

ИССЛЕДОВАНИЕ АЛГОРИТМОВ РАСПОЗНАВАНИЯ ОБРАЗОВ НА ОСНОВЕ ГЕОМЕТРИИ ТОЧЕК И СКРЫТЫХ МАРКОВСКИХ МОДЕЛЕЙ

Хлопенкова А.Ю.¹, Рыбкин С.В.¹

¹*Калужский филиал федерального государственного бюджетного образовательного учреждения высшего образования «Московский государственный технический университет имени Н.Э. Баумана (национальный исследовательский университет)», Калуга, e-mail: fn1-kf@mail.ru;*

В статье рассматриваются основные алгоритмы для автоматического распознавание лиц. Анализируются этапы реализации каждого алгоритма. Перечисляются основные преимущества и недостатки системы распознавания образов по сравнению с другими биометрическими системами идентификации человека. Излагается основная концепция процесса поиска и распознавания. Выявляется общая иерархическая структура реализации алгоритмов. Описывается структура поиска узловых точек для дальнейшего построения геометрической инвариантности с применением триплетов. Перечисляются этапы получение абсолютных инвариантов через относительные посредством аффинного преобразования. Определяются основные зоны скопления опорных точек. Описывается процесс построения 3D-моделей с использование двумерных проекций. Демонстрируется пример вычисления расстояния через метрический тензор Римана, называемого геодезическим расстоянием. Указываются преимущества в процессе идентификации по сравнению с более простыми 2D-моделями. Рассматривается топология построения линейного алгоритма на основе скрытых Марковских моделей и нейронных сетей. Приводится алгоритм инициализации скрытой Марковских модели. Описываются различные сферы применения системы распознавания лиц. Прогнозируется дальнейшее развитие и применение системы в различных областях, таких как медицина, таможенный контроль и госбезопасность. Рассматриваются варианты неправомерного использования технологии в рамках защиты персональных данных.

Ключевые слова: распознавание лица, узловые точки, 3D распознавание, скрытые Марковские модели, нейронные сети

INVESTIGATION OF ALGORITHMS OF FACIAL RECOGNITION BASED ON THE POINT GEOMETRY AND HIDDEN MARKOV MODELS

Khlopenkova A.U.¹, Rybkin S.V.¹

¹*Bauman Moscow State Technical University (Kaluga Branch) Kaluga, e-mail:fn1-kf@mail.ru*

The article considers the basic algorithms for automatic facial recognition. The stages of implementation of each algorithm are analyzed. The main advantages and disadvantages of the facial recognition system are listed in comparison with other biometric systems of human identification. The basic concept of the tracing and recognition process is presented. The general hierarchical structure of the algorithm implementation is revealed. The structure of the searching landmarks for the further construction of geometric invariance using triplets is described. Listed are the stages of obtaining absolute invariants via relative ones through affine transformation. The main zones of clustering landmarks are determined. The process of building 3D models using two-dimensional projections is described. An example of calculating the distance through the Riemann metric tensor, called the geodesic distance, is shown. The advantages in the process of identification person are indicated in comparison with the simpler 2D models. The topology of building a linear algorithm based on hidden Markov models and neural networks is considered. The algorithm of initialization the hidden Markov model is given. The various scope of application of the face recognition system are described. Further development and application of the system is predicted in various areas, such as medicine, customs control, and state security. The options for the abuse of technology within the personal data protection are considered.

Keywords: facial recognition, landmarks, 3D recognition, hidden Markov models, neural networks

Введение. Человек имеет врожденную способность различать лица, и лишь в недавнее время компьютеры смогли достичь того же результата. В середине 1960-х годов ученые

начали работу по использованию компьютера для распознавания человеческих лиц. С тех пор программное обеспечение распознавания лиц прошло долгий путь [1].

Технология распознавания лиц является наименее интрузивной и наиболее быстрой биометрической технологией. Она работает с наиболее очевидным индивидуальным идентификатором - человеческим лицом. Используя человеческое лицо как ключ к безопасности, биометрическая технология распознавания лиц получила значительное внимание в течение последних нескольких лет из-за ее потенциала для широкого спектра приложений как в правоохранительных, так и других органах [1]. В настоящее время распознавание лиц уже используется во многих правоохранительных органах в качестве перспективного средства предотвращения террористических преступлений. Программное обеспечение также разработано для автоматизации идентификации банковских служащих, используя распознавание лиц в качестве проверки пользователей.

