АНАЛЬТЕРНАТИВНЫЕ МЕТОДЫ РЕШЕНИЯ

Метод главных компонент

Метод главных компонент (Principal Component Analysis, {PCA}) применяется для сжатия информации без существенных потерь информативности. Он состоит в линейном ортогональном преобразовании входного вектора размерности в выходной вектор размерности , . При этом компоненты вектора являются некоррелированными и, следовательно, общая дисперсия после преобразования остается неизменной. Матрица состоит из всех примеров изображений обучающего набора. Решив уравнение , получаем матрицу собственных векторов , где - ковариационная матрица для , а - диагональная матрица собственных чисел. Выбрав из подматрицу , соответствующую наибольшим собственным числам, получим, что преобразование , где - нормализованный вектор с нулевым математическим ожиданием, характеризует большую часть общей дисперсии и отражает наиболее существенные изменения . Выбор первых главных компонент разбивает векторное пространство на главное (собственное) пространство , содержащее главные компоненты, и его ортогональное дополнение . В качестве индикаторов принадлежности в методе главных компонент используют:

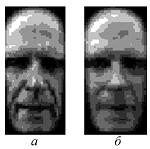
DIFS - distance in feature space, расстояние от образа анализируемого изображения в собственном пространстве, до эталонного образа;

DFFS - distance from feature space, расстояние от представления анализируемого изображения в пространстве наблюдения до проекции эталона в собственном пространстве.



Пример изображений собственных векторов (собственные лица)

Применение для задачи распознавания человека по изображению лица имеет следующий вид. Входные векторы представляют собой отцентрированные и приведенные к единому масштабу изображения лиц. Собственные векторы, вычисленные для всего набора изображений лиц, называются собственными лицами (eigenfaces). Метод главных компонент в применении к изображениям лиц также называют методом собственных лиц (рис. 1). С помощью вычисленных ранее матриц входное изображение разлагается на набор линейных коэффициентов, называемых главными компонентами. Сумма первых главных компонент, умноженных на соответствующие собственные векторы, является аппроксимацией изображения порядка (рис. 2).



Нормализованное изображение лица (*а*) и его реконструкция по 85 главным компонентам (*б*)

Для каждого изображения лица вычисляются его главные компоненты. Обычно берется от 5 до 200 главных компонент. Остальные компоненты кодируют мелкие различия между лицами и шум. Процесс распознавания заключается в сравнении главных компонент неизвестного изображения с компонентами всех остальных изображений. Для этого обычно применяют какую-либо метрику (простейший случай - Евклидово расстояние). Дополнительное повышение надежности достигается за счет дополнительного применения анализа главных компонент к отдельным участкам лица таким, как глаза, нос, рот.

Также метод главных компонент применяется для обнаружения лица на изображении. Для лиц значения компонент в собственном пространстве имеют большие значения, а в дополнении собственного пространства - близки к нулю. По этому факту можно обнаружить, является ли входное изображение лицом. Для этого проверяется величина ошибки реконструкции; чем больше ошибка, тем больше вероятность, что это не лицо. При наличии в наборе изображений лиц вариаций таких, как раса, пол, эмоции, освещение, будут появляться компоненты, величина которых в основном определяется этими факторами. Поэтому по значениям соответствующих главных компонент можно определить, например, расу или пол человека. Основные недостатки PCA таковы. Метод собственных лиц требует для своего применения идеализированных условий таких, как единые параметры освещенности, нейтральное выражение лица, отсутствие помех вроде очков и бород. При несоблюдении этих условий главные компоненты не будут отражать межклассовые вариации. Например, при различных условиях освещенности метод собственных лиц практически неприменим, поскольку первые главные компоненты преимущественно отражают изменения освещения, и сравнение выдает изображения, имеющие похожий уровень освещенности.

Вычисление набора собственных векторов отличается высокой трудоемкостью. Один из способов - это свертка изображений по строкам и столбцам; в такой форме представление изображения имеет на порядок меньший размер, вычисления и распознавание происходит быстрее, но восстановить исходное изображение уже невозможно.

Линейный дискриминантный анализ

Линейный дискриминантный анализ (линейный дискриминант Фишера Linear Discriminant Analysis, {LDA}) использует такую проекцию пространства изображений на пространство признаков, которая минимизирует внутриклассовое и максимизирует межклассовое расстояние в пространстве признаков. В этих методах предполагается, что классы линейно разделимы.

