

КОМПЬЮТЕРНОЕ ЗРЕНИЕ. БИБЛИОТЕКА OPENCV И ЕЁ ПРИМЕНЕНИЕ В ПРИКЛАДНЫХ ЗАДАЧАХ РОБОТОТЕХНИКИ

Суков А. В., студ.; Суков С. Ф., доц., к.т.н.

(ГОУВПО «Донецкий национальный технический университет», г. Донецк, ДНР)

Компьютерное зрение является одной из самых важных технологий для будущего развития робототехники. Данная область развивается уже на протяжении 40 лет, однако развивающиеся в последнее время технологии цифровой визуализации делают проблему автоматизированной интерпретации изображений всё более актуальной. Как научная дисциплина, компьютерное зрение относится к теории и технологии создания искусственных систем, которые получают информацию из изображений. Видеоданные могут быть представлены множеством форм, таких как видеопоследовательность, изображения с различных камер или трехмерными данными, например, с устройства Kinect или медицинского сканера.

Как технологическая дисциплина, компьютерное зрение стремится применить теории и модели компьютерного зрения к созданию систем компьютерного зрения. Примерами применения таких систем могут быть:

1. Системы управления процессами (промышленные роботы, автономные транспортные средства).
2. Системы видеонаблюдения.
3. Системы организации информации (например, для индексации баз данных изображений).
4. Системы моделирования объектов или окружающей среды (анализ медицинских изображений, топографическое моделирование).
5. Системы взаимодействия (например, устройства ввода для системы человеко-машинного взаимодействия).
6. Системы дополненной реальности.
7. Вычислительная фотография, например, для мобильных устройств с камерами.

Одним из наиболее распространённых программных обеспечений для реализации компьютерного зрения является OpenCV.

OpenCV — библиотека алгоритмов компьютерного зрения, обработки изображений и численных алгоритмов общего назначения с открытым кодом. Реализована на C/C++, также существуют версии для Python, Java, Ruby, Matlab, Lua и других языков. Она разрабатывается с 1998 г., сначала в компании Intel, затем в Itseez при активном участии сообщества.

Библиотека распространяется по лицензии BSD, что означает, что ее можно свободно и бесплатно использовать как в открытых проектах с открытым кодом, так и в закрытых, коммерческих проектах. Библиотеку не обязательно копировать целиком в свой проект, можно использовать куски кода. Единственное требование лицензии – наличие в сопровождающих материалах копии лицензии OpenCV.

Из-за либеральной лицензии библиотека используется многими компаниями, организациями, университетами, например, NVidia, Willow Garage, Intel, Google, Stanford University. Компании NVidia и WillowGarage частично спонсируют ее разработку.

Среди проектов, в которых используется OpenCV, можно отметить следующие:

- Система зрения робота PR2, разработанного компанией WillowGarage. Проект PR2 нацелен на решение сложной задачи – создание платформы для персональной робототехники.
- Аудио-визуальная инсталляция в Музее Современного Искусства в Сан-Франциско.
- Контроль качества монет, изготавливаемых Центробанком Китая.
- Курсы компьютерного зрения в Стэнфорде.
- Панорамы улиц в картах Google.

Многомерная архитектура проекта представлена на рис. 1. Библиотека состоит из 16 модулей, реализовано около 1000 алгоритмов. Поддерживаются основные операционные системы: MS Windows, Linux, Mac, Android, iOS. Есть возможность использования сторонних библиотек, например, для работы с устройством Kinect (OpenNI), разработки параллельных программ (TBB) и др.

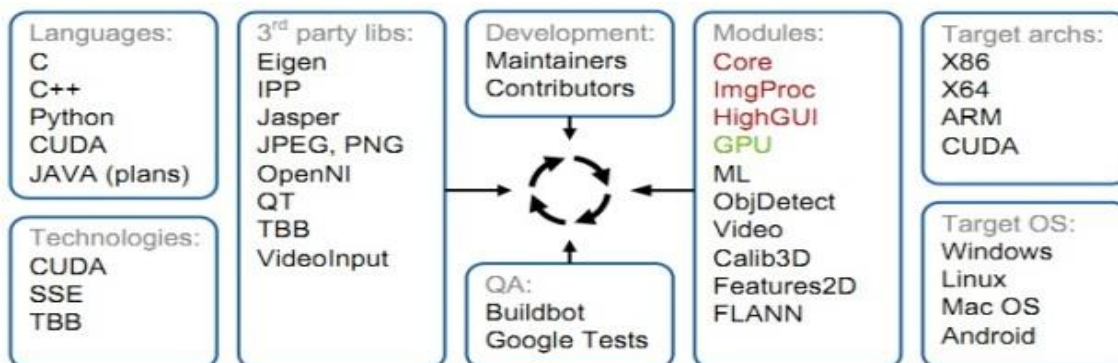


Рисунок 1 - Архитектура OpenCV

Основные модули библиотеки можно отнести к 4 группам (разделам):

- Модули Core, HighGUI, реализующие базовую функциональность (базовые структуры, математические функции, генераторы случайных чисел, линейная алгебра, быстрое преобразование Фурье, ввод/вывод изображений и видео, ввод/вывод в форматах XML, YAML и др.).
- Модули ImgProc, Features2D для обработки изображений (фильтрация, геометрические преобразования, преобразование цветовых пространств, сегментация, обнаружение особых точек и ребер, контурный анализ и др.).
- Модули Video, ObjDetect, Calib3D (калибровка камеры, анализ движения и отслеживание объектов, вычисление положения в пространстве, построение карты глубины, детектирование объектов, оптический поток).
- Модуль ML, реализующий алгоритмы машинного обучения (метод ближайших соседей, наивный байесовский классификатор, деревья решений, бустинг, градиентный бустинг деревьев решений, случайный лес, машина опорных векторов, нейронные сети и др.).

