

РЕАЛИЗАЦИЯ СИСТЕМЫ ДОПОЛНЕННОЙ РЕАЛЬНОСТИ ДЛЯ МАКЕТА Г. КАЛУГИ

И.В. Винокуров, Д.О. Климов

КГУ им. К.Э. Циолковского, г. Калуга

Разработанная в КГУ им. К.Э. Циолковского система дополненной реальности позволяет осуществить интерактивное взаимодействие с 3-х мерными моделями, визуализирующими значимые объекты городской среды г. Калуги. Реализация такого взаимодействия заключается в сегментации изображения макета, распознавании фрагментов на сегментированном изображении, замене фрагментов изображения 3-х мерными моделями объектов и последующее интерактивное взаимодействие с ними – вращение, масштабирование и трансляция. Дополнительный неинтерактивный слой позволяет отобразить связанную с объектом информацию. Система дополненной реальности реализована как приложение для iPad и в настоящее время используется для более глубокого погружения в городскую среду с целью ознакомления гостей университета с муниципальными, культурными и историческими объектами г. Калуги.

Ключевые слова: дополненная реальность, свёрточная нейронная сеть.

ВВЕДЕНИЕ

Системы дополненной реальности (СДР) позволяют заменить реальные объекты их виртуальными аналогами с целью более глубокой интеграции с объективной реальностью [1]. Современные информационные технологии позволяют реализовать такие системы и на планшетах и смартфонах, что существенно расширяет области применения СДР и является стимулирующим фактором их развития за счёт массового использования персональных мобильных устройств.

Летом прошлого года в КГУ им. К.Э. Циолковского был создан макет центральной части г. Калуги, рис. 1.



Рис. 1. Макет г. Калуги в КГУ им. К.Э. Циолковского

Естественным этапом его развития явилась разработка СДР, которая позволяет любому гостю университета получить более подробную информацию о выбранном объекте и его архитектурных особенностях.

СДР является собственной разработкой КГУ им. К.Э. Циолковского, позволяющей осуществлять интерактивное взаимодействие с 3-х мерными моделями значимых объектов макета г. Калуги. СДР установлена на планшетах iPad, предлагаемых к использованию всем гостям университета при проведении некоторых мероприятий.

В процессе разработке СДР исследовалась применимость и эффективность использования известных методов решения следующих задач:

1. выявление объектов на изображении;
2. распознавание или классификация выявленных объектов;
3. формирование 3-х мерных моделей для значимых объектов городской среды;
4. замена выявленных на изображении объектов на их 3-х мерные модели (в текущей реализации СДР модели отображаются рядом с объектами);
5. реализация пространственных операций с 3-х мерными моделями.

Все перечисленные выше задачи были с той или иной мерой эффективности решены в процессе работы над СДР.

ОСНОВНАЯ ЧАСТЬ

Как было отмечено во введении, основными этапами работы СДР являются распознавание значимых архитектурных объектов на изображении макета, их контурное выделение и последующее взаимодействие с 3-х мерными моделями таких объектов. На текущем этапе развития программной системы возможны масштабирование 3-х мерных объектов до определённых размеров их трансляцию по всем осям и вращение вокруг вертикальной оси, имитирующее обход объекта по его периметру.

Главными задачами, стоящими перед коллективом разработчиков, были не разработка

новых методов распознавания, последующего выделения объектов на изображениях и их замена на 3-х мерные модели, а обоснованный выбор существующих и эффективная интеграция таких методов в единую СДР макета г. Калуги.

В результате работы над СДР была принята целесообразная следующая последовательность действий, рис. 2.



Рис. 2. Основные этапы работы СДР

Выявление и распознавание объектов на изображении макета г. Калуги было реализовано с использованием свёрточной нейронной сети (СНН). Эффективность и, как следствие, целесообразность использования СНН для реализации сегментации изображений показана в [2,3]. Сегментами, предлагаемыми к выделению СНН, являются области изображения, предположительно содержащие наиболее интересные для гостей города объекты – административные, культурные и исторически значимые. Архитектура СНН, использованной в СДР, во многом аналогична архитектуре U-Net [2,3] – она так же предполагает наличие сужающего и расширяющего изображение путей, реализующих применение свёрточных фильтров и восстановление исходного сегментированного изображения.