Матрица для проецирования пространства изображения на пространство признаков выбирается из следующего условия:

Где - матрица межклассовой дисперсии, - матрица внутриклассовой дисперсии.

Может существовать до векторов составляющих базис пространства признаков, где - общее число классов. С помощью этих векторов пространство изображений переводится в пространство признаков.

Поскольку работа непосредственно с матрицей затруднительна из-за ее размерности, используется предварительное уменьшение размерности с помощью метода главных компонент, и затем вычисления производятся в пространстве меньшей размерности:

Где - матрица для проецирования в пространство меньшей размерности (пространство главных компонент).

Обычно тренировочный набор содержит изображения лиц при нескольких базовых условиях освещенности, на основе которых при помощи линейных комбинаций можно получить любые другие условия освещенности. Этот метод дает высокую точность распознавания (около 96{\%}) для широкого диапазона условий освещенности, различных выражений лица и наличия или отсутствия очков. Однако остаются невыясненными вопросы, применим ли этот метод для поиска в больших базах данных, может ли метод работать, когда в тренировочной выборке для некоторых лиц имеется изображение только в одних условиях освещенности. Для задачи детектирования лица с помощью LDA-классы лиц и "не лиц" разбивают на подклассы. Вышеописанный метод основывается на предположении о линейной разделимости классов в пространстве изображений. В общем случае такое предположение несправедливо. Инструмент для построения сложных разделяющих поверхностей предлагают нейросетевые методы.

Синтез объектов линейных классов

Данный метод позволяет синтезировать новые изображения объекта (и в частности, изображения лица) для разных ракурсов. Имеется тренировочный набор изображений лиц и только одно изображение нового объекта в определенном ракурсе. Тренировочный набор состоит из изображений объектов того же класса (класс лиц в данном случае), что и новый объект, и включает в себя изображения различных лиц, причем для каждого лица имеются его изображения в широком диапазоне ракурсов. Для нового объекта, имеющего изображение в ракурсе , осуществляется линейное разложение на изображения объектов из тренировочного набора в том же ракурсе, с вычислением коэффициентов , где - количество объектов в тренировочном наборе. Синтез изображения в новом ракурсе для нового объекта осуществляется сложением изображений из тренировочного набора в ракурсе с теми же коэффициентами: . Таким образом, метод позволяет синтезировать изображения нового объекта в различных ракурсах по изображению в одном ракурсе без привлечения сложных трехмерных моделей. Данный метод является перспективным для синтеза изображений в новых ракурсах без привлечения сложных трехмерных моделей, однако вопрос о качестве и количестве примеров в тренировочном наборе остается открытым.

Гибкие контурные модели

В данных методах распознавание производится на основе сравнения контуров лица. Контуры обычно извлекаются для линий головы, ушей, губ, носа, бровей и глаз. Контуры представлены ключевыми позициями, между которыми положение точек, принадлежащих контуру, вычисляются интерполированием. Для локализации контуров в различных методах используется как априорная информация, так и информация, полученная в результате анализа тренировочного набора. Обычно ключевые точки размещаются вручную на наборе тренировочных изображений. При поиске контуров нового лица используется метод симуляции отжига с целевой функцией из двух составляющих. Для первой из них ищется максимум при соответствии интенсивностей пикселов, извлеченных на перпендикулярной контуру линии, аналогичным пикселам из тренировочной выборки, для второй - при совпадении контура с формой контуров тренировочных примеров. Таким образом, извлекается контур черт лица. Для сравнения изображений используются значения главных компонент, вычисленные на наборе векторов, представляющих собой координаты ключевых точек. Главной задачей при распознавании по контурам является правильное выделение этих контуров. В общем виде эта задача по сложности сравнима непосредственно с распознаванием изображений.

Сравнение эластичных графов

В этом методе (Elastic Bunch Graph Matching) лицо представляется в виде графа, вершины которого расположены на ключевых точках лица таких, как контуры головы, губ, носы, и их крайних точках. Каждая грань помечена расстояниями между ее вершинами. В каждой такой точке вычисляются коэффициенты разложения по функциям Габора для пяти различных частот и восьми ориентаций. Набор таких коэффициентов называется *джетом* (jet). Джеты характеризуют локальные области изображений и служат для двух целей: во-первых, для нахождения точек соответствия в заданной области на двух различных изображениях; во-вторых - для сравнения двух соответствующих областей различных изображений. Каждый коэффициент для точек из одной области различных изображений характеризуется амплитудой , которая медленно меняется с изменением положения точки, и фазой , которая вращается со скоростью, пропорциональной частоте волнового вектора базисной функции. Поэтому в простейшем случае для поиска на новом изображении точки с аналогичными характеристиками в функции подобия фазу не учитывают:

