Создание системы динамического взаимодействия с объектами дополненной реальности

Калинин Даниил Евгеньевич, Радькин Кирилл Алексеевич

МАОУ «Лицей №97 г. Челябинска», 10м1 класс

Научный руководитель:

Саканов Дамир Муратович

педагог дополнительного образования

МАОУ «Лицей №97 г. Челябинска»

Работа к защите допущена \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ // Саканов Дамир Муратович

Оглавление

[Введение 3](#_Toc7890836)

[Глава 1. Изучение литературы 5](#_Toc7890837)

[1.1. Разработка в Processing 5](#_Toc7890838)

[1.2. Разработка в Unity-3D 8](#_Toc7890839)

[1.3. Изучение способов отслеживания перемещения 9](#_Toc7890840)

[1.3.1. Акустические методы 10](#_Toc7890841)

[1.3.2. Радиочастотные методы 10](#_Toc7890842)

[1.3.3. Магнитные методы 11](#_Toc7890843)

[1.3.4. Оптические методы 11](#_Toc7890844)

[1.3.5. Инерциальный трекинг 12](#_Toc7890845)

[Глава 2. Разработка принципа работы модели 14](#_Toc7890846)

[2.1. Выбор оборудования 14](#_Toc7890847)

[2.1.1. Микроконтроллер Arduino 14](#_Toc7890848)

[2.1.2. Модуль трекинга 15](#_Toc7890849)

[Глава 3. Написание программного обеспечения 17](#_Toc7890850)

[3.1. Программирование на Unity и рендеринг модели на экране 17](#_Toc7890851)

[3.1.1. Динамическое позиционирование модели на изображении-метке 17](#_Toc7890852)

[3.1.2. Перемещение объекта посредством управления виртуальным джойстиком. 19](#_Toc7890853)

# Введение

Прогресс человечества растет с невероятной скоростью. Прямо сейчас по всему миру происходит столько же научных открытий, сколько произошло за весь 15 век. Ежеминутно создаются новые системы, призванные упростить повседневную жизнь человека. В качестве примера, приведем несколько таких систем. IOT[[1]](#footnote-1)-устройства, расположенные как в доме, так и в промышленном производстве, могут послужить хорошим примером, поскольку получили широкое распространение относительно недавно. Такие устройства позволяют осуществлять автоматический сбор и анализ данных, например, IOT-холодильник способен автоматически заказывать в интернет-магазине, а IOT-склад способен автоматически вести учет товара. Еще одним примером может стать технология компьютерного зрения и машинного обучения в целом. Данная система хоть и была разработана довольно давно, в жизнь обычного человека она вошла так же недавно, как и предыдущая. Данная технология используется для анализа больших баз данных и выведения закономерностей. Подборка фильмов, из учета ваших интересов, распознавание номера автомобильного знака по фото, сделанным камерой контроля скорости, распознавание лица для разблокировки смартфона – это алгоритмы машинного обучения. И последняя в списке примеров, но далеко не последняя по значимости технология – технология AR[[2]](#footnote-2)– дополненной реальности. Технология представляет дополнение нашей с вами реальности (отсюда и название) различными объектами из виртуального мира. От технологий, приведенных выше, эта имеет одно отличие: эта технология почти не используется в нашей повседневной жизни. Эта особенность связана, в первую, по нашему мнению, очередь с тем, что не создано по-настоящему удобной и универсальной системы взаимодействия с AR, а без такой системы, область применения данной технологии урезается мобильными играми и разными развлекательными приложениями. Отсюда вытекает цель нашей работы: разработать и реализовать универсальную интуитивно-понятную систему управления AR-объектами. Актуальность работы заключается в том, что создание такой системы позволит повсеместно распространить AR, а также значительно расширить область применения технологии, кроме того, далее будет рассмотрено, что подобных работ в свободном доступе найдено не было, а, значит, данную работу можно считать уникальной в своем роде. По окончании разработки проекта, все материалы будут выложены в open-source-источники[[3]](#footnote-3) для того, чтобы каждый желающий мог собрать данное устройство, и технология дополненной реальности получила максимально большое распространение.

# Изучение литературы

Перед началом работы необходимо было изучить множество различной литературы. Поскольку наш проект представляет собой систему взаимодействия дополненной реальности и физических объектов, первое, чему было уделено время, – выбор языка программирования и среды разработки для отрисовки объектов дополненной реальности. После тщательного изучения различных IDE[[4]](#footnote-4) и языков программирования, среди которых были, в том числе: С++, Kotlin, Java, C# и Processing, – были отобраны два варианта: Среда Unity-3D или язык программирования Processing. Чтобы использовать в своем проекте наиболее удобный вариант, было решено опробовать обе среды.

