

МИНИСТЕРСТВО ОБРАЗОВАНИЯ РЕСПУБЛИКИ БЕЛАРУСЬ
БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
Факультет прикладной математики и информатики
Кафедра многопроцессорных систем и сетей

ПЕРЕСТОРОНИН ПАВЕЛ СЕРГЕЕВИЧ
КАЛАШНИКОВ ИГОРЬ АНАТОЛЬЕВИЧ

**Использование распределенных вычислений в оптическом
распознавании символов**

Курсовой проект
студентов 4 курса 1 группы

«Допустить к защите»
с предварительной оценкой
Руководитель работы

Руководитель
Гусейнова Анастасия Сергеевна
ассистент кафедры МСС

«__» _____ 2011 г

Минск 2011

РЕФЕРАТ

Курсовая работа, 12 с., 3 рис., 4 ист.

Ключевые слова: НЕЙРОННЫЕ СЕТИ, РАСПОЗНАВАНИЕ ОБРАЗОВ, РАСПАРАЛЛЕЛИВАНИЕ, КОМПЬЮТЕРНОЕ ЗРЕНИЕ, СЕГМЕНТАЦИЯ ИЗОБРАЖЕНИЙ.

Объект исследования — распознавание текста с использованием распараллеливания в распределенной сети.

Цель работы — разработка программного комплекса для распознавания текста с использованием технологий параллельных вычислений.

Результатами являются изучение и освоение основных технологий и алгоритмов распознавания образов – нейронных сетей, сегментации изображений, распараллеливания алгоритмов в распределенных сетях, создание программного комплекса для распознавания текста.

СОДЕРЖАНИЕ

Реферат	2
Содержание	3
1 Введение.....	4
2 Проблема.....	5
3 Архитектура программного комплекса.....	6
3.1 Общие сведения	6
3.2 Типы узлов вычислительной сети.....	6
3.2.1 Master-узел сети.....	6
3.2.2 Slave-узел сети	7
3.2.3 Gatherer-узел сети	7
3.2.4 Storage-узел сети.....	7
3.2.5 CLI-узел сети.....	7
3.3 Взаимодействие узлов сети.....	7
4 Детали реализации	8
4.1 Общие сведения	8
4.2 Сетевое взаимодействие и распараллеливание.....	8
4.3 Сегментирование изображений.....	8
4.4 Распознавание	9
4.5 Общее описание процесса работы сети	10
5 Заключение	11
6 Список использованных источников	12

1 ВВЕДЕНИЕ

В последнее время в мире очень большое внимание уделяется технологиям распознавания изображений и в частности распознавания текста. Сложно представить сферу человеческой деятельности, где бы ни могли найти применение эти технологии – так как информация сейчас собирается в первую очередь именно в виде фотоснимков и видео. И собранную таким образом информацию часто необходимо приводить к иному виду в целях хранения, демонстрации или поиска.

В настоящее время распознавание обширно используется в целях обеспечения безопасности. Так, распознаванию подвергаются кадры, на которых запечатлены предполагаемые преступники и нарушители. Идет поиск лиц по базам данных, что помогает раскрывать преступления. Также распознаванию подвергаются автомобильные номера, что позволяет отыскивать угнанные автомобили и разыскивать скрывающихся на них людей.

Распознавание текста используется во многих приложениях, целевой аудиторией которых является среднестатистический пользователь. Такие приложения умеют распознавать текст с фотографий, отсканированные документы и текст на рисунках. Эта информация может использоваться для формирования текстовых документов в различных форматах из исходных изображений, может особым образом сохраняться в базу данных для дальнейшего поиска текста. Такие программы и программные комплексы позволяют также осуществлять перевод целых библиотек из книжного формата в формат электронный, что увеличивает доступность литературы для любого человека, имеющего доступ к базе данных оцифрованных книг.

Для распознавания текста сейчас существует огромное число технологий и алгоритмов – это серьезная проблема, затрагивающая области распознавания образом, компьютерного зрения и искусственного интеллекта. Очень многие из алгоритмов основаны на использовании нейронных сетей, которые позволяют обучать себя и постоянно повышать качество распознавания изображений и текста.

В то же время в распознавании текста есть очень много нерешенных проблем и проблем, которые в ближайшее время решить не представляется возможным. Так, человек распознает символы на основе многолетнего опыта, сравнивая картинку с множеством виденных ранее по очень сложным алгоритмам. И все равно часто человек не может понять написанное другим. Компьютер же не обладает таким опытом и такой вычислительной мощностью, как человеческий мозг – потому ошибки появляются достаточно часто.

