электронный журнал
МОЛОДЕЖНЫЙ НАУЧНО-ТЕХНИЧЕСКИЙ ВЕСТНИК
Издатель ФГБОУ ВПО "МГТУ им. Н.Э. Баумана". Эл No. ФС77-51038.

УДК 681.327.1

 # 12, декабрь 2015

Проектирование системы распознавания динамического поведения людей

***Демин Н.А.**, студент*
Россия, 105005, г. Москва, МГТУ им. Н.Э. Баумана,
кафедра «Информационные системы и телекоммуникации»

Научный руководитель: Алфимцев А.Н., к.т.н, доцент
Россия, 105005, г. Москва, МГТУ им. Н.Э. Баумана,
кафедра «Информационные системы и телекоммуникации»
iu3@bmstu.ru

Введение

В последние годы область цифровой обработки сигналов интенсивно развивается. В особенности, большой прогресс наблюдается в цифровой обработке изображений и видеосигналов. Одно из главных направлений развития цифровой обработки – это охранное видеонаблюдение. Необходимость в обеспечении безопасности привела к широкому распространению систем безопасности и соответствующим исследованиям в данной области, многие из которых связаны с разработкой разнообразных алгоритмов детектирования, применяющихся в охранных системах. К ним относится и автоматическое распознавание людей в видеопотоке.

Идеальный детектор человеческих фигур должен обнаружить всех присутствующих на цифровом изображении людей независимо от их расположения. В связи с этим, при решении задачи автоматического детектирования человека возникает множество проблем. Система должна успешно справляться с задачей распознавания и подсчетом человеческих фигур под воздействием различных негативных факторов (изменения освещения, различные позы людей, перекрытия фигур людей и т.д.).

Анализ существующих алгоритмов

Задача детектирования фигуры человека на настоящий момент очень актуальна, а также алгоритмически сложна. В связи с этим, существует множество разнообразных подходов к ее решению. Для их классификации возможно использовать несколько различных критериев. В представленной работе будет рассматриваться классификация

алгоритмов распознавания по используемой модели (см. рис. 1). Данный подход позволяет наиболее широко и детально обозреть существующие алгоритмы, определить сферу применения, а также выявить их сильные и слабые стороны.

Исходя из выбранного подхода к классификации алгоритмов распознавания, существуют два глобальных класса данных алгоритмов: основанные на выявлении движения и основанные на анализе и распознавании образов. Последний подразделяется на два подкласса: распознавание силуэтов человека и анализ областей изображения. Анализ областей основан на использовании информации о внешнем виде объектов изображения и делится на два класса. Первый класс – это системы, основанные на поиске силуэта, которые фокусируются на выделении контуров объектов и на основе полученных данных создают или обновляют классификационную модель. Второй класс – системы, основанные на анализе областей, которые не требуют применения алгоритма выделения контуров, но производят поиск участков изображения, которые могут быть сходными с искомыми объектами, после чего приводятся в действие те или иные классификаторы.

Одним из наиболее популярных настоящий момент подходов к детектированию объектов на изображении путем анализа областей изображения является метод бегущего окна. Его принцип заключается в последовательном сканировании всего изображения блоками различных размеров с целью выявить искомый объект. Самыми известными примерами реализации данного подхода являются каскады Хаара и HOG-дескриптор.

Алгоритмы с распознаванием движения основываются на сравнении кадров в видеопотоке и выявлением областей с совпадающим вектором перемещения. Дальнейший анализ может проводиться с целью выявить закономерность движения, присущую человеку, либо для того, чтобы определить, соответствует ли внешний вид движущегося объекта человеческой фигуре. Распознавание движения может использоваться также и в качестве вспомогательного инструмента, например, как предварительный этап для локализации области интереса перед применением других методов анализа изображения.

