**«Компьютерное зрение и графика в современном мире»**

Подготовила:

студентка 2 курса 1 группы

факультета информационных технологий

Кашперко Василиса Сергеевна

Руководитель: Игнаткова Янина Алексеевна

Цель:  
  
ознакомление с компьютерным зрением, его задачами и практическим применением  
  
Задачи:  
  
- Дать определение компьютерному зрению;  
- Выяснить, какие задачи стоят перед компьютерным зрением;  
- Изучить доступные Интернет-ресурсы;  
- Изучить области применения компьютерного зрения;  
- Рассмотреть примеры применения компьютерного зрения на практике.

**Что такое компьютерное зрение?**

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 не стоит путать машинное и компьютерное зрения. Компьютерное зрение является более общей областью исследований, тогда как машинное зрение является инженерной дисциплиной, связанной с производственными задачами.

**История развития компьютерного зрения**

Компьютерное зрение сформировалось как независимая дисциплина еще к концу 60-х годов.

Это течение появилось в пределах искусственного интеллекта тогда, когда еще велись жаркие дискуссии о вероятности сотворения мыслящей машины. Оно появилось из трудов по распознаванию образов.

Пионеры в области искусственного интеллекта начали интенсивнее обсуждать вопросы распознавания образов с помощью компьютерных алгоритмов. Тогда ученые считали, что имитация зрительной системы человека поможет наделить роботов разумным поведением.

**Первые эксперименты**

1958 г. - Фрэнк Розенблатт, психолог из Корнеллского университет, сотворил компьютерное воплощение перцептрона (*от «perception» - восприятие*) - приспособления, имитирующего схему толкования образов человеческим мозгом.

**Перцептрон**

Перцептрон был впервые создан в 1958 году, вдобавок его подготовка занимала около получаса машинного времени на ЭВМ IBM-704. Аппаратный вариант - Mark I Perceptron – был сконстурирован в 1960 г. и применялся для толкования зрительных образов.

В 1966 году они предложили 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.

В следующем десятилетии ученые работали над более строгим математическим анализом и количественными аспектами технологии.

К концу 1990 годов произошли значительные изменения с усилением взаимодействия между областями компьютерной графики и компьютерного зрения. Это включало рендеринг на основе изображения, интерполяцию вида, сшивание панорамных кадров и прочее.

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. Однако настоящая революция произошла лишь с развитием области глубокого обучения, точность которого превзошла все существующие на тот момент подходы.

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.

Цель **компьютерного зрения** - принятие решений о реальных физических объектах и сценах, основываясь на воспринимаемых изображениях.

Миссия компьютерного зрения — научить вычислительную машину видеть и понимать окружение с помощью цифровых фотографий и видеозаписей. Для достижения этой цели используются три компонента:

* получение изображений;
* обработка информации;
* анализ данных.

Как правило, обработка данных осуществляется с помощью сложных математических алгоритмов. Популярными методами низкоуровневого анализа являются:

* выделение границ, или edge detection;
* классификация и обнаружение объектов;
* cегментация.

**Источники изображений**

Источниками изображений для компьютерного зрения являются такие устройства ввода как: видеокамеры, веб-камеры, фотоаппараты, сканеры и скрытые камеры видеонаблюдения.

Компьютерное зрение теснейшим образом взаимодействует с областью обработки изображений, часто трудно однозначно отнести возникающие задачи и применяемые методы решения к одной из этих областей.

