Распознавание лиц — Мифы и реальность. Основные понятия

Интерес к теме распознавания лиц неуклонно растет, особенно в сфере систем безопасности. Эта тема обросла множеством мифов, некорректно влияющих на выбор и внедрение систем распознавания лиц. Чтобы мифов становилось меньше, а качественных решений — больше, важно, чтобы интегратор и заказчик говорили на одном языке.

Сегодня даже мобильные телефоны умеют распознавать лица своих владельцев и создается ощущение, что распознавание лиц — одна из стандартных ИТ-технологий, которая уже давно является частью нашей повседневной жизни. Как будто это просто обычная опция, которую можно получить в любой сфере — с покупкой нового телефона или как дополнительную функцию в системе видеонаблюдения. Такой подход порождает завышенные ожидания от системы распознавания лиц и массу мифов. Мы же обратимся к фактам и определим границы реальных возможностей систем распознавания лиц, на которые может рассчитывать интегратор и заказчик.

Массовое представление о системах распознавания лиц сформировано рекламой и голливудскими фильмами. Киноиндустрия приучила нас к тому, что с помощью систем распознавания лиц возможно почти все, а реклама — что это уже давно существует. Но на практике еще даже не созданы отраслевые стандарты, обеспечивающие совместимость решений различных разработчиков.

Мировой рынок и рейтинг производителей

Согласно отчету Facial Recognition Market компании MarketsandMarkets от июня 2019 года, мировой рынок распознавания лиц к 2024 году принесет $7 млрд. прибыли при ежегодном росте в 16%. До конца 2019 года этот сегмент рынка ожидает получить прибыль в $3,2 млрд. Согласно отчету, самыми большими игроками рынка распознавания лиц являются следующие компании: NEC (Япония), Aware (США), Gemalto (Нидерланды), Ayonix Face Technologies (Япония), Cognitec Systems GmbH (Германия), NVISO SA (Швейцария), Daon (США), StereoVision Imaging (США), Techno Brain (Кения), Neurotechnology (Литва), Innovatrics (Словакия), id3 Technologies (Франция), IDEMIA (Франция), Animetrics (США) и MEGVII (Китай).

Самым авторитетным источником информации о производителях алгоритмов распознавания лиц является NIST (Американский национальный институт стандартов

и технологий). Каждый год NIST проводит Face Recognition Vendor Test и публикует результаты в своем отчете. Тестирование NIST очень серьезное, каждый алгоритм прогоняется через базу в 26,6 млн лиц, в числе которых портретные фотографии людей в хорошем качестве и фото с веб-камер (телефонов) в различном качестве и ракурсах. Именно тестирование NIST на практике показывает значительное улучшение качества алгоритмов распознавания лиц, потому что они стали работать на нейронных сетях. Если еще 6 лет назад для распознавания лица требовалось время не менее 2–3 секунд при неподвижном лице в кадре и качество изображения не менее 500 пикселей на метр, то сейчас многие решения по распознаванию лиц могут работать «на лету» и в потоке (даже в условиях города) при разрешении от 200 пикселей на метр и выше. Чтобы любой читатель понял разницу в этих параметрах, мы рассмотрим базовые понятия по распознаванию лиц для двухмерных моделей, так как именно 2D-модели обеспечивают доступность применения в системах IP-видеонаблюдения.

Распознавание лиц является одной из биометрических технологий, наряду с такими параметрами, как отпечаток пальца, геометрия руки, радужная оболочка глаза, голос, подчерк и т.п. Системы биометрической идентификации людей на основании распознанных лиц исторически начали свое бизнес-применение в проектах, связанных с пограничным контролем в США и Европе. Традиционными «потребителями» таких систем являются государственные службы: полиция, пограничные и таможенные службы. Самими передовыми странами по внедрению технологий распознавания лиц являются США и Китай.

Системы распознавания лиц активно используются в Китае, где коммунистическое правительство применяет китайские технологии распознавания лиц от компаний Megvii, SenseTime и CloudWalk в системах Skynet и Sharp Eyes для задач так называемого «социального контроля» — системы оценки граждан Китая на лояльность к общественному и политическому строю, по итогам работы которой неблагонадежным гражданам могут отказать в ряде сервисов и услуг, например, в выдаче кредита или получении социальных выплат.

