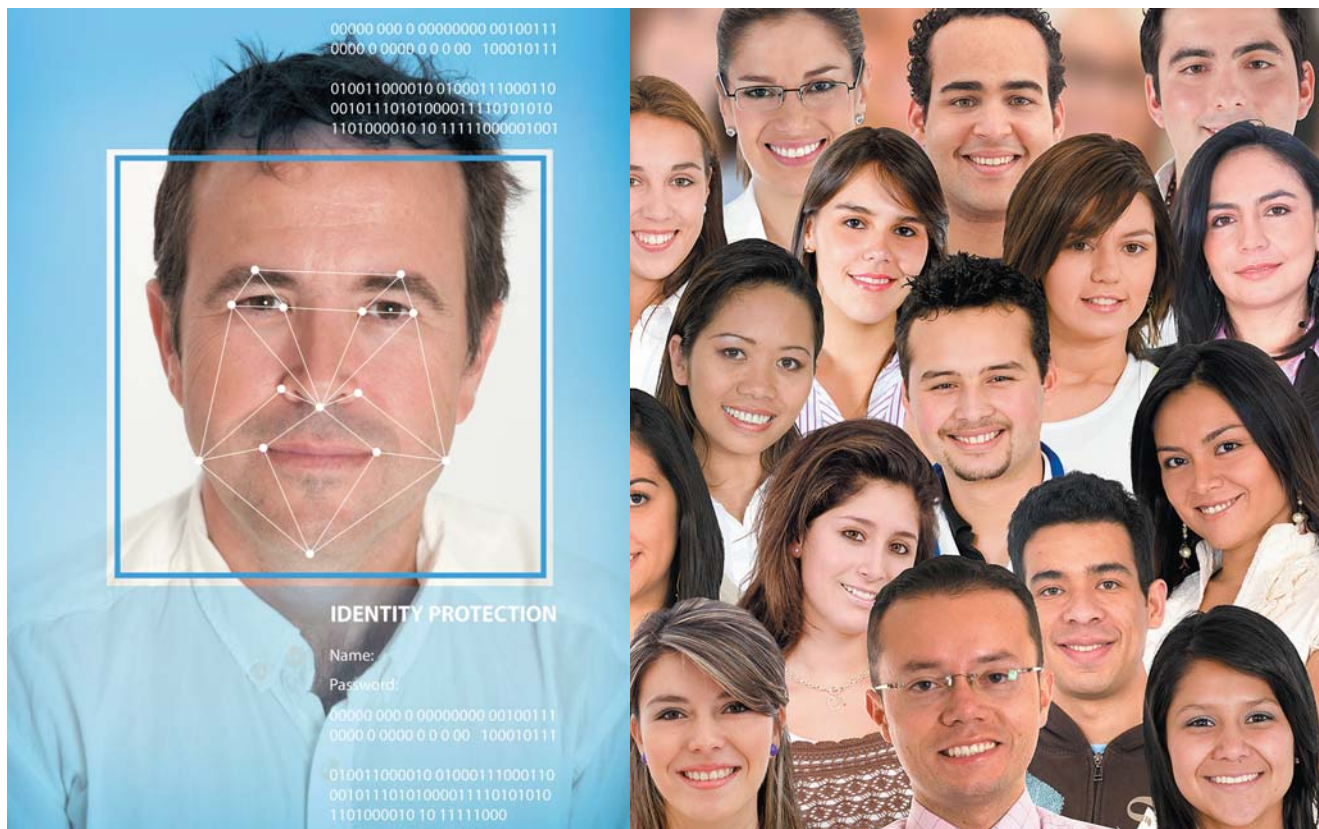


ТЕХНОЛОГИЯ И РЕШЕНИЯ ДЛЯ РАСПОЗНАВАНИЯ ЛИЦ И ЖЕСТОВ ОТ КОМПАНИИ OMRON



За последние годы произошли значительные изменения, связанные с разработкой новых технологий взаимодействия между человеком и машиной (компьютером), позволившие существенно улучшить качество и комфортность управления, расширить возможности взаимодействия, позволяя создавать и внедрять более интуитивные интерфейсы пользователя. Компания Omron в настоящее время активно работает в данном направлении.

