

Методы обнаружения и выделения областей на текстурных изображениях

Г. Б. Абдикеримова, А. Л. Бычков, Вей Синьюй, Ф. А. Мурзин,
Н. Е. Русских, Е. И. Рябчикова, С. С. Хайрулин¹

В статье речь идет о методах анализа текстурных изображений. Рассматриваются микрофотографии растительного сырья, полученные просвечивающей электронной микроскопией. Работа выполнена для Института химии твердого тела и механохимии СО РАН. Основной целью исследований является разработка и реализация алгоритмов, позволяющих обнаруживать и выделять на изображении области, представляющих интерес для специалистов-химиков. Например, области, в которых происходит разупорядочение исходной структуры материала после применения различных механохимических методов обработки. Для решения поставленной задачи используются текстурные признаки, кластеризация, R/S-анализ, ортогональные преобразования, вейвлет-анализ. Большое внимание было уделено разработке программных инструментов, позволяющих осуществлять выбор признаков, описывающих текстурные различия, чтобы сегментировать текстурные области на подобласти. То есть исследуется вопрос о применимости наборов текстурных признаков и других параметров для анализа экспериментальных данных с целью выявить на микрофотографиях характерные участки, которые в будущем можно будет увязать с пористостью, химической реактивностью и т.д.

Ключевые слова: обработка изображений, микрофотографии, текстурные признаки, кластеризация, R/S-анализ, ортогональные преобразования, электронная микроскопия, растительное сырье.

1. Введение

Исследования направлены на создание научно-технического задела в области обработки изображений текстурного типа. Целью проводимых исследований является разработка и поиск алгоритмов анализа изображений, полученных из различных источников, например, при помощи современных электронно-микроскопических методов. В том числе анализируются изображения ультраструктуры растительных клеточных стенок, полученные просвечивающей электронной микроскопией растительного сырья, после различной физико-химической и/или механохимической обработки. В настоящее время данная информация востребована учеными из различных областей – химиками, биологами, технологами, но обрабатывается на качественном (редко – полуколичественном) уровне вручную. Переход к алгоритмам, позволяющим оперировать большими объёмами данных, позволил бы перечисленным областям науки сделать значительный шаг вперёд, к совершенствованию существующих и созданию новых технологических процессов.

В работе используются различные наборы текстурных признаков (около 20) и спектральные преобразования на основе ортогональных матриц (наибольший интерес представляют 6

¹ Работа поддержана Российским научным фондом (грант № 16-13-10200) и Российским фондом фундаментальных исследований (грант № 18-08-01284).

преобразований). Также проводились эксперименты по применению R/S-анализа и вейвлет-анализа.

Программные продукты, позволяющие детально анализировать текстуры, могут успешно применяться в различных областях науки и промышленности. Прежде всего, это химия и материаловедение. Можно анализировать материалы органического происхождения, срезы металлов и минералов, керамику и др.

Спектр задач не ограничивается анализом микрофотографий. Так, при обработке аэрокосмических снимков исследователи также имеют дело с различными текстурами. По текстурным признакам можно определить хвойный или лиственный лес, поля, засеянные зерновыми или бобовыми растениями, и др. Также можно выделить лес, пораженный вредителями, опустыненные территории. Другая область исследований, где эти методы могут быть эффективно использованы, – это диагностика внутренних патологий человека, в том числе злокачественных, методом анализа изображений, полученных с помощью тепловизора.

Принципиальное отличие идей проекта от существующих аналогов состоит в корректном применении математических методов и в более глубокой их проработке. Например, текстурных признаков известно более двухсот, в научных обзорах приводят обычно около пятидесяти. В то же время на практике, как правило, используют 3–4 признака, например, при обработке космических снимков. То есть исходные изображения остаются не исследованными в полной мере. То же самое можно сказать о применении интегральных преобразований. Например, при исследовании прочности металлов под нагрузками используют преобразование Хаара, чтобы охарактеризовать трещиноватость. Вопрос, какую информацию можно получить на основе других преобразований, почти не изучен. В литературе по химии древесины говорится о полезности R/S-анализа и фрактального анализа для соответствующих исследований, но сведения носят отрывочный характер.

2. Методы анализа текстурных признаков

Несмотря на повсеместное присутствие текстур в изображениях, единого и формального подхода к описанию текстуры и строгого её определения на данный момент не существует. Методы анализа текстур, как правило, разрабатываются для каждого отдельного случая.

