

Разработка параллельной реализации метода формирования информативных признаков в области пространственного спектра

Н.С. Кравцова^а, Р.А. Парингер^{а,б}, А.В. Куприянов^{а,б}

^а Самарский национальный исследовательский университет имени академика С.П. Королева, 443086, Московское шоссе, 34, Самара, Россия

^б Институт систем обработки изображений Российской академии наук – филиал ФНИЦ «Кристаллография и фотоника» РАН, 443001, Молодогвардейская, 151, Самара, Россия

Аннотация

В работе предлагается параллельная реализация метода выделения информативных сегментов изображений. Изображения сегментируются в области пространственного спектра. Признаками является среднее значение энергии в каждом выбранном сегменте. Для снижения временных затрат была разработана параллельная реализация этапа вычисления признаков. Экспериментальное исследование разработанной программной реализации было проведено на высокопроизводительной многоядерной вычислительной системе.

Ключевые слова: текстурные изображения; пространственный спектр; локальные признаки; дискриминантный анализ; параллельная реализация

1. Введение

В настоящее время компьютерные методы обработки медико-диагностических изображений используются в качестве одного из важнейших инструментов исследований и повышения эффективности ранней диагностики различных заболеваний. Одним из информативных показателей состояния здоровья человека является изменение состава биологических жидкостей [1-4]. Изменение метаболизма, происходящее при появлении патологического состояния, влияет на состав жидкости, происходят многочисленные изменения молекулярного состава тканей и биологических жидкостей. Одним из способов выявления этих изменений является перевод жидкостей из одного фазового состояния в другое. В лабораторной диагностике наиболее удобным способом изменения фазового состояния жидкости является её кристаллизация. Изменение свойств кристаллов является следствием изменения физико-химических свойств биологической жидкости. Медицинские кристаллограммы – это структуры, образованные при кристаллизации солей, вследствие высушивания биологических жидкостей.

В клинической практике кристаллограммы анализируют по их изображениям. Визуально не всегда возможно точно определить изменение основных параметров кристаллограмм, таких как преимущественное направление полос, их густота, и т.д., из которых складываются ведущие патологические признаки. Преимуществом компьютерного анализа изображений можно назвать возможность количественного анализа и объективность.

Информация, содержащаяся на изображении, является структурно избыточной. Известно, что, если на исходном изображении преобладали параллельные полосы определённого направления, тогда и на Фурье-преобразовании исходного изображения будут преобладать полосы с таким же направлением. Это свойство можно использовать для анализа кристаллограмм [5-8], а также других изображений, имеющих ветвистую структуру.

В работе для определения характеристик исходных изображений кристаллограмм используется информативный набор признаков, для формирования которого применяется разработанный метод, основанный на алгоритме дискриминантного анализа.

Основным способом повышения быстродействия программной реализации метода является использование параллельной реализации вычислений.

2. Метод формирования информативных признаков

2.1. Описание используемых признаков

В данной работе используются признаки, полученные путём вычисления общей энергии на выбранных областях изображения спектра. Большая часть спектра содержит информацию, не пригодную для определения характеристик исходного изображения.

Для получения признаков нужно рассматривать функцию исходного изображения в пространственной области и её преобразование Фурье $F(u, v)$. Тогда величина $|F(u, v)|^2$ определяет энергетический спектр изображения.

В качестве признаков применяются значения, полученные путём вычисления общей энергии на выбранной области изображения спектра. Сегментация изображения спектра на области осуществляется по формуле:

$$C_{r_1 r_2 \theta_1 \theta_2} = \sum_{r=r_1}^{r_2} \sum_{\theta=\theta_1}^{\theta_2} |F(r, \theta)|^2$$

где r_1 , r_2 , θ_1 и θ_2 – углы, ограничивающие сектор.

Так как изображение спектра симметрично относительно центра, то для выделения признаков будет использоваться только половина изображения пространственного спектра [9].

2.2. Технология построения эффективного набора признаков для различения изображений

В данной работе описан способ выделения информативных сегментов на изображениях спектра (на рис. 1 изображен как этап интеллектуального анализа признаков). Оценка информативности сегментов проводилась с использованием критериев разделимости алгоритма дискриминантного анализа.

Для формирования новых признаков, направленных на решение конкретной задачи, хорошо зарекомендовали себя методы, основанные на алгоритме дискриминантного анализа [10]. Данные методы позволяют повысить достоверность классификации данных [11, 12].

