

Министерство образования и науки Российской Федерации
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ Н.Г. ЧЕРНЫШЕВСКОГО»

Кафедра дискретной математики и
информационных технологий

**Разработка и программная реализация алгоритма обнаружения и
сопровождения заданного объекта в потоковом виде**

АВТОРЕФЕРАТ БАКАЛАВРСКОЙ РАБОТЫ

Студентки 4 курса 421 группы
направления 09.03.01 – Информатика и вычислительная техника
факультета компьютерных наук и информационных технологий
Пиуновой Екатерины Ивановны

Научный руководитель
ассистент кафедры ДМиИТ

Н.Е. Тимофеева

Зав. кафедрой
доцент, к.ф.-м.н., доцент

Л.Б. Тяпаев

Саратов 2017

ВВЕДЕНИЕ

С изобретением ЭВМ перед человеком открылся глобальный мир новых возможностей. Одни из которых – проведение сложных расчётов и решение различных задач. Компьютеры отличаются от человека скоростью, точностью и отсутствием потребности в отдыхе, поэтому, если необходимо решить математическую задачу, применяется искусственный интеллект. Но есть области, в которых компьютер уступает человеку. Одна из таких областей – распознавание образов.

В момент попадания отраженного света в глаза, человек может определить, какой объект находится перед ним, его приблизительные размеры, скорость передвижения, цвет и дать ему оценку. Компьютер же работает с массивом пикселей, которые ему необходимо обработать и сравнить с массивом образов, которые ему известны. Точность и скорость распознавания зависят от задачи, мощности компьютера и данных, предложенных для анализа.

Целью бакалаврской работы является программная реализация модифицированного алгоритма обнаружения и сопровождения заданного объекта в потоковом виде.

Для достижения поставленной цели необходимо выполнить следующие задачи:

- изучить и проанализировать теоритические аспекты, методы и задачи распознавания образов;
- обучить каскады Хаара и выявить зависимость точности работы каскада от обучающей выборки;
- программно реализовать стандартный алгоритм Виолы-Джонса;
- модифицировать алгоритм Виолы-Джонса и программно реализовать его;
- провести тестирование разработанного алгоритма и сравнить его характеристики со стандартным.

Основное содержание работы

1 Общая характеристика алгоритма Виолы-Джонса

Метод Виолы-Джонса – алгоритм, позволяющий обнаруживать объекты на изображении и видео в режиме реального времени. Метод был предложен Майклом Джонсоном и Паулом Виолой в 2001 году. Первоначальной задачей метода было обнаружение лиц, но алгоритм способен также распознавать любые классы изображений.

Метод использует представление изображения в интегральном виде, которое позволяет вычислять быстро, необходимые объекты, с помощью признаков Хаара. Алгоритм применяет бустинг (от англ. boost – улучшение, усиление) для выбора наиболее подходящих признаков для искомого объекта на данной части заданного изображения. Метод использует классификатор, на вход которого поступают все признаки, затем выдаётся результат «верно» либо «ложь», то есть, принадлежит ли выделенный объект искомому классу или нет. Для быстрого отбрасывания окон, где найден объект, используются каскады признаков.

Процесс обучения классификаторов занимает большое количество времени, но результаты нахождения лица очень быстры, именно поэтому был выбран данный метод распознавания лиц на изображении. Алгоритм Виолы-Джонса является одним из лучших по соотношению показателей эффективности распознавания и скорости работы. Алгоритм распознает черты лица под небольшим углом, примерно до 30 градусов, но при увеличении угла наклона процент обнаружений резко падает.

Рассмотрим далее более подробно принципы, на которых основан алгоритм Виолы-Джонса.

Метод Виолы-Джонса в общем виде ищет лица и черты лица по общему принципу сканирующего окна. Рассмотрим задачу обнаружения лица на цифровом изображении.