## Разработка в Processing

Processing – это подъязык программирования, основанный на java с простым и понятным синтаксисом. Он дает возможность быстро и легко создавать мультимедиа приложения[[5]](#footnote-5). Данный язык используется в основном дизайнерами и художниками. Ниже приведены примеры арт-объектов, созданных, с помощью processing.

  
Рис. 1. Мультимедийный арт-объект, созданный в processing.

  
Рис. 2.   
Мультимедийная динамическая скульптура, созданная в processing.

Разработчики processing создали интерактивный курс, позволяющий изучить основы и синтаксис языка людям, не работавшим ранее с программированием вообще[[6]](#footnote-6). Этот видеокурс удобен тем, что для написания программ, которые используются в качестве примеров, не нужно ничего скачивать. При просмотре курса человеку будет доступно окно ввода кода и вывода результата, код в которых будет помещаться автоматически, в зависимости от того примера, который рассматривается в данный момент в видео, но обучающийся сможет изменять его и наблюдать за результатом.

  
Рис. 3. Страница видеокурса Processing

Таким образом, видно, что Processing имеет большие возможности в реализации графических приложений, а также прост в обучении. К сожалению, Processing – это только подъязык, иными словами, – обертка для Java, специализирующийся на графике, то есть его возможности в других областях довольно узки. В этом мы убедились на собственном опыте, попытавшись написать AR-приложение со следующим принципом работы: программа получает изображение с веб-камеры, и если на этом изображении присутствует некий маркер – специальное изображение, которое будет распознавать программа, например, купюра определенного достоинства, или уникальный рисунок (вообще, AR-маркером может служить любое изображение, главное – обучить программу отличать его от остального фона) – то программа к изображению с веб-камеры добавляет некую 3D-модель. После позиционирования модель на маркере, она должна следовать за маркером, как если бы была настоящей: при повороте маркера – поворачиваться, при перемещении маркера – перемещаться вслед за ним. В следствии узконаправленности Processing, часть с отрисовкой и перемещением 3D-модели была написана нами без особых трудностей, а часть с добавлением изображения с веб-камеры, распознаванием маркера и вообще внедрением AR вызвала такое множество проблем, что мы решили перейти к Unity-3D, опробовать разработку там, и только в случае неудачи вернуться к Processing, либо при необходимости внедрить часть кода на Processing в наш проект, но, в любом случае, отказаться от него, как от основного средства разработки.

## Разработка в Unity-3D

Следующим нашим шагом был переход к Unity-3D. Unity представляет из себя не конкретный язык программирования, а систему, объединяющую среду программирование скриптов[[7]](#footnote-7), инструменты для дизайна приложения, добавления анимации, и т.д. Данная IDE пользуется успехом среди разработчиков игр и мобильных приложений за простоту использования, многофункциональность и широкую область применения.

|  |  |
| --- | --- |
|  |  |
| Рис. . Иерархия обьектов внутри проекта в Unity | Рис. . Иерархия файлов, доступных для использования |

У системы Unity-3D нет официальных систематизированных обучающих курсов, как у Processing, но, благодаря популярности этой среды разработки, существует множество неофициальных уроков о работе в данной среде. Воспользовавшись одним из таких уроков, мы написали приложение, принцип работы которого изложен в предыдущем параграфе. Вкратце, приложение позиционирует 3D-модель на специальном маркере, а также отслеживает положение маркера и перемещает модель вслед за ним. Подробнее об этом будет сказано далее.

  
Рис. 6.  
Функионал Unity на примере нашего приложения.

## Изучение способов отслеживания перемещения

Отслеживание перемещения (positional tracking) объектов в пространстве – достаточно сложная задача, и, чтобы делать это с достаточно большой точностью, необходимо реализовать сочетание аппаратных средств и программного обеспечения.

Методы решения такой задачи можно поделить на несколько групп:

1. Акустические
2. Радиочастотные
3. Магнитные
4. Оптические
5. Инерциальные
6. Гибридные

Рассмотрим базовые принципы, на которых построены вышеперечисленные методы.

### Акустические методы

Акустические приборы слежения используют высокочастотные волны для определения положения объекта в пространстве. А именно, они замеряют время пролета волны от объекта к одному из датчиков системы, которое пропорционально расстоянию до этого датчика. Измерив эти расстояния, можно однозначно определить положения объекта относительно датчиков.