Функция подобия с одним джетом в фиксированной позиции и другим с переменной позицией является достаточно гладкой, для того чтобы получить быструю и надежную сходимость при поиске с применением простейших методов таких, как диффузия или градиентный спуск. Более совершенные функции подобия привлекают информацию о фазе. Для различных ракурсов соответствующие ключевые точки отмечаются вручную на тренировочном наборе. Кроме того, чтобы для одного и того же лица представить различные вариации его изображения в одном и том же графе, для каждой точки используются несколько джетов, каждый из которых может соответствовать различным локальным характеристикам данной точки, например, открытому и закрытому глазу. Процесс распознавания неизвестного лица состоит в сравнении графа изображения лица со всеми остальными графами из набора при помощи функции подобия

Левая сумма характеризует подобие джетов, вычисленное с применением фазочувствительной функции, правая - топографическое соответствие, которое пропорционально квадрату разности расстояний между соответствующими вершинами сравниваемых изображений, - количество вершин, - количество граней, - коэффициент относительной важности топографической информации.

В представленном выше виде метод способен достаточно надежно распознавать при изменениях ракурса до 20 ; при больших углах точность распознавания резко уменьшается, функция подобия оказывается более чувствительной к ракурсу, чем к межклассовым различиям. Дальнейшее развитие метода заключается в извлечении коэффициентов важности на основе анализа обучающей выборки. Для каждого джета симплекс-методом вычисляется коэффициент важности, который затем используется в функции подобия. Коэффициенты важности вычисляются из условия максимизации функции подобия для одного и того же лица и минимизации - для различных лиц. Существуют также более ранние разновидности этого метода, которые не используют изначально определенные ключевые точки и структуры графа. Одни из них используют для сравнения решетки джетов, наложенные на изображение, рис. 3. В неизвестном изображении отыскиваются точки соответствия, и затем по найденным точкам строится искаженная решетка и измеряется мера ее искажения для определения наиболее похожего изображения. В других методах точки извлечения джетов изначально образуют решетку, а затем наименее пригодные для распознавания точки отсеиваются в процессе обучения.



Наложенная на изображение эластичная решетка и ее искаженная версия

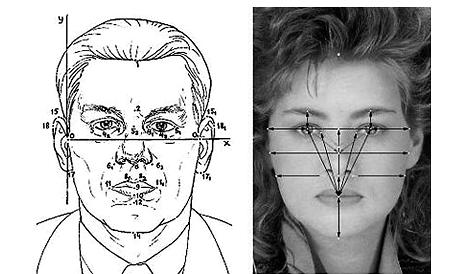
Методы, основанные на геометрических характеристиках лица

Один из самых первых методов - это анализ геометрических характеристик лица. Изначально он применялся в криминалистике и был там детально разработан. Потом появились компьютерные реализации этого метода. Суть его заключается в выделении набора ключевых точек (или областей) лица и последующем выделении набора признаков. Каждый признак является либо расстоянием между ключевыми точками, либо отношением таких расстояний. В отличие от метода сравнения эластичных графов, здесь расстояния выбираются не как дуги графов. Наборы наиболее информативных признаков выделяются экспериментально.

Ключевыми точками могут быть уголки глаз, губ, кончик носа, центр глаза и т. п. рис. 4. В качестве ключевых областей могут служить прямоугольные области, включающие в себя: глаза, нос, рот.

В процессе распознавания сравниваются признаки неизвестного лица с признаками, хранящимися в базе. Задача нахождения ключевых точек приближается к трудоемкости непосредственно распознавания, и правильное нахождение ключевых точек на изображении во многом определяет успех распознавания. Поэтому изображение лица человека должно быть без помех, мешающих процессу поиска ключевых точек. К таким помехам относят очки, бороды, украшения, элементы прически и макияжа. Освещение желательно равномерное и одинаковое для всех изображений. Кроме того, изображение лица должно иметь фронтальный ракурс, возможно, с небольшими отклонениями. Выражение лица должно быть нейтральным. Это связано с тем, что в большинстве методов нет модели учета таких изменений.