Обе страны — и США, и Китай — уже отличились серьезными скандалами на тему распознавания лиц. В США поднят вопрос о правомерности и законности распознавания лиц людей в публичных местах. Сан-Франциско и еще несколько городов законодательно запретили применение распознавания лиц в публичных местах. А мировое сообщество обеспокоено и жестко осудило тот факт, что китайское правительство использует технологию распознавания лиц с целью ущемления прав человека.

Обнаружение, распознавание, идентификация в СВН

Напомним читателям базовые термины систем видеонаблюдения (СВН), так как они являются основой для успешной работы алгоритмов распознавания лиц. Это детекция, распознавание и идентификация.

Детекция или обнаружение — это изображение от камеры такого качества, которое позволит оператору визуально детектировать/обнаружить в кадре движущийся объект.

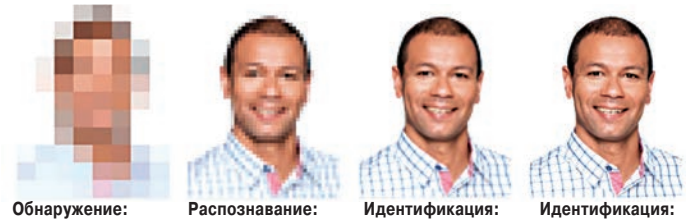
Распознавание — это изображение от камеры такого качества, которое позволит оператору визуально распознать тип объекта – человек, машина, животное и определить его цвет, параметры и направление движения.

Идентификация — это изображение от камеры такого качества, которое позволит оператору визуально узнать/идентифицировать/опознать незнакомого человека, транспортное средство или другие объекты.

Воспользуемся данными одного из ведущих производителей IP-видеонаблюдения по параметрам детекции, распознавания и идентификации (**рис.1** и **Табл.1**).

Таблица 1. Пример требований для обнаружения, распознавания и идентификации

Эксплуатационные требования	Количество пикселей по горизонтали для лица	пикс./см
Идентификация в неблагоприятных условиях	80 пикс./лицо	5 пикс./см
Идентификация в благоприятных условиях	40 пикс./лицо	2,5 пикс./см
Распознавание	20 пикс./лицо	1,25 пикс./см
Обнаружение	4 пикс./лицо	0,25 пикс./см



Обнаружение:
4 пикс./лицо
25 пикс./метр

Распознавание:
20 пикс./лицо
125 пикс./метр

Идентификация:
(хорошие условия)
40 пикс./лицо
250 пикс./метр

Идентификация:
(сложные условия)
80 пикс./лицо
500 пикс./метр

Рис. 1. Минимальные требования для обнаружения, распознавания, идентификации человека системой видеонаблюдения

Данные параметры регулирует стандарт ***ДСТУ EN 62676–4:2017*** «Системи відеоспостереження охоронного призначення. Частина 4. Правила застосування», который соответствует международному стандарту EN 62676–4:2015 Video surveillance systems for use in security applications Part 4: Application guidelines.

Для проектирования СВН применяют программные калькуляторы, которые помогают просто и наглядно определить плотность пикселей в сцене с учетом модели камеры, высоты установки и поля ее обзора (**рис. 2**).

Основополагающим в успешной работе алгоритмов распознавания лиц является именно ***возможность идентификации человека***. Опираясь на стандарт EN 62676–4, мы можем легко произвести расчет плотности пикселей, воспользовавшись калькуляторами, которые есть практически у всех вендоров IP-камер.

Для примера возьмем онлайн-калькулятор одного из ведущих мировых брендов IP-камер и решим две задачи:

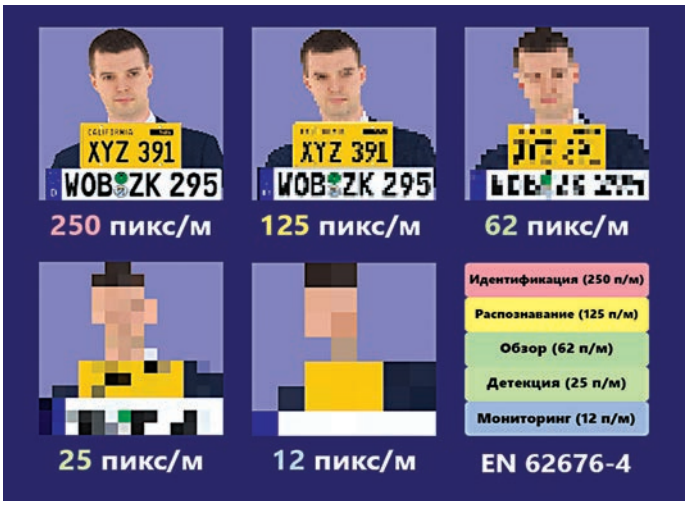


Рис. 2. Плотность пикселей для различных целей видеонаблюдения в IP Video System Design Tool

