Анотация

Одной из востребованных областей поиска с Интернете является поиск по изображениям. Современные методы индексации данных показывают отличный результат поиска в многомиллиардных коллекциях и повсеместно применяются для данной задачи. В этой этой работе предлагается изучить существующие подходы индексации больших коллекций изображений и посмотреть на поведение этих алгоритмов в задаче распознавания лиц. Будем считать, что современные нейросетевые алгоритмы достаточно хорошо выделяют признаки лица на изображении и лишь немного затронем этот вопрос. Основной задачей ставим сравнение скорости поиска индексных методов на наборах данных лиц. Начнем рассмотрение с простейших алгоритмов поиска, таких как поиск по точному расстоянию и простая индексная структура. Затем на более продвинутых алгоритмах будем увеличивать скорость поиска и исследовать ухудшение точности. Эксперименты показывают, что в больших коллекция изображений лиц можно добиться приемлемой скорости поиска с хороших показателем точности. Сильное же ускорение приводит к большим потерям.

Введение

Социальные медиа стали широко распространенными в нашем обществе. Одним из популярных аспектов социальных сетей является обмен личными фотографиями. Facebook в официальном документе 2013 года показал, что его пользователи загрузили более 250 миллиардов фотографий и загружают 350 миллионов новых фотографий каждый день1. Для включения автоматической пометки этих изображений необходимы точные и надежные возможности распознавания лиц. На основе загруженной фотографии системы предложения тегов Facebook и Google автоматически обнаруживают лица, а затем предлагают возможные теги имен на основе сходства между шаблонами лиц, сгенерированными из входной фотографии, и ранее помеченными фотографиями в их наборах данных. Что касается правоохранительных органов, то ФБР планирует включить более 50 миллионов фотографий в свой набор данных «Идентификация следующего поколения» (NGI) 2 с целью предоставления следственных указаний путем поиска в галерее изображений, похожих на фотографии подозреваемого. И предложение тега в социальных сетях, и поиск подозреваемого в уголовном расследовании являются примерами масштабного поиска лица (рис. 1). Мы решаем проблему крупномасштабного поиска лиц в контексте социальных сетей и других веб-приложений, где изображения лиц, как правило, не ограничены с точки зрения позы, выражения и освещения [1], [2].

**Введение**

В последнее десятилетие эффективный миллиардный поиск ближайших соседей стал серьезной исследовательской проблемой [1–6], вдохновленной потребностями современных приложений компьютерного зрения, например, крупномасштабный визуальный поиск [7], классификация снимков с малого выстрела [8] и распознавание лиц [9]. В частности, поскольку количество изображений в Интернете растет чрезвычайно быстро, мультимедийным поисковым системам требуются масштабируемые и эффективные алгоритмы поиска, чтобы отвечать на запросы к базам данных миллиардов элементов за несколько миллисекунд. Все существующие миллиардные системы избегают неосуществимого исчерпывающего поиска, ограничивая часть базы данных, которая рассматривается для запроса. Это ограничение выполняется с помощью структуры индексации. Структуры индексации разбивают пространство признаков на большое количество непересекающихся областей, и процесс поиска проверяет только точки из областей, наиболее близких к конкретному запросу. Проверенные пункты организованы в короткие списки кандидатов, а поисковые системы исчерпывающе рассчитывают расстояния между запросом и всеми кандидатами. В сценариях, когда база данных не помещается в ОЗУ, используются сжатые представления точек базы данных. Сжатые представления обычно получают с помощью квантования продукта [10], что позволяет эффективно вычислять расстояния между запросом и сжатыми точками. Этап вычисления расстояний имеет сложность, которая является линейной по количеству кандидатов, поэтому короткие списки, предоставляемые структурами индексации, должны быть краткими. Первая структура индексации, способная работать с наборами данных в масштабе миллиарда, была представлена ​​в [1]. Он был основан на инвертированной структуре индекса, которая разбивает пространство признаков на области Вороного для набора центроидов K-средних, изученных на основе набора данных. Было показано, что эта система достигает разумных скоростей повторного вызова за несколько десятков миллисекунд. Позже обобщение структуры перевернутого индекса было предложено в [2]. В этой работе был представлен инвертированный мультииндекс (IMI), который разбивает пространство признаков на несколько ортогональных подпространств и разбивает каждое подпространство на районы Вороного независимо. Тогда декартово произведение областей в каждом подпространстве образует неявное разбиение всего пространства признаков. Из-за огромного количества регионов космический раздел IMI очень мелкозернистый, и каждый регион содержит только несколько точек данных. Следовательно, IMI формирует точные и краткие списки кандидатов, при этом эффективно используя память и время выполнения. Однако структурированная природа регионов в разделе IMI также оказывает негативное влияние на конечную производительность поиска. В частности, в [5] было показано, что в большинстве регионов IMI нет точек, а эффективное число регионов намного меньше теоретического. Для определенных распределений данных это приводит к тому, что процесс поиска тратит много времени на посещение пустых регионов, в которых нет кандидатов. Фактически, причина этого недостатка состоит в том, что IMI изучает кодовые книги K-Means независимо для различных подпространств, в то время как распределения соответствующих подвекторов данных не являются статистически независимыми на практике. В частности, существуют значительные корреляции между различными подпространствами созданных CNN дескрипторов, которые наиболее актуальны в наши дни. В этой статье мы утверждаем, что предыдущие работы недооценивают простую структуру инвертированного индекса и защищают ее использование для всех типов данных.

Поиск приблизительного ближайшего соседа (ANN) имеет большое значение во многих проблемах компьютерного зрения, таких как поиск изображений / видео [4], классификация изображений [5] и распознавание объектов / сцен [6]. Векторное квантование (VQ) [7] - это популярный и успешный метод поиска по ANN. Метод векторного квантования для ANN используется двумя способами: (i) для построения инвертированного индексирования [4] для неисчерпывающего поиска или (ii) для кодирования векторов в компактные коды [1], [8], [9] для исчерпывающий поиск.

Литература (По номерам статей из файла TextNotes):

[1]

[2]

[3]

[4]

[5]

[6]

[7]

[8]