ручном либо интерактивном режиме. Для того чтобы добиться автоматического режима нужно использовать другие средства, например мягкие вычисления.

Термин «Мягкие вычисления» (Soft computing) впервые ввел Л. Заде в своей работе [1]. На сегодняшний день под мягкими вычислениями понимается совокупность методов, использующихся для решения задач с высокой степенью неопределенности. Это искусственные нейронные сети (ИНС), нечеткая логика (НЛ) и эволюционные вычисления, в различных комбинациях.

Рассмотрим возможности данных методов на примере сегментации космических изображений.

### Искусственные нейронные сети

Искусственные нейронные сети (ИНС) для обработки изображений и распознавания образов используются достаточно длительное время и многими исследователями [2, 3]. Данное обстоятельство обусловлено особыми и важными свойствами ИНС, в первую очередь положительными качествами в распознавании образов [2, 6].

Применение ИНС непосредственно в задачах обработки аэрокосмических изображений достаточно подробно приведено в [4, 5]. Предложено использовать модель многослойного персептрона в сочетании с алгоритмом обратного распространения ошибки, предназначенным для обучения (с учителем) ИНС. Подобный под-

*В статье рассматривается вопрос о применении методов и алгоритмов относящихся к «мягким вычислениям» (искусственные нейронные сети, нечеткая логика, генетические алгоритмы) в задаче автоматической сегментации космических изображений Земли.*

ход положительно зарекомендовал себя при решении разного рода задач искусственного интеллекта в нашей стране и за рубежом. Однако для решения задачи автоматической сегментации изображений он подходит слабо, т.к. требует участия оператора. Существуют и другие типы ИНС, в работе которых участие оператора не требуется.

В качестве примера рассмотрим задачу сегментации облачности на космических изображениях с использованием помощью ИНС «модель Кохонена» [2].

Вектором  $\vec{f} = [f_R, f_G, f_B]$  обозначим многозональное спутниковое изображение, где  $f_Z = f_Z(x, y)$ ;  $x = \overline{1, N_x}$ ;  $y = \overline{1, N_y}$ ;  $Z \in \{R, G, B\}$ . ИНС задается матрицей  $W = w(i, j)$  коэффициентов сети, в которой  $w_j$  - некоторый  $j$ -й столбец матрицы весовых коэффициентов.

На вход ИНС подается вектор  $\vec{X} = (X_R, X_G, X_B)$  - входной вектор признаков, описывающий область изображения  $r \times r$ , где  $f_Z = f_Z(x+s, y+t)$ ;  $s = \overline{1, r}$ ;  $t = \overline{1, r}$ ;  $x = \overline{1, (r)N_x}$ ;  $y = \overline{1, (r)N_y}$ ;  $Z \in \{R, G, B\}$ .

При введенных обозначениях алгоритм обучения ИНС может быть представлен в виде последовательности следующих шагов.

1. На обрабатываемом изображении случайным образом выделяется область размером  $r \times r$  и формируется вектор  $\vec{X}$ . Этот шаг выполняется многократно, что позволяет сформировать обучающую выборку (без привлечения оператора).

2. Для каждого вектора  $\vec{X}$  из обучающей выборки определяется функция

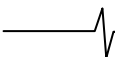
$$i(x) = \arg \min_j \| \vec{X} - w_j \|,$$

где  $j = \overline{1, q}$ ;  $q$  - число нейронов сети в выходном слое.

1. Для нейрона, вектор весов которого наиболее близок к входному вектору, корректируются весовые коэффициенты по формуле:  $w_j(l+1) = w_j(l) + \eta(l)(\vec{X} - w_j(l))$ , где  $\eta(l)$  - параметр скорости обучения;  $w_j(l)$  - вектор синтаксических весов  $j$ -го нейрона сети на шаге обучения  $l$ .

2. Обработка, реализуемая шагами 2 и 3, повторяется заданное число раз.

