

СРАВНЕНИЕ БИБЛИОТЕК КОМПЬЮТЕРНОГО ЗРЕНИЯ ДЛЯ ПРИМЕНЕНИЯ В ПРИЛОЖЕНИИ, ИСПОЛЬЗУЮЩЕМ ТЕХНОЛОГИЮ РАСПОЗНАВАНИЯ ПЛОСКИХ ИЗОБРАЖЕНИЙ

Булатников Евгений Владиславович

*доцент кафедры информатики и информационных технологий, кандидат технических наук
Московский государственный университет печати имени Ивана Федорова
127550 Россия, г. Москва, ул. Прянишникова, д. 2А
bev@mgup.ru*

Гоева Анастасия Андреевна

*студентка института принтмедиа и информационных технологий
Московский государственный университет печати имени Ивана Федорова
127550 Россия, г. Москва, ул. Прянишникова, д. 2А
nastyu.goeva@yandex.ru*

Аннотация. В работе проводится сравнение популярной на сегодняшний день библиотеки компьютерного зрения OpenCV и других библиотек компьютерного зрения, в функционал которых входит технология распознавания плоских изображений. Кроме того, рассмотрены само понятие компьютерного зрения и проблемы, с которыми можно столкнуться при работе с технологиями, относящимися к нему.

Ключевые слова: распознавание, библиотеки компьютерного зрения, Open CV, компьютерное зрение, LTI, VXL, AForge.NET.

В процессе разработки игрового мобильного приложения использовались различные технологии. Для распознавания плоских изображений — библиотека OpenCV, для внедрения карт в приложение применялось Google Maps Api, для визуализации трехмерной графики — библиотека OpenGL.

В качестве основной технологии разработанного приложения выступает технология распознавания плоских изображений. Для данной технологии существуют разнообразные библиотеки. В данном приложении использовалась наиболее популярная библиотека компьютерного зрения, а именно, библиотека OpenCV.

OpenCV (Open Source Computer Vision Library) — это библиотека компьютерного зрения, которая поставляется с открытым исходным программным кодом. Данную библиотеку можно скачать по адресу <http://SourceForge.net/projects/opencvlibrary>. Спектр возможностей данной библиотеки очень широк. В ней собрано большое количество алгоритмов для использования технологий компьютерного зрения. После подключения данной библиотеки к своему проекту пользователь получает доступ к более чем 500 функций, предназначенных для решения разнообразных задач. Помимо алгоритмов для работы с технологиями компьютерного зрения, данная библиотека применяется и для обработки изображений, содержит большое число численных алгоритмов и многое другое. Данная библиотека реализована на языках программирования C/C++. Однако поставляется также и для других языков,

таких как Python, Java, Ruby, Matlab, Lua и т.д.. Библиотека может использоваться на различных операционных системах, к числу которых относятся Linux, Mac OS X, iOS, Android и др.

Целью разработки данной библиотеки является повышение эффективности вычислений в приложениях реального времени. Язык C, на котором была написана библиотека, является оптимизированным. Библиотека OpenCV способна использовать многоядерные процессоры. В случае, если понадобится автоматическая оптимизация на различных аппаратных платформах Intel, возможно дополнительное приобретение библиотеки IPP, с английского Integrate Performance Primitives. В состав данной библиотеки входят процедуры с низкоуровневой оптимизацией, которые могут применяться для разнообразных алгоритмических областей. В спектр возможностей OpenCV входит автоматическое применение IPP во время выполнения какой-либо программы.