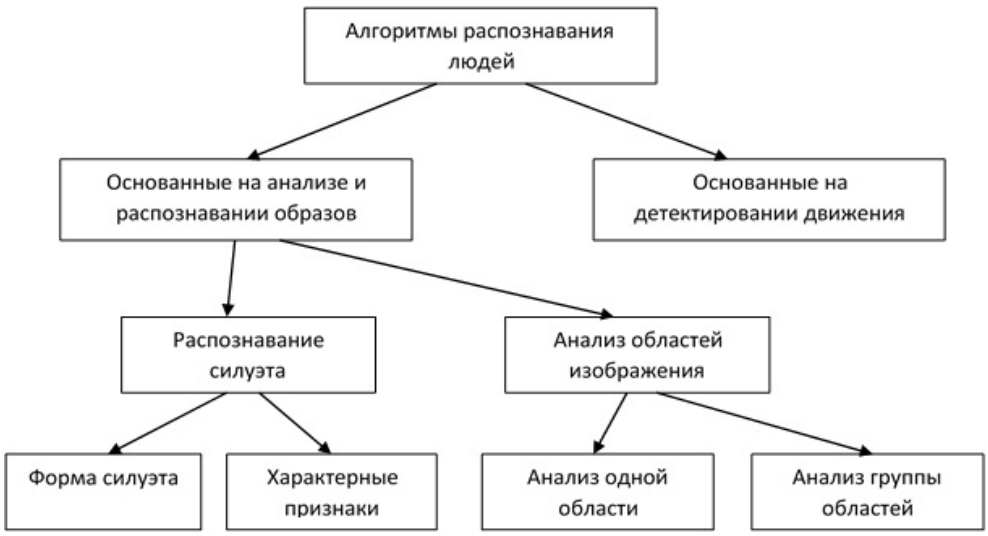


Рис. 1. Классификация алгоритмов распознавания людей по модели, лежащей в их основе

Подход к распознаванию человека через обнаружение его лица является одним из самых точных путей решения проблемы распознавания людей, однако вся точность сходит на нет в том случае, если отсутствует гарантия того, что все присутствующие в видеопотоке люди обращены лицом к камере. Однако некоторые из рассмотренных алгоритмов распознавания лиц могут быть адаптированы под распознавание фигуры в целом. В особенности это относится к обучаемым алгоритмам, к примеру, нейронной сети.

В таблице приведены результаты сравнения по основным рассмотренным параметрам: точность распознавания, устойчивость к изменению освещения, алгоритмическая сложность и необходимость обучения. Также указаны базовые методы, на которых основывается каждый алгоритм, и классификация распознанных объектов.

Методы на основе детектирования движения проигрывают по эффективности и точности методам, анализирующим внешний вид изображения, также они не справляются с задачей распознавания в случае перекрытия объектов. Однако, данный класс методов нечувствителен к изменениям окружающей среды. Исходя из этого, алгоритмы детектирования движения выгодно применять в качестве дополнительного шага в тех случаях, когда не справляются алгоритмы анализа внешнего вида.

Что касается алгоритмов, использующих анализ цельного силуэта, результаты их применения в распознавании точны, однако их недостаток в высокой вероятности возникновения ошибок при выявлении силуэта целиком, так как во многих случаях его

получение осложнено факторами окружающей среды (условия изменяющегося освещения, наличие теней, активные перемещения объектов и т.п.).

Системы распознавания на основе анализа отличительных черт силуэта показывают худшие результаты в точности распознавания по сравнению с методами, основанными на поиске цельного силуэта, однако обладают меньшей вычислительной сложностью. Анализ отличительных черт силуэта является быстрой и простой альтернативой поиску силуэта целиком.

Методы, анализирующие области изображения, являются самым мощным средством распознавания, но в то же время затратным по вычислительным мощностям. Нейронные сети требуют больших усилий по настройке и обучению, а также требовательны к вычислительным ресурсам. То же самое относится и к детектору на основе каскадов Хаара. Что касается HOG-дескриптора, он проигрывает остальным в точности распознавания.

