

ВЫСШАЯ ШКОЛА МАЙНОР
Институт инфотехнологии
Веб программирование

Александр Мочёнов
IT-3-Q-V-Tal

**Система (? Метод ?) обнаружения и слежения за положением лица
человека на основе нейронных сетей и сегментации по цвету
кожного покрова (? по цвету кожи?)**

Дипломная работа

Руководитель: Jelena Faronova, MSc

Таллинн 2010

Оглавление

Резюме	3
Введение	4
1 Введение в предметную область	6
1.1 Компьютерное зрение	6
1.2 Распознавание и обнаружение лиц	7
1.3 Существующие методы обнаружения лиц	7
1.3.1 Методы предварительной подготовки данных	8
1.3.2 Методы представления изображения	10
1.3.3 Задача классификации	12
1.4 Вспомогательные методы обнаружения лиц	14
2 Предлагаемый метод решения	16
2.1 Модуль нахождения лица	16
2.1.1 Нормализация контраста и баланс белого	16
2.1.2 Поиск зон с цветом кожного покрова	16
2.1.3 Выделение и объединение областей с цветом кожного покрова .	17
2.1.4 Фильтрация по пропорциям и заполненности	17
2.1.5 Классификация	18
2.2 Выбор цели для слежения	19
2.3 Механическая часть	19
2.3.1 описание установки для демонстрации	19
2.3.2 Подсчёт вектора движения	19
2.3.3 Arduino	19
3 Результаты работы (Испытания?)	20
3.1 Автоконтраст и баланс белого	20
3.2 Поиск зон с кожным покровом	20
3.3 Объединение областей	20
3.4 (Результаты) работа с ИНС	20
3.4.1 Различные представления	20
3.4.2 (Результаты) обучения и тестирования	21
3.5 Выбор лица и arduino	21

3.6 Испытание всей системы	21
Заключение и выводы	22
А Приложение. Отчёт по курсовой практике	23
Литература	23

РЕЗЮМЕ

TODO:

ВВЕДЕНИЕ

Роботы в различных вариациях являются частью жизни человека. Робототехника уже давно применяются, например, в индустриальном производстве, в детских игрушках, авиации и многих других местах. Так же роботы применяются в военными (беспилотные самолёты, роботы-сапёры), медицине и даже в космосе¹.

Тем не менее, применение роботов в сфере обслуживания сегодня не так распространено. Оно находится на рубеже науки робототехники и пока ещё широко не применяется. В данной работе автор разрабатывает небольшую часть робота, функционирующего в сфере обслуживания, главной целью которого является общение с человеком.

В частности цель работы - создать интерактивную систему слежения за человеческим лицом подобием головы робота, которая оборудована веб-камерами на месте глаз и серво-приводами, способными поворачивать её по двум осям. Вся система состоит из 3 модулей:

- Нахождение местоположения и размеры лиц людей на изображении с веб-камеры
- Выбор лица из найденных, за которым необходимо следовать
- Вычисления вектора движения и само общение с серво-приводами

Самой сложной из задач является поиск лица человека. В работе автор предлагает последовательный алгоритм поиска, который состоит из 3 подзадач, где результат предыдущей является источником данных для последующего:

- Предварительная обработка и подготовка изображения;
- Поиск, сегментация и кластеризация участков кадра, в которых высока вероятность обнаружения лица;
- Применение искусственных нейронных сетей для окончательной классификации (лицо или нет) по нескольким представлениям данного изображения;

TODO: Почему именно такой?

TODO: Про real-time сюда?

¹<http://robonaut.jsc.nasa.gov/default.asp>

Подобная система может применяться в любых роботах, обладающих подобием головы. Например: робот-консьерж в отеле, робот-официант или робот-домохозяйка. Это может упростить и улучшить впечатление от общения человека с машиной.

TODO: Содержание глав

1 ВВЕДЕНИЕ В ПРЕДМЕТНУЮ ОБЛАСТЬ

1.1 Компьютерное зрение

Основной частью данной работы является обработка изображений поступающих с веб-камеры. Трансформация данных с видео камеры или из статичных изображений в новое представление или принимаемое решение называется - *Компьютерным Зрением* (Computer Vision или CV)(Bradski and Kaehler, 2008) Т.е. программы и алгоритмы, которые в своей работе используют визуальную информацию - всё это компьютерное зрение.