Главной целью библиотеки OpenCV является предоставление легкого в использовании интерфейса, который поможет облегчить использование технологий компьютерного зрения в довольно сложных приложениях. Функции, которые поддерживает библиотека, охватывают разнообразные сферы компьютерного зрения, от медицины, безопасности и до стереозрения и робототехники. Все это благодаря тому, что компьютерное зрение и машинное обучение — два неразрывно связанных понятия. Кроме того, библиотека OpenCV содержит библиотеку MLL, с

английского Machine Learning Library. Данная библиотека является библиотекой общего назначения и ориентирована на распознавание статических образов и технологию кластеризации. Данная библиотека является очень эффективной для решения задач компьютерного зрения, которое как раз и является основой OpenCV. Однако для решения конкретных задач машинного обучения данная библиотека не приспособлена и является довольно обобщенной.

Многие люди понимают, насколько важно компьютерное зрение на сегодняшний день. Оно применяется в разнообразных областях. Всем известно, что компьютерное зрение применяется в системах видео наблюдения, при работе с графическими и видео файлами в Интернете. Однако на этом области применения компьютерного зрения не ограничиваются. Не многие знают, насколько широко популярна данная технология в компьютерных играх, в аэросъемке и уличных картах. Вряд ли многие догадываются, что компьютерное зрение работает и в сервисе Google's Street View. Данный сервис использует технологию калибровки камеры и технологию сшивания нескольких изображений. Кроме того, компьютерное зрение очень популярно в таких областях как безопасность, конструирование беспилотных самолетов и в разработке биомедицинских анализов. Также широкое распространение компьютерное зрение получило в производстве — в определенный момент технологического процесса практически любое массовое производство использует технологию компьютерного зрения.

Библиотека OpenCV имеет лицензию, которая была разработана таким образом, что у разработчика появляется возможность создавать коммерческий продукт, при этом использовать функциональные возможности библиотеки либо частично, либо полностью. Основным условием использования OpenCV является тот факт, что разработчик обязан делать свой проект с открытым исходным кодом, а также делать доступными для всех разработанные им улучшения для библиотеки. На сегодняшний день существует сообщество, участники которого делятся опытом и обсуждают различные темы и проблемы, связанные с OpenCV. В данное сообщество входят разработчики таких компаний, как IBM, Microsoft, Intel, Google и многие другие. Также участниками сообщества являются члены научно-исследовательских центров, например, Stanford, Cambridge и т.д. Данная библиотека становится все более популярной по всему миру. На данный момент ей уже пользуются в таких странах, как Китай, Япония, Россия и др.

Альфа-версия библиотеки была выпущена в январе 1999 года. И уже на сегодняшний день она использовалась в огромном числе приложений и научно-исследовательских работах. OpenCV применялась для сшивки спутниковых карт, уменьшения шума медицинских изображений, выравнивания отсканированных снимков, анализа разнообразных объектов, для системы обнаружения

вторжения, автоматического мониторинга, системы контроля, калибровки камеры, в приложениях для военно-вооруженных сил, в беспилотных аппаратах, наземных и подводных аппаратах. Кроме того, библиотека имеет применение при распознавании различных звуков и музыки, при помощи анализа спектрограмм.

Как было сказано ранее, библиотека OpenCV является библиотекой компьютерного зрения. Компьютерное зрение — это процесс преобразования данных, которые были получены с устройств фото или видео съемки, в новое представление. Данные преобразования выполняются с какой-то определенной целью. В качестве входных данных может быть какая-либо контекстная информация, например, «камера установлена в машине» или «датчик глубины определил объект в радиусе 1 метра». Выходными параметрами для таких входных данных могут быть следующие решения: «есть ли человек в данной сцене» или «есть 14 клеток опухоли на изображении». Кроме того, новым представлением, полученным из данных с камеры, может быть черно-белое изображение, полученное из цветного или устранение эффекта движения камеры из последовательности кадров.

