

БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

Факультет прикладной математики и информатики

Кафедра дискретной математики и алгоритмики

Аннотация к магистерской диссертации

**РАЗРАБОТКА АЛГОРИТМИЧЕСКИХ И ПРОГРАММНЫХ
РЕШЕНИЙ ДЛЯ ВЫСОКОПРОИЗВОДИТЕЛЬНЫХ СИСТЕМ
ОБРАБОТКИ, АНАЛИЗА И РАСПОЗНАВАНИЯ ИЗОБРАЖЕНИЙ НА
БАЗЕ OPENCV И APACHE SPARK.**

Николайчик Александр Евгеньевич

Научный руководитель — кандидат технических наук,
В. А. Ковалёв

2016

Реферат

Магистерская диссертация, 45 страниц, 15 рисунка, 10 источников.

ПОИСК ПО ИЗОБРАЖЕНИЯМ, РЕГИСТРАЦИЯ ИЗОБРАЖЕНИЙ, APACHE SPARK, OPENCV, ШАРД, РАСПРЕДЕЛЁННОСТЬ, ОТКАЗОУСТОЙЧИВОСТЬ, РАССТОЯНИЕ БХАТТАЧАРИА, LBP .

Объект исследования — алгоритмические и программные решения для высокопроизводительных систем обработки, анализа и распознавания изображений на базе OpenCV и Apache Spark.

Цель работы — разработать легко конфигурируемую систему для поиска похожих регионов изображений. Компоненты в этой системе должны быть легко отчуждаемы, дабы допускать их повторное использование для других целей. Оценить качество предоставляемое полученной системой, её плюсы и минусы.

Результат работы — была создана библиотека, позволяющая путём лёгкой конфигурации получить поиск по своей коллекции изображений, обладающий некоторым базовым качеством и гарантиями работоспособности.

В ходе построения системы были рассмотрены различные варианты расчёта признаков по региону изображения и вычисления расстояния между признаками, проведено сравнение качества получаемых на их основе результатов. Особое внимание уделено одному из расстояний, значительно улучшающему полученный результат и причинам стоящим за этим. Также опробованы различные нормализации и выравнивание изображений для улучшения качества.

Полученная система может быть использована для исследования новых данных в условиях их большого объёма либо же для задач связанных с поиском похожих регионов изображений в однородных базах.

Abstract

Master thesis, 45 pages, 15 figures, 10 references.

IMAGE SEARCH, IMAGE REGISTRATION, APACHE SPARK, OPENCV, SHARDING, DISTRIBUTABILITY, FAULT TOLERANCE, BHATTACHARYYA DISTANCE, LBP.

Object of research is algorithmic and program solutions for high-performance systems for processing, analysis and recognition of images basing on OpenCV and Apache Spark.

Purpose is to develop easy-configurable system suited for same image region searching. Such system components should be easy unplugable from system, allowing their later reuse for different purposes. Measure search quality of such system, find out its pros and cons.

Result of current work is library that allows you easily configure search for your image collection. Such search will provide you some base search quality and availability guarantees.

Different approaches to defining image regions features and distance between them where considered and described. Quality of such approaches where measured and reviewed. Special attention is paid to one distance measure, which greatly improve search quality, and reason behind such improvement. Also different variants of images normalization and alignment where described and compared.

Search engine, made during magister work, can be used for studying new big image collections or in tasks related to image region searching in homogeneous image collections.