Изображение - многозначный термин.

Видеоданные - информация представленная визуально(на экране, терминале и т.п.).

Задача классификации изображений - классифицировать конкретное изображение в соответствии с набором возможных категорий(критериев).

Задача детектирования объектов на изображении - определить наличие объекта на изображении и определить его положение.

# Обзор методов детектирования и локализации лица на изображении.

Задача обнаружения лица на изображении часто является первым шагом в процессе решения задачи более высокого уровня — распознавания лица, деталей лица или его мимики.

Татаренков, Д. А. Анализ методов обнаружения лиц на изображении / Д. А. Татаренков. — Текст : непосредственный // Молодой ученый. — 2015. — № 4 (84). — С. 270-276. — URL: https://moluch.ru/archive/84/15524/ (дата обращения: 23.02.2022).

Алгоритмы обнаружения лиц можно разбить на четыре категории:

1) эмпирический метод;

2) метод характерных инвариантных признаков;

3) распознавание с помощью шаблонов, заданных разработчиком;

4) метод обнаружения по внешним признакам.

Эмпирический метод предполагает создание алгоритма, реализующего набор правил, которым должен удовлетворять фрагмент изображения, чтобы его признали человеческим лицом.

Пример набора правил - центральная часть лица имеет однородную яркость и цвет, лицо содержит в себе нос, рот и два симметрично расположенных глаза.

Метод характерных инвариантных признаков не зависит от угла наклона и положения изображения и включает в себя три этапа:

1) детектирование на изображении явных черт лица - глаза, носа, рта;

2) обнаружение границ лица, его формы, текстуры и цвета;

3) объединение всех найденных инвариантных признаков и их верификация.

Обнаружение лица с помощью шаблона состоит в проверке каждой области изображения на соответствие заданному шаблону.

В методе обнаружения по внешним признакам необходимо провести этап обучения системы путем обработки тестовых изображений, далее для изображения вычисляется набор

признаков и проводится классификация изображения на два класса - лицо/не лицо.

# Обзор методов выделения признаков портретных изображений

Not implemented…

# Обзор существующих решений по задаче матчинга натуралистического изображения лица и наброска.

Метод матчинга натуралистического изображения лица и наброска использующий контрольную сумму локального градиента (LGCS) состоит из нескольких этапов:

1) набросок и фотоизображения выравниваются в соответствии с координатами глаз, обрезаются и изменяются в размерах;

2) система преобразовывает эскиз в градиентное изображение, а затем в LGCS изображение

3) LGCS изображение наброска сопоставляется с LGCS изображениями из базы данных и на выход дается результат из четырех лучших совпадений в соответствии с расстоянием Евклида. Это так называемая система распознавания эскизов лиц и фотографий (FSPR).

Метод матчинга основанный на самоподобии лица заключается в сопоставлении небольшого участка поверхности с его локально окрестностью. Такой подход позволяет избежать необходимость преобразования модальности неявно сокращая разрыв между ними.

Дескриптор предлагаемого самоподобия лица демонстрирует высокую точность, а также устойчив к аномалиям, таким как сопоставление набросков с размытыми фотографиями.

Метод матчинга на основе признаков включает в себя получение вектора признаков, который содержит сведения о изображении лица(как на наброске, так и на фотографии) в терминах гистограммы направленных градиентов и объектов матрицы совпадения уровней серого.

Вычисление признаков увеличивает шансы на правильные совпадения.

Метод требует предварительной обработки как наброска так и фотографий, вычисления признаков, а также сопоставления наброска с фотографиями с использованием расстояния Евклида.

Еще один метод матчинга на основе признаков использующий гибрид двух методов матчинга:

1) масштабно-инвариантной трансформации признаков дескриптора (SIFT)

2) "общего представления" - меры сходства наброска и фотографии сравниваются по расстоянию с обучающим набором пар набросок/фото.

Данный метод требует обучающего набора соответствия набросков/фотографий , а также преобразования наброска в фотографию или наоборот.

Этот метод повышает точность распознавания за счет взаимодополняющего характера двух подходов (один метод использует локальные различия между двумя лицами,

в то время как другой учитывает глобальные различия).

# Обзор методов распознавания человеческих лиц.

Методы распознавания человеческих лиц можно поделить на несколько категорий:

1. основанные на признаках;
2. основанные на изображениях.

# Методы основанные на признаках.

Методы основанные на признаках подразделяются на три подраздела:

1. режим активной формы;
2. анализ низкого уровня;
3. анализ признаков.

Системы использующие режим активной формы генерируют маску на основе найденных черт лица таких как нос, рот и уши. Сгенерированные маски можно изменять. Лучший результат можно получить, тренируя систему большим количеством изображений. Данный метод можно разделить на четыре группы – змеи(snakes), модель деформируемого шаблона (DTM), модель деформируемой детали (DPM) и модель распределения точек (PDM).

