**РУКОВОДСТВО ПОЛЬЗОВАТЕЛЯ**

«Подсистема построения карты глубины изображений для системы стереозрения»

Авторство: Батракова Елена Сергеевна, Батонова Оксана Юрьевна

Содержание

[Аннотация 3](#_Toc132222523)

[Введение 4](#_Toc132222524)

[Инструкция 6](#_Toc132222525)

[Предметный указатель 15](#_Toc132222526)

Аннотация

Данное руководство содержит подробную инструкцию по использованию авторского алгоритма для автоматической генерации баз данных в среде Blender. Данный алгоритм обеспечивает быстрое и качественное решение главной проблемы при работе с глубоким обучением, а именно качественная база и реалистичные данные, приближенные к реальным измерениям. Это станет хорошей отправной точкой для работы с картами глубины в исследованиях смежных тематик, а также может послужить отличным обучающим ресурсом для студентов, которые хотели бы погрузиться в сферу стереозрения и моделирования. Качественный датасет – это также всегда экономия времени для разработчика, а значит задействование меньшего количества ресурсов для тех же объемов работы, что выполнялась ранее.

Также стоит отметить, что результатами могут воспользоваться любые исследователи, увлеченные темой, поскольку код будет использоваться в свободном доступе. Следовательно, он может быть использован в качестве основы для дальнейшей работы и дать идею для того, чтобы модифицировать код для других нужд.

Введение

В последнее время использование синтетических данных при обучении моделей значительно возросло в связи с их повышенной эффективностью и скоростью. Архитектуры глубокого обучения требуют большого объема данных для обучения, и несмотря на то, что крупномасштабный сбор визуальных данных относительно прост в реализации, аннотирование собранных данных требует большого количества времени. Хотя при использовании предварительно обученных моделей требуется только небольшой набор данных для тонкой настройки, маркировка всего нескольких сотен или нескольких тысяч изображений для сегментации объектов или экземпляров по-прежнему является трудоемкой задачей, если она выполняется вручную.

Далее будет обоснован выбор определенного программного обеспечения. В качестве основного инструмента был выбран Blender, поскольку это популярное программное обеспечение с открытым исходным кодом, которое может предоставить ряд преимуществ для разработки. Некоторые из них:

а) Создание синтетических изображений: Blender может создавать 3D-модели объектов и сцен и генерировать синтетические изображения с широким диапазоном ракурсов камеры и форм объектов, что позволяет обучать подсистему.

б) Реалистичное моделирование освещения и материалов: Моделирование различных условий освещения, свойств материала и факторов окружающей среды помогает в создании реалистичных изображений.

в) Экономичность: это экономически выгодно по сравнению с использованием изображений реального мира, для чего может потребоваться дорогостоящее оборудование, логистика и рабочая сила. Используя Blender, исследовательская группа экономит ресурсы и распределяет их на другие важные аспекты проекта.

г) Эффективный рабочий процесс: Удобный интерфейс и полный набор инструментов упрощают рабочий процесс, в частности, подключаемый механизм для расширения встроенных функциональных возможностей.

Python API для Blender позволяет разработчикам получать доступ к функциональности Blender и управлять ею с помощью языка программирования Python, а исследователям предоставляет возможность автоматизировать задачи, настраивать рабочие процессы и расширять функциональность Blender путем создания пользовательских скриптов и дополнений.

Пакет состоит из набора модулей и функций, которые обеспечивают доступ к внутренним структурам данных Blender, таким как объекты, материалы, текстуры и анимации, для создания, изменения и удаления объектов, а также взаимодействия с его интерфейсом. Например, сценарий может быть написан для случайно размещенных объектов в сцене и изменять их размер, ориентацию и свойства материала. Можно написать другой скрипт для автоматического создания точек обзора камеры с заданного диапазона углов и расстояний.

Python API позволяет автоматизировать повторяющиеся задачи, такие как настройка сцен, импорт и экспорт файлов и генерация синтетических изображений. Такая автоматизация экономит время и снижает вероятность ошибок. В контексте разработки API используется для автоматизации генерации обучающих данных и экспорта файлов.

Инструкция

В данном разделе приведен пошаговый план эксплуатации разработанной подсистемы построения карты глубины изображений для системы стереозрения.

1. Откройте Blender на вашем компьютере. Если ранее вы не сталкивались с этим ПО, его необходимо скачать и установить на вашем устройстве. Используйте официальный сайт <https://www.blender.org/download/> (рис.1). Blender является бесплатным ПО с открытым программным кодом, что делает его доступным для всех. В данном руководстве рассмотрена новейшая на сегодняшний день версия Blender 3.5.

