

МОДЕЛЬ ОПРЕДЕЛЕНИЯ КЛЮЧЕВЫХ ТОЧЕК ЛИЦА НА ОСНОВЕ ПРЕДСКАЗАНИЯ ГЛАВНЫХ КОМПОНЕНТ МОДЕЛИ PDM ПРИ ПОМОЩИ СВЕРТОЧНОЙ НЕЙРОННОЙ СЕТИ

Леонтьева П. А. (Университет ИТМО)

Научный руководитель – **Галюк В. Е.** (ООО «ЦРТ»)

Построена модель определения ключевых точек лица. В основе работы лежит предсказание главных компонент для модели распределения точек (PDM). В отличие от существующих методов, предлагается для определения значений главных компонент использовать сверточную нейронную сеть. Полученная в результате модель, позволяет учитывать связи между ключевыми точками лица и избегать нежелательных отклонений/выбросов. Сделаны выводы относительно возможности дальнейших исследований.

В настоящее время сопоставление изображений является одной из важнейших задач в сфере компьютерного зрения, и считается передовой областью развития лицевых биометрических систем. Существует несколько методов для сравнения изображений, в основе которых лежит сопоставление знаний об изображении. При непосредственно прямом сопоставлении выполняются некоторые преобразования, позволяющие привести большинство пикселей картинки к одинаковым значениям. Однако данный метод не инвариантен относительно точки съемки, света, масштаба и т. д.

Для решения этой проблемы существуют методы, основанные на поиске ключевых точек на изображении. В ходе анализа вычисляются характеризующие объект значения. Это могут быть координаты некоторых значимых для конкретной задачи объектов: расположение центров глаз, формы рта, суставов скелета. В дальнейшем полученные ключевые точки могут применяться в зависимости от задачи: для получения из изображения конкретной области для анализа (Lip reading, Eye tracking, etc..) или для извлечения характеризующих признаков (SURF, SIFT).

Однако существующие детекторы сталкиваются с рядом проблем, вызванных недостаточной устойчивостью к искажению изображения: вращение, смещение, положение камеры, размер изображения, яркость. В дополнение к этому, не учитывается связь между ключевыми точками, что может приводить к нежелательным выбросам и к невозможным в реальной жизни положениям ключевых точек относительно друг друга.

Так, одной из актуальных проблем нахождения ключевых точек является их устойчивость. Многие базы идентификации/верификации содержат фото сотни различных людей, которые применительно к задаче являются классами (Celeb500k, MSCeleb1M, etc.), необходимыми для определения. В таких условиях необходимо максимально облегчить работу с изображениями для нейронной сети, в частности привести все изображения к примерно одинаковым условиям. Именно для этого используют ключевые точки.

Как ясно из вышесказанного, в случае если ключевые точки определяются некорректно, то полученное после выравнивания по ним изображение будет отличаться от остальных, что приводит к увеличению ошибки распознавания. Это ставит исследователей перед необходимостью поиска актуального для практического применения решения особенно в рамках задачи распознавания лиц на изображении.

В данной работе проблему устойчивости ключевых точек предлагается решить, при помощи модели распределения точек (Point Distribution Model). Модель распределения точек представляет собой модель, описывающую изменения формы объекта на основе некоторого входного набора данных. Имея обучающий набор данных при помощи PDM модели мы можем представить любой семпл из набора при помощи среднего семпла и добавки в виде вектора главных компонент с заданными коэффициентами. Поскольку в данной модели учитывается взаимосвязь точек из входного набора данных, то тем самым обеспечивается их устойчивость к выбросам и корректность положения относительно друг друга.

Построение PDM модели тесно связано с методом главных компонент (PCA). С математической точки зрения PCA представляет собой ортогональное линейное преобразование, которое отображает данные из исходного пространства признаков в новое пространство меньшей размерности. При этом первая ось новой системы координат строится таким образом, чтобы дисперсия данных вдоль неё была бы максимальной. Вторая ось строится ортогонально первой так, чтобы дисперсия данных вдоль неё, была бы максимальной из оставшихся возможных и т. д. Соответственно первая главная компонента отвечает за характеристику с наибольшей изменчивости входного набора данных, вторая менее и так далее по убыванию. В результате оптимальным для представления данных с помощью PDM модели было выбрано количество главных компонент равное 10.

В существующих методах (Active shape models, active appearance model) для получения ключевых точек объекта применяется предобработка изображения для выделения границ объектов, и затем итерационными методами находят необходимые коэффициенты главных компонент. Основное отличие рассматриваемого в данной работе подхода заключается в том, что для предсказания коэффициентов используется свёрточная нейронная сеть. В результате исчезает необходимость предварительно применять алгоритмы выделения границ и итерационные подходы. Ввиду доступных на данный момент мобильных архитектур нейронных сетей это позволяет существенно повысить производительность методов.

Предложенная модель позволяет учитывать взаимосвязь между ключевыми точками и обеспечивает большую робастность точек к выбросам и некорректным смещениям. При применении с мобильной архитектурой свёрточной сети позволяет также обеспечить хороший уровень производительности, несмотря на необходимость большого разнообразия обучающих семплов для построения PDM модели. Данная модель может применяться для получения ключевых точек лица для задач распознавания лиц.