
Обнаружение и локализация лица на изображении

Автор: Вежнев Владимир vvp@graphics.cs.msu.su

Задача обнаружения лица на изображении (face detection) часто является "первым шагом", предобработкой в процессе решения задачи "более высокого уровня" (например узнавания лица, распознавания выражения лица). Однако и сама информация о присутствии и, возможно, количестве лиц на изображении или в видеопотоке может быть полезна для таких приложений, как охранные системы и содержательная индексация базы данных изображений или видеофрагментов.

В контексте построения естественного интерфейса человек-компьютер задача локализации лица находит свое применение в подсистеме отслеживания лица и его характерных черт в видеопотоке для распознавания выражения лица, определения его положения в 3D пространстве для генерации команд компьютеру.

В машинном зрении часто встречаются две модификации задачи обнаружения лица - локализация лица (face localization) и отслеживание перемещения лица (face tracking). Локализация лица является упрощенным вариантом задачи обнаружения, так как опирается на знание о том, что на изображении присутствует одно и только одно лицо. Задачу отслеживания перемещения лица в видеопотоке можно сформулировать как задачу локализации лица на текущем кадре, опираясь на информацию о его положении на предыдущих кадрах.

Задача	Исходные данные	Результат
Обнаружение лица	Изображение	Вынесение решения о наличии (и, возможно, количестве) лиц на изображении, определение их положения
Локализация лица	Изображение (или его фрагмент), содержащее ровно одно лицо	Положение лица на изображении
Отслеживание лица	Текущий кадр видеопотока, положение лица на предыдущих кадрах	Позиция лица в текущем кадре видео

Задача обнаружения лица на изображении является более чем простой для человеческого зрения, однако при попытке построения автоматической системы обнаружения лиц приходится столкнуться со следующими сложностями:

- Сильно варьирующийся внешний вид лица у разных людей;
- Даже относительно небольшое изменение ориентации лица относительно камеры влечет за собой серьезное изменение изображения лица;
- Возможное присутствие индивидуальных особенностей (усы, бороды, очки, морщины и т.д.) существенно осложняет автоматическое распознавание;
- Изменение выражения лица может сильно сказаться на том, как лицо выглядит на изображении;
- Часть лица может быть невидима (закрыта другими предметами) на изображении;
- Условия съемки (освещение, цветовой баланс камеры, искажения изображения, привносимые оптикой системы, качество изображения) в значительной степени влияют на получающееся изображение лица;

Существующие алгоритмы обнаружения лиц можно разбить на две широкие категории. К первой категории относятся методы, отталкивающиеся от опыта человека в распознавании лиц и делающие попытку формализовать и алгоритмизовать этот опыт, построив на его основе автоматическую систему распознавания. Вторая категория опирается на инструментарий распознавания образов, рассматривая задачу обнаружения лица, как частный случай задачи распознавания.

1.1 Методы первой категории - Эмпирическое распознавание

Человеческий мозг справляется с задачей обнаружения лиц на изображениях более чем успешно. Естественно было бы попробовать определить и использовать принципы, которыми руководствуется мозг при решении задачи распознавания. Среди методов, делающих такую попытку, можно выделить два направления: методы распознавания "сверху-вниз" основанные на знаниях и методы распознавания "снизу-вверх" основанные на особенностях.

Распознавание "сверху-вниз" означает построение некоторого набора правил, которым должен отвечать фрагмент изображения, для того чтобы быть признанным человеческим лицом. Этот набор правил является попыткой формализовать эмпирические знания о том, как именно выглядит лицо на изображениях и чем руководствуется человек при принятии решения лицо он видит или нет. Довольно легко построить набор простых и очевидных (как кажется) свойств изображения лица, например: лицо обычно симметрично, черты лица (глаза, носа, рот) отличаются от кожи по яркости (обычно им также соответствуют области резкого изменения яркости), черты лица расположены вполне определенным образом. Опираясь на перечисленные свойства, можно построить алгоритм проверяющий их наличие на фрагменте изображения. К этому же семейству методик можно также отнести распознавание с помощью шаблонов, заданных разработчиком (predefined template matching). Шаблоны задают некий стандартный образ изображения лица, например, путем описания свойств отдельных областей лица и их возможного взаимного расположения. Обнаружение лица с помощью шаблона заключается в проверке каждой из областей изображения на соответствие заданному шаблону.