Мультисенсорные интерфейсы окончательно перестали быть футуристической дизайнерской концепцией и перешли в наступление во всех сегментах ценового диапазона. К стандартным и привычным интерфейсам на основе клавиатуры, мыши, тачпада, сенсорного экрана активно внедряются интерфейсы распознавания речевых команд, синтез речи для обратной связи с оператором, системы виртуальной реальности с применением методов видеообработки изображения в рабочем поле с выделением нужных объектов, их идентификацией, трекингом выделенных объектов для применения в пользовательском интерфейсе для управления (жестовый

интерфейс, интерфейс мимики лица). Интерфейсы нового типа активно внедряются и в бытовую технику, мобильные устройства, автоматизированные системы управления. В интерфейсах мобильных устройств все большее применение находят биометрические системы, в которых используется распознавание индивидуальных параметров человека, строения его тела, деталей лица, отпечатков пальцев и сетчатки глаза.

VFRS — первый шаг в распознавании лиц

Японская компания Omron, ведущий мировой производитель электронных компонентов и систем, в настоящее время активно проводит разработки в данном направлении. Разработки новых технологий для биометрических систем, и в частности, распознавания на изображении тел человека, анализ движения и идентификации лица человека начались еще в начале 2000-х гг. Возможности новой технологии распознавания лиц в кадрах изображения, полученных с видеосенсоров мобильных устройств, впервые были продемонстрированы компанией Omron на выставке

«Безопасность Японии» в 2005 г. Тогда был представлен датчик с программным обеспечением (ПО) Vision Face Recognition Sensor (VFRS). До этого времени на рынке были слабо представлены коммерческие продукты данного направления.

Новая технология Omron ориентирована в основном на сектор мобильных устройств, например для реализации в сотовых телефонах со встроенными камерами. Применение биометрического метода распознавания владельца устройства по лицу является хорошей альтернативой таким методам безопасности и защиты от несанкционированного доступа посторонних лиц к компьютерам или объектам, как пароли и системы распознавания отпечатков пальца. Основное достоинство технологии в том, что она не требует дополнительных аппаратных устройств и работает по биометрическому принципу. Система распознает лица на фоне других фоновых объектов, классифицирует особенности лица по форме глаз, носа и рта. VFRS позволяет выделить до 89 характерных параметров, которые могут сличаться при идентификации личности. Длительность процедуры — 1–2 с.

Первыми коммерческими продуктами для распознавания лица и его деталей на изображении стали цифровые камеры и системы охранного видеонаблюдения.

Процесс обработки видеоизображения

Техническое, или машинное, зрение реализует сложный процесс выделения, идентификации и преобразования видеоинформации, который содержит шесть основных этапов:

- получение (восприятие) информации с датчика;
- предварительная обработка изображения;
- сегментация;
- описание;
- распознавание;
- интерпретация.

После восприятия информации в виде визуального изображения производится ее предварительная обработка для снижения посторонних помех, улучшения изображений отдельных элементов объекта или сцены. Затем происходит сегментация, заключающаяся в подразделении сцены на составляющие части или элементы для выделения на изображении интересующих объектов.

Проблемы и методы автоматического распознавания лиц

Идентификация и распознавание лиц — одна из первых практических задач, которая стимулировала становление

и развитие теории распознавания и идентификации объектов. Существует девять категорий объектов, которые вызывают ассоциативные зрительные образы:

- объекты, которыми можно манипулировать;
- объекты, которыми можно частично манипулировать;
- объекты не манипулируемые;
- лица;
- выражения лиц;
- живые существа (животные, фигура человека);
- печатные знаки (буквы, символы, знаки);
- рукописные изображения;
- характеристики и расположение источников света (луна, солнце).

Интерес к процедурам, лежащим в основе процесса узнавания и распознавания лиц, всегда был значительным, особенно в связи с возрастающими практическими потребностями: охранные системы, верификация, криминалистическая экспертиза, телеконференции и т. д. Несмотря на ясность того житейского факта, что человек хорошо идентифицирует лица людей, совсем не очевидно, как научить ЭВМ проводить эту процедуру, в том числе — как декодировать и хранить цифровые изображения лиц. Еще менее ясными являются оценки схожести лиц.

