РЕАЛИЗАЦИЯ ПРОЦЕДУР ВЗАИМОДЕЙСТВИЯ ЧЕЛОВЕКА С МАШИНОЙ, ПОСРЕДСТВОМ ОБНАРУЖЕНИЯ И РАСПОЗНАВАНИЯ ЛИЦ

А.В. Скрыпников, *д-р техн. наук, профессор*
В.В. Денисенко, *канд. техн. наук, доцент*
С.Е. Демичев, *студент*
А.В. Самохин, *студент*
Воронежский государственный университет инженерных технологий (Россия, г. Воронеж)

DOI:10.24412/2500-1000-2021-12-3-100-104

Аннотация. *В этой статье рассмотрена процедура взаимодействия человека с машиной. Приведены примеры применения системы обнаружения и распознавания лиц. Показаны различные алгоритмы обработки изображений и возможные решения ряда проблем, возникающих при аутентификации пользователей в информационной системе. Продемонстрировано, что распознавание лиц в реальном времени зависит от большого количества факторов и качества поступающих данных.*

Ключевые слова: *распознавание лиц, обработка изображений, аутентификация, ИИ, OpenCV.*

За последнее десятилетие обнаружение и распознавание лиц вышло из эзотерической области в область компьютерного зрения и стало одним из лучших и успешных прикладных решений анализа изображений и понимания алгоритмов. Системы компьютерного зрения являются областью не только компьютерных исследований, но и объектом исследования в нейробиологических и психологических тематиках, в основном из-за общего мнения, что достижения в области компьютерной обработки изображений прольют свет на то, как работает человеческий мозг, и наоборот.

Цель этой статьи – упростить процедуру взаимодействия человека с машиной при аутентификации пользователя посредством обнаружения и распознавания лица.

Общая постановка проблемы распознавания лица (в компьютерном зрении) может быть сформулирована следующим образом: «провести поиск и идентифицировать одно или несколько лиц в кадре на фото или видеофрагменте, на основании существующей базы данных лиц».

Распознавание лиц обычно состоит из двух этапов:

- Обнаружение лица – на фотографии выполняется поиск лица, затем изображение обрабатывается для кадрирования и выделения лица человека для более легкого распознавания.

- Распознавание лица – сравнение обнаруженного и обработанного в первом этапе лица с базой данных известных лиц для идентификации.

В данный момент одним из самых популярных решений для работы с компьютерным зрением является библиотека OpenCV [1]. В этом фреймворке реализовано достаточно много различных распознаваний. Встроенный детектор лица распознает лица на более 90% фотографий человека, смотрящего напрямую в камеру (анфас). Однако определение лица человека под углом требует оценки позы головы в 3D. Помимо этого на распознавание лица влияет недостаточная яркость изображения, увеличенный контраст, отсутствие точно фокусировки камеры или наличие посторонних предметов на лице.

Несмотря на то, что обнаружение лиц на фото имеет разбег в 10%, распознавание таких показателей достичь не может – 30-70% в целом. Распознавание лиц было серьезной областью исследований с 1990-х годов, но все еще далеко от надежного метода аутентификации пользователей. С каждым годом разрабатывается все больше и больше техник. Самый простой метод точного распознавания лиц – Eigenface –

она использует многомерное векторное пространство совместно с ковариантной матрицей из распределения вероятностей.

Фреймворк компьютерного зрения OpenCV с открытым исходным кодом значительно упрощает программирование ПО, основанного на распознавании объектов. Он имеет широкий функционал:

- Обнаружение объектов в пространстве в 2D-пространстве;
- Отслеживание объектов;
- Распознавание заданных образов;
- Фильтрацию Калмана;
- Различные методы искусственного интеллекта.

в уже готовой к использованию форме. Помимо этого, OpenCV предоставляет множество базовых алгоритмов компьютерного зрения через специальный API нижнего уровня.

Благодаря этой функциональности OpenCV руками энтузиастов был доработан под мультиплатформенность и поддерживает все доступные ныне системы. Местами функционал OpenCV может показаться подавляющим поскольку имеет слишком обширные возможности. Но это является и его достоинством, поскольку отличная работа реализованных методов позволяет достигать хороших результатов.

