

Д.О. МИНЖИЛИЙ, С.В. ЕРЕМЕЕВ

**Исследование работы алгоритмов
топологического анализа данных**

Муромский институт
(филиал) ФГБОУ ВО
«Владимирский
государственный
университет имени
А.Г. и Н.Г. Столетовых»,
г. Муром

В статье проводится анализ изображений с разными дефектами, их классификация и выявления общих признаков с использованием разработанного программного обеспечения.

Исследование выполнено при финансовой поддержке РФФИ и администрации Владимирской области в рамках научного проекта № 17-47-330387.

Введение

Топологический анализ данных на сегодняшний день является одним из эффективных способов анализа и сравнения различных изображений [1].

Уже несколько десятилетий проводятся попытки разработать общий алгоритм, который позволяет воспринимать изображение так, как его видит человек [2]. На данный момент существует множество различных методов, выполняющих узконаправленные задачи и основанные на особенностях самого изображения, его геометрических и топологических характеристик [3,4].

Топологический анализ данных основан на изучении топологических свойств фигур, т.е. таких свойств, которые не изменяются при любых деформациях [5]. Используя данный метод можно проводить анализ расположенных на изображении объектов различной сложности и размерности [6].

В настоящей статье приводится анализ изображений с различного рода дефектами с применением разработанного программного обеспечения, основанного на методе топологического


анализа данных. Настоящее программное обеспечение позволяет выявить дефекты, а из полученных результатов выявить закономерность.

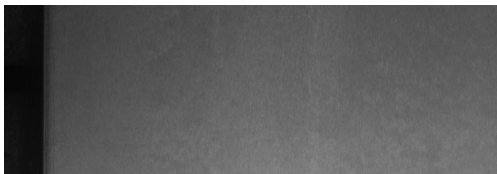
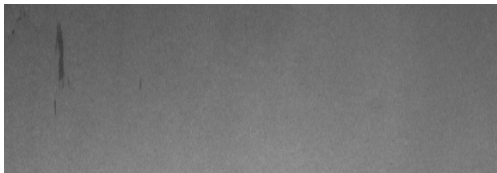
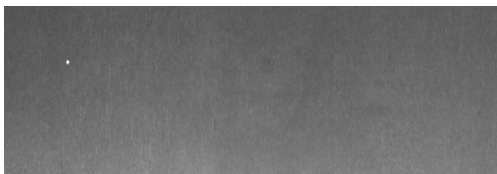
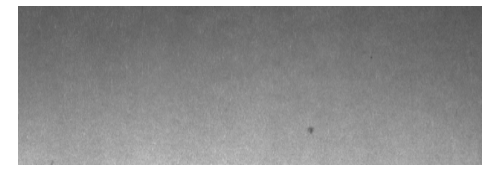

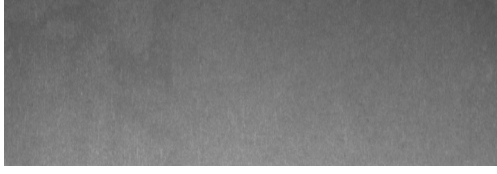

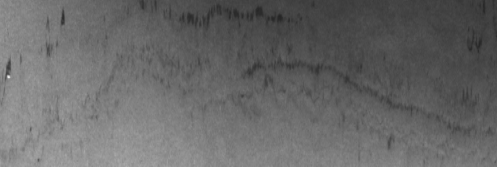
Анализ исходных данных

В ходе анализа использовались различного рода растровые изображения, которые посредством визуального отбора были классифицированы на разные группы. Группы формировались по идентичности изображений относительно друг другу и именовались в соответствии с общим признаком их группировки. Основным фактором отбора являлись дефекты, отображаемые на исследуемых изображениях.

После проведения классификации определены следующие классы, которые продемонстрированы в таблице 1.

Таблица № 1. Классификация анализируемых изображений в зависимости от расположенных на них дефектов

№ п/п	Наименование класса	Визуальный пример изображения
1.	Серый объект на черном фоне с двумя и более пятнами	
2.	Серый объект на черном фоне с темной точкой	
3.	Серый объект на черном фоне со светлой точкой	
4.	Серый объект на черном фоне с темной полосой	

5.	Серый объект на черном фоне	
6.	Серый фон с двумя или более пятнами	
7.	Серый фон со светлой точкой	
8.	Серый фон с темной точкой	
9.	Серый фон с темной полосой	
10.	Серый фон	
11.	Серый фон со светлой полосой	
12.	Изображения с волнообразными темными пятнами	

После проведения классификации анализируемых изображений путем визуального перебора отобрано по 30 наиболее отчетливых примеров каждого класса для дальнейшего анализа.

Анализ изображений при помощи разработанного программного обеспечения

Разработанное программное обеспечение реализует метод персистентной гомологии. Основная суть метода заключается в выявлении таких структур, которые будут устойчиво сохраняться при топологических деформациях и искажениях. Данный метод описан в [7,8].

Используя данное программное обеспечение выявлены закономерности, на каких этапах и итерациях (один полный проход при циклической обработке данных) наиболее характерно выявляются дефекты на изображениях одного конкретного класса. Средние значения результатов исследований представлены в таблице 2.