Акустические трекеры, как правило, имеют низкую скорость обновления, вызванную относительно низкой скоростью звука в воздухе.

### Радиочастотные методы

Методов, основанных на радиоволнах множество, и по принципам работы они схожи с акустическими, только работают с другими типами волн. Также существует система устройств, измеряющие время пролета волны от объекта до датчика. Данный способ, подобно предыдущему, является недостаточно точным (точность порядка сантиметров, что является некомфортным для восприятия человеческим глазом).

### Магнитные методы

Магнитный трекинг[[8]](#footnote-8) основан на измерении интенсивности магнитного поля в различных направлениях. Как правило, в таких системах существует генератор переменного или постоянного магнитного поля.

Точность такого метода может быть достаточно высока в контролируемых условиях, однако в обычных помещениях точность может нарушаться из-за различных проводящих поверхностей или других электронных устройств, например, из-за металлических стержней в блоках железобетона, из которого строятся некоторые здания.

### Оптические методы

Оптические методы представляют собой комбинацию из алгоритмов машинного обучения (в частности, алгоритмы сверточных нейронных сетей, или, как их еще называют, компьютерного зрения) и различных отслеживающих устройств, в роли которых могут выступать различные камеры: видимого или инфракрасного диапазона, стерео-камеры[[9]](#footnote-9) и камеры глубины.

В оптических методах трекинга выделяют два подхода:

* *Outside-in[[10]](#footnote-10)* – в качестве точки отсчета присутствует внешняя камера, определяющая положение объекта по определенным маякам[[11]](#footnote-11), расположенным на нем самом.
* *Inside-out[[12]](#footnote-12)* – обратный метод, точкой отсчета является сам объект, на котором закреплен определенный оптический сенсор, а маяки расположено около отслеживаемого тела.

  
Рис. 7.Наглядная демонстрация обоих методов трекинга

### Инерциальный трекинг

Современные инерциальные измерительные системы позволяют отслеживать ориентацию и положение в пространстве с большой точностью и минимальными задержками с помощью таких датчиков, как гироскоп и акселерометр.

Акселерометр – прибор, измеряющий проекцию разности истинного ускорения объекта и гравитационного ускорения по каждой из трех осей. Многие акселерометры оснащены электронной составляющей, способной передавать их показания на различные контроллеры.

Гироскоп – устройство, способное реагировать на изменение углов ориентации тела, на котором оно установлено, относительно инерциальной системы отсчета. Термин впервые был введен Ж. Фуко в своем докладе в 1852 году во Французской академии наук.

Рассмотрев таким образом, возможные способы отслеживания перемещения, мы пришли к выводу, что для нас наиболее удобен способ инерциального трекинга. Объясним свою позицию. Во-первых, этот способ является достаточно точным, его точность сравнима с точностью оптического метода, который использует большинство корпораций, разрабатывающих VR-шлемы, а именно: HTC, Valve, Oculus VR, Sony PlayStation, Microsoft и другие. Во-вторых, он имеет достаточно высокую частоту опроса[[13]](#footnote-13). В-третьих, это одно из самых дешевых решений в современных реалиях. Существуют специальные модули, которые содержат в себе все датчики, необходимые для инерциального трекинга, один из таких модулей был использован в нашей работе.

# Разработка принципа работы модели

## Выбор оборудования

Как было сказано выше, для реализации нашей работы нами был выбран способ инерциального трекинга. Соответственно, было необходимо подобрать оборудование, которое бы позволило нам все реализовать.

### Микроконтроллер Arduino

Arduino – торговая марка аппаратно-программных средств для создания различных систем автоматики и робототехники. В частности, Arduino выпускает одноименные микроконтроллеры, линеек которых существует великое множество, приведем самые популярные из них:

1. Uno – классический вариант, один из самых дешевых, но обладающий достаточной производительностью
2. Mega – вариант, обладающий большей, чем Uno мощностью, но и уступающий ему в цене и по размерам
3. Due – самый мощный контроллер, однако обладающий отличной от остальных архитектурой, соответственно не приспособлен для использования некоторых библиотек и работы в связке с другими контроллерами
4. Nano – миниатюрная плата размером около 5-ти сантиметров, используется для проектов, которым не нужны большие мощности

Изначально мы не знали, насколько большими окажутся затраты вычислительных ресурсов у контроллера, поэтому мы решили использовать Arduino Mega, достаточно мощную, имеющуюся у нас на руках, плату, к работе с которой мы привыкли и, пожалуй, все «подводные камни» нам уже известны.