Как показали экспериментальные исследования, для приемлемого качества сегментации изображения макета вполне достаточно преобразовать его цветное изображение в черно-белое и выбрать фильтры для детектирования объектов по их основным характерным признакам – наличие колонн, куполов, наличия фасадов, особенностям архитектурного стиля и т.п. Все такие признаки могут быть выявлены следующими стандартными фильтрами для свёрточных слоёв нейронной сети [5,6]:

1. фильтр повышения резкости, позволяющий выявить неровности и общие характеристики контура объекта;

2. контурный фильтр, выделяющий основные характеристики контура объекта;

3. фильтр детектирования вертикальных границ и переходов;

4. фильтр детектирования горизонтальных границ и переходов;

5. фильтр размытия Гаусса – крайне редко, но может заменять собой один из перечисленных выше фильтров, за исключением контурного фильтра с целью повышения точности определения объекта.

Приемлемые для решения большинства задач размеры ядер всех указанных выше фильтров – 3x3 [5] определили размеры ядер свёрточных слоёв СНН. Использование того или иного фильтра вариативно, и определяется сложностью распознавания объекта макета [6].

С целью уменьшения размерности пространства признаков и, как следствие, вычислительной нагрузки на мобильное устройство, после каждого из свёрточных слоёв реализован слой усредняющего пуллинга с размером ядра 2x2. Общая структура СНН, используемая для сегментации изображения макета, приведена на рис. 3.

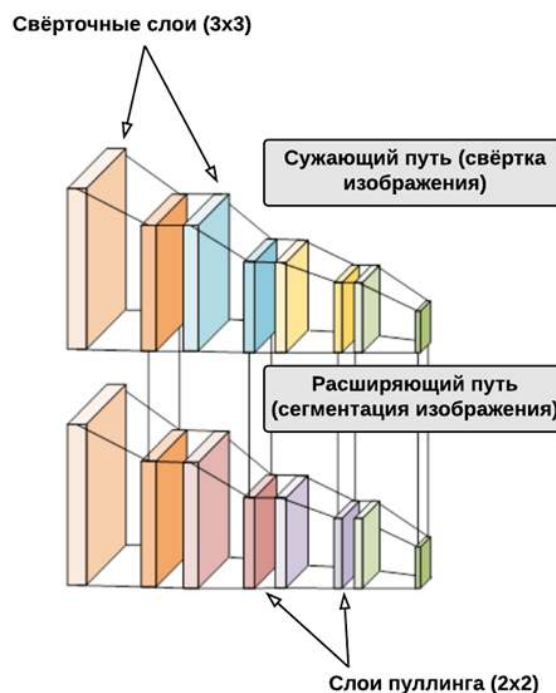


Рис. 3. U-образная структура СНН для сегментирования изображения макета

Как и собственно U-Net, СНН для сегментации изображения макета обучается методом стохастического градиентного спуска на основе входных изображений макета и соответствующих им карт сегментации [4,7].

На рис. 4 и 5 приведены фрагмент исходного изображения макета и результаты его сегментации с использованием СНН.



Рис. 4. Фрагмент изображения макета со Свято-Троицким собором Калужской епархии

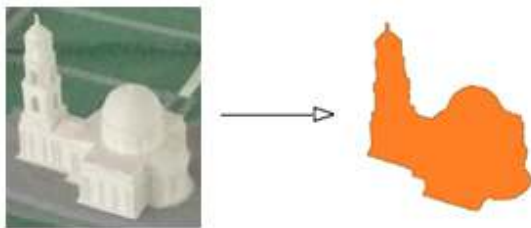


Рис. 5. Реализуемая СНН сегментация фрагмента изображения

Контуры всех сформированных СНН сегментов приводятся к единому размеру, определяя тем самым размер скользящего окна [8]. Скользящее окно было выбрано по причине развитых фреймворков машинного обучения iOS, использование которых позволяет получить приемлемую точность распознавания (классификации) объектов без применения полносвязных нейронных сетей (НС). НС такого типа предполагают наличие эвристических этапов формирования обучающих и проверочных наборов данных и этапа обучения [9]. Кроме этого, при достаточно длительном времени обучения полносвязных НС точность распознавания может быть неприемлемой.