Проблема распознавания лиц рассматривалась еще на ранних стадиях компьютерного зрения. Ряд компаний на протяжении более 40 лет активно разрабатывают автоматизированные, а сейчас и автоматические системы — современные технологии распознавания лиц позволяют производить автоматический поиск и распознавание лиц в графических файлах и видеопотоке.

Алгоритмы распознавания лиц на изображении

Стоит отметить, что коммерциализация продуктов данного сектора стала возможна в первую очередь благодаря появлению на рынке дешевых, компактных датчиков изображения высокого разрешения, а также появлению дешевой элементной базы для цифровой обработки изображения: процессоров с высокой производительностью и большой памятью, дешевой КМОП-ОЗУ, flash-памяти для хранения больших объемов графической информации. Наличие на рынке данных компонентов позволило перейти от чисто исследовательских работ к выпуску массовых и доступных для широкого применения устройств. На рис. 1 показана структура реализации алгоритма распознавания лица человека.

Изображение с видеосенсора (микровидеокамеры) сначала преобразуется в цифровую форму, далее производится фильтрация артефактов изображения, выделение зоны лица,

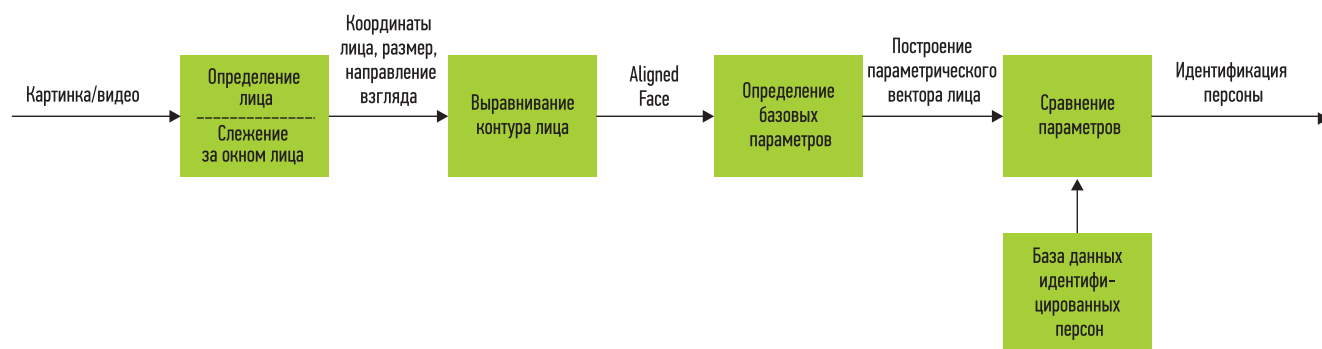


Рис. 1. Алгоритм распознавания лиц на изображении

выравнивание контура лица, выделенного из кадра, построение графической модели лица для параметрической оценки деталей (размеров, положения и т. п.). Характерный набор параметров персоны затем сохраняется в базе данных и может быть использован для идентификации личности

Из истории цифровых камер

Первым цифровым фотоаппаратом считается зеркальная камера Sony Mavica, прототип которой был представлен в 1981 г.: она оснащалась 0,28-Мпикс. ПЗС-матрицей и позволяла делать снимки с разрешением до 570×490 точек, которые хранились на 2" магнитном диске VF. Камера в серию так и не пошла — по причине высокой цены и ограниченных функций. За прошедшие с тех пор более 30 лет цифровые камеры стали сложными электронными устройствами, выполняющими широкий набор функций, включающих фильтрацию шумов, удаление эффекта «красных глаз», конвертирование высококачественных изображений из видеопотока, стабилизацию изображений и видео, способность снимать видео Full HD со стереозвуком, редактирование изображения, выделение человеческих лиц, беспроводную передачу фотографий.

Тем не менее технологии продолжают совершенствоваться, и сегодня невозможно даже предположить, на что будет способен фотоаппарат еще через несколько лет.



персоны. Таким образом, система распознавания состоит из двух компонентов: аппаратной части для захвата, обработки и сохранения изображения и отдельной программы, которая и обеспечивает необходимые процедуры для выделения и идентификации лиц в захваченном кадре изображения.

