

АЛГОРИТМЫ ИДЕНТИФИКАЦИИ ЛИЧНОСТИ В СИСТЕМАХ МАССОВОГО ИНТЕРАКТИВНОГО ПОЛЬЗОВАТЕЛЬСКОГО ВЗАИМОДЕЙСТВИЯ

В. М. Медведев

Белорусский государственный университет, г. Минск;

tom.stanley19981996@gmail.com;

науч. рук. – А. В. Курочкин

В данной статье представлена интерактивная пользовательская система, основанная на алгоритмах распознавания и идентификации личности по лицу. На основе поставленных требований и технических возможностей системы проанализированы существующие алгоритмы распознавания и их недостатки, существенно влияющие на качество. В работе предложена архитектура разрабатываемой системы с разделением по ключевым модулям, а также представлены рекомендации по подбору параметров исследуемых моделей для повышения точности распознавания и идентификации лица в реальном времени по видеопотоку.

Ключевые слова: нейронные сети; глубокое обучение; идентификация личности; интерактивные системы; компьютерное зрение.

ПРОЕКТИРОВАНИЕ ИНТЕРАКТИВНОЙ СИСТЕМЫ

В качестве системы для проектирования был выбран интерактивный домашний помощник (рис. 1). К числу основных требований относятся: многопоточность (одновременная обработка лиц нескольких пользователей), хорошее качество распознавания, устойчивость к шумам во входных данных, производительность и защищенность данных.

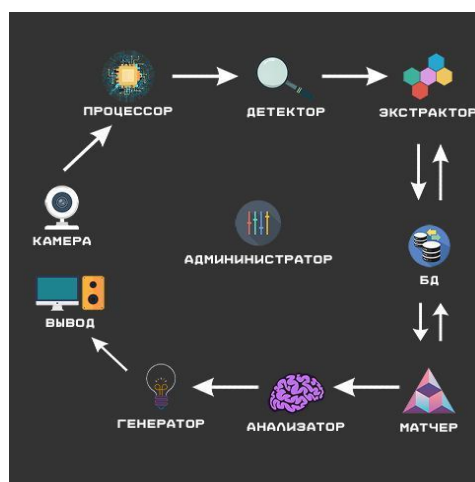


Рис. 1. Архитектура будущей интерактивной системы

Архитектура системы состоит из следующих компонентов:

- Камера. Используется для подачи непрерывного видеопотока на процессор.

- Процессор. Получает видеопоток с камеры, преобразовывает его в кадры, нормализует, центрирует, и осуществляет другую предварительную обработку для улучшения качества дальнейшего детектирования. При этом от видеопотока отделяется аудио и формируется спектрограмма.

- Детектор. Получает кадры и спектрограмму, осуществляет поиск лица в кадре и выделяет его с использованием метода Виолы-Джонса и каскадов Хаара [1]. В спектрограмме осуществляется отделение говорящих. Обычно детектор содержит обнаружение торцевых кромок изображения, сегментацию и локализацию.

- Экстрактор. Здесь выделяются шаблоны для каждого лица и речь говорящего, которые затем передаются на вход в матчер. Шаблон для лица – это некоторое представление ключевых точек или текстуры, а для речи – фонемы и форманты [2]. В зависимости от используемой в системе модели экстрактор и матчер могут быть объединены.

- Матчер. Представляет собой алгоритм для сопоставления лица и речи с данными в базе. Получает на вход шаблон лица и речь говорящего от экстрактора и сопоставляет его с данными в базе. На выходе образуется метка, например, имя человека, и границы лица. Отправляет результат работы модулю администрирования.

- Анализатор. После успешной идентификации анализирует полученные данные: эмоции, окружение, речь. Анализатор состоит из двух частей. Первая часть отвечает за преобразование данных в текстовый вид, а вторая – за структурирование полученной информации и выделение значимых частей. После этого он создает приоритет генерации интерактива и передает данные генератору в виде готового набора команд.

- Генератор. На основе полученной информации от анализатора взаимодействует с пользователем при помощи устройств вывода.

- Модуль администрирования. Контролирует весь процесс работы системы, отвечает за добавление новых пользователей в систему и дает команду к анализу полученной информации. Обрабатывает нештатные ситуации и взаимодействует с пользователем через приложение.

Система должна уметь распознавать действия в ситуации, когда личность не распознана, и немедленно оповещать об этом владельца через администратора. Для ускорения работы алгоритма в системе распознавания можно использовать структуры данных на основе k-d или VP деревьев.

