

УДК 004.023

ПРИМЕНЕНИЕ ЭВОЛЮЦИОННЫХ ВЫЧИСЛЕНИЙ ДЛЯ СЕГМЕНТАЦИИ ТРЕХМЕРНЫХ МЕДИЦИНСКИХ ИЗОБРАЖЕНИЙ

Дрындик Р.В., Привалов М.В.

Донецкий национальный технический университет г. Донецк
кафедра автоматизированных систем управления
E-mail: rman-dryndik@rambler.ru

Аннотация

Дрындик Р.В., Привалов М.В. Применение эволюционных вычислений для сегментации трехмерных медицинских изображений. Проанализировано два эволюционных алгоритма сегментации трехмерных изображений. Рассмотрена адаптация алгоритма кластеризации K-средних под эволюционную парадигму и алгоритм муравьиных колоний. Показаны особенности реализации эволюционного алгоритма кластеризации.

Общая постановка проблемы. Задача анализа изображений, видео и других данных является очень актуальной в настоящее время, особенно в медицине. Зачастую, постановка правильного диагноза болезни основывается на информации, полученной посредством магнитно-резонансного исследования, позитронно-эмиссионного исследования, рентгена и прочих. Под сегментацией изображения понимают разбиение изображения на непохожие области по некоторому признаку. Все многообразие алгоритмов сегментации трехмерных изображений можно разделить на 3 основные группы [1].

- 1) структурные алгоритмы;
- 2) стохастические алгоритмы;
- 3) комбинированные алгоритмы.

Данные алгоритмы относятся к алгоритмам поиска оптимальных решений. Следует отметить, что хотя данные методы в той или иной мере решают поставленную перед собой задачу, они не лишены недостатков. К основным недостаткам можно отнести:

- 1) высокую вычислительную сложность при большом объеме обрабатываемых данных;
- 2) проигрыш в использовании ресурсов памяти;
- 3) чувствительность некоторых алгоритмов к шумам;
- 4) необходимость в задании дополнительной информации;
- 5) необходимость в настройке некоторых параметров;
- 6) плохие результаты работы с изображениями, на которых яркостные характеристики области различаются незначительно.

Решение задачи и результаты исследований. Один из вариантов преодоления части из изложенных проблем – использования алгоритмов, дающих субоптимальные решения, а именно – эволюционных алгоритмов. Рассмотрим возможность применения генетических и муравьиных алгоритмов для сегментации трехмерных изображений.

Под сегментом подразумевается некоторая область, такая, что расстояние между ее элементами минимально, а расстояние между двумя соседними областями – максимально. Такой факт дает возможность построить генетический алгоритм, в основе которого может использоваться известный метод кластеризации. В рамках данной статьи рассмотрена адаптация алгоритма K-средних для применения в генетическом алгоритме сегментации.

Требуется разбить множество, состоящее из n элементов, на K групп так, чтобы расстояние между элементами внутри группы было минимальным, а расстояние между

группами - максимальным. Каждый элемент рассматриваемого множества представляет собой вектор длины d .

Пусть дано множество $A = \{x_1, x_2, \dots, x_{n-i+1}, \dots\}$ - множество векторов длины d и пусть элемент x_{ij} - это j -е свойство i -го элемента множества A . Введем следующую величину:

$$\mathbf{W} = \begin{pmatrix} w_{11} & \dots & w_{1K} \\ \vdots & \ddots & \vdots \\ w_{N1} & \dots & w_{NK} \end{pmatrix} = \|w_{ik}\| = \begin{cases} 1, x_i \in G_k \\ 0, x_i \notin G_k \end{cases}. \quad (1)$$

Матрица \mathbf{W} обладает следующим важным свойством:

$$w_{ij} \in \{0,1\} \text{ и } \sum_{1 \leq j \leq K} w_{ij} = 1 \quad (2)$$

Обозначим центры кластеров следующим образом $c_k = (c_{k_1}, c_{k_2}, c_{k_3}, \dots, c_{k_d})$, тогда

$$c_{k_j} = \frac{\sum_{1 \leq i \leq n} w_{ik} x_{ij}}{\sum_{1 \leq i \leq n} w_{ik}}. \quad (3)$$

Тогда внутрикластерное отклонение может быть определено по формуле:

$$S^k(W) = \sum_{1 \leq i \leq n} w_{ik} \sum_{1 \leq j \leq d} (x_{ij} - c_{kj})^2 \quad (4)$$

Требуется найти такую матрицу $W^* = \|w_{ik}^*\|$, что $S(W^*) = \min_W \{S(W)\}$.