В задачи, рассматриваемые в рамках компьютерного зрения, в частности, входят:

* Обнаружение/распознавание/отслеживание объектов, обладающих определенными свойствами (в самом широком смысле) на статическом изображении и в видеопотоке;
* Восстановление 3D формы по 2D изображениям с помощью:
  + Стереореконструкции карты глубины
  + Реконструкции поля нормалей и карты глубины по закраске полутонового изображения
  + Реконструкции карты глубины по текстуре
  + Определению формы по перемещению
* Выделение на изображениях структур определенного вида (например краев - резких переходов яркости), cегментация изображений (например на области, однородные по определенному признаку);
* Анализ оптического потока (нахождения перемещения пикселей между двумя изображениями);

Одна из задач это **Распознавание –**определить содержит ли увиденное компьютером некоторый объект, особенность или активность. Эта задача легко решена человеком, но до сих пор не решена полностью в компьютерном зрении. Следующая задача – **Движение.**Пример такой задачи - Слежение, то есть следование за перемещениями объекта (например, машин или людей)

**Задачи Восстановление сцены и восстановление изображений** они еще более сложные, чем первые.

**Область применения CV в современном мире**

Ежедневно нас окружают удивительные вещи, к которым мы привыкли. Например, считывание штрих-кода при покупке, автоматическое определение лиц при фотографировании, разблокировка мобильного телефона по лицу, фильтры и маски при съемке мобильным телефоном – это лишь небольшой список повседневных приложений, при реализации которых решаются задачи компьютерного зрения. А есть еще такие отрасли, как промышленность, медицина, производство охранных систем, роботов, умных автомобилей и домов, в которых умение получать информацию из изображения является важной (если не главной) функцией.

**Системы безопасности**

Приложения методы Computer Vision находят в следующих областях:

* Охранные системы (идентификация личности, детекторы движения, распознавание и отслеживание движущихся объектов, распознавание автомобильных номеров и т.д.);

**Распознавание жестов**

* Естественный интерфейс человек-компьютер, основанный на жестах, мимике и голосовых командах пользовтаеля, не обремененный устройствами "удобными компьютеру" (клавиатурой, мышью) - human-centered computing;

**Распознавание текста**

* Системы распознавания рукописного и печатного текста;

**Отслеживание перемещения человека**

* Промышленность (дефектоскопия, контроль качества, промышленные роботы)
  + автомобилестроение;
  + электроника;
  + машиностроение;

**Отслеживание перемещения человека**

**Рендеринг на основе изображения**

* сверхвысокое разрешение  
  расширенный динамический диапазон  
  усиленное воздействие (экспозиция)  
  ретушь объекта  
  удаление объекта
* Сжатие видео с учетом содержимого передающегося видеопотока (content-aware video compression);

**Дополненная реальность**

* Мультимедия-приложения, игры (реальный пример - автоматическое создание модели игрока для Quake3), расширенная реальность (комбинация виртуального окружения и реальных объектов) ;
* Медицина (автоматический анализ медицинских изображений – рентген, томография, УЗИ);

**Headtracking**

Здесь используется замечательный пример обнаружения веб-камерой лица человека и определения его положения в пространстве. Человек буквально может посмотреть на изображение с другой стороны. Давайте разберемся, каким образом это происходит.

Для начала нам надо понять, как машина различает объекты и фон. К примеру, для нахождения веб-камерой каких-то геометрических фигур надо, что бы камера отличала следующие атрибуты объекта:

* Форму объекта (круг, квадрат, треугольник…);
* Цвет определяемого объекта;
* Размер объекта и положение его относительно других объектов.

Для нахождения формы объекта используется эффект размытия контуров заданной формы. Это делается для того, чтобы устройство точно могло определить, где закончился контур нашего объекта и начался фон.

Говоря о положении определяемого объекта в зоне видимости, мы переходим к главному принципу работы, который демонстрируется на видео. Устройство находит объект, фиксирует его положение в реальном времени и, следуя алгоритму, выполняет нужные нам действия. Одной из ключевых позиций является размер нашего объекта. Мы приближаем объект, изменяя его координаты по осям, и видим изображение уже под другим углом. В случае с человеческим лицом, а не простой геометрической фигурой, есть ряд нюансов:

— Форма лица у всех людей разная, хотя есть определенные типы этой формы.

— Цвет кожи человека, личные особенности, такие как прическа, украшения, и тому подобное.

— Лицо человека, хоть и симметрично, имеет различия, если смотреть под разными углами.