До сих пор нет точного способа определения слов, символов по определенным изображениям – и такие алгоритмы, вероятно, не существуют из-за сложности и неоднозначности человеческого письма. Без анализа почерка конкретного человека часто невозможно сказать, какой символ он писал.

Проблемы недостатка вычислительных мощностей решаются разными способами – как то использование суперкомпьютеров, создания более эффективных алгоритмов. Достаточно эффективным является использование технологии распределенных сетей, позволяющая одновременно использовать ресурсы многих компьютеров. Это позволяет быстро и качественно получить требуемую вычислительную мощность. Среди недостатков этой технологии – медленная коммуникация между машинами.

2 ПРОБЛЕМА

В качестве направления для исследований нами выбрано оптическое распознавание, которое в настоящее время является очень актуальным. Развитие и распространение компьютерной обработки информации привели к возникновению потребностей в технологиях, позволяющих машинам осуществлять распознавание в обрабатываемой ими информации.

Разработка методов машинного распознавания позволяет расширить круг выполняемых компьютерами задач и сделать машинную переработку информации более интеллектуальной. Несмотря на то, что некоторые из этих задач распознавания решаются человеком на подсознательном уровне с большой скоростью, до настоящего времени ещё не создано компьютерных программ, решающих их в столь же общем виде. Существующие системы предназначены для работы лишь в специальных случаях со строго ограниченной областью применения.

Примерами задач распознавания могут служить:

- распознавание букв;
- распознавание автомобильных номеров;
- распознавание лиц;
- распознавание речи.

Очевидно, что многие из этих задач могут иметь широкое практическое применение. Например, с помощью системы распознавания автомобильных номеров вы без труда автоматизируете въезд на любую территорию, будь то стоянка у торгового центра, территория промышленного предприятия или платная автострада.

Установка системы распознавания на въездах и выездах позволит контролировать присутствие транспортных средств на территории. Как владелец парковки, вы получите полную статистическую информацию о ее загруженности, проанализировав которую сможете принять меры по повышению эффективности вашего бизнеса.

На сегодняшний день в основе наиболее распространенных технологий верификации и идентификации лежит использование паролей и персональных идентификаторов. Однако такие системы слишком уязвимы и могут легко пострадать от подделки, воровства и просто человеческой забывчивости. Поэтому все больший интерес вызывают методы биометрической идентификации, позволяющие определить личность человека по его физиологическим характеристикам путем распознавания по образцам. Классический пример биометрии – анализ отпечатков пальцев, а к новейшим технологиям относятся распознавание сетчатки и радужной оболочки глаза, что применяется и для идентификация личности, и для расследования преступлений – так как имеется определенная уникальность этих частей тела..

3 АРХИТЕКТУРА ПРОГРАММНОГО КОМПЛЕКСА

В работе поставлена цель разработать распределенную систему распознавания текста на различных изображениях. Так, программная система должна уметь получить изображение на вход, найти на нем вхождения текста и распознать его. При этом распределение нагрузки должно заметно увеличить возможное быстродействие обработки изображения.

3.1 Общие сведения

Предполагается, что разрабатываемый программный комплекс будет состоять из пяти различных типов приложений (рисунок 1):

- Master – центральный узел сети.
- Slave – вычислительный узел сети.
- Gatherer – узел, передающий изображения из внешней среды внутрь сети.
- Storage – база данных, сохраняющая результаты обработки изображений.
- CLI (command-line interface) – утилита для администрирования сети.