Таким образом, данный метод предъявляет достаточно строгие требования к условиям съемки и нуждается в надежном механизме нахождения ключевых точек для общего случая. Кроме того, требуется применение более совершенных методов классификации или построения модели изменений. В общем случае этот метод не является самым оптимальным, однако для некоторых специфических задач все же перспективен. К таким задачам можно отнести документный контроль, когда требуется сравнить изображение лица, полученного в текущий момент, с фотографией в документе. При этом других изображений этого человека не имеется, и, следовательно, механизмы классификации, основанные на анализе тренировочного набора, недоступны.



Идентификационные точки и расстояния: *а* - используемые при криминалистической экспертизе; *б* - наиболее часто применяемые при построении автоматизированных систем идентификации

Сравнение шаблонов.

Сравнение шаблонов (Template Matching) заключается в выделении областей лица на изображении рис. 5, и последующем сравнении этих областей для двух различных изображений. Каждая совпавшая область увеличивает меру сходства изображений. Это также один из исторически первых методов распознавания человека по изображению лица. Для сравнения областей используются простейшие алгоритмы вроде попиксельного сравнения.

Недостаток этого метода заключается в том, что он требует много ресурсов как для хранения участков, так и для их сравнения. Ввиду того, что используется простейший алгоритм сравнения, изображения должны быть сняты в строго установленных условиях: не допускается заметных изменений ракурса, освещения, эмоционального выражения и пр.



Области, входящие в шаблон лица

Скрытые Марковские модели.

Марковские модели являются мощным средством моделирования различных процессов и распознавания образов. По своей природе Марковские модели позволяют учитывать непосредственно пространственно-временные характеристики сигналов, и поэтому получили широкое применение в распознавании речи, а в последнее время - изображений (в частности, изображений лиц). Каждая модель , представляет собой набор состояний , между которыми возможны переходы. В каждый момент времени система находится в строго определенном состоянии. В наиболее распространенных Марковских моделях *первого* *порядка* первого порядка полагается, что следующее состояние зависит только от текущего состояния. При переходе в каждое состояние генерируется наблюдаемый символ, который соответствует физическому сигналу с выхода моделируемой системы. Набор символов для каждого состояния , количество символов . Выход, генерируемый моделью, может быть так же непрерывным. Существуют так же модели, в которых набор символов для всех состояний одинаков. Символ в состоянии в момент времени генерируется с вероятностью . Набор всех таких вероятностей составляет матрицу .

Матрица определяет вероятность перехода из одного состояния в другое: . Считается, что не зависит от времени. Если из каждого состояния можно достичь любого другого за один переход, то все , и модель называется эргодической. Модель имеет вероятность начальных состояний , где . Обычно в реальных процессах последовательность состояний является скрытой от наблюдения и остается неизвестной, а известен только выход системы, последовательность наблюдаемых символов , где каждое наблюдение - символ из , и - число наблюдений в последовательности. Поэтому такие модели называют *скрытыми* Марковскими моделями (Hidden Markov Models, ).

Модель с настроенными параметрами может быть использована для генерирования последовательности наблюдений. Для этого случайно, в соответствии с начальными вероятностями выбирается начальное состояние, затем на каждом шаге вероятность используется для генерации наблюдаемого символа, а вероятность - для выбора следующего состояния. Вероятность генерирования моделью последовательности состояний :

Где - последовательность состояний. Предполагается, что наблюдения статистически независимы.

В распознавании образов скрытые Марковские модели применяются так. Каждому классу соответствует своя модель . Распознаваемый образ (речевой сигнал, изображение и т. д.) представляется в виде последовательности наблюдений . Затем для каждой модели вычисляется вероятность того, что эта последовательность могла быть сгенерирована именно этой моделью. Модель , получившая наибольшую вероятность, считается наиболее подходящей, и образ относят к классу .

В связи с этим появляются несколько вопросов, называемых тремя основными задачами скрытых Марковских моделей.

1. Имея последовательность наблюдений

и настроенную модель , как оценить вероятность генерации этой моделью данной последовательности наблюдений? Эта задача называется задачей распознавания.

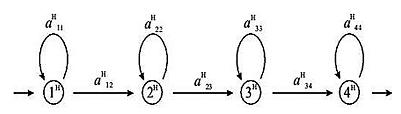
1. Имея последовательность наблюдений

и настроенную модель , как подобрать последовательность состояний , чтобы она была оптимальной (в соответствии с некоторым критерием, аналитически эта задача неразрешима)? Другими словами, это задача объяснения. Она нужна для последующей коррекции параметров модели.