Параметр  $\eta(l)$  позволяет варьировать скорость изменения весовых коэффициентов нейронной сети и в процессе обучения может изменяться по некоторому закону [6].



Процесс работы обученной ИНС “модель Кохонена” описывается следующей последовательностью шагов.

1. На обрабатываемом изображении последовательно выделяется область размером  $r \times r$  и формируется вектор  $\vec{X}$ .

2. Для каждого вектора  $\vec{X}$  определяется функция  $i(x) = \arg \min_j \|\vec{X} - w_j\|$ ,

где  $j = \overline{1, q}$ ;  $q$  - число нейронов сети в выходном слое.

3. На основании результатов анализа определяется нейрон “победитель”, т.е. исследуемая область относится к одному из классов (число определяемых классов, сегментов равно числу нейронов выходного слоя ИНС).

4. Пункты 1-3 повторяются последовательно для всех фрагментов  $r \times r$ , на которые разбивается обрабатываемое изображение.

С помощью данного алгоритма может быть решена задача локализации облачности на космических многозональных изображениях. Результаты работы алгоритма представлены на рис. 1. Приведенные примеры свидетельствуют о том, что алгоритм сегментации позволяет определять как плотную, так и легкую облачность.

Однако сравнительный анализ полученных результатов обработки спутниковых снимков показал ряд недостатков, которые наблюдаются у ИНС в случае применения её для решения рассматриваемой задачи. К недостаткам рассмотренного алгоритма можно отнести следующее:

- метод обучения является чисто эвристическим, и завершение процедуры обучения не основывается на строгих оптимизационных математических моделях;

- обучение реализуется на основе случайного выбора, поэтому итоговые весовые коэффициенты выходных нейронов зависят от входной последовательности;

- различные начальные условия, как правило, приводят к различным результатам обработки.

В таблице 1 приведены результаты расчета ошибок ложного обнаружения и ложного пропуска, величины которых обеспечивает предложенный алгоритм. Отметим, что расчет ошибок может быть выполнен вручную оператором путем отнесения того или иного выделенного участка к некоторому классу и сравнения полученных результатов с результатами автоматической обработки. Совершенно очевидно, что полученные результаты сильно зависят от случайно набранной обучающей выборки.

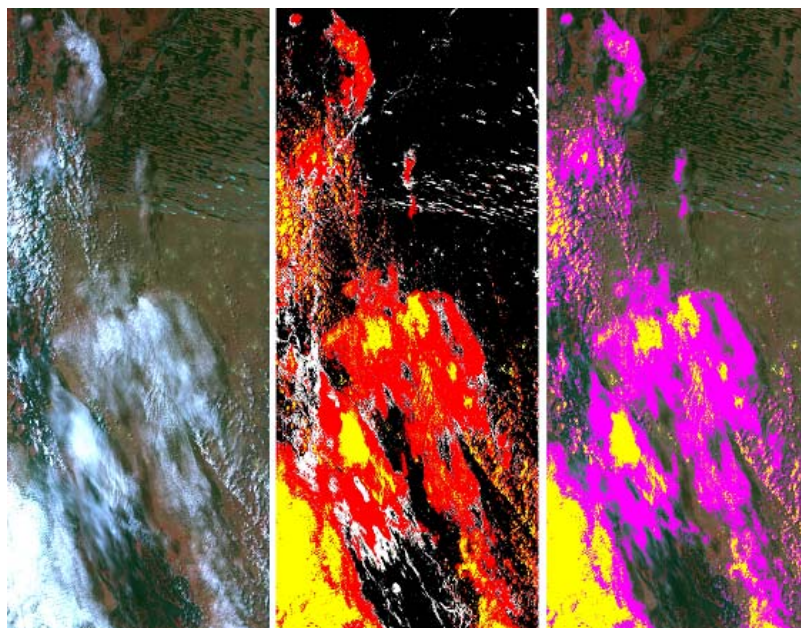


Рис. 1.

Таблица 1.

