# Сравнение методов распознавания образов в задачах поиска характерных зерновых микроструктур на фотографиях шлифов металлов

## Постановка задачи

Решаемая задача анализа шлифов металлов и классификации наблюдаемых на них зерновых структур была разбита на несколько достаточно независимых друг от друга задач:

* Автоматическая кластеризация шаблонных фотографий шлифов и автоматизированное определение характеристик, наиболее достоверно определяющих принадлежность шлифа тому или иному классу.
* Классификация поступившего на анализ шлифа с пояснением, почему классификация была проведена таким образом и указанием точности классификации.
* Подсчёт метрических характеристик по анализируемому шлифу.

На данный момент работа была сосредоточена на решение последних двух подзадач, об одной из которых будет рассказано более подробно в этой статье.

В данной работе речь пойдёт о поиске определённых зерновых вкраплений на анализируемой фотографии шлифа. В работе создана система, которая находит на фотографии области с одинаковыми зерновыми структурами (интересующими, прежде всего, исследователя), которые впоследствии могут быть использованы для более точной классификации шлифа. Сделать это необходимо по ряду причин:

* В силу неоднородности распределения зерновых структур, на основании которых было принято решение о принадлежности шлифа конкретному классу, необходимо их визуализировать на фоне шлифа.
* В случае спорной принадлежности шлифа конкретному классу (если на шлифе присутствуют несколько разноплановых структур, ни одна из которых не является преобладающей) эксперт, основываясь на собственном опыте, может отклонить решение системы и указать своё собственное.
* Для зерновых структур разного вида ,возможно, потребуется подсчёт различных численных характеристик.

## Характер входных данных

На вход системе подаётся фотография шлифа в электронном виде. Формат данных не столь важен. Ключевыми характеристиками являются:

* масштаб увеличения металла на фотографии,
* разрешение фотографи, с каким она была отсканирована.

Эти данные необходимы для того, чтобы система адаптировала свои априорные данные под эти показатели. По большей части это необходимо по той причине, что одни и те же структуры, сфотографированные с разным уровнем увеличения, выглядят совершенно по-разному.

В дальнейшем можно попытаться избавиться от необходимости задания этих данных (часть из них может быть взята из самого файла с фотографией, часть из EXIF-метаданных).

Также известно, что фотографии на вход подаются с различного рода шумами и искажениями. Наиболее распространены среди них:

* цифровой шум,
* затемнения некоторых областей фотографии,
* различные диапазоны между светлыми и тёмными пятнами на фотографии.

На данном этапе работы критичным является лишь разница в диапазонах между светлыми и тёмными пятнами. Цифровой шум не приводит к неправильному поведению системы. Затемнение лишь определённы областей на фотографии присутствует лишь на небольшом количестве доступных для обработки фотографий. Поэтому, на данном этапе работы происходит корректировка лишь цветовых диапазонов путём масштабирования их с тем, чтобы гистограмма занимала весь диапазон поддерживаемой глубины цвета (используются 8 бит на канал, фотографии преобразуются в градации серого – наиболее эффективного для обработки алгоритмами распознавания).

## Подходы к поиску закономерностей

Существует несколько подходов для поиска областей с одинаковыми зерновыми структурами:

* Поиск по образцу. База знаний системы содержит шаблоны (паттерны) зерновых структур и использует их для поиска областей, наиболее похожих на один из классов паттернов.
* Определение зёрен, их форм и взаимного расположения и их дальнейшая классификация на основе этих данных.

Первые метод является наиболее простым как с точки зрения постановки, так и с точки зрения реализации. Второй вариант приближен к тому, как классификация производится самим экспертом. Это снижет вероятность ложного срабатывания. Тем не менее, данный метод является более ресурсоёмким и сложным для поддержки (эксперт уже не может так настроить систему для работы в других условиях). Кроме того, как показала практика, первый вариант работает достаточно эффективно как с точки зрения производительности, так и с точки зрения релевантности выдаваемого результата.

Поиск закономерностей разбивается на два больших этапа:

1. Система обучается структурам, которые она может в дальнейшем распознавать.
2. Системе на вход подаётся фотография шлифа, и она ищет на ней закономерности.

На данном этапе работы используется обучение системы «с учителем». Как было описано выше, было решено использовать распознавание структур по их образцу. «Учитель» готовит паттерны структур (небольшие участки фотографий шлифов, на которых отображены наиболее характерные участки микроструктуры) и подаёт их на вход системе. При этом, он также сообщает о том, к какому классу структур относится данный паттерн.

После того как база знаний системы заполнена (в зависимости от используемого алгоритма распознавания это может быть матрица весовых коэффициентов нейросети, или это может быть база с предобработанными фотографиями) система может начинать распознавать структуры, которым она была обучена. Распознавание структур происходит в несколько этапов:

Предобработка изображения

Поиск паттернов

Отображение результата

База  
знаний

Рис.1 Этапы распознавания микроструктур

1. Предобработка фотографий.