В [1] под текстурой понимают «пространственную организацию элементов в пределах некоторого участка поверхности». Там же объясняется, что эта организация обусловлена определенным статистическим распределением интенсивности серых тонов или тонов различного цвета. Участок может считаться текстурным, если количество отмечаемых на нем перепадов интенсивности или изменений цвета достаточно велико. В [2] текстурой называют «некоторым образом организованный участок поверхности». В [3] текстура определяется как матрица или фрагмент пространственных свойств участков изображений с однородными статистическими характеристиками.

Текстуры можно разделить на несколько классов следующим образом:

- 1) по происхождению: на искусственные – например, графические узоры, и естественные – например, трава, лес, земля;
- 2) по структуре поверхности: структурные, состоящие из геометрически правильных повторяющихся элементов, и стохастические, сформированные последовательностью случайных элементов; по относительным размерам элементов текстуры: мелкозернистые и крупнозернистые;
- 3) по форме элементов текстуры: волнистые, пятнистые, неправильные, линейчатые и так далее [1].

Из приведенных выше определений и характеристик следует, что текстура – это некоторый участок изображения, тот, который имеет однородные статистические характеристики. Это значит, что каждую текстуру данного класса можно описать с помощью характерного свойства, общего для всех текстур данного класса [1]. Такие свойства называют текстурными

признаками [4]. Текстуальные признаки играют важную роль при разделении изображения на отдельные области.

В качестве таких признаков можно использовать статистические характеристики пространственных распределений, вычисляемые как меры однородности по одномерной гистограмме значений сигналов (характеристики 1-го порядка) и по двумерным гистограммам значений сигналов (характеристики 2-го порядка).

Второй класс – это признаки, учитывающие взаимное расположение. Для формирования текстурных признаков, учитывающих взаимное расположение пикселей внутри скользящего окна, применяется подход, основанный на использовании матрицы смежности (другое название – матрица распределения градиентов [4]).

В рассматриваемой задаче исследуются серые (полутонные) изображения. Таким образом, изображение задается в виде матрицы, значениями которой являются значения яркости пикселей в интервале от 0 до 255.

Стандартный подход для вычисления текстурных признаков следующий. Необходимо выбрать так называемое бегущее окно с нечетной стороной: 3, 5, 7 пикселей. Признак вычисляется внутри бегущего окна. Размер локального фрагмента является носителем текстурных свойств. Значение признака записывается в новую матрицу того же размера, что и исходная. В новой матрице значение записывается в точку с координатами, равными координатам центра бегущего окна. Элементы новой матрицы получаются в некотором интервале $[A, B]$. Далее обычно этот интервал линейно отображается в отрезок $[0, 255]$. После этого имеется возможность визуализировать результат вычисления текстурного признака.

Эксперименты показали, что стандартный подход в нашем случае малоинформативен. Поэтому было решено использовать нестандартный подход. А именно, текстурные признаки вычисляются по большим окнам (в том числе по неквадратным), которые пользователь может задавать, выбирая область, которая может представлять для него интерес. То есть речь идет о вычислении числовых характеристик, относящихся к обширным областям, включающим различного рода артефакты.

Рассмотрены 18 наиболее важных текстурных признаков, все они реализованы в программе. Ниже в таблице приведены некоторые из них.

Таблица 1. Примеры текстурных признаков

Признаки, основанные на статистических характеристиках	Признаки, учитывающие взаимное расположение
k-th начальный момент $T_1^k = n^{-2} \sum_{i=1}^n \sum_{j=1}^n [f(i, j)]^k.$ Энтропия $T_2 = - \sum_{g=0}^{N-1} F(g) \log_{10} F(g).$ Энергия $T_3 = \sum_{k=0}^{N-1} [F(g)]^2.$ Вариация $T_4 = - \sum_{g=0}^{N-1} (g - \mu)^2 F(g).$	Среднее $T_1 = \mu_i = \mu_j = \sum_{i=0}^{N-1} \left[i \sum_{j=0}^{N-1} P(i, j) \right].$ Энергия $T_2 = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} [P(i, j)]^2.$ Вариация $T_3 = \sigma_i^2 = \sum_{i=0}^{N-1} \left[(i - \mu_2)^2 \sum_{j=0}^{N-1} P(i, j) \right].$ Однородность $T_4 = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} P(i, j) / (1 + i - j).$

