

«УТВЕРЖДАЮ»

Проректор Московского государственного
университета имени М.В. Ломоносова, доктор
физико-математических наук, профессор



А.А. Федянин
А.А. Федянин

« » апреля 2017 г.

ОТЗЫВ ВЕДУЩЕЙ ОРГАНИЗАЦИИ

**федерального государственного бюджетного образовательного учреждения
высшего образования «Московский государственный университет
имени М.В. Ломоносова»**

на диссертацию Бабенко Артема Валерьевича

«Эффективные алгоритмы поиска по большим коллекциям изображений»,
представленную на соискание ученой степени кандидата физико-
математических наук по специальности 05.13.11 – Математическое
и программное обеспечение вычислительных машин, комплексов
и компьютерных сетей.

Актуальность темы диссертационной работы

Задача визуального поиска по большим коллекциям изображений крайне востребована на сегодняшний день: современным поисковым системам, таким как Яндекс и Google, необходимо осуществлять поиск среди миллиардов изображений в Интернете. Помимо применения в поисковых системах задача визуального поиска также возникает при распознавании лиц, в диагностике на основе медицинских изображений, в задаче локализации. В настоящее время во всех вышеперечисленных приложениях количество изображений в поисковых коллекциях стремительно увеличивается, поэтому задача эффективного визуального поиска по большим объемам, безусловно, является актуальной.

Структура работы и ее содержание

На отзыв ведущей организации была представлена диссертационная работа, состоящая из введения, трех глав, заключения, списка терминов и списка литературы. Текст диссертации изложен на 137 страницах, список литературы содержит 71 наименование.

Во введении обосновывается актуальность и формулируется цель диссертационного исследования, аргументируется научная новизна исследования и показывается практическая значимость полученных результатов.

В первой главе автор предлагает два идейно новых подхода к формированию векторных описаний изображений для задачи визуального поиска. Предложенные подходы, в отличие от существующих аналогов, используют глубокие нейронные сети. Так как описания строятся с использованием нейросети, обученной в режиме обучения с учителем на большом объеме размеченных данных, качество поиска на основе таких описаний существенно выше аналогов, формирующих описания в режиме обучения без учителя.

Во второй главе автор решает задачу компактного представления поисковой базы из миллиардов векторных описаний. Здесь автор описывает существующие решения этой задачи, а также демонстрирует, что передовой метод мультиквантизации неявно делает ряд предположений о свойствах распределения имеющихся векторных описаний, которые зачастую не выполняются на практике. В связи с этим автор предлагается два способа обобщения метода мультиквантизации, которые избегают подобных предположений, и следовательно являются более универсальными и применимыми к векторным описаниям, сформированными различными способами. Экспериментально продемонстрировано, что предложенные методы сжатия позволяют снизить ошибку сжатия на 20%, что приводит к существенному повышению точности поиска.

В третьей главе автор решает задачу эффективного поиска ближайших соседей среди миллиардов векторов. Автор приводит описание существующих алгоритмов для решения этой задачи и демонстрирует, что

время работы передовых методов составляет сотни миллисекунд, что может быть nepозволительно долгим для многих приложений. Автором предлагается новая структура данных – инвертированный мультииндекс, которая является обобщением известного подхода инвертированной индексации больших коллекций. Автор доказывает, что асимптотика алгоритмической сложности основных операций в мультииндексе логарифмическая, а также демонстрирует высокую скорость работы предложенной структуры данных экспериментально. Показано, что на коллекциях из миллиарда векторов мультииндексу достаточно десяти миллисекунд для достижения высокой точности поиска, что на порядок быстрее скорости работы передовых существующих методов.

В заключении приведены основные результаты работы.

Научная новизна работы

Основная новизна работы состоит в следующем:

1. Разработано два метода построения векторных описаний изображений. В отличие от существующих подходов к построению векторных описаний для задачи визуального поиска, разработанные методы используют предобученные глубокие нейронные сети, что позволяет достичь существенного более высокого качества поиска.

2. Разработано два новых метода сжатия векторов высокой размерности, позволяющие хранить поисковую коллекцию векторных описаний в сжатом виде в условиях ограниченной оперативной памяти, например, на мобильном устройстве. В отличие от существующих подходов к задаче сжатия, разработанные методы не делают никаких априорных предположений о распределении сжимаемых векторов, что делает их более универсальными и позволяет достичь меньшей ошибки компрессии.

3. Разработано две структуры данных для эффективного поиска ближайших соседей среди миллиардов векторов. Предложенные структуры данных позволяют достигать приемлемой точности поиска за несколько

миллисекунд, в то время как существующим аналогам для этого требуются сотни миллисекунд.