Задачи компьютерного зрения не являются простыми задачами. Например, процесс поиска на изображении с камеры автомобиля по его образу — не такая уж и легкая задача. Если сравнивать компьютерное зрение с человеческим, то процесс нахождения автомобиля на снимке человеком можно представить как следующую последовательность действий. Для начала человеческий мозг разделяет сигнал, который поступил от зрительного аппарата, на множество каналов. Каждый канал передает информацию различного рода в человеческий мозг. Устройство мозга таково, что он концентрируется на главной информации снимка, при этом остальная информация исключается. На следующем этапе появляется ответный сигнал, перемещающийся в зрительном канале. Кроме того, существуют сигналы, которые поступают на ассоциативные входы. Такие сигналы, которые пришли от датчиков контроля мышц и других чувств, дают возможность мозгу опираться на перекрестные ассоциации, которые были накоплены за все прожитые годы. Из-за наличия обратной связи в мозговом отделе, процесс повторяется снова и снова и включает в себя глаз, который механически управляет освещением с использованием радужной оболочки и настраивает прием на поверхности сетчатки. Из описанного выше видно, что процесс распознавания глазами объекта на снимке является довольно сложным процессом и требует выполнения немалого числа действий человеческим мозгом.

Однако технология компьютерного зрения устроена несколько иначе. Входными данными, которые получает компьютер, является сетка с числами от камеры, либо с диска. У компьютерного зрения нет тех возможностей, которыми управляет человеческий глаз. Отсутствует возможность встроенного распознавания образов, нет

автоматического управления фокусом и диафрагмой, нет перекрестных ассоциаций с опытом работы, накопленным за всю человеческую жизнь. Существующие на сегодняшний день системы компьютерного зрения являются довольно простыми.

Можно привести пример. Возьмем какую-либо картину. Смотри на нее, человек видит именно картину с различными цветами и линиями на ней. Компьютер в свою очередь видит только сетку с различными числами. Каждое значение в данной сетке имеет большую шумовую составляющую и, если рассматривать его отдельно от всех остальных чисел, то оно может дать нам очень маленькое количество информации. Однако именно такую сетку и видит компьютер. Задача, которая ставится перед разработчиком — это преобразование сетки, которую видит компьютер, в картину, которую видит человек.

Кроме того, существует еще одна проблема, связанная с компьютерным зрением. Данная проблема связана с тем, что компьютер воспринимает только двумерные изображения, в свою очередь как каждый предмет обладает трехмерностью. Восприятие трехмерного объекта может радикально меняться в зависимости от позиции, с которой на него смотреть. Данная проблема не имеет однозначного и окончательного решения, имея только двумерное представление трехмерного мира, невозможно однозначно восстановить трехмерный сигнал. 3d сцену можно представить как очень большое количество двумерных изображений. Однако данное представление не является совершенным. Это связано с тем, что данные могут подвергаться различным шумам и искажениям. А также возможны изменения в окружающем пространстве, например, изменения погоды, освещения, отражения, перемещения и многого другого. Также проблемы могут возникнуть из-за каких-либо недостатков в объективе, механических установках, конечном времени интеграции на датчике, которое может привести к размытию движения, электрическим шумам на датчике или разнообразной другой электронике, артефактам, появившимся после сжатия и захвата изображения. Чтобы объект правильно распознавался, необходимо учитывать все перечисленные выше проблемы.

Очень часто при проектировании практической системы могут учитываться дополнительные знания контекста. Эти знания могут применяться для того, чтобы обойти ограничения, которые были наложены визуальными датчиками. В качестве примера можно рассмотреть следующую ситуацию. Например, есть робот, задачей которого является нахождение ручки на сцене, полученной с камеры. Робот может опираться на следующий факт, стол — это объект, который очень часто располагается в офисах и рабочих пространствах, и именно на столе можно найти ручку. Знание данного факта дает ссылку на размер ручки, она явно должна располагаться на столе, т.е. ее размер должен быть много меньше размеров стола. Кроме того, использование данного факта помогает устранить некоторые ложные

распознавания ручки в местах, где она не может располагаться, например, на кровати или на стене.