Инновационные технологии и решения Omron для распознавания лиц

Компания Omron в настоящее время выпускает два продукта для данного сектора: программное обеспечение (ПО) OKAO Vision и аппаратный модуль Human Vision Components (HVC). Это функционально законченный модуль, имеющий все необходимое, в том числе ПО. На выставке Elektra Awards, проходившей 26 ноября 2014 г. в Лондоне, модуль HVC (рис. 2) был удостоен престижной премии за инновационность.

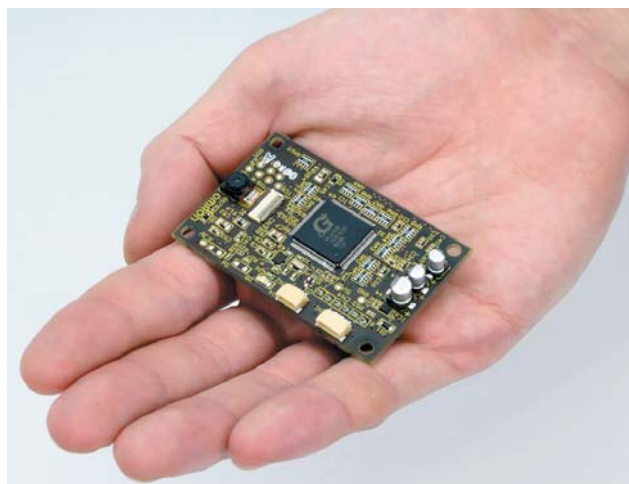


Рис. 2. Модуль HVC

OKAO Vision

Технология OKAO Vision может с успехом применяться в следующих областях:

- системы автоматизации зданий для управления освещением и доступом;
- системы безопасности;
- медицинское оборудование;
- автоматизация производства;
- оборудование для торговых залов с целью регулирования количества открытых касс в зависимости от количества людей, оценки удовлетворенности людей обслуживанием;
- оборудование для маркетинговых исследований — определение количества подошедших к рекламному стенду людей, оценка их реакции на рекламу, фиксация пола и возраста тех, кого привлёк рекламный щит.

Функции, реализуемые OKAO Vision (рис. 3):

- распознавание лица;
- идентификация пользователя;
- распознавание пола пользователя;
- определение возраста;
- определение выражения лица (нейтральное, радостное, удивленное, раздраженное, расстроенное);

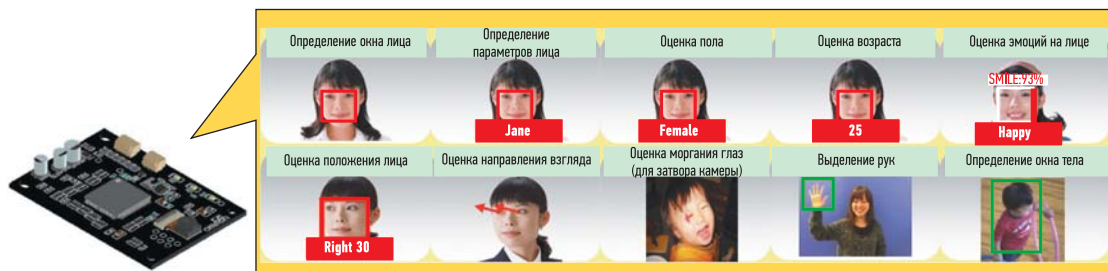


Рис. 3. Основные функции распознавания параметров лица модулем HVC-P

- определение направления взгляда;
- распознавание моргания;
- распознавание жестов рук;
- распознавание человеческого тела и его положения.

Все функции предусматривают отображение вероятности распознавания того или иного факта.

При использовании технологии ОКАО в фотокамерах или смартфонах доступны также и дополнительные функции:

- улучшение качества и однородности кожи лица на фото;
- ретуширование артефактов;
- удаление эффекта «красных глаз»;
- увеличение размера глаз (по установке параметра в меню клиента);
- отбеливание зубов;
- исправление и удаление нежелательных артефактов на лице.