Как было описано выше, фотографии на входе системы могут быть зашумлены и на этом шаге происходят следующие операции:

* преобразование фотографии в формат градации серого с глубиной цвета 8 бит (наиболее оптимальный для большинства алгоритмов распознавания),
* корректировка контраста изображения путём масштабирования цветов до получения гистограммы изображения, полностью покрывающей диапазон от 0 до 255.

Такие же действия выполняются с фотографиями, используемыми при обучении системы.

1. На втором шаге алгоритма выделяются все возможные подобласти фотографии шлифа с указанным шагом и размерностью, зависящим от размера паттернов в базе знаний (размер подобластей должен быть равен размеру паттерна из базы знаний для того же самого масштаба). После этого происходит классификация выделенной подобласти. Подобласть маркируется найденным классом.
2. На третьем шаге изображения происходит объединение разрозненных пересекающихся прямоугольных областей, принадлежащих одному классу, в одну N-угольную фигуру которую можно вывести на экран.

## Архитектура системы поиска закономерностей

В системах распознавания трудно предугадывать какой алгоритм является более адекватным для использования в конкретной ситуации, поэтому архитектура системы была разработана с учётом того, что используемый алгоритм распознавания можно было легко поменять прямо по ходу выполнения программы (с целью более простого сравнения результатов их работы).

Максимальная гибкость была достигнута за счёт использования единого интерфейса для алгоритмов распознавания, который поддерживает три метода:

* Загрузка изображения на вход алгоритма.
* Обучение алгоритма загруженному изображению, указывая к какому классу это изображение относится.
* Классификация образа, загруженного на вход алгоритма.

Этот интерфейс был представлен в виде абстрактного класса ImageRecognizer, от которого были наследованы новые классы с реализацией таких алгоритмов, как:

* Нейросеть Хэмминга.
* НейросетьХопвилда.
* SAD (Sum of Absolute Differences – сумма абсолютных разностей).

Более подробная диаграмма классов представлена на следующем рисунке, но именно на сравнении методов распознавания (классов, наследованных от ImageRecognizer)и было сосредоточено внимание.

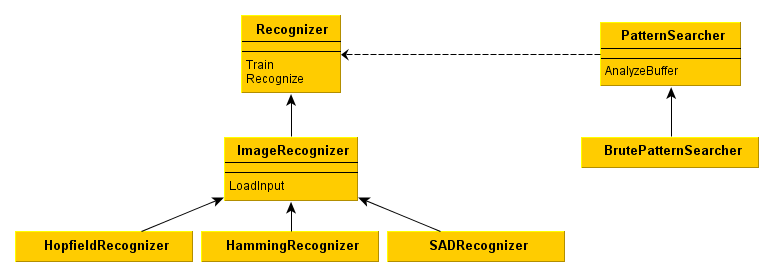


Рис.2 Даграмма классов

На данной диаграмме представлены лишь классы, непосредственно участвующие в процессе распознавания и не отображены вспомогательные классы и классы с реализацией интерфейса пользователя программы.

PatternSercher является абстрактным классом поиска всех «знакомых» паттернов на изображении. BrutePatternSearcher наиболее простая реализация, которая просто перебирает все возможные области изображения и классифицирует их.

## Результаты сравнение алгоритмов

Сравнение алгоритмов производилось по таким критериям как точность и адекватность выдаваемого результата, скорость работы. Под адекватностью предполагается определение характерных структур там, где они действительно есть и определение отсутствия структур, там их где действительно нет.

Протестировать качество работы нейронной сети Хопфилда на реальных данных не удалось, так как он требует O(N4) дополнительной памяти, а так как тестирование производилось на ноутбуке с 4 гигабайтами памяти, матрица весовых коэффициентов не помещалась в памяти.

При сравнении нейронной сети Хэмминга и алгоритма SAD наиболее показательным примером является различие сорбитов и троститов на шлифах. Эти два вида структур являются наиболее сложными для распознавания друг от друга этим алгоритмам, так как оба представляют собой монотонную структуру.

Тем не менее, удалось добиться адекватной работы обоих алгоритмов на таких текстурах, причём SAD на них выдаёт меньший процент неправильно определённых структур.

а) б)

|  |  |
| --- | --- |
|  |  |

Рис. 3 Результаты определения закономерностей в микроструктуре сплава …. а) сеть Хемминга б) SAD

На левом рисунке представлен результат работы нейронной сети Хэмминга, на правом – SAD при распознавании шлифа с сорбитом. Зелёный цвет означает правильно определённый сорбит, красный – остальные виды несорбитных образований, которые были определены (в основном это был тростит).

По изображениям видно, что работа системы является неидеальной, но уже теперь стало видно в какую сторону её следует развивать.