Оценка доли ложных обнаружений и ложных пропусков

| Ложные обнаружения и пропуски | Алгоритм сегментации изображений на основе ИНС “модель Кохонена” |           |           |           |           |
|-------------------------------|------------------------------------------------------------------|-----------|-----------|-----------|-----------|
|                               | Пример 1                                                         | Пример 2  | Пример 3  | Пример 4  | Пример 5  |
| обнаружение                   | 0,0328                                                           | 0,096152  | 0,000955  | 0,031198  | 0,016091  |
| пропуск                       | 0,037897                                                         | 0,027064  | 0,037756  | 0,042287  | 0,037647  |
|                               | Пример 6                                                         | Пример 7  | Пример 8  | Пример 9  | Пример 10 |
| обнаружение                   | 0,17092                                                          | 0,21026   | 0,19331   | 0,09027   | 0,17265   |
| пропуск                       | 0,005958                                                         | 0,0039645 | 0,005845  | 0,031599  | 0,0020351 |
|                               | Пример 11                                                        | Пример 12 | Пример 13 | Пример 14 | Пример 15 |
| обнаружение                   | 0,1411                                                           | 0,18016   | 0,22046   | 0,055055  | 0,19998   |
| пропуск                       | 0,0082896                                                        | 0,0003105 | 0,026456  | 0,034144  | 0,0006707 |
|                               | Пример 16                                                        | Пример 17 | Пример 18 | Пример 19 | Пример 20 |
| обнаружение                   | 0,04038                                                          | 0,040127  | 0,082472  | 0,051243  | 0,12407   |
| пропуск                       | 0,02234                                                          | 0,077534  | 0,013937  | 0,037225  | 0,017542  |

Улучшением данного алгоритма является модификация процедуры генерации обучающей выборки, которая заключается в следующем. Предварительно изображение обрабатывается с помощью колориметрического алгоритма [7] с целью получения “грубой” карты облачности.

После чего на сегментируемом изображении, с учетом грубой карты облачности, выделяются фрагменты для обучающей выборки. Это происходит с таким расчетом, чтобы в равных пропорциях в обучающую выборку попали области соответствующие облачности, несоответствующие облачности и находящиеся на границе.

Использование такого подхода позволяет значительно стабилизировать результаты работы ИНС Кохонена. Однако необходимость в предварительной обработке изображения отрицательно сказывается на быстродействии алгоритма в целом. Кроме того, из-за использования предварительной маски облачности алгоритм практически теряет возможность идентификации легкой облачности [5].

В целом для решения задач сегментации космических изображений наиболее подходящей оказалась ИНС многослойный перцептрон с обучением по алгоритму обратного распространения ошибки. При использовании такой ИНС полностью исключить работу оператора невозможно. ИНС на основе самоорганизации нуждаются в дополнительных алгоритмах и процедурах пред- и постобработки данных, иначе результаты носят стохастический характер. Исходя из этого, можно говорить о том, что в «чистом» виде для создания системы автоматической сегментации космических изображений ИНС не пригодны. ИНС могут быть составной частью комплексной системы сегментации космических изображений Земли.

**Нечеткая логика**

Методы и алгоритмы нечеткой логики [1] стали использоваться в обработке изображений сравнительно недавно. Практически все предлагаемые методы сегментации изображений основаны на алгоритме нечетких  $k$ -средних (Fuzzy C Means, FCM) или его модифицированных версиях. Алгоритм FCM, в свою очередь, основан на алгоритме  $k$ -средних. Оба алгоритма достаточно подробно описаны в литературе [8, 9].

Исходными данными алгоритма FCM служат:  $X$  – матрица наблюдений обрабатываемого изображения ( $X = [x_{ij}]$ ;  $i = 1, N$ ;  $j = 1, p$ );  $k$  – количество кластеров;  $m$  – экспоненциальный вес;  $\varepsilon$  – параметр останова алгоритма.

