

УДК 004.48+004.932

**ЭВОЛЮЦИОННЫЙ АЛГОРИТМ ПОИСКА ЗАВИСИМОСТИ МЕЖДУ ТЕКСТУРНЫМИ ПАРАМЕТРАМИ УЗИ И КАЧЕСТВОМ ОБРАБОТКИ УЗИ РАЗЛИЧНЫМИ МЕТОДАМИ****Т. А. Махно**Институт прикладной математики и механики НАН Украины  
ул. Розы Люксембург, 74, Донецк, 83114, Украина. E-mail: belikova.taisija@gmail.com

Одной из задач систем, призванных облегчать оценку степени атеросклеротического поражения артерий на основании УЗИ, является разделение плотных тканей и просвета артерий на УЗ снимке. В большинстве работ эта задача решается либо при помощи алгоритмов бинаризации изображения, либо подобную работу выполняет эксперт вручную. В предыдущих работах был предложен генетический алгоритм, позволяющий автоматически построить последовательность обработки изображений. Алгоритм находил различные решения при заранее заданном результате обработки. Однако эти решения не были одинаково хороши для любого изображения. Для выбора метода обработки необходимо предварительно рассчитать некоторые параметры изображения. Эти параметры должны находиться в зависимости с последовательностью обработки изображения. В данной работе предложен генетический алгоритм поиска соответствий между значениями текстурных параметров изображений и подходящими для них методами обработки. Алгоритм был протестирован на 375 реальных УЗ изображениях сонных артерий человека.

**Ключевые слова:** обработка изображений, генетический алгоритм, матрицы смежности яркостей, УЗИ.

**ЕВОЛЮЦІЙНИЙ АЛГОРИТМ ПОШУКУ ЗАЛЕЖНОСТІ МІЖ ТЕКСТУРНИМИ ПАРАМЕТРАМИ УЗЗ І ЯКІСТЮ ОБРОБКИ УЗЗ РІЗНИМИ МЕТОДАМИ****Т. О. Махно**Институт прикладної математики та механіки НАН України  
вул. Розі Люксембург, 74, Донецьк, 83114, Україна. E-mail: belikova.taisija@gmail.com

Однією з задач систем, покликаних полегшувати оцінку ступеня атеросклеротичного ураження артерій на підставі УЗД, є поділ щільних тканин і просвіту артерій на УЗ знімку. У більшості робіт ця задача вирішується або за допомогою алгоритмів бінаризації зображень, або таку роботу виконує експерт вручну. У попередніх роботах ми пропонували генетичний алгоритм, що дозволяє автоматично побудувати послідовність обробки зображень. Алгоритм знаходив різні рішення при заздалегідь заданому результаті обробки. Однак ці рішення не були однаково добрими для будь-якого зображення. Для того щоб обрати метод обробки необхідно попередньо розрахувати деякі параметри зображення. Ці параметри повинні знаходитися в залежності з послідовністю алгоритмів обробки зображення. У даній роботі запропонований генетичний алгоритм пошуку відповідностей між значеннями текстурних параметрів зображень і придатними для них методами обробки. Алгоритм був протестований на 375 реальних УЗ зображеннях сонних артерій людини.

**Ключові слова:** обробка зображень, генетичний алгоритм, матриці суміжності яскравості, УЗЗ.

**АКТУАЛЬНОСТЬ РАБОТЫ.** Задачи разработки систем автоматизации процесса обработки цифровых изображений остаются актуальными и активно исследуемыми во всём мире уже несколько десятков лет. К настоящему моменту существует множество программных продуктов и приложений, позволяющих облегчить работу и/или сократить участие человека в процессах обработки и анализа изображений. Разнообразие используемых при создании вышеупомянутых программных продуктов подходов объясняется сложностью формализации и обобщения исходных данных, а также требуемых результатов обработки.

При решении задач формализации изображений некоторого типа в настоящее время существует целый ряд подходов, некоторые из них [1–3] находят применение и в задачах обработки УЗ изображений.

При выборе того или иного метода обработки изображения, специалист зачастую пользуется существующими общепризнанными рекомендациями [4], иногда дополняемыми собственным опытом и наблюдениями [5, 6]. Также существуют работы, в которых для решения некоторых прикладных задач применяют

алгоритмы автоматической генерации способа обработки изображения [7]. Подобные алгоритмы могли бы позволить создавать множества различных комбинаций методов обработки изображений и затем применять их в зависимости от автоматически рассчитанных характеристик обрабатываемых изображений. В предыдущих работах [8, 9] были предложены алгоритмы поиска подходящих способов обработки изображений при определённом заранее результате обработки. При наличии некоторой репрезентативной выборки изображений заданного типа, алгоритмы [8, 9] позволяют составить наиболее эффективные схемы (последовательности) их обработки. В данном случае, последовательности (схемы) обработки изображений называются эффективными, если они удовлетворяют некоторому желаемому результату. Желаемым результатом применения такой схемы, в текущей задаче сегментации УЗ изображений артерий человека, было заранее подготовленное экспертом изображение-шаблон, задающее области и границы плотных тканей, а также области просветов артерий и полостей тканей.