Таблица № 2. Средние значения результатов анализа дефектов для различного рода изображений

Класс	Кол-во дыр	Кол-во компонент	Кол-во симплексов	Кол-во ребер	Кол-во точек	Кол-во итераций
Серый объект на черном фоне с двумя и более пятнами	2	3	194814	293119	98304	5
Серый объект на черном фоне с темной точкой	1	1	194810	293117	98304	15
Серый объект на черном фоне со светлой точкой	1	1	194810	293117	98304	45
Серый объект на черном фоне	5	1	194814	293117	98304	14
Серый объект на черном фоне с темной полосой	4	1	194810	293117	98304	10
Серый фон с двумя или более пятнами	1	1	194814	293119	98304	20
Серый фон со светлой точкой	2	1	194810	293117	98304	50
Серый фон с темной точкой	1	1	194810	293117	98304	19
Серый фон с темной полосой	5	1	194815	293117	98304	13
Серый фон	2	1	194810	293117	98304	12
Серый фон со светлой полосой	1	1	194810	293117	98304	10
Изображения с волнообразными темными пятнами	2	1	194810	293117	98304	47
Среднее значение	2	1	194811	293117	98304	22

Пример графика кривой, демонстрирующей развития дыр (ось абсцисс) при увеличении количеств итераций (ось ординат) продемонстрирован на рисунке 1.

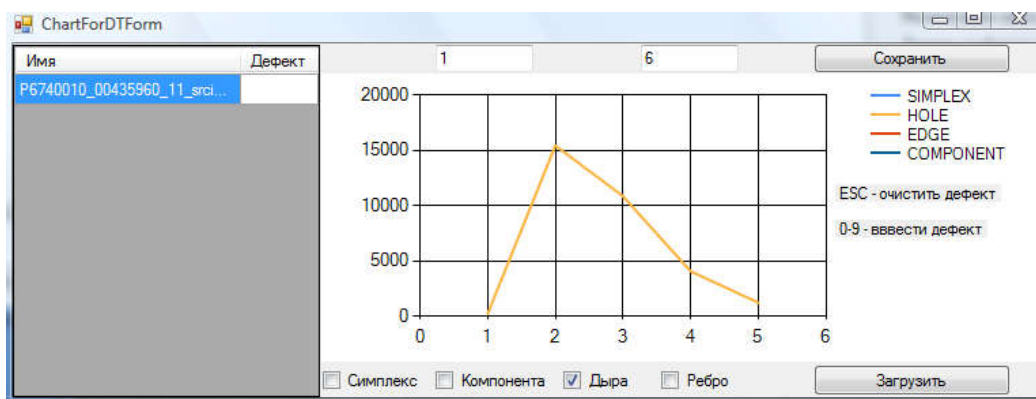


Рис.1. График демонстрации развития дыр

Из полученных данных можно сделать вывод, что при проведении классификации посредством визуального анализа изображения, отнесенные к категориям «Серый фон» и «Серый объект на черном фоне», показали отсутствие каких-либо дефектов. Это в свою очередь показывает, что в процессе первичной классификации допущена погрешность, которая выявлена при использовании программного средства.

По данным, полученным при анализе с использованием программного средства, можно выявить закономерность, что если количество итераций превышает 21, то на изображении находятся пятна или точки белого цвета.

Данное предположение было проверено. Произвольно было выбрано 30 изображений, которые поочередно анализировались при помощи программы. В соответствии с полученными результатами, основываясь только на числовых характеристиках, изображения классифицировались на две группы:

- изображения с темными пятнами;
- изображения со светлыми пятнами.

После программного анализа проведен визуальный осмотр изображений, который показал 100% точность в выборе определенного класса.

Заключение

В статье рассмотрена задача анализа изображений с разного рода дефектами. К таким дефектам относились точки и линии разного типа и цвета, их размер, размещение и количество на изображении. В ходе анализа классифицированы изображения по

общим признакам, в зависимости от цвета и формы дефектов на этих изображениях.

Данная классификация проверена при помощи разработанного программного обеспечения. В ходе анализа полученных данных составлена зависимость цвета дефектов от количества итераций. При помощи данной зависимости, не используя визуального анализа, можно сделать вывод, какого рода, цвета дефекты имеются на данном изображении.

Исследование выполнено при финансовой поддержке РФФИ и администрации Владимирской области в рамках научного проекта № 17-47-330387.

Литература

1. Еремеев С.В., Сельцова Е.А. Алгоритм топологического анализа пространственных структур в геоинформационных системах // Алгоритмы, методы и системы обработки данных. 2017. №1(35) С.16-27.
2. Патин М.В., Коробов Д.В. Сравнительный анализ методов поиска особых точек и дескрипторов при группировке изображений по схожему содержанию // Молодой ученый. 2016. №11. С. 214-221.
3. Robert J. Adler, Omer Bobrowski, Matthew S. Borman, Eliran Subag, Shmuel Weinberger, *Persistent homology for random fields and complexes*, Institute of Mathematical Statistics Collections. С.124–143.
4. Еремеев С. В., Ковалев Ю. А., Андрианов Д. Е. Алгоритм идентификации временных эволюций пространственно-распределенных объектов на основе Баркодов // Геоинформатика. 2018. №4. С. 23-29.
5. Розенберг И.Н. Обработка пространственной информации // Перспективы науки и образования. 2015. №1. С. 17-24.
6. Еремеев С.В., Андрианов Д.Е., Купцов К.В., Ковалев Ю.А. Алгоритм идентификации пространственных объектов растровой карты на основе топологического анализа данных // Телекоммуникации. 2017. № 7. С. 39-44.
7. Eremeev S., Kuptsov K., Romanov S. An Approach to Establishing the Correspondence of Spatial Objects on Heterogeneous Maps Based on Methods of Computational Topology. In: van der Aalst W. et al. (eds) *Analysis of Images, Social Networks and Texts*. AIST 2017. Lecture Notes in Computer Science, vol 10716. Springer, pp. 172–182.
8. A. Zomorodian, G. Carlsson, *Computing persistent homology*, *Discrete Comput. Geom.* P. 249–274.