Дискриминантный анализ используется для устранения корреляции между признаками и, как следствие, уменьшения размеров набора признаков. Использование алгоритма позволяет с одной стороны сохранить информативность набора признаков для классификации, а с другой уменьшить число признаков, что в свою очередь позволяет использовать более простые методы классификации и снижает значение ошибки классификации.

Для каждого признака был рассчитан индивидуальный критерий разделимости, вычисляемый по формуле: $J = \text{tr}((\mathbf{T})^{-1}\mathbf{B})$, где $\mathbf{T} = \mathbf{B} + \mathbf{W}$, \mathbf{B} – матрица межгруппового рассеяния, \mathbf{W} – матрица внутригруппового рассеяния.

Дальнейшее формирование информативного набора признаков производилось следующим образом:

- Признаки ранжируются по уменьшению значений индивидуальных критериев разделимости.
- Начальный набор признаков состоит из признака с наибольшим критерием. Проводится классификация.
- Добавляется признак, следующий по значению критерия разделимости. Проводится классификация по новому набору признаков.
- Повторять пункт 3, пока в набор не будут добавлены все признаки.

Информативным будет считаться набор признаков, для которого результат классификации будет давать наименьшее значение ошибки. Ошибка классификации, определяющая, в скольких случаях классификатор принял неправильное решение, вычисляется по формуле: $\varepsilon = (m/n) \cdot 100\%$, – где m – число ошибок классификации, n – общее число тестируемых изображений.

3. Параллельная реализация метода формирования информативных признаков

Весь алгоритм формирования информативных признаков можно представить в виде схемы, изображенной на рис. 1. При анализе времени работы данного алгоритма было выявлено, что большая часть времени работы приходится на первый этап вычисления признаков.



Рис. 1. Схематическое изображение алгоритма формирования информативных признаков.

Как было сказано ранее, наибольшие временные затраты при работе алгоритма приходятся на этап вычисления признаков. На этом этапе нужно произвести предобработку обучающей выборки, сформировать изображения пространственного спектра для каждого обучающего изображения и рассчитать значение признаков. Количество расчётов можно сократить, если принять во внимание тот факт, что изображение пространственного спектра симметрично относительно центра изображения, поэтому расчёт признаков можно и нужно производить только на половине изображения пространственного спектра. Для увеличения скорости работы алгоритма, в данной работе будет использовано разделение по задачам. В данной работе будет использована не технология MPI, а распределение задач по потокам, поэтому неважно как именно нужно делить изображения на задачи. В случае предобработки выборки и формирования изображений пространственного спектра, на каждый поток подаётся одно изображение. Далее, на этапе вычисления признаков, на каждый поток подаётся отдельный элемент – сегмент изображения, вычисленный на основе

предложенного разбиения. Сразу же после окончания работы потока с элементом, на него передаётся следующая задача.

Применение данной схемы разбиения позволяет значительно сократить время выполнения первого этапа алгоритма. На этом наиболее ресурсоёмком этапе данный способ разделения задач на потоки позволил получить ускорение в три раза при использовании четырёх потоков. График зависимости ускорения от количества признаков представлен на рис. 2.

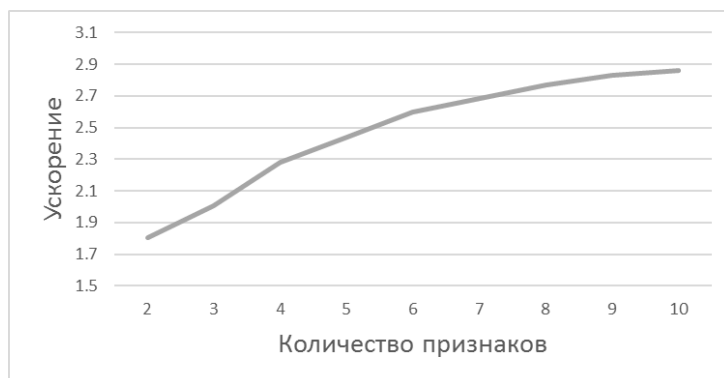


Рис. 2. Зависимость ускорения от количества признаков.