Принципы шаблонов и другие методы распознавания "сверху-вниз" использовались, в основном, в ранних работах по обнаружению лица [3], [4], [5], [6], [7]. Это были первые попытки формализации признаков изображений лица, к тому же вычислительные мощности компьютеров в те годы не позволяли эффективно использовать более сложные методы распознавания изображений. Несмотря на некоторую наивность алгоритмов, не стоит недооценивать значение этих работ, поскольку многие методики, успешно применяемые в настоящее время, были разработаны или адаптированы к данной конкретной проблеме именно в них.

Распознавание "снизу-вверх" использует инвариантные свойства (invariant features) изображений лиц, опираясь на предположение, что раз человек может без усилий распознать лицо на изображении независимо от его ориентации, условий освещения и индивидуальных особенностей, то должны существовать некоторые признаки присутствия лиц на изображениях, инвариантные относительно условий съемки. Алгоритм работы методов распознавания "снизу-вверх" может быть кратко описан следующим образом:

1. Обнаружение элементов и особенностей (features), которые характерны для изображения лица;
2. Анализ обнаруженных особенностей, вынесение решения о количестве и расположении лиц;

1.1.1 Обнаружение элементов и особенностей (features), которые характерны для изображения лица.

Края (edges) - резкие переходы яркости. Края обычно соответствуют границам объектов на изображении. Используя этот факт и то, что лицо представляет собой эллипс определенных пропорций (близких для разных людей) были сделаны попытки распознавания лица с помощью *карты краев* (изображения, на котором обозначены резкие переходы яркости) и характерной формы лица. Резкие переходы яркости также часто соответствуют *чертам лица* (facial features) - границам глаз, бровей, рта, носа. Это свойство также используется в ряде работ, которые рассматривают края на изображении как признаки потенциального присутствия лица.

Яркость. Области изображения, соответствующие чертам лица, зачастую темнее, чем окружающая их кожа. Воспользовавшись этим наблюдением, ряд исследователей использует алгоритмы обнаружения и подчеркивания областей локальных минимумов яркости, рассматривая их как потенциальные черты лица. В некоторых работах делается попытка использовать определенные схемы взаимоотношений яркостей, характерных для некоторых черт лица.

Цвет- несмотря на то, что яркость обычно является основным источником информации во многих

задачах машинного зрения, цвет (благодаря дополнительной информации об оттенке объекта) является более мощным средством распознавания и различения объектов на изображении. Как показали эксперименты, цвет кожи разных людей занимает достаточно небольшую ограниченную подобласть цветового пространства, даже при рассмотрении цветов кожи различных рас. Причем основное отличие заключается в яркости, а не оттенке цвета, что позволяет сделать вывод о близости оттенка цвета кожи разных людей и использовать характерный цвет кожи как признак для распознавания лиц.

Характерная форма черт лица. Исходя из того, что процессам распознавания визуальных образов высокого уровня в мозгу предшествует некая низкоуровневая организация визуальной информации, было предложено несколько операторов, подчеркивающих области изображения, обладающими свойствами, характерными для черт лица. Такими, например, как симметричность, близость границы черт лица по форме к параболе. Результатом применения таких операторов является набор точек на изображении, с высокой вероятностью относящиеся к чертам лица. Другой близкий вариант распознавания - использование жестких или деформируемых шаблонов для обнаружения черт лица (например, глаз).

Более подробное описание алгоритмов, опирающихся на описанные особенности изображения лица, а также большое количество ссылок на литературу содержится в статьях [2, секция 3], [1, секция 2.2].

1.1.2 Анализ обнаруженных особенностей, вынесение решения о количестве и расположении лиц.