Функциональность OpenCV для распознавания лиц состоит из нескольких модулей:

- CXCORE – пространство имен для основных типов данных, содержит методы линейной алгебры и статистики, функции сохранения и обработчики ошибок, графический функционал для отрисовки на изображениях.
- CV – пространство имен, содержит в себе методы обработки изображений и калибровки используемой камеры.
- CVAUX – пространство имен для хранения уже deprecated и экспериментального кода, а также простые интерфейса для распознавания лиц.
- ML – содержит в себе интерфейсы машинного обучения.
- HighGUI – основные интерфейсы ввода/вывода и мультиплатформенные реализации для работы с окнами.
- CVCAM – интерфейсы для доступа к видео через DirectX для платформы Windows x86.

Когда принимается во внимание качество изображения, существует множество факторов, которые влияют на точность распознавания. Важно применять различные методы предварительной обработки изображений для стандартизации, которые поступают в систему распознавания. Большинство алгоритмов чрезвычайно чувствительны к условиям освещения, поэтому, если алгоритм распознавания обучен распознавать лица в темном помещении, то он вряд ли покажет хорошие результаты для изображения с хорошей освещенностью и наоборот. Эта проблема называется «зависимость от освещенности».

Существует множество других факторов, которые влияют на успешность распознавания: примерно одинаковые координаты глаз, одинаковое соотношение лица с общим размером фото, поворот лица в пространстве, наличие прически или макияжа, мимика лица, расположение источника света.

Основываясь на параметрах выше следует разрабатывать также алгоритм применения фильтров на предварительной обработке изображений. Для еще более точного распознавания имеется возможность применять эллиптическую маску, для вычитания с фото меняющихся объектов (прическа, фон и прочие).

Для разработки детектора лиц в данных условиях подойдет классификатор каскада Хаара.

При получении изображения (фото с камеры, фото с устройства, стоп-кадр видео) детектор лиц исследует каждое местоположение изображения и классифицирует его как «Лицо» или «Не лицо». Классификация предполагает фиксированный масштаб лица, скажем, 50на50 пикселей. Поскольку лица на изображении могут быть меньше или больше обычного, классификатор проходит по изображению несколько раз для поиска лиц в различных масштабах. Это может показаться огромным объемом обработки, но благодаря функциональным алгоритмам, классифи-

кация выполняется очень быстр, даже когда она применяется в различных масштабах. Классификатор использует данные, хранящиеся в файле XML, чтобы решить, как классифицировать каждое местоположение изображения. Загрузка OpenCV включает четыре разновидности XML-данных для фронтального распознавания лиц и одну для «лиц в профиль». Помимо этого, существует множество способов улучшить качество распознавания изображения – обработка цвета, обнаружение краев и прочие.

Благодаря такой реализации существует возможность использовать поток вебкамеры в качестве входных данных для системы вместо списка файлов. По сути, просто нужно захватить кадры с камеры, а не из файла, которые могут поступать бесконечно в потоке пока системе не придет сигнал на остановку и не привязываться к ограниченному количеству файлов.

Обычно точность распознавания лиц прямолинейно зависит от входных изображений для человека (обычно ~50 на человека), делая больше снимков каждого человека с разных ракурсов и различной освещенности. Если сделать большое количество снимков не представляется возможным, есть несколько иных методов, которые доступны для получения большего количества обучающих изображений путем создания новых изображений из уже существующих. Простое отзеркаливание по горизонтали увеличит количество образов для изучения в два раза. Преобразование фото и изменения масштаба, размера и угла наклона фотографии также предоставит множество альтернативных изображений для обучения, чтобы они были менее чувствительны к точным условиям.

Дополнительным условиям для улучшения обучения выступает «изменение качества изображения». К примеру, чтобы настроить для детектора лиц нечувствительность к шуму, следует обучить его на изображениях, где этот шум уже частично присутствует. В любом случае классификатор для распознавания лиц в разных условиях должен быть сначала обучен на похожих, чтобы иметь устойчивость к ним.

Не стоит забывать про переобучение, поскольку тогда классификатор становится слишком общим и результаты значительно ухудшаются. Вот почему часто получаются очень плохие результаты при первых прогонах алгоритма обучения, которая улучшается со временем.

Существуют также спорные решения в виде – выравнивание гистограммы. В данной области – это очень простой метод предварительной обработки изображений, который в некоторых ситуациях может ухудшить ситуацию, поэтому следует использовать несколько комбинаций с тестами для поиска наилучшей.

Важно понимать, что в основе алгоритма лежит сопоставление изображений, в основном выполняя эквивалент вычитания тестового изображения из обучающего изображения, чтобы увидеть, насколько они похожи. Это срабатывает неплохо, если это выполняет человек, но компьютер мыслит только в пикселях и числах, поэтому иногда это бывает затруднительно.

