

УДК 677.026.4.019

## ПРИМЕНЕНИЕ МЕТОДОВ КЛАСТЕРНОГО АНАЛИЗА ДЛЯ ОЦЕНКИ ЗАСОРЕННОСТИ ДВУМЕРНЫХ ВОЛОКНИСТЫХ МАТЕРИАЛОВ

*В.И.ЛЕБЕДЕВА, П.А.СЕВОСТЬЯНОВ**(Московский государственный текстильный университет им.А.Н. Косыгина)*

Автоматизация обнаружения сорных примесей и оценки их содержания в плоских волокнистых материалах (ватке-прочесе, нетканых материалах и др.) включает решение задач получения информации о неравномерности анализируемого образца и выделения из этой неравномерности составляющих, связанных с наличием сорных примесей.

Задача исследования неровноты плоских волокнистых материалов решалась рядом авторов [1]. Современные компьютерные средства и технологии позволяют получить информацию о неравномерности образца, используя электронные цифровые сканирующие устройства, причем обычные компьютерные сканеры, как показали эксперименты по их использованию для этих целей, обладают достаточной чувствительностью и обеспечивают воспроизводимость результатов без специальных мер по подготовке изучаемого образца. Сканеры позволяют регистрировать и сохранять в дискретном виде информацию об отражательной способности поверхности образца.

Оставляя открытым вопрос о коррелированности этого показателя с плотностью материала по массе, отметим, что для выявления сорных примесей анализ отражательной способности поверхности образца может быть даже более полезным, чем обработка информации о его поверхностной или объемной плотности по массе.

Таким образом, можно считать, что первая задача – получение информации о неравномерности образца плоского волокнистого материала, включая и инфор-

мацию о наличии сорных примесей (выделяющихся цветом и отражательной способностью), решается сканирующими устройствами.

Вторая задача – выделения сорных примесей и оценки их содержания в материале – требует разработки специальных алгоритмов обработки полученной от сканера информации. Перспективным, на наш взгляд, для этих целей является использование теории и методов кластер-анализа алгоритмов на их основе [2], [3]. Изображения образца двумерного волокнистого материала, полученные в результате сканирования, содержат также и изображения включенных в него сорных примесей, которые выделяются цветом и интенсивностью окраски, что проявляется в изменении значений матрицы, описывающей поле неровноты образца. Для выявления таких включений осуществляются сечения поля на различных уровнях значений неровноты.

В результате отсечения изображения образца волокнистого материала по пороговому значению образуются локализованные пятна (объекты), внутри которых значение оттенка серого превышает пороговое значение.

Если пороговое значение выбрано недостаточно точно или примеси волокнистых материалов имеют неравномерную структуру, количество выделенных объектов может превышать количество реально существующих пороков. Для группирования объектов был применен метод кластерного анализа [2], который позволяет

разделить множество полученных пятен на подмножества.

В качестве показателя близости объектов может быть выбрано евклидово расстояние между геометрическими центрами (центрами "масс") выделенных объектов:

$$R_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2},$$

где  $i, j = 1, \dots, n$  – индексы соседних объектов;  $n$  – число объектов.

В [3] был предложен алгоритм определения расстояния между кластерами, основанный на вычислении центра "массы" каждого из кластеров и объединении в один кластер тех из них, которые на данном этапе классификации наиболее близки по этой метрике. Применить методы "ближайшего соседа" и "удаленного соседа" невозможно, так как расстояние между точками, принадлежащими одному объекту, того же порядка, что и расстояние между различными объектами.

Принципиальным является и то, что пятна, получаемые при отсечении, имеют неправильную форму, и расстояния от двух крайних точек объекта до соседнего объекта существенно отличаются. Это обстоятельство и то, что объекты, объединяемые в кластеры, имеют неодинаковую площадь, потребовало внести изменения в существующий алгоритм.

Опишем алгоритм разбиения объектов на кластеры, используя следующие обозначения:  $n$  – число объектов, образованных в результате отсечения изображения волокнистого материала по пороговому значению;  $x_j, y_j, S_j$  – экранные координаты центра площади  $j$ -го объекта;  $NC_k$  – число кластеров на  $k$ -м шаге классификации пятен  $k=1, \dots, n$ ;  $x_c, y_c$  – экранные координаты центра "массы"  $i$ -го кластера,  $j=1, \dots, NC_k$ ;  $R_{ij}$  – расстояние между центрами "массы"  $i$ -го и  $j$ -го кластеров,  $i, j=1, \dots, NC_k$ .