Алгоритм	Классификация объектов	Поддержка перекрытия объектов	Точность распозна-вания	Устойчивость к изменению освещения	Алгоритми-ческая сложность
Детектирова-ние паттернов движения	Человек, посторонний объект	нет	низкая	высокая	средняя
Вычитание фона	--	нет	--	высокая	низкая
Выделение характерных черт силуэта	Человек, группа людей, транспорт	нет	низкая	низкая	низкая
Поиск цельного силуэта	Человек, посторонний объект	нет	средняя	низкая	низкая
Поиск частей тела	Человек, посторонний объект	да	средняя	низкая	средняя
Каскады Хаара	Лицо, человек, посторонний объект	нет	высокая	низкая	средняя

HOG-дескриптор	Человек, посторонний объект	нет	средняя	низкая	средняя
Нейронная сеть	Лицо, человек, посторонний объект	да	высокая	средняя	высокая

Таким образом, можно сделать вывод, что наилучших результатов в распознавании можно достигнуть путем комбинирования различных методов, каждый из которых успешно справляется с определенной задачей, но не может быть использован отдельно в связи с накладываемыми ограничениями. Таким образом, полученный алгоритм будет обладать преимуществами каждого используемого метода, и вследствие этого будет более универсальным.

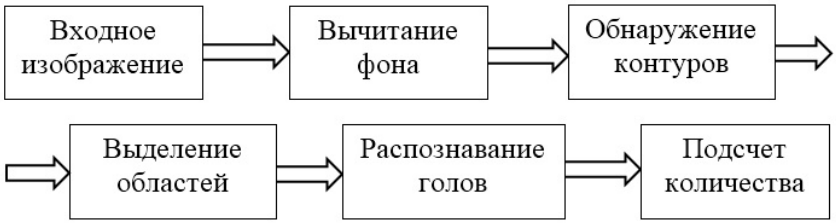


Рис. 2. Структура алгоритма распознавания и подсчета

Разработка алгоритма распознавания и подсчета людей

На рис. 2 представлена схема разработанного алгоритма. Первым этапом применяется вычитание фона для локализации операционной области, затем алгоритмы выделения контуров и областей. Обнаружение контуров способствует более точному выделению областей интереса и отсечению ненужных фрагментов изображения. В завершение применяется распознавание лиц и голов людей, после чего производится подсчет их количества.

Для повышения точности и быстродействия детектора людей необходимо исключить из входного изображения фон и выбрать операционную область. В качестве фона используется специально подготовленный кадр, представляющий собой снимок пустого помещения. Данный фон будет неизменным на протяжении всего процесса распознавания.

Схема получения маски операционной области представлена на рис. 3. Из каждого обрабатываемого кадра видеопотока вычитается фон, после чего к абсолютному значению полученной разности применяется пороговая функция с настраиваемым параметром λ.

Завершающим этапом к полученному изображению применяется преобразование из цветового пространства RGB в оттенки серого с целью получения маски. После вычисления маски M она накладывается на входной кадр, в результате чего происходит выделение операционной области:

$$ROI_{ij} = \begin{cases} I_{ij} - \text{ при } M \neq 0; \\ 0 - \text{ при } M = 0. \end{cases}$$