Преимущества технологии:

- широкие возможности для интеграции модуля HVC в любое устройство;
- компактные размеры модуля HVC (ДxШxВ): 60x40x14,2 мм;
- возможность приобретения только ПО;
- 10 уникальных функций распознавания;
- высокая точность распознавания;
- возможность бесплатного предоставления демо-кита во временное пользование.

Перспективы применения ОКАО Vision

ПО ОКАО не требует установки дополнительного оборудования, поэтому эта технология может стать самой доступной системой снятия биометрических данных из всех ныне существующих. Компания Omron разработала алгоритм, позволяющий фиксировать особенности глаз, носа и рта (всего 80 особенностей лица). Создание этого ПО стало возможно благодаря разработанной Omron технологии ОКАО Vision, которая позволила серьезно уменьшить размер рабочих файлов. Таким образом удалось решить проблему ограниченных возможностей сотовых телефонов. Новая программа полностью совместима с самыми распространенными операционными системами для сотовых телефонов — Symbian, BREW, Linux и ITRON. Размер фото при работе сенсора не превышает 1,5 кбайт, минимальны и требования к памяти: ПЗУ — 450 кбайт, ОЗУ — 370 кбайт. На MSM 6500 весь процесс займет примерно 1 с.

Принцип работы системы идентификации доступа довольно прост. Пользователь снимает свое лицо с помощью встроенной камеры, затем технология ОКАО Vision преобразует картинку в набор биометрических данных, для хранения которых требуется всего 1,5 кбайт. Чтобы провести процесс идентификации, необходимо снова сфотографироваться. При этом ОКАО не требует никаких особых настроек камеры — лицо будет распознано в любом случае, если оно поместилось на фотографии. Предполагаемый объем рынка мобильных устройств, в которых может быть использовано ПО ОКАО, по предварительным оценкам составляет 500 млн устройств.

Для реализации 3D-эффектов Amazon.com, американская компания, крупнейшая в мире по обороту среди продающих товары и услуги через Интернет, собирается использовать технологии японской компании Omron Okao Vision, доработанную и адаптированную для работы с операционной системой Android. На тыльной стороне смартфона будут расположены сразу четыре ИК-камеры и одна фронтальная, которые будут распознавать координаты лица пользователя и его черты, определяя пол и возраст.

А в связке с гироскопом и акселерометром система будет мгновенно подстраиваться под пользователя и выводить на экран 3D-картинку, для просмотра которой не требуются специальные очки. Amazon возлагает большие надежды на сторонних разработчиков, которые получают доступ к пакету инструментов для использования возможностей 3D-системы нового смартфона в своих приложениях. Сам же смартфон на момент запуска будет поставляться лишь с парой предустановленных жестов. Лица гарантировано распознаются на дистанциях до 1,3 м от камеры, а тело — на расстоянии до 2,8 м.

Технология распознавания жестов для сектора мобильных устройств

Бесконтактный ввод команд стал уже привычным средством управления, по крайней мере, для тех, кто уже знаком со спортивными играми на приставке Wii или с развлечениями на базе приставки Xbox 360 и системы Kinect. Тем не менее пока эти технологии предусматривают стационарное размещение самого приемника. Японские инженеры предлагают сделать систему подвижной за счет новых алгоритмов и поддержки массового оборудования, включая смартфоны и веб-камеры. Сердцем новой технологии является фирменная программная разработка OKAO Vision, изначально предназначенная для распознавания лиц. На ее основе компания Omron создала систему распознавания жестов, которая может работать и на обычных ПК, и на мобильных устройствах с операционными системами Android или iOS. Именно низкое потребление ресурсов является ключевым фактором в возможном переносе этой технологии на мобильные платформы. Для распознавания жестов в новой системе Omron достаточно снимка руки всего 40 пикселей в высоту. Даже по таким данным технология может определить движение руки вверх, вниз, влево и вправо, а также различные движения пальцев. Жесты распознаются на расстоянии от 10 см до нескольких метров от камеры. При использовании популярного мобильного процессора Snapdragon с тактовой частотой 1 ГГц система распознает жесты с частотой до 30 кадр/с. Важная особенность технологии — анализ взаимосвязи между позицией/движением лица и руки, что позволяет более точно определять намерения пользователя.