➢ на каком расстоянии от камеры можно будет распознавать лицо, если у заказчика имеется 2 Мп камера и она установлена на уровне 2 метра;
➢ какое разрешение камеры потребуется, если необходимо распознавать лица на улице города при высоте установки камеры не менее 3 метров.

Возьмем для примера 2 Мп IP-камеру с варифокальным объективом и сделаем расчет для граничных значений фокусного расстояния при f=2,8 мм и f=8,5мм, установив параметр плотности 250 пикселей на метр, так как у нас оптимальная высота установки камеры 2 метра. Очень быстро получаем четкие ответы: при f=2,8 мм камера сможет распознать лица на расстоянии до 4 метров, при этом ширина зоны обзора составит 7 метров (**рис. 3**), а при f=8,5 мм камера сможет распознать лица на расстоянии до 11 метров при той же ширине сцены.

На основании расчетов делаем общий вывод по первой задаче: стандартная 2 Мп камера с широким углом обзора позволит идентифицировать человека в условиях, когда камера установлена невысоко и человек находится достаточно близко к самой камере. Такими зонами являются проходные, входы в офис и участки, которые имеют небольшую ширину сцены. Если 2 Мп камера установлена на высоте 2,5 метра и более, то для получения нужной плотности пикселей для идентификации человека варифокальной камерой необходимо уменьшить угол обзора, а если это камера с фиксированным объективом, то придется поменять установленную 2 Мп камеру на камеру с меньшим углом обзора.

Если нам требуется идентифицировать людей в условиях города (улица шириной 3 метра) при высоте установки камеры не менее 3 метров, то необходимо делать расчет для неблагоприятных условий и устанавливать параметр 500 пикселей на метр. Расчет показывает, что 2 Мп и даже 3 Мп камеры не подойдут для решения задачи идентификации человека в условиях города, качественную картинку мы сможем получить только при использовании камер более высокого разрешения, например — 4Мп и выше.

На основании расчетов делаем общий вывод по второй задаче и развеиваем самый распространенный миф: ***2 Мп или 3 Мп камеры, использующиеся для оценки общей обстановки в городе, не подойдут для задач распознавания лиц.*** *Если в условиях города требуется решать задачу распознавания лиц, необходимо либо заменить имеющиеся камеры на модели более высокого разрешения, либо установить дополнительные камеры специально для задач распознавания лиц.* Выбор камер высокого разрешения необходимо делать параллельно с выбором самого решения по распознаванию лиц, учитывая рекомендации разработчиков решений по распознаванию лиц, так как у некоторых IP-камер большого разрешения может быть низкая светочувствительность, которая в условиях сложного освещения (сумерки, пасмурная погода и т.п.) не позволит качественно распознавать лица.

Благодаря несложным расчетам, которые можно сделать на калькуляторах производителей IP-камер, интегратор и сам заказчик смогут четко определить, пригодны ли установленные у него камеры для задачи распознаваниц лиц или нет. Помимо плотности пикселей в расчетах также важно учесть рекомендуемые разработчиком решений по распознаванию лиц параметры — высоту установки камеры, предельные отклонения лица от горизонтальной и вертикальной плоскости и т.п.

Алгоритм распознавания лиц

Если геометрия установки и разрешение IP-камер позволили получить нужную для идентификации людей плотность пикселей, интегратор и/или заказчик смогут протестировать работу системы распознавания лиц в реальных условиях. Рассмотрим еще несколько базовых понятий, с которыми нередко наблюдается путаница как у заказчиков, так и у некоторых интеграторов. Речь идет о таких терминах, как «детекция лиц» и «распознавание лиц».

