УДК xxx.xxx.x

студ. В. С. Кашперко

Науч. рук. ст. преп. Я. А. Игнаткова

(каф. информатики и веб-дизайна, БГТУ)

**АНАЛИЗ ИЗОБРАЖЕНИЙ НА ОСНОВЕ РЕАЛИЗАЦИИ АЛГОРИТМОВ И МЕТОДОВ КОМПЬЮТЕРНОГО ЗРЕНИЯ**

Компьютерное зрение (Computer Vision, CV) — это область машинного обучения и компьютерных наук, помогающая вычислительным машинам понимать мир путем распознавания визуальных образов и обнаружения объектов, как это делают люди. Как научная дисциплина, компьютерное зрение относится к теории и технологии создания искусственных систем, которые получают информацию из изображений. Также широко используются и другие названия для компьютерного зрения, например: техническое, машинное зрение. Однако компьютерное зрение является более общей областью исследований, тогда как машинное зрение является инженерной дисциплиной, связанной с производственными задачами.

Компьютерное зрение сформировалось как независимая дисциплина еще к концу 60-х годов. Это течение появилось в пределах искусственного интеллекта еще тогда, когда еще велись дискуссии о вероятности сотворения мыслящей машины. Тогда ученые считали, что имитация зрительной системы человека поможет наделить роботов разумным поведением [1].

В 1958 году Фрэнк Розенблатт, психолог, сотворил компьютерное воплощение перцептрона (от «perception» – восприятие) – приспособления, имитирующего схему толкования образов человеческим мозгом. В 1966 году поступило предложение подключить камеру к компьютеру и заставить машину «описывать увиденное», однако технологии того времени не позволили реализовать задуманное.

Исследования 1970 годов заложили ранние основы для многих алгоритмов компьютерного зрения, существующих сегодня, включая выделение границ на изображениях, маркировку линий, оценку движения и прочее.

К концу 1990 годов произошли значительные изменения с усилением взаимодействия между областями компьютерной графики и компьютерного зрения. Это десятилетие также ознаменовалось первым использованием методов статистического обучения на практике для распознавания лиц на фотографиях.

В начале XXI века наблюдалось возрождение основанных на функциях методов, которые начали использовать в сочетании с машинным обучением и сложными структурами оптимизации. В 2012 году на конкурсе ImageNet сверхточная нейронная сеть AlexNet вошла в топ-5 алгоритмов с уровнем ошибок 15,3%. В 2015 году нейросеть победила в конкурсе. Именно это событие считается отправной точкой в современной истории компьютерного зрения [2].

Цель компьютерного зрения – принятие решений о реальных физических объектах и сценах, основываясь на воспринимаемых изображениях. Задача компьютерного зрения – научить вычислительную машину видеть и понимать окружение с помощью цифровых фотографий и видеозаписей при помощи методов низкоуровневого анализа: выделения границ, классификации, обнаружения объектов и сегментации.

Ежедневно нас окружают удивительные вещи, к которым мы привыкли. Например, автоматическое определение лиц при фотографировании, разблокировка мобильного телефона при помощи распознавания лица, использование фильтров и масок при съемке мобильным телефоном – это лишь небольшой список повседневных приложений, при реализации которых решаются задачи компьютерного зрения. Также существуют такие отрасли, как промышленность, медицина, производство охранных систем, роботов, умных автомобилей и домов, в которых умение получать информацию из изображения является самой основной функцией.

Таким образом, задача детектирования лица на изображении является одной из самых используемых задач компьютерного зрения. Для решения этой задачи успешно применяется метод Виолы-Джонса – алгоритм, позволяющий обнаруживать объекты на изображениях в реальном времени. Хотя алгоритм может распознавать объекты на изображениях, основной задачей при его создании было обнаружение лиц. Метод Виолы-Джонса является одним из лучших по соотношению распознавание/скорость работы. Благодаря скорости обработки изображения можно с легкостью обрабатывать потоковое видео.