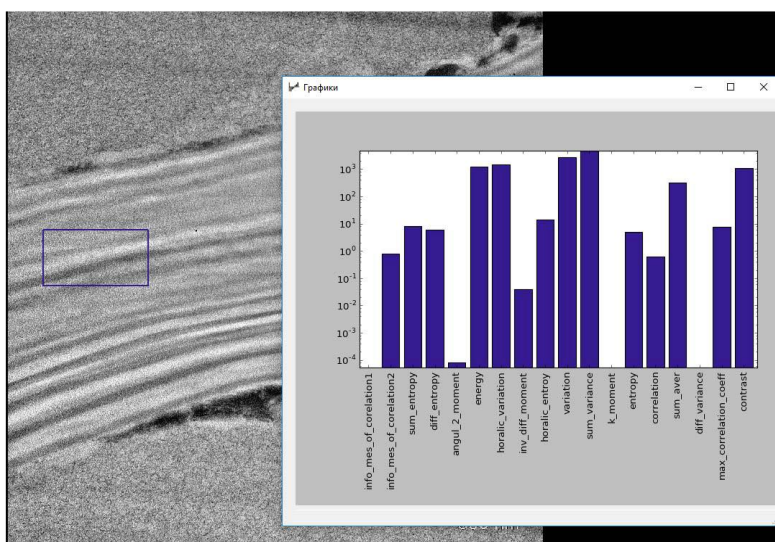


Рис. 1. Вычисление текстурных признаков

3. Алгоритмы кластеризации

Сегментация изображения может быть осуществлена локальным и глобальным способом. Локальный способ оценивает наличие границы между областями по поведению признаков в окрестности точки изображения. Глобальный способ предполагает предварительную кластеризацию пространства признаков и затем установление соответствия между пикселем изображения и кластером, в который попадает его вектор признаков [5].

Одни из эвристических методов кластеризации – методы, основывающиеся на последовательной агломеративной процедуре. Достоинством этих методов является простота вычислительной процедуры и алгоритмов.