Наиболее важным параметром алгоритма FCM является количество кластеров ( $k$ ). Выбор количества кластеров, как правило, осуществляется для различных задач по результатам анализа априорной информации. Существуют методы автоматического выбора оптимального числа кластеров, большинство из них основаны на различных критериях оценки компактности кластеров, среди которых можно выделить индекс Хие–Бени [10] как наиболее оптимальный.

Другим параметром алгоритма FCM является экспоненциальный вес ( $m$ ). При больших  $m$  конечная матрица нечеткого разбиения  $F_H$  “размывается”. При

$m \rightarrow \infty$  она принимает вид  $F_H = [1/k]$  (что является плохим решением, т.к. все объекты принадлежат всем кластерам с одной и той же степенью). На сегодня теоретически обоснованного решения по правилам выбора  $m$  не существует. Практически значение экспоненциального веса устанавливается в пределах [2:5].

В «чистом» виде алгоритмы сегментации изображений, основанные на FCM, обнаруживает несколько недостатков:

- решение может сводиться к нахождению локального оптимума, поэтому получаемые результаты могут зависеть от начальных параметров алгоритма;
- приемлемое решение находится в результате многократного прогона итеративных вычислений, что в значительной степени усложняет процесс обработки в реальном или квазиреальном масштабе времени.

Можно предположить, что использование нечетких множеств 2-го типа, позволит решить хотя бы часть проблем.

**Нечеткие множества второго типа**

Одной из последних тенденций в развитии нечеткой логики стало применение в системах искусственного интеллекта и принятия “мягких” решений нечетких множеств второго типа (type 2 fuzzy set, T2FS). По мнению некоторых зарубежных и отечественных специалистов использование методов обработки на основе T2FS может привести к заметному улучшению получаемых решений. Впервые идея нечетких множеств второго типа предложена Заде [1] и получила развитие в работах Менделя [11].

Понятие нечеткого множества 2-го типа в алгоритме нечеткой кластеризации элементов изображения следует использовать следующим образом. Экспоненциальный вес  $m$  определяется не точечным значением, а некоторым “нечетким” значением из интервала  $(m_1, m_2)$ , т.е.  $m \in (m_1, m_2)$ .

Обоснованность такого подхода рассмотрена в [5].

T2FS  $A$  представляется в форме нечеткой системы

$$A = \{((x, u), \mu_A(x, u)) \mid \forall x \in A,$$

$$\forall u \in J_x \subseteq [0, 1], \mu_A(x, u) = 1\},$$

в которой первичная функция принадлежности  $J_{x_i}$  образца  $x_i$  представляется интервальной функцией принадлежности со всеми вторичными степенями первичных функций принадлежности, равными 1.

Для задания интервальной первичной функции принадлежности определим нижнюю и верхнюю интервальные функции принадлежности при значениях фаззификатора  $m_1, m_2$ . Первичные функции принадлежности, представляющие образец  $x_i$ , определяются следующим образом:

$$\mu_j^{(B)}(x_i) = \begin{cases} \frac{1}{\sum_{l=1}^k \left(\frac{D_{ji}}{D_{li}}\right)^{\frac{2}{m_1-1}}}, & \text{если } \frac{1}{\sum_{l=1}^k \left(\frac{D_{ji}}{D_{li}}\right)^{\frac{2}{m_1-1}}} > \frac{1}{\sum_{l=1}^k \left(\frac{D_{ji}}{D_{li}}\right)^{\frac{2}{m_2-1}}} \\ \frac{1}{\sum_{l=1}^k \left(\frac{D_{ji}}{D_{li}}\right)^{\frac{2}{m_2-1}}}, & \text{в противном случае} \end{cases} \quad (1)$$

$$\mu_j^{(H)}(x_i) = \begin{cases} \frac{1}{\sum_{l=1}^k \left(\frac{D_{ji}}{D_{li}}\right)^{\frac{2}{m_1-1}}}, & \text{если } \frac{1}{\sum_{l=1}^k \left(\frac{D_{ji}}{D_{li}}\right)^{\frac{2}{m_1-1}}} \leq \frac{1}{\sum_{l=1}^k \left(\frac{D_{ji}}{D_{li}}\right)^{\frac{2}{m_2-1}}}; \\ \frac{1}{\sum_{l=1}^k \left(\frac{D_{ji}}{D_{li}}\right)^{\frac{2}{m_2-1}}}, & \text{в противном случае} \end{cases} \quad (2)$$

В (1) и (2) фаззификаторы  $m_1$  и  $m_2$  представляют различные нечеткие степени принадлежности, использование некоторых приводит к различным целевым функциям, которые должны быть минимизированы с помощью алгоритма FCM.