Под системой распознавания лиц подразумевают технологию, позволяющую идентифицировать человека с цифрового изображения или видеоисточника. Существует несколько различных методов для распознавания, но в целом они работают путем сравнения выбранных черт лица с заданным изображением лиц из базы данных [2]. Так же метод может быть основан на биометрическом искусственном интеллекте, который может однозначно идентифицировать человека, анализируя шаблоны, основанные на лицевых текстурах и форме лица [3]. По сравнению с другими биометрическими системами, использующими отпечатки пальцев / ладони или радужную оболочку, распознавание лица имеет определенные преимущества из-за его бесконтактного процесса. Системы распознавания лиц ненавязчиво сканируют лица людей, как они входят в определенную область, поэтому в большинстве случаев субъекты вообще не знают о процессе распознавания.

Алгоритмы распознавания. Все алгоритмы распознавания можно разделить на два основных подхода: геометрический, заключающийся в фотометрическом сравнении, когда изображение представляется в виде значений и сравнивается со значениями из шаблонов для устранения дисперсий; линейный, основанный на скрытой Марковской модели и динамическом сопоставлении ссылок нейронов [4].

Среди большого разнообразия алгоритмов, выделяется общая структура процесса распознавания лиц:

Рис. 1 Общий алгоритм распознавания лица

Когда система обнаруживает лицо, она определяет положение, размер и позу головы. Используя глубину и ось измерения, на которые не влияет освещение, трехмерное распознавание лица может даже использоваться в темноте и имеет возможность распознавать объект под разными углами зрения с возможностью распознавания до 90 градусов (лицо в профиле). В случае работы с 2D, голова должна быть повернута не менее чем на 35 градусов к камере. На этапе непосредственного распознавания система сравнивает вычисленные признаки с заложенными в базу данных эталонами. Поэтому в качестве основного отличия среди алгоритмов будет являться вычисление признаков и сравнение их совокупностей между собой [3, 5].

Алгоритм распознавания по геометрии точек. Каждое лицо имеет многочисленные, отличимые ориентиры, которые составляют черты лица. Такие ориентиры часто называют узловыми точками. Человеческое лицо в среднем может иметь около 80 узловых точек. Алгоритм производит анализ относительного положения, размера и / или формы глаз, носа, скул и челюсти и затем использует эти параметры для поиска других изображений с соответствующими параметрами [5].

Существуют множество реализаций данного алгоритма. Наиболее современным является алгоритм, предложенный сотрудниками технического института имени Короля Монгкута (King Mongkut's Institute of Technology) [6]. Алгоритм основан на ортогональных проекциях и евклидовом преобразовании. Основное скопление точек расположено вокруг глаз, носа и рта. С извлеченными ориентирами формируются триплеты области и связанная с ними геометрическая инвариантность. Поскольку область является относительной инвариантностью при евклидовом преобразовании, применяются абсолютные отношения площадей между последовательными треугольными зонами.

Рис.2 Границы и внутренние углы лицевых треугольников

Относительный инвариант можно описать следующей формулой:

$$A_a(k) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} A(k), \quad k = 1, 2, 3, \dots, n$$

где $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$ - является определителем аффинного преобразования, $A(k)$ - фрагмент последовательности треугольников на шаблоне [6].

Поскольку матрица линейного преобразования неизвестна, абсолютные аффинные инварианты построены из области относительных инвариантов, взяв соотношение двух треугольников для того, чтобы вынести за скобки зависимости относительного инварианта области в определителе матрицы аффинного преобразования. Принимая отношение последовательных элементов в последовательности, набор абсолютных инвариантов может быть представлен в виде:

$$I(k) = \frac{A(k)}{A((k+1) \bmod n)}, \quad k = 1, 2, 3, \dots, n$$

$$I_a(k) = \frac{A(k)}{A_a((k+1) \bmod n)}, \quad k = 1, 2, 3, \dots, n$$

В случае измерения без помех абсолютный инвариант запроса совпадает с абсолютным инвариантом шаблона, т.е. $I(k) = I_a(k)$ при наличии помех и окклюзии, каждый из $I_a(k)$ будет иметь счетную часть $I(k)$, причем этот аналог легко определяет через круговой сдвиг, включающий n сравнений, где n - число инвариантов [6].