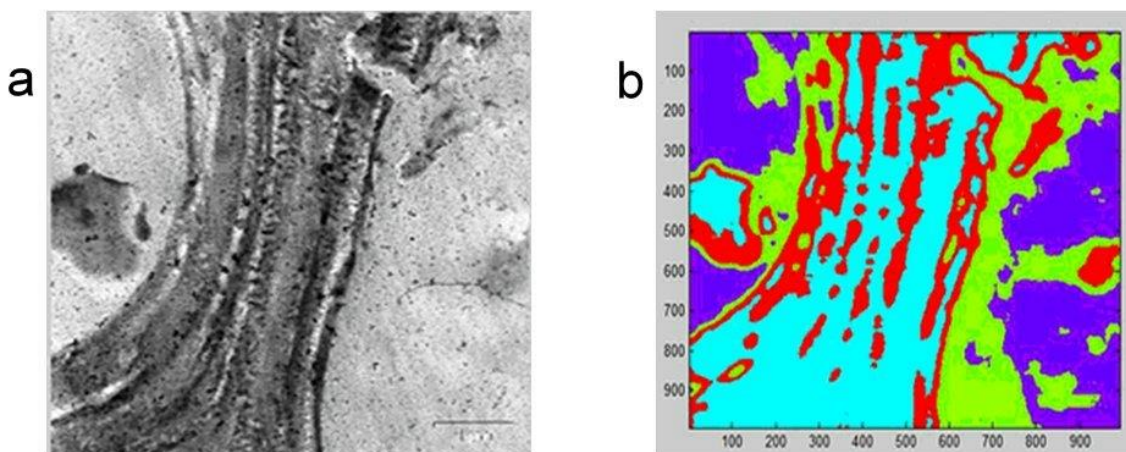


Рис. 2. (a) исходное изображение; (b) результат сегментации

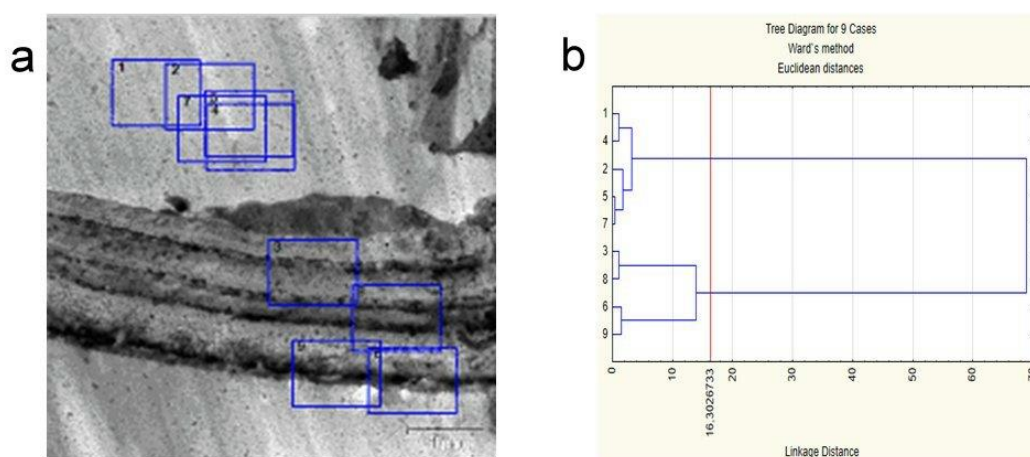


Рис. 3. (a) выбор фрагментов изображения; (b) иерархическая кластеризация

4. R/S-анализ для изображений

Метод R/S-анализа – это статистический метод анализа временных рядов или числовых последовательностей (преимущественно финансовых), позволяющих определить некоторые важные их характеристики, такие как наличие непериодических циклов, «памяти» у процесса, степень хаотичности и т.д. [6]. Метод применим и для анализа изображений, т.к. можно рассматривать последовательность значений функции яркости вдоль некоторой прямой или кривой линии. Важной характеристикой, вычисляемой методом R/S-анализа, является показатель Хёрста, обычно обозначаемый H , который характеризует степень хаотичности процесса [7]. Показатель Хёрста также называют иногда фрактальной размерностью. Интересно, что метод применяется в химии для анализа микрофотографий (например, лигнина) [8, 9] и он показал определенную эффективность.

Метод программно реализован в среде MATLAB, проводится анализ изображений. Изображения программными средствами делится на отдельные участки. Если участок изображения принадлежит так называемой области скейлинга, то это позволяет выявить с помощью оценки фрактальной размерности соседних участков более мелкие электроноплотные (чёрные) образования, которые, скорее всего, являются остатками внутриклеточного содержимого липидной или белковой природы (дробный метод исследования). При этом фрактальная размерность участков микрофотографий клеточных стенок (например, соломы пшеницы) будет лежать в интервале нормы, а фрактальная размерность участков с лигнифицированными слоями – ниже границы нормы. Было произведено разделение изображения на 1024 части. Для каждой части изображения выполнялся расчет фрактальной размерности.

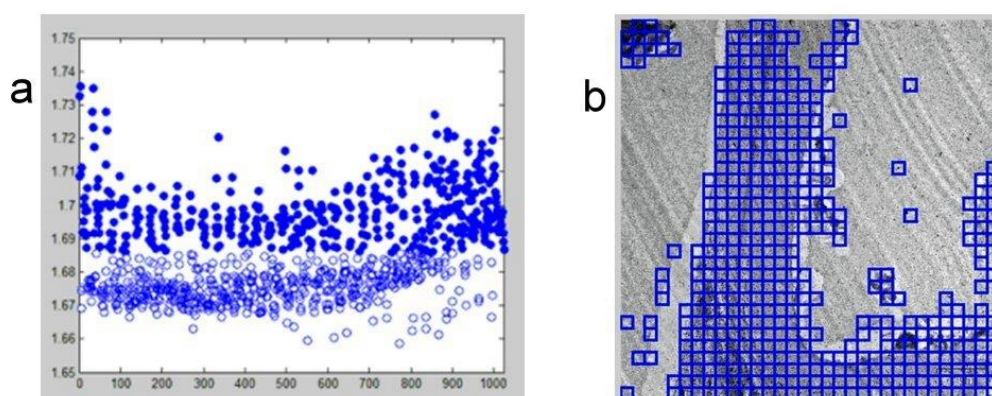


Рис. 4. (a) значения показателя Хёрста; (b) выделенная область

Результаты применения дробного метода изображаются в виде множества кружков, отражающих вычисленную фрактальную размерность. Синие кружки соответствуют синим квадратам на исходном изображении, белые кружки указывают на фон (рис. 4).

5. Ортогональные преобразования

Спектральный анализ является мощным инструментом анализа сигналов и изображений. Уже давно замечено, что спектр очень чутко реагирует на различные изменения в структуре сигналов и изображений [10].