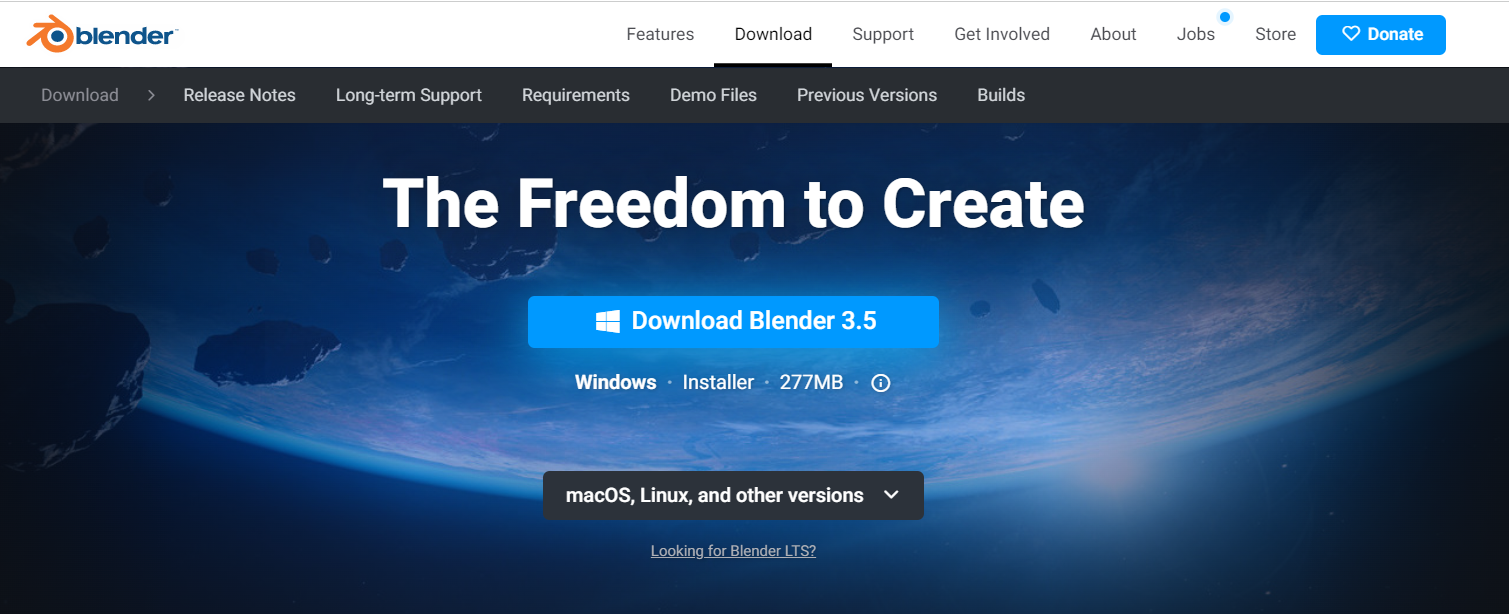


Рисунок 1 – Загрузка Blender

1. После установки и успешного запуска Blender обратимся к настройкам. В панели управления, в левом верхнем углу, выберите «Edit», в открывшемся списке выберите пункт «Preferences» (рис. 2).

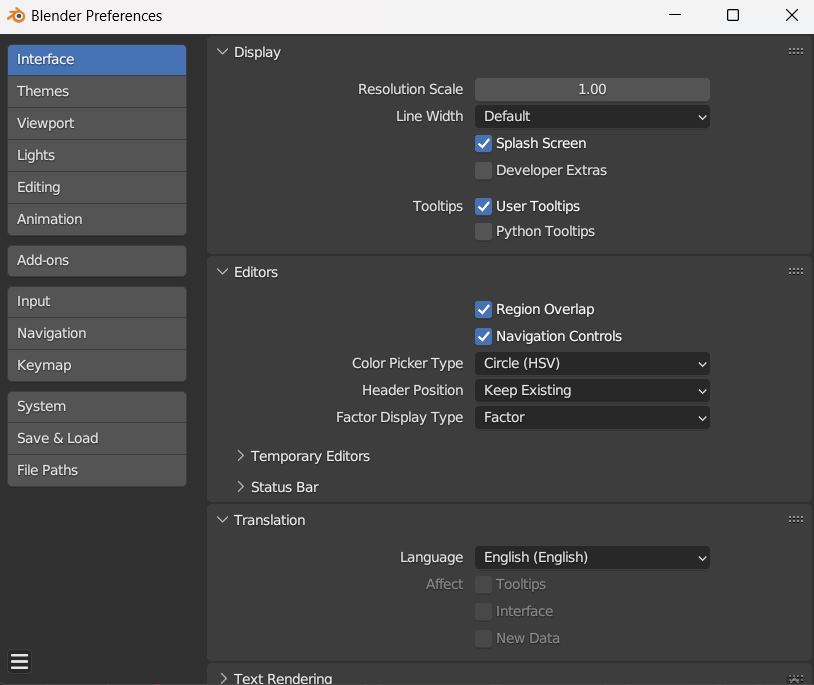


Рисунок 2 – Всплывающее окно «Preferences»

1. В открывшемся окне из списка слева выберите раздел «Safe & load». Найдите пункт «Auto Run Python Scripts» и активируйте галочку с левой стороны (рис. 3).