Если, например, за окном есть рекламный постер с изображением ручки, то робот может его проигнорировать, это связано с тем, что на фоне ручки нет древесины стола. Существуют также случаи, когда изображение приходится принимать из базы данных, в которой ручка может оказаться очень больших размеров или, например, быть необычной формы. Данный факт повлечет за собой то, что данные изображения будут исключены из рассмотрения. Также необходимо учитывать, что обычно люди снимают предметы в центре изображения и спереди, например, никто не будет снимать человека вверх ногами. Все это дает неявную информацию по предмету, который необходимо распознать.

Информацию по контексту можно также смоделировать, используя методы машинного обучения. Данный способ является способом явного моделирования. Скрытые переменные, к числу которых можно отнести, размер, ориентацию, положение предмета и многое другое, могут быть соотнесены с их значениями в маркированном обучающем множестве. Также можно попытаться изменить скрытые переменные смещения с помощью некоторых дополнительных датчиков. С помощью датчика глубины можно довольно точно измерить размер объекта.

Еще одна проблема компьютерного зрения — это шум. Очень часто с понятием шумов можно столкнуться, когда используются статистические методы. Часто становится невозможным определение контура путем сравнения соседних точек. Однако если собирать статистику по локальной области, то задача обнаружения контура облегчается. Контур получается в виде строки ответов по локальной области, каждый из которых ориентирован в соответствии со своими соседями. Также существует возможность компенсации шума путем получения статистических данных в течение довольно долгого периода времени. Однако существуют и другие методы учета шумов и искажений. Данные методы основаны на создании четких моделей, позволяющих распознавать предмет из имеющихся данных. В качестве примера искажений, которые можно так исправить, можно привести искажения от объектива. Для того чтобы полностью избавиться от таких искажений, необходимо только знать параметры простой полиномиальной модели.

Все действия, которые выполняются в компьютерном зрении, происходят из расчета тех данных, которые были получены с камеры в контексте поставленной цели и задачи. Например, можно избавиться от шума и повреждений на изображении таким образом, что система безопасности будет оповещать, если кто-то пересек поставленную границу. Однако программное обеспечение, которое будет использовать технологии компьютерного зрения для роботов, перемещающихся, например, по территории офиса,

будет пользоваться несколько иными подходами, чем стационарные камеры. Это связано с тем, что две эти системы имеют различные контексты и конечные цели. Очень часто контекст задач компьютерного зрения является ограниченным и, учитывая все эти ограничения, поставленную задачу можно во многом упростить, что приведет к тому, что конечное решение будет более правильным.

Библиотека OpenCV направлена как раз на обеспечение работы инструментов для устранения описанных выше проблем компьютерного зрения. Функционала библиотеки будет достаточно для решения довольно сложных задач в области компьютерного зрения. Однако есть такие задачи, для решения которых функционала библиотеки недостаточно. В связи с этим, можно на основе компонентов библиотеки создать собственное решение задачи почти любой сложности. Существуют несколько проверенных и довольно надежных методов, которые используют компоненты библиотеки. После создания проекта можно найти некоторые слабые места и исправить путем дописания собственного кода. Полученное решение можно использовать в качестве эталонного и применять в дальнейшем. Таким образом, любое слабое место можно исправить, используя встроенное решение.

OpenCV появилась благодаря Intel Research с целью продвижения ресурсоемких приложений. С этой целью компанией был запущен ряд проектов. Одному из сотрудников компании пришла идея, что разработчику не нужно изобретать функции самому, а использовать те, что были написаны до него и лишь дописывать к ним необходимые дополнения. Идея OpenCV состояла в том, чтобы сделать компьютерное зрение доступным каждому. Таким образом, была поставлена задача реализации кода ядра и алгоритмических спецификаций OpenCV. На начальных этапах при разработке OpenCV ставились следующие задачи. Было необходимо разработать не только открытый, но и хорошо оптимизированный код, что поспособствовало бы исследованиям в области компьютерного зрения. Еще одной задачей было распространение знаний через общую инфраструктуру, чтобы разработчики могли разрабатывать приложения с легко читаемым и передаваемым кодом. Также задачей стало продвижение коммерческих приложений, использующих компьютерное зрение и собранных с общедоступным оптимизированным кодом.