Реализация генетического алгоритма. Так как пространство поиска представляет собой множество всех существующих двоичных матриц W , удовлетворяющих условию (2), то можно представить хромосому в виде строки длины n , при этом каждый аллель может принимать значения из множества $\{1, 2, \dots, K\}$. Инициализация популяции может быть осуществлена одним из классических способов.

Селекция предполагает выбор случайной хромосомы из популяции с помощью метода рулетки, либо какого-либо другого метода. Решение в текущей популяции может задавать некоторый вес, который влияет на выживание в следующем поколении. Это означает, что каждой особи задается некоторое число, либо вычисляется значение специальной фитнес-функции. Существует много способов вычисления этой фитнес-функции [3]. Авторы [2] предлагают следующий вариант решения. Пусть $f(s_W) = -S(W)$, $g(s_W) = f(s_W) - (\bar{f} - c \cdot \sigma)$, где \bar{f} , c - среднее значение и стандартное отклонение $f(s_W)$ соответственно, $c \in [1..3]$ - некоторая константа. Тогда значение фитнес-функции $F(s_W)$ определяется следующим образом:

$$F(s_W) = \begin{cases} g(s_W), g(s_W) \geq 0 \\ 0, g(s_W) < 0 \end{cases}. \quad (5)$$

Оператор мутации меняет значение аллеля в зависимости от расстояния от данной точки до центра кластера. Зададим его следующим образом: аллель заменяется значениями, которые выбираются случайно, согласно следующему распределению вероятностей:

$$p_j = \Pr\{s_W(i) = j\} = \frac{c_m d_{\max} - d_j}{\sum_{1 \leq i \leq K} (c_m d_{\max} - d_i)}, \quad (6)$$

где c_m - некоторая константа, $d_{\max} = \max_j \{d_j\}$.

Алгоритм с представленным выше оператором селекции может сходиться достаточно долго. Кроме того, вероятность мутации должна быть достаточно мала, так как большие

значения ведут к колебаниям решений. Чтобы исключить такие ситуации вводится новый оператор, называемый K-means operator (КМО) [2]. Выполнение данного оператора состоит из следующих двух шагов:

1) вычисляются центры кластеров по формуле (3);

2) каждая точка данных переназначается кластеру с ближайшим центром, тем самым формируя новую матрицу \tilde{W} .

Имея технику построения начальной популяции и генетические операторы, можно построить генетический алгоритм, блок-схема которого представлена на рис. 1.