Рис.3 Вектор лицевого ориентира

Углы вычисляются по формуле [6]:

$$A \times B = |A||B| \sin \theta$$

$$\theta = \arcsin \left(\frac{A \times B}{|A||B|} \right)$$

Площадь каждого треугольника будет вычисляться как:

$$S = \frac{1}{2} |A \times B|$$

Алгоритм трехмерного распознавание лица. Трехмерное распознавание лиц - это механизм распознавания лиц, в котором используется трехмерная геометрия. Доказано, что методы распознавания трехмерного лица могут достигать значительно более высокой точности, чем их 2D-копии, позволяя идентифицировать человека независимо от угла наклона лица и условий освещения. Основным технологическим ограничением методов распознавания 3D-поверхностей является получение 3D-изображения, которое обычно требует использование камеры широкого диапазона. В качестве альтернативы для создания 3D-модели с последующей обработкой можно использовать несколько изображений под разными углами [3].

Первым этапом алгоритма является моделирование человеческого лица в виде двумерной гладкой параметрической поверхности, обозначаемой S и представляемой координатной диаграммой из выбранного подмножества [7].

$$x(\Omega) = (x^1(\xi^1, \xi^2), x^2(\xi^1, \xi^2), x^3 + (\xi^1, \xi^2))$$

Предположим, что функции $x^1, \dots, x^3 = C^r$ (при достаточно большом r) и векторы $\partial_i x = \frac{\partial}{\partial \xi^i} x$ линейно независимы.

В координатном пространстве элемент расстояния поверхности может быть выражен через метрический тензор как [7]:

$$ds^2 = g_{ij}d\xi^i d\xi^j, \quad i, j = 1, 2$$

где g_{ij} – тензор, равный

$$g_{ij} = \partial_i x \cdot \partial_j x$$

Нормалью S в точке x , является вектор, ортогональный касательному пространству $T_x S$ и может быть записан в виде произведения [7]:

$$N(x) = \frac{\partial_1 x \times \partial_2 x}{\|\partial_1 x \times \partial_2 x\|_2}$$

Поскольку наша поверхность является связанной и компактной, риманова метрика инициализирует метрику расстояния. Чтобы определить его, пусть $x, y \in S$ - две точки на поверхности и $c: [0, 1]$. Тогда длина определяется формулой

$$l(c) = \int_0^1 \left\| \frac{dc(t)}{dt} \right\| dt$$

Тогда расстояние от x до y на поверхности S можно получить по формуле

$$d(x, y) = \inf l(c)$$

Траектория пути минимальной длины, полученная в результате из экстремума функционала $[c]$, называется минимальными геодезическим, а $d(x, y)$ называется геодезическим расстоянием [7].

Нейронные сети и скрытые Марковские модели. Скрытые Марковские модели - это статистический инструмент, используемый при распознавании лиц [3]. Они используются совместно с нейронными сетями. Они генерируются в нейронной сети, которая создает псевдо 2D НММ (Hidden Markov Models). В качестве входных параметров для этого процесса 2D НММ являются выходные данные ANN (Artificial Neural Networks – искусственная нейронная сеть) [8].

Инициализация НММ:

$$\lambda = (A, B, \Pi)$$

$A = \{a_{ij}\}$ – матрица вероятности перехода состояния, где

$$a_{ij} = P[q_t = S_j | q_{t-1} = S_i], \quad 1 \leq i, j \leq N, \quad q_t \in S, \quad 1 \leq t \leq T, \quad 0 \leq a_{ij} \leq 1$$

$$\sum_{j=1}^N a_{ij} = 1, \quad 1 \leq i \leq N$$

где N - количество состояний в модели, $S = \{S_1, S_2, \dots, S_N\}$ – набор состояний, T – длина рассматриваемых последовательностей (количество фреймов).

$B = \{b_j(k)\}$ – матрица вероятности рассматриваемых символов, где

$$b_j(k) = P[O_t = v_k | q_t = S_j], \quad 1 \leq j \leq N, \quad 1 \leq k \leq M$$

где O_t - рассматриваемый символ во время t , M - количество различных рассматриваемых символов, $V = \{v_1, v_2, \dots, v_M\}$ – набор возможных рассматриваемых символов [8].

$\Pi = \{\pi_i\}$ – распределение исходного состояния, где

$$\pi_i = P[q_1 = S_i], \quad 1 \leq i \leq N$$

После инициализации Марковской модели начинается работа с изображением. Изображения просматривается сверху вниз, и модель последовательно переходит из одного состояния в другое.

Рис. 5 Разделение области лица на значимые части

Для каждого изображения из базы строится своя Марковская модель [8]. Задача нейронной сети будет заключаться в сравнении исходного изображения с шаблонным, посредством вычисления вероятности с помощью алгоритма прямого-обратного хода. С помощью такого метода успех идентификации посредством нейронной сети очень высок.