Имеется изображение, на котором есть объекты, среди которых есть искомые. Оно представляется двумерной матрицей пикселей размером $w \times h$ (w – ширина, h – высота изображения). Каждый пиксель в матрице принимает значение от 0 до 255 для черно-белого изображения, и от 0 до 255^3 для цветного изображения (компоненты R, G, B).

Алгоритм должен определить лица и пометить их. Поиск осуществляется с помощью прямоугольных признаков в активной области изображения. С помощью признаков, описываемых формулой, и обозначается найденное лицо:

$$rectangle[i] = \{x, y, w, h, a\}, \quad (1)$$

где i – номер прямоугольника, x и y – координаты центра i -го прямоугольника, w – ширина, h – высота, a – угол наклона прямоугольника к вертикальной оси изображения.

Иными словами, изображение сканируется окном поиска (окно сканирования), а затем применяется классификатор к каждому положению.

Признаки Хаара (Haar-Like features) – признаки цифрового изображения, которые используются в распознавании образов. Свое название признаки получили благодаря сходству с вейвлетами Хаара.

Вейвлет – математическая функция, которая позволяет анализировать частотные компоненты данных. Вейвлет Хаара – один из первых и наиболее простых вейвлетов, основанный на ортогональной системе функций, которую в 1909 году предложил Альфред Хаар.

Алгоритмы, которые работают только с интенсивностью изображения, имеют большую вычислительную сложность. В работе Папагеоргиу, была рассмотрена и проанализирована работа с множеством признаков, которые основаны на вейвлетах Хаара. Виола и Джонс адаптировали идею использования данных вейвлетов и разработали то, что впоследствии было названо признаками Хаара.

Признак Хаара имеет форму прямоугольника или квадрата и состоит из смежных прямоугольных областей. Признаки Хаара представлены на рисунке 1.

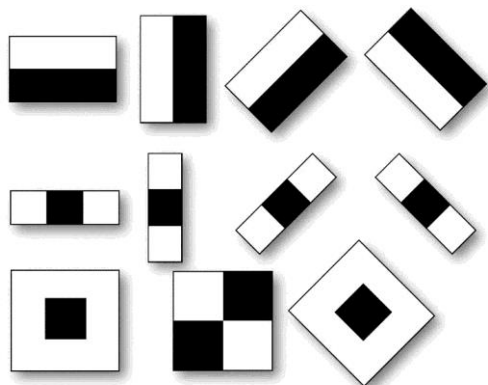


Рисунок 1 – Признаки Хаара

Они располагаются на изображении, после чего суммируются интенсивности пикселей в областях, далее вычисляется разность между суммами. Посчитанная разность и будет значением определенного признака, определенного размера, определенным образом расположенного на изображении.

В библиотеке компьютерного зрения OpenCV в расширенном методе Виолы-Джонса используются дополнительные признаки.

Каскадная модель классификаторов – это дерево принятия решений, где каждый узел дерева построен так, чтобы находить большинство интересных образов и отклонять регионы, которые не являются образами. При этом узлы дерева размещены таким образом, что чем ближе узел находится к корню дерева, тем из меньшего количества примитивов он состоит и тем самым требует меньше времени на принятие решения.

Рассмотрим подробнее каскад Хаара. Данный классификатор формируется на примитивах Хаара путём расчёта значений признаков. Чтобы обучить классификатор, на его вход подаётся набор «хороших» (правильных) изображений, которые имеют предварительно выделенную область на изображении. Далее происходит перебор примитивов и расчёт значения признака, после чего вычисленные значения сохраняются в файле в формате xml.

2 Обучение каскада Хаара

Рассмотрим обучение каскада Хаара.

Для обучения необходимо собрать следующий набор файлов:

- Реальные фотографии объекта. Чем более похожа выборка будет на то, что необходимо распознать, тем лучше будут результаты;
- Выборка отрицательных фотографий. Фотографии должны быть сделаны в той же среде где будет распознавание.