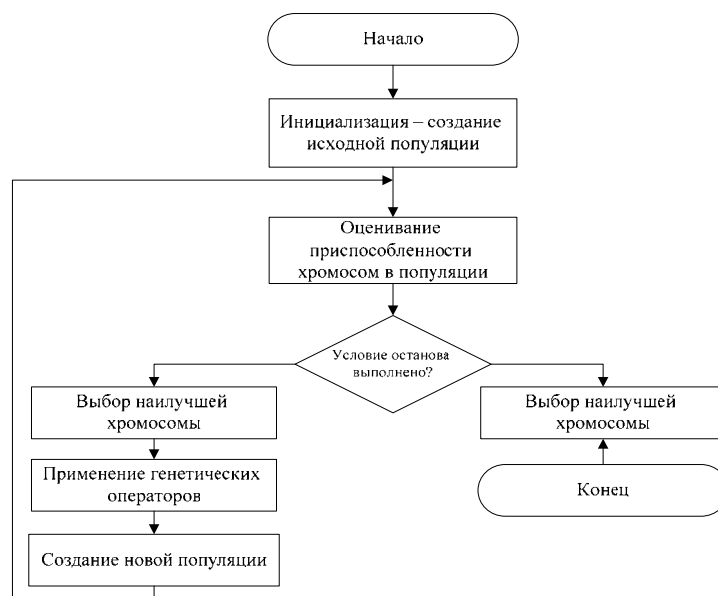


Рисунок 1 - Общая схема работы генетического алгоритма

Муравьиные алгоритмы. Задача сегментации изображений также может быть сведена к задаче оптимизации на графах [1]. Такая специфика задачи позволяет использовать алгоритм муравьиных колоний для ее решения.

Идея муравьиного алгоритма состоит в следующем. Происходит моделирование поведения муравьёв, связанного с их способностью быстро находить кратчайший путь от муравейника к источнику пищи и адаптироваться к изменяющимся условиям, находя новый кратчайший путь. При своём движении муравей оставляет на своем пути феромон, и эта информация используется другими муравьями для выбора пути. Это элементарное правило поведения и определяет способность муравьёв находить новый путь, если старый оказывается недоступным.

Происходит так же моделирование обогащения пути феромоном и его испарения. Данные шаги алгоритма позволяют искать все возможные оптимальные пути в пространстве поиска, а не какой-то один конкретный путь.

В независимости от модификации общая идея работы муравьиного алгоритма может быть продемонстрирована блок-схемой, приведенной на рис. 2.

Инициализация муравьёв - стартовая точка, куда помещается муравей, зависит от ограничений, накладываемых условиями задачи, потому что для каждой задачи способ размещения муравьёв является определяющим. Либо все они помещаются в одну точку, либо в разные с повторениями, либо без повторений. На этом же этапе задаётся начальный

уровень феромона. Он инициализируется небольшим положительным числом для того, чтобы на начальном шаге вероятности перехода в следующую вершину не были нулевыми.

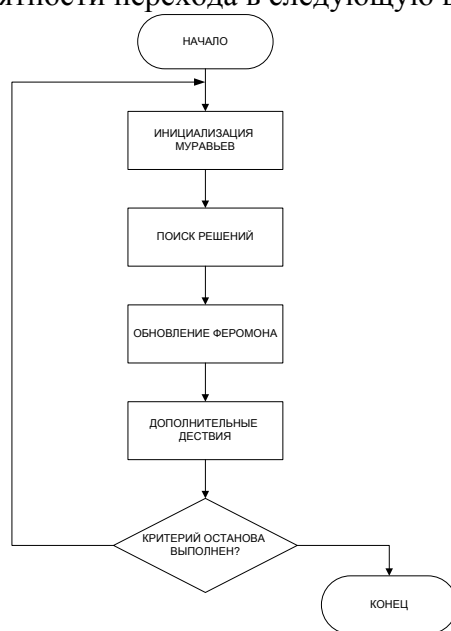


Рисунок 2 - Общая схема работы муравьиного алгоритма

Вероятность перехода из вершины i в вершину j определяется по следующей формуле:

$$P_{ij}(t) = \frac{\tau_{ij}(t)^\alpha \left(\frac{1}{d_{ij}}\right)^\beta}{\sum_{j \in N_i^k} \tau_{ij}(t)^\alpha \left(\frac{1}{d_{ij}}\right)^\beta}, \quad (7)$$

где N_i^k представляет множество возможных вершин, связанных с вершиной i , для муравья k , τ_{ij} - уровень феромона, d_{ij} - эвристическое расстояние, α и β - константные параметры.

Обновление феромона происходит следующим образом:

$$\tau_{ij}(t+1) = (1 - \rho)\tau_{ij}(t) + \sum_{k \in n_k} \frac{Q}{L_k(t)}, \quad (8)$$

где n_k - количество муравьев, ρ - интенсивность испарения, $L_k(t)$ - длина пути, построенного k -м муравьем в момент времени t , а Q - параметр значимости, т. е. $\frac{Q}{L_k(t)}$ - феромон, откладываемый k -м муравьем, использующим ребро (i, j) .