Кроме очевидного потенциала новой технологии в смартфонах и планшетах, Omron считает, что она также может найти применение и в других типах техники. Например, подобные системы можно использовать для управления компьютером, для удаленного управления затвором фотокамеры или для переключения каналов на телевизоре.

Модуль HVC

На рисунках 4 и 5 представлены конструкция и рабочий угол обзора встроенной камеры, а также внешний вид модуля.



Рис. 4. Внешний вид модуля B5T-001001(G) HVC

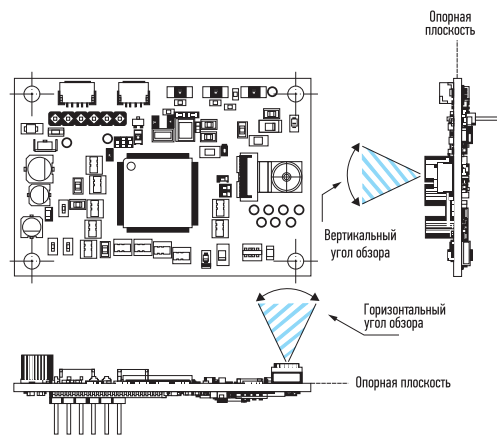


Рис. 5. Конструкция модуля и рабочий угол обзора встроенной камеры

Технические характеристики HVC:

- диапазон рабочих температур 0...+50 °С;
- угол обзора камеры по горизонтали 49°;
- угол обзора камеры по вертикали 37°;
- ток потребления 0,45 А;
- потребляемая мощность 2,25 Вт;
- подключение по интерфейсу UART (3,3 В).

Диапазон рабочих дистанций, на которых эффективно применение HVC, — до 2,8 м. База — не более 1,3 м.

Структура и применение

В состав модуля HVC-P входит модуль камеры с разрешением 640x480 пикс., а также модуль захвата и обработки изображения по алгоритмам, разработанным Omron. Напряжение питания +5 В поступает от пользовательского устройства (рис. 6). Обмен между пользовательской платой, на которой реализовано приложение, и модулем HVC-P производится по специальному протоколу через асинхронный последовательный интерфейс UART. Скорость обмена устанавливает сам пользователь (максимально — 921 600 бод).

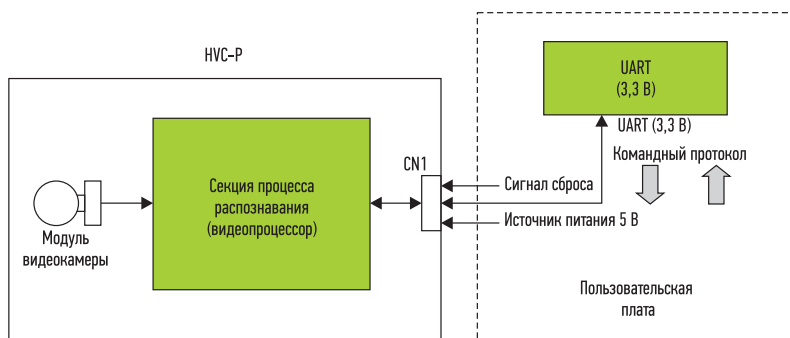


Рис. 6. Структура системы на базе модуля HVC-P

На рис. 7 показаны ориентировочные промежутки времени, затрачиваемые программными модулями на выполнение конкретных процедур распознавания: захват изображения, определение тел персон в кадре, их наличия и положения, обнаружение лиц, определение направления взгляда и эмоциональной окраски выражения лица, примерного возраста и пола.