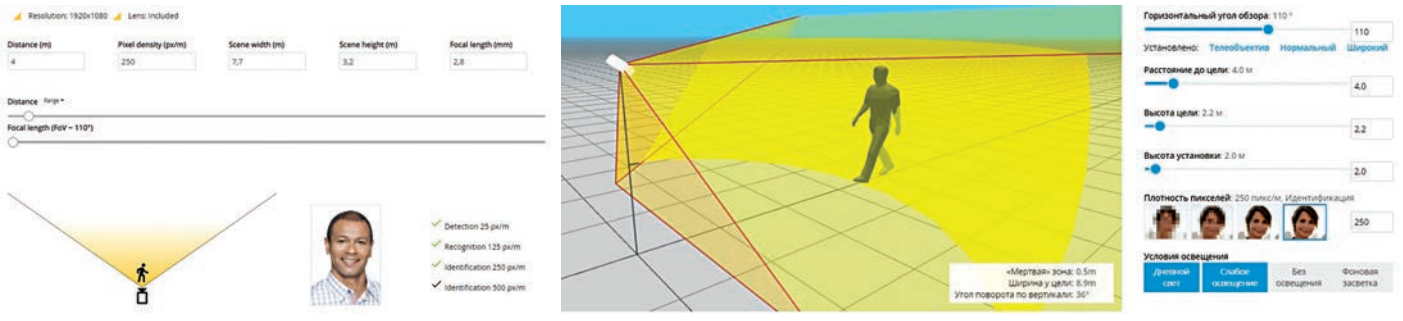


Рис. 3. Пример обеспечения идентификации человека 2Мп камерой при f=2.8мм.

Термины «детекция лиц» и «распознавание лиц» применимы только к изображению, которое позволит нам идентифицировать человека, т.е. в сценах с разрешением более 250 пикселей на метр или свыше 40 пикселей на лицо.

Детекция лиц – это результат работы видеоаналитики, позволяющий обнаружить факт наличия одного или несколько лиц в кадре (**рис. 4**). Сама по себе детекция лиц не определяет кто эти люди, сколько их, какого они пола или возраста. Детекция лиц доступна у некоторых вендоров, как встроенная видеоаналитика на борту IP-камер.

Рис. 4. Пример детекции лиц в кадре

Детекция лиц может стать бюджетным способом решения ряда задач, например, если это позволяет IP-камера, делать снимок экрана (скриншот) каждый раз при появлении людей перед камерой, сохраняя эти изображения в заданной папке.

Распознавание лиц — это результат работы видеоаналитики, позволяющей определить соответствует ли лицо/лица в кадре изображению лица/лиц в базе данных. Профессиональные решения по распознаванию лиц разрабатываются специализированными компаниями, а не производителями камер.

Оба параметра тесно связаны между собой — детекция лица является ключевым фактором для алгоритма распознавания лица. Если «пойманные» камерой лица будут иметь невысокое качество (а это происходит при качестве изображения менее 250 пикселей на метр или 40 пикселей на лицо), ни о каком распознавании лиц речь идти не может (либо алгоритмы распознавания лиц будут работать со значительными ошибками и решение будет не эффективным).

Помимо плотности пикселей в кадре на качество работы алгоритма распознавания лиц влияют такие факторы, как наличие теней на лице, отклонение, поворот от горизонтальной и вертикальной плоскости, засветка лица либо его частичное сокрытие (умышленное или случайное).

Стоит уточнить разницу между такими понятиями как «алгоритм» и «система распознавания лиц».

Алгоритм распознавания лиц представляет собой программное обеспечение, которое способно осуществлять детекцию лица в кадре (face detection), оценивать качество изображения и строить математическую модель лица (или биометрический шаблон — (face capture). На следующем этапе построенный шаблон сравнивается с шаблонами из базы данных (face match).

Система распознавания лиц — это готовое к работе ПО (имеющее пользовательскую оболочку), с помощью которого можно создать базу данных лиц, получить результат работы алгоритма распознавания и соответствующие отчеты. Результатом работы алгоритма/системы распознавания лиц является вероятность идентификации конкретного человека, выраженная в процентах (**рис. 5**).

Ключевыми точками для построения биометрического шаблона являются расстояние между глазами, форма лица и другие параметры (**рис. 6**) — основные принципы формирования такого шаблона, типы и количество ключевых точек, а также другие параметры являются секретом каждой компании-разработчика алгоритма. Для построения биометрического шаблона лица активно используют машинное обучение на основе нейронных сетей, что, согласно данным NIST, позволило улучшить алгоритмы распознавания лиц в 20 раз за последние 5 лет.

Показателем качества работы алгоритмов/системы распознавания лиц являются параметры предельно допустимых ошибок первого и второго рода.