Таким образом, для некоторых наперед заданных изображений по заранее подготовленным шаблонам при помощи алгоритма [8, 9] могут быть найдены схемы фильтрации изображений, дающие качественный результат обработки. Следовательно, на данном этапе, задачу выбора последовательности фильтрации для произвольного УЗ изображения, при отсутствии шаблона-результата, вышеупомянутый алгоритм решить не может. Результатом работы данного алгоритма является ряд решений, дающих достаточно близкий к желаемому результат. При необходимости автоматического выбора конкретной схемы обработки из множества ранее построенных, возникает задача классификации или предварительной оценки обрабатываемого изображения. В работах [7, 10] для быстрой классификации произвольного изображения применяют алгоритм SIFT [2, 3], затем выполняют кластеризацию в «ручном» режиме и текстурный анализ [1]. В работах С. Loizou [5, 6] акцент делается на последовательность, состоящую из порогового фильтра Отсу и группы фильтров, объединённых С. Loizou под названием *lsmv*. В некоторых случаях автор использовал морфологический фильтр дилатации, с целью устранения «узких впадин» значений яркости на изображении. Однако подбор последовательностей производился вручную, на основании предложенных автором параметров «качества» изображения и экспертных оценок.

Цель работы – поиск зависимости между текстурными параметрами УЗИ и качеством УЗИ различными методами.

#### МАТЕРИАЛ И РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЙ.

Была поставлена задача выбора параметров изображения или его фрагмента на основании которых можно принимать решение о выборе того или иного метода или последовательности методов обработки. При выборе текстурных параметров сложно предположить какова будет зависимость методов фильтрации (обработки) от значений той или иной текстурной характеристики. Также не определён выбор из самого множества параметров [1] и их комбинаций. Кроме того, при построении двумерных графиков зависимости результатов фильтрации изображений от их параметров, визуально не определяются ни кластеры, ни функциональные зависимости. Следовательно, при параметризации изображений с целью быстрого выбора методов его фильтрации, необходимо свыше двух параметров и на момент реализации и начала тестирования алгоритма невозможно определить, какие это параметры и какова зависимость между ними.

Для решения задач поиска в пространстве сложно формализуемых данных широко используются эволюционные алгоритмы. Эволюционный алгоритм, подходящий для решения данной задачи, должен находить некоторые зависимости, позволяющие по значению параметров однозначно определить последовательность (одну или несколько), которая потребуются для успешного решения задачи сегментации. Таким

образом, алгоритм должен находить решение в виде пар множеств. Первое множество – это набор диапазонов параметров, определяющих класс изображений. Второе множество – это набор последовательностей фильтрации (из одного и более элементов), позволяющих получить качественный результат обработки изображений данного класса.

Решением или особью популяции генетического алгоритма (ГА), в таком случае, будет набор из двух не гомологичных хромосом. Первая хромосома – это упорядоченная последовательность генов, длина которой зависит от количества параметров изображений. Каждый ген первой хромосомы представляет собой некоторый числовой диапазон, заданный парой вещественных чисел, принимающих значения от  $-1$  до  $1$ . Начальная популяция генерируется случайным образом, при этом большинство генов (диапазонов) хромосомы первого вида остаются парой  $(-1;1)$ , и только некоторая часть произвольно выбранных генов приобретает случайно выбранные значения отличные от  $-1$  и  $1$ , т.е. в хромосоме первого вида решения может использоваться любое число параметров из подготовленного вектора. Вторая хромосома представляет собой подмножество множество схем фильтрации, ранее построенных при помощи алгоритма [8, 9]. При генерации начальной популяции подмножество хромосомы второго вида составляется случайным образом из различных последовательностей фильтрации, для увеличения скорости вычислений мощность такого множества устанавливается близкой к  $1$ .

Имеющиеся в распоряжении начальные данные, составляющие пространство поиска, были получены по формулам, приведенным в [1], а также в результате многочисленных экспериментов по автоматическому поиску эффективных методов фильтрации изображений [8, 9]. В данной реализации они представляют собой множество наборов значений текстурных параметров и схем обработки, отношение между которыми выражается некоторым вещественным числом, принимающим значение от  $0$  до  $1$ . Это значение выражает степень приемлемости (эффективности) применения данных схем обработки к изображениям с текущими значениями текстурных параметров. При построении векторов параметров, задающего хромосомы первого вида, для каждого заданного изображения параметры [1] вычисляются на основании различных матриц смежности яркостей. Определение эффективности каждой последовательности обработки производится путём вычисления близости результата фильтрации данного изображения и соответствующего ему шаблона. Исходное пространство поиска генетического алгоритма (ГА) сформировано следующим образом: номер изображения; номер схемы обработки; показатель качества фильтрации текущего изображения данной схемой; значения текстурных параметров текущего изображения.

Не углубляясь в подробности реализации, процесс работы генетического алгоритма можно описать как поиск соответствий различных цепочек фильтров из

заранее выделенного множества и соответствующих им диапазонов параметров, описывающих изображения, поддающиеся качественной фильтрации данной последовательностью фильтров.

Упрощенная схема используемого генетического алгоритма приведена ниже:

- 1) случайным образом сгенерировать популяцию G размера M;
- 2) выполнить отбор родительских особей (оператор селекции);
- 3) применить оператор кроссинговера к родительским особям;
- 4) применить оператор редукции (сокращение размера популяции до исходного);
- 5) применить оператор мутации;
- 6) перейти к шагу 1, если количество поколений меньше N, иначе выход.

Распределённый генетический алгоритм (РГА) представляет собой несколько параллельно работающих генетических алгоритмов, каждый из которых аналогичен описанному выше. В зависимости от настроек алгоритма, обмен особями параллельно развивающихся популяций происходит синхронно, с удалением из исходной популяции мигрирующей особи. Миграция однонаправленная, по кольцу. Мигрирующая особь выбирается случайным образом, не зависимо от значения фитнес функции.

При построении новой «гипотезы», алгоритм должен ставить в соответствие некоторому множеству диапазонов параметров определённые последовательности фильтров. Затем «гипотеза» проверяется на всём множестве данных и считается её фитнес.