Человеку, в силу своей зрительной природы, может показаться, что обработка визуальной информации - это очень просто. Но эта представление крайне ошибочно. Наш мозг разделяет визуальную информацию на множество каналов, в которых зашифрованы различные виды информации и посылает их мозг человека. В мозгу есть системы распределения внимания, в ходе работы корой, часть информации обрабатывается, а часть остаётся незамеченной. (Bradski and Kaehler, 2008) Даже сетчатка глаза - внутренняя поверхность глаза, заполненная светочувствительными клетками (колбочками и палочками), отвечает за предварительную обработку сигнала. На поверхности глаза около 130 миллионов светочувствительных элементов, а нервных окончаний идущих к мозгу в 100 раз меньше. Это говорит о том, что сетчатка сжимает информацию. В частности одной из её функций является *обнаружение границ* (Edge detection) (RetinaOnWiki)

А что “видит” компьютер? Матрицу из чисел, представляющих собой интенсивность света в разных участках светочувствительной матрицы. При этом каждая ячейка этой матрицы кроме полезной информации содержит ещё и шум. И в этом наборе чисел надо найти машину или, например, идущего человека.

Компьютерное зрение широко применяется в медицине, где оно помогает человеку анализировать визуальные данные. Например по снимку с МРТ¹ указать на опухоль или другие патологии. Это задача называется *распознаванием образов* (Pattern recognition). Распознавать можно так же другие объекты и образы. Например, система автоматического замера скорости на дорогах распознаёт автомобили и их регистрационные номера, а система безопасности распознаёт передвижение людей. (TODO Тут картинку показать?)

¹Магнитно-резонансная томография

Распознавание лица (Face recognition) человека является одной из наиболее популярных задач в области компьютерного зрения. Она заключается в обнаружение и определении по изображению лица кому именно оно принадлежит. Решения этой проблемы применяются в системах безопасности (авторизация) и системах управления базами данных лиц людей. С быстрым развитием более развитых методов в этой области, распознавание лица человека, как средство авторизации, представляет из себя более дешёвое решение по сравнению с системами распознавания сетчатки глаза или отпечатка пальца. (Kumar and Bindu, 2006)

В популярной программе для работы с цифровыми фотографиями Picasa (<http://picasa.google.com/>) есть эффективная система распознавания, маркировки и каталогизирования найденных на фотографиях лиц людей. (TODO Тут картинку показать?)

1.2 Распознавание и обнаружение лиц

Но для того, что бы распознать лицо, сначала необходимо найти его место положение и границы на изображении. Эта задача называется *обнаружение лица* (Face detection), что является частным случаем более общей проблемы *обнаружение объекта* (Object detection). Почти все алгоритмы распознавания лиц в качестве входных данных используют изображение, содержащее только лицо, которое надо распознать. По-этому обнаружение лица есть предварительная и очень важная задача, которую надо выполнить, перед распознаванием. Следовательно от точности и быстроты определения местоположения лица зависит эффективность всей задачи по распознаванию.

Но, обнаружение не обязательно должно вести к распознаванию. Обнаружение и слежение за лицом без определения принадлежности его к конкретному человеку является основной задачей как данной работы, так и схожих по своей сути работ: (Carp et al., 2010), (Luo et al., 2007), (Saxena et al., 2008).

Нахождение лица заключается в том, что бы по данному изображению определить, количество, место положения и размеры всех имеющихся лиц. (Liu et al., 2010) (TODO Тут картинку показать?)

1.3 Существующие методы обнаружения лиц

Почти все современные подходы к решению задачи обнаружения лица, по мнению автора, содержат так или иначе 3 составляющих:

(TODO Не не. способы тут - 15, 14

Подготовка данных

Для успешного обнаружения лица поступившие данные необходимо подготовить для дальнейшей работы. Сюда входят и различные методы *предварительные обработки* изображения в целом (т.н. preprocessing) и методы позволяющие уменьшить область поиска или *область интереса* (Region of interest или ROI) для ускорения всего процесса обнаружения.