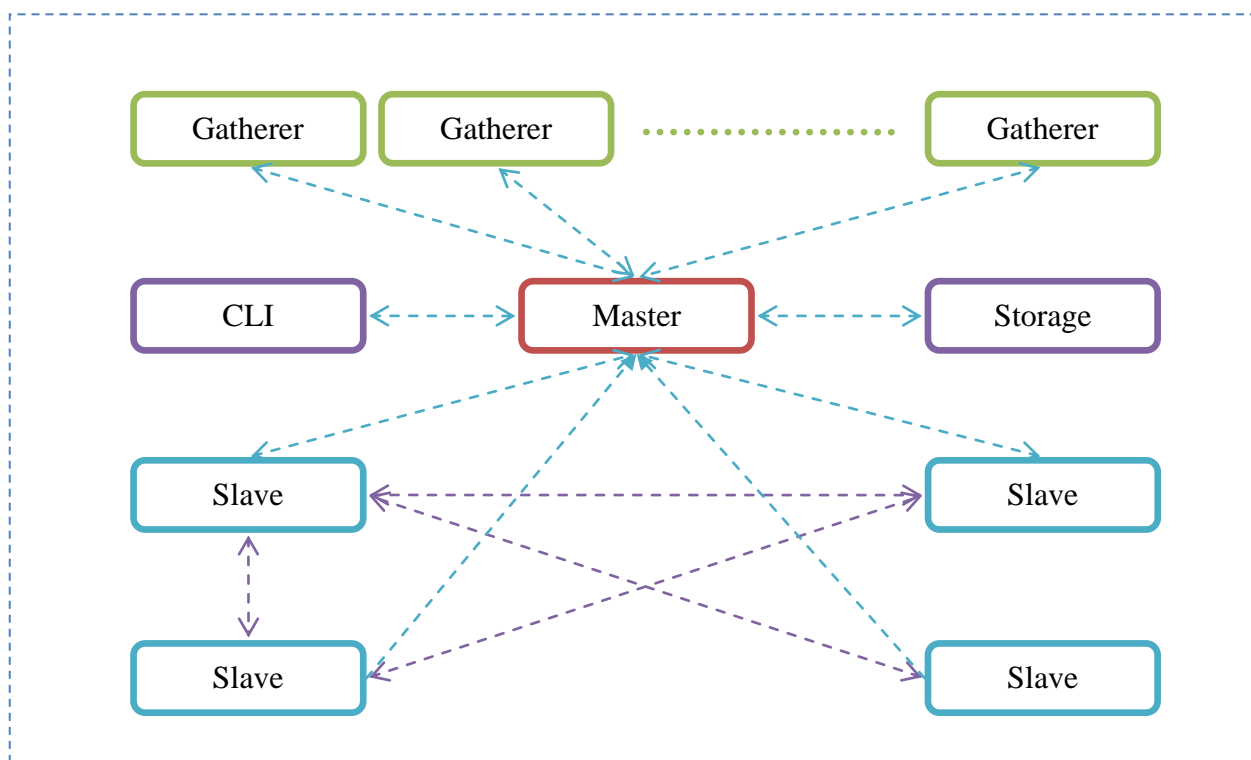


Рисунок 1 – Архитектура приложения

Предполагается, что все узлы сети могут находиться как на одном компьютере, так и на разных, находящихся в пределах одной локальной сети, или разнесенных в разные локальные сети, имеющие доступ к глобальной сети Интернет.

3.2 Типы узлов вычислительной сети

3.2.1 Master-узел сети

Задачами центрального узла сети являются:

- Отслеживание входа и выхода других узлов в и из сети.
- Запуск и останов работы всех вычислительных узлов.
- Поддержка виртуальной топологии сети.
- Обеспечение возможности соединений между узлами сети в режиме «точка-точка» без дальнейшего участия центрального узла для передачи данных.
- Отслеживание ошибок в работе сети.

В одной вычислительной сети должен присутствовать ровно один центральный узел.

3.2.2 Slave-узел сети

Узел сети, выполняющий основную вычислительную работу. Предполагается, что он будет поддерживать соединение с Master-узлом и с каждым из соседних (в заданной виртуальной топологии) Slave-узлов.

Один Slave-узел может содержать несколько рабочих потоков выполнения. Потоки в этом случае являются независимыми, хотя и разделяют общий канал связи. По существу, Slave-узел является «контейнером» потоков, обеспечивающим для потоков возможности коммуникации.

В сети предполагается наличие большого количества Slave-узлов.

3.2.3 Gatherer-узел сети

Вся задача этого узла – пересылать изображения из внешней среды (например, с файловой системы, на которую их сохранили) на обработку на вычислительные узлы.

В сети должен присутствовать хотя бы один такой узел.

3.2.4 Storage-узел сети

Узел должен уметь сохранять информацию о результате обработки изображений в постоянное хранилище, например, в базу данных.

В сети должен присутствовать ровно один такой узел.

3.2.5 CLI-узел сети

Представляет собой узел с запущенной командной строкой для администрирования вычислительной сети, то есть, для выполнения таких операций, как запуск, останов, конфигурирование.

В сети может присутствовать несколько таких узлов, хотя для полноценного управления будет достаточно ровно одного.

3.3 Взаимодействие узлов сети

Узлы сети взаимодействуют по протоколу UDP на сетевом уровне и по протоколу Protocol Buffers, описанном в [1], на уровне приложений. Контроль за доставкой сообщений берут на себя приложения.

Для взаимодействия узлов, находящихся в разных сегментах глобальной сети Интернет предполагается применять технологию UDP Hole Punching, описанная в [2].