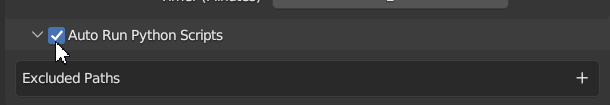


Рисунок 3 – Настройка автоматического запуска скриптов

1. Далее перейдем к установке архива. Для этого, если необходимо, повторим ранее проделанный путь: «Edit» >>> «Preferences». Далее кликните на раздел «Add-ons», а затем на кнопку «Install» в правом верхнем углу. Вы увидите раздел «Документы» и список файлов, хранящихся на вашем устройстве. Среди них найдем нужный zip-архив. Наберите в поиске: «stereovision\_dataset.zip» (рис. 4), после чего кликните кнопку синего цвета «Install add-on» в правом нижнем углу.

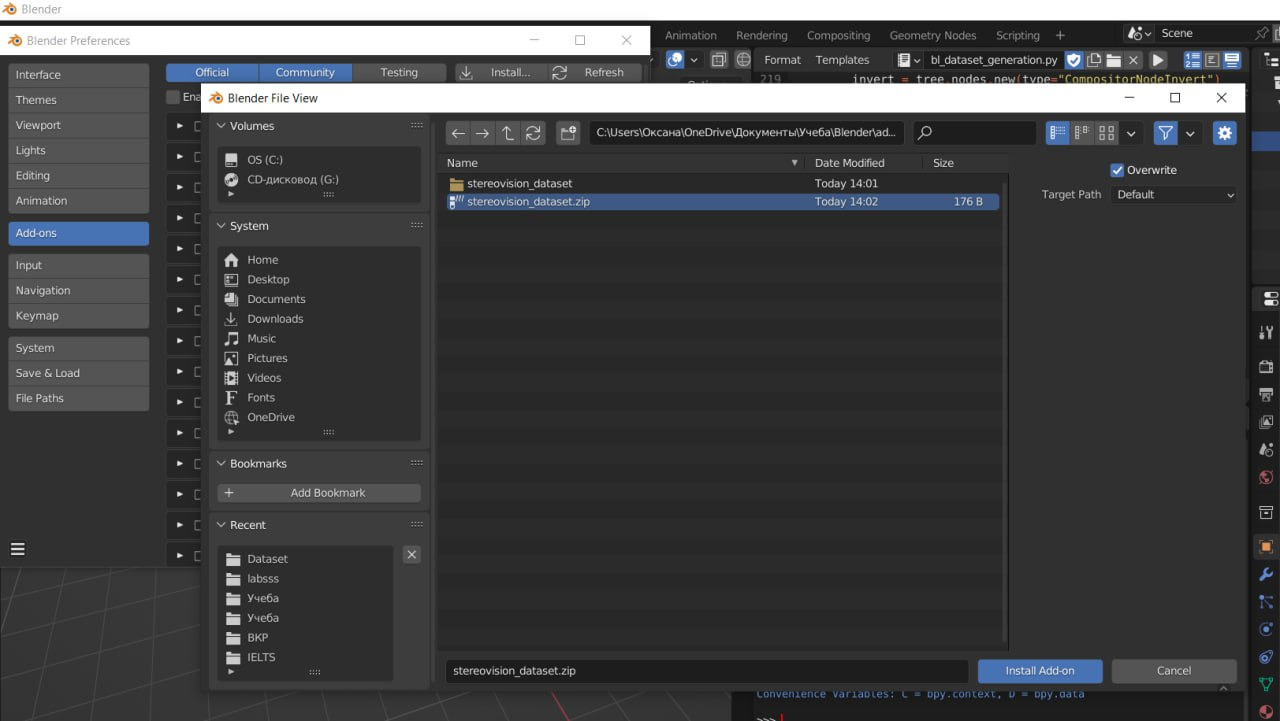


Рисунок 4 – Загрузка скрипта в Blender

1. После загрузки скрипта в Blender вы сможете увидеть его в списке всех аддонов (рис. 5), если вы не видите данный аддон, то закройте Blender, а затем запустите приложение снова.

