

Бекбулатов Р.Н. Анализ средств детектирования и трекинга тела человека // Академия педагогических идей «Новация». – 2020. – №1 (январь). – АРТ 14-эл. – 0,2 п. л. – URL: <http://akademnova.ru/page/875548>

РУБРИКА: ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

УДК 004.93

Бекбулатов Руслан Наильевич
студент 2 курса магистратуры факультета информатики и систем
управления
ФГБОУ ВПО «Московский государственный технический
университет имени Н.Э.Баумана»
г. Москва, Российская Федерация
e-mail: ruslan075@mail.ru

**АНАЛИЗ СРЕДСТВ ДЕТЕКТИРОВАНИЯ И
ТРЕКИНГА ТЕЛА ЧЕЛОВЕКА**

Аннотация: В данной работе приведен анализ современных средств детектирования и трекинга тела человека. Описаны и проанализированы наиболее распространённые на сегодняшний день методы решения данной задачи. Выявлены недостатки, которым подвержены системы, определены наиболее перспективные пути решения задачи.

Ключевые слова: детектирование, трекинг, компьютерное зрение, дополненная реальность.

Bekbulatov Ruslan Nailevich
2th year magistracy students, faculty of informatics and control systems
FGBOU VPO «Bauman Moscow State Technical University»
Moscow, Russian Fedration
e-mail: ruslan075@mail.ru

ANALYSIS OF HUMAN BODY DETECTION AND TRACKING SOLUTIONS

Abstract: Article is provided analysis of modern ways for human body detection and tracking solutions. The most common methods for solving this problem are described and analyzed. The disadvantages that the systems are subject to are identified, and the most promising ways to solve the problem are defined.

Keywords: detection, information technologies, computer vision, body tracking, augmented reality.

Все больше компаний по всему миру начинают интересоваться технологиями компьютерного зрения(CV) и дополненной реальности(AR). Такие IT-гиганты, как Microsoft, Apple, Intel давно ведут разработки в этом направлении.

На сегодняшний день существует много решений по детектированию и трекингу человека. Однако, проблема применения в промышленных масштабах до сих пор актуальна.

Прежде чем приступать к разработке любой системы необходимо рассмотреть преимущества и недостатки возможных методов решения поставленной задачи. Поэтому целью данной работы является анализ существующих аппаратных и программных средств в области детектирования и трекинга человека.

Анализ аппаратных решений

Ниже в таблице 1 представлена сравнительная таблица существующих сенсоров. На основании полученной информации можно сделать вывод, что Kinect V2 (Kinect for XBOX One) имеет лучшие

характеристики. Датчик глубины Orbbec Astra (PRO) располагается на втором месте.

Таблица 1 – Сравнительная таблица сенсоров

Сенсор	Преимущества	Недостатки
Kinect V2 (Kinect for XBOX One)	<ul style="list-style-type: none">– FOV выше аналогов;– Более высокое разрешение глубины и RGB видео;– Низкий уровень шума, лучшее качество и точность облака точек;– Можно ожидать более высокое общее качество отслеживания;– Намного лучше отслеживание ног;– Низкие взаимные помехи датчиков;– Более терпимым к условиям освещения;– Имеет крепление для штатива.	<ul style="list-style-type: none">– Не любит определенные ткани (например, блестящие);– Требуется контроллер USB 3.0 и внешний источник питания.
Orbbec Astra (PRO)	<ul style="list-style-type: none">– Обеспечивает качество отслеживания, сравнимое с качеством Kinect 2, но требования к оборудованию не так высоки;– Требуется контроллер USB 2.0;– Несколько датчиков могут быть подключены к одному компьютеру;	<ul style="list-style-type: none">– FOV ниже Kinect V2– Разрешение ниже– Точность ниже– Менее терпимый к условиям освещения– Более высокая взаимная интерференция датчиков приводит к большему шуму в случае использования нескольких датчиков

	<ul style="list-style-type: none">– В продаже.	<ul style="list-style-type: none">– Нет крепления для штатива
Intel RealSense D435	<ul style="list-style-type: none">– До 90 кадров в секунду;– Высокое разрешение карты глубины;– Хорошее поле зрения (FOV);– Очень компактный;– Отличная техническая поддержка, быстрые исправления, частые обновления.	<ul style="list-style-type: none">– Очень низкое качество карты глубины (высокий уровень шума, много артефактов);– Конфигурация с несколькими датчиками нецелесообразна с этими датчиками из-за точности и качества карты глубины (не может быть откалибрована);– RGB камера имеет значительно меньшее поле зрения.