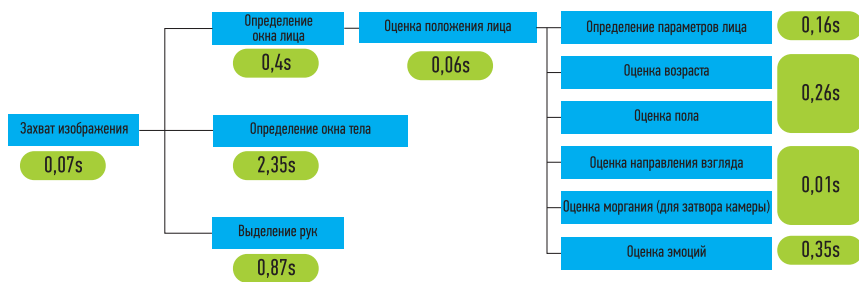


Рис. 7. Время, затрачиваемое на выполнение функций

Протокол обмена «модуль HVC–компьютер»

Протокол обмена двусторонний. Со стороны компьютера (приложения пользователя) в модуль HVC передаются команды режима работы и пороговые значения параметров распознавания — например, размеры окна изображения тела или лица. Из модуля в компьютер передаются данные результата проведенной по команде процедуры распознавания заданных объектов в кадре. Это может быть и само изображение (картинка урезана до форматов 427x320 или 107x80 для уменьшения трафика) или файл результатов идентификации заданных

объектов в кадре (рис. 8). Определяется число персон/лиц (не более 12), координаты центров окон лиц по номерам, вероятный пол и возраст объекта. Передача изображения в устройство пользователя по интерфейсу UART в форматах 320x240 или 160x120 (для экономии трафика интерфейса). При выборе максимальной скорости передачи через UART (921 600 бод) изображение передается за 1 с.

Связь с функциями интерфейса пользователя (например, жестового) обеспечивается посредством ПО ОКАО, которое содержит алгоритмы распознавания. Пользователь получает набор

команд для управления модулем и должен самостоятельно написать программу верхнего уровня для графического интерфейса приложения. Работа со статистикой, установка порогов по вероятности и всего того, что нужно для адаптации функций модуля к конкретным задачам пользователя, должна быть выполнена разработчиком приложения. Разрешение изображения на входе зависит от самой камеры: если применять модуль Omron, то это 640x480, а если это камера заказчика, то используется ее собственное разрешение (можно даже использовать камеру с режимом ночного видения, если необходимо). Масштабируемость изображения в модуле HVC не предусмотрена.

На рис. 9 показан результат работы пользовательского приложения, обеспечившего сортировку фото лиц и создание фотоархива идентифицированных персон с именами с помощью модуля HVC.

Основные области применения модуля HVC

- Промышленное оборудование (в частности, машины для приготовления продуктов);
- управление процессами без уча-

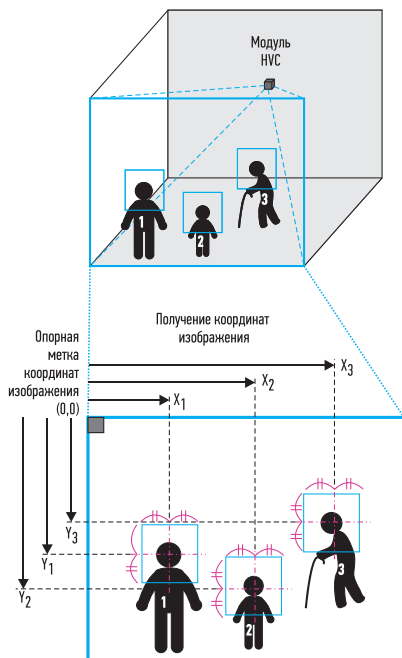


Рис. 8. Результат обработки: координаты окна лица на карте изображения и определение координат окон для тел объектов, попавших в рабочее поле камеры модуля



Рис. 9. Пример идентификации персон с подписью имен в пользовательском приложении

Имеется возможность бесплатного предоставления демо-кита модуля распознавания лиц Omron HVC во временное пользование или его продажа.

В состав демо-кита входят модуль HVC, плата-переходник UART-USB, ПО для работы под Windows. По запросу доступны протоколы сопряжения, 3D-модель модуля B5T, техническая поддержка.