Приложения, в которые были включены технологии компьютерного зрения, становились более востребованы в быстрых процессорах. Обновление драйверов приносило компании Intel большой доход. Возможно, именно по этой причине идея открытого кода возникла у компании, поставляющей оборудование, а не у компаний, производящих дополнительное программное обеспечение. Кроме того, компании-производители оборудования менее ограничены в создании инновационного программного обеспечения.

На данный момент библиотека OpenCV скачана около 2 млн раз. Число пользователей, активно

работающих с ней, достигается 20 тыс. Разработчики вносят все больший вклад в развитие библиотеки.

Постепенно значение OpenCV в области приложений, использующих компьютерное зрение, стало расти. На сегодняшний день библиотека активно развивается, постоянно выходят обновления. В скором времени появятся обновления, связанные с калибровкой сразу нескольких камер, восприятием глубины сцены, методами смешивания зрения с датчиками глубины, улучшенными возможностями распознавания образов и большей поддержкой для роботов.

Библиотека OpenCV имеет функцию ручной настройки на использование высокооптимизированного кода IPP. На рис. 1 представлено сравнение OpenCV и OpenCV с IPP с двумя другими библиотеками компьютерного зрения LTI и VXL. Производительность являлась основной задачей OpenCV, в связи с тем, что требовалось использовать компьютерное зрение в режиме реального времени. Тот факт, что OpenCV была написана с помощью высокопроизводительного кода на языках C и C++ никак не связано с IPP, который автоматически подключается к OpenCV для улучшения производительности.

Как видно из диаграммы, представленной на рис. 1, по всем четырем критериям эффективности библиотека OpenCV имеет лучшие результаты, OpenCV + IPP превосходит OpenCV без IPP (диаграмма показывает результаты, пропорциональные времени выполнения для каждой из библиотек по четырем критериям эффективности).

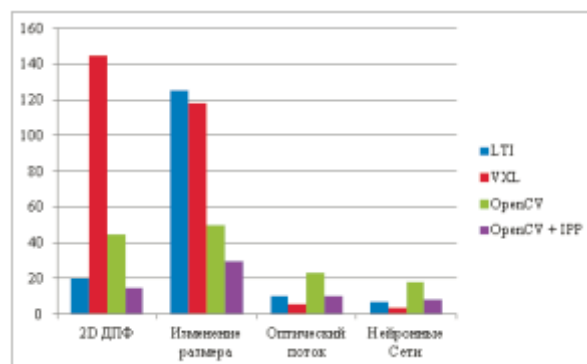


Рис. 1. Сравнение OpenCV (с и без IPP) и библиотек компьютерного зрения LTI и VXL

Еще одним преимуществом использования библиотеки OpenCV в своем проекте является наличие HTML-документации, которая поставляется вместе с исходным кодом библиотеки. В папке `.../opencv/docs` можно найти документ `index.html`, загрузив данный документ можно увидеть список ссылок, каждая из которых отвечает за определенные группы алгоритмов, содержащихся в библиотеке, например, ссылка `CV` содержит обработку изображений, анализ структуры изображений, движение слежения, распознавание образов и калибровку камеры, Machine Learning (ML) — функции кластеризации, классификации и анализа

данных, CVCAM — работа с камерой и т.д. Документация, представленная в таком виде, помогает при разработке проекта. Кроме того, существует более новый тип документации по библиотеке OpenCV, а именно, документация в формате Wiki. Документация содержит инструкции по сборке OpenCV, используя среду разработки Eclipse IDE, распознавание лиц с OpenCV, библиотеку видеонаблюдения, список литературы, совместимые камеры и ссылки на сообщества. Данную документацию можно найти по адресу <http://opencvlibrary.SourceForge.net>.