После того, как на изображении выделены области, обладающие свойствами, характерными для человеческого лица, производится их комплексная проверка для выявления областей, действительно являющихся лицами. Сущность этой проверки зависит от характера используемых признаков, а также от избранной исследователями стратегии. Например, если в качестве признаков выступают потенциальные черты лица, обнаруженные с помощью анализа карты краев, то проверкой будет анализ их взаимного расположения с целью определения, могут ли они образовывать человеческое лицо. Если используется также распознавание по цвету, то может быть добавлено дополнительное условие, что рассматриваться в качестве потенциальных лиц будут только области близкие по цвету к оттенку кожи, причем форма областей должна быть эллиптической. Проверка соотношения обнаруженных признаков лица может быть основана на: некотором эмпирическом алгоритме [8], статистике взаимного расположения признаков, собранной по изображениям лиц [10], моделировании процессов, происходящими в человеческом мозгу при распознавании визуальных образов [9], применении жестких или деформируемых шаблонов расположения черт лица [11] и т.д.

1.2 Методы второй категории, моделирование изображения лица

Второе семейство методов подходит проблеме с другой стороны, и, не пытаясь в явном виде формализовать процессы, происходящие в человеческом мозге, стараются выявить закономерности и свойства изображения лица неявно, применяя методы математической статистики и машинного обучения. Методы этой категории опираются на инструментарий распознавания образов, рассматривая задачу обнаружения лица, как частный случай задачи распознавания. Изображению (или его фрагменту) ставится в соответствие некоторым образом вычисленный вектор признаков, который используется для классификации изображений на два класса - лицо/не лицо. Самый распространенный способ получения вектора признаков это использование самого изображения: каждый пиксель становится компонентом вектора, превращая черно-белое изображение $n \times m$ в вектор пространства $R^{n \times m}$. Недостатком такого представления является чрезвычайно высокая размерность пространства признаков. Достоинство заключается том, что используя все изображение целиком вместо вычисленных на его основе характеристик, из всей процедуры построения классификатора (включая выделение устойчивых признаков для распознавания) полностью исключается участие человека, что потенциально снижает вероятность ошибки построения неправильной модели изображения лица вследствие неверных решений и заблуждений разработчика.

Обычно поиск лиц на изображениях с помощью методов, основанных на построении математической модели изображения лица, заключается в полном переборе всех прямоугольных фрагментов изображения всевозможных размеров и проведения проверки каждого из фрагментов на наличие лица. Поскольку схема полного перебора обладает такими безусловными недостатками, как избыточность и

большая вычислительная сложность, авторами применяются различные методы сокращения количества рассматриваемых фрагментов.

1.2.1 Моделирование класса изображений лиц с помощью Метода Главных Компонент (Principal Components Analysis, PCA)

Метод главных компонент [12] применяется для снижения размерности пространства признаков, не приводя к существенной потере информативности тренировочного набора объектов (в данном случае - изображений лиц). Применение метода главных компонент к набору векторов линейного пространства R^n , позволяет перейти к такому базису пространства, что основная дисперсия набора будет направлена вдоль нескольких первых осей базиса, называемых *главными* осями (или главными компонентами). Таким образом, основная изменчивость векторов тренировочного набора представляется несколькими главными компонентами, и появляется возможность, отбросив оставшиеся (менее существенные), перейти к пространству существенно меньшей размерности. Натянутое на полученные таким образом главные оси подпространство размерности $m \ll n$ является оптимальным среди всех пространств размерности m в том смысле, что наилучшим образом (с наименьшей ошибкой) описывает тренировочный набор изображений.

В приложении к задаче обнаружения лиц, МГК обычно применяется следующим образом. После вычисления главных осей тренировочного набора изображений лиц, вектор признаков тестового изображения проецируется на подпространство, образованное главными осями. Вычисляются две величины: расстояние от проекции тестового вектора до среднего вектора тренировочного набора - Distance in Feature Space (DIFS), и расстояние от тестового вектора до его проекции в подпространство главных компонент - Distance From Feature Space (DFFS). Исходя из этих расстояний выносится решение о принадлежности тестового изображения классу изображений лиц [13].

1.2.2 Моделирование класса изображений лиц с помощью Факторного анализа (Factor Analysis, FA)

Факторный анализ (ФА) [15], как и многие методы анализа многомерных данных, опирается на гипотезу о том, что наблюдаемые переменные являются косвенными проявлениями относительно небольшого числа неких скрытых факторов. ФА, таким образом, это совокупность моделей и методов ориентированных на выявление и анализ скрытых (латентных) зависимостей между наблюдаемыми переменными. В контексте задач распознавания, наблюдаемыми переменными обычно являются признаки объектов. Факторный анализ можно рассматривать как обобщение метода главных компонент.