  
Рис. 8. Arduino Mega

### Модуль трекинга

В качестве модуля трекинга, т.е. датчика для отслеживания перемещений, мы решили использовать плату MPU-9250. Этот модуль содержит в себе такие датчики, как гироскоп, акселерометр и магнитометр.

Магнитометр – прибор для измерения характеристик магнитного поля и магнитных свойств материалов, например, направления поля, градиента поля[[14]](#footnote-14), магнитной индукции, магнитного потока и т.д.

  
Рис. 9. Модуль MPU-9250

Магнитометры применяются в таких областях, как:

1. Геология
2. Археология
3. Навигация
4. Военная разведка
5. Биология и медицина
6. Сейсмология

Используя данные с этого модуля, мы можем отслеживать перемещение и вращение объекта в пространстве с учетом некоторой погрешности, от которой можно избавиться с помощью программы-фильтра, которая подробнее будет описана в следующей главе.

# Написание программного обеспечения

## Программирование на Unity и рендеринг модели на экране

### Динамическое позиционирование модели на изображении-метке

Как было сказано в введении, цель нашего проекта – создать удобную систему взаимодействия с AR-объектами. Соответственно, первой нашей задачей было научиться генерировать AR-объекты на экране компьютера.

Unity – сложная, многогранная и постоянно развивающаяся среда разработки и, разумеется, среди asset-ов[[15]](#footnote-15) к Unity, нашлось то, что нам необходимо.

  
Рис. 10. Интерфейс разработки на Unity

Asset, который мы использовали для нашего проекта, называется Vuforia. С помощью нескольких простых шагов он позволяет создать необходимую нам программу, а именно, алгоритм проецирования выбранной пользователем модели на специальные маркеры, в качестве которых может выступать любое изображение. Чем больше на выбранном изображении различных уникальных элементов, расположенных в разных его частях, тем эффективнее оно будет работать в качестве маркера. При работе с Vuforia, необходимо создавать базы данных, в которых, кроме самой 3D-модели, должно находиться наше изображение-маркер. Vuforia автоматически оценивает эффективность вашего изображения по пятибалльной шкале и показывает вам местонахождение тех уникальных мест по которым будет ориентироваться программа.

  
Рис. 11. Функционал Vuforia на примере 50-рублевой купюры.

Как мы видим на рисунке выше, изображение 50-рублевой купюра имеет рейтинг 4/5 и достаточно большое количество уникальных точек, которые отмечены желтыми крестиками, что, впрочем, логично, ведь это необходимо для того, чтобы отличить настоящие купюры от поддельных.

Вникнем в логику приложения чуть глубже. Вначале программа получает изображение с веб-камеры на устройстве и просто выводит его на экран. При появлении в области видимости камеры AR-маркера, программа определяет перспективное искажение изображения, на основе чего определяет и ориентацию маркера в пространстве, и затем позиционирует 3D модель на маркер. Далее алгоритм продолжает отслеживать перспективное искажение маркера и передвигать объект в соответствии с положением маркера, т.е. динамически позиционирует объект на AR-метке.

### Перемещение объекта посредством управления виртуальным джойстиком.

После разработки динамического позиционирования нужно было дать объекту перемещаться вне зависимости от маркера, т.к. система управления AR-объектами подразумевает взаимодействие с ними с помощью некого устройства, а не посредством передвижения маркера. Для начала, было решено перемещать объект с помощью виртуального джойстика.

Для реализации задуманного мы написали скрипт, который добавили в проект динамического позиционирования. Общая логика скрипта такова. Существуют две встроенные функции: одна из них вызывается только при запуске скрипта, а вторая вызывается при каждом обновлении кадра видеоряда программы. В первой производится настройка программы, углубляться в которую нет смысла в тексте данной работы.[[16]](#footnote-16) Вторая функция считывает отклонение джойстика, записывает пару координат джойстика в трехмерный вектор (трехмерный вектор (x, y, z) был выбран потому, что тогда мы можем перемещать объект во всех плоскостях, записывая в соответствующие позиции вектора координаты отклонения джойстика), после чего умножает вектор на некоторый коэффициент скорости и перемещает объект в соответствии с вектором. Поясним. Например, мы отклонили джойстик от начального положения (координаты (0, 0)) в положение с координатами (x', y'), а затем заполнили вектор перемещения следующим образом: (x', 0, y') – тогда отклонение джойстика вверх повлечет перемещение объекта в направлении от нас, а перемещение вниз повлечет перемещение объекта «к нам». Отклонение джойстика вправо и влево будут перемещать объект в соответствующем направлении. Запись данных в следующей форме: (x', y', 0) – повлечет за собой перемещение объекта в направлении «вверх-вниз» при отклонении джойстика в соответствующем направлении, и перемещение объекта «вправо-влево», как в предыдущей форме записи, при отклонении джойстика «вправо-влево». Таким образом у нас была возможность определять плоскость перемещения объекта.