В результате выполнения алгоритма для каждого кластера сегментируемого изображения определяются два значения центра кластера  $V_j^{(H)}$  и  $V_j^{(B)}$  (в пространстве признаков сегментируемого изображения), поэтому в модернизированный алгоритм FCM предлагается ввести следующую процедуру расчета центров кластеров:

$$V_j = \frac{V_j^{(H)} + V_j^{(B)}}{2}. \quad (3)$$

Модернизацию центров кластеров следует выполнять в ходе реализации алгоритма FCM. Для этого используются операции "понижения типа" с помощью обобщенного центроида, применение которых подобно методу центра тяжести в базовом алгоритме кластеризации. Понижение типа во время поиска центров кластеров реализуется с помощью итерационного алгоритма Карника-Менделя [12]. Основным недостатком данного подхода приходится признавать резкое увеличение сложности вычислений и, как следствие, возрастание времени обработки, что особенно критично при обработке многозональных спутниковых снимков. Предлагается следующая модификация алгоритма FCM: определяются два значения центров каждого кластера  $V_j^{(H)}$  и  $V_j^{(B)}$ ,  $j = \overline{1, k}$ . Полученные значения участвуют в модернизации центров кластера в соответствии с формулой (3). Подобная процедура модернизации центров кластеров будет более "грубой" по сравнению с предлагаемым в [12] использованием итерационного алгоритма Карника-Менделя. Алгоритм Карник-Менделя обеспечивает последовательную модернизацию центров кластеров, однако в основном (итерационном) цикле алгоритма FCM необходима реализация еще одного итерационного алгоритма. Подобное обстоятельство резко увеличивает время обработки. Предлагаемая модернизация алгоритма позволяет заметно улучшить результаты обработки по сравнению с базовым алгоритмом FCM (уступая по эффективности алгоритму из [12]), сохраняя сопоставляемую сложность вычислений.

Качественный анализ результатов обработки изображения показывает, что с увеличением разницы между значениями фаззификаторов  $m_1$  и  $m_2$  результаты обработки сглаживаются.

### Гибридные системы сегментации изображений

Естественным следует признать применение в алгоритмах одного из методов эвристической оптимизации, например генетического алгоритма, моделирующего процессы природной эволюции и относящегося к эво-

люционным методам поиска [5, 13].

В общем случае генетическим алгоритмом называется следующий объект:

$$GA(P^0, r, l, sl, Fit, cr, mu, ot), \quad (4)$$

где  $GA$  – генетический алгоритм;  $P^0$  – исходная популяция;  $r$  – количество элементов популяции;  $l$  – длина битовой строки, кодирующей решение;  $sl$  – оператор селекции;  $Fit$  – фитнес-функция (функция приспособленности), определяющая "пригодность" решения;  $cr$  – оператор кроссинговера ("скрещивания") определяющий возможность нового решения;  $mu$  – оператор мутации;  $ot$  – оператор отбора.

Эффективный генетический алгоритм может быть построен только в том случае, если параметры и операторы в (4) будут выбраны на основе качественного анализа предметной области.

При решении задачи сегментации космических изображений структура хромосомы может быть представлена битовой строкой, в которой закодированы координаты центров кластеров. На рис. 2 приведен пример структуры хромосомы.

|           |   |   |           |   |   |     |   |   |           |   |   |
|-----------|---|---|-----------|---|---|-----|---|---|-----------|---|---|
| Cluster 1 |   |   | Cluster 2 |   |   | ... |   |   | Cluster k |   |   |
| R         | G | B | R         | G | B | R   | G | B | R         | G | B |

Рис. 2.