В работе проверяется влияние следующих принципов построения гипотез на свойства построенных алгоритмом решений:

- 1) случайный выбор диапазонов параметров и случайный выбор последовательностей фильтров;
- 2) случайный выбор диапазонов параметров из имеющихся в начальных данных и случайный выбор последовательностей фильтров.

Кроссинговер хромосом первого вида может осуществляться:

- 1) по принципу дискретной рекомбинации (многоточечного кроссинговера);
- 2) аналогично правилам промежуточной и линейной рекомбинации для нижней и верхней грани диапазона, в зависимости от начальных настроек алгоритма;
- 3) комбинированием двух вышеописанных методов.

В случае применения третьего вида кроссинговера, в процессе скрещивания выбираются гены потомков из родительских особей по маске, аналогично дискретной рекомбинации. Отличие от первого вида кроссинговера состоит в том, что только часть генов копируются из одной родительской особи. Значения остальных генов вычисляются аналогично правилам промежуточной рекомбинации, а не копируются из второй родительской особи, как в случае дискретной

рекомбинации. Третий вид кроссинговера был предложен с целью уменьшения изменчивости особей при скрещивании, сравнительно со вторым видом кроссинговера. К тому же при этом сохраняется возможность изменения при помощи кроссинговера диапазонов параметров, сгенерированных в начальной популяции.

Кроссинговер хромосом второго вида происходит с заданной вероятностью при наличии несовпадающих элементов у родительских особей, в противном случае не производится.

Мутация хромосом первого вида может происходить следующим образом:

- 1) замена одной из границ диапазона на случайное допустимое вещественное число;
- 2) изменение верхней и нижней границы диапазона на случайные допустимые числа;
- 3) изменение границ диапазонов на случайно выбранные из множества начальных данных допустимые значения текущего параметра.

Мутация хромосом второго вида происходит путём замены любой случайно выбранной схемы фильтрации, входящей в состав мутирующей хромосомы, на любую другую произвольно выбранную схему из множества начальных данных.

В алгоритме предложено несколько вариантов расчёта фитнес функции:

- 1) на основании среднего значения качества фильтрации всеми схемами всех изображений класса, описываемого параметрами особи;
- 2) по наихудшему значению качества фильтрации, полученному по всем схемам всех изображений класса, описываемого параметрами особи;
- 3) аналогично второму варианту, но с поправкой на количество изображений класса, описываемого параметрами особи.

Отбор родительских особей осуществляется по принципу турнирного отбора. Сокращение промежуточной популяции осуществляется методом селекции.

При поиске оптимальной модификации генетического алгоритма была выполнена проверка работы различных вышеописанных типов операторов ГА и РГА и вариантов их настроек.

При сравнении эффективности работы различных модификаций генетического алгоритма, на начальных этапах проводилось не менее 100 запусков каждой модификации и результаты, приведенные в описании и на графиках, являются усреднёнными.

Изменение мутации в пределах 0,05–0,15 в среднем влияет на значение целевой функции в пределах 0,5 %. Такое изменение значения целевой функции соответствует ошибке эксперимента. Таким образом, считаем, что большая часть модификаций данного алгоритма нечувствительна к изменениям вероятности мутации в вышеуказанных пределах.

На рис. 1 отображено влияние размера турнирной группы при выборе родителей. На графике отображены средние значения фитнес функции по 100 запускам ГА, а также максимальное и минимальное

значение лучшего решения, среди полученных. Первые три эксперимента проводились при вероятности мутации 0,01 и вероятности кроссинговера 0,5, а также для третьего вида оператора мутации и второго вида оператора кроссинговера, описанных ранее. Изменение размера турнирной группы: 2, 4, 6 соответственно. Для 3, 4 и 5 эксперимента размер турнирной группы одинаков. Отличие четвертого эксперимента в вероятности мутации 0,1 и вероятности кроссинговера 0,7, а также в 4-9 экспериментах используется первый тип фитнес функции. В 6-м эксперименте численность турнирной группы 8. В 7-м и 8-м численность группы 2, в 9-м – 6. В 8-м и 9-м экспериментах использовался РГА. Таким образом, при определённых значениях вероятности мутации и кроссинговера численность турнирной группы не существенно влияет на качество работы алгоритма. При тестировании РГА размер турнирной группы не влиял на качество работы алгоритма.

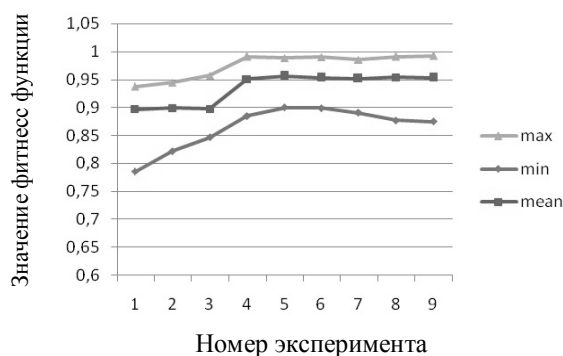


Рисунок 1 – Влияние размера турнира на качество работы алгоритма

Частота обмена особями между под популяциями РГА считается оптимальной в пределах вероятности обмена 0,2–0,4 для каждого поколения. Уменьшение этого значения ведёт к излишним затратам вычислительных ресурсов, а увеличение – к сходимости различных под популяций к одним и тем же решениям.

Влияние численности популяции и количества поколений на качество работы алгоритма отображено на рис. 2. Численность популяции для 1-5 эксперимента составляла: 10, 10, 15, 30, 50 особей соответственно. Количество поколений: 10, 20, 15, 20, 50. Остальные параметры были одинаковы, тестирование проходило на модификации РГА.