В итоге, было решено воспользоваться первым вариантом записи – (x', 0, y') – поскольку он являлся интуитивно-понятным при управлении объектом с помощью джойстика. Здесь стоит отметить, что пока что мы можем управлять объектом только в одной плоскости, поскольку джойстик двигается только в двух осях. При управлении непосредственно с помощью устройства, способного перемещаться во всех трех осях, объект, конечно, будет передвигаться во всех возможных плоскостях.

Кроме указанного, в функции, вызывающейся при смене кадра, была также вложенная функция определения поворота объекта. Объясним ее необходимость методом «от обратного». Представим, что функции поворота модели нет. В таком случае, при отклонении джойстика вправо, объект, не поворачиваясь, двигался бы вправо. Данный эффект кажется логичным для симметричных объектов, например, куба или сферы. Но если объект несимметричен, как, например, моделька человека, то ее перемещение вправо «боком» кажется, как минимум странным. Поэтому нами была написана функция, определяющая поворот модели. Работает она следующим образом.

Список литературы

[В Интернете] // Википедия. - 24 Февраль 2019 г.. - 21 Апрель 2019 г.. - https://ru.wikipedia.org/wiki/Стереоскопический\_фотоаппарат.

**Hello, Processing** [В Интернете] / авт. Shiffman Daniel // Processing. - Ben Fry, Casey Reas, 2001 г.. - https://hello.processing.org/.

**Знакомство с Processing 1.0** [В Интернете] / авт. Frexin (@sindrom) // Habrahabr. - 27 Апрель 2009 г.. - https://habr.com/ru/post/58314/.

**Обзор методов и технологий отслеживания положения для виртуальной реальности** [В Интернете] / авт. Сайфуллин Дамир // habr.com. - 20 Сентябрь 2016 г.. - 20 Апрель 2019 г.. - https://habr.com/ru/post/397757/.

1. Internet of things (англ.) – Интернет вещей. [↑](#footnote-ref-1)
2. Augmented reality (англ.) – дополненная реальность. [↑](#footnote-ref-2)
3. Open-source-источники (англ.) – источники свободного доступа. [↑](#footnote-ref-3)
4. IDE – Integrated Development Environment (англ.) – интегрированная среда разработки. Набор средств (специализированные программы) для разработки программного обеспечения. [↑](#footnote-ref-4)
5. (Frexin, 2009) [↑](#footnote-ref-5)
6. (Shiffman, 2001) [↑](#footnote-ref-6)
7. Скрипт (от англ. Script – сценарий) – профессионализм, обозначающий исполняемый файл, содержащий набор инструкций, которым следует программа. [↑](#footnote-ref-7)
8. Трекинг (англ. Tracking – отслеживание) – определение местоположения [↑](#footnote-ref-8)
9. Стереокамеры (или стереоскопический фотоаппарат) – тип фотоаппарата с двумя или более объективами, которые позволяют фотоаппарату симулировать человеческое (бинокулярное) зрение. [↑](#footnote-ref-9)
10. Outside-in (англ.) – дословно переводится как «снаружи – внутри». Здесь и далее профессиональные термины будут использоваться на родном языке с приложением перевода. [↑](#footnote-ref-10)
11. Маяк (здесь и далее) – какие-то характерные отличительные знаки или точки, радиометки, и другие подобные устройства, с помощью которых определяется положение объекта. [↑](#footnote-ref-11)
12. Inside-out (англ.) – дословно переводится как «Внутри – снаружи». [↑](#footnote-ref-12)
13. Частота опроса – величина, показывающая, как часто датчик (или другое любое устройство) может передавать информацию на вычислительную машину. [↑](#footnote-ref-13)
14. Речь идет о градиент-функции, векторе (матрице, в общем виде), заключающим в себе все частные производные данной функции, т.е. непосредственно показывающей направление к экстремумам функции. [↑](#footnote-ref-14)
15. Asset (англ.) – дословно – ценный вклад, достояние. В Unity – созданный кем-то (организацией или одним человеком) набор скриптов и моделей (т.е. готовый пакет) для выполнения определенной задачи [↑](#footnote-ref-15)
16. Код программы содержится на сайте GitHub по ссылке в приложении [↑](#footnote-ref-16)