Цель ФА в контексте задачи обнаружения лиц - получить модель изображения лица (с обозримым числом параметров), с помощью которой можно провести оценку близости тестового изображения к изображению лица [16].

1.2.3 Проблема сбора контрпримеров для тренировки классификаторов

Методы, использующие МГК и ФА требуют для тренировки классификатора только набора положительных случаев распознавания (изображений лиц), им не требуются контрпримеры (изображения без лиц). Методы описанные ниже нуждаются также и в контрпримерах, что поднимает еще одну проблему - как найти репрезентативный набор изображений "не-лица" для успешной тренировки классификатора? В работе [17] предложено решение этой проблемы методом самонастройки - он заключается в постепенном формировании набора контрпримеров, по результатам проводимых тестов. На первом шаге для тренировки классификатора используется небольшого тренировочного набора изображений-контрпримеров. Затем производится тестирование на некоторой случайной выборке из базы данных изображений. Все изображения, в ходе теста ошибочно распознанные, как лица, добавляются в набор контрпримеров и тренировка повторяется.

1.2.4 Моделирование распределения векторов лиц с помощью смеси многомерных нормальных распределений (mixture of Gaussians)

МГК и Факторный Анализ являются мощными и удобными способами получения подпространства для

эффективного представления класса объектов во многих случаях, однако они не обязательно являются оптимальными инструментами для моделирования многообразия изображений лиц. Попытка построить модель, состоящую из нескольких кластеров изображений лиц и "не-лиц", смоделировав каждый из них с помощью многомерной нормальной плотности распределения была сделана в [17]. Рассматривая черно-белые изображения размером 19x19 пикселей как вектора в 361-мерном пространстве, был найден набор кластеров, образованных изображениями лиц и "не-лиц". Расстояния до этих кластеров передавались нейронной сети (многослойному перцептрону), выносившему решение о наличии лица на изображении.

1.2.5 Линейный Дискриминантный Анализ (Linear Discriminant Analysis, LDA)

Линейный Дискриминантный Анализ [14], в отличие от МГК и ФА не ставит своей целью найти подпространство меньшей размерности, наилучшим образом описывающее набор тренировочных изображений. Его задача - найти проекцию в пространство, в котором разница между различными классами объектов максимальна. Это требование формулируется как получение максимально компактных кластеров, соответствующих различным классам, удаленных на максимально возможное расстояние. С помощью ЛДА удастся получить подпространство небольшой размерности, в котором кластеры изображений лиц и "не-лиц" пересекаются минимально. Производить классификацию в таком пространстве значительно проще [16].

1.2.6 Метод Опорных Векторов (Support Vector Machines, SVM)

Цель тренировки большинства классификаторов - минимизировать ошибку классификации на тренировочном наборе (называемую *эмпирическим риском*). В отличие от них, с помощью метода опорных векторов [18] можно построить классификатор минимизирующий верхнюю оценку *ожидаемой* ошибки классификации (в том числе и для неизвестных объектов, не входивших в тренировочный набор). Применение метода опорных векторов к задаче обнаружения лица заключается в поиске гиперплоскости в признаковом пространстве, отделяющей класс изображений лиц от изображений "не-лиц".

Возможность линейного разделения столь сложных классов, как изображения лиц и "не-лиц" представляется маловероятной. Однако, классификация с помощью опорных векторов позволяет использовать аппарат *ядерных функций* [19] для неявного проецирования векторов-признаков в пространство потенциально намного более высокой размерности (еще выше, чем пространство изображений!), в котором классы могут оказаться линейно разделимы. Неявное проецирование с помощью ядерных функций не приводит к усложнению вычислений, что позволяет успешно использовать линейный классификатор для линейно неразделимых классов [20].

1.2.7 Искусственные Нейронные Сети (Neural Networks, NN)

Нейросети давно и успешно применяются для решения многих задач распознавания. Для решения задачи обнаружения лица применялось большое количество нейронных сетей различных архитектур [26], в частности: многослойные перцептроны [24], probabilistic decision-based neural networks (PDBNN) [25], и т.д. Достоинством использования нейросетей для решения задачи обнаружения лица является возможность получения классификатора, хорошо моделирующего сложную функцию распределения изображений лиц $p(x | \text{face})$. Недостатком же является необходимость в тщательной и кропотливой настройке нейросети для получения удовлетворительного результата классификации.