На рис. 3 приведено сравнение качества работы простого ГА и РГА для одинаковых типов операторов и их настроек. Пунктиром на рисунке отображаются результаты тестирования простого ГА и сплошной линией отображаются результаты тестирования распределённого ГА. Не останавливаясь подробно на настройках алгоритмов при каждом эксперименте можем сделать вывод о большей устойчивости РГА по сравнению с простым ГА.

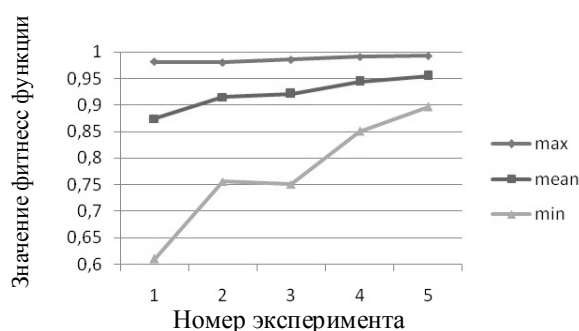


Рисунок 2 – Влияние численности популяции и количества поколений на качество работы алгоритма

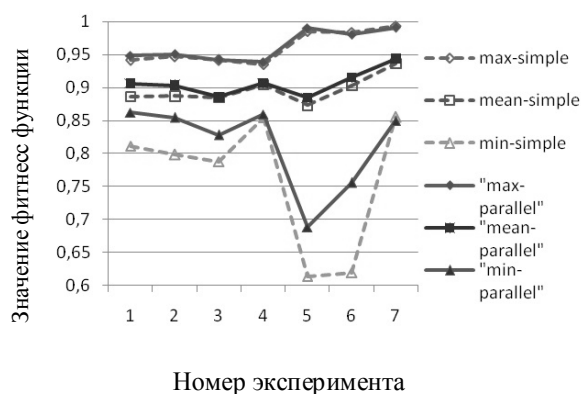


Рисунок 3 – Сравнение качества работы простого ГА и РГА

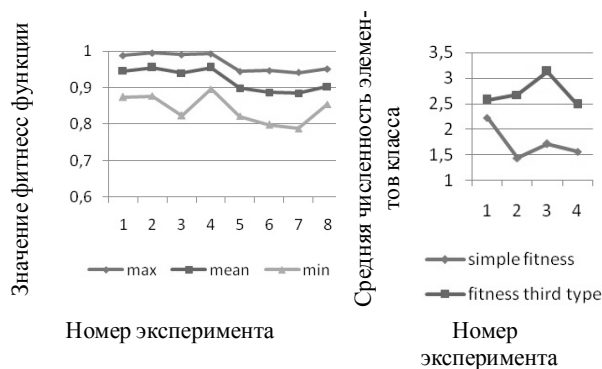


Рисунок 4 – График качества работы ГА в зависимости от типа фитнес функции. График среднего количества наблюдений, описываемых найденными решениями

На рис. 4 приведены график качества работы ГА в зависимости от типа фитнес функции, а также график, отображающий среднее количество наблюдений, описываемое найденными решениями. Эксперименты попарно проводились для одинаковых настроек ГА (1 и 5, 2 и 6, 3 и 7, 4 и 8), за исключением типа фитнес функции. Так эксперименты 1–4 осуществлялись для фитнес функции первого типа, а эксперименты 5–8 для фитнес функции третьего типа. Не трудно заметить, что для более простого вида фитнес функции

алгоритм находит лучшие решения, однако при этом численность наблюдений, на основании которых строится решение, значительно меньшая, нежели в случае фитнес функции третьего типа.

На рис. 5 приведены примеры экспериментов изменения средних значений фитнес функции в зависимости от типа кроссинговера. Как показано на примере содержащем 8 экспериментов, для различных модификаций алгоритмов может быть более эффективен обычный двухточечный кроссинговер либо кроссинговер третьего типа. Первые 2 значения приведены для простого генетического алгоритма, а вторые 2 для РГА. В первом и третьем экспериментах процент изменяемых при генерации начального поколения параметров 10 % от всего возможного, а во втором и четвертом это значение было 5 %.

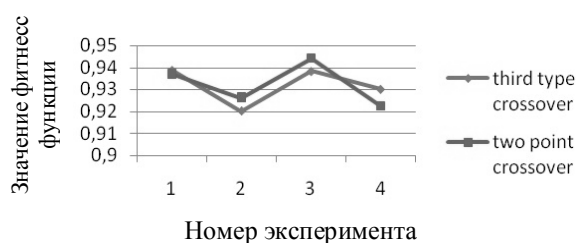


Рисунок 5 – Пример тестирования алгоритмов при различных видах кроссинговера

В данном описании приведены лишь некоторые графики, полученные по экспериментальным данным. Описание экспериментов приведено для общего представления об особенностях работы алгоритма. Целью тестирования, кроме поиска оптимальной модификации алгоритма, был также поиск модификации, позволяющей получить хороший результат проверки полученных решений на другом тестовом пространстве. По причине того, что добавлять проверку решения на тестовом множестве в расчёт фитнес функции нельзя, т.к. множество в таком случае не будет тестовым, результат проверки некоторых решений может быть не удовлетворителен (совпадение с желаемым результатом менее 80 %).