Для стабильно работающего детектора необходимо 3000 – 4000 положительных примеров и столько же отрицательных. Чем больше и разнообразнее выборка, тем стабильнее работает каскад, но при этом дольше обучается.

Обучение происходит в два этапа. Первый этап – все положительные изображения приводятся к общему формату, расположенной в папке библиотеки компьютерного зрения OpenCV программой `opencv_createsamples.exe`.

Результатом работы программы является файл `samples.vec`, который будет содержать положительные изображения в формате, близком к bmp и с размером `w*h`.

Следующий шаг обучения – создание итогового каскада, для которого используется программа `opencv_traincascade.exe`, лежащая в той же папке, что и `opencv_createsamples.exe`.

После продолжительного обучения, на выходе будет получен каскад Хаара.

Для точного распознавания объектов необходимо использовать хорошо обученный каскад Хаара. Качество работы каскада зависит от количества положительных и отрицательных образцов, а так же выбираемых параметров при обучении. Обучим несколько каскадов для обнаружения одного и того же класса объектов, после чего оценим количество ложных и правильных срабатываний, а также время обучения каскада. В процессе обучения будем изменять объем обучающей выборки.

После обучения, отберем 50 изображений, на которых присутствуют искомые объекты. Будем выбирать изображения, на которых присутствует как один объект, так и несколько, а также фотографии разного размера и качества.

Введем следующие обозначения:

- N_p – количество положительных образцов;
- N_n – количество отрицательных образцов;
- t – примерное время обучения каскада;
- f – количество ложных срабатываний;
- tr – количество правильных срабатываний.

Результаты обучения занесены в таблицу 1.

Таблица 1 – Таблица параметров для обучения каскадов

№	N_p (штук)	N_n (штук)	t	f (%)	Tr (%)
1	50	50	1 час	40	60
2	80	80	1 час 20 минут	37,5	62,5
3	200	80	1 час 40 минут	39	61
4	200	200	2 часа	28,7	71,3
5	500	200	3 часа	30	70
6	500	500	4 часа 10 минут	19,2	80,8
7	1000	500	6 часов	19	81
8	1000	1000	8 часов 30 минут	12	88

Рассмотрим результаты обучения каскадов. Как видно из таблицы 1, количество ложных срабатываний уменьшается при увеличении объемов выборки. Стоит заметить, что количество положительной и отрицательной выборки должны примерно быть равны. Иначе, как показал анализ,

количество правильных срабатываний будет расти незначительно, или вовсе может уменьшиться. Каскад под номером шесть приведен в приложении А.

Таким образом, для достижения максимально точного распознавания, необходимо проводить обучение на выборках, которые содержат 3000 – 4000 изображений каждая. При этом время обучения будет значительно увеличиваться.

3 Разработка программного продукта

Реализуем стандартный и модифицированный алгоритм для распознавания изображения, видеофрагментов и потокового видео. Общая структура программы показана на рисунке 2.