Представление данных

Изначальное изображение в виде матрицы интенсивностей светочувствительного элемента камеры часто трансформируют в иные, более компактные отображения или *представления* (Representation). Такие изменения чаще всего ведут к потере информации, но облегчают процесс классификации.

Классификация

Само определение наличия или отсутствия в данном участке картинки лица человека. Методов классификации в принципе (не только лиц) на сегодняшний день существует огромное количество и все они применимы для данной задачи.

1.3.1 Методы предварительной подготовки данных

Коррекция цвета и освещённости

Для эффективной работы с изображением его необходимо подготовить. (О проблеме освещения в принципе)

Некоторые механизмы применяются для всего изображения в целом. Примером таких изменений могут служить коррекция контраста, интенсивности и баланса белого. Другие подготовительные процессы касаются изображений, которые поступают непосредственно в классификатор. Речь идёт о небольших частях изображения, которые получают методов скольжения окна (sliding window) и масштабирования, которые называют *образцами* (об этом подробнее в (TODO)). Процессы коррекции контраста и баланса белого часто накладывают и на образцы, т.к. зачастую после исправления контрастности всего изображения, распределение интенсивности в образце всё ещё остаётся узким. (TODO Картинку с тусклым изображением и после наложения контраста?)

В работе (Rowley et al., 1998), в частности, применяется два метода предварительной обработки образцов. Во-первых производится нелинейное *уравновешивание гистограммы* (Histogram equalization), которое “растягивает” гистограмму, что компенсирует недостающие уровни интенсивности изображения, в результате чего изображение становится более контрастным. Во-вторых производится линейная компен-

сация интенсивности. Для этого линейная функция аппроксимирует общую яркость в каждой из частей образца, после чего она может быть вычтена из образца в случае сильной разницы в освещённости разных частей образца. Это приводит к уравниванию контрастности в разных частях образца. (TODO Пример работы такого подхода)

Похожий алгоритм нормализации интенсивности применяется и в работе (Lin et al., 2005). Перед классификацией каждый образец нормализуется по формулам 1.1 и 1.2.

$$\bar{I} = \frac{1}{N} \sum_{i=1}^N I_i \quad (1.1)$$

$$I'_i = (I_i - \bar{I}) + 128 \quad (1.2)$$

где

- N – количество пикселей в образце
- \bar{I} – средняя интенсивность по всем пикселям образца
- I_i – интенсивность i 'ого пикселя образца
- I'_i – нормализованная интенсивность

Таким образом если какие-то образцы были слишком тёмными или слишком светлыми они все становятся единообразно освещены, что упрощает процесс машинного обучения. (TODO Пример. Картинка)

В данной работе применяются схожие методы предварительной обработки образцов и изображения в целом. Подробнее в разделе(TODO)

Сегментация по цвету кожного покрова

Чем меньше изображение необходимо сканировать на наличие лица, тем меньше процессорного времени необходимо затрачивать, что ведёт к ускорению процесса обнаружения. Это можно достигнуть, например, уменьшением изображения, что приведёт к потере разрешающей способности данного подхода. Одним из способов позволяющих уменьшить площадь сканирования и одновременно не потерять разрешающей способности является эвристическое знание о том, что все лица людей покрыты кожным покровом. Это можно использовать для нахождения областей изображения содержащих цвета схожие с цветами кожного покрова (skin-color detection), и в дальнейшем осуществлять сканирование только этих областей.