Для оценки решений, найденных на текущем тренировочном пространстве, по окончании каждого запуска ГА и РГА сохраняются лучшие особи последней популяции (каждой под популяций). По окончании серии запусков, производится проверка полученных решений на тестовом множестве изображений. Каждое решение представляет собой вектор диапазонов значений некоторых статистических параметров изображения и поставленная в соответствие этому вектору цепочка фильтров (одна или несколько). Для тестирования полученного множества решений, для тестового множества изображений рассчитываются такие же статистические характеристики, которые были ранее рассчитаны для тренировочного множества. Тестовое множество изображений подвергается фильтрации теми же последовательностями

фильтров, которыми обрабатывались изображения базового множества, для каждого изображений и каждой цепочки фильтров рассчитывается значение параметра «качества фильтрации». При проверке, для каждого решения (каждой особи) выбираются изображения, удовлетворяющие ему по значениям статистических параметров. Затем проверяются значения параметра «качества фильтрации» фильтрационными последовательностями, задаваемыми особью, изображений, удовлетворяющих диапазонам параметров. По множествам значений параметра «качество фильтрации» на тестовом множестве для каждой особи делается вывод о результатах применения решения в задаче автоматического выбора последовательности фильтрации.

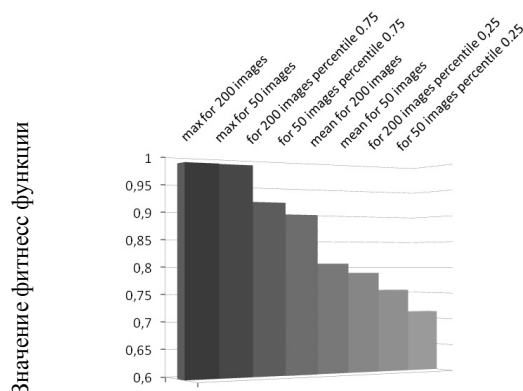


Рисунок 6 – Статистические оценки решений на тестовом множестве изображений.

Сравнение качества работы различных модификаций генетического алгоритма может производиться по нескольким критериям, позволяющим в той или иной мере описывать найденные множества фильтров и параметров. Т.к. результатом поиска являются пары множеств, мощности которых заранее не известны, оценивать их можно по численности, минимальным элементам каждого множества на тестовой выборке, среднему значению элементов на тестовом множестве, отношению минимального значения множества на тестовой выборке к минимальному значению соответствующего множества на тренировочной выборке. В данном случае под значением элемента подразумевается «качество фильтрации» изображения фильтром из некоторого решения, при условии, что параметры текущего изображения удовлетворяют этому решению.

Кроме свойств алгоритма необходимо учитывать сильную зависимость результатов поиска от пространства поиска. Количество наблюдений и репрезентативность тренировочного множества сильно влияют на качество решений на тестовом множестве. С целью определения влияния численности тренировочного множества на качество решений на тестовом множестве была проведена серия экспериментов, а именно: не менее 1000 запусков одинаковых модификаций алгоритма на тренировочных множествах раз-

личной мощности, (состоящих из 50 и 200 изображений соответственно). Найденные решения тестировались на одном и том же тестовом множестве, построенном на основании 100 изображений, отличных от используемых в тренировочных множествах. На рисунке 6 в качестве примера приведены результаты серии запусков алгоритма, состоящей из 3000 запусков простого генетического алгоритма. В данном случае уровень мутации был 0,1, кроссинговера – 0,7, численность популяции 30, число поколений – 20, третий тип фитнес функции, двухточечный кроссинговер. По приведенным примерам видно, что разность 0,75 и 0,25 квантили превышает 15 % диапазона значений случайной величины. Для сравнения, подобное значение на тренировочной выборке менее 3 %. Кроме того, для каждых 1000 запусков при одних и тех же параметрах ГА, статистические характеристики численных значений «качества» решений на тестовом множестве отличаются на 5 % диапазона значений случайной величины. Подобное значение на тренировочной выборке менее 0,25 %. На практике это означает недостаточную уверенность в качестве работы некоторого решения, найденного алгоритмом для произвольного изображения. Качество работы решений на тестовом множестве сильно зависит от состава тренировочного и тестового множества. Так в данной совокупности всех изображений присутствуют такие, попадание которых в тестовое множество резко увеличивает ошибку тестирования, ввиду отдалённости данных элементов выборки от прочих в пространстве поиска. Так, при наличии в тестовой выборке таких удалённых точек (рис. 6), увеличение тренировочного множества в четыре раза не даёт ожидаемого улучшения качества решений. На данный момент дать оценку отдалённости таких элементов выборки от прочих представляется возможным только на основании значения ошибки тестирования. Для устранения столь сильной зависимости необходимо расширить множество изображений обучающей выборки.

С целью улучшения работы предложенного метода было также изучено влияние набора параметров изображений, используемого в хромосомах первого вида, на работу ГА. Используемые в ранее описанных экспериментах параметры таковы, что коэффициент взаимной линейной корреляции некоторых из них порой достигает значения 0,9. Коррелирующие параметры на момент начала тестирования не были исключены из рассмотрения, т.к. не представляется возможным заранее определить какие из них лучше выбрать для решения поставленной задачи. Т.е. неизвестно какие параметры наиболее удачно описывают свойства изображения, которые значительно влияют на выбор методов его фильтрации. В процессе тестирования алгоритма и анализе статистики использования различных текстурных параметров в удачных и неудачных решениях, не было найдено «приоритетных» параметров. Вероятность попадания того или иного параметра, из множества коррелирующих, в «хорошее» решение имеет равномерное распределе-

ние. Таким образом, был сделан вывод, что исключение пар взаимозависимых параметров из множества, используемых при работе ГА, только увеличит скорость работы алгоритма не уменьшая его эффективности.

Первоначально для анализа было выделено некоторое множество векторов:

$$\{\delta = (a, b)\} = \{(0;1), (1;1), (1;-1), (-1;1), (1;-4), (2;-3), (-2;4), (2;5), (3;4), (3;5), (3;7), (3;11), (5;13), (7;13), (7;17), (13;17)\},$$

определяющее множество матриц смежности яркостей  $\{M_\delta = M_{(a,b)}\}$ .