Существует также документация в формате Wiki с более уникальной информацией по различным вспомогательным функциям. Данная документация содержит описание следующих понятий: стерео соответствие, точка зрения морфинга камер, 3d слежение в режиме стерео, объект соответствия (PCA), функции для распознавания объектов и внедрение скрытой Марковской модели (СММ).

Основная документация по библиотеке OpenCV представлена на английском языке, однако на сегодняшний день существуют и русскоязычные аналоги данной документации, что во многом способствует тому факту, что большинство русскоговорящих разработчиков используют именно OpenCV при создании своих проектов.

Библиотека OpenCV является структурированной библиотекой. Она разделена на пять основных компонентов: алгоритмы обработки изображений, высокоуровневые алгоритмы компьютерного зрения, MLL — библиотека машинного обучения, HighGUI — процедуры и функции ввода и вывода для хранения и загрузки видео и изображений, CXCore — основные структуры данных.

Структура библиотеки представлена на рис. 2. На данном рисунке не представлен еще один — пятый компонент библиотеки, а именно CvAux (экспериментальные алгоритмы). CvAux содержит: 1D и 2D скрытая Марковская модель, техника статистического распознавания с использованием метода динамического программирования, Встроенная СММ, Распознавание жестов благодаря стерео зрению, Стереозрение, Deskriptor текстур, Поиск скелета (центральных линий) объектов сцены, видеонаблюдение и т.д.



Рис. 2. Базовая структура OpenCV

Еще одним преимуществом использования OpenCV является ее переносимость. OpenCV разрабатывалась как портативная библиотека. В самом начале разработка велась на Borland C++,

MSVC++, и Intel compilers. Все это означало тот факт, что код на языках C и C++ должен был быть стандартным, чтобы создать облегченную поддержку кроссплатформенности. В начале разработки поддерживаемая архитектура была 32-битной Intel с операционной системой Windows, немного позже с поддержкой ОС Linux. После того как компания Apple начала использовать процессоры Intel, появилась также поддержка MAC OS X. Немного позже появилась поддержка и 64-битной архитектуры Intel. На сегодняшний день библиотека OpenCV адаптирована почти под все коммерческие системы, кроме того, она работает и на процессорах AMD, в которых также доступна IPP для улучшения производительности (при использовании MMX — мультимедийного расширения).

Помимо библиотеки OpenCV, существуют также и другие библиотеки компьютерного зрения. Например, библиотеки Matrox Imaging Library, Camellia Library, Open eVision, HALCON, VXL, libCVD, IVT, LTI, AForge.NET и некоторые другие.

AForge.NET является библиотекой с открытым исходным кодом, созданной на языке C#, которая предназначена для разработчиков и исследователей в области компьютерного зрения. Кроме того, в библиотеке есть функционал для разработчиков в области искусственного интеллекта. Спектр возможностей библиотеки довольно широк: обработка изображений, нейронные сети, генетические алгоритмы, нечеткая логика, машинное обучение, робототехника и многое другое.

Библиотека включает несколько основных компонентов. AForge.Imaging — библиотека подпрограмм для обработки изображений и фильтров. AForge.Vision — библиотека компьютерного зрения. AForge.Video — набор библиотек для работы с видео информацией. AForge.Neuro — библиотека для выполнения разнообразных действий и операций с нейронными сетями. AForge.Genetic — библиотека подпрограмм для использования генетических алгоритмов для решения различных задач. AForge.Fuzzy — библиотека для работы с нечеткой логикой. AForge.Robotics — библиотека, обеспечивающая поддержку некоторых методов, применяемых в сфере робототехники. AForge.MachineLearning — библиотека для работы с элементами машинного обучения. AForge.NET постоянно улучшается и прогрессирует. По данной библиотеке есть большое количество примеров, демонстрирующих ее работу, а также актуальная html-документация, которая помогает начинающим разработчикам, которые хотят применять данный фреймворк в своих проектах. Кроме того, как и у библиотеки OpenCV существует сообщество, где можно задавать вопросы и делиться своими наработками по функциям и компонентам AForge.NET. Однако количество участников сообщества по данной библиотеке все-таки меньше, чем в сообществах по библиотеке OpenCV. Данный фреймворк уступает по популярности OpenCV. Еще одним недостатком является тот факт, что вся документация по библиотеке написана только на