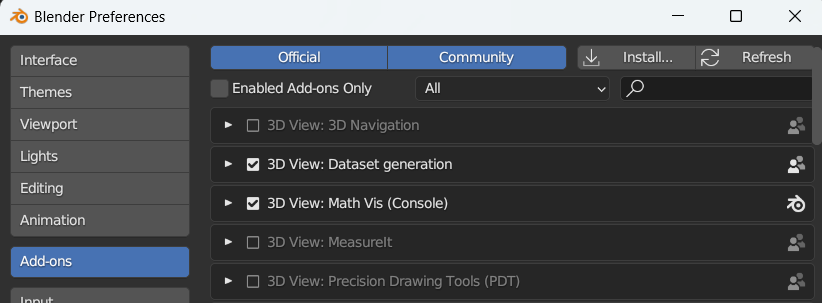


Рисунок 5 – Список аддонов в Blender

1. Чтобы воспользоваться загруженным аддоном, воспользуйтесь стрелкой в правом верхнем углу рабочего пространства (рис. 6).

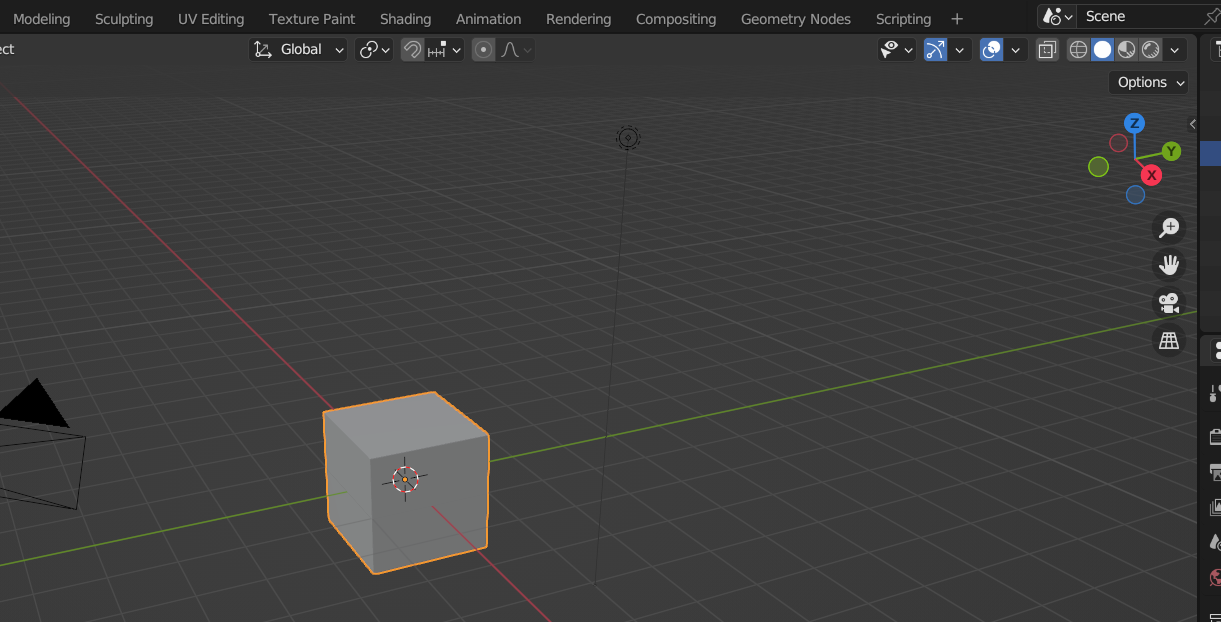


Рисунок 6 – Рабочее пространство Blender

1. В появившемся окне выберите вкладку «Datasets» (рис. 7).

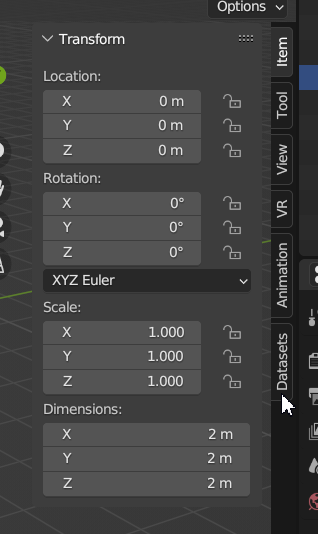


Рисунок 7 – Окно для работы с аддонами

1. Приступим непосредственно к созданию датасета. Для этого заполните все поля в появившемся окне «Dataset Generation». Если какие-то из полей не будут заполнены, кнопка «Generate» останется некликабельной и надпись на ней будет окрашена в тусклый серый цвет (рис. 8).

