

УДК 004.93

РАСПОЗНАВАНИЕ ПОЗ ЧЕЛОВЕКА С ПРИМЕНЕНИЕМ ГЛУБИННЫХ КАМЕР В КОНТЕКСТЕ УПРАВЛЕНИЯ  
ОКРУЖЕНИЕМ ИНТЕРНЕТА ВЕЩЕЙ

Медведев Д.С. (Университет ИТМО), Шматков В.Н. (Университет ИТМО)

Научный руководитель – к. т. н., доцент Муромцев Д.И.  
(Университет ИТМО)

Исследование посвящено решению задачи распознавания позы человека применительно к управлению окружением Интернета Вещей. Распознавание производится с применением глубинной камеры и алгоритмов RF и CNN.

**Введение.** В настоящее время управление окружением Интернета Вещей становится всё более продвинутым и разнообразным. Постепенно внедряются такие методы управления как, например, голосовое управление или управление жестами. При имплементации последнего требуется понимание конфигурации позы человека и его жестов. Данная задача в общем виде носит название Human Pose Estimation (HPE) – определение позы человека. Цель – определение ориентацию тела в пространстве и построение 3d-скелета.

Задача может быть решена с применением комплекса аппаратных средств (например, с помощью глубинных, RGB-камер или носимых устройств). Существующие алгоритмы используемые с визуальными сенсорами (камерами) можно поделить на два класса: алгоритмы, основанные на Random Forest Classifier, либо алгоритмы, основанные на свёрточных нейронных сетях (CNN). RF-алгоритмы хорошо зарекомендовали себя на практике (использовался в Microsoft Kinect), но требуют больших объёмов данных для обучения, поэтому постепенно вытесняются CNN-алгоритмами.

**Основная часть.** В данной работе сделан упор на распознавание с применением глубинных камер и существующих алгоритмов на основе свёрточных нейронных сетей.

В качестве визуального сенсора была использована глубинная камера Intel Depth Camera D435i. В некоторых экспериментах использовались несколько таких камер. В качестве хоста были опробованы несколько платформ: персональный компьютер, а также несколько одноплатных компьютеров Raspberry Pi различных моделей.

Для обучения моделей были взяты синтетические и реальные данные из датасетов ITOP, UBC3V и PDT.

**Выводы.** Практическая имплементация распознавания поз открывает путь к реализации полноценного управления окружением Интернета Вещей с использованием жестов.

Распознавание поз человека с применением глубинных камер имеет высокую точность. Использование свёрточных нейронных сетей для анализа видеопотока на одноплатных компьютерах осложняется относительно невысокой производительностью последних. Несмотря на это, их производительности может быть достаточно для прикладных задач управления окружением Интернета Вещей.

Медведев Д.С. (автор)

Подпись

Муромцев Д.И. (научный руководитель)

Подпись