Необычное применение системе распознавания лиц корпорации Omron придумали управляющие японской железнодорожной компании Keihin Electric Express Railway. Они решили приучить своих сотрудников улыбаться чаще. Теперь каждое утро перед началом рабочего дня 530 железнодорожников проходят обязательный тест на «улыбчивость». Программное обеспечение OKAO Vision проверяет выражение лица каждого и оценивает его по шкале от 0 до 100%. Для этого компьютер анализирует контрольные точки, определяющие положение и форму рта, бровей, глаз, носа и контура лица. Затем вся эта информация сравнивается с базой данных, и выдается результат. Тем работникам, которые «не дотянули» до некоторой положительной отметки, компьютер выдает сообщение, которое, по идее, должно стимулировать человека: «Вы выглядите слишком серьезно» или «Поднимите краешки своих губ». Чтобы не растерять полученное таким образом положительное настроение, каждому работнику выдается распечатка его улыбающегося лица, которую он должен хранить весь день, якобы для вдохновения.

ствия рук (Hands-free) посредством использования жестов или направления взгляда;

- распознавание лиц в камерофонах и системах домашних охранных систем;
- робототехника (роботы для выбора своей ответной реакции могут использовать движение рук или выражение на лице пользователя);
- системы пассивной безопасности автомобилей для оценки эмоционального состояния водителя автотранспорта (сон или алкогольное опьянение).

Улыбнитесь, вас снимает скрытая камера

В октябре 2007 г. японская Omron продемонстрировала применение системы OKAO Vision для распознавания и анализа искренней улыбки на человеческом лице. Система распознает такие признаки улыбки, как сужение глаз, изменение формы рта, складки вокруг рта и т.д. Кроме того, определяется пол, примерный возраст человека и выполняется проверка на наличие изображения лица в базе данных фотографий. Изначально эту программу разрабатывали для пользователей цифровых фотоаппаратов. Новинка позволяла камере сделать снимок в то время, когда все позирующие улыбаются, и даже автоматически улучшала фотографию, подкорректировав

цвет лица, например. Получая двухмерное изображение человеческого лица, OKAO Vision составляет его трехмерную проекцию. Анализ проходит в режиме реального времени и занимает около 44 мс при использовании процессора Pentium 4 (3,2 ГГц). Система работает с изображениями величиной от 60 пикс. и имеет специальные инструменты для распознавания плохо освещенных лиц. Для обучения системы инженеры Omron использовали более 10 тыс. изображений лиц со спонтанными, тренированными и притворными улыбками. Технология была продемонстрирована на японской выставке передовых технологий Ceatec (Combined Exhibition of Advanced Technologies). Предполагается, что с помощью OKAO Vision работодатели смогут обучать персонал улыбаться клиентам наиболее «искренне», так как система позволяет моделировать разные улыбки на одном и том же лице. Пригодится новая технология и всем, кто работает с людьми и хочет производить на них наиболее приятное впечатление: менеджерам по продажам, врачам, учителям, политикам. Также новой технологией уже заинтересовались производители цифровых камер. Затвор камеры срабатывает в тот момент, когда фотографируемый улыбается наиболее искренне. Новая система станет существенным шагом вперед в общении человека с машинами. Есть сведения, что в приложении iPhoto '09 также используется ПО Omron.

Другим перспективным направлением, по мнению японских разработчиков системы распознавания улыбок, могут стать маркетинговые исследования. Модули распознавания лиц и улыбок устанавливаются в торговых залах вдоль полок с товарами, а также у кассовых терминалов. Программно производится учет эмоционального состояния покупателей по отношению к определенным товарам и выставляются оценки качества или привлекательности товара. Реакция покупателя, который кладет в корзину эти товары, сохраняется, при этом используется информация о поле и возрасте.

Интерфейсы будущего

Сегодня понятие «интерфейсы будущего» в значительной мере ассоциируется со средствами воздействия на продолжающие оставаться неохваченными органы обоняния и осязания, а также вкусовые рецепторы. Во всяком случае, японские эксперты уже предвидят появление рисоварок, позволяющих пользователям получать представление о вкусе риса в процессе его приготовления. Традиционные средства устройства ввода, как клавиатура или система распознавания речи, будут применяться для получения сведений о текущем эмоциональном состоянии пользователя. Это, в частности, позволит устранить «бестактность» систем голосового общения, встраиваемых, например, в современные бортовые компьютеры автомобилей.



Литература

1. Gabriel Sikorjak. HVC-P: Human Vision Component. Sales presentation. Electronic & Mechanical Components Company, 2014.
2. Модуль B5T-001001(G) Human Vision Components. Datasheet Omron.