Размер популяции выбирается из следующих соображений. При использовании  $r > 20$  резко возрастает время работы алгоритма и требования к размеру оперативной памяти. Если  $r < 20$ , возникают случаи нестабильной работы генетического алгоритма, т.е. решение может быть неустойчивым.

В разработанном алгоритме применяется схема, так называемого классического ГА, в котором предполагается использовать строки постоянной длины (в мобильном ГА строки имеют переменную длину, т.е. могут быть переопределены и не доопределены). Применение оператора мутации  $mu$  в процессе реализации генетического алгоритма предназначено для выхода из локальных экстремумов. Практика применения генетических алгоритмов для решения задачи оптимизации рекомендует использовать классические процедуры генной мутации: изменение величины случайно выбранного гена. Мутация подвергается, как правило, до 5% особей популяции. В предлагаемом алгоритме реализуется стохастическое изменение нескольких бит хромосомы. Мутация выполняется с вероятностью 0,01..0,05 для каждой хромосомы и предполагает изменение от 1 до 5 случайно выбранных бит на их противоположенное значение.

Выбор оператора селекции  $sl$  родительской пары определяет качество поколений потомков. Наиболее часто в генетических алгоритмах используются следующие виды  $sl$  – операторов:

- случайный выбор пар;
- выбор пар на основе определения фитнес-функции  $Fit$ .



В алгоритме используется второй подход. В качестве фитнес-функции, при решении задачи сегментации, предлагается использовать индекс Хие-Бени [14]. Индекс Хие-Бени представляет из себя функцию следующего вида:

$$HB = \frac{\sum_{i=1}^k \sum_{j=1}^N \mu_{ij}^2 \|V_i - X_j\|^2}{N \cdot \min_{i,j} \|V_i - V_j\|^2}, \quad (5)$$

где  $k$  – число кластеров;  $N$  – число объектов;  $\mu_{ij}$  – функция принадлежности;  $V_i$  – координаты центра  $i$ -го кластера;  $X_j$  – некоторый  $j$ -й объект.

С помощью индекса Хие-Бени можно оценивать качество нечеткой кластеризации: чем меньше значение  $HB$ , тем разбиение элементов на классы выполнено качественнее. Та хромосома, которая обеспечивает наименьшее значение функции (5), признается более приспособленной, чем другие сравнимые варианты.

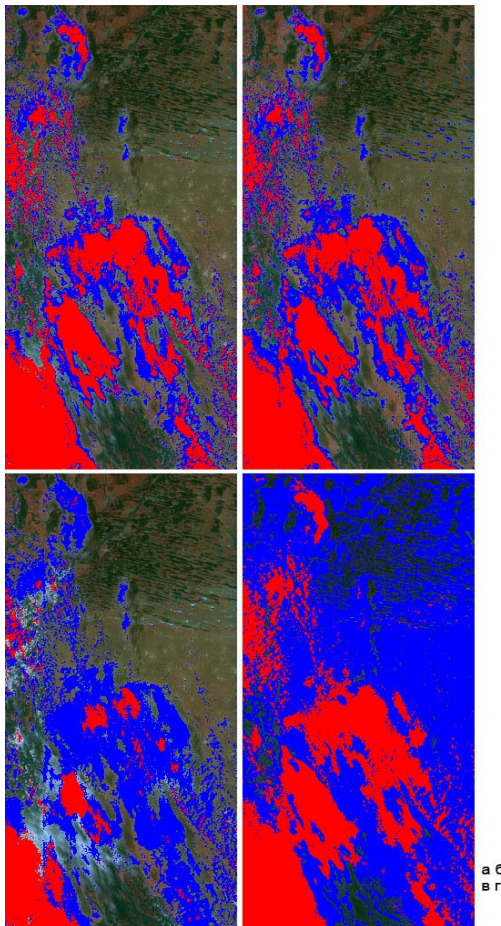


Рис. 3.