Для проведения спектрального анализа необходимо предварительно разложить сигнал или изображение по частотам. Для этого применяются [11] различные наборы базисных функций. Соответствующие алгоритмы называют преобразованиями: косинусное, Адамара, Хаара, наклонное и др. Отметим, что преобразования Хаара и Добеши являются простейшими вейвлет-преобразованиями.

Программа реализована в среде MATLAB и позволяет осуществлять спектральные преобразования шести видов: 1) косинусное, 2) Адамара порядка 2^n , 3) Адамара порядка $n = p + 1$, $p \equiv 3 \pmod{4}$ – простое число, т.е. на основе символа Лежандра, 4) Хаара, 5) наклонное, 6) Добеши-4. Пример преобразования приведен ниже (рис. 5).

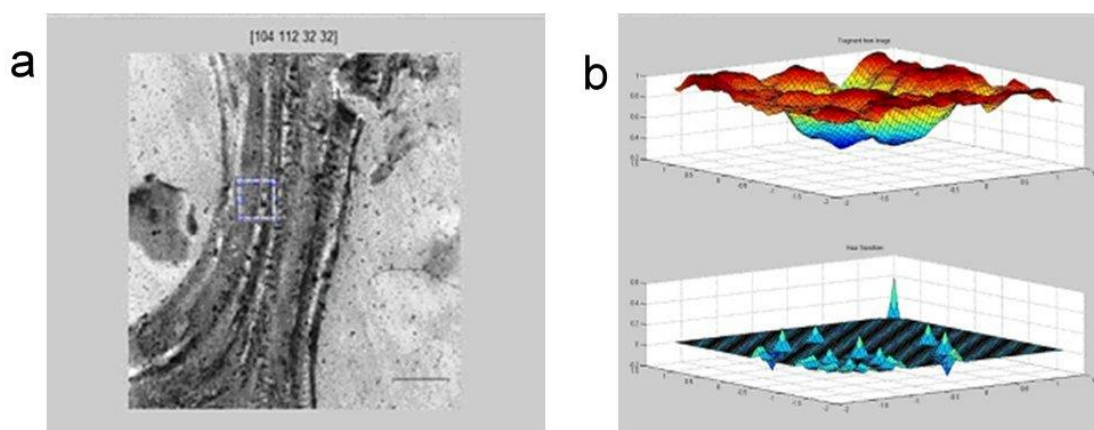


Рис. 5. (a) выделенное окно; (b) функция яркости в окне и результат преобразования Хаара

Для изучения текстур можно применять разные нестандартные подходы с использованием ортогональных преобразований. Например, исходное изображение разбивается на непересекающиеся квадратные окна. Как показывают эксперименты, целесообразно взять размер окна достаточно большим, например, 32×32 , 64×64 и т.д. Затем в каждом окне мы делаем ортогональное преобразование. Далее некоторые спектральные коэффициенты могут быть отброшены, например, высокочастотные или в нескольких определенных частях спектра. В двумерном случае частотные спектры представляют собой двумерные матрицы. Естественно, что можно расположить элементы матриц в векторах. Например, строки матрицы могут быть расположены последовательно одна за другой. Фактически позиции в векторе, в которых расположены спектральные коэффициенты, которые мы обнулили, могут быть вычеркнуты из результирующего вектора. Далее процедура кластеризации может быть выполнена применительно к полученным векторам.

Другой подход – это вейвлет-декомпозиция [12]. Можно сказать, что вейвлет-разложение [8] выполняется с помощью ассоциированных с деревом блоков двухканальных фильтров. В вершине этого дерева, высота которого равна d , должен находиться сигнал длины, равной

2^d . Соответственно, размер изображения должен быть $2^d \times 2^d$. Если это нарушено, то недостающие значения обычно добавляются равные нулю.

На каждом шаге преобразования изображение разбивается на 4 матрицы. Одна из них представляет собой изображение, похожее на оригинал, но оно меньше по горизонтали и по вертикали в 2 раза. Можно сказать, что это «грубая» версия исходного изображения. В трех других матрицах «кодируются» различные перепады яркости исходного изображения. Таким образом производится «детализация» информации, содержащейся в изображении.

Значения элементов могут не попадать в интервал $\{0, \dots, 255\}$. Для каждой из матриц мы находим значения минимального и максимального элементов и отображаем соответствующий интервал в интервал $\{0, \dots, 255\}$. Это позволяет визуализировать результаты в виде серых или цветных изображений.