4. Заключение

В данной работе был представлен метод формирования набора информативных локальных признаков пространственного спектра для классификации изображений медицинских кристаллограмм, а также вариант его параллельной реализации. Кроме того, в работе было проведено экспериментальное исследование разработанной программной реализации, которое показало, что использование параллельной реализации алгоритма даёт на этапе вычисления признаков ускорение, стремящееся к 3. В дальнейшем планируется для повышения скорости вычисления информативного набора признаков введение параллельных вычислений на этапе интеллектуального анализа признаков.

Благодарности

Работа выполнена при частичной поддержке Министерства образования и науки РФ в рамках реализации мероприятий Программы повышения конкурентоспособности СГАУ среди ведущих мировых научно-образовательных центров на 2013–2020 годы; грантов РФФИ № 15-29-03823, № 15-29-07077, № 16-41-630761, № 16-29-11698; № 17-01-00972; программы № 6 фундаментальных исследований ОНИТ РАН «Биоинформатика, современные информационные технологии и математические методы в медицине» 2017 г.

Литература

- [1] Куприянов, А.В. Изучение кристаллограмм слезной жидкости на основе оценивания локальных статистических признаков о применении метода мер текстурной энергии. / А.В. Куприянов, Т.П. Чухман, Н.Ю. Ильясова, С.М. Свердлин // VI-я Международная научная конференция "Кинетика и механизм кристаллизации. Самоорганизация при фазообразовании" – Иваново, Россия – 21-24 сентября 2010 – С.366.
- [2] Камышев, И.А. Анализ планарных кристаллограмм лазерно облученных растворов биожидкостей / И.А. Камышев, А.Н. Малов, А.В. Неупокоева // Конференция – Иркутск – 2012.
- [3] Пьянкова, О.Б. Применение кристаллографии биологических жидкостей в диагностике генеза механической желтухи / Пьянкова О.Б. Карпунина Т.И. Карпунина Н.С. // Гастроэнтерология Санкт-Петербурга – 2010 – №2-3 – С.259.
- [4] Булкина, Н.В. Сравнительная характеристика кристаллографической картины ротовой жидкости и жидкости десневой борозды или пародонтальных карманов при диагностике воспалительных заболеваний пародонта / Н.В. Булкина, Г.Е. Бриль, Д.Э. Постнов, В.Т. Поделинская, О.В. Еремин // Российский стоматологический журнал – 2012 – №4 – С.12-16.
- [5] Paringer, R.A. The Method for Effective Clustering the Dendrite Crystallogram Images / R.A. Paringer, A.V. Kupriyanov // 9th Open German-Russian Workshop on Pattern Recognition and Image Understanding «OGRW 2014» – Electronic on-site Proceedings – University of Koblenz-Landau – 2014.
- [6] Парингер, Р.А. Метод формирования признаков для классификации изображений дендритных кристаллограмм / Р.А. Парингер, А.В. Куприянов // Статья в сборнике трудов конференции «Информационные и математические технологии в науке и управлении» – Иркутск, Россия – 2015 – С.50-55.
- [7] Paringer, R.A. Research methods for classification of the crystallogramms images / R.A. Paringer, A.V. Kupriyanov // Proceedings of the 12th international conference «PRIP'2014» – Minsk, Belarus – 2014 – Vol. 1 – pp. 231-234.
- [8] Парингер, Р.А. Разработка методов повышения качества классификации изображений дендритных кристаллограмм / Р.А. Парингер, А.В. Куприянов // Сборник научных трудов международной научно-технической конференции «ПИТ-2014» – Самара, Россия – 2014 – С.282-286.
- [9] Kravtsova, N. Development of methods for crystallogramms images classification based on technique of detection informative areas in the spectral space / N. Kravtsova, R. Paringer, A. Kupriyanov // CEUR Workshop Proceedings – 2016 – Vol.1638 – P. 357-363 DOI: 10.18287/1613-0073-2016-1638-357-363
- [10] Fukunaga, K. Introduction to statistical pattern recognition / San Diego: Academic Press – 1990 – p.592
- [11] Ильясова, Н.Ю. Формирование признаков для повышения качества медицинской диагностики на основе методов дискриминантного анализа / Н.Ю. Ильясова, Р.А. Парингер, А.В. Куприянов // Компьютерная оптика – 2014 – Т.38, №4 – С.81-856
- [12] Biryukova, E. Development of the effective set of features construction technology for texture image classes discrimination / E Biryukova, R. Paringer, A. Kupriyanov // CEUR Workshop Proceedings – 2016 – Vol.1638 – P. 263-269 DOI: 10.18287/1613-0073-2016-1638-357-363