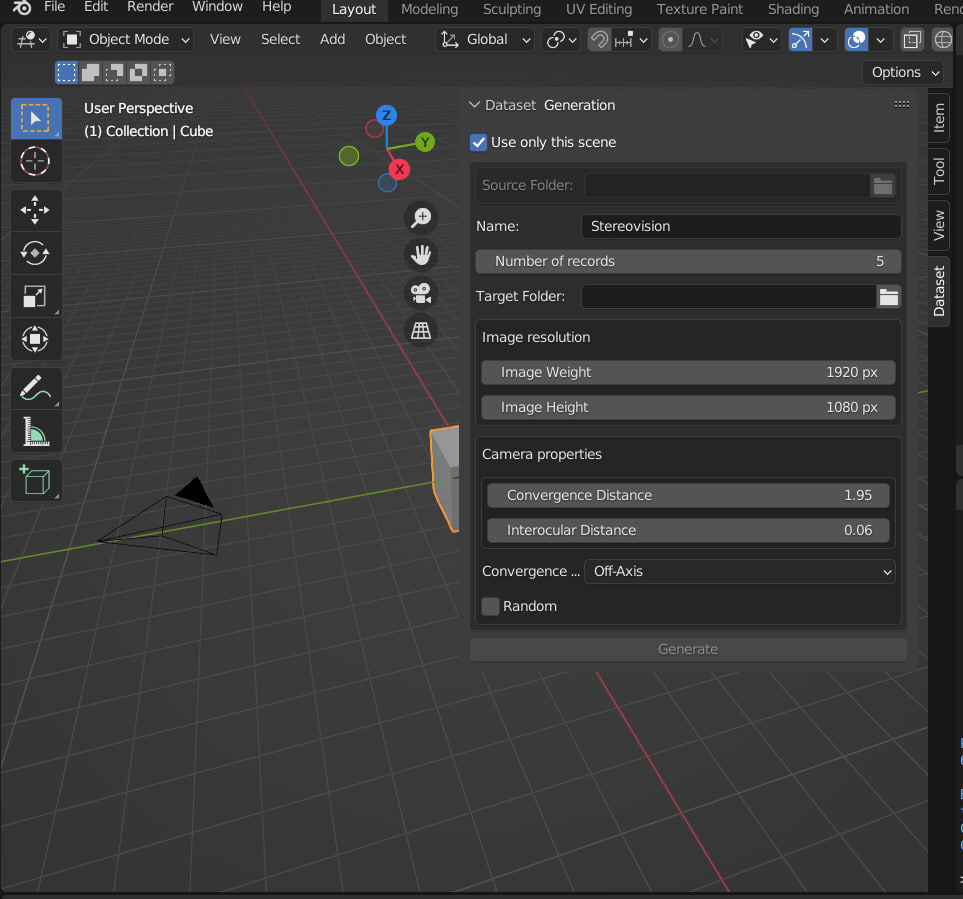


Рисунок 8 – Заполнение параметров в окне «Dataset Generation»

1. Рассмотрим каждый параметр подробнее. Если кликнуть и поставить галочку напротив пункта «Use only this scene» (рис. 9), потребуется заполнить только два поля: «Name» (до 30 символов) и «Target Folder», а именно имя датасета и указать папку, куда датасет будет сохранен. Можно заметить, что после заполнения этих полей кнопка «Generate» стала активной, а значит датасет можно генерировать.