1.2.8 Sparse Network of Winnows (SNoW)

Sparse Network of Winnows (SNoW) [22] можно попытаться перевести как "разреженная сеть просеивающих элементов". SNoW для обнаружения лиц [21] представляет собой двухслойную сеть, входной слой которой состоит из узлов, каждый из которых соответствует некоторой характеристике входного изображения (генерирует 1 при наличии некоторой особенности и 0 в случае ее отсутствия на изображении), выходной же состоит всего из двух узлов, каждый из которых соответствует распознаваемым классам изображений ("лицо", "не-лицо"). В качестве характеристик изображения используются флаги равенства определенным величинам среднего значения и дисперсии яркости в каждом из прямоугольных фрагментов изображения размером 1x1, 2x2, 4x4 и 10x10 (все изображения

имеет размер 20x20 пикселей). Это дает пространство признаков размерности 135424. При проведении классификации на входные узлы подается информация о присутствии определенных характеристик в обрабатываемом изображении. Узлы выходного слоя вычисляют линейную комбинацию сигналов, генерируемых входными узлами. Коэффициенты линейной комбинации задаются весами связей между входными и выходными узлами. При превышении заданного порога, принимается решение о наличии лица на изображении.

SNoW специально разработана для случаев классификации, когда потенциальное число характеристик объектов, важных для классификации может быть очень велико, но неизвестно заранее. Разреженная архитектура сети позволяет использовать огромное количество свойств изображения в качестве входных данных, поскольку в процессе тренировки все несущественные характеристики отбрасываются, и не замедляют, в конечном итоге, функционирование классификатора.

1.2.9 Скрытые Марковские Модели (Hidden Markov Models, HMM)

Скрытые Марковские Модели (СММ) [27] являются одним из способов получения математической модели (описания свойств) некоторого наблюдаемого сигнала. СММ относятся к классу *стохастических моделей*. Стохастические модели пытаются охарактеризовать только статистические свойства сигнала, не обладая информацией о его специфических свойствах. В основу стохастических моделей положено допущение о том, что сигнал может быть описан некоторым параметрическим случайным процессом и что параметры этого процесса могут быть достаточно точно оценены некоторым, вполне определенным способом. Настроенную СММ можно рассматривать как источник некоторого случайного сигнала со вполне определенными характеристиками. Для настроенной СММ есть возможность подсчитать вероятность генерации тестового сигнала данной моделью. В приложении к задаче распознавания, представив вектор признаков объекта в виде сигнала (набора последовательных наблюдений), можно смоделировать класс объектов с помощью СММ. Вероятность принадлежности тестового объекта классу, заданному СММ оценивается как вероятностью генерации сигнала, соответствующего его вектору признаков. Настройка (обучение) СММ - состоит в модификации ее параметров для того, чтобы добиться максимальной вероятности генерации сигналов, соответствующих векторам тренировочного набора.

Для применения СММ к задаче обнаружения лиц, нужно определить способ, которым изображения лица преобразуется в сигнал (набор последовательных наблюдений). Изображение лица можно естественным образом разделить на несколько горизонтальных областей: лоб, глаза, рот и подбородок. Лицо может быть представлено в виде сигнала, в котором передаются эти области в определенном порядке (обычно сверху-вниз, слева-направо). Таким образом, изображение лица представляется в виде последовательности наблюдений векторов (каждый из векторов представляет собой горизонтальную полосу пикселей лица), которые во время тренировки и распознавания последовательно передаются случайному процессу, моделируемому СММ [28].