Среди методов обнаружения лиц основанных на поиске признаков (feature based) подходы использующие информацию о цвете кожи, как признак обнаружения, получают всё большую популярность. Цвет легко и быстро обрабатывать и он инвариантен

к геометрическим особенностям образов лиц людей. К тому же опыт подсказывает, что цвет кожи человека имеет отчётливый характерный цвет, который легко узнаваем людьми. (Vezhnevets et al., 2003) Цвет кожи - это один из тех признаков, что не зависит от положения лица, частичной закрытости и контраста, по-этому именно этот метод часто используют для локализации лиц. (Ruanyam and Covavisaruch, 2009)

Такой метод возможен благодаря тому факту, что различные цветовые вариации кожных покровов людей (даже среди представителей различных этнических групп) лежат в достаточно узком диапазоне и отличаются только яркостью (Luminance), в то же время цветность (Chrominance) практически не меняется. Цвет кожи лежит в основном в красной части цветового спектра и определяется цветом крови. А яркость определяется прозрачностью эпидермиса (верхнего слоя кожи), за что в свою очередь отвечает концентрация меланина. (Xu and Zhu, 2006) Подробнее в разделе ... (TODO)

1.3.2 Методы представления изображения

Одно и тоже изображение можно представить разными способами. Это нужно для того, что бы обучаемые классификаторы (см. 1.3.3) могли обучаться на различных характеристиках или признаках, имеющихся в различных представлениях.

Пиксельная интенсивность

Самым простым и популярным представлением является информация об интенсивности в каждом пикселе изображения. Такой метод представления считается не имеющим потерь (lossless). Т.е. в нём присутствует вся информация об оригинальном изображении. (Bojkovic and Samcovic)

Информация о контурах

Любой метод обнаружения лица (да как и все другие алгоритмы в компьютерном зрении) должен быть стойким к таким вещам как поза голов, угол обзора, освещение и многим другим факторам. Для решения проблемы с освещением используется информация о контурах (edginess) или градиент изображения. Одно и то же лицо под разными источниками света с точки зрения пиксельного представления совершенно отличны друг от друга, что делает проблемой для классификатора найти среди них нечто общее. С другой стороны информация о пограничных областях в большей степени неизменна. (Ahmadyfard et al., 2008)

Это представление, в отличие от пиксельного представления, не содержит всей изначальной информации. Сохраняется информация лишь о пограничных контрастных зонах лица (глаза, брови, губы)

Собственные лица

Собственное лицо (eigenface) - это набор *собственных векторов* (eigen vectors), используемых для описания “стандартизированных компонентов лица”. Один образец лица принимается как точка в многомерном пространстве. Много образцов образуют некую область в этом пространстве. Задача заключается в нахождении *главных компонент* (principal components) этой области, параметрами которых её можно описать используя значительно меньшее количество переменных, нежели для описания всех точек изначального пространства. (Turk and Pentland, 1991)

Собственный вектор представляет из себя набор параметров, которыми можно описать лицо человека. Их же можно хранить например в базе данных, или использовать в качестве представления группы изображений. Этот набор параметров является своего рода выжимкой из многих тренировочных лиц и может использоваться для дальнейшей классификации.

В работе (Tsai et al., 2006) автор для обнаружения лиц использует собственные лица для поиска кандидатов областей с лицами, а нейронная сеть проводит конечную валидацию.

Характеристики типа Хаара

Характеристики типа Хаара (Haar-like features) представляют изображения при использовании платформы по обнаружению объектов Виола-Джонса (Viola-Jones object detection framework). (Viola and Jones, 2001) Этот метод и применяемое в нём представление является одним из самых быстрых на сегодняшний день методов обнаружения объектов.

Использование нескольких представлений

Многие решения обнаружения лица используют сразу несколько представлений. Это позволяет использовать более широкий спектр характеристик, что делает классификатор более чувствительным к особенностям конкретного вида объектов (в данном случае к лицам).

Так в работе (BojkoVIC and Samcovic) используются три представления: пиксельное, коэффициенты собственных лиц и профильные коэффициенты. В работе (Ahmadyfard et al., 2008) авторы используют две представления: пиксельное и информацию о контурах. По их словам “объединение информации об интенсивности и о контурах даёт более описательные характеристики для представления изображения с лицом”.

1.3.3 Задача классификации

Задача *классификации* - это проблема определение класса из всех возможных, к которому относятся классифицируемые объекты или наблюдения. Классификация тесно связана с *машинным обучением* (Machine Learning или ML), задачей которого является превращение данных в информацию (Bradski and Kaehler, 2008)

Задача обнаружения лица является классическим примером задачи классификации, где есть всего два класса “лицо” или “не лицо”. Классификатор должен по данному ему изображению уметь определить к какому классу из двух оно относиться.