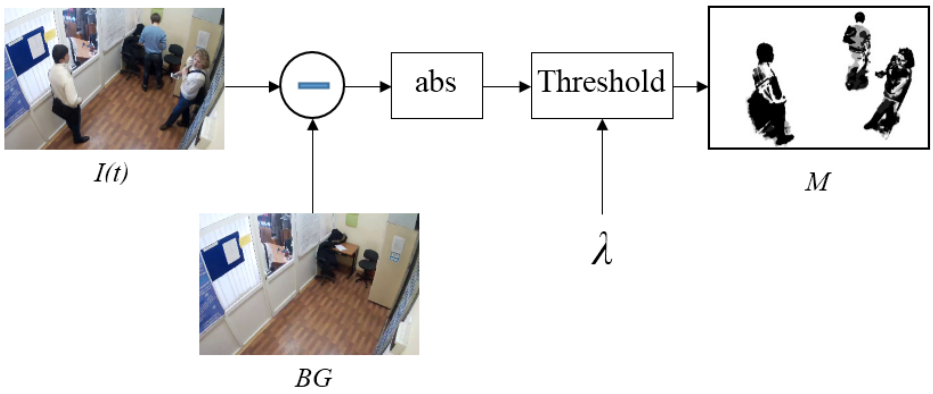


Рис. 3. Схема получения маски операционной области

Следующим этапом в цепочке методов предложенного алгоритма распознавания идет обнаружение контуров объектов, находящихся в операционной области. Обнаружение контуров производится с целью разделения объектов на области для дальнейшего их анализа. Для реализации метода обнаружения контуров выбран оператор Кэнни. Работа оператора Кэнни состоит из следующих шагов:

1. Применение фильтра Гаусса для удаления шума и сглаживания изображения;
2. Поиск градиентов цветовой интенсивности на изображении;
3. Локализация максимумов с целью уменьшения толщины контуров;
4. Применение пороговой функции;
5. Трассировка контуров для подавления побочных контуров, несвязанных с основными.

Результат применения оператора Кэнни с прорисованными на операционной области контурами показан на рис. 4. Как видно из рисунка, обнаруженные контуры разделяют операционную область на несколько фрагментов, среди которых находятся интересующие нас лица и головы. Преимущество такого разбиения в том, что с помощью пороговых значений возможно отсеять те области, которые заведомо не могут содержать интересующий объект в связи с малым размером.

Рис. 4. Контуры объектов, выделенные оператором Кэнни

Выделение областей основано на применении оператора Лапласиан Гауссиана (Laplacian of Gaussian). В результате свертки изображения с оператором Лапласиан Гауссиана фиксируются отклики, соответствующие различным областям. Пример результата выделения областей приведен на рис. 5. Выделенные области для наглядности закрашены различными цветами. В результате операционная область делится на набор областей, соответствующих частям тела, элементам одежды и т.д. на исходном обрабатываемом изображении.

Следующим этапом после выделения областей применяется распознавание лица или головы человека в зависимости от того, видно лицо или нет. Так как точность распознавания лица выше из-за большего количества характерных черт, приоритет отдается данному методу. Однако если в пределах области лицо не распознано, производится попытка обнаружить голову.

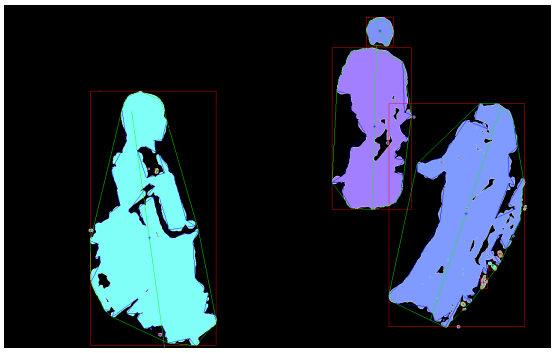


Рис. 5. Области в пределах обнаруженных контуров

Распознавание лиц и голов производится только в выделенных областях операционной области, полученных на предыдущем шаге. Вначале производится поиск лица в пределах одной области. В случае, если лицо не найдено, в данной области запускается распознавание голов обученным классификатором Хаара. Классификатор Хаара обучен на положительном и отрицательном тестовых массивах: первый содержит

изображения голов в профиль и сзади (см. рис. 6), второй – изображения посторонних объектов, которые не должны вызывать положительный отклик классификатора. Обнаружение лиц производится с помощью алгоритма Виолы-Джонса.