1.2.10 Active Appearance Models (AAM)

С помощью Active Appearance Models (что можно перевести как "активные модели внешнего вида") [29] можно моделировать изображения объектов, подверженных как жесткой (rigid) так и нежесткой (non-rigid) деформации. Жесткая деформация - любая деформация, которая может быть представлена в виде композиции переноса, поворота и масштабирования. ААМ состоит из набора параметров, часть из которых контролируют форму объекта, остальные задают его текстуру. Параметры модели выбираются автоматически, исходя из наиболее характерных деформаций формы и изменений текстуры, присутствующих в тренировочном наборе изображений объекта. Активная модель внешнего вида лица задает изменения формы лица и его характерных черт (формы глаз, бровей, рта, носа, подбородка), а также возможные изменения текстуры лица. Для решения задачи обнаружения лица на изображении, делается попытка найти параметры (расположение, форма и текстура) ААМ, которые задают изображение наиболее близкое к наблюдаемому. Степень близости внешнего вида модели в оптимальной конфигурации к наблюдаемому изображению дает возможность оценить видим мы лицо или нет.

1.3 Достоинства и недостатки методов первой и второй категории

Основа методов первой категории - эмпирика, является одновременно их сильной и слабой стороной. Большая изменчивость объекта распознавания, зависимость вида лица на изображении от условий съемки и освещения позволяют без колебаний отнести обнаружение лица на изображении к задачам высокой сложности. Применение эмпирических правил позволяет построить некоторую модель изображения лица и свести задачу к выполнению некоторого количества относительно простых проверок. Однако, несмотря на безусловно разумную посылку - попытаться использовать и повторить уже успешно функционирующий инструмент распознавания - человеческое зрение, методы первой категории пока далеки по эффективности от своего прообраза, поскольку исследователи, решившие избрать этот путь, сталкиваются с рядом серьезных трудностей. Во-первых, процессы, происходящие в мозгу во время решения задачи распознавания изображений изучены далеко не полностью, и тот набор эмпирических знаний о человеческом лице, которые доступны исследователям на "сознательном уровне", далеко не исчерпывает инструментарий, используемый мозгом "подсознательно". Во-вторых, трудно эффективно перевести неформальный человеческий опыт и знания в набор формальных правил, поскольку чересчур жесткие рамки правил приведут к тому, что в ряде случаев лица не будут обнаружены, и напротив, слишком общие правила приведут к большому количеству случаев ложного обнаружения.

Можно перечислить следующие проблемы, общие для методов второй категории:

1. Зависимость от ориентации и масштаба лица. Большинство классификаторов не являются инвариантными к повороту лица в плоскости изображения и изменению его размера. Поэтому для успешного обнаружения лица, отличного по размеру или ориентации от лиц в тренировочном наборе, требуется дополнительная обработка входного изображения (масштабирование, поворот). Проблему изменения масштаба решают, обычно, путем полного перебора всех возможных прямоугольных фрагментов изображения всех возможных размеров. Попытка же рассматривать еще и все возможные углы поворота лиц в плоскости изображения приведет к тому, что время выполнения и без того долгой процедуры перебора фрагментов превысит все мыслимые пределы. Если говорить о повороте головы вне плоскости изображения, то это является проблемой для многих методов из обеих категорий, поскольку при значительном повороте лицо на изображении изменяется настолько сильно, что многие признаки и правила (заданные разработчиком или полученные неявно) распознавания фронтального изображения лица становятся совершенно непригодными.
2. Неявный способ определения признаков для распознавания лица таит в себе потенциальную опасность: классификатор, обладающий недостаточно репрезентативным набором изображений лиц, теоретически может выделить вторичные или ложные признаки в качестве важных. Одно из следствий - потенциальная зависимость от освещения, которое преобладало в тренировочном наборе. В ряде случаев [26] применяется дополнительная предобработка изображения для компенсации влияния освещения.
3. Высокая вычислительная сложность. Во-первых, сами классификаторы часто включают в себя большое количество достаточно сложных вычислений; во-вторых, полный перебор всех возможных прямоугольных фрагментов изображения сам по себе занимает большое количество времени. Это затрудняет использование некоторых методов в системах реального времени (например - отслеживании перемещения лица в видеопотоке).

Сравнивать между собой качество распознавания методов разных категорий достаточно тяжело, поскольку в большинстве случаев, опираться можно только на данные испытаний, предоставляемые самими авторами, поскольку провести крупномасштабное исследование по реализации большинства известных методов и сравнения их между собой на едином наборе изображений не представляется возможным по причине невообразимой трудоемкости этой задачи.