Все методы можно грубо разделить на две части: классификаторы основанные на знаниях (knowledge-based) и статистически обучаемые. Основанные на знаниях используют эмпирические знания о местоположении и распределении органических характеристик лица. (BojkoVIC and Samcovic) Например, факт, что глаза находятся в верхней половине картинки и они симметричны, а под ними находится рот, с которым они образуют равнобедренный треугольник. Такие алгоритмы легко реализовать, но они неустойчивы ко многим факторам (например наклон головы).

К статистически обучаемым алгоритмам классификации относятся те, что получают знания из данных при обучении с учителем или без него. Т.е. работа таких алгоритмов состоит из двух этапов: обучение и само использование. Самих методов очень много. Ниже приведены те, что чаще всего встречаются в обнаружении лиц.

Метод опорных векторов

Метод опорных векторов (Support vector machine или SVM) заключается в нахождении разделяющей гиперплоскости в более высокомерном пространстве, чем классифицируемые вектора. Конечная плоскость должна быть максимально удалено от представителей разных классов. Для этого находятся наиболее близкие к этой плоскости и тем самым более влиятельные векторы, которые называют опорными векторами. (Shavers et al., 2006)

Ограничением SVM подхода является проблематичность использования для классификации более чем по двум классам. Впрочем в данной задаче речь как раз и идёт о двух классах и SVM хорошо справляется с ней во многих работах. (Shavers et al., 2006), (Jee et al., 2004), (Saxena et al., 2008)

Искусственные нейронные сети

Одним из самых популярных (но не самым эффективным) методом классификации до сих пор является искусственные нейронные сети (далее ИНС, Artificial Neural Network или ANN). Этот способ применяется в большинстве научных работ изученных в ходе подготовки данной работы, что стало одним из факторов выбора её в качестве классификатора.

Появление ИНС было вдохновлено биологическими системами, в частности нервной системой. До определённой степени структура и функциональность ИНС напоминают структуры нейронных связей головного мозга. Биологические нейроны это - нервная клетка, у которой есть сеть окончаний *дендритов*, через которые она получает сигнал от других нейронов. В ядре клетки происходит суммирование всех сигналов, после чего он передаётся по длинному выходному окончанию *аксону* на входы следующих нейронов. В ИНС всё похоже. У нейрона есть входной вектор (сеть дендритов), все значения которых умножаются на вес соответствующей связи и суммируются, после чего суммарный сигнал проходит через *активационную функцию* (ядро) и результат передаётся к последующему нейрону как один из компонентов его входного вектора. (тут рисунок, где рядом нейрон живой и искусственный)

Определяющими факторами работы ИНС являются: топология или структура сети, виды активационных функций и веса соединений. Веса соединений - это параметры сети. Они меняются во время обучения сети, после чего сеть может описывать модель, которую её учат классифицировать. Главной задачей активационной функции является сжатие сигнала до определённых границ, а так же усиление слабых сигналов и гашение сильных. Самой популярной функцией является сигмовидная (TODO картинка). Так же она обладает простой производной, что очень удобно во время обучения.

Концептуально есть две структуры ИНС: с прямым распространением сигнала (feedforward) и рекуррентные (т.е. с обратными связями). Второй тип применяется для обучения без учителя (unsupervised learning) и в обнаружении используется реже, чем первый тип, к которому применяют метод обучения с учителем (supervised learning).

Популярным методом обучения ИНС с прямым распространением является алгоритм *обратного распространения ошибки* (backpropagation) “Он имеет солидное мате-

матическое обоснование. Несмотря на некоторые ограничения, процедура обратного распространения сильно расширила область проблем, в которых могут быть использованы искусственные нейронные сети, и убедительно продемонстрировала свою мощь.” (Wasserman, 1989) Принцип метода такой: сначала на вход сети подаётся входной вектор, который проходит через сеть и сеть выдаёт результат; разница полученного результата и заранее известного необходимого выхода сети посылается в обратном направлении от выходов к входам; по мере прохождения высчитывается ошибка по методу наименьших квадратов; по этим ошибкам методом градиентного спуска высчитываются изменения весов соединений. Таким образом с каждым новым образцом веса меняются всё лучше описывая желаемую модель.