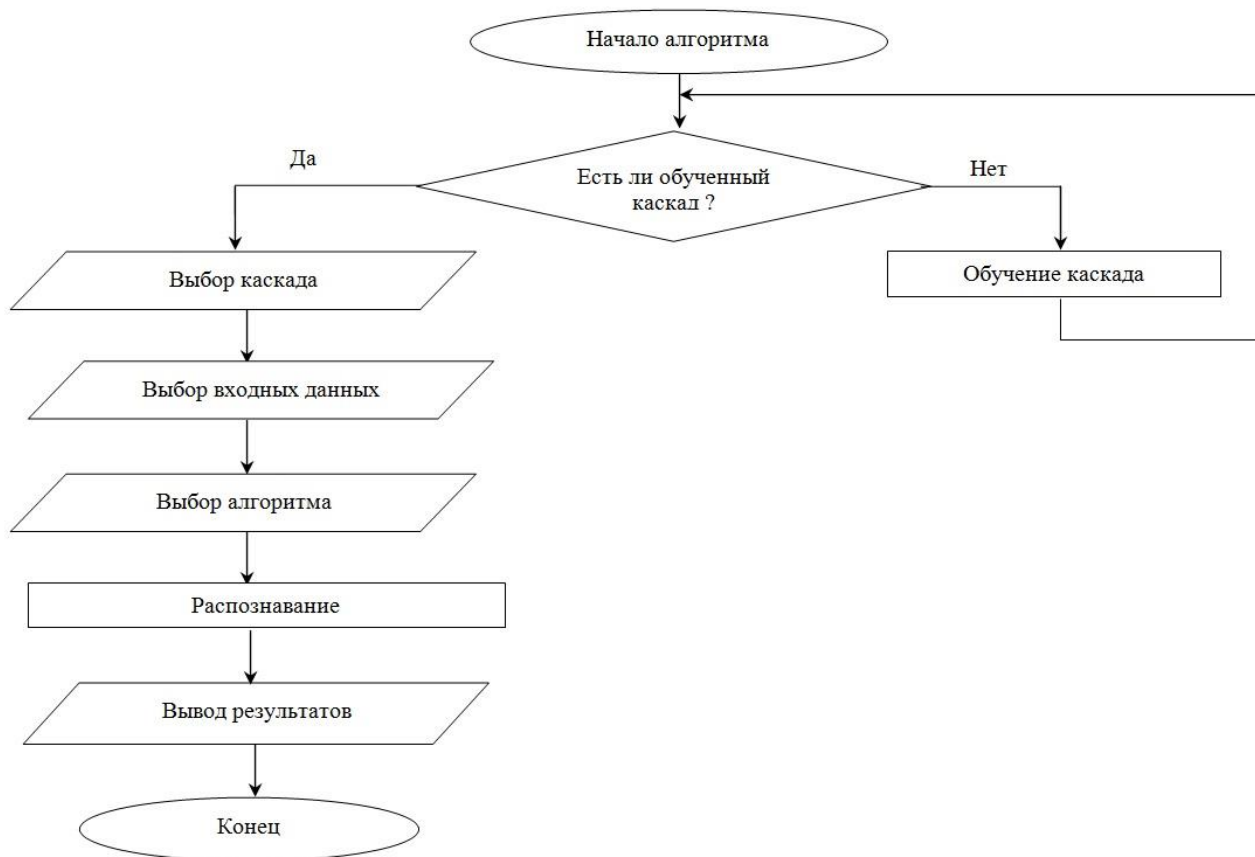


Рисунок 2 – Общая структура программы

Программа написана на объектно-ориентированном языке высокого уровня C#, с подключением библиотеки Emgu CV. Emgu CV это кросс-платформенное .Net дополнение для библиотеки OpenCV, которое позволяет обрабатывать изображения. Дополнение разработано для работы с .NET совместимыми языками, например, такими как C #, VB, VC ++. Поэтому может быть использовано в Visual Studio, Xamarin, и работает с различными операционными системами: Windows, Linux, Mac OS X, IOS, Android и Windows Phone.

Для корректной работы программы, необходимо наличие библиотеки OpenCV версии 2.4.1 в корне диска C и установленная библиотека Emgu 2.9.

После запуска программы появляется главное окно, представленное на рисунке 3.