На начальных этапах проектирования и разработки СДР предполагалось использование полносвязной НС, обученной по методу прямого распространения ошибки. Однако в результате работы над СДР было принято решение отказаться от её использования в пользу скользящего по изображению окна. Размер скользящего окна в СДР макета г. Калуги равен 212×328 пикселей. Такой размер объясняется усреднёнными пропорциями объектов на изображении и позволяет обеспечить приемлемую точность их распознавания.

Выделенные СНН контуры всех объектов определяют интерактивные области изображения макета г. Калуги, позволяющие визуализировать их 3-х мерные модели и реализовать с ними пространственное взаимодействие, рис. 6.

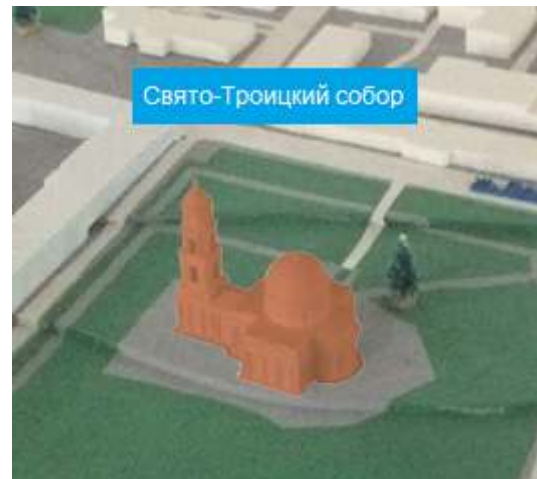


Рис. 6. Визуализация интерактивного фрагмента изображения

Выбор (тап) интерактивной области приводит к отображению рядом с ним его 3-х мерной модели, рис. 7.

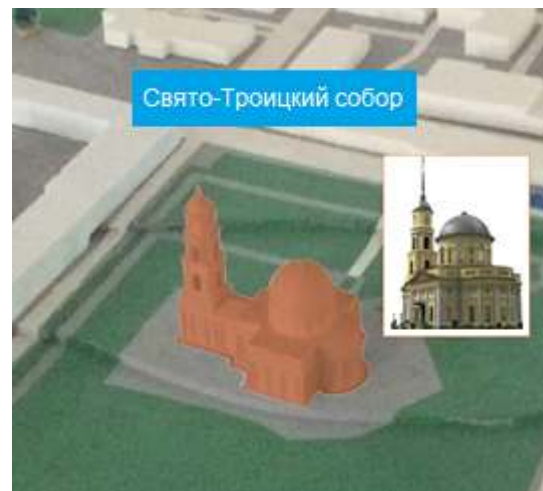


Рис. 7. Отображение 3-х мерной модели Свято-Троицкого собора

Модели всех значимых объектов городской среды были выполнены в 3ds MAX 2019. Координаты их вершин и текстур сохранялись в отдельных файлах, из которых 3-х мерные модели средствами фреймворка Metal [10] визуализировались рядом с выбранным на изображении макета объектом. Размер окна с моделью в текущей реализации СДР фиксирован и составляет 836×1004 пикселя. В последующих версиях СДР, с целью более детальной визуализации 3-х мерных моделей, предполагается реализовать их полноэкранное отображение в отдельном контроллере.

Как было отмечено выше, окно с 3-х мерной моделью объекта позволяет реализовать основные пространственные операции – вращение, масштабирование и трансляция. Все эти операции выполняются в результате использования соответствующих им жестов на экране мобильного устройства, Примеры реализации таких операций приведены на рис. 8 и 9.