У ИНС метода есть и проблемы. Что бы получить результат выше среднего сеть должна быть очень точно отрегулирована (число слоёв, число узлов, скорость обучения и т.д.) (Jee et al., 2004) Авторы работы (Capi et al., 2010) так же считают, что эффективность работы зависит от качества подстроки всей системы, и что это трудоёмкая работа. Результат данной работы лишний раз подтверждает эти выводы.

1.4 Вспомогательные методы обнаружения лиц

На тему обнаружения лиц написано огромное количество научных работ. В базе данных IEEE Xplore за период с 2000 года можно найти свыше 6000 работ, так или иначе связанных с данной тематикой. В каждой из них есть какая-то своя, оригинальная идея.

Так в работе (Jee et al., 2004) одной из главных характеристик поиска являются глаза, которые обнаруживаются на изображении при помощи SVM классификатора.

В работе (Capi et al., 2010) решается проблема, подобная основной задаче данной работы, а именно создание навигационной системы робота, способной следовать за человеком. Для этого сначала обнаруживается лицо человека, после чего определяется основной цвет одежды человека. Далее человек может развернуться, и робот будет следовать за ним ориентируясь по одежде, а не по лицу.

Похожие на предыдущую и данную работы стоят задачи в работе (Luo et al., 2007), где решают проблему взаимодействия человека и робота. Как и в данной работе стоит задача обнаружения лица человека и дальнейшего слежения за ним. Особенностью данной работы является способ уменьшения области интереса, в которой высока вероятность обнаружения лица. Для этого информации о положении лица берётся из предыдущих кадров и высчитывается траектория движения лица и прогнозируемое следующее положение. После эта область расширяется на константу и в полученном месте изображения производится уже поиск лица.

В работе (Zhang and Liang, 2010) применяется простая идея о том, что лицо человека, как и сама голова, скорее всего в кадре будет двигаться. Так для области интереса можно определить все движущиеся области. Но, такой метод предполагает статичную камеру и следовательно статичный фон.

(про real-time)

2 ПРЕДЛАГАЕМЫЙ МЕТОД РЕШЕНИЯ

Краткое описание всей системы. Диаграмма.

Модульность. Какие-то части могут быть реализованы по разному - но сами модули такие как тут.

используемые программные библиотеки: OpenCV, PyBrain, ...

2.1 Модуль нахождения лица

2.1.1 Нормализация контраста и баланс белого

(предобработка)

Описание алгоритма.

Много примеров, гистограмм, псевдокод. (без особых результирующих картинок)

2.1.2 Поиск зон с цветом кожного покрова

Общая информация о проблеме.

- зачем применяют

—про поиск картинок для взрослых

—для face detection

Возможные пути решения

Уже существующие различные методы нахождения цвета кожного покрова. Описание тут.

Проблема выбора цветового пространства

Про цветовые пространства. Про информативность каждого из них.

Метод статического диапазона

В работе реализуется он. Почему? (просто, быстро, достаточно эффективный) Описание метода, псевдокод?

Сравнение двух реализованных моделей.

2.1.3 Выделение и объединение областей с цветом кожного покрова

Общие слова переходного характера.

Выделение найденных областей

Описание алгоритма (сжатие, расширение) - избавление от шума, более адекватные замкнутые области.

Кластеризация

Обоснование необходимости. Зачем объединять.

Почему обычный k-mean не подходит? примеры.

Описание метода кластеризации через минимальное оставное дерево. Что такое оставное дерево?

Примеры мест (ситуаций), где это необходимо. Где лицо состоит из нескольких небольших участков и полезно объединять.

Псевдокод, диаграммы процесса.

2.1.4 Фильтрация по пропорциям и заполненности

Описание возможного постпроцессинга для отфильтровывания неподходящих участков.

2.1.5 Классификация

Описание проблемы классификации в целом. Опять о том какие методы бывают. О том, что сейчас применяют чаще.