На основе информации, предоставляемой авторами методов, также сложно провести корректное сравнение, поскольку проверка методов часто производится на разных наборах изображений, с разной формулировкой условий успешного и неуспешного обнаружения. К тому же проверка для многих методов первой категории производилась на значительно меньших наборах изображений.

Заметное различие между первой и второй категорией описанных методов заключается еще и в том, что

эмпирические методы часто довольно просты в реализации (особенно относительно методов второй категории), и предоставляют возможность гибкой настройки под конкретную задачу путем модификации интуитивно понятных параметров. Методы, опирающиеся на инструментарий распознавания образов, требуют значительных усилий по формированию тренировочных наборов изображений и обучению классификатора. Влияние параметров, контролирующих классификатор, на его поведение часто далеко неочевидно. Однако трудоемкость создания работающих прототипов методов второй категории частично компенсируется высокими заявленными показателями качества распознавания на больших коллекциях изображений.

Что касается рекомендаций по выбору метода для решения задачи обнаружения лиц, то можно сказать, что выбирать подходящий метод, исходя из цифровых показателей качества распознавания вряд ли целесообразно. Скорее, все зависит от конкретной задачи и условий в которых должен функционировать разрабатываемый алгоритм. Построение универсального метода, обеспечивающего высокий уровень распознавания при отсутствии ограничений на исходные изображения в настоящее время не представляется возможным, однако для большинства конкретных задач можно создать методы, предоставляющие достаточный уровень распознавания.

В качестве условий, влияющих на выбор метода решения задачи, можно перечислить следующие:

- Предполагаемое разнообразие лиц: ограниченный набор людей, ограничения на возможный тип лица (раса, присутствие растительности на лице, очков и т.д.), отсутствие ограничений;
- Ориентация лиц на изображении: строго вертикальная (или наклон под известным углом), в определенных границах вблизи известного угла наклона, любая;
- Цветное или черно-белое изображение;
- Масштаб лиц, разрешение и качество изображения (зашумленность, степень сжатия);
- Предполагаемое количество лиц, присутствующих на изображении: известно, примерно известно, неизвестно;
- Условия освещения: фиксированные известные, приблизительно известные, любые;
- Фон: фиксированный, контрастный однотонный, слабоконтрастный зашумленный, неизвестный;
- Что важнее - не пропустить ни одного лица или минимизировать количество случаев ложного обнаружения?

Bibliography

[1]

M. H. Yang, D. J. Kriegman, N. Ahuja, "[Detecting faces in images: A survey](#)," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 24, no. 1, pp. 34-58, Jan. 2002.

[2]

E. Hjelm and B.K. Low, "[Face detection: A survey](#)," Journal of Computer Vision and Image Understanding, vol. 83, pp. 236-274, 2001.

[3]

G. Yang and T. S. Huang, "Human Face Detection in Complex Background," Pattern Recognition, vol. 27, no. 1, pp. 53-63, 1994.

[4]

C. Kotropoulos and I. Pitas, "[Rule-Based Face Detection in Frontal Views](#)," Proc. Int'l Conf. Acoustics, Speech and Signal Processing, vol. 4, pp. 2537-2540, 1997.

[5]

T. Sakai, M. Nagao, and S. Fujibayashi, "Line Extraction and Pattern Detection in a Photograph," Pattern Recognition, vol. 1, pp. 233-248, 1969.

[6]

I. Craw, H. Ellis, and J. Lishman, "Automatic Extraction of Face Features," Pattern Recognition Letters,

vol. 5, pp. 183-187, 1987.

[7]

V. Govindaraju, "Locating Human Faces in Photographs," Int'l J. Computer Vision, vol. 19, no. 2, pp. 129-146, 1996.

[8]

K. Sobottka and I. Pitas, "[A novel method for automatic face segmentation, facial feature extraction and tracking](#)," Signal Processing: Image Communication, Vol. 12, No. 3, pp. 263-281, June, 1998.

[9]

F. Smeraldi, O. Carmona, and J. Bigun, "[Saccadic search with Gabor features applied to eye detection and real-time head tracking](#)," Image Vision Comput. 18, pp. 323-329, 2000.

[10]

M. C. Burl and P. Perona, "Recognition of planar object classes," in IEEE Proc. of Int. Conf. on Computer Vision and Pattern Recognition, 6, 1996.