Степень обоснованности и достоверности результатов исследования

Все положения и выводы диссертации достоверны и научно обоснованы. Полученные результаты подтверждены большим количеством проведенных вычислительных экспериментов на множестве различных наборов данных. Достоверность результатов работы подтверждается использованием стандартных способов оценки качества визуального поиска и согласованностью полученных результатов с результатами других авторов.

Практическая значимость

На основе разработанных методов можно реализовать систему визуального поиска, способную искать по миллиардам изображений, как на высокопроизводительных серверах, так и на мобильных устройствах с ограниченными ресурсами. Разработанные методы являются универсальными и могут быть использованы в множестве приложений.

Рекомендации по использованию результатов работы

Результаты диссертационной работы представляют большой практический интерес. Они могут быть использованы в различных организациях, разрабатывающих системы визуального поиска. Рекомендуется также использовать их в научной работе и при организации учебного процесса в Московском государственном университете им. М.В. Ломоносова и в других ведущих вузах и научно-исследовательских институтах России.

Замечания

- 1) В работе формулируется цель разработать «эффективные» методы поиска изображений по содержанию, но при этом в постановке и введении не формулируется четко критерии «эффективности» методов, при достижении каких показателей система визуального поиска становится эффективной.
- 2) В первой главе для экспериментальной оценки дескрипторов используются тестовые коллекции, размер которых на несколько

порядков меньше чем размер целевых коллекций, на которых должна работать вся система в итоге. Не обсужден вопрос, насколько оценка методов на таких коллекциях соответствует поставленной задаче, как методы будут масштабироваться на коллекции большего размера.

- 3) Некоторая небрежность в использовании терминов. Например, термин СПОК вводится сразу без расшифровки, хотя из статей соискателя известно, что это перевод с английского аббревиатуры SPoC (Sum Pooling of Channels). В первой части автор пишет «тестовые коллекции», затем этот термин заменяется на термин «датасет» в других частях. На странице 40 пишется про «кропы» изображений. В одном месте говорится от «списке кандидатов», а в других местах о «шорт-листах».
- 4) Нет экспериментальной оценки всех предложенных методов в едином комплексе. В первой части сказано, что наилучшие результаты делаются за счёт СПОК признаков, но собственная эталонная коллекция для оценки скорости поиска сделана на основе полностью связанных признаков сверточной сети, причём была использована архитектура “Inception”, отличная от той, которая рассматривалась в первой части работы.
- 5) Практически ничего не сказано про программную реализацию методов.


Заключение

Указанные замечания не влияют на общую положительную оценку работы. Диссертация Бабенко А.В. выполнена на высоком научном уровне и обладает высокой практической значимостью. Результаты диссертации полно представлены в его публикациях и правильно отражены в автореферате. Стоит отдельно отметить высокий уровень выполненных научных публикаций, апробацию работы на ведущих международных конференциях. Автореферат полностью соответствует диссертации. Считаю, что работа «Эффективные алгоритмы поиска по большим коллекциям изображений» удовлетворяет требованиям, предъявляемым к диссертациям на соискание

ученой степени кандидата наук, а ее автор – Бабенко Артем Валерьевич – заслуживает присуждения искомой степени кандидата физико-математических наук по специальности 05.13.11 – Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей.

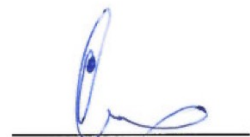
Отзыв заслушан и одобрен 13 апреля 2017 г. на научном семинаре лаборатории компьютерной графики и мультимедиа Московского государственного университета имени М.В. Ломоносова (протокол № 20170413-1)

Отзыв на диссертацию подготовил
зав. лабораторией компьютерной графики и мультимедиа
Московского государственного университета имени М.В. Ломоносова
кандидат физико-математических наук
(05.13.11 – Математическое и программное
обеспечение вычислительных машин,
комплексов и компьютерных сетей),
доцент

 Конушин Антон Сергеевич

24.04.2017

Подпись А.С. Конушина заверяю:
Зам. декана факультета ВМК
МГУ им. М.В. Ломоносова,
доктор физико-математических наук,
профессор

 С.А. Ложкин

Федеральное государственное бюджетное образовательное учреждение высшего образования «Московский государственный университет имени М.В. Ломоносова»,

Адрес: 119991, Российская Федерация, Москва, Ленинские горы, 1,
тел.: +7 (495) 939-10-00, e-mail: info@rector.msu.ru, веб-сайт: www.msu.ru