Выбор метода ИНС для классификации

Почему выбрал ann? (real-time, простота понимания и использования)

Как это делают другие?

Описание сети

Несколько сетей для разных представлений. V/w, Edges

О проблемах недофитинга и overfit'инга.

Описание структуры ИНС. Почему именно такая.

bias'ы, преждевременная остановка, ...

Код с созданием сети.

Обучение сети

Первый этап.

Проблема и важность выбора примеров для обучения.

Применяемые базы лиц, усреднённые лица, возможная рамочка

Проблема выбора "не лиц".

Описание подготовки выборок для тренировки и тестирования.

Код PyBrain по тренировке сети.

Применение сети

Сохранение и загрузка обученной сети.

Понятие порога.

sliding window алгоритм. диаграммы, код.

Кластеризация всех найденных лиц в группы, что бы отсеять случайные False positives. Overlap'ы и всё такое.

2.2 Выбор цели для слежения

найти наибольшее лицо

искать ближайшее к тому, за которым уже следим

2.3 Механическая часть

Работа с сервоприводами

2.3.1 описание установки для демонстрации

arduino, сервоприводы, камеры

2.3.2 Подсчёт вектора движения

2.3.3 Arduino

коммуникация с PC

листинги кода, диаграммы (этого нет =/)

3 РЕЗУЛЬТАТЫ РАБОТЫ (ИСПЫТАНИЯ?)

3.1 Автоконтраст и баланс белого

Когда работает? А когда не очень? примеры и того и того.

Возможные пути решения.

3.2 Поиск зон с кожным покровом

Когда работает? А когда не очень? примеры и того и того.

Пути решение. (Другой метод, выбор более узкой области диапазонов. - пример возможного приложения для сбора образцов)

О том что хорошо, что больше false negative, чем false positive

3.3 Объединение областей

Примеры хорошей и плохой работы.

Как можно улучшить. (выбор другого алгоритма выбора цвета кожи, подгонка параметров кластеризации)

3.4 (Результаты) работа с ИНС

3.4.1 Различные представления

Почему представление с пограничными областями не работает. Усреднённые морды где видно проблему. Как-то улучшить алгоритм выявления пограничных областей? Какие-то другие представления?

3.4.2 (Результаты) обучения и тестирования

Цифры, проценты результативности на тестовых данных. False positive, False negative. Примеры неузнанных лиц, примеры узнанных не лиц. Усреднённые нелица, усреднённые ненайденные лица. Как можно улучшить?

- Правильная структура сети, выделяющая характеристики (features)
- Более тщательный подбор примеров (глаза на одном месте, одна ориентация)
- Разные классы для разных поз (направление взгляда - прямо, вверх, вниз, вправо, влево)
- Икрементальный процесс обучения (где все false из тестового набора добавляются обратно в набор обучения)

3.5 Выбор лица и arduino

Так и не успел закончить эту часть. Что писать в результатах пока не знаю.

3.6 Испытание всей системы

результаты испытаний.

небольшие ошибка на всех уровнях в итоге дают неудовлетворительный результат. улучшать необходимо каждый из элементов в отдельности.

- о проблеме 2-7 процентах на 97 000 примерах с одного кадра. Много false positive. Сложно настроить правильные порог.
- о проблеме со скоростью. Решение - оптимизация и использование pyuco.