1.  $k=1, NC_k=n$  (число кластеров равно числу выделенных объектов).

2.  $x_{c_i} = x_i, y_{c_i} = y_i, i=1, \dots, NC_k$ ; (центры "масс" кластеров совпадают с центрами "масс" объектов).

3. Определение расстояния между центрами "массы" кластеров:

$$R_{ij} = \sqrt{(x_{c_i} - x_{c_j})^2 + (y_{c_i} - y_{c_j})^2}, i, j=1, \dots, NC_k.$$

4. Определение номеров  $i_m$  и  $j_m$  двух наиболее близких друг к другу кластеров, для которых минимально  $R_{ij}$ .

5. Сохранение найденного минимального расстояния и соответствующего ему числа кластеров.

6. Объединение найденных кластеров в один с меньшим порядковым номером.

7. Вычисление координат центра "массы" вновь сформированного кластера. Смещение нового центра зависит от коэффициента, равного отношению площадей объединяемых кластеров, и определяется следующим образом:

$$b_k = \frac{S(i_m)}{S(j_m)}, \text{ если } S(i_m) < S(j_m),$$

или

$$b_k = \frac{S(j_m)}{S(i_m)}, \text{ если } S(j_m) < S(i_m).$$

Если выбранные объекты объединяются в кластер с номером  $i_m$ , то координаты нового центра "масс" по  $X$  и по  $Y$  по формулам:

$$\begin{aligned} y_{c(i_m)} &= y_{c(i_m)} + b_k(y_{c(j_m)} - y_{c(i_m)}), \\ x_{c(i_m)} &= x_{c(i_m)} + b_k(x_{c(j_m)} - x_{c(i_m)}). \end{aligned}$$

Если объекты объединяются в кластер с номером  $j_m$ , то по формулам:

$$\begin{aligned} y_{c(j_m)} &= y_{c(j_m)} + b_k(y_{c(i_m)} - y_{c(j_m)}), \\ x_{c(j_m)} &= x_{c(j_m)} + b_k(x_{c(i_m)} - x_{c(j_m)}). \end{aligned}$$

8. Исключение кластера с большим номером из множества кластеров и уменьшение числа кластеров  $NC_k = NC_k - 1$ .

9. Пока  $NC_k \leq 1$  (все объекты не объединятся в одно подмножество) переход к п.3.

Результатом работы данного алгоритма является последовательность значений минимальных расстояний между кластерами при различном их количестве, образуемом из исходного множества объектов. Последовательность минимальных расстояний между формируемыми кластерами должна обладать следующими свойствами. На начальных этапах алгоритма в каждый кластер будут входить элементы, принадлежащие одному и тому же компактному объекту. Поэтому минимальное расстояние будет нарастать монотонно и достаточно медленно.

Процесс будет продолжаться до тех пор, пока множество элементов не разделится на кластеры, равные компактным подмножествам исходного множества элементов. В этом случае минимальным будет расстояние между центрами "масс" ближайших объектов, которое существенно больше расстояния между элементами, об-

разующими один объект. В последовательности минимальных расстояний произойдет резкий скачок. Число кластеров в момент скачкообразного роста минимальных расстояний будет равно количеству объектов, характеристики которых существенно отличаются от заданных.

Эффективность алгоритма была проверена на моделируемом поле неровности волокнистого материала. В качестве включений к базовому полю были добавлены две поверхности, имеющие колоколообразную форму, описанную функцией:

$$f(x, y) = \beta \exp \left[ -\lambda \left( x^2 + \mu xy + y^2 \right) \right],$$

где  $\mu, \beta$  и  $\lambda$  – параметры.

Производились отсечения моделируемых изображений при разных пороговых значениях. В результате образовывалось разное количество объектов с разными площадями и расположением в пространстве.