Для некоторого заданного монохромного изображения  $I$ , имеющего  $L$  квантованных значений яркости можно построить матрицу смежности яркостей размером  $L \times L$  следующим образом:

$$M_{(a,b)}(i, j) = \left\{ \begin{array}{l} (x, y), (x+a, y+b) : \\ I(x, y) = i, I(x+a, y+b) = j \end{array} \right\} \quad (1)$$

где  $I(x, y)$  – яркость пикселя изображения  $I$  с координатами  $(x, y)$ .

По значениям матриц смежности яркостей (Spatial Gray Level Dependence (SGLD) или co-occurrence matrix) рассчитываются следующие параметры: контраст (2), второй угловой момент матрицы смежности (3), энтропия (4), корреляция (5).

$$CON = \sum (i - j)^2 p(i, j) \quad (2)$$

$$ASM = \sum p(i, j)^2 \quad (3)$$

$$ENT = -\sum p(i, j) \log p(i, j) \quad (4)$$

$$COR = \sum [ijp(i, j) - \mu_x \mu_y] / (\sigma_x \sigma_y) \quad (5)$$

где  $\mu_x$  и  $\sigma_x$  математическое ожидание и дисперсия сумм строк матрицы  $M_\delta$ , а  $\mu_y$  и  $\sigma_y$  мат. ожидание и дисперсия сумм столбцов матрицы  $M_\delta$ , а значение  $p(i, j) = M_\delta(i, j) / \sum M_\delta(i, j)$ .

Пусть заданно монохромное изображение  $I$ , размера  $M \times N$  тогда положим:

$$I_\delta(x, y) = |I(x, y) - I(x+a, y+b)| \quad (6)$$

где  $I(x, y)$  – яркость пикселя изображения  $I$  с координатами  $(x, y)$ .

Тогда  $p_\delta(i)$  вычисляем следующим образом:

$$p_\delta(i) = \frac{|\{(x, y), (x+a, y+b) : I_\delta(x, y) = i\}|}{\sum_{\substack{x=0 \\ y=0 \\ x=M-a \\ y=N-b}} I_\delta(x, y)} \quad (7)$$

Кроме вышеописанных параметров (2-5), рассчитываются следующие текстурные параметры: контраст (8), второй угловой момент матрицы смежности (9), энтропия (10), математическое ожидание разности яркостей (11).

$$CON = \sum i^2 p_\delta(i) \quad (8)$$

$$ASM = \sum p_{\delta}(i)^2 \quad (9)$$

$$ENT = -\sum p_{\delta}(i) \log p_{\delta}(i) \quad (10)$$

$$MEAN = \frac{1}{m} \sum ip_{\delta}(i) \quad (11)$$

Таким образом, для каждого изображения, по каждой матрице смежности было рассчитано значение каждого параметра. После чего были рассчитаны коэффициенты линейной корреляции для всех пар параметров всех матриц смежности. Построены графики зависимостей, модели линейных регрессий для всех пар параметров, рассчитаны такие параметры как коэффициент детерминации, скорректированный коэффициент детерминации, t-критерий Стьюдента, р-значение.

Наиболее сильно, среди всех рассчитанных параметров, коррелировали между собой параметры (3) и (9) не зависимо от принципа выбора векторов матриц смежности. Коэффициент линейной корреляции для большинства проверенных значений превышал 0,9. Взаимная степень корреляции значений прочих параметров находится в более тесной зависимости от векторов, задающих исходные матрицы смежности. А именно от длины вектора и его направления. Так, в приведенном на рисунке 7 двумерном графике, коэффициент линейной корреляции значений параметра (2), рассчитанного по матрицам, заданным векторами (0,1) и (1,1), был равен 0,2716486. Когда коэффициент корреляции значений того же параметра, по матрицам, заданным векторами (1,-1) и (1,1), был равен 0,9328412. В тоже время коэффициент корреляции значений параметра (3) для этих векторов был равен 0,9318660 и 0,9981008 соответственно. На основании этого можно сделать вывод, что параметры (3) или (9) имеет смысл рассчитывать по матрицам смежности, заданным векторами, значительно отличающимися между собой по длине и ориентации в пространстве. При этом не произойдет потери какой либо значимой информации об изображении.

Для всех параметров, кроме параметров (3) и (9), коэффициент корреляции превышает порог 0,9 лишь для близких, с точки зрения принципа расчёта параметра, значений векторов. Так, на приведенном на рис. 8 графике, отображаются значения параметра (8), рассчитанного для 250 изображений. Значения коэффициента линейной корреляции 0,9691838 и 0,7914600 для пар векторов (2;5) и (3;7), и (2;5) и (-2;4) соответственно.

Расчёты проводились для различных матриц смежности яркостей. Кроме ранее упомянутых векторов, задающих матрицы смежности, в работе были рассмотрены также следующие: (0;2), (2;2), (0;3), (3;3), (0;5), (5;5), (0;7), (7;7), (0;11), (11;11), (0;13), (13;13), (0;17), (17;17), (2;1), (3;1), (1;4), (2;3), (2;4), (0;19), (0;23), (0;29), (0;31), (0;37), (0;41), (0;43), (0;47) и др.