Поэтому, в отличие от простых фигур, для определения лица требуется использовать немного другой подход.

Оптимальный вариант – это xml файлы, содержащие необходимые нам сведения по всем пунктам, которые были описаны выше. Это множество изображений лиц разных размеров и форм, сделанных заранее под разными углами. В сети уже есть готовые xml файлы, которые можно использовать в работе.

Работать с такой технологией легко и просто. Открывается множество интересных решений для работы с изображением, которые можно использовать как в повседневной жизни, так и в больших проектах.

**Пример разработки программы**

**Пример разработки программы**

**Пример разработки программы**

Задача детектирования лица на изображении, как правило, является подзадачей при решении более сложных проблем (например, распознавания лица или определения выражения лица). Для решения этой задачи успешно применяется метод Виолы-Джонса.

Метод Виолы– Джонса – алгоритм, позволяющий обнаруживать объекты на изображениях в реальном времени. Его предложили Паул Виола и Майкл Джонс в 2001г. 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. Однако у данного метода есть ограничение – он хорошо работает и распознает черты лица под небольшим углом, примерно до 30.

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.

**Пример разработки программы**

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.

Именно использование при вычислениях интегрального представления изображения, которое рассчитывается один раз и в дальнейшем используется для расчета признаков, позволило создать быстрый алгоритм поиска объектов.

В алгоритме Виолы-Джонса для обучения классификатора используется технология бустинга, а точнее вариация бустинга – AdaBoost. Бустинг (от англ. boosting – повышение, усиление, улучшение) означает «усиление» «слабых» моделей. Это процедура последовательного построения композиции алгоритмов машинного обучения, когда каждый следующий алгоритм стремится компенсировать недостатки композиции всех предыдущих алгоритмов.

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.

**Пример разработки программы**

Программа обнаружения на видео **лица и глаз человека**, используя каскады Хаара, которая проверяет **наличие маски на лице** (один из способов: проверить наличие рта внутри контура лица). Для работы понадобились готовые обученные классификаторы XML, которые можно найти на официальной странице OpenCV на GitHub или в директории установки библиотеки на локальном компьютере.

**Пример использования программы**

**Компьютерное зрение Vs Обработка изображений**

Обработка изображений изучает преобразование изображения в изображение. Вход и выход обработки изображения – оба изображения.  
  
Компьютерное зрение – это построение явных, значимых описаний физических объектов по их изображению. Результатом компьютерного зрения является описание или интерпретация структур в трехмерной сцене.

**Разница между компьютерным зрением (CV) и компьютерной графикой (CG)**

Компьютерная графика (Computer Graphics) и Computer Vision (Computer Vision) - это два направления в одном процессе.  
  
CG преобразует абстрактную семантическую информацию в изображения, а CV извлекает абстрактную семантическую информацию из изображений.

Компьютерная графика: входные данные - это описание виртуальной сцены, обычно это массив многоугольников, каждый многоугольник состоит из трех вершин, и каждая вершина включает трехмерные координаты, координаты текстуры, цвета RGB и т. Д. На выходе получается изображение, представляющее собой двумерный массив пикселей.

Компьютерное зрение: ввод представляет собой изображение или последовательность изображений, обычно с камеры, видеокамеры или видеофайла. Результатом является понимание реального мира, соответствующего последовательности изображений, например, распознавание лиц и распознавание автомобильных номеров.

**Заключение**

В заключение я постараюсь дать ответ на вопрос: «Зачем компьютеру зрение?».

Научить компьютер видеть мир — значит получить надежного помощника, который день и ночь будет занят решением важных задач. Некие опасения, конечно, присутствуют, но объективная польза от технологий и будущий прогресс, которые они могут дать, перевешивают доводы из теории «восстания машин». Компьютерное зрение – уникальный инструмент, и, рассмотренные примеры использования компьютерного зрения — это лишь малая доля того, на что способен «прозревший» компьютер!

**Спасибо за внимание!**