

Абакумов А.В.

Научный руководитель: д.т.н., доц. Д.Е. Андрианов

*Муромский институт (филиал) федерального государственного образовательного учреждения высшего образования «Владимирский государственный университет имени Александра Григорьевича и Николая Григорьевича Столетовых»
602264, г. Муром, Владимирская обл., ул. Орловская, 23*

Использование баркодов для обработки цифровых изображений

Спутниковая инфраструктура позволяет получать данные о каждом уголке земли, поэтому анализ спутниковых геоснимков является важным и нужным процессом. Ввиду специфики снимков, первичный анализ исходного изображения с использованием классических алгоритмов весьма затруднителен. В этом случае может помочь использование алгоритмов персистентной гомологии, особенно алгоритм сравнения объектов с использованием баркодов.

В случае растрового изображения предполагается использование компонент-связностей разного уровня. Для простоты первый уровень можно называть компонентами, а второй - дырами.

Дыры являются более надежной характеристикой, поэтому вместо использования связки компонента-дыра можно использовать связку дыра-дыра. При этом, вторая компонента будет строиться на инвертированном изображении - это позволит зацепить то, что было пропущено при первой итерации.

Ряд тестов выявил, что после определенного порога новые значимые дыры практически перестают появляться, поэтому две характеристики можно объединить в одну, а под порогом понимать цвет фона. В итоге, алгоритм будет работать следующим образом:

- От нулевой яркости до порога начинают строиться дыры.
- По достижению порога, дыры начинают строиться от 255 яркости, при этом время их жизни продолжают записываться на предыдущей шкале, начиная с порога.
- По достижению порога — алгоритм заканчивается, на выходе получается баркод.
- Может производиться отсеивание слишком маленьких линий.

Сравнение же происходит по принципу отношения. Линии жизни дыр обоих баркодов сортируются по длине, а также находится общая сумма всех длин. Затем пары линий из сравниваемых баркодов начинают перебираться одна за другой. Находится площадь их общего пересечения и их общая площадь, а затем находится отношение одного к другому. Соответственно, чем больше линии накладываются друг на друга, т.е. схожи, тем больше будет значение. Затем это значение умножается на отношение суммы длин этих линий на общую сумму всех линий для определения процента вклада этих линий. Результат добавляется к общему счетчику. Соответственно, если линии вообще не пересекаются, то вклад будет нулевым. Результатом будет значение от 0 до 1.

Данный алгоритм не учитывает одну особенность - если взять два одинаковых изображения, повысить на одном яркость, а затем сравнить, то результат схожести будет обратно пропорционален проценту увеличения яркости. Чтобы этого избежать, перед сравнением нужно нормировать время жизни всех линий относительно самой первой линии, т.е. у первой линии время начала всегда будет нулевым, а у остальных оно будет равно разнице между их временем и временем первой.