Рассмотрим способ применения алгоритма муравьиных колоний к сегментации изображений. В этом алгоритме муравьи представляются простыми агентами, которые случайным образом перемещаются по дискретному массиву. Пиксели, которые рассеяны в элементах массива могут быть перемещены из одного элемента массива в другой, формируя кластеры. Перемещение пикселей косвенным образом определяется поведением муравьев. Каждый муравей может присоединять или исключать пиксель согласно функции, которая вычисляет сходство пикселя с другими пикселями в кластере. Таким образом, муравьи динамически кластеризируют изображение на независимые кластеры, которые включают только сходные между собой пиксели. Кроме этого агент представляет новые вероятностные

правила для присоединения или исключения пикселей, а так же стратегию локального перемещения для ускорения сходимости алгоритма.

Предположим, размерность массива равна N . Каждый элемент этого массива связан с гнездом муравьев в определенном порядке, что позволяет муравьям легко переходить от одного элемента к другому. В процессе работы алгоритма муравьи могут изменять, строить или убирать существующие кластеры пикселей. Кластер представляется соединением двух или более пикселей и пространственно расположен в отдельной ячейке массива, что упрощает его идентификацию.

Инициализируем N пикселей $\{P_1, \dots, P_n\}$ для кластеризации, расположенные в массиве таким образом, что каждый элемент массива связан только с одним пикселем. Каждый муравей a_i из колонии с K муравьями $\{a_1, \dots, a_k\}$ присоединяет случайно выбранный пиксель из элементов массива и возвращается в гнездо.

После начального этапа инициализации происходит этап кластеризации. Выбор муравья производится случайным образом. Он перемещается от своего гнезда к элементам массива и определяет с помощью вероятностного правила, присоединять ли этот пиксель к кластеру. Каждый муравей знает список расположения пикселей, которые не были присоединены другими муравьями. Случайным образом муравей определяет следующий пиксель из списка свободных. Этот алгоритм повторяется для каждого муравья. Останов происходит при выполнении заданного критерия останова.

Основная преследуемая цель выполняемой работы состоит в том, чтобы построить хорошо распараллеливаемый эволюционный алгоритм, который способен решить свою поставленную задачу за приемлемое время, обрабатывая при этом достаточно большие массивы данных, а так же комбинируя при необходимости ресурсы CPU и GPU. Легко заметить, что наличие множества простых, практически независимых друг от друга агентов дает хорошие потенциальные возможности для распараллеливания алгоритма.

Вывод

На основании вышеизложенного можно сделать вывод, что применение эволюционных и роевых вычислений может помочь устранить некоторые недостатки известных методов при обработки трехмерных изображений. Показано, что одним из подходов может быть применение метода кластеризации в совокупности с генетическим алгоритмом. В качестве альтернативы предлагается использовать алгоритм муравьиных колоний для решения задачи кластеризации, сведенной к задаче оптимизации на графах.

Анализ предлагаемых решений показал, что генетические алгоритмы и алгоритмы муравьиных колоний основаны на ряде независимых подзадач, что позволяет значительно повысить скорость их работы, выполнив распараллеливание с применением вычислений на CPU, а также GPGPU. Таким образом, направлением дальнейшей работы является разработка и реализация параллельных вычислительных схем для эволюционного варианта алгоритма K-средних и муравьиного алгоритма кластеризации изображений, а также их проверка на задачах кластеризации и классификации КТ и МРТ изображений.

Список литературы

1. 3D Segmentation Techniques for Medical Volumes Sarang Lakare Center for Visual Computing Department of Computer Science State University of New York at Stony Brook Research Proficiency Exam Advisor: Prof. Arie Kaufman
2. Genetic K-Means Algorithm K. Krishna and M. Narasimha Murty
3. D. Goldberg, Genetic Algorithms in Search, Optimization and Machine Learning. Reading, MA: Addison-Wesley, 1989.
4. C. von der Malsburg. The correlation theory of brain function. Technical report, Max-Planck-Institute Biophysical Chemistry, 1981.