Kinect Azure DK

Azure Kinect DK – это набор разработчика и периферийное компьютерное устройство с усовершенствованными датчиками на основе искусственного интеллекта (ИИ) для создания комплексных моделей компьютерного зрения и речи.

Компоненты Azure Kinect DK представлены на рисунке 1.

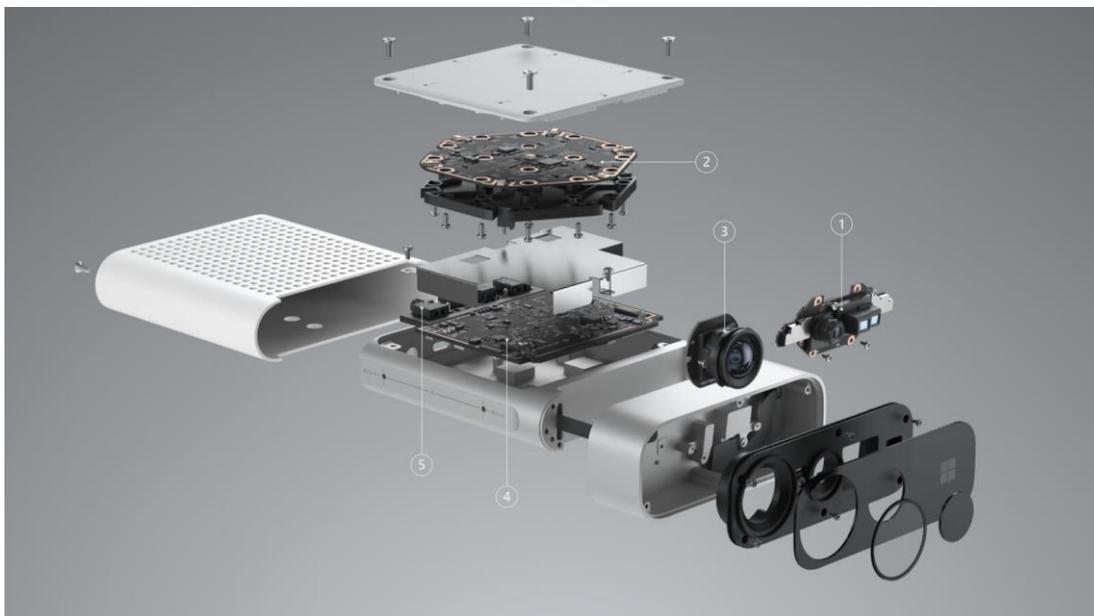


Рисунок 1 – Компоненты Azure Kinect DK

1. Датчик глубины разрешением 1 Мп с возможностью выбора широкого или узкого поля зрения позволяет пользователю оптимизировать устройство для работы с приложением;
2. Массив из 7 микрофонов, обеспечивающий обнаружение голоса в дальней зоне и захват звука;
3. RGB-камера разрешением 12 Мп для дополнительного цветового потока, совмещенного с потоком глубины;
4. Акселерометр и гироскоп, обеспечивающие ориентацию датчиков и пространственное отслеживание;
5. Внешние синхронные фиксации, обеспечивающие простую мгновенную синхронизацию потоков датчиков разных устройств Kinect.

Данное решение совсем недавно поступило в продажу. По заявленным характеристикам уже можно сделать вывод, что Azure Kinect DK станет флагманом среди аппаратных решений по компьютерному зрению.

Анализ программных решений

Рассмотрим популярные программные решения по детектированию и трекингу тела человека:

- OpenPose;
- AlphaPose.