английском языке, поэтому у некоторых разработчиков могут возникнуть трудности с изучением и работой с данной библиотекой.

VXL, от английского the Vision-something-Libraries, — это набор библиотек, написанных на языке C++, которые предназначены для научных исследований и реализации технологий компьютерного зрения. VXL была написана в ANSI/ISO C++ и предназначена для портативных платформ. Библиотека состоит из нескольких основных составляющих: VNL (числа) — численные алгоритмы и контейнеры, например, матрицы, векторы, оптимизаторы и т.д., VIL (изображения) — загрузка, сохранение и редактирование изображений во многих наиболее распространенных форматах (также существует возможность работы с очень большими изображениями), VGL (геометрия) — геометрия точек, кривых и других элементарных объектов в одно-, двух- и трехмерном пространствах, VSL (входный и выходной потоки), VBL (основные шаблоны), VUL (утилиты) — разный функционал для независимых платформ. Кроме основных библиотек входящих в состав VXL, есть также и дополнительные. Дополнительные библиотеки отвечают за такие понятия, как численные алгоритмы, обработка изображений, системы координат, геометрия камер, стерео, манипуляции с видео потоком, восстановление структуры при движении камеры, графический дизайн, функции отслеживания, топология, классификаторы, 3d визуализация и многое другое. Особенность библиотеки заключается в том, что каждый ее компонент может использоваться отдельно, не ссылаясь на другие компоненты библиотеки. Таким образом, в приложении можно использовать только те библиотеки, которые действительно необходимы. VXL используется по всему миру. Библиотека применяется в сфере обучения и промышленности, некоторые ведущие мировые эксперты в сфере компьютерного зрения пользуются данной библиотекой. Существует документация по VXL с описанием каждого класса и функций. Однако недостатком является тот факт, что аналогично AForge.NET вся документация написана на английском языке.

LTI или LTI-lib является объектно-ориентированной библиотекой алгоритмов и структур данных, часто применяется при обработке изображений и в сфере компьютерного зрения. LTI-lib была разработана в техническом университете как часть научно-исследовательских проектов в области компьютерного зрения с технологиями робототехники, распознавания объектов, голоса и жестов. Основной целью разработки данной библиотеки является создание объектно-ориентированной библиотеки на языке C++, что во многом упрощало бы использование кода и его обслуживание, но при этом были бы обеспечены быстрые алгоритмы, которые можно было бы использовать в реальных приложениях.

Библиотека была разработана с применением GCC (набор компиляторов, применяемый для

разнообразных языков программирования) под Linux и Visual C++ по Windows NT. Многие классы инкапсулируют Windows/ Linux функциональность для того, чтобы упростить решение системных или аппаратных задач (например, классы для многопоточности и синхронизации, измерения времени и доступ к последовательному порту).

Остальные классы, количество которых превышает 500, выполняют одну из следующих функций. Функции линейной алгебры: предоставление матриц, векторов, решения линейных уравнений, собственные векторы, расчет статистики и многое другое. Функции классификации и кластеризации: радиально базисные функции-классификаторы, методы опорных векторов, метод K-средних, нечеткие C-средства, классификация статистических данных. Функции для обработки изображений: разнообразные подходы к сегментации, линейные фильтры, вейвлет-преобразования и многое другое.