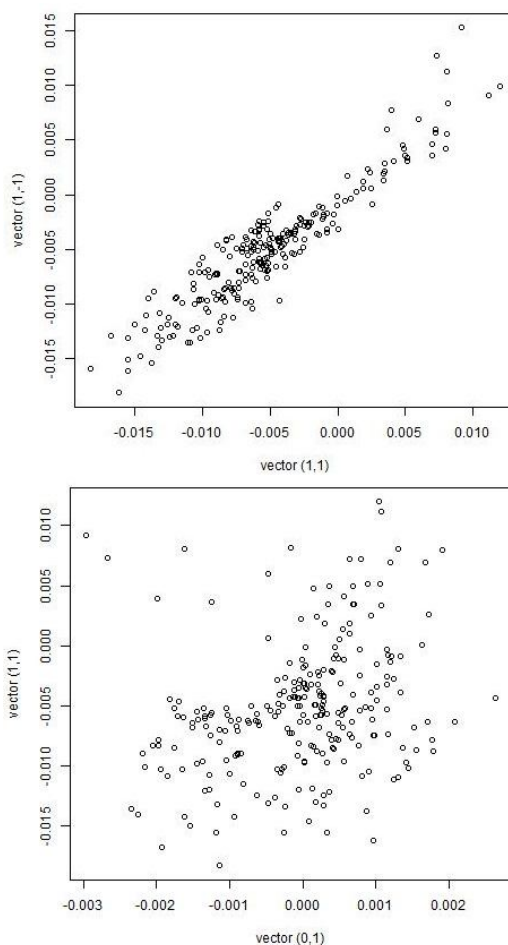


Рисунок 7 – Экспериментально полученные значения параметра «контраст» (2) для векторов (1;-1), (0;1) и (1;1)..

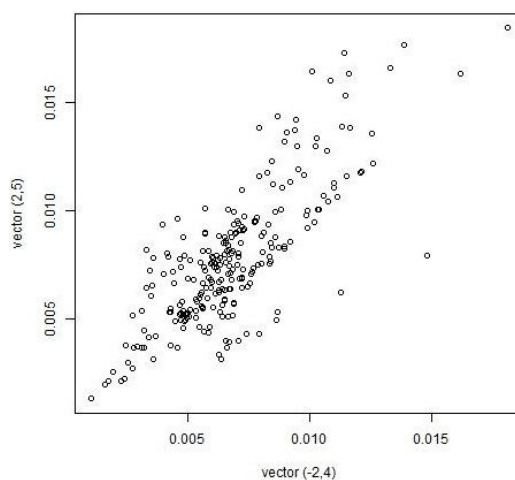


Рисунок 8 – Экспериментально полученные значения параметра «контраст» (8) для векторов (2;5) и (-2;4).

На основании полученных данных, множество используемых ГА текстурных параметров было сокращено, что повысило эффективность работы ГА. Зави-

симости значений параметров и при значении коэффициента корреляции большего 0,8 кажутся на первый взгляд достаточными, для исключения многих векторов из рассмотрения. Однако это увеличивает вероятность ошибки разделения множеств и приводит к ухудшению показателей качества работы алгоритма на таких сокращённых множествах

**ВЫВОДЫ.** В результате данной работы был предложен алгоритм поиска соответствий между эффективными методами обработки изображений и значениями текстурных параметров изображений. Алгоритм позволяет с большой степенью уверенности получить желаемый результат обработки, при условии, что параметры заданного изображения достаточно близки к параметрам какого-либо изображения из тренировочной выборки. Для произвольного тестового изображения значение вероятности получения удовлетворительного качества обработки была от 0,55 до 0,9 для различных вариантов тренировочной и тестовой выборки (совпадение с шаблоном более 80%).

Разработанный алгоритм затрачивает гораздо меньше времени и вычислительных ресурсов сравнительно с ранее предложенными в [8, 9] алгоритмами. Кроме того, не представляется возможным решить задачу выбора последовательности фильтрации для произвольного УЗ изображения при помощи алгоритмов, описанных в [8, 9]. Однако совместное использование алгоритмов [8, 9] и алгоритма, предложенного в данной работе, позволяет параметризовать произвольное УЗ изображение и выбрать последовательность его обработки. Последовательность выбирается из списка ранее найденных алгоритмом [8, 9]. Этот список может быть расширен путём поиска новых последовательностей обработки для новых УЗ изображений.

В результате работы было изучено влияние принципов выбора набора векторов, задающих матрицы смежности яркостей и рассчитанных на их основании параметров изображений, используемого в хромосомах первого вида, на работу ГА. Для построения регрессионных моделей использовалась программная среда R. С целью улучшения качества работы алгоритма из рассмотрения были исключены сильно коррелирующие параметры. Уменьшение числа параметров обеспечивает более высокую скорость работы алгоритма, сокращает время предварительной оценки нового изображения. В целях поиска более эффективных модификаций генетического алгоритма в дальнейшем планируется использовать не простые диапазоны параметров, а функциональные зависимости параметров. Это могло бы позволить строить логически аргументированные теории взаимосвязи параметров и методов обработки, не прибегая к случайному поиску решений. Кроме того, изучение текстурных характеристик УЗ изображения тканей паци-

ентов, на различных стадиях заболевания, может помочь ставить более точные диагнозы на основании УЗ исследования. Такие исследования требуют большого объёма обучающей выборки, состоящей из самих УЗ исследований и диагноза, подтверждённого после хирургического вмешательства.