ЗАКЛЮЧЕНИЕ И ВЫВОДЫ

А ПРИЛОЖЕНИЕ. ОТЧЁТ ПО КУРСОВОЙ ПРАКТИКЕ

Литература

- A. Ahmadyfard, B. Yousefi, and S.M. Mirhassani. A Hierarchical Fuzzy Based Approach for Face Detection. In *Bioinformatics and Biomedical Engineering, 2008. ICBBE 2008. The 2nd International Conference on*, pages 2327–2330. IEEE, 2008.
- Z. Bojkovic and A. Samcovic. Face Detection Approach in Neural Network Based Method for Video Surveillance. In *Neural Network Applications in Electrical Engineering, 2006. NEUREL 2006. 8th Seminar on*, pages 44–47. IEEE. ISBN 1424404339.
- G.R. Bradski and A. Kaehler. *Learning opencv*. O'Reilly, 2008. ISBN 0596516134.
- G. Capi, H. Toda, and T. Nagasaki. A vision based robot navigation and human tracking for social robotics. In *Robotic and Sensors Environments (ROSE), 2010 IEEE International Workshop on*, pages 1–6. IEEE, 2010.
- H. Jee, K. Lee, and S. Pan. Eye and face detection using SVM. In *Intelligent Sensors, Sensor Networks and Information Processing Conference, 2004. Proceedings of the 2004*, pages 577–580. IEEE, 2004. ISBN 0780388941.
- CNR Kumar and A. Bindu. An efficient skin illumination compensation model for efficient face detection. In *IEEE Industrial Electronics, IECON 2006-32nd Annual Conference on*, pages 3444–3449. IEEE, 2006. ISBN 1424403901.
- H.J. Lin, S.Y. Wang, S.H. Yen, and Y.T. Kao. Face detection based on skin color segmentation and neural network. In *Neural Networks and Brain, 2005. ICNN&B'05. International Conference on*, volume 2, pages 1144–1149. IEEE, 2005. ISBN 0780394224.
- X. Liu, G. Geng, and X. Wang. Automatically face detection based on BP neural network and Bayesian decision. In *Natural Computation (ICNC), 2010 Sixth International Conference on*, volume 3, pages 1590–1594. IEEE, 2010.
- RC Luo, AC Tsai, and CT Liao. Face Detection and Tracking for Human Robot Interaction through Service Robot. In *Industrial Electronics Society, 2007. IECON 2007. 33rd Annual Conference of the IEEE*, pages 2818–2823. IEEE, 2007. ISBN 1424407834.
- RetinaOnWiki. Retina on wikipedia.org, 2011. URL <http://en.wikipedia.org/wiki/Retina>.
- H.A. Rowley, S. Baluja, and T. Kanade. Neural network-based face detection. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 20(1):23–38, 1998. ISSN 0162-8828.

- P. Ruangyam and N. Covavisaruch. An efficient region-based skin color model for reliable face localization. In *Image and Vision Computing New Zealand, 2009. IVCNZ'09. 24th International Conference*, pages 260–265. IEEE, 2009.
- V. Saxena, S. Grover, and S. Joshi. A real time face tracking system using rank deficient face detection and motion estimation. In *Cybernetic Intelligent Systems, 2008. CIS 2008. 7th IEEE International Conference on*, pages 1–6. IEEE, 2008.
- C. Shavers, R. Li, and G. Lebby. An SVM-based approach to face detection. In *System Theory, 2006. SSST'06. Proceeding of the Thirty-Eighth Southeastern Symposium on*, pages 362–366. IEEE, 2006. ISBN 0780394577.
- CC Tsai, WC Cheng, JS Taur, and CW Tao. Face detection using eigenface and neural network. In *Systems, Man and Cybernetics, 2006. SMC'06. IEEE International Conference on*, volume 5, pages 4343–4347. IEEE, 2006. ISBN 1424400996.
- M. Turk and A. Pentland. Eigenfaces for recognition. *Journal of cognitive neuroscience*, 3 (1):71–86, 1991. ISSN 0898-929X.
- V. Vezhnevets, V. Sazonov, and A. Andreeva. A survey on pixel-based skin color detection techniques. In *Proc. Graphicon*, volume 3. Citeseer, 2003.
- P. Viola and M. Jones. Rapid object detection using a boosted cascade of simple features. 2001. ISSN 1063-6919.
- P.D. Wasserman. *Neural computing: theory and practice*. Van Nostrand Reinhold Co. New York, NY, USA, 1989. ISBN 0442207433.
- Z. Xu and M. Zhu. Color-based skin detection: survey and evaluation. In *Multi-Media Modelling Conference Proceedings, 2006 12th International*, pages 10–pp. IEEE, 2006. ISBN 1424400287.
- L. Zhang and Y. Liang. A fast method of face detection in video images. In *Advanced Computer Control (ICACC), 2010 2nd International Conference on*, volume 4, pages 490–494. IEEE, 2010.