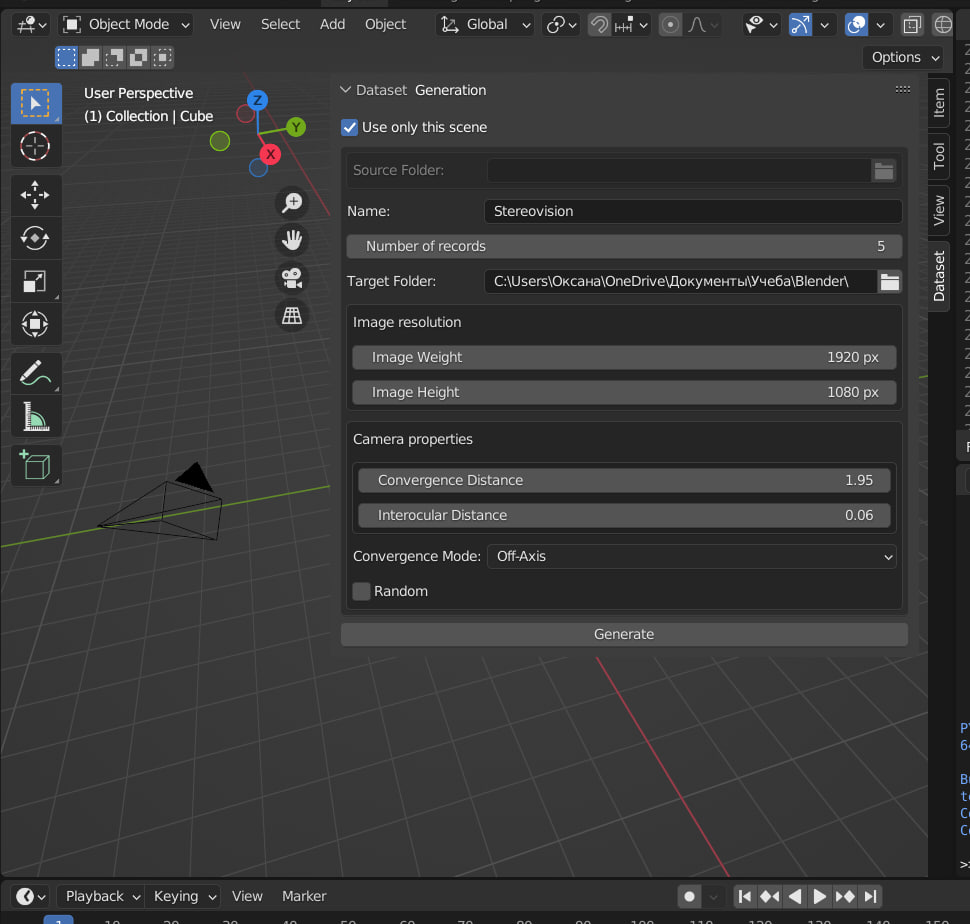


Рисунок 9 – Заполнение полей формы

1. Снятие галочки с пункта «Use only this scene» подразумевает, что будут использованы дополнительные сцены, а значит, они должны быть загружены. Сделать это можно, указав путь к файлам в графе «Source Folder» (рис. 10). Обратите внимание, что расширение файлов, принимаемых к загрузке, записывается как .blend.

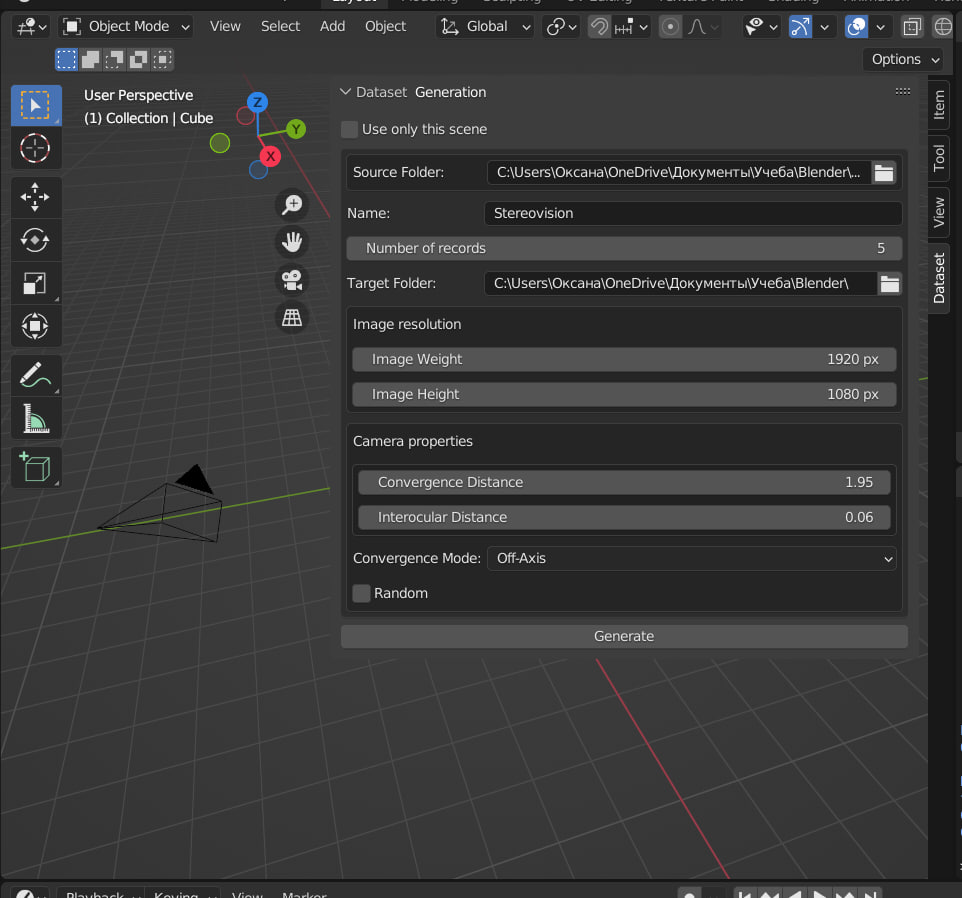


Рисунок 10 – Выбор дополнительных сцен

1. В нижней части окна вы можете видеть пункт «Random» (рис. 11). Он отвечает за настройки камер системы стереозрения, перечисленные в разделе «Camera properties». Если поставить галочку – значения устанавливаются рандомно, а если ее убрать, параметры можно настроить вручную.