*Работа проводится в сотрудничестве с Институтом неотложной и восстановительной хирургии им. В.К. Гусака АМН Украины.*

#### ЛИТЕРАТУРА

1. Joan S. Weszka, Charles R. Dyer, Azriel Rosenfeld. A Comparative Study of Texture Measures for Terrain Classification // IEEE Transactions On Systems, Man, And Cybernetics, Vol. Smc-6, No. 4, April 1976.
2. Lowe D. Distinctive image features from scale-invariant keypoints // International Journal of Computer Vision. – 2004. – Vol. 60(2). – PP. 91–110.
3. Lowe, D.G. Object recognition from local scale-invariant features. In International Conference on Computer Vision, Corfu, Greece. – 1999. – PP. 1150–1157.
4. Gonzalez R.C. and Woods R.E. Digital Image Processing 3rd edition, Pearson Prentice Hall, 2008.
5. Loizou C. P., Pattichis C. S., Pantziaris M., Tyllis T., Nicolaidis A. Quality evaluation of ultrasound imaging in the carotid artery based on normalization and speckle reduction filtering // Med Biol Eng Comput. – 2006. – Vol. 44, № 5. – PP. 414–426.
6. Loizou C.P., Pattichis C.S., Pantziaris M., Nicolaidis A. Snakes based segmentation of the common carotid artery intima media // Med Bio Eng Comput. – 2007. – № 45. – PP. 35–49.
7. Yu-Hsiang Wu, Jhu-Yun Huang, Shyi-Chyi Cheng Evolutionary Feature Construction for Ultrasound Image Processing and Its Application to Automatic Liver Disease Diagnosis // IEEE International Conference on Complex, Intelligent, and Software Intensive Systems 2011.
8. Беликова Т.А., Скобцов В.Ю. Эволюционный поиск эффективных последовательностей фильтров в задаче бинаризации УЗ изображений // Труды ИПММ НАН Украины. – 2011. – Т. 23. – С. 21–34.
9. Беликова Т.А., Скобцов В.Ю. Генетический алгоритм в задаче фильтрации УЗ изображений и анализ эффективности его модификаций // Вестник Херсонского Национального Технического Университета. – 2012. – № 1 (44). – С. 331–338.
10. John Stoitsis, Spyretta Golemati, Konstantina S. Nikita A Modular Software System to Assist Interpretation of Medical Images—Application to Vascular Ultrasound Images // IEEE Transactions On Instrumentation And Measurement, Vol. 55, no. 6, December 2006.



## EVOLUTIONARY METHODS FOR SEARCH DEPENDENCIES BETWEEN TEXTURE PARAMETERS OF UI AND QUALITY OF IMAGE PROCESSING METHODS

**T. Mahno**

Institute Of Applied Mathematics And Mechanics

vul. Rozy Luksemburg, 74, Donetsk, 83114, Ukraine. E-mail: belikova.taisija@gmail.com

There are a lot of systems which can facilitate an extent assessment of arteries atherosclerotic lesions by ultrasound scan. One of the tasks of these systems is segmentation of dense tissue and lumen of human artery on ultrasound image. Usually, this problem is solved by binarization algorithms, or carried out by experts manually. There are a lot of works, in which different separation algorithms are proposed in the task of dense tissue and artery lumen separation on ultrasound images of human carotid artery. In the previous works we developed genetic algorithm which can automatically construct sequences of filtration algorithms. The genetic algorithm found different solutions for predetermined processing result. However, these solutions were not equally good for all images. In this paper the genetic algorithm is developed which allows to conduct a preliminary analysis of an image and the selection of appropriate processing methods. We performed an analysis to determine if there is an association between filters sequences and images texture parameters. The algorithm was tested on 375 real ultrasound images of human carotid arteries.

**Key words:** image processing, genetic algorithm, Spatial Gray Level Dependence Matrices, ultrasound.

### REFERENCES

1. Gonzalez, R.C., Woods, R.E., (2008) Digital Image Processing: 3rd edition, Pearson Prentice Hall, USA.
2. Weszka, J.S., Dyer, C.R., Rosenfeld, A., (1976) "A Comparative Study of Texture Measures for Terrain Classification", *IEEE Transactions On Systems, Man, And Cybernetics*, vol. Smc-6, no. 4, pp. 269–285.
3. Lowe, D., (2004), "Distinctive image features from scale-invariant keypoints", *International Journal of Computer Vision*, vol. 60(2), pp. 91–110.
4. Lowe, D.G., (1999), "Object recognition from local scale-invariant features", *Transactions of International Conference on Computer Vision*, Corfu, Greece, pp. 1150–1157.
5. T.Belikova, V.Skobtsov (2011), "Evolutionary search for effective sequences of filters in the problem of ultrasonic images binarization", *Transactions of IAMM of NAS of Ukraine*, vol 23, pp. 21–34.
6. T.Belikova, V.Skobtsov (2012), "Analysis of genetic algorithm modifications in the problem of filtering ultrasound images" *Vestnik Khersonskogo Natsionalno-go Tehnicheskogo Universiteta*, no. 1(44), pp. 331–338.
7. Wu, Yu-H., Huang, J.-Y., Cheng S.-C., (2011), "Evolutionary Feature Construction for Ultrasound Image Processing and Its Application to Automatic Liver Disease Diagnosis", *IEEE International Conference on Complex, Intelligent, and Software Intensive Systems*, Seoul, June 30 -July 2 2011, pp. 565–570.
8. Stoitsis, J., Golemati, S., Nikita K. S., (2006), "A Modular Software System to Assist Interpretation of Medical Images—Application to Vascular Ultrasound Images", *IEEE Transactions On Instrumentation And Measurement*, vol. 55, no. 6, pp. 135–140.
9. Loizou, C. P., Pattichis, C. S., Pantziarism, M., Tyllis, T., Nicolaidis, A., (2006), "Quality evaluation of ultrasound imaging in the carotid artery based on normalization and speckle reduction filtering", *Med Biol Eng Comput*, vol. 44, no. 5, pp. 414–426.
10. Loizou, C. P., Pattichis, C. S., Pantziaris, M., Nicolaidis, A., (2007), "Snakes based segmentation of the common carotid artery intima media", *Med Bio Eng Comput*, no. 45, pp. 35–49.

Стаття надійшла 21.01.2015.