На рис. 3 приведены примеры идентификации облачности с помощью генетического алгоритма. Получены результаты обработки при 10 (r), 20 (v), 50 (б) и 100 (а) сменах поколений хромосом. Из приведенных рисунков видно, что при числе популяций 100 достаточно точно и надежно удается выявить образования типа “облачность” (красный цвет) и “полуоблачность” (синий цвет). Попытка увеличения числа популяций до 200 и выше не обеспечивает значительных преимуществ, но при этом резко увеличивается время работы алгоритмов и задействованный объем оперативной информации,

т.к. для каждой хромосомы требуется генерировать свою матрицу функций нечеткой принадлежности.

В целом анализ полученных результатов позволяет сделать вывод о том, что применение генетического алгоритма, обеспечивает нахождение субоптимального решения, но не приводит к какому-либо выигрышу по критерию быстродействия.

### Заключение

Алгоритм FCM показывает достаточно хорошие характеристики по точности, однако с учетом указанных недостатков, не может считаться основным алгоритмом при решении задачи сегментации космических изображений. Использование генетических алгоритмов позволяет обеспечить снижение ошибки ложного обнаружения, но повышает ошибку ложного пропуска. Алгоритм, основанный на T2FS, обеспечивает наиболее сбалансированные показатели по соотношению ошибок ложного обнаружения и ложного пропуска. Следует отметить, что одна из основных проблем заключается в скорости обработки. Если не использовать современные подходы к увеличению скорости обработки, то реальное использование предложенных алгоритмов в задачах автоматической сегментации изображений будет затруднено.

### Литература

1. Заде Л.А. Понятие лингвистической переменной и его применение к принятию приближенных решений – М.: Мир, 1976.-165 с.
2. Оссовский С. “Нейронные сети для обработки информации” М.: Финансы и статистика 2004.
3. Нейрокомпьютеры в системах обработки изображений. Под ред. Ю.В. Гуляева и А.И. Галушкина. М.: Радиотехника 2003.
4. Atkinson P.M., Tatnall A.R.L. Neural Networks in Remote Sensing, Int. J. of Remote Sensing. 1997, vol. 18, no 4, pp. 699-709
5. А.Н. Пылькин, Р.В. Тишкин Методы и алгоритмы сегментации изображений. М.: Горячая линия- Телеком, 2010 – 92с.
6. Хайкин С., “Нейронные сети”, Полный курс, второе издание, Изд. дом “Вильямс”, 1104 стр.
7. Кузнецов А.Е., Кочергин А.М. Обработка изображений в задаче каталогизации данных дистанционного зондирования. Известия вузов: Геодезия и аэрофотосъемка. 2002г. №1. с. 166-172.
8. Штовба С.Д. Введение в теорию нечётких множеств и нечёткую логику. <http://matlab.exponenta.ru/fuzzylogic/book1/>
9. Форсайт Д. А., Понс Ж. Компьютерное зрение. Современный подход, М.: Издательский дом “Вильямс”, 2004г.
10. Xei X.L., Beni G.A. Validity Measure for Fuzzy Clustering // IEEE Trans. on Pattern Anal. and Machine Intell. 3 (8). - 1991. - P. 841 - 846.
11. Mendel, J.M., Type-2 Fuzzy Sets and Systems and Overview // IEEE Computational Intelligence Magazine 2(1):20-29, February 2007.
12. Mendel Uncertain Rule-Based fuzzy logic systems: introduction and new direction, Prentice Hall, upper saddle river, NJ, 2001
13. Комарцова Л.Г., Максимов А.А. Нейрокомпьютеры, М.: МГТУ им. Н.Э. Баумана, 2004. – 400с.
14. Xei X.L., Beni G.A. Validity Measure for Fuzzy Clustering // IEEE Trans. on Pattern Anal. and Machine Intell. 3 (8). - 1991. - P. 841 - 846.