Предоставление инструментов для визуализации и рисования. Самое сложное при разработке алгоритмов обработки изображений в C++ это представление временных образов во время отладки. Благодаря объектно-ориентированной архитектуре библиотеки LTI-lib, для предпросмотра необходимо создать объект представления и дать ему образ, который необходимо показать. Кроме того, если необходимо нарисовать некоторую дополнительную информацию на изображении, например, текст, эллипс, прямоугольник, линию или точку, можно использовать один из объектов рисования. Такой подход позволяет экономить время.

По данной библиотеке также существует документация, которая предоставляется в свободном доступе в Интернете. Документация по библиотеке написана на английском языке. Что касается лицензии, то библиотека LTI, также как и OpenCV, является свободно распространяемым программным обеспечением. Все пункты и ограничения по лицензии значатся в GNU Lesser General Public License, кроме того, там же можно посмотреть более подробную информацию по лицензии. Существует большое количество проектов, разработанных с применением технологий компьютерного зрения, для которых использовалась именно библиотека LTI.

В качестве библиотеки для создания игрового мобильного приложения с использованием технологии распознавания плоских изображений применялась наиболее популярная на сегодняшний день библиотека OpenCV. Такой выбор связан с тем, что по сравнению с некоторыми другими библиотеками компьютерного зрения, данная библиотека обладает некоторыми преимуществами.

Первым и наиболее значительным преимуществом библиотеки является ее производительность. Как видно из диаграммы на рис. 1, производительность OpenCV превышает производительность таких библиотек компьютерного зрения как VXL и LTI. Кроме того, используя дополнительный компонент IPP, можно в разы увеличить производительность библиотеки OpenCV.

Наличие большого числа функций (более 500), решающих разнообразные задачи от обработки изображений, компьютерного зрения и до обучения машин, также сыграло роль при выборе в качестве библиотеки компьютерного зрения OpenCV. Кроме того, библиотека имеет открытый исходный код и лицензию, которая позволяет даже с использованием функционала библиотеки выпускать коммерческие продукты.

Еще одним преимуществом, которое учитывалось при выборе библиотеки компьютерного зрения, является наличие русскоязычной документации по библиотеке. Помимо этого, существует огромное число tutorиалов, научных материалов и книг, в которых описываются функции и методы работы с OpenCV. Также в Интернете можно найти огромное число сообществ, в которых разработчики и пользователи библиотеки делятся опытом и отвечают на вопросы, связанные с библиотекой.

Последним и наиболее важным моментом при выборе библиотеки компьютерного зрения стал тот факт, что библиотека OpenCV имеет реализацию для мобильной платформы Android — OpenCV for Android SDK.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Bradski G.R. Learning OpenCV/ Dr. Gary Rost Bradski, Adrian Kaehler* — Изд.: O'Reilly Media, Inc., 2008. — 556 с.
2. *Laganière R. OpenCV 2 Computer Vision Application Programming Cookbook/ R.Laganière* — Изд.: Packt Publishing, 2011. — 298 с.
3. *D.L. Baggio Mastering OpenCV with Practical Computer Vision Projects/ Baggio D. L., Emami S., Escrivá D.M.* — Изд.: Packt Publishing, 2012. — 321 с.

COMPARING COMPUTER VISION LIBRARIES FOR USING IN APPLICATION THAT USES THE RECOGNITION TECHNOLOGY OF FLAT IMAGES

Evgeny Vladislavovich Bulatnikov

*Moscow State University of Printing Arts
127550 Russia, Moscow, Pryanishnikova st., 2A*

Anastasiya Andreevna Goeva

*Moscow State University of Printing Arts
127550 Russia, Moscow, Pryanishnikova st., 2A*

Annotation. *The paper compares the popular today OpenCV computer vision library, and other computer vision libraries, which have the functional part of the flat image recognition technology. In addition the paper conducts a review of the concepts of computer vision and the challenges you might face when working with computer vision technologies.*

Keywords: recognition, computer vision libraries, Open CV, computer vision, LTI, VXL, AForge.NET.