АНАЛИЗ АЛГОРИТМОВ РАСПОЗНАВАНИЯ

Алгоритм Виолы-Джонса

Алгоритм разработан для фронтальных изображений лица, поэтому плохо определяет лица, повернутые вбок, вверх или вниз, а при изменении освещения лица алгоритм и вовсе перестает работать (рис. 2). Несмотря на эти недостатки, алгоритм до сих пор является одним из самых быстрых и популярных для задач классификации, но медленным при обучении.

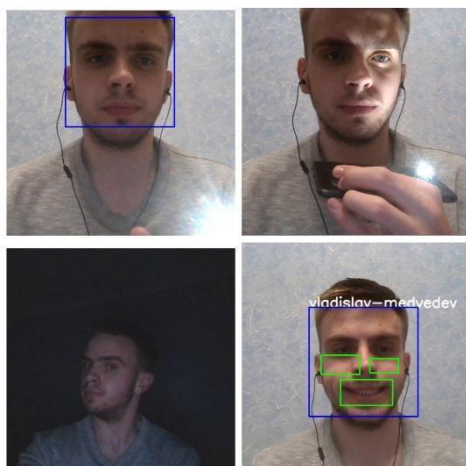


Рис. 2. Результат распознавания алгоритмом Виолы-Джонса

Метод главных компонент

Получаемый вектор признаков в результате работы метода главных компонент может использоваться для последующего сравнения с вектором, имеющимся в базе данных. Проблемы распознавания схожи с проблемами в алгоритме Виолы-Джонса за исключением того, что метод главных компонент неплохо работает в случаях, когда лицо частично скрыто. Но при плохой освещенности он начинает ошибаться, неверно идентифицируя человека.

Методы на основе глубокого обучения. FaceNet

FaceNet инвариантна к освещению, позе и другим изменениям. Хорошее качество распознавания наблюдается также при наличии на лице посторонних предметов и в случае отдаления от камеры (рис. 3). Если использовать в базе данных небольшое количество изображений, сеть может колебаться между пользователями, что не очень хорошо.

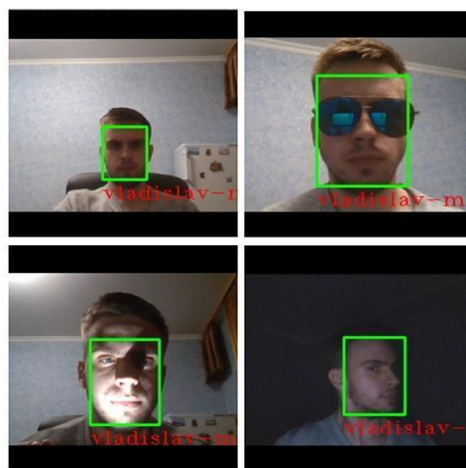


Рис.3. Результат распознавания FaceNet

Наиболее существенный минус FaceNet – это ошибочная идентификация при появлении нового пользователя, лицо которого воспринимается как одно из лиц, существующих в базе.

ЗАКЛЮЧЕНИЕ

Алгоритмы распознавания по лицу позволяют перейти к интерактивным системам нового уровня, не требующим прямого воздействия пользователя. В таких системах не удастся добиться полного совпадения лица в базе данных с лицом, поступающим на вход распознавателя, из-за постоянного перемещения человека в пространстве. Поэтому рациональнее применять алгоритмы, основанные на поиске по подобию. Помимо проблем с перемещением, распознавателю на стадии обучения нужно подавать различные модификации лица (положение головы, лицо с аксессуарами и под разными углами освещения) для получения хорошего качества идентификации, а в качестве алгоритмов распознавания стоит использовать модели глубокого обучения, которые зарекомендовали себя как модели, имеющие высокую точность распознавания.

Библиографические ссылки

1. Viola P., Jones M. Rapid Object Detection using a Boosted Cascade of Simple Features [Электронный ресурс]. URL: <https://www.cs.cmu.edu/~efros/courses/LBMV07/Papers/viola-cvpr-01.pdf> (data of access: 06.04.2020).
2. Coxhead P. Phones and Phonemes [Электронный ресурс]. URL: <https://www.cs.bham.ac.uk/~pxc/nlp/NLPA-Phon1.pdf> (data of access: 17.05.2020).