Змеи(snakes) обычно используются для определения границ головы, они относительно нечувствительны к шуму, так как интегральный оператор, использующийся в них для как внутренних, так и внешних энергетических функций является фильтром шума. Как правило, этот метод способен определять границы объектов, но обладает некоторыми ограничениями, например, контуры часто попадают в ловушку ложных объектов изображения, а также плохо подходят для нахождения невыпуклых объектов. Также, стоит отметить, что данных алгоритм требует много времени ввиду особенности его реализации.

В модели деформируемого шаблона задача стоит не только в обнаружении границ, как в змеях(snakes), но и в обнаружении таких черт лица как глаза, рот, брови, нос и уши. Эта модель работает путем формирования деформируемых форм лица, что достигается заранее заданными формами, которые могут быть как многоугольными шаблонами, так и иерархическими шаблонами. Модель обеспечивает лучшее выделение нужных черт за счет объединения локальной информации с глобальной, но, как и модель змея, требует чрезмерное время обработки.

Модель деформируемой детали использует графическую структуру для распознавания объектов, обычно используется для распознавания лиц, а также для обнаружения лиц в комиксах. Эта модель основана на машинном обучении. Она хорошо справляется с обнаружением лиц различной формы, различным ракурсом и плохим освещением, однако имеет проблемы со скоростью распознавания.

Модель распределения точек описывает форму лица точками, используя для этого машинное обучение. Метод может использоваться и для поиска объемных(3D) данных. Одним из недостатков является сложность построения тренировочного набора путем указания границ лица и черт лица ввиду неизбежно тяжелой работы и возникновения ошибок.

Анализ низкого уровня не зависит от типа детектируемого объекта и от положения камеры, применяется ко всему изображению и может быть разделен на четыре подкатегории:

1. движение;
2. цвет;
3. информация о серых тонах;
4. границы.

Анализ движения обеспечивает надежное и точное отслеживание, работает на уменьшенном пространстве поиска, так как в значительной степени фокусируется на движении, и хорошо себя показывает в среде реального времени. Однако система не всегда способна обнаружить глаза, а лица с бородой могут искажать положительные результаты.  
Обработка цвета выполняется намного быстрее по сравнению с обработкой других черт лица.

Кроме того, цветовая ориентация неизменна при определенных условиях освещения. Тем не менее, информация о цвете чувствительна к изменению яркости, и разные камеры производят значительно различные значения цвета.

Информация о сером представляет собой двухмерную (2D) обработку, а информация о цвете — трехмерную. Следовательно, это менее сложно в вычислительном отношении. Однако обработка серой информации менее эффективна, а отношение сигнал/шум меньше.

Определение границ было одним из первых методов в компьютерном зрении. Границей считается резкое изменение яркости. Для обнаружения границ на изображении используется множество различных фильтров и операторов:

1. оператор Собеля;
2. реберный оператор Марра–Хилдрета;
3. управляемый фильтр.

В системе обнаружения лиц на основе границ лицо может быть обнаружено с использованием минимального количества сканирования и, кроме того, система относительно надежна и экономична. Несмотря на это распознавание лиц по краям не подходит для зашумленных изображений.

Методы поиска признаков используют довольно традиционную технику, заключающуюся в том, что сначала ищется примечательная черта лица, затем другие, менее примечательные черты лица, основываясь на черты, найденные первыми.

Анализ созвездия, который также относится к методам анализа признаков состоит в том, что при анализе созвездия черта лица формируется группа для поиска лица на изображении. Этот метод может обнаруживать лица на изображениях с шумным фоном, чего не могли сделать большинство алгоритмов, рассмотренных ранее.

# Методы основанные на изображениях.

Большинство таких методов работают используют оконное сканирование, окно сканируется попиксельно, чтобы классифицировать лицо/не лицо. Подходы, основанные на изображениях, являются самыми последними методами, появившимися в обнаружение лиц и подразделяются на три основные области: нейронные сети, линейное подпространство методы и статистические подходы.

Нейронные сети можно разделить на искусственные нейронные сети (ИНС), нейронная сеть, основанная на принятии решений (DBNN) и нечеткая нейронная сеть (FNN).

Искусственная нейронная сеть основана на наборе связанных узлов. Факт изучения закономерностей в данных позволяет ей давать лучшие результаты при наличии большого количества данных. Наиболее часто используемые ИНС для распознавания лиц являются такие нейронные сети как:

1. нейронная сеть, связанная с сетчаткой;
2. нейронная сеть с прямой связью;
3. нейронная сеть обратного распространения;
4. нейронная сеть радиальной базисной функции;
5. нейронная сеть, инвариантная к вращению;
6. быстрая нейронная сеть;
7. полиномиальная нейронная сеть;
8. сверточная нейронная сеть.