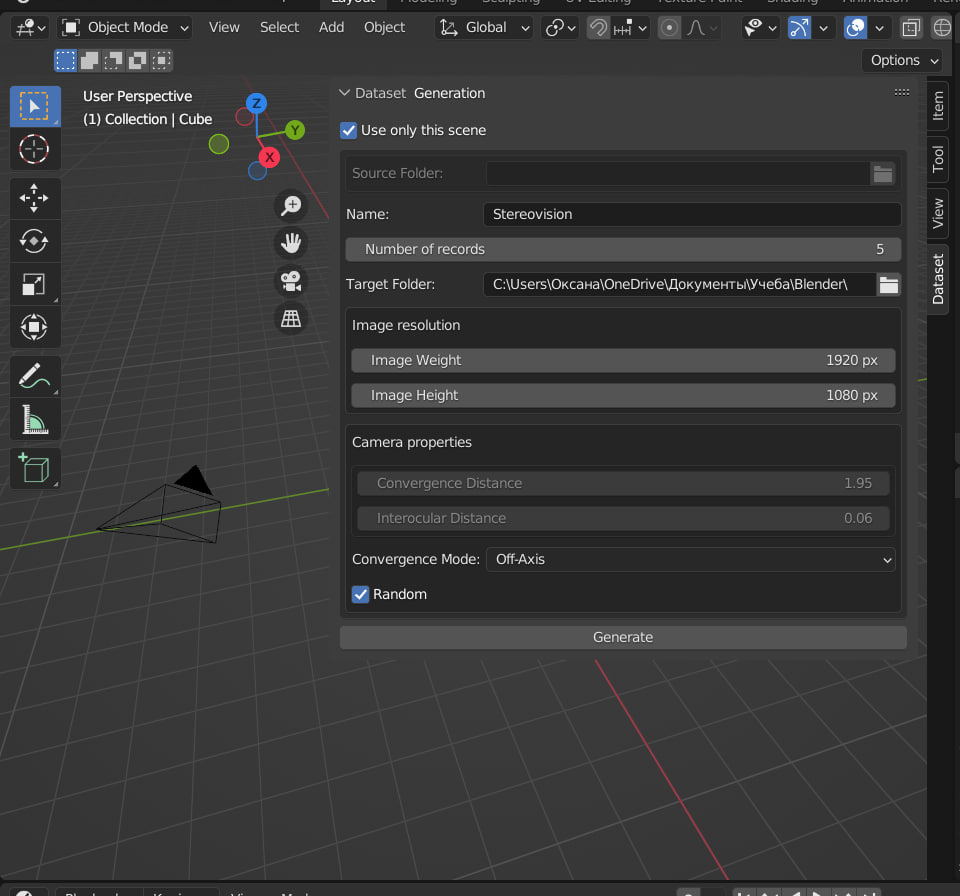


Рисунок 11 – Рандомная установка значений параметров камер

1. Обратите внимание на предустановленные параметры в разделах «Image resolution» и «Camera properties», отвечающих за размер изображений и настройку камер соответственно. Для «Image Weight» это 1920 px, для «Image Height» это 1080 px. Для «Convergence Distance» это 1.95, для «Intercocular Distance» это 1.06. Также в разделе «Number of records», отвечающем за количество записей в датасете, первоначально указано число 5.
2. У каждого параметра есть свои максимальные и минимальные значения, которые могут быть указаны. Число записей от 1 до 10.000. Разрешение изображения до 50.000 пикселей. Для двух дистанций минимум составляет и 0 соответственно. Ввести параметры можно после двойного клика по необходимой строке.
3. После генерации датасета заглянем в папку, чтобы оценить результат (рис. 12). В папке «images» находятся фотографии каждой сцены с двух ракурсов. В «depthmap» находятся карты глубины. В «Readme.md» содержится инструкция по расшифровке результатов работы аддона. Ознакомьтесь с ним, прежде чем переходить к следующему файлу. Файл с расширением .txt имеет название, указанное вами ранее в графе «Name», здесь это «Stereovision.txt». Он содержит текстовые описания полученных карт глубины с параметрами, координатами и именами объектов на изображениях.