Простейшими формами вейвлет-преобразований для изображений являются преобразования Хаара и Добеши-4. Ниже приведены результаты такого рода эксперимента для преобразования Хаара (рис. 6).

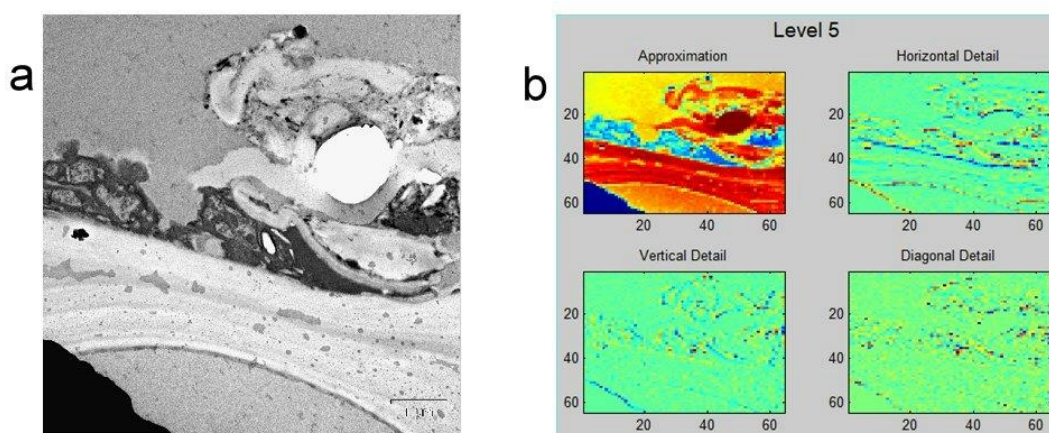


Рис. 6. (а) исходное изображение;
(б) 5-й уровень вейвлет-декомпозиции посредством преобразования Хаара

6. Заключение

Работа посвящена изучению текстурных изображений, сегментации текстур, ортогональным преобразованиям, R/S-анализу и т.д. Исследуются микрофотографии растительного сырья, измельченного на специальных мельницах и подвергнувшегося нагреванию или охлаждению, в том числе до криогенных температур. Работа выполнена для Института химии твердого тела и механохимии СО РАН.

В основном речь идет о расчете различных видов текстурных признаков и других характеристик изображений. Был проанализирован большой набор методов исследования текстур, из рассмотренных методов были выбраны наиболее информативные. В итоге были реализованы следующие программные инструменты:

1. Программа для расчета различных текстурных признаков.

2. Программы для кластерного анализа в стандартном и иерархическом вариантах. Были проведены эксперименты. В первом случае использовалось 18 текстурных признаков с окном прокрутки 9×9 пикселей. Во втором – 5 текстурных признаков с окном 400×300 пикселей. Точнее, во втором случае после проведения многочисленных экспериментов были выбраны пять наиболее информативных текстурных признаков, по которым и была проведена кластеризация.

3. Прототипы программ, использующих R/S-анализ. Эксперименты показали интересные результаты, являющиеся полезными на практике. Фон и клеточная стенка четко отделяются

друг от друга. Отметим, что подобные высказывания встречаются в научной литературе по изучению лигнина, но без особых деталей.

4. Программа, позволяющая осуществлять спектральные преобразования шести видов. Рассмотрены некоторые их приложения. Например, вейвлет-декомпозиция изображений.

В дальнейших исследованиях векторы показателей могут быть связаны с химической реактивностью, пористостью и другими характеристиками. Программная система может быть обучена на примерах с использованием алгоритмов, основанных на нейροкомпьютерных или других подходах, обычно применяемых в машинном обучении. После обучения система сможет предсказывать значения параметров.