[11]

L. C. De Silva, K. Aizawa, and M. Hatori, "Detection and tracking of facial features by using a facial feature model and deformable circular template," IEICE Trans. Inform. Systems E78-D(9), pp. 1195-1207, 1995.

[12]

"Метод главных компонент," Цифровая библиотека лаборатории компьютерной графики и мультимедиа при факультете ВМиК МГУ, <http://library.graphicon.ru/catalog/19>.

[13]

B. Moghaddam and A. Pentland, "Probabilistic visual learning for object representation," IEEE Trans. Pattern Anal. Mach. Intell. 19(1), pp. 696-710, 1997.

[14]

"Линейный дискриминантный анализ," Цифровая библиотека лаборатории компьютерной графики и мультимедиа при факультете ВМиК МГУ, <http://library.graphicon.ru/catalog/184>.

[15]

"Факторный анализ," Цифровая библиотека лаборатории компьютерной графики и мультимедиа при факультете ВМиК МГУ, <http://library.graphicon.ru/catalog/217>.

[16]

M.-H. Yang, N. Ahuja, and D. Kriegman, "[Face Detection Using Multimodal Density Models](#)," Computer Vision and Image Understanding (CVIU), vol. 84, no. 2, pp. 264-284, 2001.

[17]

K.-K. Sung and T. Poggio, "Example-Based Learning for View- Based Human Face Detection," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 20, no. 1, pp. 39-51, Jan. 1998.

[18]

"Метод опорных векторов," Цифровая библиотека лаборатории компьютерной графики и мультимедиа при факультете ВМиК МГУ, <http://library.graphicon.ru/catalog/217>.

[19]

K.-R. Muller, S. Mika, G. Ratsch, K. Tsuda, and B. Scholkopf. "[An introduction to kernel-based learning algorithms](#)," IEEE Transactions on Neural Networks, 12(2), pp. 181-201, 2001.

[20]

Jochen Maydt and Rainer Lienhart. "[Face Detection with Support Vector Machines and a Very Large Set of Linear Features](#)," IEEE ICME 2002, Lousanne, Switzerland, pp. xx-yy, Aug. 2002

[21]

D. Roth, M.-H. Yang, and N. Ahuja, "[A SNoW-based face detector](#)," in Advances in Neural Information Processing Systems 12 (NIPS 12), MIT Press, Cambridge, MA, pp. 855-861, 2000.

[22]

D. Roth, "[The SNoW Learning Architecture](#)," Technical Report UIUCDCS-R-99-2102, UIUC Computer Science Department, 1999.

[23]

N. Littlestone, "Learning Quickly when Irrelevant Attributes Abound: A New Linear-Threshold Algorithm," Machine Learning, vol. 2, pp. 285-318, 1988.

[24]

P. Juell and R. Marsh, "A hierarchical neural network for human face detection," Pattern Recog. 29, pp. 781-787, 1996.

[25]

S.-H. Lin, S.-Y. Kung, and L.-J. Lin, "Face recognition/detection by probabilistic decision-based neural network," IEEE Trans. Neural Networks 8, pp. 114-132., 1997

[26]

H. A. Rowley, S. Baluja, and T. Kanade, “ [Neural network-based face detection](#),” IEEE Trans. Pattern Anal. Mach. Intell. 20, pp. 23-38., January 1998

[27]

Л.Р. Рабинер, “ [Скрытые марковские модели и их применение в избранных приложениях при распознавании речи: Обзор](#),” Труды ИИЭР, т. 77, Номер 2, февраль 1989.

[28]

S. Marchand-Maillet and B. Merialdo, “ Pseudo two-dimensional hidden markov models for face detection in colour images,” in Proceedings Second International Conference on Audio- and Video-based Biometric Person Authentication (AVBPA), 1999.

[29]

G.J.Edwards, C.J.Taylor, T.F.Cootes, “ [Interpreting Face Images using Active Appearance Models](#),” Int. Conf. on Face and Gesture Recognition 1998. pp. 300-30, 1998.

(c) Graphics & Media lab (webmaster@graphics.cs.msu.su)

При использовании материалов в сети Интернет или бумажной прессе ссылка на сайт (cgm.graphicon.ru) обязательна.