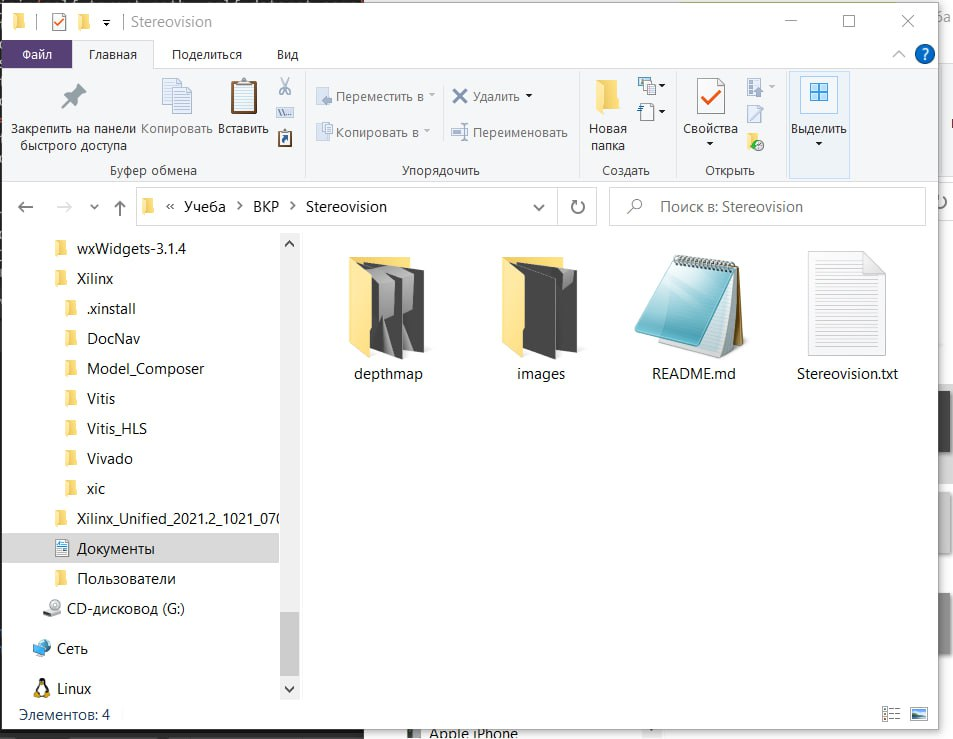


Рисунок 12 – Полученный датасет, выходные файлы

Глоссарий

1. **Blender** – профессиональное свободное и открытое программное обеспечение для создания трёхмерной компьютерной графики, включающее в себя средства моделирования, скульптинга, анимации, симуляции, рендеринга, постобработки и монтажа видео со звуком, компоновки с помощью «узлов», а также создания 2D-анимаций.
2. **Python** – высокоуровневый язык программирования общего назначения с динамической строгой типизацией и автоматическим управлением памятью, ориентированный на повышение производительности разработчика, читаемости кода и его качества, а также на обеспечение переносимости написанных на нём программ.
3. **Python API** (application programming interface) – программный интерфейс приложения.
4. **База данных** – совокупность данных, хранимых в соответствии со схемой данных, манипулирование которыми выполняют в соответствии с правилами средств моделирования данных.
5. **Глубокое обучение** – совокупность методов машинного обучения, основанных на обучении представлениям, а не специализированных алгоритмах под конкретные задачи.
6. **Датасет** – это коллекция данных. В случае с табличными данными набор данных соответствует одной или нескольким таблицам баз данных, где каждый столбец таблицы соответствует отдельной переменной, и каждая строка соответствует записи в наборе данных.
7. **Карта глубины** – это изображение или канал изображения, который содержит информацию, касающуюся расстояния поверхностей объектов сцены от точки обзора.
8. **Программное обеспечение** (ПО) – программа или множество программ, используемых для управления компьютером.
9. **Синтетические данные** – это информация, созданная искусственно, а не в результате реальных событий. Синтетические данные, обычно создаваемые с использованием алгоритмов, можно использовать для проверки математических моделей и обучения моделей машинного обучения.
10. **Стереозрение** (стереоскопическое зрение) – вид [зрения](https://ru.wikipedia.org/wiki/%D0%97%D1%80%D0%B5%D0%BD%D0%B8%D0%B5), при котором возможно [восприятие](https://ru.wikipedia.org/wiki/%D0%92%D0%BE%D1%81%D0%BF%D1%80%D0%B8%D1%8F%D1%82%D0%B8%D0%B5) формы, размеров и расстояния до предмета, например благодаря [бинокулярному зрению](https://ru.wikipedia.org/wiki/%D0%91%D0%B8%D0%BD%D0%BE%D0%BA%D1%83%D0%BB%D1%8F%D1%80%D0%BD%D0%BE%D0%B5_%D0%B7%D1%80%D0%B5%D0%BD%D0%B8%D0%B5) (количество глаз может быть и больше 2-х, как, например, у ос — два сложных глаза и три простых глаза (глазка), скорпионов — 3-6 пар глаз) или другим типам зрения.

Предметный указатель

Blender, 3, 4, 5, 6

Python, 5

Python API, 5

база данных, 3, 4

глубокое обучение, 3, 4

датасет, 3, 8, 9, 12

карта глубины, 3, 6, 12

программное обеспечение, 4, 6

синтетические данные, 4

стереозрение, 3, 6, 10