На рисунке 2 представлена общая схема типичного приложения, предназначенного для решения той или иной задачи компьютерного зрения.



Рисунок 2 - Общая структурная схема приложения на OpenCV

Все начинается с захвата изображений (модуль HighGUI). Читается изображение из файла или видео с сетевой камеры через сетевой протокол.

Далее осуществляется предварительная обработка (модуль ImgProc), такая, как устранение шума, выравнивание яркости, контраста, выделение и удаление бликов, теней. Например, один и тот же объект при разном освещении выглядит по-разному. В ярком свете

красная машина, движение которой необходимо отслеживать, будет ярко-оранжевой. В пасмурную погоду та же машина будет выглядеть красно-розовой. В этом случае на изображении необходимо выполнить выравнивание цветов. Предобработка может быть как простой, так и может заключать в себе целую сложную технологию.

Следующий этап – выделение особенностей (модули `ImgProc`, `Features2D`). Например, в задаче слежения за объектом это может быть поиск специальных точек на объекте, за которыми легко наблюдать; для задачи детектирования (т. е. обнаружения на изображении) лица – вычисления описания каждого пиксела.

Далее происходит детектирование интересующих нас объектов, выделение значимых частей, сегментация изображения (модули `ImgProc`, `ObjDetect`). Если, например, камера неподвижна, а изображение подвижное, можно использовать алгоритмы вычитания фона.

После этого решается основная задача, такая, как вычисление расположения объекта в 3d, реконструкция 3d структуры, анализ структуры, регистрация и т. п. (модули `Calib3D`, `Contrib`, `Video`, `Stitching`, `Videostab`, `ML`). Например, в задаче склейки панорам изображений – это сопоставление частей разных кадров, определение нужного преобразования. В задаче видеонаблюдения это восстановление траекторий объектов и т. д.

В конце происходит распознавание и принятие конкретных решений (модуль `ML`). Например, в системе видеонаблюдения: появился нежелательный объект в кадре или нет. В задаче детектирования текста – детектирован ли текст, что именно за текст и т. д.

Для иллюстрации принципов обработки изображений с помощью библиотеки `OpenCV` следует рассмотреть реализацию некоторых типовых задач.

- 1.) Программа загружает изображение, имя которого указывается в параметрах запуска программы, применяет к нему серию преобразований и отображает результат. Исходное изображение и результат показаны на рисунке 3.



Рисунок 3 - Преобразование рисунка в монохромное с помощью `OpenCV`

- 2.) Настоящее компьютерное зрение начинается, когда появляется работа с "живыми" камерами. В следующей программе добавим обработку видео. Для каждого кадра будем выполнять те же операции, что и в предыдущей программе для одного изображения (рисунок 4).

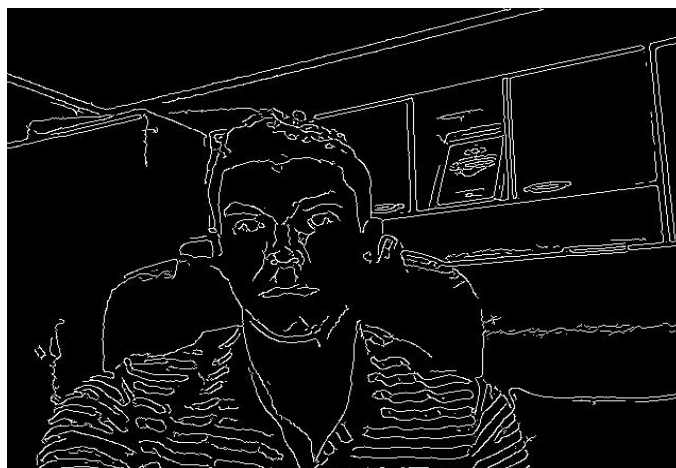


Рисунок 4 – Монохромное изображение с камеры

- 3.) Третья программа: добавляется детектирование лиц к предыдущей программе. Результат представлен на рисунке 5.



Рисунок 5 - Детектирование лица с помощью OpenCV

В условиях, когда качественные устройства захвата изображений становятся вполне доступными для их не только промышленного или военного использования, но и бытового, а вычислительные мощности позволяют совершать миллиарды операций в секунду, давая тем самым возможность их использования в системах реального времени, становится возможным применение компьютерного зрения в робототехнике. Перспективным направлением развития машинного зрения является получение и дальнейшая обработка данных с использованием пассивных устройств захвата изображения, к которым относится камера. В отличие от применения различных датчиков, такой метод вплотную приближает работу системы машинного зрения (особенно, при использовании стереозрения) к той, что использует человек для решения зрительных задач.

Предполагается рассмотреть возможность применения OpenCV для построения системы визуального ориентирования балансирующего робота.

Перечень ссылок

1. Википедия – свободная энциклопедия. OpenCV [Электронный ресурс]. Режим доступа - <https://ru.wikipedia.org/wiki/OpenCV>, свободный.
2. RoboCraft. Компьютерное зрение на Wi-Fi роутере [Электронный ресурс]. Режим доступа - <http://robocraft.ru/blog/computervision/2852.html>, свободный
3. Instructables. Face detection and tracking with Arduino and OpenCV [Электронный ресурс]. Режим доступа - <http://www.instructables.com/id/Face-detection-and-tracking-with-Arduino-and-OpenC/>, свободный.