Основу метода Виолы-Джонса составляют примитивы Хаара, представляющие собой разбивку заданной прямоугольной области на наборы разнотипных прямоугольных подобластей. Например, из двух признаков Хаара строится первый каскад системы по распознаванию лиц, который имеет вполне осмысленную интерпретацию: зона глаз должна быть темнее, чем зона носа и щек, а переносица светлее, чем глаза.

При наложении примитивов Хаара на изображение рассчитываются признаки, которые равны разности суммы пикселей под белым прямоугольником и суммы пикселей под черным прямоугольником. Для увеличения скорости расчетов используется интегральное представление изображения. Интегральное представление изображения – это матрица, такого же размера, как исходное изображение. Каждый элемент интегрального изображения содержит в себе сумму значений всех пикселей левее и выше данного пикселя. Интегральное представление позволяет быстро вычислить сумму пикселей произвольного прямоугольника с помощью всего лишь четырех ссылок на массив, независимо от размера массива.

В результате обучения формируется XML-файл, содержащий каскад обученных классификаторов с выбранными признаками и со значениями пороговых величин для классификаторов. Для определения принадлежности объекта к классу в каждом каскаде находится сумма значений слабых классификаторов этого каскада. Каждый слабый классификатор выдает два значения в зависимости от того, больше или меньше заданного порога значение признака, принадлежащего этому классификатору. В конце сумма значений слабых классификаторов сравнивается с порогом каскада и выносится решения, найден объект или нет данным каскадом [3].

С целью проведения исследования, насколько данный алгоритм эффективен, была реализована программа обнаружения на видео лица и глаз человека в реальном времени, при помощи каскадов Хаара, на языке Python. Программа проверяет наличие маски на лице (наличие рта внутри контура лица). Для работы понадобились готовые обученные классификаторы XML. Создавать собственные классификаторы XML нецелесообразно, так как всё необходимое обучение с выбранными признаками и значениями пороговых величин было пройдено ранее, используя множество примеров лиц реальных людей. Для работы с компьютерным зрением использовалась OpenCV — библиотека алгоритмов компьютерного зрения и обработки изображений [4].

После реализации программы были проведены соответствующие тесты, результаты которых приведены ниже:

Таблица

**Результаты тестирования программы для проверки наличия маски на лице человека**

|  |  |  |  |
| --- | --- | --- | --- |
|  | Естественное освещение | Искусственное освещение | Темнота (полумрак) |
| 1 лицо | + | + | + |
| 2 лица | + | + | - |
| 3 лица | + | - | - |
| Считывание лица с фотографии | + | + | + |
| Закрывание лица рукой, имитируя наличие маски | - | - | + |
| Проверка наличия маски на лице у человека | + | + | - |
| Наличие рта внутри контура лица | + | + | - |

На основании проведенного тестирования сделаем вывод: данный алгоритм хорошо справляется с поставленными задачами, может быть успешно применим для достижения целей в области распознавания и определения лиц, частей лица и выявления наличия необходимых признаков, в зависимости от поставленной задачи.

Научить компьютер видеть мир — значит получить надежного помощника, который день и ночь будет занят решением важных задач. Компьютерное зрение – уникальный инструмент, и, рассмотренные примеры использования компьютерного зрения — это лишь малая доля того, на что способен «прозревший» компьютер!

**ЛИТЕРАТУРА**

1. Что такое компьютерное зрение? (машинное обучение) [Электронный ресурс]// Forklog. – Режим доступа: <https://forklog.com/chto-takoe-kompyuternoe-zrenie-mashinnoe-obuchenie/>. – Дата доступа: 04.04.2022.
2. История компьютерного зрения [Электронный ресурс]// Telecom & IT. – Режим доступа: <https://shalaginov.com/2020/05/16/computer-vision-history/>. – Дата доступа: 08.04.2022.
3. Методы компьютерного зрения: учеб.-метод. пособие для студентов специальностей 1-40 01 01 «Программное обеспечение информационных технологий», 1-98 01 03 «Программное обеспечение информационной безопасности мобильных систем» / сост. И. Г. Сухорукова. – Минск : БГТУ, 2021. – 71с.
4. OpenCV шаг за шагом [Электронный ресурс]// Robocraft. – Режим доступа: http://robocraft.ru/page/opencv/. – Дата доступа: 12.04.2022.