Рис. 8. Вращение 3-х мерной модели Свято-Троицкого собора

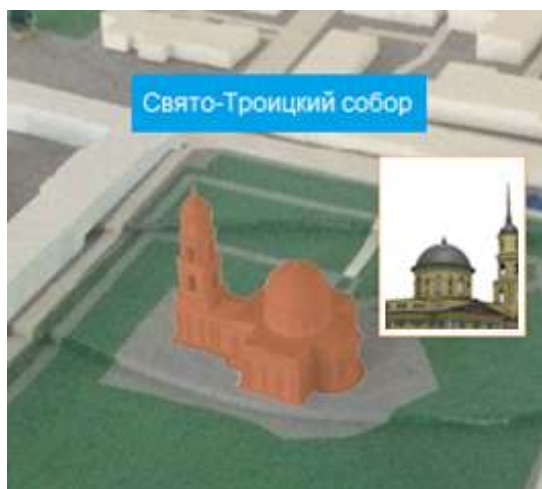


Рис. 9. Масштабирование 3-х мерной модели Свято-Троицкого собора

Для всех распознанных объектов на изображении макета отображается его название, см. рис. 7-9. Названия всех объектов образуют неинтерактивный информационный слой.

ЗАКЛЮЧЕНИЕ

Разработанная в КГУ им. К.Э. Циолковского СДР позволяет выявить на изображении макета значимые объекты городской среды и отобразить рядом с ними их 3-х мерные модели. Для всех моделей могут быть реализованы пространственные операции вращения, масштабирования и трансляции.

Использование СДР заключается в наведении планшета на макет г. Калуги и выделении тапом распознанных СНН объектов. Простота использования СДР является её главным достоинством. На текущем этапе развития СДР существуют и определённые недостатки:

1. ограниченность диапазона расстояния от пользователя СДР до макета 1-5 м. – при меньшем или большем расстояниях возможны ложные сегментации изображения макета;
2. сравнительно небольшой размер скользящего окна, приводящий в ряде случаев к заметным задержкам выделения распознанных объектов;
3. отображение 3-х мерной модели объекта в небольшом окне, приводящем к сравнительно малой детализации его отображения.

Все указанные недостатки будут устраняться в процессе дальнейшей работы над СДР, которая, как предполагается, помимо работы на планшетах iPad будет работать и на смартфонах iPhone с большими экранами. Основные программные составляющие СДР реализованы на современном языке программирования Swift в среде разработки Apple Xcode 10.2 с использованием высокоуровневых фреймворков векторной графики Metal [10], машинного обучения CoreML [11], Vision и реализации дополненной реальности ARKit 3 [12].

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Дополненная реальность. [Электронный ресурс]. Режим доступа: <https://www.apple.com/ru/ios/augmented-reality>
2. Convolutional Neural Networks [Электронный ресурс]. Режим доступа: <https://www.coursera.org/learn/convolutional-neural-networks>
3. U-Net: нейронная сеть для сегментации изображений. [Электронный ресурс]. Режим доступа: <https://neurohive.io/ru/vidy-nejrosetej/u-net-image-segmentation>
4. Neural networks and Deep Learning [Электронный ресурс]. Режим доступа: <http://neuralnetworksanddeeplearning.com>
5. Матричные фильтры обработки изображений. [Электронный ресурс]. Режим доступа: <https://unick-soft.ru/article.php?id=39>
6. Viola P., and M. Jones, Rapid object detection using a boosted cascade of simple features [Text] // Computer Vision and Pattern Recognition. 2001. – vol. 1. – pp. 511-518.
7. Deep Learning for Image Segmentation: U-Net Architecture. [Электронный ресурс]. Режим доступа: <https://heartbeat.fritz.ai/deep-learning-for-image-segmentation-u-net-architecture-ff17f6e4c1cf>
8. Object Category Detection. [Электронный ресурс]. Режим доступа: <http://www.robots.ox.ac.uk/~az/lectures/aims-cv/detection-part1.pdf>
9. Основы ИНС. [Электронный ресурс]. Режим доступа: <https://neuralnet.info/chapter/основы-инс>
10. Metal. Accelerating graphics and much more. [Электронный ресурс]. Режим доступа: <https://developer.apple.com/metal>
11. Apple Machine Learning Journal [Электронный ресурс]. Режим доступа: <https://machinelearning.apple.com>
12. Get Ready for ARKit 3. [Электронный ресурс]. Режим доступа: <https://developer.apple.com/augmented-reality/arkit>