Нейронные сети основанные на принятии решений используют статический процесс для неподвижных изображений и временную стратегию для видео. На этапе обучения рисунок лица вырезался, чтобы сделать плоскость глаза горизонтальной и создать структуру, в которой расстояние между глазами постоянное. Данный алгоритм очень эффективен с точки зрения производительности и времени вычислений. Основным недостатком является использование небольшого набора данных.

Нечеткие нейронные сети(ННС) называемые также нейро-нечеткой гибридизацией для обучения используют метод обратного распространения ошибки – метод вычисления градиента, который используется при обновлении весов многослойного перцептрона. Когда обучение закончено, по сети запускается оценка, которая определяет, в какой степени данное окно содержит лицо. ННС требует меньше скрытых нейронов и может обрабатывать фоновые шумы. Несмотря на это, система ННС требует лингвистических правил вместо обучения на примерах в качестве априорного знания.

Линейное подпространство - это векторное пространство, которое является подмножеством большего векторного пространства. В терминах обработки изображений подпространством называется меньшая часть кадра. Линейное подпространство делится на четыре группы: собственные грани, вероятностные собственные пространства, фишерфейсы и тензорфейсы.

Подход с собственными гранями не требует знания геометрии и отражательной способности граней. Кроме того, сжатие данных достигается за счет низкоразмерного представления подпространства. Однако подход очень чувствителен к масштабированию изображения. Кроме того, этап обучения очень трудоемок и показывает эффективность только при условии, что классы граней больше по размерности по сравнению с пространствами граней.

Fisherface реализован для поиска лицевое подпространство, которое называется Fisherface. Алгоритм очень полезен, когда изображения лица имеют большие различия в освещении и в выражении лица. Кроме того, частота ошибок при обнаружении лиц с очками стеклом очень мала при сравнению с собственным лицом. Кроме того, метод Fisherfaces требуют меньше времени для вычислений, чем метод собственных граней.

Метод Tensorfaces – это полилинейный подход, в котором тензор является обобщением матрицы в многомерном базисе. Tensorfaces могут быть реализованы как унифицированный фреймворк для решения нескольких задач области компьютерного зрения. Кроме того, сообщалось, что производительность тензорных граней значительно лучше по сравнению с собственными гранями.

Статистические подходы можно разделить на несколько видов:

1. анализ основных компонентов (АОК);
2. метод опорных векторов (МОВ);
3. дискретное косинусное преобразование (ДКП);
4. проекция с сохранением локальности (ПСЛ);
5. анализ независимых компонентов (АНК).

Анализ основных компонентов очень хорошо работает в ограниченной среде, и, как сообщается, он работает быстрее, чем другие статистические подходы, а его усовершенствованная версия имеет высокую скорость распознавания и простоту вычислений.

Метод опорных векторов очень эффективен при работе с многомерными данными. Кроме того, модели этого метода имеют практическое обобщение, таким образом, риск переобучения в МОВ довольно мал. Сообщается, что МОВ также эффективно использует память. Однако МОВ не подходит для большого набора данных и плохо работает с зашумленным набором данных изображений.

В методе дискретного косинусного преобразования положение глаз на изображении необходимо вводить вручную, но это не считается серьезным недостатком алгоритма, так как алгоритм можно использовать с системой локализации. ДКТ демонстрирует значительное улучшение скорости обнаружения из-за нормализации и является менее затратным в вычислительном отношении по сравнению с преобразованием Карунена-Лоева, а также обеспечивает более простой способ борьбы с трехмерными искажениями лица и предоставляет обширную информацию о дескрипторах лица. Тем не менее, для принятия некоторых решений в ДКТ требуется квантование (игнорирование высокочастотных составляющих).

Метод проекций с сохранением локальности является быстрым и подходящим для практических приложений, частота ошибок алгоритма намного меньше по сравнению с методом анализа основных компонентов. Несмотря на это, построение графа в методе ПСЛ чувствительно к шуму и выбросам, что является главным недостатком алгоритма.

Метод анализа независимых компонентов во многих отношениях работает лучше, чем АОК. АНК чувствителен к данным более высокого порядка, в то время как АОК ищет только более высокую дисперсию. Более того, алгоритм ICA является итеративным, но алгоритм, не смотря на все его преимущества, показывает трудности в обработке больших объемов данных.

Вывод: технологии распознавания лиц имеют ряд серьезных проблем, которые снижают время и точность обнаружения лица. Различные алгоритмы борются с этими проблемами по-разному, чтобы повысить точность и скорость обнаружения.