1. Каким образом корректировать параметры модели

, для того чтобы максимизировать ? То есть как сделать так, чтобы модель больше соответствовала своему классу, одним из образов которого является данная последовательность наблюдений (или несколько различных последовательностей)? Это задача обучения.

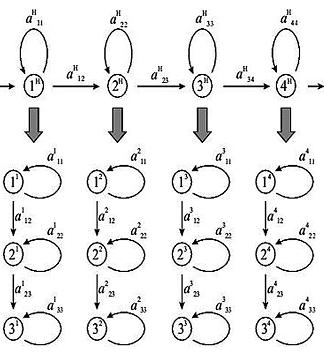
Первая задача имеет точное аналитическое решение, называемое процедурой прямого-обратного прохода. Последующие две задачи не имеют точного аналитического решения. Для решения второй задачи используется алгоритм Витерби, для третей - алгоритм Баума-Вельча. Оба этих метода являются разновидностями градиентного спуска и решаются оптимизационными методами.



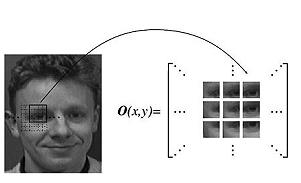
Линейная Марковская модель

Для того чтобы сократить вычисления, в распознавании речи используются линейные модели рис. 6. В таких моделях каждое состояние имеет только одно последующее, так же переход возможен обратно в то же состояние. Такие модели учитывают временные характеристики речевого сигнала: определенный порядок следования участков сигнала, их взаимное расположение, возможность локальных растяжений или сжатий. Это позволяет их применять и в распознавании изображений.

Суть двумерных Марковских моделей заключается в том, что, в отличие от одномерных линейных СММ, они позволяют моделировать искажения изображения и взаимное расположение участков не отдельно по горизонтали или вертикали, а в обоих направлениях одновременно. Для уменьшения вычислительной сложности применяются псевдодвумерные СММ (Pseudo-2D Hidden Markov Models, P2D-HMM). Такая модель состоит из нескольких линейных вертикальных моделей нижнего уровня и одной линейной горизонтальной модели верхнего уровня, на вход которой поступают выходы моделей нижнего уровня, рис. 7. Каждое



Псевдодвумерная скрытая Марковская модель



Извлечение участков-образцов наблюдения

состояние модели верхнего уровня включает в себя последовательность состояний соответствующей модели нижнего уровня. Модели нижнего уровня не связаны между собой. Изначально в модели верхнего уровня были вертикальными. В последующих работах модели верхнего уровня были сделаны горизонтальными (как это и изображено на рисунке), для того чтобы вертикальные модели нижнего уровня могли учесть тот факт, что глаза могут находиться на разной высоте. Таким образом, псевдодвумерная модель позволяет учесть локальные деформации и взаимное расположение участков изображений. Но в отличие от оптических потоков и других методов сопоставления деформациями, псевдодвумерная модель учитывает характер деформаций, а то, какими именно могут быть возможные деформации, псевдодвумерные СММ усваивают в процессе обучения. Другими словами, участок, соответствующий глазу, никогда не будет сопоставлен, например, участку на месте рта.

Пример функционирования СММ. Входом СММ являются квадратные участки изображений (рис. 8). Было обнаружено, что участки, извлекаемые с 75{\%} перекрытием друг с другом, дают наилучшую точность распознавания.

Для СММ важное значение имеет начальная инициализация модели. В качестве начальной инициализации всех моделей используются все изображения из тренировочного набора. Затем модель каждого класса настраивается на свои изображения.

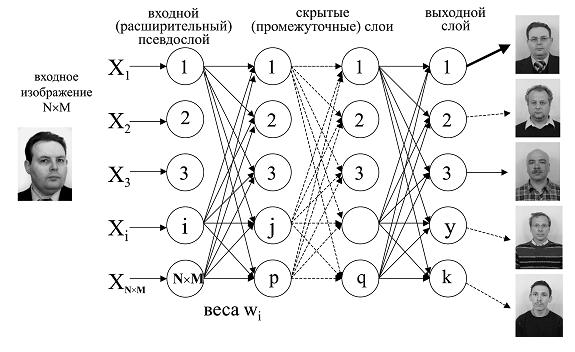
Недостатком СММ является то, что СММ не обладает различающей способностью, т. е. алгоритм обучения только максимизирует отклик каждой модели на свои классы, но не минимизирует отклик на другие классы и не выделяются ключевые признаки, отличающие один класс от другого. Таким образом, похожие классы могут оказаться слабо различимыми и при увеличении объема базы или использования в более широких условиях СММ может оказаться ненадежными.