Литература

1. Андреев Г. А., Базарский О. В., Глауберман А. С., Колесников А. И., Коржик Ю. В., Хлявич Я. Л. Анализ и синтез случайных пространственных текстур // Зарубежная радиоэлектроника. 1984. № 2. С. 3–33.
2. Харалик Р. М. Статистический и структурный подходы к описанию текстур // ТИИЭР. 1979. Т. 67, № 5. С. 98–119.
3. Потапов А. А. Новые информационные технологии на основе вероятностных текстурных и фрактальных признаков в радиолокационном обнаружении малоконтрастных целей // Радиотехника и электроника. 2003. Т. 48, № 9. С. 1101–1119.
4. Колодникова Н. В. Обзор текстурных признаков для задач распознавания образов // Доклады ТУСУРа. Автоматизированные системы обработки информации, управления и проектирования. 2004. С. 117–118.
5. Sidorova V. S. Hierarchical Cluster Algorithm for Remote Sensing Data of Earth // Pattern Recognition and Image Analysis. 2012. V. 22, № 2. P. 373–379.
6. Федер Е. Фракталы. М.: Мир, 1991. 192 с.
7. Петерс Э. Хаос и порядок на рынках капитала. М.: Мир, 2000. 90 с.
8. Карманов А. П. Лигнин. Структурная организация и самоорганизация // Тезисы докладов III Всероссийской конференции «Химия древесины и органический синтез». Институт химии Коми научного центра Уральского отделения РАН, Сыктывкар, 1999. С. 66–67.
9. Карманов А. П., Матвеев Д. В. Проблемы химии древесины и лесохимии // Институт химии Коми научного центра Уральского отделения РАН, Сыктывкар. 2001. С. 50–52.
10. Rani R. Performance analysis of different orthogonal transform for image processing application // Inter. J. of Applied Research. 2015. V. 1, № 12. P. 844–847.
11. Shahdoosti H. R., Mirzapour F. Spectral-spatial feature extraction using orthogonal linear discriminant analysis for classification of hyperspectral data // European J. of Remote Sensing. 2017. V. 50, № 1. P. 111–124.
12. Воробьев В. И., Грибунин В. Г. Теория и практика вейвлет-преобразования. Санкт-Петербург: Изд.-во Военного университета связи, 1999. 92 с.

Статья поступила в редакцию 27.08.2019.

Абдикеримова Гульзира Бахытбековна

ст. преподаватель, Евразийский национальный университет им. Л. Н. Гумилева (010000, Казахстан, Нурсултан, ул. Сатбаева, 2), e-mail: gulzira1981@mail.ru.

Бычков Алексей Леонидович

к.х.н., с.н.с., Институт химии твердого тела и механохимии СО РАН (630128, Новосибирск, ул. Кутателадзе, 18), e-mail: bychkov.a.l@gmail.com.

Вей Синьюй

аспирант, Хэйлунцзянский университет (150000, Харбин, Китай, ул. Сюефу 74), e-mail: xinyuwei2016@163.com.

Мурзин Федор Александрович

к.ф.-м.н., зам. директора по научной работе, Институт систем информатики им. А. П. Ершова СО РАН (630090, Новосибирск, пр. Академика Лаврентьева, 6), e-mail: murzin@iis.nsk.su.

Русских Николай Евгеньевич

аспирант, Институт систем информатики им. А. П. Ершова СО РАН, e-mail: russkikh.nikolay@gmail.com.

Рябчикова Елена Ивановна

д.б.н., профессор, руководитель группы микроскопических исследований, Институт молекулярной биологии и фундаментальной медицины СО РАН (630090, Новосибирск, пр. Академика Лаврентьева, 8), e-mail: lenryab@niboch.nsc.ru.

Хайрулин Сергей Сергеевич

м.н.с., Институт систем информатики им. А. П. Ершова СО РАН, e-mail: s.khayrulin@gmail.com.

Methods for detecting and highlighting areas in textural images

**G. B. Abdikerimova, A. L. Bychkov, Wei Xinyu, F. A. Murzin, N. E. Russkikh,
E. I. Ryabchikova, S. S. Khayrulin**

The article deals with methods for analyzing textural images. Micrographs of plant materials obtained by the transmission electron microscopy are considered. This work was carried out for the Institute of Solid State Chemistry and Mechanochemistry SB RAS. The main goal of the research is the development and implementation of algorithms allowing us to detect and highlight areas of interest to chemists in the image. For example, areas, in which the initial structure of the material is disordered after applying various mechanochemical methods of processing. To solve this problem, we use: analysis of textural features, clustering, R/S-analysis, orthogonal transformations, wavelet analysis. Much attention was paid to the development of software tools that allow us to select features describing textural differences in order to segment textural regions into subregions. So, the question of the applicability of sets of textural features and other parameters for the analysis of experimental data is being investigated in order to identify characteristic areas in microphotographs that can be associated in future with the porosity, chemical reactivity, etc.

Keywords: image processing, microphotographs, textural features, clustering, R/S-analysis, orthogonal transformations, electron microscopy, plant materials.