К ошибке первого рода «ошибочная идентификация» (false positive, ложноположительное решение) или FPIR (False Positive Identification Rate) или FAR (False Acceptance Rate) относят распознавание постороннего человека, как будто он есть в заданной базе данных или «пропуск чужого».

Рис. 5. Пример распознавания лица в кадре. Вероятность совпадения 98%

Рис. 6. Пример биометрического шаблона лица

К ошибке второго рода «пропуск цели» (false negative, ложноотрицательное решение) или FNIR (False Negative Identification Rate) или FRR (False Rejection Rate) является пропуск (не распознавание) человека из заданной базы данных или «пропуск своего».

Параметры FAR и FRR связаны между собой. Чем выше точность идентификации и меньше FAR, тем более высока вероятность FRR и не распознавания человека из заданной базы данных, что добавит службе безопасности работу по ручной идентификации человека. И наоборот, чем меньше показатель FRR, тем более высока вероятность пропуска чужого человека, а это для многих отраслей очень критично и может представлять реальную угрозу.

По данным отчета NIST Face Recognition Vendor Test (FRVT) 2018 года, алгоритм компании NEC Corporation признан самым лучшим в мире с параметром FAR = 0,5 % при базе данных в 12 миллион человек. В разные годы тот же NEC приводил следующие пары параметров: FAR = 0,1 % при FRR = 0,13 %, FAR = 0,01 % при FRR = 0,26 %, FAR = 0,001 % при FRR = 0,64 % и FAR = 0,0001 % при FRR = 1,8 %.

Важно отметить, что NIST проводит так называемые синтетические тесты — т.е. тесты с использованием заранее подготовленной базы. Тестирование распознавания лиц в условиях реального объекта с реальными камерами, ракурсами, освещением, засветками и потоком людей не дают нормативных результатов, поэтому в некоторых странах на законодательном уровне приняты базовые значения вероятности идентификации не менее 85%, а величины FAR — не более 1%.

Отличие системы распознавания лиц от алгоритма

Алгоритм распознавания лиц можно сравнить с операционной системой, а систему распознавания лиц с пользовательским приложением. Последняя может иметь вид программно-аппаратной платформы, отдельного ПО, дополнительного модуля системы видеонаблюдения или комплексной системы безопасности. В основе разработки может лежать как проприетарный алгоритм производителя, так и алгоритмы распознавания лиц из открытых источников (open source).

Выбор решения и вендора по распознаванию лиц зависит от конкретной задачи и целей биометрической ***идентификации*** или ***верификации*** людей на объекте. Биометрическая идентификация отвечает на вопрос: ***«Кто ты?»*** и использует для идентификации человека только его лицо. В свою очередь верификация отвечает на вопрос: ***«Действительно ли это ты?»*** и позволяет использовать два или более параметров идентификации конкретного человека, например, лицо в сочетании с картой доступа, паролем, отпечатком пальца и т.д.

Биометрическая идентификация по лицу позволяет операторам СВН предотвращать инциденты и реагировать на тревожные сообщения, если среди распознанных лиц появятся злоумышленники, а верификация дает возможность решить задачи системы контроля доступа или комплексной системы безопасности. В последнем случае система распознавания лиц должна быть интегрированной с СВН и СКД.

Системы распознавания лиц с биометрической идентификацией или верификацией находят свое применение в следующих сферах:

- пограничный контроль;
- правоохранительные органы, криминалистика;
- безопасные города;
- объекты транспортной инфраструктуры (аэропорты, вокзалы и т.п.);
- контроль доступа к зонам/объектам повышенной важности;
- учет рабочего времени;
- казино;
- сфера гостеприимства (отели, клубы и т.п.);
- банки;
- маркетинговые исследования и реклама;
- ритейл, продажи и программы лояльности;
- социальные сети;
- VAaaS (Video-Analytics-as-a-Service, видеоаналитика как сервис).

В статье затронуты лишь базовые понятия систем видеонаблюдения и распознавания лиц, призванные помочь интегратору и заказчику найти общий язык. Как выбрать и протестировать систему распознавания лиц, оценить бюджет проекта, что является мифом и реальностью с точки зрения конкретных задач, какие интересные проекты уже реализованы — осталось еще много вопросов, на которые можно дать четкий и профессиональный ответ.

Алена ШВЕЦОВА,
эксперт по системам видеонаблюдения, #cctvMadonna