Необходимо проанализировать представленные средства более подробно.

OpenPose

OpenPose является одним из самых популярных подходов снизу вверх для оценки позы человека с несколькими людьми, в том числе из-за их хорошо документированной реализации GitHub.

Как и во многих подходах «снизу вверх», OpenPose сначала обнаруживает части (ключевые точки), принадлежащие каждому человеку на изображении, а затем присваивает детали отдельным лицам. Ниже показана архитектура модели OpenPose (Рисунок 2).

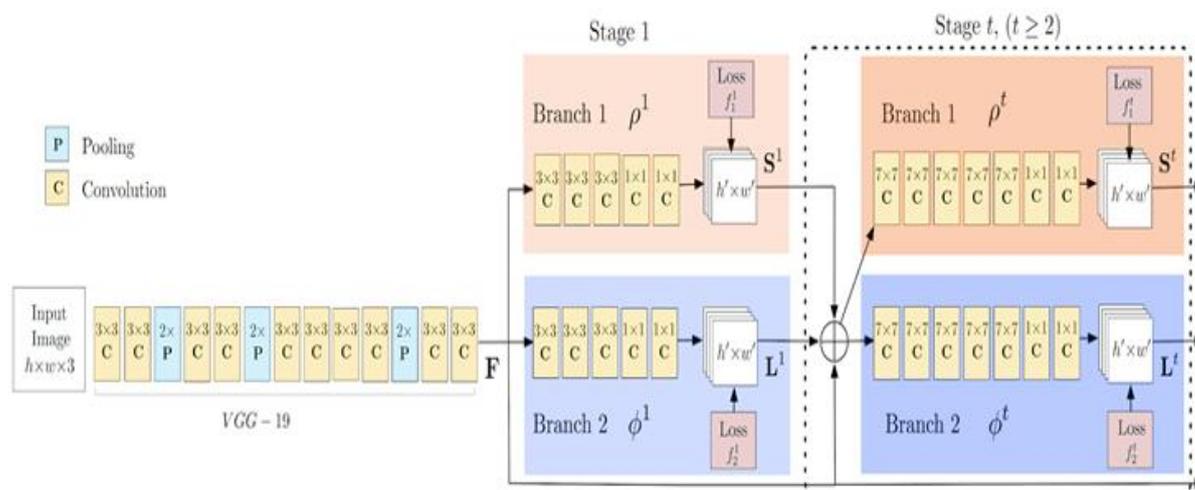


Рисунок 2 – Блок-схема архитектуры модели OpenPose

Сеть OpenPose сначала извлекает элементы из изображения, используя первые несколько слоев (VGG-19 на приведенной выше блок-схеме). Затем элементы подаются в две параллельные ветви сверточных слоев. Первая ветвь предсказывает набор из 18 карт достоверности, каждая из которых представляет определенную часть позы человека. Вторая ветвь предсказывает набор из 38 полей сходства частей (PAF), который представляет степень ассоциации между частями.

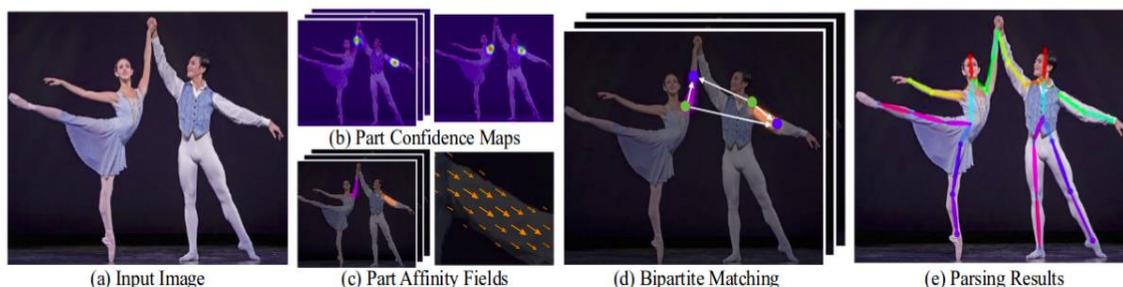


Рисунок 3 – Промежуточные результаты работы алгоритма

Последовательные этапы используются для уточнения прогнозов, сделанных каждой ветвью. Используя карты достоверности деталей, двусторонние графики формируются между парами деталей. Используя значения PAF, более слабые связи в двудольных графах удаляются. Посредством вышеупомянутых шагов человеческие позы могут быть оценены и назначены каждому человеку в изображении.