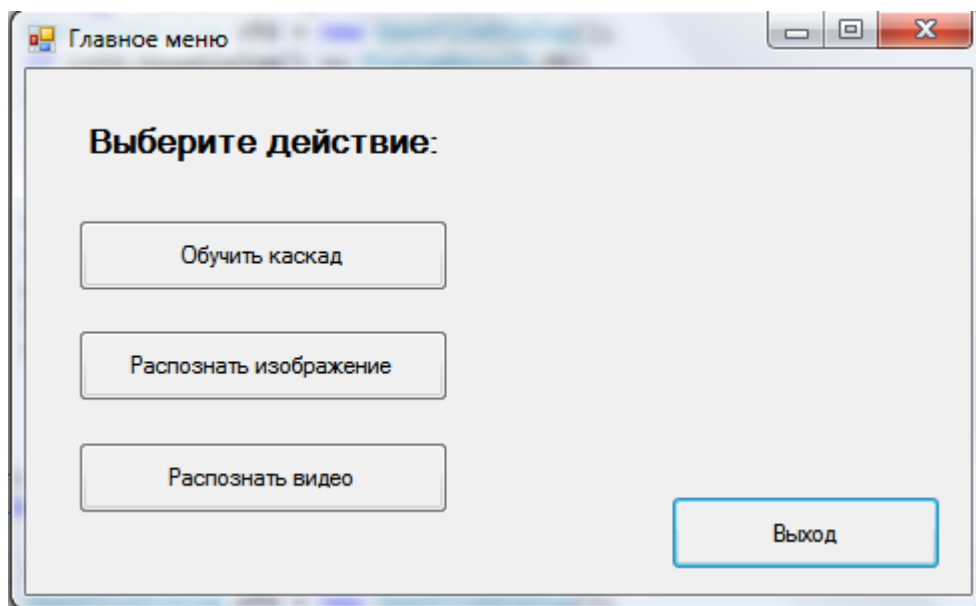


Рисунок 3 – Главное окно программы

В главном окне необходимо выбрать действие. Действие «обучить каскад» позволяет получить новый классификатор Хаара для дальнейшего распознавания изображения. Кнопку «распознать изображение» и «распознать видео» позволяют перейти к распознаванию изображения и видео с готовыми каскадами. Кнопка выход позволяет закрыть программу.

ЗАКЛЮЧЕНИЕ

В ходе выполнения данной работы были изучены материалы по распознаванию образов и компьютерному зрению. Рассмотрены проблемы, которые могли возникнуть в процессе распознавания объектов на изображениях, а также предложены возможные методы их решения. В результате чего был разработан и реализован алгоритм с использованием языка C#, библиотекой компьютерного зрения OpenCV.

Разработан и программно реализован алгоритм, позволяющий находить и отслеживать объекты определенного цвета. Алгоритм был применен к изображению, и потоковому видео, на котором выделенный объект сопровождался.

Также проведен сравнительный анализ стандартного и модифицированного алгоритмов.

Таким образом, поставленные цель и задачи, бакалаврской работы были выполнены полностью.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

- 1 Википедия [Электронный ресурс] : свободная энциклопедия / текст доступен по лицензии Creative Commons Attribution-ShareAlike ; Wikimedia Foundation, Inc, некоммерческой организации Электрон. дан. (1 395 000 статей, 5 267 970 страниц, 199 383 загруженных файлов). Wikipedia, 2001 - 2017. URL: https://ru.wikipedia.org/wiki/Компьютерное_зрение (дата обращения: 13.04.2017). Загл. с экрана. Последнее изменение страницы: 19:50, 9 января 2017 года. Яз. рус.
- 2 Дмитриев, И.Н. Использование систем машинного зрения в проектировании охранных приборов / И. Н. Дмитриев // Сборник научно-практической конференции студентов и аспирантов ВГУ 2012. С. 120 – 128.
- 3 Википедия [Электронный ресурс] : свободная энциклопедия / текст доступен по лицензии Creative Commons Attribution-ShareAlike ; Wikimedia Foundation, Inc, некоммерческой организации Электрон. дан. (1 395 000 статей, 5 267 970 страниц, 199 383 загруженных файлов). Wikipedia, 2001 - 2017. URL: https://ru.wikipedia.org/wiki/Теория_распознавания_образов (дата обращения: 16.05.2017). Загл. с экрана. Последнее изменение страницы: 9:27, 14 мая 2017 года. Яз. рус.
- 4 Черногорова Ю. В. Методы распознавания образов / Ю. В. Черногорова // Молодой ученый. 2016. №28. С. 40-43.
- 5 Энциклопедический фонд России [Электронный ресурс]. // Потокое видео URL: <http://www.russika.ru/ef.php?s=4262> (дата обращения: 27.04.2017). Загл. с экрана. Яз. рус.
- 6 Вязовых М.В., Зайцев К.И., Мухортов М.В., Перов А.Н. Обнаружение лиц на цветном растровом изображении с применением модифицированного метода Виолы-Джонса / М.В.