4 ДЕТАЛИ РЕАЛИЗАЦИИ

4.1 Общие сведения

Приложения разрабатываются на языке C# на платформе .NET 3.5. В дальнейшем, возможно, потребуется ускорить отдельные участки кода на языке Managed C++.

4.2 Сетевое взаимодействие и распараллеливание

Для распараллеливания работы сети было решено отказаться от готовой библиотеки MPI и разработать свою собственную, что в перспективе может дать дополнительные возможности:

- Поддерживать собственный код значительно легче, чем стороннюю разработку.
- Предполагается, что узлы вычислительной сети могут находиться в разных сегментах сети Интернет. Подобной функциональности в библиотеке MPI найти не удалось.
- Предполагается, что сеть не должна терять работоспособность при отключении одного или нескольких вычислительных узлов. MPI в данном случае останавливает весь процесс.
- Реализация на языке C# позволит отказаться от «оберточных» библиотек для использования MPI и использовать все преимущества платформы.

На текущий момент реализован необходимый базовый функционал для запуска и останова задач в сети, обработки ошибок, назначения порядков потокам в виртуальной топологии, маршрутизации сообщений между потоками и выполнения операций типа «точка-точка», реализована топология «звезда».

4.3 Сегментирование изображений

В компьютерном зрении, сегментация — это процесс разделения цифрового изображения на несколько сегментов (множество пикселей, также называемых суперпикселями). Цель сегментации заключается в упрощении и/или изменении представления изображения, чтобы его было проще и легче анализировать. Сегментация изображений обычно используется для того, чтобы выделить объекты и границы (линии, кривые, и т. д.) на изображениях. Более точно, сегментация изображений — это процесс присвоения таких меток каждому пикселю изображения, что пиксели с одинаковыми метками имеют общие визуальные характеристики.

Первый модуль нашего программного комплекса — модуль выделения на изображении сегментов. Под сегментом мы понимаем часть изображения, на которой изображен предположительно отдельный объект в реальном мире.

Пример сегментации произвольного изображения (фотографии) можно видеть на рисунке 2.

Для сегментации нами используется алгоритм на графах, описанный в [3]. Его преимуществом является то, что он работает относительно быстро по сравнению с

другими алгоритмами сегментации, что позволит в дальнейшем использовать его для обработки непрерывного потока изображений.