Заключение. Технология распознавания лиц добилась больших успехов в последние годы. Теперь компьютеры могут идентифицировать лица, которые не находятся в области с хорошим освещением и широким радиусом обзора. Бокового обзора движущегося изображения уже может быть достаточно для искусственного интеллекта, чтобы идентифицировать. Страны, включая Китай и США, разрабатывают, тестируют и используют технологию распознавания лиц. В Международном аэропорту Лос-Анджелеса Администрация безопасности транспорта США пытались использовать различные устройства безопасности, включая технологию распознавания лиц, которая должна была фотографировать пассажира и сравнивать его с фотографией паспорта непосредственно перед тем, как он пройдет охрану аэропорта [9]. Так же ведутся разработки по внедрению системы распознавания лиц в медицине. Вместо того, чтобы распознавать человека через FaceTech, медицинские работники могли бы идентифицировать болезни, глядя на лицевые особенности пациента. Это помогло бы снизить постоянную нагрузку в медицинских центрах, сократив очереди и упростив процесс назначения лечения.

Организации по защите прав и сторонники неприкосновенности частной жизни, такие как «Фонд электронных рубежей», «Большой брат смотрит» и ACLU (Американский союз защиты гражданских свобод) выражают озабоченность по поводу нарушения конфиденциальности за счет использования технологий наблюдения. Некоторые опасаются, что это может привести к «обществу тотального надзора», когда правительство и другие органы власти будут иметь возможность знать местонахождение и деятельность всех граждан круглосуточно [10].

Распознавание лиц можно использовать не только для идентификации человека, но и для извлечения других персональных данных, связанных с отдельными людьми, таких как фотографии с участием отдельных лиц, сообщения в блогах, профили социальных сетей, поведение в Интернете, шаблоны путешествий и т.д. Это в корне меняет динамику повседневной конфиденциальности, позволяя любому маркетологу, правительственному агентству или случайному незнакомцу тайно собирать личность и связанную с ними персональную информацию любого человека, захваченного системой распознавания лиц.

Список литературы

- [1] Тихонова Т.С., Белов Ю.С. Введение в системы отслеживания взгляда // Электронный журнал: наука, техника и образование. 2016. № 1 (5). С. 94-101.
- [2] Сакович И.О., Белов Ю.С. Обзор основных методов контурного анализа для выделения контуров движущихся объектов // Инженерный журнал: наука и инновации. 2014. № 12 (36). С. 11.
- [3] Фисенко, В.Т. Компьютерная обработка и распознавание изображений [Электронный ресурс]: учебное пособие / В.Т. Фисенко, Т.Ю. Фисенко. — Электрон. дан. — Санкт-Петербург: НИУ ИТМО, 2008. — 192 с. — Режим доступа: <https://e.lanbook.com/book/40795> (дата обращения 15.10.2018)
- [4] Тихонова Т.С., Белов Ю.С. Основные подходы к отслеживанию и распознаванию лица // Электронный журнал: наука, техника и образование. 2016. № 2 (6). С. 111-115.
- [5] Нестеров А.Ю., Бурмистров А.В. Белов Ю.С. Метод определения положения объекта в задачах распознавания образов // Электронный журнал: наука, техника и образование. 2016. № 2 (6). С. 82-89.
- [6] Aniwat Juhong, C. Pintavirooj Face Recognition Based on Facial Landmark Detection // Biomedical Engineering International Conference, IEEE, 2017 DOI: 10.1109/BMEiCON.2017.8229173
- [7] Alexander M. Bronstein, Michael M. Bronstein, Ron Kimmel Three-Dimensional Face Recognition // International Journal of Computer Vision, August 2005, Volume 64, Issue 1, pp 5–30
- [8] A.V. Nefian, M.H. Hayes Hidden Markov models for face recognition // Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP '98 (Cat. No.98CH36181) DOI: 10.1109/ICASSP.1998.678085
- [9] Юрко И.В., Алдобаева В.Н. Области применения и принципы работы систем распознавания и идентификации лиц по видеофиксации в реальном времени // Международный студенческий научный вестник. – 2018. – № 2.; URL: <http://eduherald.ru/ru/article/view?id=18416> (дата обращения: 15.10.2018).
- [10] Cackley, Alicia Puente Facial recognition technology Commercial Uses, Privacy Issues, and Applicable Federal Law // United States Government Accountability Office, GAO-15-621