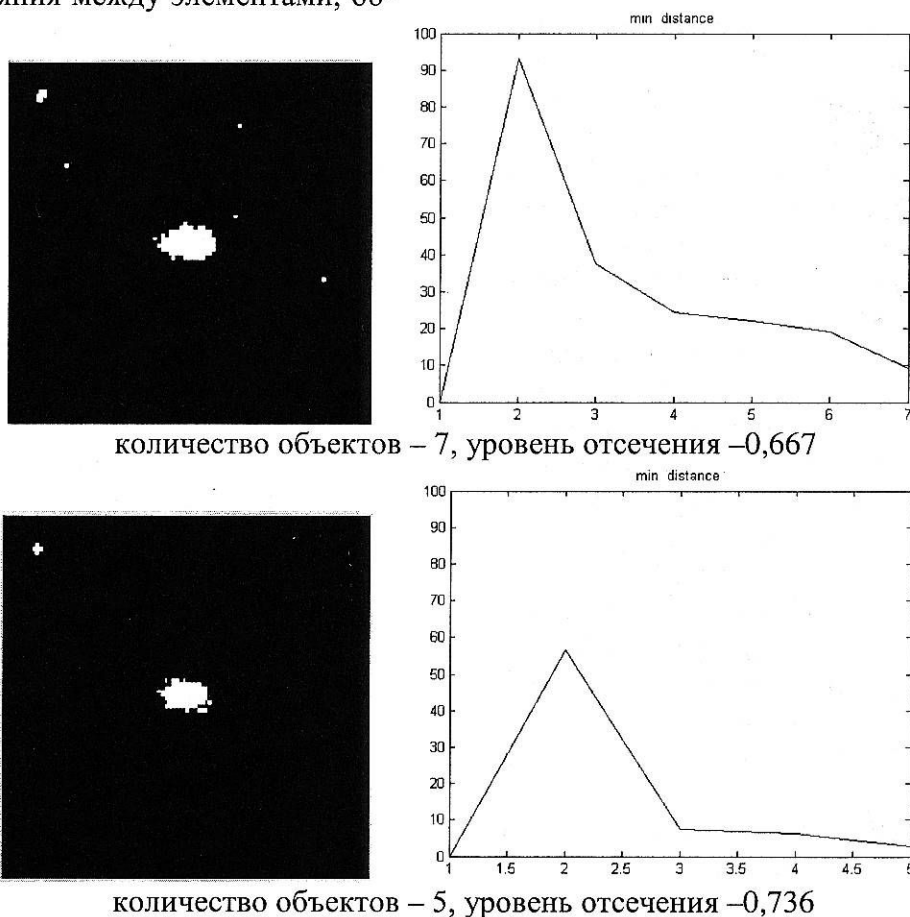


Рис.1

На графике (рис.1) представлены последовательности минимальных расстояний между образованными кластерами при различных значениях уровня порога отсека. Видно, что пока число кластеров не снижается до двух, минимальное расстояние между кластерами нарастает монотонно, независимо от количества объектов, образованных при отсеке. Когда же элементы объединяются в два локальных подмножества, происходит скачкообразный рост и минимальное расстояние между кластерами принимает максимальное значение.

Полученные результаты говорят о том, что при любом исходном количестве объектов, получаемых после отсека в образце, они группируются в два кластера, что соответствует количеству включений в моделируемое поле. Малые объекты, расположенные на больших расстояниях от значительно больших по площади объектов, как бы "притягиваются" к крупным объектам. Их общий центр "масс" будет находиться внутри крупного объекта или в непосредственной близости от него.

Такие объекты не образуют нового кластера, а площадь вновь образованного кластера будет увеличиваться только в том случае, если присоединяемый объект находится на расстоянии менее половины

максимального линейного размера существующих включений. Это ограничение делает невозможным увеличение площади кластера за счет тех точек, значение в которых превосходит пороговое, но которые не принадлежат к подмножеству сорных примесей.

## ВЫВОДЫ

Поставлена и решена задача оценки площади посторонних пороков и сорных примесей плоских волокнистых материалов с применением методов кластерного анализа на основе средств вычислительной техники.

## ЛИТЕРАТУРА

1. Черкасский А.Е. Неровнота нетканых материалов. – М.: Легпромбытиздат, 1989.
2. Дюран Б., Одел П. Кластерный анализ. – М.: Статистика, 1977.
3. Горячая И.С., Севостьянов П.А. Использование кластерного анализа для оценки неравномерности распределения волокон в поперечном сечении пряжи // Сб. тр. аспирантов. – М.: МГТУ им. А.Н. Косыгина, 2002, №4.

Рекомендована кафедрой информационных технологий и вычислительной техники. Поступила 01.07.05.