Рис. 6. Пример изображений из обучающего массива

В результате выполнения цепочки вышеописанных шагов получены следующие данные: количество людей N, вычисленное как сумма распознанных лиц и голов людей на изображении (см. рис. 7), а также оценка количества людей Z, которая вычисляется по следующей формуле:

$$Z = \frac{P_{\text{обл}}}{\beta P_{\text{общ}}},$$

где $P_{\text{обл}}$ – количество пикселей в выделенных областях, $P_{\text{общ}}$ – общее количество пикселей изображения, β – коэффициент, который показывает, какой процент пикселей изображения в среднем приходится на одного человека (определяется экспертной оценкой при настройке системы).

Для вычисления итоговой оценки количества людей в помещении выбран метод группового учета аргументов. В результате экспериментов была получена следующая модель:

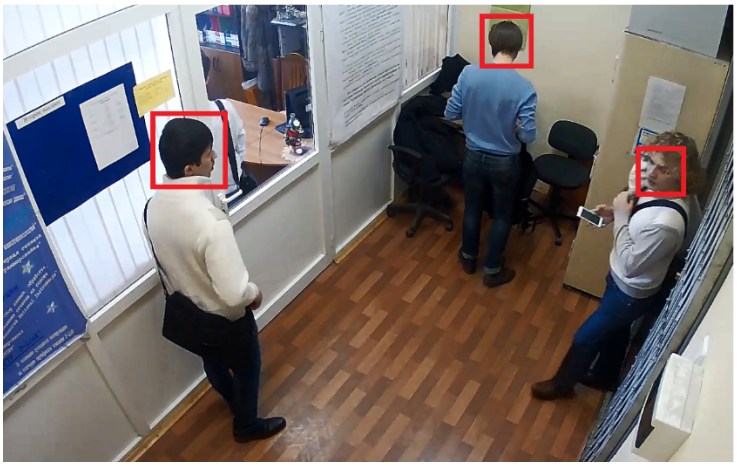


Рис. 7. Результат работы алгоритма распознавания лиц и голов

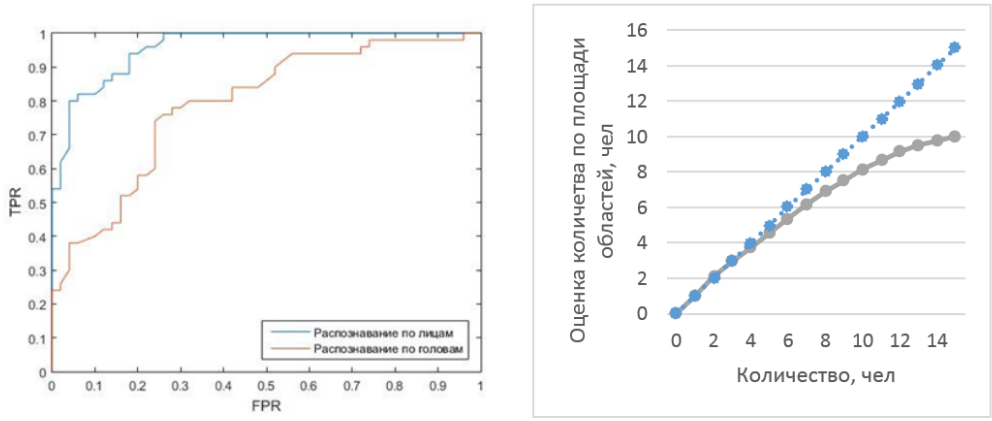
$$Y(N,Z) = a_0 + a_1N + a_2Z + a_3N^2 + a_4Z^2 + a_5NZ + a_6N^3 + a_7Z^3 + a_8N^2Z + a_9NZ^2 + a_{10}N^4 + a_{11}Z^4 + a_{12}N^3Z + a_{13}NZ^3 + a_{14}N^2Z^2.$$

Коэффициенты $a_0,\ldots,a_{14}$ определяются с помощью нелинейной регрессии по методу наименьших квадратов с использованием метода Гаусса-Ньютона. Полином для расчета итоговой оценки количества людей с найденными коэффициентами выглядит следующим образом:

$$Y(N,Z) = 0.2496 + 0.6659N + 0.2409Z + 0.2361N^2 - 0.168Z^2 - 0.1133NZ - 0.2728N^3 + 0.5046Z^3 + 0.9497N^2Z - 1.1696NZ^2 - 0.0041N^4 - 0.0547Z^4 + 0.0446N^3Z + 0.1392NZ^3 - 0.1256N^2Z^2.$$

Анализ качества работы полученной системы распознавания и подсчета людей

Качественные характеристики оценок количества людей, полученных от модулей распознавания и выделения, проиллюстрированы на рис. 8.