- Вязовых, К.И. Зайцев, М.В. Мухортов, А.Н. Перов // Молодежный научно-технический вестник 2016 № 5. С. 78 – 85.
- 7 Habrahabr [Электронный ресурс]. // Метод Виолы-Джонса (Viola-Jones) как основа для распознавания лиц. URL: <https://habrahabr.ru/post/133826/> (дата обращения: 7.02.2017). Загл. с экрана. Яз. рус.
 - 8 Местецкий Л. М., «Математические методы распознавания образов», МГУ, ВМиК, Москва, 2002–2004., с. 42 – 44.
 - 9 MachineLearning.ru [Электронный ресурс]. // Бустинг. URL: <http://www.machinelearning.ru/wiki/index.php?title=Бустинг> (дата обращения: 2.03.2017). Загл. с экрана. Яз. рус.
 - 10 Википедия [Электронный ресурс] : свободная энциклопедия / текст доступен по лицензии Creative Commons Attribution-ShareAlike ; Wikimedia Foundation, Inc, некоммерческой организации Электрон. дан. (1 395 000 статей, 5 267 970 страниц, 199 383 загруженных файлов). Wikipedia, 2001 - 2017. URL: <https://ru.wikipedia.org/wiki/AdaBoost> (дата обращения: 01.02.2017). Загл. с экрана. Последнее изменение страницы: 13:19, 11 марта 2017 года. Яз. рус.
 - 11 Википедия [Электронный ресурс] : свободная энциклопедия / текст доступен по лицензии Creative Commons Attribution-ShareAlike ; Wikimedia Foundation, Inc, некоммерческой организации Электрон. дан. (1 395 000 статей, 5 267 970 страниц, 199 383 загруженных файлов). Wikipedia, 2001 - 2017. URL: https://ru.wikipedia.org/wiki/C_Sharp (дата обращения: 20.04.2017). Загл. с экрана. Последнее изменение страницы: 13:19, 11 марта 2017 года. Яз. рус.
 - 12 Википедия [Электронный ресурс] : свободная энциклопедия / текст доступен по лицензии Creative Commons Attribution-ShareAlike ; Wikimedia Foundation, Inc, некоммерческой организации Электрон.

дан. (1 395 000 статей, 5 267 970 страниц, 199 383 загруженных файлов). Wikipedia, 2001 - 2017. URL: <https://ru.wikipedia.org/wiki/OpenCV> (дата обращения: 24.04.2017). Загл. с экрана. Последнее изменение страницы: 12:07, 17 марта 2017 года. Яз. рус.

13 Emgu CV in .Net [Электронный ресурс]. // Emgu CV. URL: http://www.emgu.com/wiki/index.php/Main_Page (дата обращения: 30.04.2017). Загл. с экрана. Яз. англ.

14 Habrahabr [Электронный ресурс]. // Работа каскада Хаара в OpenCV в картинках: теория и практика. URL: <https://habrahabr.ru/company/recognitor/blog/228195/> (дата обращения: 3.02.2017). Загл. с экрана. Яз. рус.

15 Эрман Е.А., Мамдух М. Метод обнаружения лиц на изображении с использованием комбинации метода Виолы Джонса и алгоритмов определения цвета кожи / Е.А. Эрман, М. Мамдух // Вестник Астраханского государственного технического университета. 2015. № 1. С. 49-55.