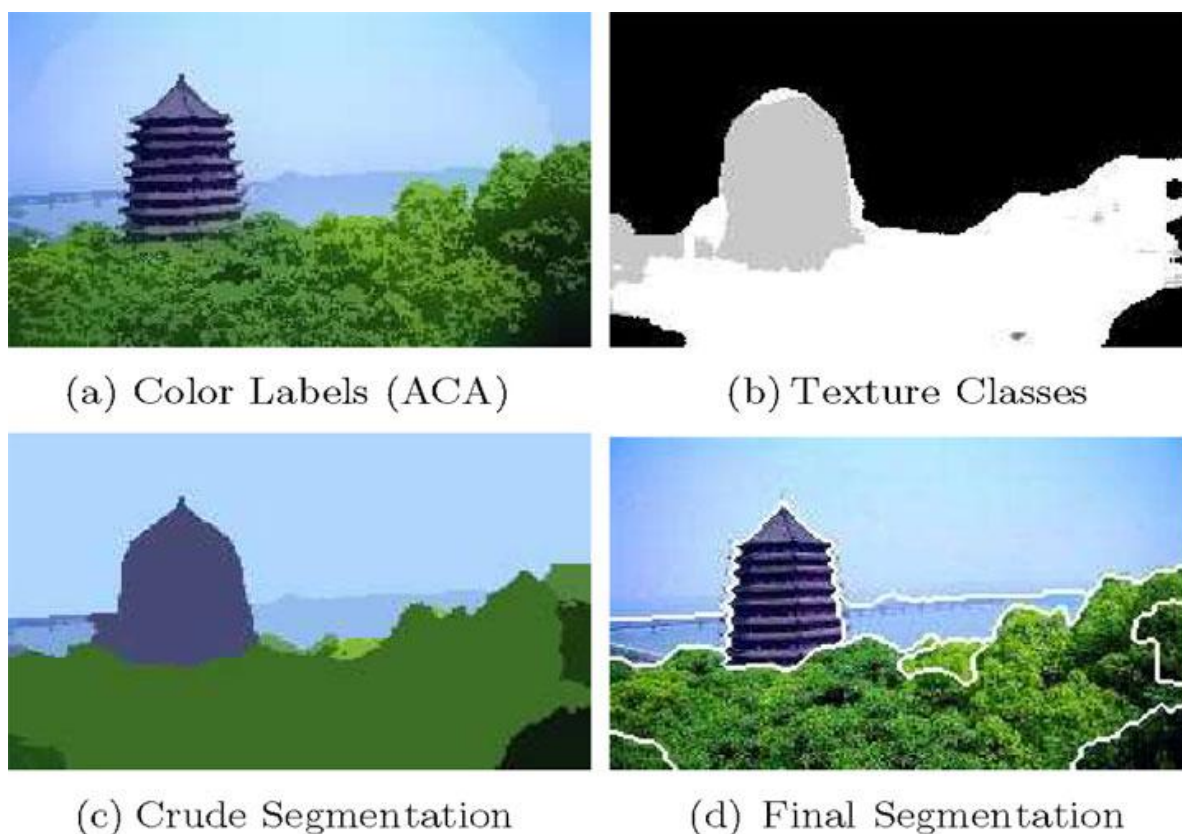


Рисунок 2 – Сегментация фотографии

4.4 Распознавание

На каждой из машин работает модуль распознавания – он получает на вход изображение и выдает строку, распознанную на нем. Для этого используется сверточная нейронная сеть описанная в [4].

Сверточные сети обладают еще одним существенным преимуществом – при распознавании они учитывают топологию входного изображения, извлекая локальные связи между пикселями изображения. Это позволяет увеличить точность распознавания, так как учитывается пространственная организация пикселей входного изображения.

В то же время сверточные сети обладают рядом недостатков, например, малой скоростью обучения. На обучение такой сети для распознавания только цифр уходит до десяти часов, время же обучения для множества символов очень велико. Сейчас рассматривается возможность распараллеливания этого обучения, но основные существующие алгоритмы не подходят для этого.

На рисунке 3 можно увидеть варианты символов для обучения нейронной сети распознаванию символа «3».

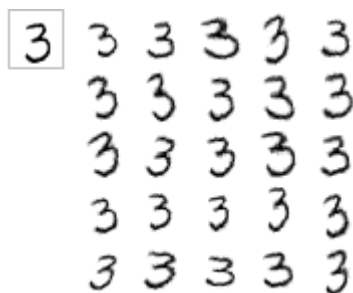


Рисунок 3 – Набор символов для обучения

Сейчас готовы общая структура нейронной сети и алгоритм получения результата распознавания. Над алгоритмом обучения ведется работа. Найдена база символов для обучения нейронной сети.

4.5 Общее описание процесса работы сети

В общем виде предполагается, что вся система будет итеративно выполнять следующие шаги:

1. Gatherer принимает очередное изображение и отправляет его на распараллеленное сегментирование. Сегментированием непосредственно должны заниматься Slave-узлы сети.
2. После сегментирования отдельные сегменты поступают на Slave-узлы для распознавания в них буквенно-цифровых символов.
3. Результаты сегментирования и распознавания поступают на Storage-узел сети и сохраняются.

5 ЗАКЛЮЧЕНИЕ

В данной курсовой работе нами были поставлены и решены следующие проблемы:

- Сформулирована задача построения программного комплекса для распределенных вычислений;
- Разработана архитектура программного комплекса;
- Частично реализована библиотека для осуществления коммуникации между удаленными проектами;
- Реализован алгоритм сегментации изображения;
- Частично реализовано распознавание изображений с помощью нейронной сети.

6 СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Encoding – Protocol Buffers [Electronic resource]. – Google, 2011. – Mode of access : <http://code.google.com/apis/protocolbuffers/docs/encoding.html>. - Date of access : 22.12.2011.
2. Peer-to-Peer Communication Across Network Address Translators [Electronic resource] / ed. Bryan Ford. – Bryan Ford's Home Page, 2005. – Mode of access : <http://www.brynosaurus.com/pub/net/p2pnat/>. – Date of access : 22.12.2011.
3. Эффективная сегментация изображений на графах [Электронный ресурс] / ред. Д.Вихарев. – Habrahabr, 2010. – Режим доступа : <http://habrahabr.ru/blogs/algorithm/81279>. – Дата доступа: 1.02.2011.
4. Neural Network for Recognition of Handwritten Digits [Electronic resource] / ed. Mike O'Neill. – The Code Project, 2006. – Mode of access : <http://www.codeproject.com/KB/library/NeuralNetRecognition.aspx>. – Date of access : 23.01.2011.