а) б)

Рис. 8. Качественные характеристики оценки количества человек:
а) по головам/лицам и б) – по площади областей

Зависимость точности данной оценки от реального количества людей приведена на графике на рис. 9.

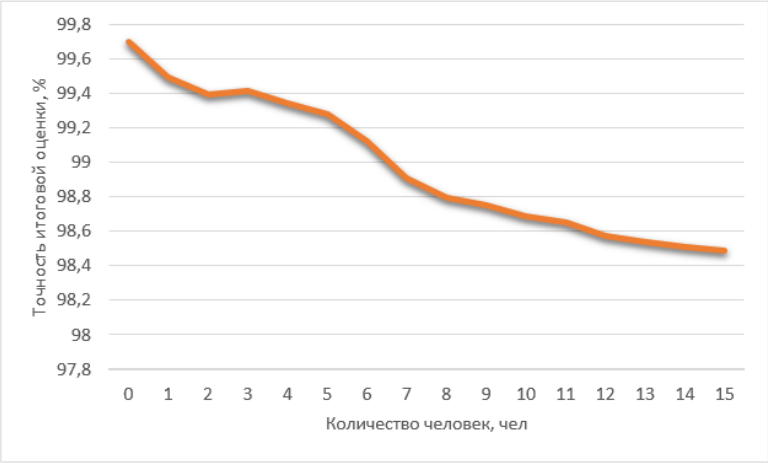


Рис. 9. Точность итоговой оценки в зависимости от числа людей

Заключение

Разработанная система распознавания динамического поведения людей позволяет вести точный подсчет людей в помещении независимо от их количества. Высокая точность итоговой оценки количетва людей достигнута засчет объединения сильных сторон различных методов распознавания и обработки изображений, а также засчет применения нелинейной мультипараметрической модели, позволяющей минимизировать ошибку итоговой оценки благодаря учету оценок от нескольких модулей. Результаты тестирования показали, что построенная система успешно справляется с поставленными задачами.

Список литературы

1. Алфимцев А.Н. Нечеткое агрегирование мультимодальной информации в интеллектуальном интерфейсе // Программные продукты и системы. 2011. № 3. С. 10.
2. Алфимцев А.Н., Демин Н.А. Захват и отслеживание удаленных объектов в видеопотоке. // Инженерный журнал: наука и инновации. 2013. Вып. 11. Режим доступа: http://engjournal.ru/catalog/it/hidden/1049.html (дата обращения 14.06.2015)
3. Lienhart, R., Maydt, J. An extended set of Haar-like features for rapid object detection // International Conference on Image Processing (ICIP). 2002. Vol. 1. P. 900-903.
4. Kilambi P., Ribnick E., Joshi A., Masoud O., Papanikolopoulos N. Estimating pedestrian counts in groups // Computer Vision and Image Understarding. 2008. Vol. 110. P. 43-59.
5. Zhou J., Hoang J. Real time robust human detection and tracking system // Computer Vision and Patter Recognition (CVPR). 25 June 2005. P. 149.

6. Bo W., Nevatia R. Detection of multiple, partially occluded humans in a single image by bayesian combination of edgelet part detector // International Conference on Computer Vision (ICCV). 17-21 October 2005. P. 90-97.
7. Koenig N. Toward real-time human detection and tracking in diverse environments // International Conference on Development and Learning (ICDL). 11-13 July 2007. P. 94-98.