

УДК 004.93'11

## РАЗРАБОТКА МЕТОДА ОПРЕДЕЛЕНИЯ ГЛУБИНЫ ДЛЯ СИСТЕМ РЕАЛЬНОГО ВРЕМЕНИ С ИСПОЛЬЗОВАНИЕМ МОНОКУЛЯРНОГО ЗРЕНИЯ И СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ

Сорокин С. С. (Университет ИТМО)

Научный руководитель – к.т.н., доцент, Быковский С. В.

(Университет ИТМО)

Работа посвящена разработке метода определения карты глубины для систем реального времени с использованием монокулярного зрения. Для решения поставленной задачи первоначально был выполнен анализ существующих методов расчета карты глубины.

Проведя эксперименты на различных нейронных сетях, были установлены их характеристики, такие как точность выдаваемого результата, время обработки изображения, потребляемая память.

На основе полученных результатов, был выбран метод, удовлетворяющий поставленным требованиям.

**Введение.** Задача расчета карты глубины является одной из важнейших задач, решаемых в области компьютерного зрения.

Данная карта представляет собой изображение, которое в каждом своем пикселе, вместо информации о цвете, хранит информацию о расстоянии до камеры. На данный момент разработано множество способов получения такого изображения, каждый из которых имеет свои преимущества и недостатки.

Целью работы является разработка метода определения карты глубины для систем реального времени с использованием монокулярного зрения. Данный метод должен выдавать результат приемлемой точности, а именно соответствовать истинной глубине сцены (отличаться максимум на 15%) и верно передавать геометрию сцены. Также, метод разрабатывается с возможностью использования во встроенных системах, поэтому должен характеризоваться низкими требованиями к вычислительным мощностям.

**Основная часть.** В ходе исследования были рассмотрены методы, которые обобщив можно разделить на две группы: использующие одно изображение и использующие пару изображений одинаковой сцены из разных точек.

Методы, использующие одно изображение, могут построить карту глубины для любого изображения, однако приемлемая точность результата будет достигаться только в случае, если окружение сцены является знакомой для нейронной сети – соответствует окружению изображений, на которых она обучалась. В противном случае сеть не способна рассчитать глубину. Ее результат сильно отличается от ожидаемого, построенная геометрия не соответствует истинной.

Вторая группа методов не имеет такого ограничения, независимо от окружения, сеть выдает одинаково точный результат. Это объясняется тем, что имея два изображения, появляется возможность использовать эффект параллакса для расчета карты глубины. Эффект заключается в том, что при перемещении камеры, объекты находящиеся вдали имеют меньшее смещение, чем объекты находящиеся вблизи.

Возможно, на первый взгляд необходимость пары изображений может показаться серьезным ограничением. Однако, работая с видеопотоком, задача извлечения двух различных кадров не составляет проблем.

Исходя из вышесказанных преимуществ метода использующего пару изображений, дальнейшую работу было решено вести именно с ним.

Сорокин С.С. (автор)

Подпись

Быковский С.В. (научный руководитель)

Подпись

Данный метод можно оптимизировать для работы со встроенными системами. Это достигается при помощи TensorFlow Lite, которая предоставляет инструменты для изменения структуры хранения модели, типов данных, используемых в вычислениях, квантования. Эти инструменты позволят уменьшить требуемые ресурсы сети без значимой потери качества.

**Выводы.** Разработанный для встроенных систем метод может быть использован в таких системах, как системы навигации беспилотного транспорта, системы зрения роботов, приборы трехмерного сканирования, дополненной и смешанной реальности.