AlphaPose

AlphaPose является точным оценщиком позы для нескольких человек и претендует на звание первой системы с открытым исходным кодом. AlphaPose выполняет оценку позы и отслеживание поз на изображениях, видео или списках изображений. Он производит различные выходные данные, в том числе изображения с отображением ключевых точек в форматах PNG, JPEG и AVI, а также вывод ключевых точек в формате JSON,

что делает его отличным инструментом для более сфокусированного использования.

В настоящее время существует как реализация TensorFlow, так и реализация PyTorch.

AlphaPose использует региональную структуру оценки позы людей (RMPE) для облегчения оценки позы при наличии неточных рамок, ограничивающих человека. Существует три компонента: сеть симметричных пространственных преобразователей (SSTN), параметрическое представление без максимального подавления (NMS) и генератор предложений, управляемых по позициям (PGPG).

Выводы по рассмотренным средствам

Программные решения требуют очень производительные графические карты. В качестве примера, можно рассмотреть карту Nvidia 1080Ti. Для алгоритма AlphaPose FPS составляет 23. Для OpenPose FPS составляет 21,7. Стоит отметить, что данные получены на основе открытых источников и могут отличаться от реальных значений.

Стоимость карты Nvidia 1080Ti составляет около 80 тысяч рублей. Типичное энергопотребление составляет: 250 Вт. Примем среднюю цену за кВт в размере 5,47 рублей по Москве. Тогда в месяц цена за работу одной карты составит около 985 рублей.

Кроме того, для полноценного тестирования одной картой не обойтись, необходимо собирать ПК.

Также для OpenPose имеется решение с использованием Tensorflow с возможностью запуска на вычислительной платформе Nvidia Jetson TX2. Цена Nvidia Jetson TX2 составляет около 50 тысяч рублей, а FPS при такой комплектации достигает приблизительно 10 FPS, что очень малоприспособно для реальных задач.

По сравнению с Kinect Azure, данные решения выглядят дорогостоящими.

Заключение

Анализ показал преимущества и недостатки возможных решений по детектированию и трекингу позы человека. Наиболее перспективным направлением на данный момент является решение от компании от компании Microsoft Kinect Azure DK.

Большие успехи были достигнуты в области детектирования тела человека. Концепции, перечисленные в данной работе, не являются исчерпывающими, а скорее стремятся представить популярные и наиболее эффективные варианты для реального применения.

Список использованной литературы:

1. Кульневич А. Д. Введение в нейронные сети // Молодой ученый. — 2017. — №8. — С. 31-36. — URL <https://moluch.ru/archive/142/40055/> (дата обращения: 20.05.2019).
2. Depth Sensors Comparison [Электронный ресурс]. URL: http://docs.ipisoft.com/Depth_Sensors_Comparison (дата обращения: 20.05.2019)
3. OpenPose [Электронный ресурс]. URL: <https://github.com/CMU-Perceptual-Computing-Lab/openpose> (дата обращения: 20.05.2019)
4. AlphaPose [Электронный ресурс]. URL: <https://github.com/MVIG-SJTU/AlphaPose> (дата обращения: 20.05.2019)
5. Azure Kinect DK [Электронный ресурс]. URL: <https://azure.microsoft.com/ru-ru/services/kinect-dk/> (дата обращения: 21.05.2019)

Дата поступления в редакцию: 25.01.2020 г.

Опубликовано: 25.01.2020 г.

© Академия педагогических идей «Новация», электронный журнал, 2020

© Бекбулатов Р.Н., 2020