Винокуров Игорь Викторович – к.т.н., доцент кафедры «Информатика и информационные технологии» КГУ им. К.Э. Циолковского, тел. 8 (4842) 22-04-06, e-mail: VinokurovIV@tksu.ru.

Климов Дмитрий Олегович – студент кафедры «Информатика и информационные технологии» КГУ им. К.Э. Циолковского, тел. 8 (4842) 22-04-06, e-mail: KlimovDO@studkg.ru.

REALIZATION OF THE AUGMENTED REALITY SYSTEM FOR THE LAYOUT OF KALUGA

I.V. Vinokurov, D.O. Klimov

Kaluga State University named after K.E. Tsiolkovski, Kaluga

Developed at KSU named after K.E. Tsiolkovski's augmented reality system allows for interactive interaction with 3-dimensional models that visualize significant objects of the urban environment of the city of Kaluga. The implementation of this interaction consists in segmenting the image of the layout, recognizing fragments in the segmented image, replacing image fragments with 3D models of objects, and subsequent interactive interaction with them – rotation, scaling, and translation. An additional non-interactive layer allows you to display information related to the object. The augmented reality system is implemented as an application for the iPad and is currently used for a deeper immersion in the urban environment in order to familiarize university guests with the municipal, cultural and historical sites of the city of Kaluga.

Index terms: augmented reality, convolutional neural network.

REFERENCES

1. Augmented reality. The access mode: <https://www.apple.com/ru/ios/augmented-reality>
2. Convolutional Neural Networks. The access mode: <https://www.coursera.org/learn/convolutional-neural-networks>
3. U-Net: neural network for image segmentation. The access mode: <https://neurohive.io/ru/vidy-nejrosetej/u-net-image-segmentation>
4. Neural networks and Deep Learning. The access mode: <http://neuralnetworksanddeeplearning.com>
5. Image processing matrix filters. The access mode: Режим доступа: <https://unick-soft.ru/article.php?id=39>
6. Viola P., and M. Jones, Rapid object detection using a boosted cascade of simple features [Text] // Computer Vision and Pattern Recognition. 2001. – vol. 1. – pp. 511-518.
7. Deep Learning for Image Segmentation: U-Net Architecture. The access mode: <https://heartbeat.fritz.ai/deep-learning-for-image-segmentation-u-net-architecture-ff17f6e4c1cf>
8. Object Category Detection. The access mode: <http://www.robots.ox.ac.uk/~az/lectures/aims-cv/detection-part1.pdf>
9. Neural Net basics. The access mode: <https://neuralnet.info/chapter/основы-инс>
10. Metal. Accelerating graphics and much more. The access mode: <https://developer.apple.com/metal>
11. Apple Machine Learning Journal The access mode: <https://machinelearning.apple.com>
12. Get Reay for ARKit 3. The access mode: <https://developer.apple.com/augmented-reality/arkit>

Vinokurov Igor Viktorovich – PhD. (IT), Associate Professor, “Informatics and Information Technologies”, Kaluga State University named after K.E. Tsiolkovski, 8 (4842) 22-04-06, e-mail: VinokurovIV@tksu.ru

Klimov Dmitry Olegovich – student of the department “Informatics and Information Technologies”, Kaluga State University named after K.E. Tsiolkovski, 8 (4842) 22-04-06, e-mail: KlimovDO@studklg.ru