Итак, классификатор смотрит на один пиксель в тестовом изображении и вычитает значение серой шкалы этого пикселя из значения пикселя в ТОЧНОМ месте каждого тренировочного изображения, и чем меньше разница, тем лучше совпадение. Поэтому если изображение сдвинуто на несколько пикселей в сторону по оси или используются изображения разных размеров, классификатор будет воспринимать их как абсолютно разные изображения.

Это также верно, если фон отличается, потому что алгоритм не знает разницы между фоном и передним планом (лицом), поскольку изображение плоское в рамках распознавания. Поэтому важно обрезать как можно больший участок фона, например, используя только небольшой участок внутри контура лица, который вообще не содержит элементов или цвета фона.

Поскольку изображения должны быть почти идеально выровнены, это фактически означает, что во многих случаях использование небольших изображений с низким разрешением (например, путем сжатия изображений до размера миниатюр) может дать лучшие результаты рас-

познавания, чем большие изображения с высоким разрешением.

Кроме того, даже если изображения идеально выровнены, если тестовое изображение немного ярче, чем тренировочное, то он все равно будет думать, что совпадения нет. Выравнивание гистограммы может помочь во многих случаях, но в других случаях оно также может ухудшить ситуацию, поэтому различия в освещении – сложная и распространенная проблема. Также существуют проблемы, например, если на тренировочном изображении была тень слева от носа, а на тестовом изображении - справа, то это часто приводит к плохому совпадению.

Вот почему распознавание лиц относительно легко сделать в режиме реального времени, если вы тренируетесь на ком-то, а затем мгновенно пытаетесь распознать его после, поскольку это будет та же камера, а фон будет таким же, их выражения будут почти такими же, освещение будет таким же, и направление, с которого вы смотрите на них, будет таким же. Так что в этот момент вы часто будете получать хорошие результаты распознавания. Но если попытаться распознать их, с другой стороны, или из другой комнаты, или снаружи, или в другое время дня, это часто даст плохие результаты.

Библиографический список

1. Бацева Д.А., Черепков Е.А., Белов Ю.С. Определение характерных точек лица для построения объемно модели головы человека // Электронный журнал: наука, техника и образование. 2018 – №1 (17). – С. 70-75.

2. Интеллектуальные информационные системы управленческой деятельности / В.В. Денисенко, К.С. Евтеева, И.И. Савченко [и др.] // Системный анализ и моделирование процессов управления качеством в инновационном развитии агропромышленного комплекса: Материалы V Международной научно-практической конференции, в рамках реализации Ассоциации «Технологическая платформа «Технологии пищевой и перерабатывающей промышленности АПК – продукты здорового питания», Воронеж, 21 мая 2021 года / Воронеж. гос. ун-т инж. технол.. – Воронеж: Воронежский государственный университет инженерных технологий, 2021. – С. 268-275.

3. Кучер М.Ю., Белов Ю.С. Этапы функционирования систем распознавания лиц // В сборнике: Наука. Исследования. Практика. Сборник избранных статей по материалам Международной научной конференции – Санкт-Петербург, 2020 – С. 91-93.

4. Скрыпников А.В. Распознавание лиц в системе управления качеством / А.В. Скрыпников, В.В. Денисенко, С.М. Орлова // Системный анализ и моделирование процессов управления качеством в инновационном развитии агропромышленного комплекса: Материалы V Международной научно-практической конференции, в рамках реализации Ассоциации «Технологическая платформа «Технологии пищевой и перерабатываю-щей промышленности АПК – продукты здорового питания», Воронеж, 21 мая 2021 года / Воронеж. гос. ун-т инж. технол.. – Воронеж: Воронежский государственный университет инженерных технологий, 2021. – С. 400-406.

IMPLEMENTATION OF PROCEDURES OF HUMAN-MACHINE INTERACTION THROUGH FACE DETECTION AND RECOGNITION

A.V. Skrypnikov, *Doctor of Technical Sciences, Professor*
V.V. Denisenko, *Candidate of Technical Sciences, Associate Professor*
S.E. Demichev, *Student*
A.V. Samokhin, *Student*
Voronezh State University of Engineering Technologies
(Russia, Voronezh)

Abstract. *This article describes the procedure for human-machine interaction. Examples of application of the face detection and recognition system are given. Various algorithms for image processing and possible solutions to a number of problems arising during user authentication in an information system are shown. It has been demonstrated that real-time face recognition depends on a large number of factors and the quality of the incoming data.*

Keywords: *face recognition, image processing, authentication, AI, OpenCV.*