

# 3D-МОДЕЛЬ ЧЕЛОВЕЧЕСКОГО ТЕЛА И ИЗМЕРЕНИЕ ЕГО ПОКАЗАТЕЛЕЙ ПРИ ПОМОЩИ КАМЕРЫ

Кулыба В. А., Гуринович А. Б.

Кафедра информационных технологий автоматизированных систем, Кафедра вычислительных методов и программирования, Белорусский государственный университет информатики и радиоэлектроники

Минск, Республика Беларусь

E-mail: kulyba.vadim@gmail.com, gurinovich@bsuir.by

В данной статье исследуются алгоритмы построения 3d моделей человеческого тела и замеров его ключевых показателей (обхват в области плеч, обхват в области талии, форма лица, положение глаз). Для решения проблемы использованы алгоритмы и модели глубокого обучения. Ключевые слова: 3d модель, машинное обучение, камера с датчиком глубины, преобразование изображений и видео, датасет. Компьютерное зрение — теория и технология создания машин, которые могут производить обнаружение, отслеживание и классификацию объектов.

## ВВЕДЕНИЕ

В настоящее время многие крупные компании работают над задачей построение точных 3d моделей человека. Существует несколько причин заинтересованности в построение 3d моделей: первая — заинтересованность в увеличении объема online продаж одежды; вторая причина — это предполагаемый AR/VR/XR рынок.

### I. СНЯТИЕ ПОКАЗАТЕЛЕЙ ТЕЛА ЧЕЛОВЕКА

В первой части статьи рассмотрим простой способ измерения параметров тела при помощи как обычной камеры, так и глубинной. При условиях близких к идеальным, то есть человек стоит ровно и смотрит в камеру можно использовать обычную камеру. Суть метода заключается в использовании методологии применяемой в глубоком обучении под название pose estimations.

Pose estimations - это одна из распространенных задач глубокого обучения, по построению модели костей человека, по его фотографии, чтобы оценить его позиционирование на изображении или видео.

На данный момент уже существует множество предобученных моделей для задачи pose estimations разного качества. На выходе их отличается лишь точность и датасет на основе которого мы получаем список точек, предполагаемого скелета. Основные датасеты для этой задачи: MPII Human Pose Dataset, COCO dataset.

При помощи этих точек на выходе модели, при условии, что человек стоит ровно можно построить зависимости опираясь на коэффициенты при помощи которых вычислить параметры тела человека. В случае же если человек стоит не ровно, стоит использовать камеру с датчиком глубины, при помощи этой камеры можно вычислить углы наклона и по схожей методике вычислить параметры тела с поправкой на углы.

Однако стоит отметить, что подход является эффективным для снятия только 2d параметров человека.

## II. ПОДХОДЫ ДЛЯ ПОСТРОЕНИЯ 3D МОДЕЛЕЙ ЧЕЛОВЕКА

Для построения 3d моделей человека будем опираться на камеры с датчиком глубины, для того чтобы эффективно строить 3d сцены вокруг человека. У таких камер по стандарту есть sdk для работы с различными видами отображений в 3d, таких как: point cloud, triangle mesh и другие. На схеме ниже представлен стек технологий и алгоритмов, которые нужны для построения 3d моделей (Рис. 1).

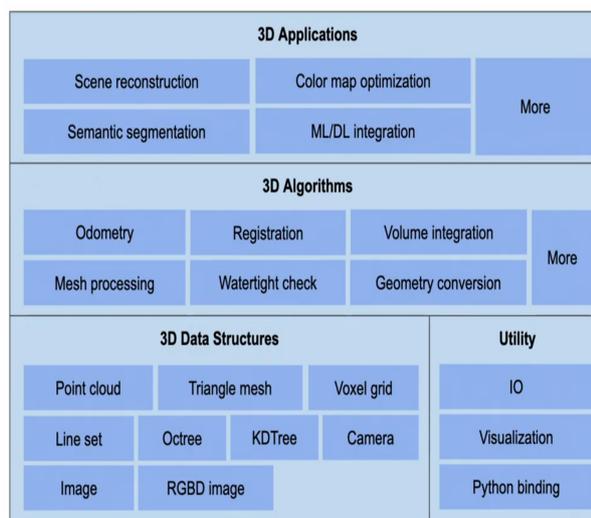


Рис. 1 – Стек технологий при построении 3d модели

Введем основные понятия блоков представленных на схеме:

Сегментация изображений (Semantic segmentation) — это одна из важнейших задач компьютерного зрения, которая представляет собой процесс разделения цифрового изображения на несколько сегментов. Цель сегментации заключается в упрощении и/или изменении представления изображения, чтобы его было проще и легче анализировать.