Многослойные нейронные сети.

Архитектура многослойной нейронной сети (МНС) состоит из последовательно соединенных слоев, где нейрон каждого слоя своими входами связан со всеми нейронами предыдущего слоя, а выходами - следующего. НС с двумя решающими слоями может с любой точностью аппроксимировать любую многомерную функцию. НС с одним решающим слоем способна формировать линейные разделяющие поверхности, что сильно сужает круг задач, ими решаемых, в частности, такая сеть не сможет решить задачу типа "исключающее или". НС с нелинейной функцией активации и двумя решающими слоями позволяет формировать любые выпуклые области в пространстве решений, а с тремя решающими слоями - области любой сложности, в том числе и невыпуклой. При этом МНС не теряет своей обобщающей способности. Обучаются МНС при помощи алгоритма обратного распространения ошибки, являющегося методом градиентного спуска в пространстве весов с целью минимизации суммарной ошибки сети. При этом ошибки (точнее, величины коррекции весов) распространяются в обратном направлении от входов к выходам, сквозь веса, соединяющие нейроны. Простейшее применение однослойной НС (называемой автоассоциативной памятью) заключается в обучении сети восстанавливать подаваемые изображения. Подавая на вход тестовое изображение и вычисляя качество реконструированного изображения, можно оценить, насколько сеть распознала входное изображение. Положительные свойства этого метода заключаются в том, что сеть может восстанавливать искаженные и зашумленные изображения, но для более серьезных целей он не подходит. МНС также используется для непосредственной классификации изображений - на вход подается или само изображение в каком-либо виде, или набор ранее извлеченных ключевых характеристик изображения, на выходе нейрон с максимальной активностью указывает принадлежность к распознанному классу (рис. 9).

Нейрон с максимальной активностью (здесь первый) указывает принадлежность к распознанному классу. Если эта активность ниже некоторого порога, то считается, что поданный образ не относится ни к одному из известных классов. Процесс обучения устанавливает соответствие подаваемых на вход образов с принадлежностью к определенному классу. Это называется "обучением с учителем". В применении к распознаванию человека по изображению лица такой подход хорош для задач контроля доступа небольшой группы лиц. Он обеспечивает непосредственное сравнение сетью самих образов, но с увеличением числа классов время обучения и работы сети возрастает экспоненциально, и поэтому для таких задач, как поиск похожего человека в большой базе данных, требует извлечения компактного набора ключевых характеристик, на основе которых можно производить поиск.

В частности, МНС может использоваться для классификации изображений лиц на основе таких характеристик, как расстояния между некоторыми специфическими частями лица (нос, рот, глаза). Также существуют гибридные системы, например объединение с марковской моделью. В классической МНС межслойные нейронные соединения полносвязны, и изображение представлено в виде одномерного вектора, хотя оно двумерно. Архитектура сверточной НС направлена на преодоление этих недостатков. В ней используются локальные рецепторные поля (обеспечивают локальную двумерную связность нейронов), общие веса (обеспечивают детектирование некоторых черт в любом месте изображения) и иерархическая организация с пространственными подвыборками (spatial subsampling). Сверточная НС (СНС) обеспечивает частичную устойчивость к изменениям масштаба, смещениям, поворотам, искажениям. Архитектура СНС состоит из многих слоев, каждый из которых имеет несколько плоскостей, причем нейроны следующего слоя связаны только с небольшим числом нейронов предыдущего слоя из окрестности локальной области (как в зрительной коре человека). Веса в каждой точке одной плоскости одинаковы (сверточные слоя). За сверточным слоем следует слой, уменьшающий его размерность путем локального усреднения. Затем опять сверточный слой, и так далее. Таким образом, достигается иерархическая организация. Более поздние слои извлекают более общие характеристики, меньше зависящие от искажений изображения. Обучается СНС стандартным методом обратного распространения ошибки. Сравнение МНС и СНС показало существенные преимущества последней как по скорости, так и по надежности классификации. Полезным свойством СНС является и то, что характеристики, формируемые на выходах верхних слоев иерархии, могут быть применимы для классификации по методу ближайшего соседа (например, вычисляя евклидово расстояние), причем СНС может успешно извлекать такие характеристики и для образов, отсутствующих в обучающем наборе. Для СНС характерны высокая скорость обучения и работы. Тестирование СНС на базе данных ORL, содержащей изображения лиц с небольшими изменениями освещения, масштаба, пространственных поворотов, положения и различными эмоциями, показало приблизительно 98{\%} точность распознавания, причем для известных лиц предъявлялись варианты их изображений, отсутствующие в обучающем наборе. Такой результат делает эту архитектуру перспективной для дальнейших разработок в области распознавания изображений пространственных объектов. МНС применяются и для обнаружения объектов определенного типа. Помимо этого, любая обученная МНС в некоторой мере может определять принадлежность образов к "своим" классам, ее можно специально обучить надежному детектированию определенных классов. В этом случае выходными классами будут классы, принадлежащие и не принадлежащие к заданному типу образов.