Реконструкция сцены (Scene reconstruction) — это процесс частичной реконструкции 3d пространства из RGB (RGBD) изображений.

Одометрия — это процесс соединения последовательности изображений в единое пространство.

Рассмотрим упрощенную последовательность действий для построения 3d модели человека. Изначально при помощи камеры получаем последовательность изображений. Следующим шагом является сегментация человека на изображениях при помощи моделей глубокого обучения. После чего зная, где находится человек на изображениях, мы можем удалить шумы с изображений. Далее при помощи библиотеки использующей одометрию мы можем объединить изображения в 3d модель человека (Рис 2).

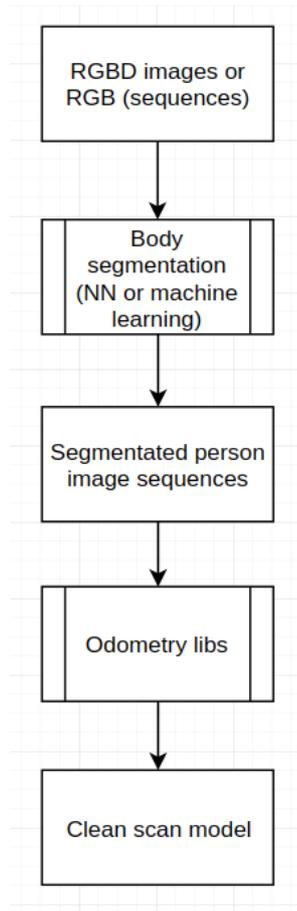


Рис. 2 – Упрощенная последовательность действий для построения 3d модели человека

В схеме были опущены некоторые пункты, которые зависят от камеры, такие как: соединение изображений с нескольких камер, так как одной будет недостаточно для детального сканирования и процесс работы с одометрией реализуется в многих библиотеках по разному при помощи многих видов преобразований.

### III. КАМЕРЫ ГЛУБИНЫ

Камеры глубины, которые снимают видео, в каждом пикселе которого хранится не только цвет, а расстояние до объекта в этой точке. Такие камеры существуют уже больше 20 лет, однако в последние годы скорость их развития выросла многократно и уже можно говорить про революцию.

Самыми известными решениями на рынке камер глубины являются:

- Microsoft Kinect;
- Intel RealSense;
- Sony DepthSensing;
- ZED Stereo Camera.

Технологий измерения глубины развиваются в следующих направлениях:

- Structured Light камеры, или камеры структурного света, когда есть проектор (часто инфракрасный) и камера, снимающая структурный свет проектора;
- Time of Flight камеры, или камеры, основанные на измерении задержки отраженного света; Depth from Stereo камеры — классическое и, пожалуй, наиболее известное направление построения глубины из стерео;
- Light Field Camera — они же камеры светового поля или пленоптические камеры;
- Камеры, основанные на Lidar-технологиях, особенно свежие Solid State Lidars, которые работают без отказа примерно в 100 раз дольше обычных лидаров и выдают привычную прямоугольную картинку.

### IV. ЗАКЛЮЧЕНИЕ

Таким образом, опираясь на пункты описанные выше можно полагать, что задачи построения качественных 3d моделей человека будут оставаться актуальными еще несколько лет, так как остаются нерешенными задачи:

- Построение моделей человеческого лица;
- Снятие параметров с 3d модели напрямую;
- Построение моделей без использования камер с дополнительными датчиками;
- Моделирование мелких деталей человеческого тела.

### СПИСОК ЛИТЕРАТУРЫ

1. Камеры глубины — тихая революция (когда роботы будут видеть) <https://habr.com/ru/post/457524/> Часть 1.
2. Камеры глубины — тихая революция (когда роботы будут видеть) <https://habr.com/ru/post/458458/> Часть 2.
3. Semantic Segmentation — <https://medium.com/nanonets/how-to-do-image-segmentation-using-deep-learning-c673cc5862ef>.