Многослойная нейронная сеть для классификации изображений

Сети Габоровых вейвлетов (GWN).

Данный метод предназначен для слежения за лицом в реальном времени с помощью Gabor wavelet template (GWT), представляющего собой дискретную линейную комбинацию Габоровых вэйвлетов. Важно отметить, что точность представления лица регулируется числом Габоровых вэйвлетов в GWN, допуская представления множества лиц одним GWT. Такое представление допускает произвольные аффинные преобразования и быструю оценку аффинных параметров методом градиентного спуска. Тем самым, выполняя слежение за лицом, метод определяет и его ориентацию, однако с ограничением на повороты: допускаются, в основном, повороты в плоскости и незначительные пространственные вокруг вертикальной оси.

Нейронные сети Хопфилда.

НС Хопфилда (НСХ) является однослойной и полносвязной (связи нейронов на самих себя отсутствуют), ее выходы связаны со входами. В отличие от МНС, НСХ является релаксационной - т. е. будучи установленной в некое начальное положение функционирует до тех пор, пока не достигнет стабильного состояния, которое и будет являться ее выходным значением. НСХ применяются в качестве ассоциативной памяти и для решения оптимизационных задач. В первом случае НСХ обучается без учителя (например, по правилу Хебба), во втором случае веса между нейронами изначально кодируют решаемую задачу. НСХ бывают синхронными, когда одновременно пересчитываются все нейроны, и асинхронными, когда пересчитывается случайно выбранный нейрон. Для исследования динамики функционирования НСХ используются методы Ляпунова. Асинхронная НСХ всегда сходится к устойчивым точкам, а аттракторами синхронной НСХ являются устойчивые стационарные точки и предельные циклы длины два. Таким образом, НСХ из начального состояния сходится к ближайшему локальному минимуму энергии сети, состояние нейронов в котором и будет восстановленным образом для задач распознавания, и решением - для оптимизационных задач. Для поиска глобального минимума применительно к оптимизационным задачам используют стохастические модификации НСХ.

Применение НСХ в качестве ассоциативной памяти позволяет точно восстанавливать образы, которым сеть обучена, при подаче на вход искаженного образа. При этом сеть "вспомнит" наиболее близкий (в смысле локального минимума энергии) образ, и распознает его. Такое функционирование также можно представить как последовательное применение автоассоциативной памяти. В отличие от автоассоциативной памяти, НСХ идеально точно восстановит образ.

Хотя разработка систем идентификации по изображениям человеческих лиц ведется уже несколько десятилетий, задача создания эффективного алгоритма идентификации по изображениям человеческих лиц еще далека от завершения.

Для большинства современных систем автоматического распознавания лиц основной задачей является задача сравнения данного изображения лица с набором изображений лиц из базы данных. Характеристики систем автоматического распознавания лиц в этом случае оцениваются путем определения вероятностей ошибочного отказа в распознавании (для изображения лица, присутствующего в базе, принимается решение как о неопознанном лице) и ошибочного распознавания. В дополнение к вероятностям ошибок для оценки системы автоматического распознавания лиц часто используется оценка устойчивости к возмущению изображений, вызываемая комбинацией со сложными фонами, изменчивостью освещения, изменению прически, и т. д.

Выбор алгоритма или группы алгоритмов идентификации лиц для создания практических систем автоматического распознавания лиц должен основываться на системе оценок рабочих характеристик конкретной системы, соответствующей ограничениям, связанным с условиями эксплуатации данной конкретной системы, и определяется непосредственно на этапе проектирования.